UNIVERSITY OF LJUBLJANA
FACULTY OF ECONOMICS

KARLA OBLAK

THE ROLE OF INCENTIVE SCHEMES ON SHAPING REFERENCE VALUES AND INFLUENCING DECISIONS ABOUT RISK AND EFFORT

DOCTORAL DISSERTATION

Ljubljana, 2018

AUTHORSHIP STATEMENT

The undersigned **Oblak Karla**, a student at the University of Ljubljana, Faculty of Economics (hereafter: FELU), author of this doctoral dissertation with the title **The role of incentive schemes on shaping reference values and influencing decisions about risk and effort**, prepared under supervision of **dr. Sergeja Slapničar**

DECLARE

1. this doctoral dissertation to be based on the results of my own research;

2. the printed form of this doctoral dissertation to be identical to its electronic form;

3. the text of this doctoral dissertation to be language-edited and technically in adherence with the FELU's Technical Guidelines for Written Works, which means that I cited and / or quoted works and opinions of other authors in this doctoral dissertation in accordance with the FELU's Technical Guidelines for Written Works;

4. to be aware of the fact that plagiarism (in written or graphical form) is a criminal offence and can be prosecuted in accordance with the Criminal Code of the Republic of Slovenia;

5. to be aware of the consequences a proven plagiarism charge based on this doctoral dissertation could have for my status at the FELU in accordance with the relevant FELU Rules;

6. to have obtained all the necessary permissions to use the data and works of other authors which are (in written or graphical form) referred to in this doctoral dissertation and to have clearly marked them;

7. to have acted in accordance with ethical principles during the preparation of this doctoral dissertation and to have, where necessary, obtained permission of the Ethics Committee;

8. my consent to use the electronic form of this doctoral dissertation for the detection of content similarity with other written works, using similarity detection software that is connected with the FELU Study Information System;

9. to transfer to the University of Ljubljana free of charge, non-exclusively, geographically and time-wise unlimited the right of saving this doctoral dissertation in the electronic form, the right of its reproduction, as well as the right of making this doctoral dissertation publicly available on the World Wide Web via the Repository of the University of Ljubljana;

10. *[please include only if your doctoral dissertation consists of several articles]* to have acquired from publishers, to whom I have previously exclusively transferred material copyrights for articles, all the necessary permissions for the inclusion of articles in the printed and electronic forms of this dissertation. These permissions enable the University of Ljubljana to save this doctoral dissertation in the electronic form, to reproduce it, as well as to make it publicly available on the World Wide Web via the Repository of the University of Ljubljana free of charge, non-exclusively, geographically and time-wise unlimited;

11. my consent to publication of my personal data that are included in this doctoral dissertation and in this declaration, when this doctoral dissertation is published.

Ljubljana, March 1st, 2018 Author's signature: _________________________

ACKNOWLEDGEMENT

»Give a man a fish and you feed him for a day; teach a man to fish and you feed him for a lifetime« (Maimonides)

I would like to express my deepest gratitude to my advisor Dr. Sergeja Slapničar who gave me her helping hand when I entered the world of sicence. I am thankful for her motivation, patience, time dedication and immense knowledge and ideas that she shared with me. The positivity and enthusiasm she has for research was inspirational, and helped me through tough times. I might have not been the best candidate, but she believed in me and guided me to the finishing line.

Statistics for lawyers are science fiction, and I would like to thank Mina, who put great effort into helping me understand and use statistics, and who encouraged me when I felt lost in the world of numbers.

I would like to thank my family for supporting me in my pursuits, for all their encouragement and love. Mum, dad, Eva and Matic. Thank you.

The role of incentive schemes on shaping reference values and influencing decisions about risk and effort

Summary

This PhD thesis focuses on the effect of incentive schemes on decisions to invest cognitive effort in a multi-period setting where the level of cognitive effort demand is increasing. The research combines theories from economics, management accounting, social psychology and insights from neuroscience to provide an insight into complex decision-making – cognitive effort. The effect of financial incentives on human behaviour is one of the fundamental questions in Economics (eg. Bénabou and Tirole, 2003; Kreps, 1997; Lazear, 1986, 2000; Prendergast, 1999), psychology and management literature (e.g., Gerhart and Rynes, 2003; Gomez-Mejia and Welbourne, 1988; Rynes et al., 2005; Vroom, 1964). Economics is, despite all the formulas, models, theories, meant to explore people's behaviour.

The dissertation aims to analyse how the incentive condition affects decisions on cognitive effort exertion, and how personality traits (achievement motivation and reward/punishment sensitivity) modulate these decisions. The hypotheses were tested experimentally. In a series of multi-period choice task experiments we manipulated individuals` reference points, creating a gain or loss domain with reward and penalty incentive treatments. Subjects had to decide between an easy and difficult cognitive task where the difficulty level was associated with outcome risk and effort level. We use a modified Sternberg task to measure cognitive risk-effort decisions.

The PhD thesis is organised as follows. In the first chapter, we explore how cognitive frames created by incentive design and the outcome's fairness influence decisions on risk and effort by affecting the subjective valuation of monetary payoffs. In a three period, 2 × 2 between subject experiment, we analyse the effects of a bonus versus a penalty contract and a fair versus an unfair outcome distribution. We hypothesise that, in the case of conflicting cues from the two frames, the cue that creates a perception of loss dominates the decision. We also hypothesise that, over time, prior performance influences current decisions by creating a new cognitive frame. We find that if the pay is unfair, neither a bonus nor a penalty seems to matter. If it is fair, high risk-effort tasks are stimulated more by a penalty than a bonus contract. The effect of prior performance eventually outweighs the effect of both incentive manipulations.

In the second chapter, we evaluate the effect of monetary incentives in choices to perform cognitively challenging tasks given individual differences in achievement motivation. In a three-period, within-subject experiment, we measured willingness to engage in a cognitively challenging task in a non-incentivised and an incentivised setting (with rewards or penalties). We provide evidence on incentive sensitivity to achievement motivation for increasing task difficulty. The same incentives affect people with different achievement motivation and failure avoidance differently. As cognitive tasks become increasingly difficult, rewards have a diminishing effect on the choice of a challenging task, significantly more so for high achievement individuals. Penalties can increase selection of challenging tasks at low and moderate task difficulty, but their effect is mitigated for individuals with high motivation to avoid failure. For high task difficulty penalties, only individuals with a low fear of failure are prepared to undertake a challenging task.

In the third chapter, we examine the differences in the effect of monetary rewards on gender in a challenging cognitive task. We are interested in how rewards motivate men and women at various difficulty levels. In a three-period, within-subject experiment, we measured participants' choices, their cognitive performance and financial outcome. We find that women and men responded differently to monetary rewards at various levels of task difficulty, but did not cofirm our hypothesis regarding the effect of the behavioural approach or avoidance behaviour. We find that the most important determinant of choices is the success or failure feedback of the prior period. The study advances the literature studying gender differences in various incentive contexts.

Key words: incentive scheme, reward, penalty, fairness, risk, effort, achievement motivation, BIS/BAS

Pomen shem nagrajevanja za oblikovanje referenčnih vrednosti in vpliv na odločitve o tveganjih in naporu

Povzetek

Doktorska disertacija raziskuje vpliv shem nagrajevanja na odločitve o kognitivnem naporu v več obdobjih, pri čemer se raven zahtevanega kognitivnega napora povečuje. Raziskava vključuje teorije s področja ekonomije, poslovodnega računovodstva, socialne psihologije in uvide s področja nevroznanosti, ki omogočajo vpogled v kompleksen proces sprejemanja odločitev o kognitivnem naporu. Učinek finančnih spodbud na človeško vedenje je eno izmed temeljnih vprašanj, ki se poraja v ekonomiji (npr. Bénabou in Tirole, 2003; Gibbons, 1998; Kreps, 1997; Lazear, 1986, 2000; Prendergast, 1999), psihologiji in poslovodenju (npr. Gerhart in Rynes, 2003; Gomez-Mejia in Welbourne, 1988; Rynes et al., 2005; Vroom, 1964), saj je ekonomija kljub vsem formulam, modelom in teorijam namenjena raziskovanju človeškega vedenja.

Namen disertacije je analizirati, kako sheme nagrajevanja vplivajo na odločitev o kognitivnem naporu in kako na to odločitev vplivajo osebnostne lastnosti: motivacija, usmerjena k uspehu (*Motivation to achieve – MTA*), in motivacija, usmerjena k izogibanju neuspeha (*Motivation to avoid failure – MTF*), ter vedenjski sistem umika (*Behavioural inhibition system* – BIS) in vedenjski sistem približevanja (*Behavioural approach system* – BAS). Hipoteze so bile preizkušene eksperimentalno. V seriji eksperimentov smo v več obdobjih z nalogo izbire manipulirali z referenčnimi točkami posameznikov, tako da smo ustvarili domeno pridobitve ali domeno izgube s shemami nagrajevanja v obliki nagrad ali kazni. Udeleženci so se morali odločiti med preprosto in zahtevno kognitivno nalogo, kjer je bila raven zahtevnosti povezana s tveganjem izida in ravnjo napora. Pri tem smo uporabili prilagojeno Sternbergovo nalogo (Sternberg, 1966) za merjenje odločitev za napore in tveganja.

Doktorska disertacija je sestavljena iz treh poglavij. V prvem poglavju raziskujemo, kako kognitivni okviri, ustvarjeni s shemami nagrajevanja in pravičnostjo porazdelitve nagrad, vplivajo na odločitve o tveganju in naporu, s tem ko vplivajo na subjektivno vrednotenje denarnih izplačil. Vzpostavili smo eksperiment z zasnovo 2 x 2 s tremi obdobji, pri čemer analiziramo učinke sheme nagrajevanja z nagrado (*reward*) v primerjavi s shemo kaznovanja (*penalty*) ter pravično porazdelitev nagrad v primerjavi z nepravično porazdelitvijo nagrad. Predpostavljamo, da v primeru protislovnih možnosti iz dveh okvirov (npr. shema kaznovanja in pravična porazdelitev nagrade ali shema nagrajevanja in nepravična porazdelitev nagrade) pri odločanju prevladuje okvir, ki ustvarja dojemanje izgube. Prav tako predpostavljamo, da predhodna uspešnost vpliva na odločitve z ustvarjanjem novega kognitivnega okvira. Ugotavljamo, da ob nepravičnem plačilu niti nagrada niti kazen nista pomembna in ne vplivata na kognitivni napor, povezan s tveganjem in naporom. Če je shema nagrajevanja pravična, pa odločitev za naloge, ki

zahtevajo veliko napora in so tvegane, spodbudimo s shemo kaznovanja. Najpomembnejši dejavnik, ki na koncu vpliva na odločitev o tveganju in naporu, je učinek predhodne uspešnosti, ki tudi prevlada nad učinkom obeh manipulativnih spodbud.

V drugem poglavju ocenjujemo učinek denarnih spodbud pri izbirah za izvedbo kognitivno zahtevnih nalog glede na individualne razlike v motivaciji, usmerjeni k dosežku (*Achievement motivation*). V eksperimentu, ki je imel tri obdobja, smo izmerili pripravljenost vložiti kognitivni napor kot izbiro med lahko in težko nalogo v pogoju brez denarnih spodbud in v pogoju z denarnimi spodbudami (shema nagrajevanja ali shema kaznovanja). Rezultati eksperimenta kažejo na občutljivost shem nagrajevanja za motivacijo, usmerjeno k dosežku pri povečanju težavnosti nalog. Enake spodbude različno vplivajo na ljudi z različno motivacijo, usmerjeno k uspehu (*MTA*), in motivacijo, usmerjeno k izogibanju neuspeha (MTF). Pri vedno zahtevnejših kognitivnih nalogah imajo nagrade manjši vpliv na izbiro zahtevne naloge, zlasti za zelo uspešne posameznike. Kazni lahko povečajo izbiro zahtevnih nalog pri lahkih in srednje težkih nalogah, vendar pa se njihov učinek ublaži pri posameznikih, ki so močno motivirani, da se izognejo neuspehu. Pri zelo težkih nalogah so samo tisti z nizko stopnjo strahu pred neuspehom pripravljeni izbrati in opraviti zahtevno kognitivno nalogo.

V tretjem poglavju obravnavamo razlike v učinku shem nagrajevanja glede na spol pri zahtevnih kognitivnih nalogah. Zanima nas, kako nagrade motivirajo moške in kako ženske pri različnih težavnostnih stopnjah. V eksperimentu v treh obdobjih smo merili odločitve udeležencev, njihovo kognitivno uspešnost in finančni izid. Ugotovili smo, da se ženske in moški različno odzivajo na denarne nagrade pri različnih stopnjah težavnosti nalog, vendar naša hipoteza glede učinka vedenjskega sistema umika (BIS) in vedenjskega sistema približevanja (BAS) ni bila potrjena. Ugotovili smo, da so najpomembnejša determinanta za odločitev o kognitivnem naporu povratne informacije, ki jih udeleženci prejmejo za uspeh ali neuspeh v predhodnem obdobju. Poglavje prispeva k literaturi, ki preučuje razlike med spoloma v različnih kontekstih shem nagrajevanja.

Ključne besede: shema nagrajevanja, nagrada, kazen, pravičnost, tveganje, napor, motivacija, usmerjena k dosežku, vedenjski sistem umika, vedenjski sistem približevanja

TABLE OF CONTENTS

LIST OF FIGURES

LIST OB TABLES

INTRODUCTION

The past decade has witnessed a turnaround in academic thought on incentive schemes that are used primarily to motivate employees, and foster effort and moderate risk behaviour. The global financial crisis brought up questions on behavioural attributes which are difficult to incorporate in models of »rational« decision-making because, as widely acknowledged, traditional bonus schemes stimulated dysfunctional risk taking. Akerlof and Shiller (2009) argue that it is impossible to understand the economic developments of recent times without psychological insights which go beyond established notions of "rationality".

Incentives affect the human decision process in various areas of life, and have the power to create a reference point that shapes the cognitive frame that further affects the decision – making process. Outcomes are assessed in relative terms rather than in absolute terms; as gains or losses relative to a reference point (Kahneman and Tversky, 1979). In order to introduce efficient incentive schemes, researchers and regulators need to understand how decision-makers value potential payoffs, monetary and non-monetary. Incentive schemes should be designed so that they guide managerial behaviour in the direction of long-term maximization of firm value and sustainability, and to promote cognitive effort that is the essence of innovation and creativity.

Cognitive effort is related to effort and risk decisions, as it refers to engagement in demanding tasks associated with high effort exertion and risk associated with the positive outcome of such tasks. Although present management and management accounting literature distinguishes the theories of risk-taking from the theories of motivation (exertion of effort), neuroscientific findings suggest that a unified theory of effort and risk may be warranted, as neurons encode value and risk of a reward simultaneously and in the same brain regions (Hughes, Yates, Morton and Smillie, 2015; Miller, Thomé and Cowen, 2013; Platt and Huettel, 2008; Wardle, Treadway and De Wit, 2012). The potential effect of a contracting frame on effort and risk is critical, as incentive contracts are offered as control mechanisms to align individual and firm goals (Church, Libby and Zhang, 2003).

We designed a research project that aims to answer the question of how to incentivise cognitively challenging tasks effectively. In three research projects we study different aspects of decision-making to understand better the effectiveness of incentive schemes and the importance of cognitive frames created by the design of incentives in relation to other variables, such as task difficulty, motivation and success feedback. Our hypotheses were tested experimentally. Despite some criticism, classroom experiments are a powerful tool to test real decision-making (Runeson, 2003; Exadaktylos, Espín and Brañas-Garza, 2013). To analyse the research problem thoroughly, we focus on three different, although related, topics. We provide a brief outline of the three chapters that follow.

In the first chapter, we explore how cognitive frames created by incentive design and the outcome's fairness influence decisions on cognitively challenging tasks (which we call risk-effort tasks), and advance the management accounting literature by integrating the organisational justice theory (Adams, 1963; 1966) and prospect/framing theory (Kahneman and Tversky, 1979; Kahneman, 2003) with the prior evidence on risk-taking and effort exertion to understand better the interactive cognitive frames in comprehensive decision-making. Organizational justice theory applies reference values to decisions about motivation to exert effort without explicit consideration of the outcome risk, prospect theory uses them to predict risk-taking behavior but in practice, decisions about risk and effort are often simultaneous. By incentive design we create two distinct cues for the formation of the reference point. The first cue comes from labeling performance pay as a bonus rather than a penalty. The second comes from the fairness or unfairness of the payoff with respect to peers. We explore whether one cue strengthens the effect of the other if they are consistent or whether one cue dominates the other if they are inconsisten. Our results help to advance the management accounting literature by integrating separate theories with the prior evidence on risk-taking and effort exertion to better understand interactive cognitive frames in comprehensive decision-making

In the second chapter, we evaluate the effect of monetary incentives on choices to perform cognitively challenging tasks given individual differences in achievement motivation and how increasing task difficulty with which successful tasks completion becomes less and less certain moderates the relation between incentives and achievement motivation. Individuals' motivational disposition influences their perception of a situation as being rewarding or punishing (Gray and McNaughton, 2000) and the source of an action may derive from motivation to achieve or to avoid a failure (Atkinson, 1957; Deci and Ryan, 1985; Heckhausen, 1963, 1991). We advance the first chapter by adding the non-incentivised treatment, and compare choices in the incentivised (reward or penalty) and non-incentivised treatment. The study advances the motivation theory and practice by an original analysis of the interaction between monetary incentives and achievement motivation in cognitively challenging tasks. The chapter contributes to a cross-disciplinary view that advances our understanding of the functionality of incentives (Merchant, Van der Stede and Zheng 2003) and to the growing body of accounting studies that examines the effects of incentive framing on behavioural outcomes and choices (e.g., Brown, Farrington and Sprinkle, 2016; Brink & Rankin 2013; Church et al., 2008; Christ et al., 2012, Hannan et al. 2005; Hartmann and Slapničar, 2015; Luft 1994, Oblak et al., 2017).

In the third chapter, we analyse how men and women make decisions in undertaking challenging cognitive tasks at increasing task difficulty, with and without rewards as one of the problems in the modern organisations is gender pay gap and under-presence of women in high profile jobs (Croson and Gneezy, 2005). As prior literature suggests, women and men respond differently to challenging tasks, and their preferences are driven by differences in personality traits, risk attitude and beliefs in their abilities (Niederle and

Yestrumskas, 2008). Women are less risk prone, more careful in making high reaching decisions, more afraid of punishment and of omission of a reward, and react differently than men to rewards where a successful outcome is not certain. Moreover, two psychological systems, Behavioural Inhibition System (BIS) and Behavioural Approach System (BAS) guide individuals' behaviour (Gray, 1981) and act as moderators of the decision-making process, as they are sensitive to rewarding or punishing stimuli. When studying gender differences in incentive sensitivity, it is important to take into account the underlying gender differences in incentive sensitivity, of which there is only limited prior evidence (De Pater et al., 2009). The study advances the literature studying gender differences in decision-making and provide evidence that selection of challeninging task is not positively related with overall better performance and financial outcome.

THE ROLE OF COGNITIVE FRAMES IN COMBINED DECISIONS ABOUT RISK AND EFFORT[1]

Abstract

Cognitive framing influences the subjective valuation of monetary payoffs and an individual's willingness to exert effort and take risk. In this chapter, we explore how cognitive frames created by incentive design and the outcome's fairness influence decisions on risk and effort. While such decisions are often combined in practice, the theories that study risk-taking and motivation to exert effort remain discrete. We set up a multiperiod, 2 × 2 experiment in which we analyze the effects of a bonus versus a penalty contract and a fair versus an unfair outcome distribution. We use a modified Sternberg task to measure risk-effort decisions. We hypothesize that in the case of conflicting cues from the two frames, the cue that creates a perception of loss dominates the decision. We also hypothesize that over time, prior performance influences current decisions by creating a new cognitive frame. We find that if the pay is unfair, neither a bonus nor a penalty seems to matter. If it is fair, high risk-effort tasks are stimulated more by a penalty than a bonus contract. The effect of prior performance eventually outweighs the effect of both incentive manipulations. Our results help to advance the management accounting literature by integrating separate theories with the prior evidence on risk-taking and effort exertion to better understand interactive cognitive frames in comprehensive decision-making.

Keywords: incentive scheme, framing, contract, bonus, penalty, fairness, effort, risk

JEL code: M41

[1] The paper is co-authored with Sergeja Slapničar and Mina Ličen (University of Ljubljana). It has been accepted for publishing in Management Accounting Review. https://doi.org/10.1016/j.mar.2017.07.001

1.1 Introduction

Notable psychological theories stress that decision-making depends on an individual's cognitive frames or mental representations of the decision problem (Birnberg et al., 2007). The design of incentive systems has an important effect on cognitive frames that influence individuals' perception of fairness, their levels of aspiration, and whether they see outcomes as gains or losses. Two leading psychology theories – organizational justice theory (Adams, 1963; 1966) and prospect/framing theory (Kahneman and Tversky, 1979; Kahneman, 2003) – propose that cognitive frames arise by comparing an outcome to a reference point. In organizational justice theory, the reference point represents a comparison with a relevant other, whereas in prospect theory the reference point is basically the status quo (Kahneman, 2003) and may be invoked by a variety of characteristics of the incentive system. An idea common to both theories is that reference points shape cognitive frames and that a deviation from them causes internal conflicts that individuals try to avoid (Birnberg et al., 2007). In more complex decision situations, individuals face several cognitive frames at the same time, and the question arises on which one plays a central role in decision-making and how they interact.

Although the two theories share a profoundly related concept, it is interesting that they remain discrete: whereas the organizational justice theory applies reference values to decisions about motivation to exert effort without explicit consideration of the outcome risk, prospect theory uses them to predict risk-taking behavior (pure monetary payoffs in the absence of any effort). Yet, in practice, decisions about risk and effort are often simultaneous: in many settings individuals face an option that requires a lot of effort, which potentially brings a high payoff, but the probability of obtaining that payoff depends on the success in completing the task. The alternative is to choose an easy option that requires little effort and has a high probability of success but results in a low payoff. Examples of these options are choosing between a more difficult or an easier field of study that leads to different future salary levels; between a demanding or a less demanding job with the corresponding pay levels and chances of success; between writing a scientific paper for a high impact journal or a low impact journal with the corresponding effort, probabilities of success, and impact factors; choosing between highly uncertain but high-yielding projects in which a lot of effort and new knowledge has to be invested or certain low-yielding projects that require an average amount of work and acquired knowledge.

The aim of this chapter is to use both theories to establish which cognitive frames dominate in simultaneous decisions on risk and effort. The literatures on neuroscience, psychology (Salamone et al., 1994; Treadway, Peterman, Zald and Park, 2015 Walton, Kennerley, Bannerman, Phillips and Rushworth, 2006; Wardle, Treadway and De Wit, 2012), and animal behavior (Hosking, Cocker, and Winstanley, 2014; Hosking, Floresco, and Winstanley, 2014) all examine the relation between risk and effort (reviewed in Salamone et al., 2012; Miller, Thomé and Cowen, 2013). This body of work reinforces the conjecture that decisions about risk and effort are related because the neural networks activated in both types of decisions tend to overlap.

We analyze the decisions on effort and risk when two features of the incentive scheme give two distinct cues for the formation of the reference point. The first cue comes from labeling performance pay as a bonus rather than a penalty. The second comes from the fairness or unfairness of the payoff with respect to peers. As the base pay is likely to be perceived as the reference point, labeling performance pay as a bonus creates a perception of a gain, and labeling it as a penalty creates a perception of a loss. Similarly, if peers receive a larger bonus or a smaller penalty for the same effort, then the peers' pay level could become the reference point, and the individual's own bonus could appear as a loss. We explore whether one cue strengthens the effect of the other if they are consistent or whether one cue dominates the other if they are inconsistent.

To understand these questions, we develop a three-period, between-subjects, 2 × 2 (bonus vs. penalty and fair vs. unfair outcome) experiment in which we test the effects of manipulations on joint risk-effort decisions. We use a modified Sternberg task (Sternberg, 1966). The Sternberg task is broadly used in psychology to measure cognitive effort (Burrows and Okada, 1973; Jansma, Ramsey, De Zwart, Van Gelderen and Duyn, 2007; Zakrzewska and Brzezicka, 2014). We operationalize the risk component by designing three periods, offering increasing incentives for rising task difficulty and probability of failure. We, thus, operationalize joint risk-effort decisions as choices between a high-yielding task that requires high effort with a higher chance of failure (a difficult task) and a low-yielding task that requires low effort with a lower chance of failure (an easy task). A temporal setting creates a third cognitive frame because a positive or a negative prior outcome affects the current decision differently (Thaler and Johnson, 1990). The experiment is tested on 100 students.

We find that the frequency of high risk-effort decisions is the lowest under a fair bonus contract and higher under either a penalty or unfair contract. In a comprehensive setting where both incentive frames are at work if the pay is unfair, it matters little whether the contract is framed as a bonus or a penalty. If the pay is fair, high risk-effort decisions are stimulated more by a penalty than a bonus contract. A fair penalty contract elicits high risk-effort decisions most frequently. A fair bonus contract seems to represent a comfort zone that invokes risk-effort decisions least frequently. In the second round, we observe that the participants' prior performance becomes relevant; and in the third round, the effect of prior performance completely overrides all others: the incentive frames are no longer important. This effect suggests that the evaluation of the probability that one can successfully complete a task based on prior performance and prior choices becomes more important than the incentive scheme or the outcome's fairness and forms a reference point on its own.

The chapter makes several contributions to the literature. The first contribution is the examination of simultaneous risk-effort decisions. Without considering such decisions, it is impossible to fully understand the effectiveness of incentive schemes. Performance is frequently a function of risk and effort, yet to our knowledge there is only one paper that explicitly addresses how managerial accounting practices affect risk and effort decisions (Sprinkle, Williamson and Upton, 2008). However, unlike our study in which risk and effort are related, Sprinkle et al. (2008) examine risk-taking independently of the participants' exertion of effort. Most management accounting studies adopt the expectancy theory's assumption about the relationship between risk and effort where the higher the probability that effort will lead to increased performance, the more motivated a person will be to exert effort (Vroom, 1964). In this decision context, an individual may affect the probability of success by exerting more effort (i.e., probability of success is endogenous). On the other hand, we study the decisions in which an individual ex ante chooses a level of a task difficulty that comprises the required effort and acceptable risk. In our decision context, the estimated probability of success is exogenously chosen. Once a level of task difficulty is chosen, the expectancy theory's assumption applies in that more effort will increase the probability of success.

The chapter's second contribution is in analyzing how individuals consider more than one cognitive frame at a time. Our findings indicate that the bonus and penalty schemes invoke cognitive frames in line with prospect theory, which adds to the evidence on how various incentive practices shape cognitive frames. We show that when multiple frames interact they stimulate different behavior to that elicited by a single cognitive frame. Third, by studying decision-making in a multiperiod setting, we show that the effect of incentive schemes fades over time as a new salient piece of information emerges (i.e., prior performance) that helps re-evaluate the probability of an outcome. Fourth, our findings hold practical implications for designing effective incentive schemes. The penalty scheme has been found to fuel high risk-effort decisions. As penalty schemes are gaining popularity via a bonus deferral system containing potential penalties and clawback clauses (Hartmann and Slapničar, 2014; Van der Stede, 2011), our findings indicate that they must be implemented with a clear awareness of their effects. Finally, this chapter integrates two influential psychological theories with the management accounting literature and practice.

The remainder of the chapter is organized as follows. The theoretical background and hypotheses are presented in section 2. Section 3 introduces the experimental design and its execution. Section 4 presents the results, while section 5 concludes with a discussion and the implications and limitations of the study.

1.2 Theoretical background and hypotheses development

The importance of cognitive frames was first described by Kahneman and Tversky (1979). In their paper on prospect theory, they showed that the utility of an outcome depends on whether it is perceived as a gain or a loss, rather than on its absolute value and probability. This perception depends on a reference value against which the outcome can be measured. The wording of a decision problem itself (i.e., framing) may change the perceived outcome's utility and influence risk choices. In general, people are risk-averse in the gain domain and risk-seeking in the loss domain: they opt for a higher but probable loss over a smaller but certain one to avoid a sure loss. Further theoretical development has resulted in the so-called theory of framing (Kahneman, 2003), which postulates that reference points may arise from various comparisons, such as with relevant others and with prior periods. The explanation of framing closely coincides with Thaler's (1999) idea of mental accounting.

Independently of the research on decision-making under risk, the organizational justice literature stresses the importance of reference values for motivation. This literature proposes that people are motivated if they perceive a balance in exchange relationships and evaluate the balance by comparing their effort and outcomes to comparable others' effort-outcome ratios (Adams, 1963). If they perceive injustice, they adjust their effort downwards. Comparison with a relevant other is hence one of the central reference points in organizational justice theory. A large body of evidence demonstrates that a perception of distributive fairness has a major impact on motivation.

While the organizational justice theory acknowledges that cognitive frames affect risk-taking and the willingness to exert effort, the questions of which cognitive frames various management accounting practices elicit and whether they are perceived as fair or unfair are less understood. What is the reference point against which one evaluates gains and losses for risk-taking, and does the same reference point impact decisions about effort? Druckman (2001) and Maule and Villejoubert (2007) find that people consider different reference points. These different points explain why the empirical findings on the effects of framing are contradictory. The management accounting literature has relatively neglected the examination of an incentive scheme's effect on risk-taking compared to some other areas (Sprinkle, 2003; Sprinkle et al., 2008). Tests of goal setting on motivation alone result in over 1,000 studies (Birnberg, Luft and Shields, 2007), whereas to our knowledge only a handful of studies examine how framing of incentive schemes influences risk-taking (Ruchala, 1999; Chow, Kohlmeyer and Wu, 2007; Sprinkle et al., 2008; Drake and Kohlmeyer, 2010; Hartmann and Slapničar, 2014). These studies show that various designs of incentive schemes create cognitive frames and influence the perceptions of gains and losses in relation to risk-taking.

The empirical studies on risk behavior are predominantly conducted on lottery gambles that isolate the decision on risk from the decision on effort, whereas the motivation literature analyzes the effect of incentives on effort but disregards the uncertainty or risk associated with increasingly large outcomes. Real-life decisions are not structured as lotteries with known probabilities of failure, and it is not always the case that probabilities of large outcomes can be increased with greater effort. Real-life alternatives are often associated with effort and uncertainty simultaneously.

The following sections will briefly overview the empirical evidence on how incentive schemes and the outcome's fairness influence the exertion of effort and risk-taking. We then hypothesize how they are expected to work in combination for joint decisions on effort and risk-taking.

1.2.1 The influences of bonus and penalty contracts on risk and effort decisions

Numerous studies on the effect of *bonus and penalty contracts* on *exerting effort* find that penalty contracts elicit a greater level of effort than bonus contracts (Brooks, Stremitzer, and Tontrup, 2013; Church, Libby and Zhang, 2008; Gose and Sadrieh, 2012; Hannan, Hoffman and Moser, 2005; Hossain and List, 2012; Van De Weghe and Bruggeman, 2006). The authors explain the effect as either because of loss aversion or a fear of losing. Their findings support the idea that penalty contracts invoke the perception of a loss domain, which affects behavior more strongly than a gain domain (Cacioppo and Berntson, 1994). However, in practice not all tasks can be governed by an incentive contract. In an incomplete contract setting, Christ, Sedatole and Towry (2012) find that penalty contracts are associated with lower trust in the principal and that they lead to lower effort than bonus contracts in all tasks not governed by a contract.

The research on the effects of bonus and penalty contracts on *risk-taking* reports consistent findings that those incentive characteristics that create the perception of a loss domain increase risk propensity. Interestingly, such perceptions may be invoked by various management accounting mechanisms. Budget levels can, for example, form a positive or a negative frame. A loss frame is induced when individuals are failing to achieve their budget goal and, to reach it, they indulge in risk-seeking behavior. On the other hand, a gain frame occurs when individuals are ahead of their budget goal. They then show risk-averse behavior (Ruchala, 1999). Chow et al. (2007) find that high budget targets promote higher risk-taking. However, Sprinkle et al. (2008) find a more complex U-shaped relationship between budget levels and risk-taking: low budget levels stimulate risk-seeking behavior, higher budget levels suppress such behavior, and stretch budget levels again promote risk-seeking behavior as the only way to potentially meet budget targets. Based on these findings we propose that a penalty contract leads to greater effort and higher risk-taking than a bonus contract:

H1: The frequency of high risk-effort decisions is greater under a penalty than under a bonus contract.

1.2.2 The influence of the outcome's fairness on risk and effort decisions

The feeling that an outcome distribution is unfair can create a cognitive conflict that influences the motivation to exert effort. The effect of the *outcome's fairness* on *exerting effort* has been extensively empirically investigated. Laboratory and field studies provide robust evidence that unfair treatment results in decreased effort (Akerlof and Yellen, 1990; Blau, 1993; Byrne, Stoner, Thompson and Hochwarter, 2005; Cohn, Fehr and Goette, 2014; Cohn, Fehr, Herrmann and Schneider, 2011; Gächter and Thöni, 2010; Hannan et al., 2005; Hartmann and Slapničar, 2008; Lindquist, 2010)[2]. We are aware of only one study, Charness and Kuhn (2007), that reports no social comparison effect. Whether it works the other way is unclear: the findings on unfairness from overcompensation have been less coherent than those on under-compensation (Ambrose and Kulik, 1999).

There is also the question of whether the outcome's fairness can invoke a gain or loss domain that affects risk-taking. Diecidue and van de Ven (2008) explain the role of reference points as an aspiration level. A simple loss frame may not elicit risk-seeking per se but it will elicit risk-seeking if the aspiration level can be achieved by assuming greater risk. Linde and Sonnemans (2012) test this question experimentally. They predict that a participant will be risk-seeking in the social loss frame and risk-averse in the social gain frame. However, they find the opposite: participants are more risk-averse in the social loss frame (i.e., unfair treatment) than in the social gain frame (fair treatment). They may have behaved so because they could not make up for their social loss with higher risk-taking. Haisley, Mostafa and Loewenstein (2008) also suggest that social status may invoke the perception of a social loss frame and increase risk-seeking. In their experiment, people who are shown that their social status is at the bottom of the income distribution are more willing to buy lottery tickets than those who are shown that their social status is somewhere in the middle. Schwerter (2013) analyzes the gambling decisions of participants who observe the earnings of peer participants before making a risky choice. Participants in a treatment group in which their peers' earnings are higher are more risk-seeking than those in a treatment group in which peer's earnings are lower. Overall, these findings show that an outcome's fairness also invokes a gain or loss domain that influences risk-taking, which is consistent with prospect theory.

The effect of fairness on risk-effort decisions is, thus, ambiguous. While perceived unfairness in comparison to peers leads to a decrease in effort (in the absence of risk), most studies generally find that unfairness can increase risk-taking (in the absence of effort) if

[2] The effect is not limited to people. Capuchin monkeys also demonstrate a negative response to the unequal distribution of rewards by refusing to participate in an effortful task if they witness that other participants receive equal reward for less work (Brosnan and De Waal, 2003).

higher risk taking is what it takes to achieve a more favorable social comparison. Which effect will dominate in combined risk-effort decisions may depend on the context and other cues.

1.2.3 The interaction effect of both frames on risk and effort decisions

While the theories on risk behavior and effort exertion (motivation) in humans are more or less unconnected, many studies on animal behavior take a unitary approach. An important stream of literature looks at risk-effort decisions from an evolutionary perspective. Bhatti, Jang and Jeong (2014) suggest that loss avoidance is an evolutionary conserved trait, and the studies that explore its origins may help uncover the mechanisms behind decision-making preferences. According to Thorndike (1898) and Williams (1988), human responses to risky situations derive from the same mechanisms that evolved in animals. The evolutionary process theory finds that loss aversion arises from the goal of minimizing the possibility of extinction (Robson and Samuelson, 2009). Animals exhibit risk-averse behavior and exert less effort when not in danger of energy depletion (gain domain) and risk-seeking behavior and exert more effort when in danger of starvation (loss domain) (Jentsch, Woods, Groman and Seu, 2010; Miller, Thomé and Cowen, 2013). The neuroscientific research on the relatedness of risk and effort decisions finds that the neural networks involved in risk and effort behaviors to some extent overlap (Hughes et al., 2015; Miller et al., 2013; Platt and Huettel, 2008; Salamone et al., 2012; Treadway et al., 2009; Walton et al., 2006; Wardle et al., 2012). Several neuroscientific studies report that risk-effort types of decisions are jointly moderated by the neurotransmitter dopamine, which functions as a reward and probability signal (Bardgett, Depenbrock, Downs, Points, Green, 2009; Bautista, Bardgett, Depenbrock, Downs, Points and Green, 2001; Cowen, Davis, Nitz, 2012; Kennerley, Walton, Behrens, Buckley and Rushworth, 2006; Kirshenbaum, Szalda-Petree and Haddad, 2003; Salamone, Cousins and Bucher, 1994, Salamone, Correa, Nunes, Randall and Pardo, 2012).

A perception of loss seems to be the ultimate driver of behavior. The research consistently finds that penalty contracts induce such a domain but a perception that outcome distribution is unfair also creates a loss domain that can stimulate risk-taking. A bonus contract may be perceived as a loss if an individual sees that others are getting higher pay. It is therefore highly likely that any cue that triggers a perception of a loss is dominant and overrides an alternative cue. In an interaction between a bonus or a penalty contract and the outcome's fairness, it may well be that in the case of a bonus contract, unfairness is such a cue; whereas in the case of a penalty contract and the outcome's fairness, the penalty contract itself suffices to create a perception of loss. We thus propose to test the following interaction effect:

H2: The frequency of high risk-effort decisions is lower under a fair bonus contract and higher under either a penalty or an unfair contract.

1.2.4 The influence of a prior outcome on risk and effort decisions

So far, we have predicted the effects of both incentive frames in a one-period setting, but decisions are rarely made in isolated time-periods. Introducing time into the decision framework creates new cognitive frames (Benartzi and Thaler, 1999). One aspect of framing is related to the broad versus narrow evaluation of outcomes and its consequences for risk-taking (Thaler, Tversky, Kahneman and Schwartz, 1997). Broad framing means that individuals adopt a long-term horizon and evaluate the outcomes from several periods together. This evaluation renders them less sensitive to interim outcomes. Narrow framing, on the other hand, is when individuals react to a single period outcome that causes excessive risk aversion (Barberis and Huang 2001; Barberis, Huang and Thaler, 2006). Another aspect of framing is how outcomes of prior decisions are integrated into current decision problems. It seems that people value gains and losses differently (gain-loss asymmetry), but the empirical findings are inconsistent (Ahlbrecht and Weber 1997; Green and Myerson, 2004; Shelley 1994). Furthermore, there are different mechanisms of integration between prior gains and losses with subsequent outcomes. Thaler and Johnson (1990) find that an initial loss caused increased risk aversion in a multiperiod gamble if the second choice did not provide the opportunity to break even. A loss after a larger initial gain, on the other hand, is integrated with the gain. Such integration mitigates the loss aversion and facilitates risk-seeking.

In a management accounting setting, Drake and Kohlmeyer (2010) investigate the effect of the past performance history and bonus incentive schemes on managers' framing of current decisions and their risk behavior. They report that individuals with negative past performance are motivated to engage in more risk-seeking behavior than those with positive past performance. Hartmann and Slapničar (2014) study the effect of bonus versus penalty contracts and deferred versus immediate payout in a multiperiod setting. They find higher risk-taking in the penalty scheme, but only in the first period. In the second period, the outcome from the first period outweighs the effect of the incentive design in the deferred payout scheme: a negative prior outcome suppresses risk-seeking and a positive one exacerbates it. The explanation for their different results to those in Drake and Kohlmeyer (2010) may lie in the fact that Drake and Kohlmeyer manipulate past performance as a treatment in a one-period setting, while the participants in Hartmann and Slapničar's (2014) study use two periods: earned an outcome in the first period and could increase or lose it in the second one.

We suggest that in a decision-making context in which prior outcomes help an agent to revise the probability of an outcome, a prior failure will decrease the likelihood of high risk-effort decisions and a prior success will increase it. We thus propose the following hypothesis:

H3: The frequency of high risk-effort decisions is higher following a positive prior outcome than following a negative one.

1.3 Research method

1.3.1 Participants and task

To test the hypotheses, we conducted a three-period, between-subject experiment. All participants were randomly assigned to one of four groups, in which we manipulated two types of incentives contracts (bonus and penalty) and two types of fairness outcomes (fair and unfair). Fair and unfair conditions were manipulated as comparisons of a participants' pay to other participants. A total of 100 undergraduate and graduate students from the Faculty of Economics at the University of Ljubljana participated in the study (74 female, mean age M = 23.5, SD = 3.6, range = 20–44 years; work experience M = 2.1, SD = 3.2, range = 0–20 years). The participants were compensated based on their task performance. The average compensation was 5.05 (SD = 1.89) experimental units (denoted as EU) that translate into EUR 2.525 for 20 minutes of activity, which approximately corresponds to the standard hourly rate for student work. Additionally, for their voluntary participation they were awarded credits for courses. All participants signed a written consent form prior to participating in the research and were informed that they were free to withdraw from the study at any point. The experimental design and the procedures were in accordance with the Helsinki Declaration as revised in 2013[3].

To measure risk-effort decisions, we used a modified Sternberg task (Sternberg, 1966), which is widely used in cognitive psychology to measure cognitive effort (Burrows and Okada, 1973; D'Esposito, Postle and Rypma, 2000; Jansma et al., 2007; Zakrzewska and Brzezicka, 2014). It requires the activation of short-term memory, attention, inhibitory control, and motor control (Kelly et al., 2004; Oberauer, 2001; Vinkhuyzen, van der Sluis, Boomsma, de Geus and Posthuma, 2010). We used a cognitive rather than a physical effort task because in contemporary organizations an increasing number of choices in pursuit of valuable outcomes involve cognitive rather than physical effort, particularly those managerial tasks incentivized by the contracts studied here. We operationalized the risk component by using a multiperiod setting, in which we incrementally increased the task's difficulty and hence the probability of failure.

[3] The Declaration guides ethical principles for medical research involving human participants, but the use of its principles is also rising in social science experiments. The experimental procedure needs to protect the well-being and rights of the participants that are consistent with existing ethical norms in scientific research. The research protocol in medicine must be submitted for approval to the research ethics committee before the study begins. Given the unavailability of such a committee in the Faculty of Economics at the University of Ljubljana at the time the experiment was conducted in November 2014, we obtained a positive opinion of the research ethics committee of the Faculty of Arts (the Department of Psychology) of the same university ex post.

The task was as follows: A series of random letters was presented to the participants on a computer screen. Each letter appeared alone at one second intervals, until the whole sequence was presented (e.g. the letter B, the letter C, then the letter D, to make up the sequence BCD). The sequence of letters then disappeared and a memory maintenance period followed. The end of the maintenance period was signaled by the appearance of a probe letter. The probe letter was one of the letters that was shown in the sequence. Participants were asked to indicate the place of the letter in the sequence by pressing the correct number on the keyboard (e.g., if the letter was second in the sequence, they had to press the number 2). Participants had to correctly respond to all letters in the sequence. After each response, visual feedback was given as to whether the answer was correct (a green square) or incorrect (a red square). Complete silence was maintained during the experiment to ensure the students' concentration.

The task is suitable for the research questions since the performance is easily measurable and does not require prior knowledge. Although there is little possibility of a training effect (Shiran and Breznitz, 2011; Sternberg, 2008), ample evidence reports that incentives can increase cognitive effort and performance via increased attention and proactive cognitive control strategies (Braver, 2016; Chiew and Braver, 2013; 2014; Fröber and Dreisbach, 2014; Jimura et al., 2010; Ličen et al., 2016, Padmala and Pessoa, 2011; Pessoa, 2009). Individuals also vary considerably in their sensitivity to motivational incentives (Ličen et al., 2016). Despite the fact that we calibrated the task's difficulty to each individual's achievement level, we expected that decisions on cognitive effort could vary with increasing bonuses or penalties, the participants' risk inclination, and the framing of incentives but that these effects would not be unlimited. To measure to what extent the participants would be willing to accept high risk-effort decisions, the experiment comprised three rounds in which we incrementally increased the probability of failure, required effort, and the incentives. As it is impossible to design a task in which the outcome would depend on effort with a fixed probability of success, the participants could not exactly evaluate this probability, which is akin to real-life decisions. The task is not very exciting in that it entails a positive cost for effort. After the experiment, the participants complete an exit questionnaire with manipulation checks and demographic questions. None of the participants were excluded based on a misunderstanding of the manipulation conditions[4]. The experiment was Web-based, developed in JavaScript, and designed with HTML and CSS. The data were stored in textual format and analyzed with Stata.

[4] Manipulation check questions: For incorrect answers I could not lose additional money. [True/False] In the first round, probability of failure for the choice of the difficult task was higher than for the choice of the easy task. [True/False]
In the second round, probability of failure for the choice of the difficult task was higher than for the choice of the difficult task in the first round. [True/False] I was paid the same amount of money as my peers for the same performance. [True/False]

1.3.2 Procedure and manipulation

The two manipulation treatments were an incentive contract (bonus vs. penalty) and the outcome's fairness (fair vs. unfair outcome). The bonus contract comprised fixed pay in the amount of two EU plus a bonus for correct responses, while the penalty contract comprised fixed pay (EU 8) minus a penalty for any incorrect responses (see Figure 1). The conditions, the expected total payment in the bonus and the penalty conditions for the same decisions (difficult/easy), were the same after three rounds. Participants in the penalty condition could not end up in negative territory because any losses were covered by a higher initial endowment. The expected payments differ in the interim periods because holding them equivalent would require the use of fixed endowments in each round, which would blur the effect of bonuses and penalties. Nevertheless, the comparison between bonuses and penalties in each round was reasonable because the incentive manipulation that *increases* or *decreases* the payment for correctly or incorrectly solved tasks was equivalent in both conditions in all three rounds. Prospect theory argues that our perceptual system reacts to relative changes and differences (Kahneman and Tversky, 1979; Kahneman, 2003). Accordingly, for the formation of cognitive frames, the relative changes are more important than absolute values.

Participants in the fair outcome condition were told that they would earn the same amount for the same level of effort as their peers; in the unfair condition, they were told that they would earn less. This fairness cue appeared before each session so that they were continuously reminded about the outcome's fairness. At the end of the session we explained that unequal payment was an experimental manipulation that was not actually exercised. That led to participants receiving better payment than expected, so they were not negatively affected by our unfairness manipulation. Paying all participants fairly regardless of the manipulative condition they were coincidentally assigned to is in line with the Helsinki Declaration, according to which research goals should never take precedence over the interests of the participants. It would have not been possible to compellingly manipulate unfairness in any other way. The only alternative to manipulate pay inequity and not deceive participants might be by using hypothetical incentives. Hypothetical incentives are common in psychological research (Hill and Buss, 2010; Wang et al., 2016), but are far less acceptable in the behavioral economics and accounting researches. The use of hypothetical incentives would be less powerful in particular in our design where we not only test the effect of pay inequity but also bonus versus penalty contracts. According to the Helsinki Declaration experimental research can only be conducted if the importance of the research objective outweighs the risks and burdens to the participants. Thus, analyzing the effect of unfairness is such an objective in our field, and we have minimized any potential costs to participants. We explained to the participants that participation was entirely voluntary and that they could withdraw their participation at any point without consequences. None of them exercised that right. Greenberg (1993), Libby (2001), Gächter and Thöni (2010), and Gabaldón, Vázquez Hernández and Watt (2014) treated their participants after unfairness manipulation in a similar way.

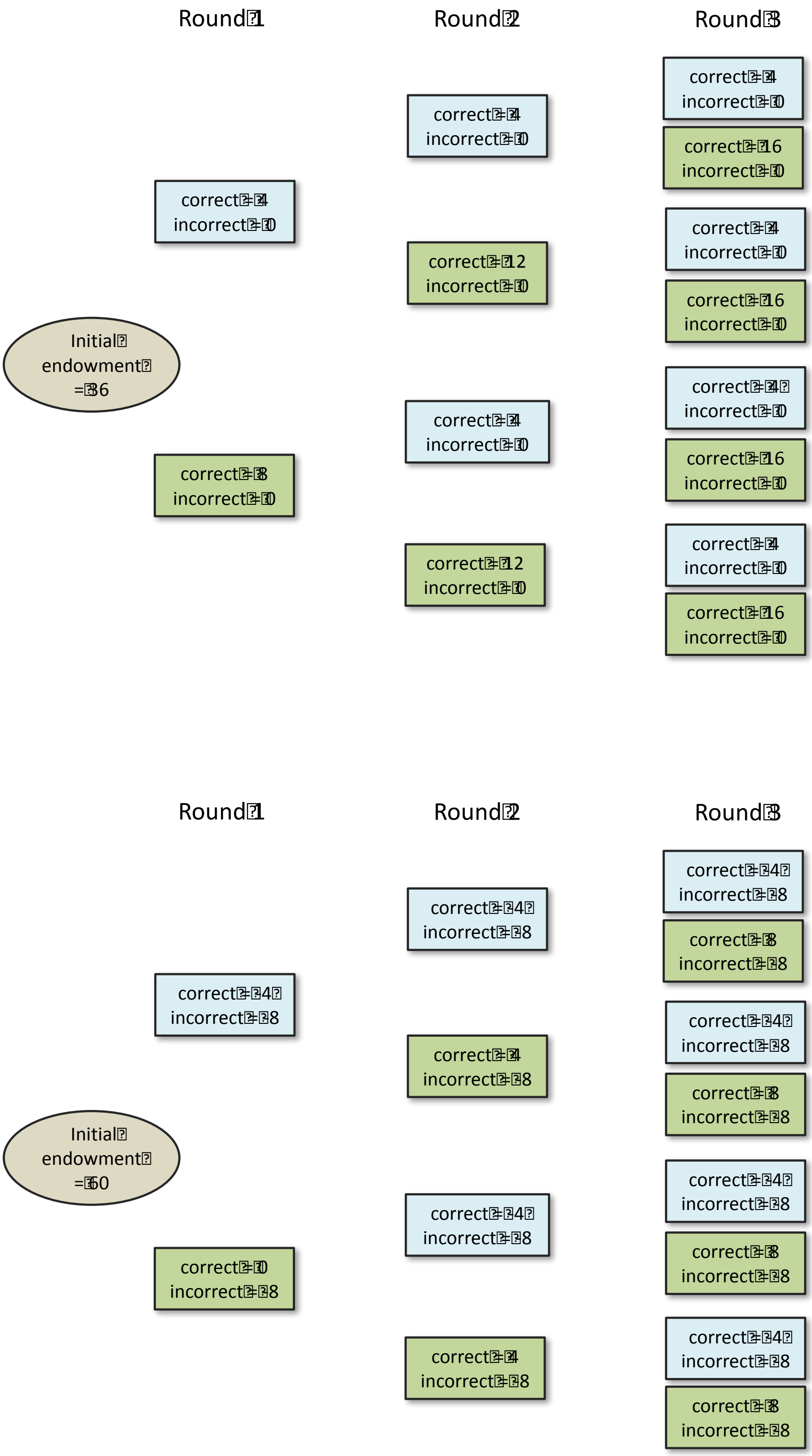
Calibration phase

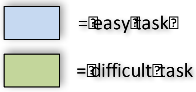
To become familiarized with the task, participants undertook two trial tests during which their performance was not recorded. To calibrate the level of difficulty in the main task, each participant then underwent a phase in which we learned about their ability to memorize the number of letters in a sequence. This phase started with a sequence of three letters. Each sequence was repeated three times. The next sequence contained an additional letter. If two sequences were incorrectly solved, the calibration phase ended and the attained number of letters was coded as the participant's maximum sequence length. Based on each individual's result, his or her difficulty level in the main task was created. We thereby eliminated the effect of an individual's working memory capacity on his or her decisions. For illustration, if one individual could solve five letters and another seven letters, then solving seven letters was not equally difficult for them. After the calibration phase, participants received fixed pay in EU. This was their playing money in the following phase. The participants were also informed of their best result to give them a sense of what for them would be an easy or difficult task and to help them estimate the effort required and the probability of success.

Decision task

In the main task participants had to choose an easy or a difficult task that was defined as the number of letters below or above their performance in the calibration phase. They thereby decided on the effort needed to solve the task and the probability that the task would be correctly solved. The easy task was defined as the sequence that was two letters shorter than their maximum achieved length in the calibration phase and was the same throughout all three rounds. In the difficult task, the sequence length in the first round had the same number of letters as the maximum achieved length in the calibration phase, while in the second and the third rounds the sequence was one and two letters longer, respectively. In the bonus condition, the reward for successfully solving the easy task was smaller than the reward for successfully solving the difficult task. In the penalty condition, the loss for successfully solving the difficult task was smaller than the loss for successfully solving the easy task (see Figure 1).

Figure 1. Decision tree in the risk-effort task for the bonus and the penalty conditions





Note. The number in the ellipse is the initial endowment for each condition. All incentives are expressed in experimental units (EU 1 = EUR 0.5). The participants could choose between the easy task with a sequence length that is two letters shorter than the sequence length achieved in the calibration phase; or the difficult task that is as long as the sequence length achieved in the calibration phase in the first round, has one letter more than the sequence length achieved in the calibration phase in the second round, and two letters more in the third round. The green squares represent the decisions on the easy task and the corresponding rewards or penalties for correctly and incorrectly solved tasks. The blue squares represent the decision on the difficult task and the rewards or penalties for correctly and incorrectly solved tasks.

Control variable

In cognitive tasks, individuals do not only respond to incentives, but they are also driven by their internal needs (Khandekar, 2012). The decision for an easy or a difficult task thus also expresses risk inclination and the need for achievement, both relatively stable personal characteristics. We control for the persistence of decisions with a variable *Prior choice* in the second and the third rounds of the experiment.

1.4 Results

1.4.1 Descriptive statistics

The average sequence length the participants completed in the calibration phase was 5.86 letters (SD = 1.48, median = 6 letters). This result is in line with a general memory span of seven items (range = 5–9 items; Miller, 1956). We found no significant effect of gender, age, years of work experience, or sequence length achieved in the calibration between the treatment groups[5] (Table 1) or the choices made[6]. Nor did we find any significant differences between the total payoffs across the four groups, which indicates a properly designed incentive structure.

[5] Pearson χ^2 tests indicate no significant difference in the frequencies for gender (χ^2 = 0.53, p = 0.912). One-way ANOVA tests indicate no significant differences in group means for age (F = 0.47, p = 0.943), work experience (F = 0.65, p = 0.796), number of letters achieved in the calibration phase (F = 0.66, p = 0.722), and the total payoff (F = 0.73, p = 0.678).

[6] Pearson χ^2 tests indicate no significant difference in the frequencies of the chosen task difficulty for gender in any of the three rounds (χ^2 = 1.53, p = 0.216; χ^2 = 2.29, p = 0.130; χ^2 = 2.02, p = 0.155). Nor do one-way ANOVAs indicate any significant differences in choices in any of the three rounds for age (F = 1.06, p = 0.407; F = 1.15, p = 0.329; F = 0.48, p = 0.937), work experience (F = 0.87, p = 0.582; F = 1.20, p = 0.294; F = 0.79, p = 0.659), and number of letters achieved in the calibration phase (F = 1.49, p = 0.173; F = 1.59, p = 0.140; F = 1.10, p = 0.372).

Table 1. Descriptive statistics by treatment group

	Penalty		Bonus		***Total***
	Unfair	**Fair**	**Unfair**	**Fair**	
N	23	26	24	27	100
Gender					
Female (N)	17	19	19	19	74
Mean	73.9%	73.1%	79.2%	70.4%	74.0%
Age					
Mean	22.52	23.93	24.13	23.41	23.51
(S.D.)	(2.09)	(3.96)	(4.67)	(3.13)	(3.59)
Years of work experience					
Mean	1.96	2.77	1.96	1.56	2.06
(S.D.)	(2.31)	(3.25)	(4.23)	(2.75)	(3.19)
Sequence length achieved in calibration					
Mean	5.52	6.23	5.71	5.93	5.86
(S.D.)	(1.38)	(1.56)	(1.49)	(1.47)	(1.48)
Variable pay (EU)					
Mean	-2.17	-3.23	3.17	2.56	0.11
(S.D.)	(2.41)	(1.82)	(1.88)	(1.25)	(3.38)
Total payoff (EU)					
Mean	5.83	4.77	5.17	4.56	5.05
(S.D.)	(2.41)	(1.82)	(1.88)	(1.25)	(1.89)

Note. Descripitive statistics for each of the manipulation treatments. Sequence length is defined as the maximum length of letters achieved in the calibration phase. Variable pay is the pay earned with chosen tasks. Total payoff is the sum of the variable pay and the initial endowment (EU 2 in the bonus scheme and EU 8 in the penalty scheme). EU 1 is worth EUR 0.5.

In H1 we predicted that the decisions on difficult tasks would be more frequent under a penalty contract. The descriptive statistics are presented in Table 2. The results show that the participants in the penalty condition decided more frequently on difficult tasks throughout all three rounds than those in the bonus condition. In the first round, 63.3% of the participants in the penalty condition decided on a difficult task, which required higher effort and brought about higher risk of failure, as opposed to 47.1% of the participants in the bonus condition. In the second round, the frequency of the decision on the difficult task was 59.2% versus 31.4% for the participants in the penalty versus the bonus condition, respectively. In the third round, it amounted to 53.1% versus 31.4% for the penalty versus the bonus condition, respectively. The results across the fairness condition show that participants in the fair condition in all rounds somewhat more frequently opted for the difficult task, but the differences between the fair and the unfair group are less pronounced. In the first round, the frequency of the difficult task was 60.4% versus 48.9% in the fair versus the unfair outcome condition, respectively. In the second round, the frequency of the difficult task choice became almost equal, 45.3% versus 44.7% in the fair versus the unfair condition, and in the third round, 50.9% versus 31.9%, respectively.

In H2 we hypothesize that decisions for difficult tasks are less frequent under a fair bonus contract and more frequent under either a penalty or unfair contract. A comparison of the four treatment groups (unfair/penalty, fair/penalty, unfair/bonus, fair/bonus) in the first round reveals that in the fair/penalty condition, participants opted for the difficult task in 76.9% of the cases compared to 44.4% in the fair/bonus, 47.8% in the unfair/penalty, and 50.0% in the unfair/bonus condition. Also in the second round the difficult task was undertaken in 73.1% of the fair/penalty condition cases compared to 18.5% in the fair/bonus, 43.5% in the unfair/penalty, and 45.8% in the unfair/bonus conditions. In the third round this pattern no longer held as 61.5% of the participants in the fair/penalty condition opted for the difficult task compared to 40.7% in the fair/bonus, 43.5% in the unfair/penalty, and 20.8% in the unfair/bonus conditions (see Table 2). Overall, the participants who by far most frequently decided on the difficult task in all rounds were those in the fair/penalty condition. This decision was least frequently adopted in the fair/bonus condition in the first and the second rounds, and in the bonus unfair condition in the third round. Figure 2 presents the frequency of the difficult task choice for all group treatments for all three rounds.

Table 2. Choice and outcome by treatment group

		Round 1				**Round 2**				**Round 3**			
		Easy	**Difficult**	% Succesfully solved task (S.D.)	% Succesfully solved difficult task (S.D.)	**Easy**	**Difficult**	% Succesfully solved task (S.D.)	% Succesfully solved difficult task (S.D.)	**Easy**	**Difficult**	% Succesfully solved task (S.D.)	% Succesfully solved difficult task (S.D.)
Penalty	**Unfair**	12	11	73.9%	63.6%	13	10	82.6%	80.0%	13	10	73.9%	40.0%
		52.2%	47.8%	(0.45)	(0.50)	56.5%	43.5%	(0.39)	(0.42)	56.5%	43.5%	(0.45)	(0.52)
	Fair	6	20	50.0%	35.0%	7	19	61.5%	47.4%	10	16	46.2%	12.5%
		23.1%	76.9%	(0.51)	(0.49)	26.9%	73.1%	(0.50)	(0.51)	38.5%	61.5%	(0.51)	(0.34)
	Total	18	31	61.2%	45.2%	20	29	71.4%	58.6%	23	26	59.2%	23.1%
		36.7%	63.3%	(0.49)	(0.51)	40.8%	59.2%	(0.46)	(0.50)	46.9%	53.1%	(0.50)	(0.43)
Bonus	**Unfair**	12	12	66.7%	41.7%	13	11	70.8%	36.4%	19	5	87.5%	60.0%
		50.0%	50.0%	(0.48)	(0.51)	54.2%	45.8%	(0.46)	(0.50)	79.2%	20.8%	(0.34)	(0.55)
	Fair	15	12	63.0%	33.3%	22	5	85.2%	60.0%	16	11	59.3%	9.1%
		55.6%	44.4%	(0.49)	(0.49)	81.5%	18.5%	(0.36)	(0.55)	59.3%	40.7%	(0.50)	(0.30)
	Total	27	24	64.7%	37.5%	35	16	78.4%	43.8%	35	16	72.5%	25.0%
		52.9%	47.1%	(0.48)	(0.49)	68.6%	31.4%	(0.42)	(0.51)	68.6%	31.4%	(0.45)	(0.45)
Total	**Unfair**	24	23	70.2%	52.2%	26	21	76.6%	57.1%	32	15	80.9%	46.7%
		51.1%	48.9%	(0.46)	(0.51)	55.3%	44.7%	(0.43)	(0.51)	68.1%	31.9%	(0.40)	(0.52)
	Fair	21	32	56.6%	34.4%	29	24	73.6%	50.0%	26	27	52.8%	11.1%
		39.6%	60.4%	(0.50)	(0.48)	54.7%	45.3%	(0.45)	(0.51)	49.1%	50.9%	(0.50)	(0.32)
	Total	45	55	63.0%	41.8%	55	45	75.0%	53.3%	58	42	66.0%	23.8%
		45.0%	55.0%	(0.49)	(0.50)	55.0%	45.0%	(0.44)	(0.50)	58.0%	42.0%	(0.48)	(0.43)

Figure 2. The decision for the difficult task of all treatment groups in all three rounds

Note. Lines represent the percentage of choices for the difficult task in all four treatments groups in each round.

In H3 we postulate that decisions for difficult tasks are more frequent following a positive prior outcome than a negative one. The descriptive statistics of the prior choice and the prior outcome are presented in Table 3. Regarding a prior outcome, the results show that 50.8% of those who had successfully completed the task in the first round opted for the difficult task in the second round, whereas only 35.1% of the participants who had incorrectly completed the task in the first round, chose the difficult one in the second round. The results are similar in the third round: 46.7% of those who had correctly solved the task in the second round decided on the difficult task in the third round, while only 28.0% of the participants who had incorrectly solved the task in the prior round decided on the difficult task again. With respect to the prior choice, the results reveal that 60.0% of the participants who had chosen the difficult task in the first round, made the same decision in the second round, whereas only 26.7% of the participants who had previously chosen the easy task decided on the difficult task in the second round. In the third round, 64.4% of the participants who had chosen the difficult task in the second round did so again, while only 23.6% of the participants who decided on the easy task in the previous round chose the difficult one. Overall, the results indicate the effect of a prior outcome and a prior choice on the decision to choose a difficult task.

Table 3. Choice and outcome by prior performance

		Round 2					Round 3				
		Easy	**Difficult**	**Total**	% Succesfully solved task (S.D.)	% Succesfully solved difficult task (S.D.)	**Easy**	**Difficult**	**Total**	% Succesfully solved task (S.D.)	% Succesfully solved difficult task (S.D.)
Prior Outcome	**Negative**	24	13	37	70.3%	38.5%	18	7	25	76.0%	14.3%
		64.9%	35.1%		(0.46)	(0.51)	72.0%	28.0%		(0.44)	(0.38)
	Positive	31	32	63	77.8%	59.4%	40	35	75	62.7%	25.7%
		49.2%	50.8%		(0.42)	(0.50)	53.3%	46.7%		(0.49)	(0.44)
Prior Choice	**Easy**	33	12	45	77.8%	41.7%	42	13	55	76.4%	15.4%
		73.3%	26.7%		(0.42)	(0.51)	76.4%	23.6%		(0.43)	(0.38)
	Difficult	22	33	55	72.7%	57.6%	16	29	45	53.3%	27.6%
		40.0%	60.0%		(0.45)	(0.50)	35.6%	64.4%		(0.50)	(0.45)
Total		55	45	100	75.0%	53.3%	58	42	100	66.0%	23.8%
		55.0%	45.0%		(0.44)	(0.50)	58.0%	42.0%		(0.48)	(0.43)

1.4.2 Results of the model estimation

The results on the model estimation are presented in Table 4. In a logistic regression where the independent variables are dichotomous, a true main effect is that of a variable across all of the observed levels of the other variable. It is only meaningful when there is no interaction, that is, when the effect of the first variable is similar at each level of the second variable. When there is significant interaction, a main effect is not unambiguously interpretable. A nonsignificant main effect of one variable could mean either that its effect is truly zero at both levels of the second variable, or that its effect is positive at one level of the other independent variable and negative at the other. It is normal, therefore, not to regard main effects as informative in of themselves when an interaction occurs. We therefore analyze the effect of each incentive variable at both levels of the second variable via a pairwise comparison. This analysis more clearly explains the model itself and provides significance levels for all four treatments.

Test of H1 and H2

In the first round (see Round 1, Table 4), we observe a significantly positive effect of fairness on the penalty condition, b = 1.29, $F = 2.05$, $p = 0.040$. In the bonus condition, the effect of fairness is insignificant, b = -0.22, $F = -0.39$, $p = 0.693$. Furthermore, the bonus/penalty coefficient is insignificant and close to zero in the unfair treatment, b = 0.09, $F = 0.15$, $p = 0.882$; whereas in the fair condition, its effect is significant and negative, b = -1.43, $F = -2.34$, $p = 0.019$, which indicates that the penalty positively influences the choice of the difficult task. The interaction term in the first round is negative, b = -1.51, $F = -1.79$, $p = 0.073$ but only marginally significant. The marginal significance is reflected in the significantly different results for only two groups. The marginal effects show how the probability of choosing the difficult task changes if the fairness condition changes from unfair to fair and from penalty to bonus.

The results of the second round (see Round 2, Table 4) are consistent with the first round: the outcome's fairness has a significant and positive effect on the penalty condition, b = 1.54, $F = 2.13$, $p = 0.033$, and a significantly negative effect on the bonus condition, b = -1.52, $F = -2.16$, $p = 0.031$. These results explain why the descriptive statistics show almost no difference in the frequency of choices for the difficult task between the fair and unfair conditions as the two effects cancel each other out (see Figure 3). Fairness has the opposite effect on the penalty and the bonus groups, whereas in the unfair condition the effect of the bonus and the penalty are roughly the same. The bonus/penalty coefficient is insignificant and close to zero in the unfair condition, b = 0.37, $F = 0.54$, $p = 0.589$; while in the fair condition, it is significant and negative, b = -2.69, $F = -3.55$, $p < 0.001$, which confirms the effect of the penalty from the first round only if the payout is fair. The interaction term is significantly negative, b = -3.06, $F = -3.07$, $p = 0.002$, which shows the reversal of the effects in both conditions. These results indicate that, in contrast to the descriptive statistics and H1, there is no main effect from the penalty, as it differs from the bonus only in the fair condition. In H2 we predict that either the penalty or unfairness will elicit decisions for the difficult task: the descriptive statistics speak for such an interpretation and so do the significant interaction term and the pairwise comparisons.

Test of H3

In H3 we predict that a positive prior outcome positively affects the choice on a difficult task. Rounds 2 and 3 in Table 4 present the analysis of the effects of the success the participants had in completing the selected task (outcome Round 1) after controlling for the choice from the first round (choice Round 1). In round 2 we find highly significant effects (the coefficient for outcome Round 1 is b = 2.83, $F = 3.17$, $p = 0.004$; for a choice in round 1 b = 2.85, $F = 3.52$, $p < 0.001$), whereas the effect of our incentive variables remained comparable to those in round 1.

Interestingly, in the third round (see Round 3, Table 4) the incentive variables become insignificant. A pairwise comparison shows that fairness is significant in only one of the four groups (in the fair/bonus condition the coefficient is significant and positive b = 1.87, $F = 2.38$, p = 0.017), but on the whole, decisions on the difficult tasks in the fair/bonus treatment still do not exceed those from the fair/penalty treatment. The interaction term is insignificant and close to zero, b = 0.97, $F = 0.85$, p = 395), which reflects similar differences in the bonus and penalty conditions for unfair compared to fair treatment (see Figure 3, Round 3). These results show that the choice of the difficult task can only be explained by the prior outcome: b = 3.40, $F = 3.68$, $p < 0.001$ and the prior choice, b = 3.80, $F = 4.65$, p < 0.001. If a prior outcome is positive (i.e., if participants are successful in whatever task they chose in the prior round), it positively influences the decision for the same level of difficulty in the next round. This is in line with the results reported by Thaler and Johnson (1990) and Hartmann and Slapničar (2014). The participants to a large extent preserve their initial choice throughout the three rounds. There are two alternative explanations for no effect from the incentive variables: either a prior outcome or a prior choice outweighs the effect of the incentive schemes and the outcome’s fairness, or a prior choice picks up a part of the effect of the manipulations and weakens the coefficients for the bonus/penalty and fairness, which means that our manipulations lose power. The increase in explanatory power of the model after the inclusion of the prior outcome and the prior choice speaks for the first interpretation (from Pseudo R^2 0.065 to 0.336, for brevity only the comprehensive model is reported). The participants in the third round seem to have no longer cared about their incentive scheme or its fairness. What they relied on most was the assessment of the probability of success based on their abilities and their inherent need for achievement, which is reflected in the consistency of the choices throughout the experiment. This corroborates the proposition of Thaler and Johnson (1990) who argue that in intertemporal settings, the prior outcome shifts the individuals’ reference points and changes their risk perceptions. The effect of the prior choice can be understood as a sign of the persistence in the choices over time and stable personality traits.

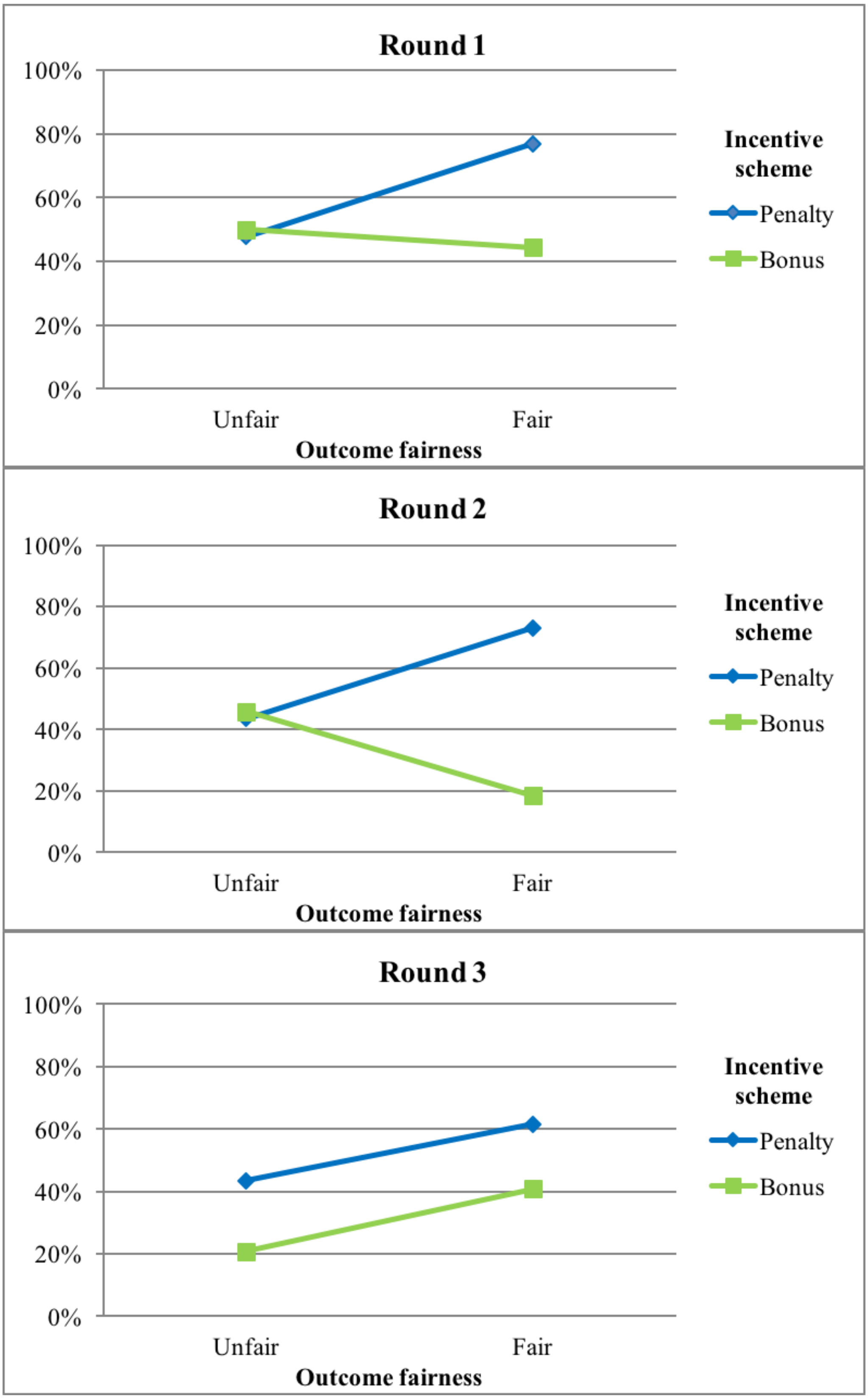
Table 4. Logistic regression predicting the decision for the difficult task

	Round 1		Round 2		Round 3	
	Model 1	**Mfx**	**Model 2**	**Mfx**	**Model 3**	**Mfx**
Intercept	-0.09 (-0.21)		**-3.79****** (-3.43)		**-4.82****** (-4.19)	
Bonus (1) / Penalty (0)	0.09 (0.15)	0.02	0.37 (0.54)	0.09	-1.03 (-1.12)	-0.25
Fair (1) / Unfair (0)	**1.29*** (2.05)	**0.31***	**1.54*** (2.13)	**0.36***	0.91 (1.09)	0.22
Bonus / Penalty * Fair / Unfair	-1.51[a] (-1.79)	**-0.36***	**-3.06**** (-3.07)	**-0.56******	0.97 (0.85)	0.24
Prior Outcome - Incorrect (0) / Correct (1)			**2.83**** (3.17)	**0.57******	**3.40****** (3.68)	**0.57******
Prior Choice – Easy (0) / Difficult (1)			**2.85****** (3.52)	**0.60******	**3.80****** (4.65)	**0.74******
Wald χ^2	6.45		22.49		27.94	
χ^2	0.092		0.000		0.000	
Pseudo R^2	0.053		0.305		0.336	

Pairwise Comparison	**Round 1**	**Round 2**	**Round 3**
Bonus / Penalty * Unfair (0)	0.09 (0.15)	0.37 (0.54)	-1.03 (-1.12)
Bonus / Penalty * Fair (1)	**-1.43*** (-2.34)	**-2.69****** (-3.55)	-0.06 (-0.09)
Fair / Unfair * Penalty (0)	**1.29*** (2.05)	**1.54*** (2.13)	0.91 (1.09)
Fair / Unfair * Bonus (1)	-0.22 (-0.39)	**-1.52*** (-2.16)	**1.87**** (2.38)

Note. * $p < 0.05$; ** $p < 0.01$; ***$p < 0.001$, [a] p=0.073; Significant coefficients less than 0.05 appear in bold. Mfx stands for marginal effects.
Pairwise comparison presents the effect of each incentive variable at both levels of the second variable: e. g. in the first line of the first round in which contracts are unfair, the effect of bonus does not differ from the effect of penalty. In the second line in which contracts are fair, the effect of bonus is significantly negative in comparison to the effect of penalty, indicating that penalty positively influences the choice of the difficult task.

Figure 3. The decisions for the difficult task according to incentive scheme and outcome fairness for round 1, round 2, and round 3



Note. The green and blue line represent the percentage of choices for the difficult task in the bonus and the penalty condition respectively in the fair and the unfair conditions.

To summarize, the results show that joint risk-effort decisions in time are influenced by all three frames: in the first two rounds both incentive scheme frames interact; but with increasing risk and effort, a prior outcome and a prior choice become dominant reference points. We find that in a multiple frame setting, the effect of the penalty condition cannot be validly observed on a standalone basis. As implied in H2, a single cue creating a loss perception is enough to significantly change the inclination toward effort and risk. We find that the fair/penalty condition most frequently elicits the decision on a high risk-effort task. The unfair/bonus condition is about the same as the unfair/penalty condition in eliciting high risk-effort decisions and more prompting than the fair/bonus condition. The latter may represent a comfort zone, which does not create a need to accept challenges. However, a penalty contract and unfair payment may no longer be perceived as loss domains if one is persuaded of one's own abilities and is confident of success in the task with a high payoff. This is already evident in the second round, and in the third round it overrides the incentive effects.

1.5 Discussion and conclusion

We investigate the effects of two incentive features – a bonus versus a penalty and an outcome's fairness – on the assumption that via cognitive framing, they influence risk and effort decisions. Our predictions are based on two influential theories: prospect theory (Kahneman and Tversky, 1979; Kahneman, 2003) and organizational justice theory (Adams, 1963; Akerlof and Yellen, 1990). These theories share a common feature: a reference point against which people determine the utility of their outcome. Both theories have been extensively tested and convincingly confirmed, but only for a single behavioral outcome – either risk-taking or exerting effort. We tested their effects on behavior that is usually beyond the scope of each theory and in a multiperiod setting. We were interested in whether social comparisons matter for risk-taking and whether gain and loss domains matter in exerting effort. Our more specific research question addressed how these frames combine in joint decisions about risk and effort. Prior empirical evidence is inconsistent showing that a variety of incentive system characteristics may invoke the same cognitive frames (i.e., gain and loss domains). But more importantly, there is no evidence on the interactive or additive effects of various frames, and whether the cues are consistent or conflicting.

The interaction effect that we find indicates that both incentive features might invoke the perception of a loss domain, which in turn stimulates risk and effort. If the level of payment is unfair, neither a bonus nor a penalty seems to matter. If individuals see that they earn less than their peers, they perceive it as a loss. The difference in frequency of high risk-effort decisions between the unfair bonus and the unfair penalty conditions is negligible. This finding contributes to the evidence documented by Hartmann and Slapničar (2012a) that the effect of distributive justice outweighs the effect of procedural justice on intrinsic motivation. In other words, people care more about how fair their pay is compared to others than how their pay is determined. If in contrast, payment is fair, then a shift from a bonus to a penalty contract significantly increases risk-taking and effort. In the fair/bonus condition the frequency of high risk-effort decisions is radically lower than in the fair/penalty condition. In the second round, the interaction effect becomes more pronounced and statistically significant. All in all, when the cues are consistent, they do not add up; when conflicting, they do not cancel out but a loss cue prevails. Fairness and bonus/penalty framing may thus be seen as complementary, in that loss perception and elicitation of risk taking behavior, accompanied with higher effort, may be created by either.

In spite of the results in this and prior studies on the effect of penalty contracts on effort and the productivity of firms (Hossain and List, 2012), in practice bonus contracts are much more widely spread than penalty contracts. This pervasiveness can be partly explained by employees' preferences (Frederickson and Waller, 2005; Hannan et al., 2005; Luft, 1994). To accept a penalty contract, they would require higher payment (Hannan et al., 2005). Respecting employees' preferences leads to positive effects in principal-agent relationships. These positive effects can explain the findings of Christ et al. (2012) that bonus contracts stimulate higher effort on tasks not governed by the contract, for which there is no performance pay at all. Many critical managerial tasks and targets that cannot be completely predefined in a contract fall into this domain, particularly because of task uncertainty, that makes preset tasks and explicit targets less controllable and relevant (Hartmann, 2005; Lau and Moser, 2008; Hartmann and Slapničar, 2012b). The third reason for a wider use of bonus contracts may be in the effect of penalty contracts on (excessive) risk-taking. However, overall, contract characteristics that invoke a perception of a loss remain under-researched, and the evidence is inconsistent.

Although penalty schemes are not very common in the nonfinancial sector, in the European Union they have been introduced by a recent banking regulation (CRD III) as a deferred bonus system that accommodates a potential penalty and a clawback. It may well be that such penalty clauses are one step closer to fairer compensation because managerial pay will become more closely aligned with the long-term performance of the firm. But the effect on risk-taking may not be what the regulators hoped for, particularly if the penalty is offset by some other forms of compensation. As found by Hartmann and Slapničar (2014), if the potential of a penalty is offset with high upside compensation, it exacerbates risk-taking. In the design of a managerial incentive scheme, it is therefore very important to consider the balance between the number of tasks that can be governed by a contract and those that cannot, and the alignment of an incentive scheme with the risk appetite of an organization. In the case of penalty contracts other mechanisms should also be in place to prevent excessive risk-taking (such as the use of appropriate and balanced performance measures, including risk measures, prudential monitoring, and an independent risk management function).

The results of the third round of the experiment show that the choice of a high risk-effort task over time becomes associated with an evaluation of the probability of success, which depends on one’s past performance. This evaluation may develop as a new reference point that delineates between the gain and loss domains in which the individual estimates the probability of winning or losing (Kahneman, 2003). With continuous success, such a reference point may prompt the adoption of increasingly demanding and risky behaviors. Rising reference points may have occurred in the financial and other sectors before the financial crisis. Fueled with prior success, managers took increasingly risky decisions. Another interesting finding is the persistence of choices throughout the rounds. It raises a question about the effectiveness of incentive schemes in the long run. Can incentives systematically influence risk and effort decisions over extended periods of time, or do people predominantly rely on their own abilities regardless of the incentive scheme? Do personality traits outweigh incentives in determining risk propensity, preparedness to exert effort, and the willingness to undertake demanding tasks in the long run? These questions could be answered by studying further periods and the effect of personality traits. It remains debatable whether organizations want to stimulate high risk-effort behavior in the long run or whether this would have too many negative consequences, such as excessive performance volatility, increased stress levels in employees, and damage to the superior-subordinate relationship.

A limitation of this study is inherent in the experimental method. The results of an experiment are not to be taken at face value: Had we varied the intensity of the penalty and bonus and the fairness manipulation, the relative relation between the four experimental groups might be different. However, our results are relatively robust and tested over three rounds of the experiment in which we intensified the rewards and the difficulty of the task; we thus believe that our main findings remain intact. Future studies could test the phenomenon in other contexts through survey methods. Another limitation of the study is that we do not control for the personality traits of the decision-makers and their autonomous motivation (Kunz, 2015). One such personality characteristic is an individual’s need for achievement (McClelland, 1987). The participants with a stronger need for achievement may be more risk and effort prone. This inclination may be evident in the significant effect of the prior choice, which we control for. The use of students is not a limitation per se as the Sternberg task is purely cognitive. Future research could also aim to develop a more comprehensive cognitive theory on risk-effort behavior, most likely involving the more advanced research methods that are emerging in neuroscience. The cognitive effects of incentives remain a fertile area for future research with great potential for practical applicability.

INCENTIVE SENSITIVITY TO ACHIEVEMENT MOTIVATION IN COGNITIVELY CHALLENGING TASKS

Abstract

Whereas many motivation theories assume that individuals evaluate incentives similarly, Gray and McNaughton (2000) propose considering individuals' motivational dispositions to incentives. Such pre-existing motivational dispositions influence individuals' perception of situations as being rewarding or punishing. In this chapter we evaluate the effect of monetary incentives in choices to perform cognitively challenging tasks given individual differences in achievement motivation. We designed a three-period, within-subject experiment in which we measured willingness to engage in a cognitively challenging task (Sternberg 1966) in a non-incentivised and an incentivised setting (with rewards or penalties). We find that that as cognitive tasks become increasingly difficult rewards have a diminishing effect on the choice of a challenging task, significantly more so for high achievement individuals. Penalties can increase selection of challenging tasks at low and moderate task difficulty, but their effect is mitigated for individuals with high motivation to avoid failure. At high task difficulty penalties had no significant main or interactive effect: only those with a low fear of failure are prepared to undertake a challenging task. The study advances the motivation theory and practice by an original analysis of the interplay among monetary incentives, achievement motivation and task difficulty.

Key words: task choice, achievement motivation, monetary incentives, cognitive effort

JEL codes: J33, G32

1.6 Introduction

The success of modern organisations depends critically on their employees' predisposition to successfully engage in cognitively challenging tasks, i.e. developing new business models, technologies, products or services, which may be far more uncertain than the usual tasks. Challenging tasks are conceptualized as problems to solve and choices to make under conditions of risk and uncertainty in dynamic settings (McCauley et al., 1999, p. 4). Although incentives are widely viewed as the motivating force of human effort, it is intriguing that compensation packages of professionals most commonly faced with such tasks, i.e. designers, engineers, scientists, rarely include any substantial performance-contingent pay. Do *ex ante* announced performance-contingent incentives discourage or stimulate engagement in cognitively challenging tasks that are reasonably attainable? Or do they encourage excessive risk taking by enagagement in tasks that are beyond an individual's capability? The research evidence on the effectiveness of monetary incentives for *cognitive* performance is inconclusive. In their review paper, Bonner, Hastie, Sprinkle and Young (2000) report that in around 50% of experiments monetary incentives had no effect on performance in cognitive tasks. As stressed by the authors, monetary incentives can become less effective when there is a growing gap between a task's difficulty and the skills needed to successfully complete it.

The idea advanced by Gray and McNaughton (2000) is that individuals' motivational disposition influences their perception of a situation as being rewarding or punishing. The source of an action may derive from motivation to achieve or to avoid a failure (Atkinson, 1957; Deci and Ryan, 1985; Heckhausen, 1963, 1991). Some individuals are highly motivated in a cognitive task itself because they are inborn high achievers, while others act in an attempt to avoid failure.

In this chapter we are interested how individuals with varying levels of achievement motivation and failure avoidance react to rewards and penalties when deciding about engaging in a challenging task. We would also like to understand how increasing task difficulty with which successful tasks completion becomes less and less certain moderates the relation between incentives and achievement motivation. People's reaction to uncertainty depends on whether incentives are framed as rewards or penalties, so we examine these two types of incentives separately.

To answer these research questions, we conducted an experiment in which we measured willingness to engage in a challenging task, i.e. the modified Sternberg task (Sternberg, 1966). The experiment has a three-round, within-subject design. We observed choices of challenging tasks in a series of rounds with increasing task difficulty in which rewards were absent (a non-incentivised setting) or present (an incentivised setting). In the incentivised setting, the incentive was framed either as a reward or a penalty. We used the cognitive effort discounting paradigm (Botvinick et al., 2009; Kool et al., 2010; Westbrook et al., 2013) in which participants selected between a challenging task for a larger monetary reward or a non-challenging task for a smaller monetary reward in economically equivalent incentive settings. Motivational orientation was measured with the revised Achievement Motivation Scale (AMS-R) (Lang and Fries, 2006). The experiment was tested on 147 students.

Our results suggest that as cognitive tasks become increasingly difficult monetary incentives have a diminishing effect on the choice of a challenging task. Whereas in the non-incentivised setting achievement motivation clearly explains choices of challenging tasks, incentives motivate people with different motivational predisposition differently. In the incentivised setting, at low and medium task difficulty rewards have no impact on willingness to select a challenging task. At high task difficulty rewards have a negative impact on engaging in a challenging task, significantly more so for high achievement individuals. Penalties can increase selection of challenging tasks at low and moderate task difficulty, but their effect is weaker for individuals that already have a high motivation to avoid failure.

Our study's most important contribution is in providing original evidence on incentive sensitivity to achievement motivation. Rewards and penalties create distinct cognitive frames in which individuals differently assess outcome risks and choose their risk behaviour accordingly. These tendencies have basis in individuals' motivational dispositions that determine how they react to potential rewards and penalties. We analyse the interplay of incentives and achievement motivation on cognitively challenging tasks in one experiment. Despite being complex only such an analysis is akin to real life situations, in which such decisions are taken. Bivariate analyses of these relationships have led to controversial results. Alone the findings on the effect of incentives on cognitive performance are polemic and do not offer an unambiguous answer as to whether cognition (e.g. increased working memory, improved reasoning, better cognitive planning, flexible adapting of behaviour to upcoming cues etc.) can be stimulated by monetary incentives. Second, earlier research rarely analysed the effects of incentives on engagement in cognitive tasks for increasing task difficulty. We find a different effect of incentives at every level of task difficulty. The studies that do not account for varying task difficulty and simply measure it at one point without calibrating it for each individual, observe the incentive–effort relationship on a limited interval. Increasing task difficulty brings about a decreasing probability of success and thus introduces the risk-behaviour aspect in these decisions. One of our most important findings is that rewards curb engagement in overly challenging tasks, whereas penalties are effective only at lower level of difficulty. We believe that the chapter contributes to a cross-disciplinary view that advances our understanding of the functionality of incentives (Merchant, Van der Stede and Zheng 2003) and to the growing body of accounting studies that examines the effects of incentive framing on behavioural outcomes and choices (e.g., Brown, Farrington and Sprinkle, 2016;

Brink & Rankin 2013; Church et al., 2008; Christ et al., 2012, Hannan et al. 2005; Hartmann and Slapničar, 2015; Luft 1994, Oblak et al., 2017).

The remainder of the chapter is organised as follows. In section two, the theoretical background for developing our hypotheses is presented. In section three, the experimental design and the procedure are described. Section four includes an outline of the results while in the last section we provide a general discussion of the results and their implications for practice and set out some limitations of the study.

1.7 Theoretical background and hypotheses

1.7.1 Achievement motivation

One of the main reasons people differ in their inclination to undertake cognitively challenging tasks is their innate need for achievement. The concept of achievement motivation dates back to Murray in 1936. It was later elaborated by McClelland and his co-authors (1953) who defined it as "*the strength of the motive to achieve success relative to the motive to avoid failure*". Achievement motivation as a personality disposition is generally divided into two tendencies, approach and avoidance tendency. These two tendencies are labelled as motivation to achieve (MTA) and motivation to avoid failure (MTF) (Atkinson, 1957; Heckhausen, 1991). Each of the two tendencies is associated with certain behavioural patterns and the anticipation of certain feelings. The MTA concept implies that feelings such as pleasure and pride are associated with goal achievement of moderate challenge, while the MTF concept suggests the anticipation of a negative affect (i.e. fear and unpleasantness) in the same circumstances (Nygard, 1981). Individuals with high motivation to achieve invest greater effort in a task and have higher cardiovascular reactivity (Beh, 1990). They direct their energy into their work to obtain professional and personal goals (Atkinson and Raynor, 1974; McClelland et al., 1953), they perform better (Heckhausen, 1991), are more persistent (Spangler, 1992) and like to work on an achievement-related task (Heckhausen, 1991). Individuals with high achievement motivation set more "realistic" goals congruent with their capacity (Kilch and Feldman, 1992; Weiner, 1982). As stated by Locke (1991), the feeling of achievement promotes higher performance and, according to Kukla (1972), high achievers usually perceive themselves as capable.

Individuals with high motivation to avoid failure tend to keep away from decisions that may bring about failure because they associate it with embarrassment, shame and humiliation (Atkinson and Raynor, 1974; McClelland et al., 1953). They are in a state of anxiety when working on an achievement-related task (Heckahausen, 1991) and set themselves less realistic and challenging goals (Birney, Burdick, and Teevan, 1969; Heckhausen, 1963, 1991). They experience negative feelings when dealing with difficult tasks, but not when dealing with easy tasks. This leads to their less frequent engagement with difficult tasks (Gendolla and Wright, 2005; Wright and Kirby, 2001).

1.7.2 Task difficulty and achievement motivation

Achievement motivation varies with task difficulty. As argued by Atkinson (1964), achievement motivation influences decisions only if the task is challenging, neither too easy nor too difficult. Individuals with high MTA consider easy tasks boring. Such tasks do not make them realise a positive affect through task mastery. They also feel aversion to very difficult tasks in which failure is almost certain and thus do not offer a chance for task mastery. Atkinson proposes an *inverse* U-shaped motivation curve for high MTA individuals: the positive incentive value of success increases with the task difficulty, but is multiplied with decreasing expectancy of success. The product of the two is maximised at medium difficulty. The curve is steeper for high-achievement individuals and flatter for low-achievement individuals. Atkinson also constructed a case for motivation to avoid failure: the negative incentive value of failure decreases with the task difficulty as it is more embarrassing to fail in an easy task than in a difficult one. Expectancy of failure rises with task difficulty. The product of both has a U-shaped function in which the choice of the medium-difficult task is the least likely. In line with the MTA logic, for individuals with a high MTF the slope of the function is steeper than for those with a low MTF. While empirical research has largely supported the inverse U-shape curve for MTA, there has been very little empirical support for the predicted U-shape curve for MTF proposed by Atkinson (Schultheiss and Brunstein, 2005). Other authors reported the least likely involvement in highly difficult tasks for high MTF individuals (Gendolla and Wright, 2005; Wright and Kirby, 2001; Heckhausen, 1963, 1991).

The achievement motive may work on its own, but it may also be triggered by a specific situation such as a reward, expectation or demand (Heckhausen, 1991; Lang and Fries, 2006). An interesting question is how it is altered if different types and magnitudes of incentives are offered.

1.7.3 Incentive sensitivity to achievement motivation under increasing task difficulty

Whereas many motivation theories assume that individuals evaluate incentives similarly, Gray and McNaughton (2000) and Corr, McNaughton, Wilson, Burch and Poropat (2017) propose considering individuals' motivational dispositions to incentives. Such pre-existing motivational dispositions influence individuals' perception of situations as being rewarding or punishing. Moreover, incentivising challenging cognitive tasks introduces a new dimension into decision-making: individuals' risk inclination related to their assessment of the probability of success relative to cognitive effort. Factors affecting the decision to select a more challenging task depend on the subjective value of the potential reward, the anticipated effort required to accomplish the task and the probability the reward will be received if the task is successfully completed (Botvinick et al., 2009). The studies that do not examine the effect of monetary incentives for increasing task difficulty provide only limited understanding of the relationship.

When studying monetary incentives, achievement motivation and task difficulty in combination, it is necessary that rewards and penalties under uncertainty are analysed separately as they create different cognitive frames and elicit different risk behaviour (Kahneman and Tversky, 1979). According to prospect theory in a *gain* domain people systematically exhibit risk averse behaviour (Kahneman and Tversky, 1979; Kahneman, 2003). They prefer smaller more certain outcomes over larger but uncertain ones. In a *loss* domain individuals are generally loss averse: they are willing to accept risky choices to avoid losses. We assume that a similar logic applies to choices of cognitively challenging tasks. As suggested by neuroscientific literature, choices of effort are associated with choices of risk (reviewed in Salamone et al., 2012; Miller et al., 2013), as neural networks activated in both types of decisions to a certain extent overlap. Let us first turn to rewards.

1.7.4 Rewards and MTA

As suggested by previous literature in a non-incentivised setting in which individuals do not take financial risk into consideration, their willingness to choose a challenging task depends on their achievement motivation, which has the strongest effect at medium task difficulty. We propose that incentivizing changes this relationship, whereby high and low achievers are not equally affected at all levels of task difficulty because of varying probability of reward and their different motivation.

At low task difficulty[7] the willingness of high and low MTA individuals to engage in a challenging task is not that different – it is low for both. We therefore assume that relatively certain rewards are likely to increase motivation and positively affect the selection of a challenging task. The effect is likely to be about the same for both types of personalities. We propose to test the following hypothesis:

H1: At low task difficulty[8], rewards have a positive effect on willingness to select a more challenging task for both, low and high MTA individuals.

With moderate task difficulty at which the positive effect of achievement motivation is the strongest and perceived success probability still high, rewards are likely to have small

[7] Please note that we refer to a *challenging* task as opposed to an *easy* task with respect to choices to be made, and to *task difficulty* with respect to the difficulty level of the challenging task whereby it increases from low, to medium and high difficulty.

[9] Manipulation check questions: For incorrect answers I could not lose additional money. [True/False]; In the first round, probability of failure for the choice of the difficult task was higher than for the choice of the easy task. [True/False]; In the second round, probability of failure for the choice of the difficult task the was higher than for the choice of the difficult task in the first round. [True/False];

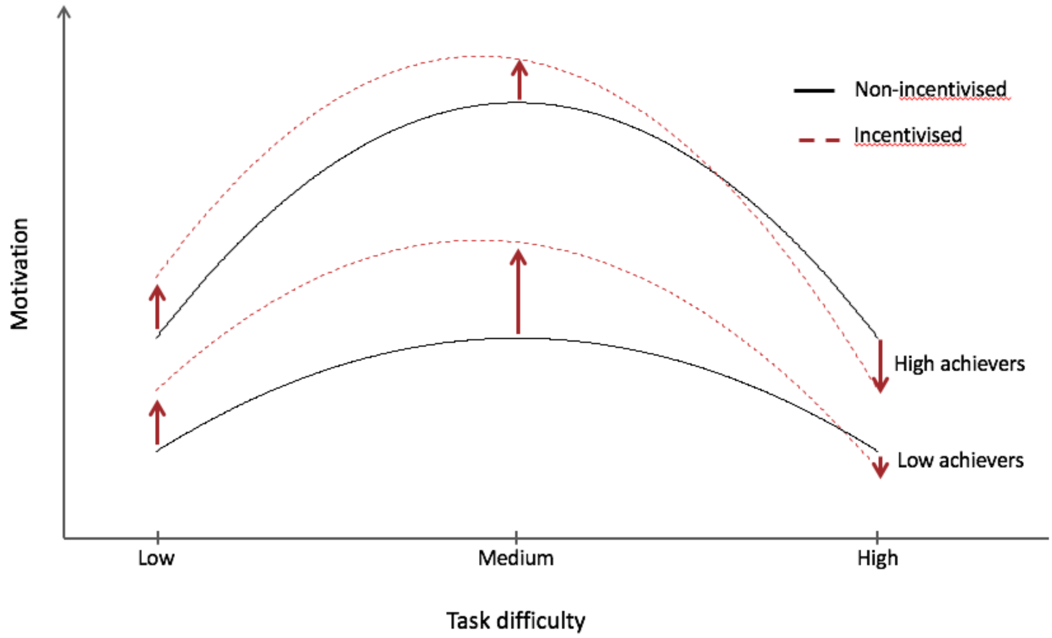
additional effect on the willingness to choose a challenging task for high achievers. The effect of rewards is limited due to cognitive constraints: if individuals are already highly motivated to take on challenging tasks and perform well, extrinsic rewards cannot significantly intensify their cognitive effort and improve performance (Ličen et al., 2016). While low MTA individuals are not so motivated to engage in a challenging task as high MTA individuals, we assume that the effect of a reward will be positive and stronger for low MTA individuals than for high MTA individuals. We thus propose the following hypothesis:

H2: At medium task difficulty, rewards have a positive effect on willingness to select a challenging task, which is stronger for low MTA individuals than for high MTA individuals.

With high task difficulty at which achievement motivation is lower, risk aversion is likely to prevail and higher yet riskier rewards are likely to have an adverse effect on the selection of a difficult task (Kahneman and Tversky, 1979). We assume that offering an uncertain reward to those high MTA individuals that would in a non-incentivized setting still engage in a more challenging task because of their genuine interest and curiosity, would discourage them. Instead of being faced with pure cognitive challenge they face a financial decision between a smaller certain reward for the choice of an easier task and a larger riskier reward for a challenging task. A failure in a non-incentivized setting is only a failure at the task, a failure in an incentivized setting is a failure coupled with the loss of the reward. As low MTA individuals are less interested in highly difficult tasks in the first place, the negative effect of the reward will not be as strong for them as for high MTA individuals. We propose the following hypothesis:

H3: At high task difficulty, rewards have a negative effect on willingness to select a challenging task, which is stronger for high MTA individuals than for low MTA individuals.

Figure 4. The relation between rewards and task difficulty



1.7.5 Penalties and MTF

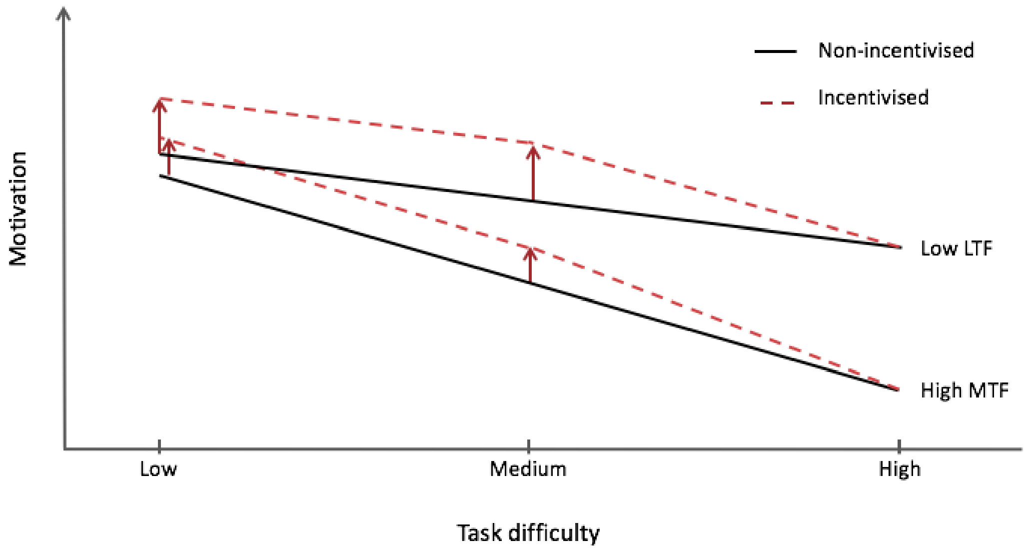
The majority of papers find that penalties elicit greater effort than if no incentives or rewards are offered (Church, Libby, and Zhang, 2008; Gose and Sadrieh, 2012; Hossain and List, 2012). The explanation is that penalties induce *a loss domain* in which people try to avoid losses by exerting a higher effort and taking on more risk (Kahneman and Tversky, 1979). The higher the potential loss, the more people want to avoid it with effort exertion. People will thus be ready to accept the outcome risk if it compensates for a loss, but only as long as successful completion is seen as probable. With high task difficulty it may no longer be so.

Individuals with high motivation to avoid failure are known to avoid decisions that could result in failure because they associate such an outcome with negative feelings, embarrassment or humiliation (Atkinson and Raynor, 1974; McClelland et al., 1953). If penalties are used to stimulate decisions for a challenging task, they are likely to be less effective for high MTF individuals than for low MTF individuals. Both start weighing between negative feelings in case of failure at task completion and the fact it could be avoided with the choice of a less difficult task. Since high MTF individuals fear failure and embarrassment more than low MTF individuals, they have weaker responsiveness to the penalty and will less likely engage in a challenging task because of the penalty. We assume that at low and moderate task difficulty penalties will have a stronger positive effect on the choice of the challenging task on low MTF individuals than on high MTF individuals. In highly difficult tasks the willingness to choose a challenging task will be low for both, because at high difficulty level there is a small chance to avoid a loss.

H4: At low and moderate task difficulty, penalties have a positive effect on willingness to select a challenging task, which is stronger for low MTF individuals than for high MTF individuals.

H5: At high task difficulty, penalties have no effect on willingness to engage in cognitive effort.

Figure 5. The relation between penalties and task difficulty



1.8 Research methodology

1.8.1 Participants

A total of 156 accounting and finance students from a European university participated in the experiment in May 2016 and January 2017, but three were excluded after a manipulation check[9] and five more after trial probes. For 147 subjects we have complete data, their mean age in years is 23.24 (*SD* = 0.48), 34.7% of them are male (N = 51) and 65.3% female (N=96).

We used a within-subject experimental design on two subsamples, each of which we incentivised with either a reward or a penalty. Subjects were randomly assigned to one of the two treatments. They were compensated with an initial endowment, an incentive based

[9] Manipulation check questions: For incorrect answers I could not lose additional money. [True/False]; In the first round, probability of failure for the choice of the difficult task was higher than for the choice of the easy task. [True/False]; In the second round, probability of failure for the choice of the difficult task the was higher than for the choice of the difficult task in the first round. [True/False];

on their choice and performance in the experiment, and granted a course credit. All participants signed an informed consent form before starting the experiment. They were informed of the possibility of withdrawing at any point during the experiment. The design and procedure of the experiment conformed with the Helsinki declaration on ethical principles for medical research involving human subjects (revised 2013 version).

1.8.2 Experimental task

To analyse cognitive effort, we used a modified Sternberg task (Sternberg, 1966) that has been widely used in psychology (Jansma et al., 2007; Locke and Braver, 2008; Zakrzewska and Brzezicka, 2014). The cognitive load in a task can be manipulated by increasing the length of the memory set (Braver, Cohen, Nystrom, Jonides, Smith, and Noll, 1997). The task is as follows: a sequence of letters (a memory set) is presented on the screen, each letter appearing for one second until all letters of the memory set appear on the screen. Then the memory set disappears and a recognition number is presented in the middle of the screen. A participant has to press the letter on the keyboard that corresponds to the position of the memory set represented by the recognition number. For example, after a memory set B H V C a recognition number 2 appears on the screen. The participant has to press the letter H because H was in the second place in the memory set. If the letter the participant indicates is correct, the square with that letter turns green, otherwise red. The task is relatively easy to understand and pre-existing knowledge is not needed. The computerised task was programmed in JavaScript and designed with HTML and CSS. The data were stored in textual format.

Figure 6. Illustration of the experimental task

Note. Figure 6 represents the experimental task. First screen shows a three letter sequence, on a second screen a probe letter appears to which a participant has to provide an answer and on a third screen a visual feedback after an incorrect response is provided.

1.8.3 Manipulation and procedure

Probe and calibration phase

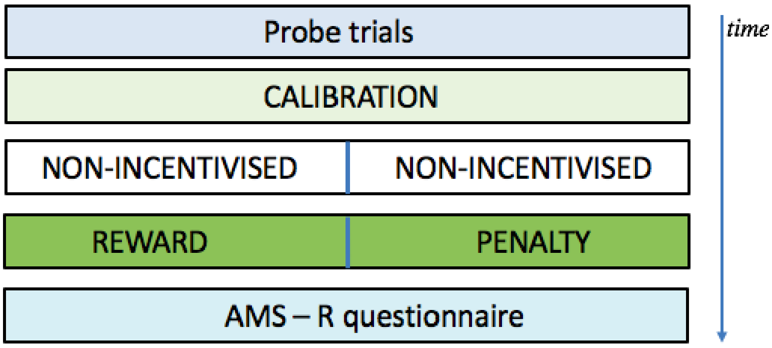
After the instructions on the procedure had appeared, the participants conducted three probe trials to ensure comprehension of the task. The probe trials were memory sets of three letters. Five subjects who did not manage to solve two of the three probe trials were excluded from the experimental results, so this also served as a manipulation check of the task comprehension.

Because individuals have different working memory capacities, we needed to ensure the results were comparable and exclude the influence of an individual's cognitive abilities (working memory) on decisions in the main experimental phase. Participants underwent a calibration phase where each individual's maximum working memory capacity was defined. Based on this result, the task difficulty was adjusted for each individual and the initial length of the memory set for each subject was established. In this phase, participants learnt about their ability to memorise the number of letters in a sequence (Bénabou and Tirole, 2003) and became acquainted with their subjective experience of the difficulty level (Westbrook et al., 2013). All participants started the calibration phase with a three letter sequence that appeared three times. In order to increase sequence length, two ouf of three presented sequences had to be solved correctly. The calibration phase ended at the sequence length where two trials out of three were not solved correctly. This length was coded as the maximum length of the memory set and was used to determine the appropriate difficulty level for each participant. At the end of the calibration phase, the subjects received their initial endowment expressed in experimental units (EU). The amount was either EU 36 or 60, depending on the incentive treatment group to which they were assigned in the following phase of the experiment. The final outcome was denoted in EU (average outcome of EU 65.6 ±7.8 in the reward treatment and EU 60.55±10.5 in the penalty treatment) and translated into euros (EU 1 = EUR 0.05). Average compensation for approximately 30 minutes of activity equalled EUR 3.28 in the reward treatment and EUR 3.03 in the penalty treatment. This amount is proportional to the average hourly rate for student work.

Experimental task

The experimental task followed the calibration phase. The participants had to decide between an easy or a challenging task over three rounds of the experiment in two different settings (i.e. two times three rounds) where the difficulty level of the challenging task increased in every round. One setting was non-incentivised and the other was incentivised. In the incentivised setting, the incentives were defined in terms of a reward or a penalty. Monetary equivalence of both conditions was ensured. Half the participants started with the incentivised setting and half with the non-incentivised setting to ensure that the order of the setting did not influence the results. In figure the experimental procedure is depicted and a decision tree showing the incentive scheme is depicted in Figure 8.

Figure 7. The experimental procedure (phases)

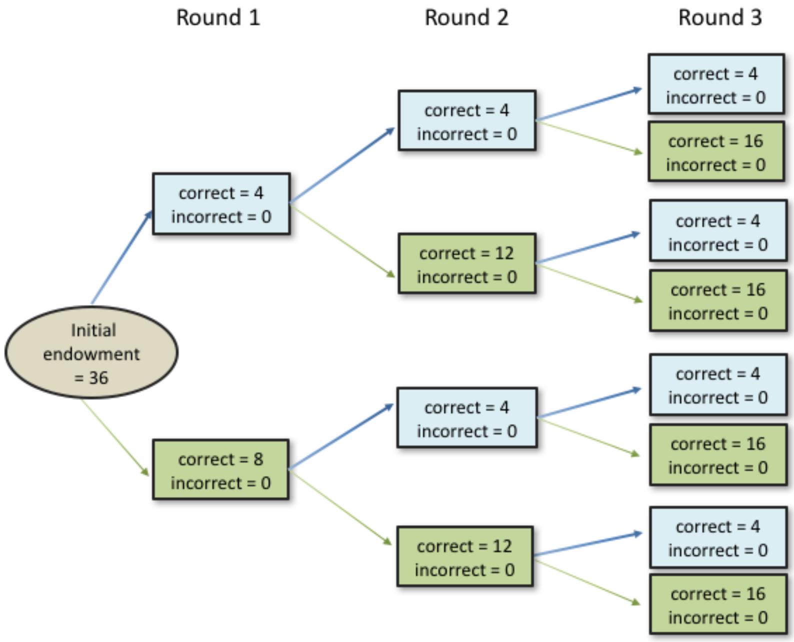


Note. The experiment started with three probe trails, followed by a calibration phase and a main experimental phase with three decisions in each treatment (incentivised and non – incentivised).

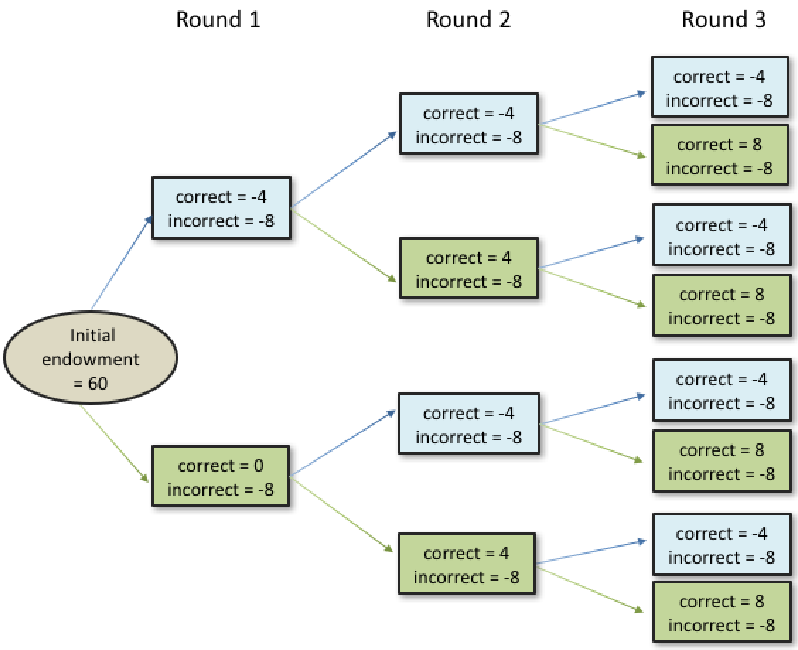
The difficulty of the task was defined as the length of the memory set below or above the length achieved in the calibration phase. In the easy task, the length of the memory set was held constant for all three periods in both settings. This length was two letters less than the length they had achieved in the calibration phase. In the first round, the length of the challenging task memory set was equal to the length achieved in the calibration phase. In the second round, the challenging task had one letter more than the length achieved in the calibration phase, while in the third round it was two letters longer than the length achieved in the calibration phase. The easy task was associated with a smaller reward in the reward treatment and a larger penalty if selected in the penalty treatment. The choice of the challenging task was associated with a higher reward in the reward treatment and a smaller penalty in average in the penalty treatment.

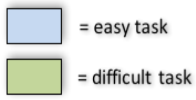
Figure 8. Decision tree in the reward and the penalty treatment

Decision tree: The ***REWARD*** *treatment*



Decision tree: The ***PENALTY*** *treatment*





Note: Figure 8 presents the decision options in the reward and the penalty treatment: the number in the circle is the initial endowment in each condition. All incentives for correct and incorrect responses are expressed in experimental units (EU). The subjects could choose between the easy task (a blue square), which in all rounds is represented by a sequence length that is two letters shorter than the sequence length achieved in the calibration phase, and the challenging task (a green square) which is as long as the sequence length achieved in the calibration phase in the first round, has one letter more in the second round and two letters more in the third round.

1.8.4 Measurement of the independent variables

Achievement motives were measured with the revised Achievement Motivation Scale (AMS-R) developed by Lang and Fries (2006). AMS-R is a shortened 10-item version of the original 30-item version of the AMS (Gjesme and Nygard, 1970) (their list is in the Appendix B.1). Participants rated the translated AMS-R items on a 4-point Likert scale (from 1 "strongly disagree" to 4 "strongly agree"). Five items measured motivation to achieve (MTA) and five items measured motivation to avoid failure (MTF). An example of an MTA item is: *"When I am confronted with a problem which I can possibly solve, I am excited to start working on it immediately"* and an example of an MTF item: *"Even if nobody notices my failure, I'm afraid of tasks which I'm not able to solve*". After completing the experiment and filling out the AMS-R scale questionnaire, the participants were paid the money they had earned in the experiment.

Control variables were the choice of a difficult task in a prior round (Diffcorrect), which is defined as 1 if in the prior round the difficult task was chosen and correctly solved, and 0 otherwise and the number of letters achieved in the calibration phase (Calibration). The concern is that the number of letters did not just reflect cognitive capacity but also achievement motivation in the calibration phase. High number of letters achieved in the calibration phase did namely not allow for the same improvement in difficulty level as low number of letters. On the other hand, some less interested participants might have not exerted as much effort in the calibration phase as was their cognitive capacity and they would have more scope for taking on a more difficult task. With Calibration we controlled for such effects.

1.8.5 Method of analysis

We estimated the effect of MTA and MTF on choices of difficult tasks in the non-incentivised setting with a logistic regression. We treated incentive and no incentive conditions as a repeated-measures design in which we analysed the effects of the incentive, time-invariant MTA, MTF and the interaction between the incentive and MTA, MTF. To estimate the logistic model, we used the Generalised Estimating Equations (GEE) method, which is suitable for correlated observations (in our case, subjects treated with or without an incentive). The model is a generalisation of GLM and based on a quasi-likelihood estimation. It considers the autoregressive correlation structure in the data, i.e. a correlation that depends on the interval of time between responses (Kleinbaum and Klein, 2010).

1.9 Results

1.9.1 Descriptive statistics

Descriptive statistics are presented in Table 5. In the calibration phase of the experiment, the average sequence length achieved was 5.67 letters (S.D. = 2.34, Min = 3, Max = 9). These results confirm a general memory span of an average 7 (+/-) 2 items (Miller, 1956). Gender distribution, age, years of work experience and payoff were not significantly different between the treatment groups (Table 5) or the choices made (Table 7). The average score of achievement motivation was: for MTA (M = 3.01, S.D. = 0.69) and for MTF (M = 2.85, S.D. = 0.74) on the scale from 1 to 5. Cronbach's alpha for MTA was 0.82 and for MTF 0.80, indicating the high level of our scale's internal consistency. Results of the independent sample t-test suggest that the average performance[10] between the reward and the penalty treatment does not differ significantly in any of the three rounds: Round 1 ($t(113) = -0.156$, $p = 0.877$), Round 2 ($t(99) = -0.685$, $p = 0.495$), Round 3 ($t(78) = 0.180$, $p = 0.857$). A paired samples t-test was conducted to compare performance in the non-incentivized and the incentivized treatment, separately for the reward and the penalty treatment in every round. We did not find any significant difference in any of the three rounds[11]. We also did not find any significant difference in the final compensation earned in the reward and the penalty treatment ($t(145) = 1.266$, $p = 0.207$).

[10] See Table 7 for the descriptives. The measure for performance is correctly solved difficult task.

[11] Reward treatment: Round 1 ($t(40) = 1.777$, $p = 0.083$), Round 2 ($t(40) = 0.274$, $p = 0.785$), Round 3 ($t(24) = -1.072$, $p = 0.294$). Penalty treatment: Round 1 ($t(42) = -0.628$, $p = 0.533$), Round 2 ($t(38) = -0.530$, $p = 0.599$), Round 3 ($t(29) = 0.254$, $p = 0.801$).

Table 5. Descriptive statistics by treatment group

	Reward treatment	**Penalty treatment**	***Total***
N	77	70	147
Gender			
Female (N)	54	42	96
%	70.1%	60.0%	65.3%
Age			
Mean	23.23	23.24	23.24
(S.D.)	(3.42)	(3.47)	(4.43)
Years of work experience			
Mean	2.37	2.30	2.33
(S.D.)	(2.97)	(4.76)	(3.90)
Sequence length achieved in calibration			
Mean	5.52	5.83	5.67
(S.D.)	(1.05)	(1.41)	(2.34)
Total payoff (EU)			
Mean	65.58	60.55	63.20
(S.D.)	(7.92)	(10.45)	(9.17)
MTA			
Mean	3.01	3.01	3.01
(S.D.)	(0.68)	(0.69)	(0.69)
MTF			
Mean	2.87	2.82	2.85
(S.D.)	(0.70)	(0.78)	(0.74)

Note: Pearson χ^2 tests indicate no significant difference in the frequencies for gender (χ^2 = - 0.120, p = 0.200). Independent *t*-tests indicate no significant differences in the group means for age $t(147) = 0.016$, $p = 0.987$, gender $t(147) = 1.283$, $p = 0.200$, work experience $t(144) = -0.112$, $p = 0.911$, number of letters achieved in the calibration phase t(147) = 1.494, $p = 0.138$, MTA $t(145) = 0.011$, $p = 0.991$ and MTF $t(145) = -0.448$, $p = 0.655$. The correlation coefficient between MTA and MTF is positive and moderate $r = 0.436$, $p \leq 0.001$.

Figure 9. Percentage of difficult task choices made in the reward vs. the no-incentive treatment

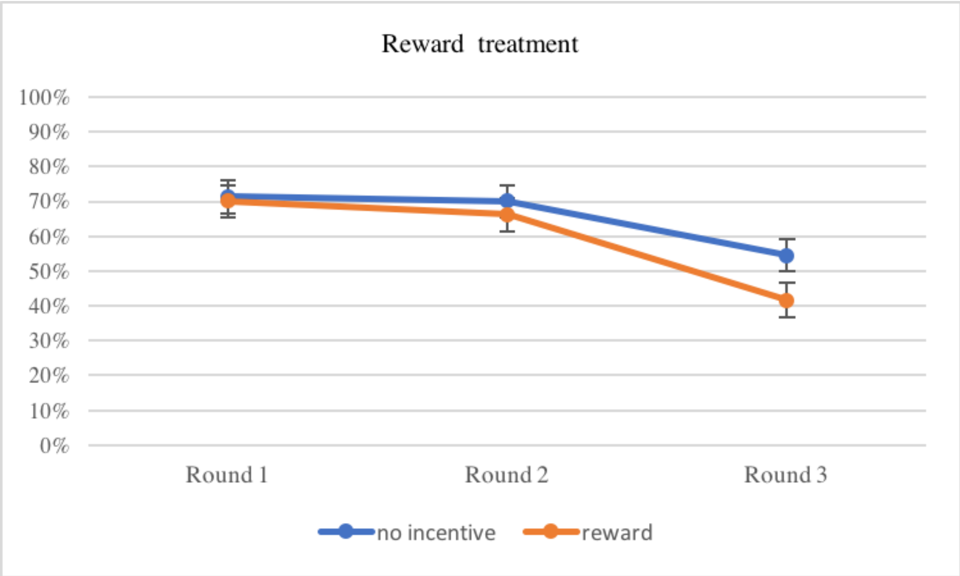
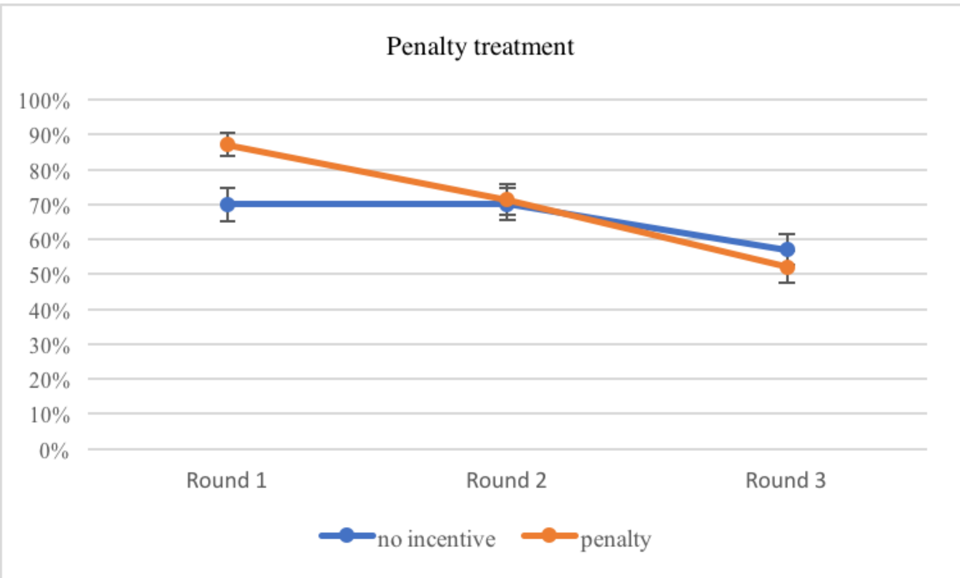


Figure 10. Percentage of difficult task choices made in the penalty vs. the no-incentive treatment



Note. Error bars represent standard errors. Y axis presents the percentage of challenging task choices each round for each treatment group.

To analyse our research questions, we first performed a logistic regression to ascertain the pattern of achievement motivation (MTF and MTA) on willingness to select a difficult task in each of the three rounds (0 indicated the choice of the easy task and 1 the choice of the difficult task) without an incentive, controlling for the sequence length achieved in the calibration phase. In the first round (see Round 1, Table 6) in which the difficult task was set equal to the subjects' performance in the calibration phase, we do not observe any significant effects of MTF (b = - 0.232, χ^2 (1) = 0.684, p = 0.408) or MTA (b = -0.441, χ^2(1) = 1.933, p = 0.164). In this round, in which the attainability of the outcome is relatively certain, MTA and MTF do not explain effort choices. In other words, an individual does not have to be a high achiever to choose a difficult task on this level, neither do those who have high MTF decide significantly differently from others for a difficult task. In the second round where the task difficulty is increased by one letter we observe a positive effect of MTA (b = 0.675, χ^2(1) = 4.218, p = 0.040). In the third round, in which the task difficulty is increased by two letters above the calibration and the outcome becomes less attainable, we find a significant negative effect of MTF (b = -0.721, χ^2(1) = 6.762, p = 0.009), suggesting that the difficult task is selected more often by those who are less afraid of failure. High MTA does not impact the choice at this level. Overall, when the difference between the easy and the difficult task is large enough, achievement motivation becomes a factor influencing the decisions on selecting a challenging task. The influence switches from MTA to MTF as the task difficulty increases.

Table 6. Logistic regression predicting the decision for the difficult task in the non-incentivised treatment

	Source	b	S.E. (b)	Wald χ^2	p value	OR	95% CI OR
Round 1	Constant	6.000	1.920	9.766	0.002	403.525	
	MTF	-0.232	0.281	0.684	0.408	0.793	[0.46, 1.38]
	MTA	-0.441	0.317	1.933	0.164	0.643	[0.35, 1.20]
	Calibration	-0.542	0.164	10.894	**0.001**	0.582	[0.41, 0.80]
Round 2	Constant	4.093	1.936	4.471	0.034	59.921	
	MTF	-0.127	0.296	0.182	0.670	0.881	[0.48, 1.58]
	MTA	0.675	0.329	4.218	**0.040**	1.964	[1.02, 3.73]
	Calibration	-0.830	0.191	18.843	**0.000**	0.436	[0.30, 0.62]
Round 3	Constant	4.287	1.751	5.993	0.014	72.715	
	MTF	-0.721	0.277	6.762	**0.009**	0.486	[0.27, 0.84]
	MTA	0.193	0.290	0.443	0.506	1.213	[0.69, 2.13]
	Calibration	-0.451	0.154	8.589	**0.003**	0.637	[0.46, 0.86]

Note. For Round 1: R^2 = 0.576 (Hosmer and Lemeshow), 0.086 (Cox and Snell), 0.122 (Nagelkerke). Model χ^2 (3) = 13.213 p=0.004; For Round 2: R^2 = 0.562 (Hosmer and Lemeshow), 0.196 (Cox and Snell), 0.277 (Nagelkerke). Model χ^2 (3) = 31.980 p<0.000; For Round 3: R^2 = 0.581 (Hosmer and Lemeshow), 0.126 (Cox and Snell), 0.169 (Nagelkerke). Model χ^2 (3) = 19.839 p<0.000.

1.9.2 Testing hypotheses

Test of H1, H2 and H3 – the effect of rewards

Next, we compared the choices in the incentivised and non-incentivised treatments for the reward and the penalty conditions. Descriptive statistics (see Table 7) reveal that the choices for the difficult task in the reward treatment are less frequent than in the non-incentivised treatment in all three rounds: the frequency amounts to 70.1% vs. 71.4% in the first round, 66.2% vs. 70.1% in the second round and 41.6% to 54.5% in the third round. Pearson χ^2 tests indicate no significant difference in the frequencies of chosen task difficulty for gender in any of the three rounds in the non-incentivised setting (χ^2 = 2.604, p = 0.107; χ^2 = 1.526, p = 0.217; χ^2 = 3.751, p = 0.053) and the incentivised setting (χ^2 = 0.214, p = 0.644; χ^2 = 0.000, p = 0.988; χ^2 = 3.330, p = 0.068).

In H1, we propose that at low difficulty level rewards positively affect the preparedness to select a challenging task for both low and high MTA individuals in comparison to a non-incentivised task. We do not anticipate any interaction effect. The results of the repeated measures (GEE) logistic regression analysis reveal that at low task difficulty rewards do not have any positive effect on the selection of a challenging task and we cannot confirm our hypothesis (see Table 4, Round 1). In H2 we predict that at medium task difficulty, rewards have a stronger positive effect on willingness to choose a challenging task for low MTA individuals than for high MTA individuals (Table 8, Round 2). The signs of coefficient for the reward and the interaction with MTA are in line with our prediction, but the effects are not significant. We find that for the medium difficulty level MTA is the only variable affecting task selection, just like in the no-incentive setting.

In H3 we propose that at high task difficulty, rewards have a negative effect on willingness to choose a challenging task, which is stronger for high MTA individuals than for low MTA individuals. We find a negative main effect of the reward (Table 8, Round 3, Model 1: b = -0.629, z = -2.070, p = 0.038). The interaction model shows that the overall negative effect of rewards can be explained by the negative interaction effect between rewards and MTA (Table 4, Round 3: Model 2: b = -0.979, z = -2.090, p = 0.037). The stronger the motivation to achieve, the stronger is the negative effect of rewards. In other words: at high task difficulty rewards undermine the motivation of high achievers more than that of low achievers. The results are in accordance with our prediction. Overall, our findings do not confirm that rewards would positively affect the willingness to choose a challenging task at easy or moderate task difficulty, but suggest that have a detrimental effect in choices where the task difficulty is high, which is particularly pronounced for high MTA individuals.

Table 7. Choice and outcome by treatment group

		No incentive	Reward	No incentive	Penalty
Round 1	**Easy**	22	23	21	9
	Difficult	55	54	49	61
	% Difficult task chosen	71.4%	70.1%	70.0%	87.1%
	(S.D.)	(0.45)	(0.42)	(0.46)	(0.34)
	% Successfully solved difficult task	72.7%	81.5%	85.7%	80.3%
	(S.D.)	(0.43)	(0.37)	(0.40)	(0.48)
Round 2	**Easy**	23	26	21	20
	Difficult	54	51	49	50
	% Difficult task chosen	70.1%	66.2%	70.0%	71.4%
	(S.D.)	(0.46)	(0.44)	(0.46)	(0.40)
	% Successfully solved difficult task	63.0%	58.8%	65.0%	52.0%
	(S.D.)	(0.45)	(0.43)	(0.50)	(0.47)
Round 3	**Easy**	35	32	30	22
	Difficult	42	45	40	48
	% Difficult task chosen	54.5%	41.6%	57.1%	52.1%
	(S.D.)	(0.50)	(0.49)	(0.50)	(0.46)
	% Successfully solved difficult task	33.3%	50.0%	42.5%	64.1%
	(S.D.)	(0.49)	(0.50)	(0.49)	(0.50)
Total		77	77	70	70

Table 8. GEE logistic regression – the reward treatment

	Difficult task choice	Model 1 b	Model 1 S. E.	Model 1 z	Model 1 p(z)	Model 2 b	Model 2 S. E.	Model 2 z	Model 2 p(z)
Round 1	Intercept	2.371	1.898	1.250	0.212	4.144	2.404	1.72	0.085
	Reward	-0.065	0.346	-0.190	0.851	-3.544	2.829	-1.25	0.210
	MTA	0.207	0.306	0.680	0.499	-0.238	0.424	-0.56	0.575
	MTF	-0.041	0.296	-0.140	0.889	-0.182	0.406	-0.45	0.653
	Reward*MTA					0.890	0.575	1.55	0.121
	Reward*MTF					0.287	0.548	0.52	0.601
	Calibration	-0.349	0.188	-1.860	0.063	-0.354	0.190	-1.86	0.063
Round 2	Intercept	2.007	1.967	1.020	0.308	0.891	2.518	0.350	0.723
	Reward	-0.248	0.404	-0.610	0.540	1.708	2.703	0.630	0.528
	MTA	0.438	0.339	1.290	0.196	0.939	0.463	2.030	**0.043**
	MTF	0.054	0.316	0.170	0.863	-0.079	0.432	-0.180	0.856
	Reward*MTA					-0.687	0.557	-1.230	0.217
	Reward*MTF					0.027	0.526	0.050	0.960
	Calibration	-0.629	0.208	-3.020	**0.003**	-0.546	0.216	-2.530	**0.011**
	Diffcorrect	2.017	0.424	4.750	**0.000**	1.154	0.431	2.670	**0.007**
Round 3	Intercept	3.137	2.101	1.490	0.135	2.292	2.374	0.970	0.334
	Reward	-0.629	0.303	-2.070	**0.038**	1.936	2.283	0.850	0.396
	MTA	0.054	0.345	0.160	0.876	0.720	0.431	1.670	0.095
	MTF	-0.207	0.326	-0.630	0.526	-0.282	0.390	-0.720	0.469
	Reward*MTA					-0.979	0.469	-2.090	**0.037**
	Reward*MTF					0.156	0.438	0.360	0.721
	Calibration	-0.549	0.211	-2.600	**0.009**	-0.646	0.219	-2.950	**0.003**
	Diffcorrect	1.247	0.412	3.030	**0.002**	0.172	0.455	0.380	0.705

Note. Diffcorrect is 1 if in the prior round the difficult task was chosen and correctly solved, otherwise 0. Model 1 and Model 2 for all rounds: Number of obs. = 154, Number of groups = 77, Periods = 2; Model 1 for round 1: Wald χ^2 (4) = 4.25, χ^2 = 0.373; For round 2: Wald χ^2 (5) = 32.06, χ^2 = 0.001; For round 3: Wald χ^2 (5) = 19.87, χ^2 = 0.001; Model 2 for round 1: Wald χ^2 (6) = 6.48, χ^2 = 0.372; For round 2: Wald χ^2 (7) = 19.33, χ^2 = 0.006; For round 3: Wald χ^2 (7) = 17.67, χ^2 = 0.014

Tests of H4 and H5 – the effect of penalties

Descriptive statistics for penalties are presented in Table 7. In the penalty scheme, the frequency of choices for the difficult task in the first round is 87.1% compared to 70.0% in the non-incentivised condition. In the second round, we observe almost no difference between the non-incentivised and incentivised settings: 70.0% vs. 71.4%, whereas in the third round the trend reverses and we observe a frequency of 57.1% for the difficult task choice in the non-incentivised setting vs. 52.1% in the incentivised setting.

In H4, we hypothesise that at low and moderate task difficulty penalties have a positive effect on willingness to choose a challenging task, which is stronger for low MTF individuals than for high MTF individuals. We hence predict a positive effect of penalties and a negative interaction effect with MTF. The GEE model shows that a penalty has a significant positive effect (b = 1.259, z = 2.710, p = 0.007) in Model 1 (Table 9), in which the interaction with MTF and MTA is not estimated. This effect corresponds to the large differences in frequencies in the first round (Descriptive statistics, Table 5). However, the positive effect of a penalty is almost completely cancelled out for high MTF individuals as shown by the significant and negative interaction term (Model 2: b = -1.479, z = -1.970, p = 0.049). In general, individuals react to a penalty as predicted, i.e. they engage more, which does not hold for high MTF individuals. The penalty does not stimulate these individuals to undertake a challenging task, presumably because they are more afraid of failure than of a penalty. The effect of a penalty in the second round is clearly distinct for high and low MTF individuals. On a stand-alone basis it is not significant as the effects for low MTF and high MTF individuals cancel out (Model 1: b = 0.087, z = 0.230, p = 0.819). In the interaction model the effect of the penalty is significant (Model 2: b = 7.490, z = 2.430, p = 0.015) but, like in the first round, high MTF considerably mitigates it (Model 2: interaction term b = -1.182, z = -2.030, p = 0.042), suggesting that high MTF individuals are less sensitive to it. Unexpectedly, we also find a significant interaction between MTA and penalty (Model 2: b = -1.336, z = -2.100, p = 0.036), implying that penalty has a smaller effect on high achievers, too, at moderate task difficulty.

In H5 we suggest that at high task difficulty, penalties have no effect on willingness to select a challenging task. We find that the effect of penalty in the third round is not significant (neither on a stand-alone basis nor in interaction), which is in line with our prediction. What remains significant is MTF (Model 2: b = -0.200, z = -2.690, p = 0.007), just like in the non-incentivised setting. The significance of MTF in both settings implies that personality traits endure in very difficult tasks regardless of incentives. Overall, our findings suggest there is no main effect of penalty because the interactive effects take place at easy and moderate task difficulty. Penalties are much more effective for low MTF individuals than for high MTF individuals, while at high task difficulty penalties no longer play a role.

Table 9. GEE logistic regression – the penalty treatment

		Model 1				Model 2			
	Difficult task choice	**b**	**S. E.**	**z**	**p(z)**	**b**	**S. E.**	**z**	**p(z)**
Round 1	Intercept	7.376	2.253	3.270	0.001	5.954	2.531	2.350	0.019
	Penalty	1.259	0.465	2.710	**0.007**	6.147	3.948	1.560	0.119
	MTA	-0.623	0.386	-1.620	0.106	-0.584	0.464	-1.260	0.208
	MTF	-0.679	0.326	-2.080	**0.037**	-0.244	0.382	-0.640	0.523
	Penalty*MTA					-0.095	0.757	-0.130	0.900
	Penalty*MTF					-1.497	0.761	-1.970	**0.049**
	Calibration	-0.465	0.164	-2.830	**0.005**	-0.453	0.168	-2.700	**0.007**
Round 2	Intercept	3.993	2.247	1.780	0.076	0.911	2.588	0.350	0.725
	Penalty	0.087	0.382	0.230	0.819	7.490	3.085	2.430	**0.015**
	MTA	-0.146	0.370	-0.390	0.694	0.420	0.460	0.910	0.361
	MTF	-0.410	0.339	-1.210	0.226	0.104	0.430	0.240	0.808
	Penalty*MTA					-1.336	0.636	-2.100	**0.036**
	Penalty*MTF					-1.182	0.582	-2.030	**0.042**
	Calibration	-0.377	0.174	-2.170	**0.030**	-0.394	0.184	-2.150	**0.032**
	Diffcorrect	1.479	0.474	3.120	**0.002**	1.601	0.505	3.170	**0.002**
Round 3	Intercept	5.246	1.956	2.680	0.007	1.500	0.456	3.290	0.001
	Penalty	0.591	0.379	1.560	0.119	-0.140	0.535	-0.260	0.794
	MTA	-0.589	0.333	-1.770	0.077	-0.025	0.084	-0.300	0.761
	MTF	-0.739	0.293	-2.520	**0.012**	-0.200	0.074	-2.690	**0.007**
	Penalty*MTA					-0.086	0.114	-0.760	0.450
	Penalty*MTF					0.181	0.101	1.790	0.073
	Calibration	-0.263	0.149	-1.760	0.078	-0.065	0.028	-2.350	**0.019**
	Diffcorrect	1.503	0.489	3.080	**0.002**	0.251	0.081	3.090	**0.002**

Note. Diffcorrect is 1 if in the prior round the difficult task was chosen and correctly solved, otherwise 0. Model 1 and Model 2 for all rounds: Number of obs. = 140, Number of groups = 70, Periods = 2; Model 1 for round 1: Wald χ^2 (4) = 16.07, χ^2 = 0.003; For round 2: Wald χ^2 (5) = 20.29, χ^2 = 0.001; For round 3: Wald χ^2 (5) = 19.82, χ^2 = 0.001; Model 2 for round 1: Wald χ^2 (6) = 16.46, χ^2 = 0.011; For round 2: Wald χ^2 (7) = 22.24, χ^2 = 0.001; For round 3: Wald χ^2 (7) = 33.34, χ^2 = 0.000

Effect of control variables in GEE models

In both models and in both treatments (Table 8 and Table 9) we consistently observe that the participants who had achieved a longer sequence length in the calibration significantly less frequently chose the difficult task in all three rounds. We suppose this occurs because it is more difficult to raise difficulty levels if one has stretched to the maximum degree already in the calibration phase. Due to cognitive limitations, such individuals have little scope to increase their performance as the task difficulty rises. We also find a consistently positive effect of the variable Diffcorrect with which we controlled for persistency of choices and success in difficult tasks throughout the three rounds. The estimation of the likelihood of success is based on past success in challenging tasks and we find that individuals' reliance on that is one of the main explanatory factors in all of the models.

1.10 Discussion and conclusion

1.10.1 The effect of achievement motivation under increasing task difficulty

Successful engagement in cognitively challenging tasks is a driving force of performance in contemporary organisations in which a large number of activities is increasingly cognitively demanding. Despite the notion that monetary incentives are an important determinant of effort choices, the results of previous studies on cognitive effort choices are extraordinarily inconsistent. To understand these inconsistencies, we analysed the interactive effects of achievement motivation and task difficulty on monetary incentives. Prior research acknowledges that achievement motivation causes different responsiveness to cognitively demanding tasks with rising task difficulty, but it does not take into account changing attractiveness of financial incentives with increasing task difficulty. Our findings of achievement motivation in the non-incentivised setting are not novel in theory or in empirical research and thus we did not explicitly hypothesise them, but used them by way of comparison to the incentivised setting.

1.10.2 The effect of incentives for varying achievement motivation and task difficulty

When comparing frequencies of challenging task choices in the incentivised setting, we find that contrary to our expectations, rewards did not have a positive impact at low task difficulty as the majority of people chose the challenging task anyway. At medium task difficulty our results suggest that motivation to achieve is more decisive for who will select a challenging task than the rewards. Rewards neither outweigh nor accentuate the effect of motivational disposition. At high task difficulty we find a negative effect of rewards on task selection which may be almost entirely attributed to the effect they had on high achievers. We predicted such an effect by assuming that when task difficulty increases, uncertain rewards are compared to the cost of cognitive effort in a gain domain in which people are risk averse (Kahneman and Tversky, 1979). High achievers are more sensitive to uncertainty of rewards as it hinders their innate interest in the task more than the one of low achievers. Our findings partly support cognitive evaluation theory (Deci et al., 1999) according to which external rewards decrease intrinsic motivation. They also concur with neuroscientific evidence that a preference for an easy cognitive task does not change even if a larger reward is offered for a more demanding task (Massar et al., 2015; Westbrook et al., 2013). We find this to be true for highly challenging tasks and for high MTA individuals, whereas for lower levels of difficulty it is not apparent.

In the penalty condition, which we assumed would induce loss aversion and greater tolerance for outcome risk, we observe a significantly higher engagement in challenging tasks only for the level at which success is reasonably attainable. The choice for the difficult task in the first round is prevalent, because it offered the option to avoid a certain penalty. While in the non-incentivised setting MTF shows no effect, in the incentivised setting the interaction between a penalty and MTF has a significant negative sign, indicating that high MTF individuals less frequently opted for the difficult task than low MTF ones. This implies that high MTF individuals are less sensitive to financial penalties than low MTF individuals, and do not undertake a more difficult task even at the expense of a small certain loss. The result also holds in the second round at moderate task difficulty. In this round, we observe also a significant interaction between MTA (not explicitly predicted) and a penalty. The negative interactive effect is so strong that penalties are actually of little effect for both high MTA and MTF individuals.

At the highest level of difficulty, at which most participants estimated that by choosing the difficult task, they cannot avoid a loss, they opted for the easier task. The penalty had no effect on a stand-alone basis nor in interaction with MTF. What is again significant, just like in the non-incentivised setting, is MTF. Overall, our results confirm the conjecture that the interaction between the type of incentive, task difficulty, and achievement motivation together affect the willingness to select a cognitively challenging task.

1.10.3 Suggested implications for practice

The practical implications of our research lie in a better understanding of the interplay between monetary incentives, an individual's estimation of the probability of success and achievement motivation on selection of challenging tasks. Our results contribute to that stream of literature which notices that promoting higher cognitive effort by using performance contingent rewards, announced prior to the task, may not be very effective. People engage in challenging tasks because of innate achievement motivation. As suggested by Kunz and Pfaff (2002), the undermining effect of rewards on intrinsic motivation takes place only in situations in which further performance improvements are not possible. Our findings support their conjecture: rewards have a negative effect on high MTA individuals at a very high difficulty level. Our findings, however, shed also a different light on the role of rewards: while a non-incentivized cognitive task triggers motivational dispositions which affect *engagement* into a challenging task, rewards bring into this decision also the estimation of successful *completion* of the task. Rewards discourage individuals to engage in tasks which are beyond their cognitive abilities and in which success is not reasonably attainable.

The finding that penalties promote higher effort is relatively robust in the literature, nevertheless, penalty contracts are much rarer in practice. We confirmed the main effect of the penalty only for low task difficulty and found that low MTF individuals are more responsive to them. Penalties may indeed stimulate higher effort in contractual tasks of low and moderate difficulty, but as reported by Christ et al. (2012) bonus contracts stimulate higher effort on all those tasks that are not governed by the contract and for which there is no performance pay at all. In highly challenging jobs this is precisely the issue: not all tasks can be pre-defined in the contract. The finding that MTF explained choices at the highest task difficulty implies that a more effective way to stimulate engagement in challenging tasks would be to alleviate agents' perception of a failure and to reduce their fear of its consequences.

How should organisations therefore stimulate engagement in cognitively challenging tasks? The first approach seems to entail a careful selection of talented candidates based on their personality traits. The importance of careful employee selection in organisations and awareness of their intrinsic motivation is also emphasised by Grabner and Speckbacher (2016). A contract that can attract highly motivated candidates should provide a competitive fixed pay and a stimulating environment. Second, agressive performance-contingent pay for highly challenging cognitive tasks increases the uncertainty of the outcome and is less likely to appeal to candidates who are intrinsically attracted by such tasks. This does not preclude *ex post* sharing of the outcome as fixed-pay contracts are not likely to keep people with different success track records motivated. Individuals' preference for a fixed-pay scheme over a performance-based pay scheme for difficult cognitive tasks is also found by Brown, Farrington and Sprinkle (2016), but the authors report that this relationship is mediated by workers' skill assessment. Workers with higher skills are less aversive to performance contingent pay. Last, the use of *ex ante* rewards will be effective in an incentive system which balances risk taking and effort exertion and promotes engagement in those challenging task in which individuals can actually attain the outcome. As rewards can not increase cognitive performance, there is little reason for them to be unlimited. In our experiment the reward for hard-to-attain level was a moderate multiplier of the reward for the easy task, not outrageously high and this is why it has prompted individuals to consider completion of the task, rather than just engagement. Unproportionately high rewards for unattainable goals might change a break-even point between a smaller certain reward and a higher uncertain one and lure individuals towards greater risk seeking as they normally accept. Engagement in unattainable goals may impose various risks upon an organisation: from the waste of resources, high variability of outcomes, to an outright misrepresentation of performance in order to get the reward (Cadsby, Song and Tapon, 2010). Moderate rewards prime individuals to think about the financial outcome of the decision and help them decide to focus on attainable tasks and outcomes. Our experiment also shows the importance of feedback on successful completion of the challenging task, which also works in the other direction: an unsuccessful completion of a challenging task has a significant negative effect in later periods.

1.10.4 Limitations of the study

As the findings were obtained in an experiment, they need to be considered in the light of the methodological limitations. Our monetary incentives were admittedly rather low, but they were within the range of our previous studies, in which we found them effective. The results of an experiment are never to be understood at face value as they depend on the experiment's specific design. Had we changed the relative value of our incentives or task difficulty, we might have observed different results. Despite low incentive amounts, we believe we carried out the experiment in line with the theory and the highest research standards and that the results are logically consistent. What had played a role was a relative difference in rewards or penalties between the easy and the difficult tasks: in the first round the reward for the difficult task was two times larger than the reward for the easy task and in the third round it was four times larger than the reward for the easy task. It needs to be acknowledged that effort is to a certain extent adapted to the magnitude of an expected reward (Pessiglione et al., 2007). If we had used more meaningful rewards, they could have increased the willingness to accept a challenging task at low and medium task difficulty. We believe that results for the effect of a penalty in the first round and a reward in the last round would remain substantively similar as subjective evaluation of probability to get those rewards would remain intact. The use of students in cognitive tasks is not seen as a limitation as the results are generalisable to the general population (Runeson, 2003; Exadaktylos, Espín and Brañas-Garza, 2013). In our experiment, the use of students made the incentives somewhat more meaningful as their total outcome in the experiment compares to the hourly rate for student work.

To increase the validity of our findings, we propose some ideas for future research. Future studies may complement ours by analysing cognitive tasks that cover other areas of cognition (i.e. judgment, reasoning, inhibitory control, planning etc.). Some sort of mapping of cognitive functions would be most warranted as it would provide answers to how people react to incentives in various cognitive tasks. Another interesting research topic would be to examine interactions between incentives and other personality traits, for example behavioural approach and inhibition systems (Corr, McNaughton, Wilson, Burch, and Poropat, 2017), risk aversion and skills, on cognitive effort. Overall, we believe that the question of incentives' sensitivity to personality traits remains an inspiring area for future research.

GENDER DIFFERENCES IN INCENTIVE SENSITIVITY

Abstract

The purpose of this chapter is to examine differences in the effect of monetary rewards on gender in a challenging cognitive task. We are interested in how rewards motivate men and women at various difficulty levels. To investigate this question, we conducted a three period between and within subject experiment using a modified Sternberg task (Sternberg, 1966) on 148 students in which we measured participants' choices, their cognitive performance and financial outcome. We find that women and men respond differently to monetary rewards at various levels of task difficulty. At low task difficulty we did not find a significant effect of rewards, either for women or for men, on engagement in a challenging task. At medium task difficulty, rewards engaged women more frequently than men. At high task difficulty, rewards had a negative effect on willingness to engage in a challenging task for women and non-negative effect for men. We found very high variation between the genders, which may be the result of behavioural approach, or avoidance behaviour (women are higher in avoidance behaviour). These personality traits did not, however significantly explain the choices in the model. The most important determinant of choices was the success or failure feedback of the prior period. The study advances the literature studying gender differences in various incentive contexts.

Keywords: gender, cognitive effort, risk, incentive

JEL M41, D91, J1

1.11 Introduction

Several choices and decisions in life, such as career challenges or educational choices, involve a selection among options that vary in required effort, success probability and financial outcome. Some people, more often than others, play it safe, choose easier tasks that require less effort and in which they are more likely to succeed. Others, on the other hand, enjoy facing challenges and choose more challenging tasks in which the reward is higher, but the probability of success lower. One of the puzzles is why women in the developed world do better at school, including a University degree, but take demanding jobs less frequently than men (Economist, 2017, October 7th, p. 13; De Pater, Annelies, Van Vianen and Ginkel, 2009). Sociological reasons (i.e. work-life balance, equal opportunities, institutional support for child care) may be main reasons, but the literature suggests that there are also psychological differences between men and women in their willingness to engage in demanding jobs (De Pater et al., 2009).

In this chapter, we analyse how men and women make decisions in undertaking challenging cognitive tasks at increasing task difficulty, with and without rewards. We compare their choices within and between subjects. It is well established in the literature that men and women make different *risk* choices. Women are more cautious and experience risk in a different manner (Croson and Gneezy, 2009), they generally enjoy performing a challenging task less, like competition less (Niederle and Vesterlund, 2007; Datta, Gupta, Poulsen and Villeval, 2013) and are more insecure (Carson and Gneezy, 2009). However, deciding between undertaking an easy or a challenging cognitive task is not just about risk, but also about cognitive effort and reliance on self-confidence of one`s own cognitive capacities. An important, although under-examined area of the decision-making process, is the role of personality profiles in responses to incentives under varying task difficulty. Behavioural Inhibition System (BIS) and Behavioural Approach System (BAS) are two psychological systems that guide individuals' behaviour (Gray, 1981). Individuals' BIS and BAS predispositions may act as moderators of the decision-making process, as they are sensitive to rewarding or punishing stimuli. When studying gender differences in incentive sensitivity, it is important to take into account the underlying gender differences in the motives to approach and avoid failure, of which there is only limited prior evidence (De Pater et al., 2009).

To analyse gender differences, BIS/BAS predispositions and the impact of incentives in cognitive effort choices, we designed a multi-period experiment, in which we observed cognitive task selection that reflected *ex ante* risk-effort preferences. Our participants were asked to perform a cognitive task at their *chosen* difficulty level. We presented them with a series of decisions between a low risk-effort (an easy task) and a high risk-effort (a difficult task). The difficult task offered a higher expected return than the easy task, but it required more effort. We operationalised the risk component by rising task difficulty (easy, moderate and high), and by increasing the reward in each period. In one setting we incentivised participants, and, in the other, we asked them to make choices without incentives. We used a modified version of the Sternberg task (Sternberg, 1966). This task measures cognitive effort and is used widely in psychology (Jansma, Ramsey, de Zwart, van Gelderen, and Duyn, 2007; Oblak, Ličen and Slapničar, 2017; Zakrzewska and Brzezicka, 2014). The experiment was tested on 148 participants. We measured BIS and BAS predispositions with The Behavioural Inhibition System/Behavioural Activation System scale (Carver and White, 1994). The scale consists of 20 items using a 4 point Likert scale.

Our descriptive results indicate that in the non-incentivised setting, men consistently chose more difficult tasks then women. In the incentivised setting, men chose more difficult tasks only at easy and high task difficulty, but not at moderate task difficulty. Responsiveness to incentives of men and women at low task difficulty does *not* differ. Somewhat surprisingly, at moderate task difficulty, rewards had a significant positive effect on women and a negative effect on men. At high task difficulty, rewards were stimulating neither for men nor for women and the effect was more pronounced for women than for men. What mattered the most and consistently across periods and gender was the effect of a prior outcome in the difficult task. This suggests that the evaluation of the probability that one can complete a challenging task successfully depends to a great extent on past track record. The finding is consistent throughout all three studies of the doctoral thesis, despite the fact that they have been made on different samples of students. Although women have significantly higher BIS, BIS and BAS were not found to explain their choices significantly, neither as main effects, nor in interaction with incentives or gender[12]. Nevertheless, they may be responsible for high standard deviations between the genders.

The chapter contributes to the literature on gender differences in incentive sensitivity from several perspectives. Firstly, while gender differences in risk taking have been explored thoroughly in the literature, we are not aware of prior studies that would investigate gender differences in combined *risk-effort* choices in a multi-period setting. Prior gender studies explored risk (Barber and Odean, 2001; Eckel and Grossman, 2008; Croson and Gneezy, 2009) or effort choices separately (Gneezy, Niederle and Rustichini, 2003; Gill and Prowse, 2010) or analysed task choices at only one difficulty level (Federer, Nehm and Pearl, 2016), not considering that incentives may have different effects when they are expected with different probabilities. Our results indicate that the effect of incentives on men and women varies at every level of task difficulty. We also show in this and both previous chapters that studies that measure the effect of incentives only at one task difficulty level without individually calibrating it, cannot explain a full incentive–effort relation. Secondly, in real life decisions, exact outcome probabilities are not known and, importantly, depend on one's estimation of one`s own abilities and on past performance (Thaler and Johnson, 1990). Prior outcomes can be taken into account only in a multi-

[12]In the interaction models we obtain robust results, but for brevity do not report them in the paper.

period setting. Thirdly, our study provides evidence that men who, more often than women, chose a task at higher difficulty level and accepted higher risk, did not achieve an overall better performance and higher financial outcome. We discuss these findings and their implication in the last section.

The rest of the chapter is organised as follows. Section 2 presents the theoretical background and hypotheses. Section 3 describes the empirical design. Section 4 presents descriptive results and the model estimation, and in section 5 we discuss our findings, conclude the chapter and describe its limitations.

1.12 Theoretical background and hypotheses

Generally, in deciding between a task with an easily attainable, smaller outcome, and a task with a less attainable, higher outcome, individuals take into consideration the value of the reward, its probability, and the cost of cognitive effort needed to accomplish a task (Kool, McGuire, Rosen and Botvinick, 2010; Kruglanski et al., 2012; Westbrook, Kester and Braver, 2013). The concrete combination of the reward magnitude, its probability and the cost of effort, will decide whether the reward will affect risk and effort choices positively, negatively or neutrally. Several studies suggest that, generally, men choose more challenging tasks than women. One of the main reasons is in women's greater risk aversion (Byrnes, Miller and Shafer, 1999), that can be attributed to differences in emotional reactions to risky situations (Loewenstein, Weber, Hsee and Welch, 2001), different valuations of outcomes and processing of information on probabilities (Fehr-Duda, De Gennaro and Schubert, 2006). Women experience negative emotions more strongly than men which, in turn, influences their utility of risky prospects (Harshman and Paivio, 1987). Lerner, Gonzalez, Small and Fischhoff (2003) suggest that women overreact to the probabilty of a failure. Because women fear a negative outcome more than men, they are also more conservative in taking risky decisions. Dreber and Hoffman (2010), suggest that risk behaviour is associated with hormone levels. High levels of testosterone and cortisol are associated with more risky behaviour (Apicella et al., 2008; Cueva et al., 2015; Garbarino, Slonim and Sydnor, 2011). The level of testosterone increases in a response to initial success, and so does the preparedness to take risks again (Bose, Ladley and Xin, 2016).

The choice between an easy and a difficult cognitive task is not purely a gamble, as it entails positive cost of effort. Therefore, it cannot simply be assumed that women being more risk averse will always prefer easier, less risky and less effortful tasks. The majority of studies examine gender differences in tournament vs. piece rate incentive schemes. Gender differences in task selection and effort exertion arise, especially in competitive environments (tournaments), which affect women's willingness to engage in challenging tasks negatively, while, in a non-competitive environment (piece rate compensation and pay-for-performance schemes), gender differences in effort exertion disappear (Gneezy et al., 2003), or women exert higher effort than men (Bracha and Fershtman, 2013; Masclet, Peterle and Larribeau, 2015). Masclet et al., (2015), examined gender differences in the context of flat wage incentives. They compared effort under the flat wage scheme and under the tournament scheme. Under the flat wage scheme women exerted greater effort than men. Under the tournament scheme, the authors do not report significant differences between men and women. The findings suggest that women feel more comfortable under non-competitive incentive schemes than men and that, under such schemes, women are motivated equally or more than men.

One explanation lies in the different effect of stress in such evaluation contexts on women and men. Stress affects decision strategies of men and women differently (Mather and Lighthall, 2012). It increases risk taking under uncertainty by men, and decreases it for women (Preston, Buchanan, Stansfield and Bechara, 2007; Van den Bos, Harteveld and Stoop, 2009). Moreover, artificially induced stress in the experiment decreased women's reward responsiveness and reward-related activation in a monetary incentive task (Bogdan and Pizzagalli, 2006; Ossewaarde et al. 2011). Decreased reward responsiveness diminishes the drive for high rewards and, thus, decreases women's risk taking.

Other compelling factors that explain choices of men and women are perception of ability and the willingness to challenge abilities (Beyer 1990, Harackiewicz and Elliot, 1993; Dweck, 2000). Women and men differ in their beliefs about their ability to perform well in harder tasks. Men, compared to women, are more certain about positive outcomes, and attribute less merit to luck in comparison to women (Datta Gupta, Poulsen and Villeval, 2013; Dweck, 2000; Eckel and Grossman, 2008; Kamas and Preston, 2012).

An overall frame that covers all these differences in personality traits is an individual's approach or avoidance behaviour (Gray, 1982). Whereas classical motivation theories assume that individuals evaluate incentives similarly, Gray and McNaughton (2000) and Corr, McNaughton, Wilson, Burch and Poropat (2017) propose that individuals' innate motivational dispositions influence their perception of situations as being rewarding or punishing. The Behavioural Approach System (BAS) represents the appetitive motivational system that responds to signals of reward or relief from punishment, and represents the driving force of behaviour. The Behavioural Inhibition System (BIS) represents the aversive motivational system that is sensitive to penalties, reward omission and novelty, and inhibits behaviour leading to negative outcomes associated with negative feelings (Corr, 2013).

The evaluation of task outcome, hence, depends largely on the motivational tendencies of an individual. In task selection, BIS/BAS affect the weighing of costs and benefits. The possibility to obtain a reward can be evaluated as an approach goal, but also as an avoidance goal, because of the possibility of failure and missing out on the reward (Spielberg, Heller and Miller, 2013). BIS activates when potential threats are detected, learned from past experience of punishment or omission of reward in case of failure at the task. High activity in BIS indicates a sharper responsiveness to punishment, or the omission of a reward, and leads to the avoidance of such situations (Corr, 2013). BIS thus plays an important role in learning to inhibit punished responses (Baskin - Sommers, Wallace, MacCoon, Curtin and Newman, 2010). Ryan and Di Domenico (2016), suggest that differences in individual motivational predisposition can also be related to intrinsic motivation. They propose that negative factors, (such as threats and pressured evaluations), have adverse effects on autonomous motivation, and that this effect will be more pronounced among people with high BIS. Tangible rewards that were found to mitigate intrinsic motivation will have a stronger effect on people who are higher in BAS.

The interest for gender differences in BIS and BAS is relatively recent and relatively scarce. Not surprisingly, the extant findings reveal systematic differences (Ding et al., 2017; Kivikangas, Kätsyri, Järvelä and Ravaja 2014; Wang et al., 2017). Men score lower on BIS (Cross, Copping and Campbell, 2011), but not higher on BAS, than women (Li et al., 2014). This reflects their lower sensitivity to punishment and reward omission and higher cognitive control over emotional states caused by punishment (Van den Bos et al., 2009). As reported in Van den Bos et al. (2009), women and men respond differently to social stressors. The stress level was measured by cortisol levels. Men were more sensitive to achievement challenges, and women were more sensitive to social rejection challenges.

Li et al. (2014), studied gender-specific the neuroanatomical basis of the BIS/BAS system. The Voxel-Based Morphometry (VBM) was used to evaluate gender differences in the correlations between regional grey matter in brain volume (rGMV) and scores on the BIS/BAS scale. Brain regions related to processing rewards and negative emotions have different patterns of activation in women and men. The authors also find that higher BIS scores in women were associated with increased sensitivity to negative information. BAS scores did not differ significantly between women and men. Ilies, De Pater and Judge (2007), find gender differences in processing punishment and rewards. They studied how performance feedback affects positive and negative affect in an 8-trial performance brainstorming task where, after each trial performance, feedback was provided, and subjects were asked to report their affective state. The authors claim that feedback information activates behavioural motivation systems. Positive feedback signals reward and activate the BAS system. Negative feedback, that is associated with punishment and negative emotions, activates the BIS system.

Since women have a stronger motive to avoid failure and a weaker motive to approach success than men, we suggest that gender differences in BIS and BAS contribute to explaining the differences in incentive sensitivity of men and women, and in their willingness to engage in challenging tasks. In deciding between a challenging task or an easy task, the more difficult the task, the more pronounced will be the gender differences in incentive sensitivity. With *low* task difficulty, women feel that even a challenging task is still a secure option. We do not predict differences in the effect of incentives on choices of a challenging task at low task difficulty. Both men and women can be stimulated by an incentive to undertake a challenging task. Low difficulty level also does not make the different effects of BIS/BAS apparent. The more difficult the task, the smaller will be the willingness of women to accept a challenging task because of higher, but more risky, reward. At *moderate* task difficulty, we expect that men react more positively to a reward than women, and are more likely to accept a challenging task. The choice between an easier and a more difficult task at *high* task difficulty level creates an articulated gain domain to which more risk averse individuals respond with increased risk aversion (Kahneman and Tversky, 1979). We expect negative effects of rewards on both men and women, but more pronounced for women. We propose the following hypotheses:

H1: At moderate task difficulty, rewards will affect men more positively than women in choices of challenging tasks.

H2: At high task difficulty, rewards will affect women more negatively than men in choices of challenging tasks.

H3: BAS will affect choices of challenging tasks positively.

H4: BIS will affect choices of challenging tasks negatively.

As first suggested by Thaler and Johnson (1990), and also found in our previous two studies, task selection in a multi-period setting is affected strongly by prior outcome. The outcome (positive or negative) changes the status quo of the decision problem and creates a new cognitive frame (Benartzi and Thaler, 1995). Thaler and Johnson (1990), propose that a negative outcome in the first period causes increased risk aversion in a multi-period gamble if the second period does not provide the opportunity to break even, whereas a negative outcome after an initial positive outcome is integrated with the reward and, thus, promotes risk-seeking. A prior positive or negative feedback also affects effort. A feedback about failure has a stronger effect on future effort exertion than a feedback about success (Anand, Oehlberg, Treadway and Nusslock, 2016). A feedback about failure lowers the effort in subsequent same, similar and unrelated tasks (Brunstein and Gollwitzer, 1996; Hutchinson, Sherman, Martinovic and Tenenbaum, 2008). A positive outcome feedback can affect motivation positively, because it acts as an attributional feedback that shapes the decision frame. There is no conclusive evidence that men would react to feedback on prior performance differently to women. Möbius, Niederle, Niehaus and Rosenblat (2013), who gave feedback to subjects in an IQ test and measured their reaction in the subsequent task, do not report any gender differences in the effect of feedback. Gill and Prowse (2014), on the other hand, report reduced effort for women after failure feedback, independently of the reward value, and reduced effort for men after failing to receive a large reward. We propose the following hypothesis:

H5: Prior outcome will affect choices of challenging tasks positively.

1.13 Research methodology

1.13.1 Participants and design

One hundred and forty-eight (148) subjects, 64.9% women and 35.1% men from 3rd year udergraduate and 1st year graduate Accounting Programme at the University of Ljubljana, Faculty of Economics, participated in the study. Despite the focus of this study on gender, we could not ensure a balanced gender structure of the sample due to the gender structure of the student population. The experiment was conducted in May, 2017 and September, 2017. It had a multi-period within-subject design, with the requirement to perform a cognitive task with and without an incentive. It consisted of a test phase, a calibration phase and the main experimental phase. Before the start, participants were informed about the confidentiality, duration, payment, and that participation is completely voluntary. Participants were informed about the possibility to withdraw at any time during the experiment. All participants signed an informed consent of participation that described briefly what the experiment is about, what is the wider scope of the research, and about the payment for partcipation.

1.13.2 Experimental task

Choices for challenging cognitive tasks were analysed with the modified Sternberg task (Sternberg, 1966), used widely in psychology (Jansma, Ramsey, de Zwart and van Gelderen, 2007; Locke and Braver, 2008; Zakrzewska and Brzezicka, 2014). Because cognitive load in the task is easy to manipulate, we were able to create different difficulty levels, which increased gradually (Braver, Cohen, Nystrom, Jonides, Smith and Noll, 1997). The modified Sternberg task is a working memory task. It requires participants to memorise a set of letters appearing in one second intervals (e.g. D H N P), and then to maintain them in working memory while responding to the probe number (e.g. 3) questioning which letter appeared on the indicated place. The participant had to press the letter on the keyboard (in this case N) that corresponds to the position of the letter in the memory set. The participants were given a visual feedback on their performance (a green colour square appeard for a correct answer, and a red colour square for an incorrect answer). Because the task is easy to understand, it does not require prior knowledge, and all participants have skills to undertake it at a level individually calibrated for them. Moreover, there is little possibility of learning effect (Sternberg, 2008; Shiran and Breznitz, 2011). The computerised task was programmed in JavaScript and designed with HTML and CSS. The data was stored in a textual format.

1.13.3 Experimental procedure

All participants initially performed a short practice session (three probe trails) that served as a test of the task comprehension[13]. Following this, they performed the calibration phase. This phase allowed us to establish what was the right difficulty level for each participant (Westbrook et al., 2013). By using this type of procedure, we ensured the comparability of results, because the difficulty level for each individual was determined on their individual cognitive ability. The calibration phase started with a three letter memory set and was repeated three times. The next memory set was one letter longer and also repeated three times. This pattern was repeated until two of three trials were solved incorrectly. The achieved length of the memory set was coded as the participant's maximum memory set. During the calibration phase, the participants learned how much effort was required to solve a determined length of the memory set correctly and estimated their working memory ability (Bénabou and Tirole, 2003).

The participants received initial financial endowment after the calibration phase. The endowment was expressed in Experimental Units (EU) and was set to EU 36 (EU1 = EUR 0.05; i. e. The initial endowment amounted to EUR 1.8). The endowment after the calibration phase reduced the house money effect (Thaler and Johnson, 1990) as we induced the participants to believe that they had earned this money and were not merely endowed with it.

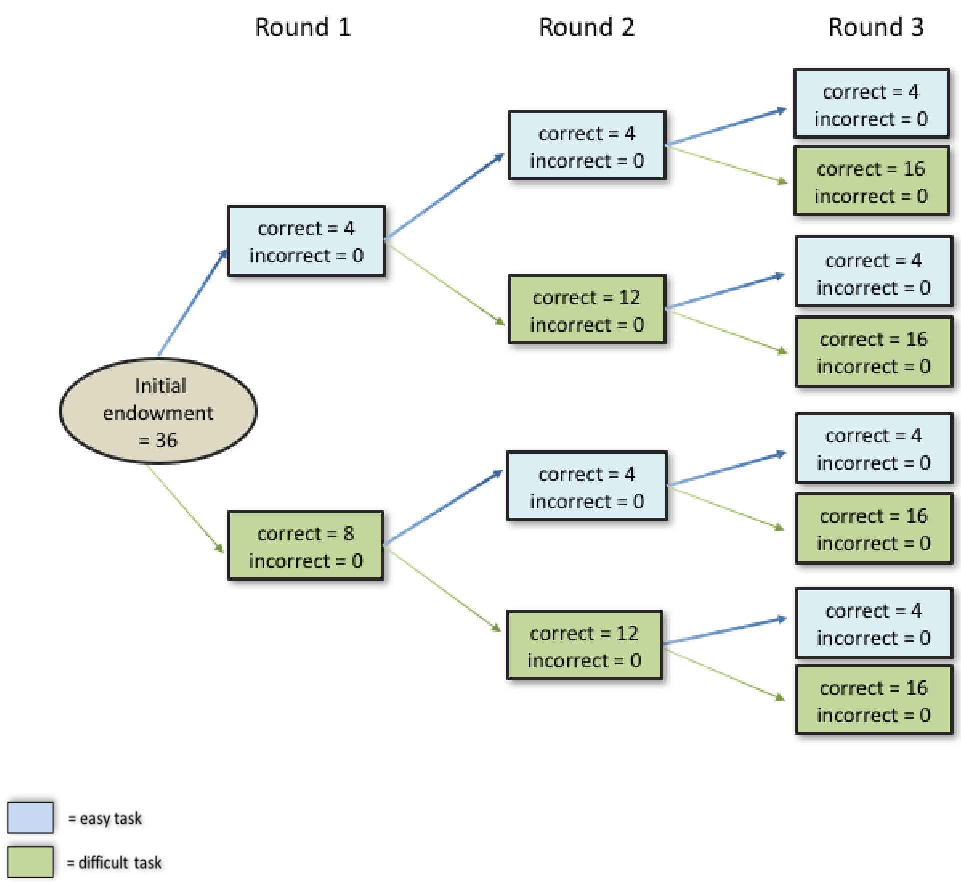
The main experimental phase that followed the calibration was divided into incentivised and non-incentivised treatments, both with three rounds. In the non-incentivised treatement, the participants made three choices between the easy or the difficult task. The difficulty was associated with the length of the memory set. The easy task in both treatments was two letters shorter than the maximum achieved length of the memory set in the calibration phase. The difficult task was as long as the maximum achieved length in the first round, one letter longer in the second round, and two letters longer in the third round. The same difficulty levels were used in the incentivised treatment. Half of the participants started with the non-incentivised treatment and half with the non-incentivised treatment, so we excluded the order effect.

In the incentivised treatment, the participants were asked to decide between the easy or the difficult task given the offered reward. The participants made choices between a lower effort - lower reward option (they could choose the option to respond for a set of letters that was two letters below the maximum memory set achieved in the calibration phase) and a higher effort - higher reward option. This option was made more difficult in each round by prolonging the number of letters they had to memorise for an increased reward. The reward associated with the difficult task choice was contingent on performance (correctly

[13] If participants managed to solve two of the three probe trials correctly, we considered that they comprehended the task.

or incorrectly solved task). In Round 1, the reward for a correctly solved task was EU 8, in Round 2 it was EU 12, and in Round 3 it was EU 16. If the task was solved incorrectly, the participants didn't get any reward. After completing the experiment the participants received their earned financial outcome. For approximately 30 minutes of activity, the average total compensation for women equalled EUR 3.63 and for men EUR 3.65 (the difference is non-significant). Despite being low, such an amount on an hourly basis would represent about the average hourly payment for student work.

Figure 11. Decison tree incentivised treatment



Note. Figure 11 presents the decision options in the incentivised treatment. The number in the circle is the initial endowment. All incentives for correct and incorrect responses are expressed in Experimental Units (EU). The subjects could choose between the easy task (a blue square), which, in all rounds, is represented by a sequence length that is two letters shorter than the sequence length achieved in the calibration phase, and the difficult task (a green square), which is as long as the sequence length achieved in the calibration phase in the first round, has one letter more in the second round and two letters more in the third round.

1.13.4 Measurement of BIS and BAS

Behavioural Inhibition System (BIS)/Behavioural Approach System(BAS) Scales were measured with the scale developed by Carver and White (1994)[14]. The questionnaire consists of 20 items; 7 items measure BIS and 13 items the BAS system. BAS can be divided into 3 subscales, although many studies apply a BAS sum score (e.g. because the subscales are not confirmed). In our analysis, we used the total BAS score combining all three BAS scales. Participants rated the Slovenian version of the questionnaire on a 4-point Likert scale (from 1 "strongly disagree" to 4 "strongly agree"). An example of a BAS item is: *"When I'm doing well at something, I love to keep at it."* and an example of a BIS item: *"If I think something unpleasant is going to happen, I usually get pretty 'worked up'."* Cronbach alpha for BAS is 0.742 and for BIS 0.732, indicating an adequate level of our scale's internal consistency.

1.13.5 Method of analysis

We analysed pairwise choices made in the incentivised and the non-incentivised treatment for each round with a repeated measure logistic regression (Generalised Estimating Equations or GEE model). GEE is suitable for auto-correlated observations (participants being treated with or without reward). The model is an extension of the GLM method and provides a semi-parametric approach to longitudinal analysis of categorical responses (Diggle, Liang and Zeger, 1994). We estimated the models in Stata. The dependent variable Choice (of the difficult task), Gender and Reward (with or without) are all dummy variables.

Next to estimation of the main effects of the reward (time-variant), gender (time-invariant) and their interaction, we controlled the length of the letter set achieved in the calibration phase. The reason why we controlled it is that those participants that had already put maximum effort in the calibration phase had less scope to accept an increased number of letters, as they were already close to their maximum working memory capacity. The outcome of the previous round if the difficult task was chosen (Diffcorrect) was measured with the value of 1 if the difficult task was solved successfully and the value of 0 otherwise. The outcome of the previous round helped the participants estimate the chances of success in the next round. In other words, if a participant had responded incorrectly in the difficult task in the previous round, it would have been less likely that he or she would decide for the difficult task in the next round, in which the difficulty was higher. The parameter estimates are provided as beta coefficients and as Odds Ratios (OR) that are suitable for binary outcome models (Homish, Edwards, Eiden and Leonard, 2010). The OR are the odds of choosing the difficult vs. the easy task in each round.

[14] The questionnaire is in Appendix D.

1.14 Results

1.14.1 Descriptive statistics

Table 10 presents descriptive statistics, and shows that age and working memory capacity are very similar for women and men. The average age of women is 22.02 (S.D. = 1.654) and of men 22.88 (S.D. = 1.023). The mean scores in the calibration phase for women (M = 5.58, S.D. = 0.966) and men (M = 5.58, S.D. = 1.334) are exactly the same, t (146) = 0.033, p = 0.976, suggesting that ability in memorising does not differ. Differences in choices are, hence, not driven by differences in cognitive capacities. The average length achieved in the calibration is in line with the general memory span of 7 (+/-) 2 items (Miller, 1956).

The average BIS score for women is 3.058 (S.D. = 0.460) and for men 2.728 (S.D. = 0.412). The scores are significantly different ($t(146) = 4.315$, $p = 0.000$). The difference in average BAS score for women 3.080 (S.D. = 0.341) and for men 3.111 (S.D. = 0.374) is non-significant, which is similar to the findings of previous studies. BIS and BAS are not correlated significantly to the results of the calibration phase.

Table 10. Descriptive statistics

	Women	**Men**
N	96	52
%	64.9	35.1
Age (mean)	22.02	22.88
(S.D.)	(1.654)	(1.023)
Sequence length achieved in calibration	5.58	5.58
(S.D.)	(0.966)	(1.334)
Total payoff (EU) (mean)	3.63	3.65
(S.D.)	(1.040)	(1.032)
BIS score	3.058	2.728
(S.D.)	(0.460)	(0.412)
BAS score	3.080	3.111
(S.D.)	(0.341)	(0.374)

From the descriptives in Table 11, in the non-incentivised treatment we can observe a robust preference of men compared to women for the difficult task in all three rounds: The relation is 69.0 % vs. 66.0% in Round 1, 69.0 % vs. 57.0% in Round 2 and 62.0% vs. 50.0% in Round 3 for men and women, respectively. In the incentivised treatment (Table 12, Figure 9) we observe a greater frequency of the difficult task choices of men in Round 1. The frequency is 77.0% vs. 61.0% in favour of men. In Round 3, where the difficulty of the task is the highest, men outperfom women strongly in difficult task selection. The frequency is 62.0% vs 36.0% in favour of men. The frequency of difficult task selection for men did not fall when incentivised, but it stayed on the same level. Contrary to our expectations, in Round 2, where the difficulty of the task is moderate, we observe a higher percentage of the difficult task choices by women, 69.0% vs. 57.0% of men. Despite quite large differences in frequencies in task selection between women and men in the non-incentivised treatment, we did not find them significant. In the incentivised treatment we find significant differences in Round 1 (χ^2= 0.198, *df* = 1, *p* = 0.056) and Round 3 (χ^2 = 8.563, *df* = 1, *p* =0.003). This suggests that standard deviations are large within both genders, and that other factors rather than gender will explain the choices.

Table 11. Percentage of difficult task choices in non-incentivized treatment

		Women	**Men**	χ^2	***p***
Round 1	**% Difficult task chosen (S.D.)**	**66.0 (0.477)**	**69.0 (0.466)**	0.198	0.656
	% Successfully solved task (S.D.)	76.0 (0.429)	85.0 (0.364)	1.497	0.221
	% Successfully solved difficult task (S.D.)	47.0 (0.502)	56.0 (0.501)	1.067	0.302
Round 2	**% Difficult task chosen (S.D.)**	**57.0 (0.497)**	**69.0 (0.466)**	2.030	0.154
	% Successfully solved task (S.D.)	76.0 (0.429)	67.0 (0.474)	1.305	0.253
	% Successfully solved difficult task (S.D.)	49.0 (0.503)	62.0 (0.490)	2.145	0.143
Round 3	**% Difficult task chosen (S.D.)**	**50.0 (0.503)**	**62.0 (0.491)**	1.808	0.179
	% Successfully solved task (S.D.)	74.0 (0.441)	69.0 (0.466)	0.376	0.540
	% Successfully solved difficult task (S.D.)	34.0 (0.481)	44.0 (0.501)	1.107	0.293
N		96	52		

Table 12. Percentage of difficult task choices in the incentivized treatment

		Women	Men	χ^2	*p*
Round 1	**% Difficult task chosen (S.D.)**	**61.0 (0.489)**	**77.0 (0.425)**	3.642	0.056
	% Successfully solved task (S.D.)	84.0 (0.365)	83.0 (0.382)	0.070	0.791
	% Successfully solved difficult task (*S.D.*)	50.0 (0.492)	61.0 (0.501)	1.808	0.179
Round 2	**% Difficult task chosen (S.D.)**	**73.0 (0.447)**	**63.0 (0.486)**	1.425	0.233
	% Successfully solved task (S.D.)	77.0 (0.423)	88.0 (0.323)	2.874	0.092
	% Successfully solved difficult task (S.D.)	52.0 (0.505)	50.0 (0.490)	0.59	0.809
Round 3	**% Difficult task chosen (S.D.)**	**36.0 (0.484)**	**62.0 (0.491)**	8.563	**0.003**
	% Successfully solved task (S.D.)	78.0 (0.416)	65.0 (0.480)	2.821	0.093
	% Successfully solved difficult task (S.D.)	18.0 (0.466)	31.0 (0.501)	3.321	0.068
N		96	52		

Figure 12. Percentage of difficult task choices made in the non-incentivised treatment

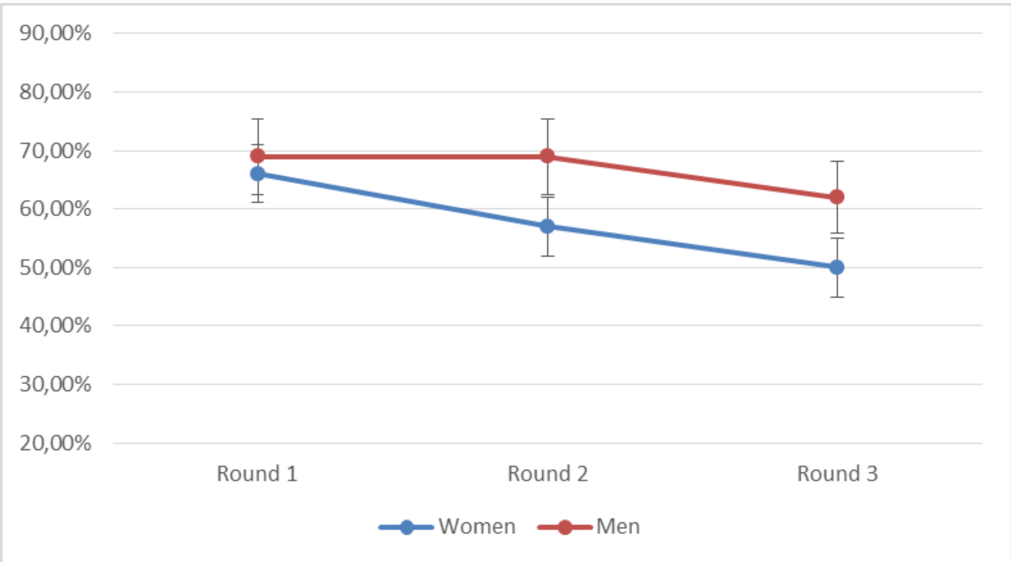
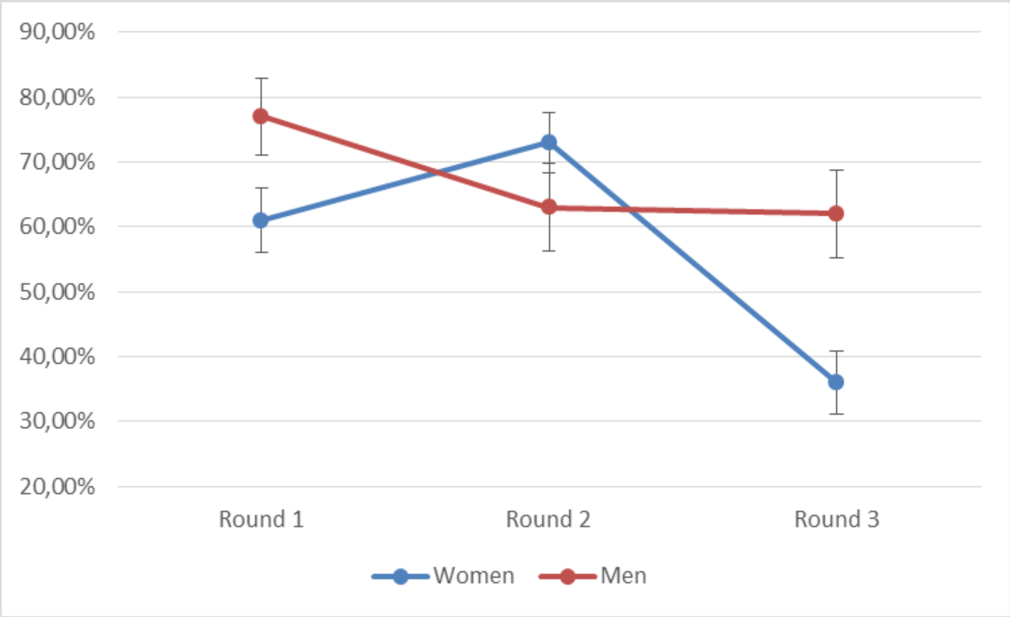


Figure 13. Percentage of difficult task choices made in the incentivised treatment



Note. Error bars represent standard errors. Y axis presents the percentage of challenging task choices each round for each treatment group.

The descriptive results of performance are presented in Table 11 and Table 12. In the non-incentivised treatment we see that men outperform women in performance of difficult tasks in all three rounds. In the non-incentivised treatment, the percentage of correctly solved difficult tasks for women vs. men is 47.0% vs. 56.0% in Round 1, 49% vs. 62% in Round 2 and 34.0% vs. 44.0% in Round 3. In the incentivised treatment, women are more successful in the difficult task in Round 1 (62.0% vs. 56%). Men perform better in difficult tasks than women in Round 2 (62% vs. 50%) and Round 3 (44% vs. 31%). Although the overall performance in difficult task in the incentivised treatment is higher for men than for women, the final financial outcome for men is not significantly higher. This occurs because, in the incentivised treatment that determines the final financial outcome, women select easy tasks more often (except in Round 2), and their overall performance (easy and difficult tasks together) is better. This is especially true for Round 3, where the difference between successfully solved tasks is the highest (78% women vs. 65% men). The results of the χ^2 square test in the cross tabs analysis indicate that the differences in overall performance and performance in difficult tasks do not differ significantly in any of the rounds and in any of the treatments. The average financial outcome of women was EUR 3.63 and of men EUR 3.65 (the difference is not significant). The results are interesting, as they reveal that, although the choices for the difficult task in the incentivised treatment were more frequent for men in two out of three rounds, the overall financial outcome is similar for both genders.

1.14.2 Testing hypotheses

The results of the GEE model are presented in Table 13. The main effects cannot be interpreted without taking into consideration the levels of the other dependent variable (Aiken and West, 1991). The beta coefficient of each variable corresponds to the effect of that variable while holding the value of the other dependent variable in the model at zero. Hence, the coefficients of Reward in the model indicate the impact of the reward on the choices of women, as gender is coded as zero for women and one for men. The impact of the reward on decisions of men is estimated by summing up the coefficient of reward and the interaction coefficient Reward*Gender. The coefficient of gender indicates the difference between women and men in a non-incentivised treatment which is coded as zero (whereas the incentivised treatment is coded as one). The sum of the coefficient for gender and the interaction coefficient indicate the difference between women and men in the incentivised treatment.

At low task difficulty, we did not hypothesise that the effect of reward would be different for women and men. Despite almost significant differences in choices between women and men in the χ^2 square analysis, Table 13 (Round 1) when we estimate a comprehensive model, the model shows no main effect of gender, nor interactive effects of gender and rewards. When the task is very easy, task selection between women and men does not differ significantly, nor it can be stimulated by (small) rewards. We also do not find BIS and BAS to influence the choices.

Our first hypothesis (H1) was that, at moderate task difficulty, rewards would affect men's willingness to undertake a difficult task more positively. Our findings are contrary to our expectations. The model confirms the descriptive statistics in Round 2, in which we observe a higher preference for a difficult task by women in the incentivised setting. In the model where we compare the decisions in the incentivised and non-incentivised setting, we find a significant positive effect of the reward ($b = 0.874$, $OR = 2.396$, $z = 2.480$, $p = 0.013$). The effect of reward on men is explained by the interaction term Reward×Gender interaction term ($b = -1.199$, $OR = 0.301$, $z = -2.030$, $p = 0.591$). The interpretation of regression coefficients cannot be done without taking into account the values of the interacting variables (Lamina et al., 2012). The effect of reward on men is a sum of the coefficent of Reward and the coefficient of the interaction term Reward×Gender interaction term: $(0.874 + (-1.199) = -0.325$. The negative value indicates the negative effect of the reward on men. We can explain the results by using OR. In the non-incentivised treatment, the odds for choosing the difficult task are 1.7 times larger for men ($OR = 1.699$), and in the incentivised treatment, the odds for choosing the difficult task are two and half times times ($OR = 2.396$) greater for women. This rejects our hypothesis that the rewards would have a greater impact on men at moderate task difficulty.

At high task difficulty, we hypothesised (H2) that rewards will have a more negative effect on women than on men. In Round 3, we find a significant negative effect of reward (b= -0.700, OR = 0.497, z = -2.590, p = 0.010). Since the Reward×Gender interaction term (b= 0.700 , OR = 2.014, z = 1.550, p = 0.122) is non-significant, we must interpret that the negative effect of Reward affects both men and women. The odds for choosing the difficult task for women in the incentivised treatment are 0.5 times lower than the odds of choosing the easy task. For men the OR equals 0, meaning that the odds for choosing the difficult task remain the same. The OR 1.714 for variable Gender indicates that the odds for choosing the difficult task in the non-incentivised setting are 1.8 times larger for men. In the incentivised setting, the odds are 3.5 times greater for men than women[15]. The figures based on descriptive statistics show main effect of gender, whereas the models show that the main effect of gender disappears when including other explanatory variables, taking into account autocorrelation of responses in the incentivized and the non-incentivized condition and the interaction effect between gender and incentives. Despite of large differences between responses of men and women in Round 3 the incentives do not significantly differently affect men and women (they decrease motivation of women and leave it intact for men, i.e. they do not increase it), this is why the interaction is insignificant.

In H3 we predict that BAS will affect choices of challenging tasks positively. We do not find a significant effect of BAS in any of the rounds, but can observe that the effect of BAS increases throughout the rounds. In H4 we predict that BIS will affect choices of challenging tasks negatively. We observe a negative main effect of BIS (b = -0.197, OR = 0.821, z = -0.560, p = 0.574) only in Round 3, but it is not significant. BAS is not significantly different between genders.

Finally, in H5, we hypothesised about the effect of failure or success in the difficult task in a a prior round on task choice. We proposed a positive association, meaning that success in a prior difficult task will affect the choice of the difficult task in the subsequent round positively, and a failure will affect it negatively. We find a strong positive effect of the variable Diffcorrect in both rounds in which it was applicable. For Round 2: b = 1.926, OR = 6.862, z = 6.720, p = 0.000 and for Round 3: b = 1.722, OR = 5.596, z = 5.480 p = 0.000, suggesting that those participants that have chosen the difficult task in the previous rounds and solved it successfully, were encouraged, and were more likely to choose a difficult task again. This is the most robust finding in the present chapter and suggests that feedback on prior outcome is more important than gender differences and BIS/BAS in the selection of challenging tasks in the future.

Additionally, we controlled the sequence of letters achieved in the calibration phase. We find persistently that subjects with high result in the calibration phase decided less

[15](0.539+0.700 = 1.293 OR = 3.452)

frequently for the difficult tasks. We assume that individuals that had already exerted maximum effort in the calibration phase felt that they had already reached their cognitive limits and had less chances to improve their performance.

Table 12. GEE model

	Difficult task choice	OR	b	S.E.	z	P(z)
Round 1	Intercept	30.205	3.408	1.796	1.900	0.058
	Gender	1.267	0.237	0.388	0.610	0.542
	Reward	0.831	-0.185	0.275	-0.670	0.501
	Reward*Gender	1.799	0.587	0.493	1.190	0.234
	BAS	0.569	-0.563	0.404	-1.390	0.163
	BIS	1.166	0.154	0.306	0.500	0.615
	Calibration	0.767	-0.265	0.125	-2.120	**0.034**
Round 2	Intercept	0.631	-0.460	1.768	-0.260	0.795
	Gender	1.699	0.530	0.421	1.260	0.208
	Reward	2.396	0.874	0.353	2.480	**0.013**
	Reward*Gender	0.301	-1.199	0.591	-2.030	**0.043**
	BAS	1.051	0.050	0.392	0.130	0.900
	BIS	1.556	0.442	0.306	1.440	0.149
	Calibration	0.747	-0.292	0.122	-2.390	**0.017**
	Diffcorrect	6.862	1.926	0.286	6.720	**0.000**
Round 3	Intercept	5.058	1.621	1.932	0.840	0.402
	Gender	1.714	0.539	0.409	1.320	0.188
	Reward	0.497	-0.700	0.270	-2.590	**0.010**
	Reward*Gender	2.014	0.700	0.453	1.550	0.122
	BAS	1.175	0.161	0.439	0.370	0.713
	BIS	0.821	-0.197	0.350	-0.560	0.574
	Calibration	0.651	-0.430	0.143	-3.020	**0.003**
	Diffcorrect	5.596	1.722	0.314	5.480	**0.000**

Note. For round 1: Wald χ^2 (6) = 9.6, χ^2 = 0.143; For round 2: Wald χ^2 (7) = 54.55, χ^2 = 0.000; For round 3: Wald χ^2 (7) = 43.58, χ^2 = 0.000.

1.15 Discussion and conclusion

Our inquiry was inspired by the question whether gender pay gap and under-presence of women in high profile jobs (Croson and Gneezy, 2005) are linked to the gender differences in undertaking challenging tasks. As prior literature suggests, women and men respond differently to challenging tasks, and that their preferences are driven by differences in personality traits, risk attitude and beliefs in their abilities (Niederle and Yestrumskas, 2008). Women are less inclined to risk and more careful in making high reaching decisions. They are more afraid of punishment and of omission of a reward, and react differently than men to rewards where asuccessful outcome is not certain. Our objective was to analyse the sensitivity to rewards on engagement in risk-effort tasks in women and men controlling avoidance and approach motivational systems, the feedback on prior performance (Diffcorrect) and the individual's working memory capacity (Calibration).

The key findings that have emerged from our work are the following. First, although the descriptive statistics suggest that, in a non-incentivised setting, men are more inclined to select the difficult task than women, the results are not significant, as the variation is high in both subsamples. We tried to see if the cause for variation can be explained with BIS and BAS, but did not find a mediation effect. In the incentivised setting, we find, surprisingly, that, at moderate task difficulty, women are more motivated by the reward than men. This finding may be paralleled with the findings in the literature that women react negatively to incentives only in competitive (tournament) environments, but not in piece rate or performance based evaluation context (Bracha and Fershtman, 2013; Gneezy et al., 2003; Masclet et al., 2015). For women, rewards may have contributed to extrinsic motivation. At the same time, this difficulty level seems not to have created perception of too high uncertainty and women's confidence issues did not arise. Women had more scope to increase performance at this level than men who, without incentives, more frequently than women chose the difficult task. The results for men are hard to explain. Perhaps the reward at moderate task difficulty for men was not large enough to trigger the approach state (Tan et al., 2017), and relatively small rewards decreased their intrinsic motivation. Replications would be needed to see if the results are robust.

At very high task difficulty, the reward had a negative effect on both, although more so on women. This result is in line with our hypothesis. Women were considerably demotivated by the incentive in the third round, whereas, without the incentive, they did not seem to perceive the setting as a financial gain domain. Women are more risk averse than men, and the outcome risk of the difficult task may have activated their BIS system that caused aversive responses. Future research may well look into the intertwine between risk aversion and BIS/BAS.

We saw that prior outcome contributes significantly to the explanation of the choices, and that it is actually the dominant explanatory variable in the model. We confirm the hypothesis proposing the importance of the prior outcome on engagement in difficult tasks. Individuals rely greatly on their own evaluation of probability of success that is created by prior performance feedback. Although BIS and BAS *per se* do not have a significant effect on task choice, as suggested in the literature, the positive feedback may have activated BAS, and the negative feedback may have activated BIS. The effect of negative prior outcome may, thus, be stronger for high BIS individuals, because the negative feedback increased the already existing sensitivity to negative stimuli. Our sample size did not allow us to estimate more elaborate models.

Our chapter's contribution to practice is to expand the awareness on the differential impact of incentives on men and women. Our descriptive statistics, that show insignificant differences in financial outcome for men and women, despite different choices, suggest that more gender balanced recruitment of managers could contribute to less volatile and more stable performance without sacrificing its magnitude. We saw that women do not differ from men in BAS, but only in BIS. Organisations will only use the full potential of women if they will create performance evaluation environments that would take into consideration the systematic differences between men and women, and encourage women to undertake challenging tasks. Utmost importance in performance evaluation systems should be given to the communication of performance feedback. Whereas positive performance feedback will motivate individuals for more demanding tasks, negative performance feedback should be formulated carefully such that it will not ruin their motivation in the future.

1.15.1 Limitations and suggestions for future research

We have already mentioned a few limitations that led to the results we obtained. There are some others that merit discussion. Methodological limitations are part of experiments. Firstly, the research was administered in a laboratory setting. This setting could raise concerns about the ecological and external validity of the conclusions, although we believe that experimental realism was achieved as the task was cognitively engaging. Participants reported to have been motivated to participate in the research and their performance was paid with real money. Using incentives on a large sample is costly and subject to budget constraints. Had we changed the magnitude of rewards, we might have obtained different results in the second round. However, we believe that, at high difficulty level, we would obtain more or less the same results, even with higher rewards and a larger sample, as they are in line with the personality traits of men and women reported robustly in the prior literature. The result on the prior outcome is strong and validated throughout the doctoral thesis, and we think that is one of the main contributions of the chapter. To overcome the problem of low rewards, we expressed them in experimental units. Despite that, the total compensation calculated for an hourly basis seems acceptable, and had worked in our previous two experiments of this thesis.

Next, using student samples in research is often seen as a limitation. As reported by (De Pater, Van Vianen, Fischer and Van Ginkel, 2009), in studies focused on task choice and job challenge, the use of pre-occupational samples is an advantage because the research shows that already, before the professional career, men and women differ with regard to the tasks they choose to perform. These differences get more accentuated with work experience (Valian, 1998). Therefore, we think that the use of a student sample is appropriate.

Overall, we find that women may well be susceptible to rewards at moderate task difficulty, but less so if they are uncertain. In the long run they take decisions which will not be financially inferior to those of men, but may even contribute to smaller volatility and their sustainability. We also believe that accounting literature should start accepting replications of experimental studies in other contexts, tasks and samples that would ensure validity and generalization of these and other experimental findings.

CONCLUSION

This PhD thesis aims to analyse the effect of incentive schemes on decisions to invest cognitive effort in a multi-period setting where the level of cognitive effort demand is increasing. The research combines theories from economics, management accounting, social psychology and insights from Neuroscience. The effect of financial incentives on human behaviour is one of the fundamental questions, not only in Economics (eg. Bénabou and Tirole, 2003; Kreps, 1997; Lazear, 1986, 2000; Prendergast, 1999), but also in psychology and management literature (e.g., Gerhart and Rynes, 2003; Gomez-Mejia and Welbourne, 1988; Rynes et al., 2005; Vroom, 1964). Generally it is assumed that incentives are the dominant force of performance and productivity, although the role of rewards has been controversial in the literature and practice (workplaces, school). Different forms of incentives are used in organisations, and we aim to analyse the effectiveness of incentive schemes in the long run and determine the importance of personality traits in the effect of incentives on cognitive effort.

The hypotheses were tested experimentally. In a series of multi-period choice task experiments we manipulated individuals` reference points by creating a gain or loss domain by labelling performance pay as a reward or as a penalty. We operationalised joint cognitive risk-effort decisions as choices between a high-yielding task that requires high effort with a higher chance of failure (a difficult task) and a low-yielding task that requires low effort with a lower chance of failure (an easy task). We use a modified Sternberg task (Sternberg, 1966) that is used broadly in psychology to measure cognitive effort.

In the first chapter, we explore how cognitive frames created by incentive design (bonus vs. penalty) and the outcome's fairness (fair outcome vs. unfair outcome) influences decisions on risk and effort. The first contribution of this chapter is examination of simultaneous risk-effort decisions building on two influential theories: The prospect theory (Kahneman and Tversky, 1979; Kahneman, 2003) and organisational justice theory (Adams, 1963; Akerlof and Yellen, 1990), and integrating these two influential psychological theories with the management accounting literature and practice. The second contribution of this chapter is in analysing how individuals consider more than one cognitive frame at a time. Our findings indicate that the bonus and penalty schemes invoke cognitive frames in line with the prospect theory, which adds to the evidence on how various incentive practices shape cognitive frames. We show that, when multiple frames interact, they stimulate different behaviour to that elicited by a single cognitive frame. The third contribution is the analysis of risk-effort decision in a multiperiod setting. We provide evidence that the effect of incentive schemes fades over time as a new salient piece of information emerges (i.e., prior performance) that helps re-evaluate the probability of an outcome. The fourth, practical contribution is in providing implications for designing effective incentive schemes.

In the second chapter, we evaluate the effect of monetary incentives on choices to perform cognitively challenging tasks given individual differences in achievement motivation. We are interested in how increasing task difficulty moderates the relation between incentives and achievement motivation. The first contribution of the chapter is in providing original evidence on incentive sensitivity to achievement motivation. The same incentives affect people with different achievement motivation and failure avoidance differently. Secondly, by using a multi-period within the subject experimental setting in two treatments (reward and penalty), we are able to analyse the interplay of incentives and achievement motivation on cognitively challenging tasks in an experimental setting that is akin to real life situations. Thirdly, we advance the research on the effects of incentives on engagement in cognitive tasks by analysing it for increasing task difficulty. Finally, we believe that the chapter contributes to a cross-disciplinary view that advances our understanding of the functionality of incentives. Because the engagement in cognitively challenging tasks is a driving force of performance in contemporary organisations in which a large number of activities are increasingly cognitively demanding, the practical implications of the second chapter lie in a better understanding of the interplay between monetary incentives, an individual's estimation of the probability of success, and achievement motivation on the selection of challenging tasks.

In the third chapter, we examine the differences in the effect of monetary rewards on gender in a challenging cognitive task controlling for avoidance and approach motivational systems, the feedback on prior performance, and the individual's working memory capacity. We are interested in how rewards motivate men and women at various difficulty levels. Gender differences in risk taking have been explored thoroughly in the literature, and we advance the literature by investigating gender differences in combined *risk-effort* choices in a multi-period setting. Secondly, we reconfirm the findings from the first and second chapter on the importance of prior outcome on decision-making. Thirdly, our study provides evidence that men who, more often than women, chose a task at higher difficulty level and accepted higher risk, do not achieve an overall significantly better performance and higher financial outcome. Our practical contribution is to expand the awareness on the differential impact of incentives on men and women, because organisations will only use the full potential of women if they will create performance evaluation environments that would take into consideration the systematic differences between genders.

In all three chapters, we find that, at low task difficulty, rewards do not increase the willingness to exert cognitive effort, as innate motivational predisposition prevails over relatively small rewards. On the other hand, penalties do have a positive effect on the selection of challenging tasks at low task difficulty. At moderate task difficulty, we observe a positive effect of rewards on women and on individuals with high motivation to achieve. Penalties also influence positively at moderate task difficulty, and even more so if individuals are not afraid of failure. At high task difficulty, rewards have a negative effect, which is attenuated for women, high achievers and individuals with high behavioural inhibition systems. Similarly, penalties have no effect on willingness to engage in cognitively challenging tasks at high difficulty. We find that prior outcome contributes significantly to the explanation of the choices in all three chapters. Overall, our findings implicate that incentives must be studied in multi-period settings, as only such settings can take into account the long-term impact of incentives.

REFERENCES

Adams, J. S. (1963). Towards an understanding of inequity. *Journal of Abnormal Psychology*, *67*(5), 422–436.

Adams, J. S. (1966). Inequity in social exchange. *Advances in Experimental Social Psychology*, *2*(C), 267–299.

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

Akerlof, G. A., & Yellen, J. L. (1990). The fair wage-effort hypothesis and unemployment. *Quarterly Journal of Economics*, *105,* 255–283.

Akerlof, G. A., & Shiller, R. J. (2009). *Animal spirits: How human psychology drives the economy and why this matters for global capitalism*. Princeton: Princeton University Press.

Apicella, C. L., Dreber, A., Campbell, B., Gray, P.B., Hoffman, L., & Little, A. C. (2008). Testosterone and financial risk preferences. *Evolution and Human Behavior, 29*(6), 384–390.

Ariely, D., Gneezy, U., Lowenstein, R., & Mazar, N. (2009). Large stakes and big mistakes.*The Review of Economic Studies, 76*(2), 451-469.

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, *64*(6), 359-372.

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

Atkinson, J. W., & Raynor, J. O. (1974). *Motivation and achievement*. Washington, D.C.: V. H. Winston & Sons.

Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics, 116*(1), 261–292.

Barberis, N., Huang, M., & Tano, S. (2001). Prospect theory and asset prices. *Quarterly Journal of Economics, 116*(1), 1–53.

Barberis, N. & Huang, M. (2006). The loss aversion/narrow framing approach to the equity premium puzzle. In Mehra, R. (Ed.), *Handbook of investments: Equity premium*. Amsterdam: Elsevier Science.

Baskin-Sommers, A. R., Wallace, J. F., MacCoon, D. G., Curtin, J. J., & Newman, J. P. (2010). Clarifying the Factors that Undermine Behavioral Inhibition System Functioning in Psychopathy. *Personality Disorders, 1*(4), 203–217.

Baumeister, R. F. (1984). Choking under pressure: Self-Consciousness and paradoxical effects of incentives on skilful performance. *Journal of Personality and Social Psychology*, *46*(3), 610-620.

Beh, H. C. (1990). Achievement motivation, performance and cardiovascular activity. *International Journal of Psychophysiology*, *10*(1), 39-45.

Bénabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. *Review of Economic Studies*, *70*(3), 489-520.

Benartzi, S., & Thaler, R. (1995). Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics, 110*(1),73–92.

Beyer, S. (1990). Gender differences in the accuracy of self-evaluations of performance. *Journal of Personality and Social Psychology, 59*(5), 960-970.

Bhatti, M., Jang, H., Kralik, J. D., & Jeong, J. (2014). Rats exhibit reference-dependent choice behavior. *Behavioural Brain Research*, *267*, 26–32.

Brink, A. G., & Rankin, F. W. (2013). The effects of risk preference and loss aversion on individual behavior under bonus, penalty, and combined contract frames. *Behavioral Research in Accounting*, *25*(2), 145–170. http://doi.org/10.2308/bria-50408

Birnberg, J. G., Luft, J., & Shields, M. D. (2007). Psychology theory in management accounting research. *Handbook of Management Accounting Research*, *1*, 113–135.

Bogdan, R., & Pizzagalli, D. A. (2006). Acute stress reduces reward responsiveness: implications for depression. *Biological Psychiatry, 60*, 1147–1154.

Bonner, S. E., & Sprinkle, G.B. (2002). The effect of monetary incentives on effort and task performance: Theories, evidence, and a framework for research. *Accounting, Organizations and Society, 27*(4–5), 303-45.

Bonner, S. E., Hastie, R., Sprinkle, G. B., & Young, S. M. (2000). A review of the effects of financial incentives on performance in laboratory tasks: implications for management accounting. *Journal of Management Accounting Research*, *12*(1), 19-64.

Bose, S., Ladley, D., & Xin, L. (2016). The role of hormones in financial markets Available at: http://dx.doi.org/10.2139 /ssrn.2743087.

Botvinick, M. M., Huffstetler, S., & McGuire, J.T. (2009). Effort discounting in human nucleus accumbens. *Cognitive, Affective, & Behavioral Neuroscience, 9*(1), 16-27.

Botvinick, M., & Braver, T. (2015). Motivation and cognitive control: From behavior to neural mechanism. *Annual Review of Psychology, 66*, 83-113.

Bracha, A., & Fershtman, C. (2013) Competitive incentives: Working harder or working smarter? *Management Science, 59*(4), 771-781.

Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E. E., & Noll, D. C. (1997). A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage, 5*(1), 49-62.

Braver, T. S., Paxton J. L., Locke H. S., & Barch D. M. (2009). Flexible neural mechanisms of cognitive control within human prefrontal cortex. *Proceedings of the National Academy of Sciences, 106*(18), 7351-7356.

Brooks, R. R. W., Stremitzer, A., & Tontrup, S. (2013). *Stretch it but don't break it: the hidden risk of contract framing* (No. 13–22). *Law-Econ Research Paper*. UCLA School of law. Retrieved from http://dx.doi.org/10.2139/ssrn.2353733.

Brosnan, S. F., & De Waal, F. B. M. (2003). Monkeys reject unequal pay. *Nature, 425*(6955), 297–299.

Brown, J. L., Farrington, S., & Sprinkle, B. G. (2016). Biased self-assessments, feedback, and employees' compensation plan choices. *Accounting, Organizations and Society*, *54*, 45-59.

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

Byrnes, J. P., Miller, D. C. & Schafer, W. D. (1999). Gender differences in risk taking: A Meta-analysis. *Psychological Bulletin, 125*(3), 367-383.

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin*, *115*(3), 401–423.

Cadbsy, C.B., Song, F., & Tapon, F. (2010). Are you paying your employees to cheat? An experimental investigation. *The B.E. Journal of Economic Analysis & Policy, 10*(1), 1-32.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology, 67*(2), 319-333.

Charness, G., and Kuhn, P. (2007). Does pay inequality affect worker effort? Experimental evidence. *Journal of Labor Economics*, *25*(4), 693–723.

Chiew, K. S., & Braver, T. S. (2013). Temporal dynamics of motivation-cognitive control interactions revealed by high-resolution pupillometry. *Frontiers in Psychol*ogy, *4*, 15. doi.org/10.3389/fpsyg.2013.00015.

Chow, C. W., Kohlmeyer, J. M., & Wu, A. (2007). Performance standards and managers' adoption of risky projects. *Advances in Management Accounting*, *16*, 63–105.

Christ, M. H., Sedatole, K. L., & Towry, K. L. (2012). Sticks and carrots: The effect of contract frame on effort in incomplete contracts. *Accounting Review*, *87*(6), 1913–1938.

Church, B. K., Libby, T., & Zhang, P. (2008). Contracting frame and individual behavior: Experimental Evidence. *Journal of Management Accounting Research*, *20*(1), 153–168.

Cocker, P. J., Hosking, J. G., Benoit, J., & Winstanley, C. A. (2012). Sensitivity to cognitive effort mediates psychostimulant effects on a novel rodent cost/benefit decision-making task. *Neuropsychopharmacology*, *37*(8), 1825–1837.

Corr, P. J., & McNaughton, N. (2012). Neuroscience and approach/avoidance personality traits: a two stage (valuation-motivation) approach. *Neuroscience & Biobehavioral Review*, *36*, 2339–2354. doi: 10.1016/j.neubiorev.2012.09.013

Corr, P. J. (2013). Approach and Avoidance Behaviour: Multiple Systems and their Interactions. *Emotion Review, 5*(3), 285–290.

Corr, P. J., McNaughton, N., Wilson, M. R., Burch, G., & Poropat, A. (2017). Putting the Reinforcement Sensitivity Theory (RST) to work. In S. Kim, J. M. Reeve & M. Bong (Eds.), *Recent developments in neuroscience research on human motivation*: *Advances in motivation and achievement* (Vol. 19, pp. 65-92). Bingley: Emerald Group Publishing.

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature, 47*(2), 1–27.

Cross C. P., Copping L. T., & Campbell, A. (2011). Sex differences in impulsivity: a meta-analysis. *Psychological Bulletin, 137*(1), 97–130.

Cueva, C., R. Roberts, R. E., Spencer, T., Rani, N., Tempest, M., Tobler, P.N., Herbert, J., & Rustichini, A. (2015). Cortisol and testosterone increase financial risk taking and may destabilize markets. *Scientific Reports, 5*, 11206.

Cullis, J., Jones, P. & Lewis, A. (2006). Tax framing, instrumentality and individual differences: Are there two different cultures? *Journal of Economic Psychology*, *27*(2), 304–320.

Datta Gupta, N., Poulsen, A., & Villeval, M. C. (2013). Gender matching and competitiveness: Experimental evidence. *Economic Inquiry, 51*(1), 816–835.

David, M., Peterle, E., & Larribeau, S. (2015). Gender differences in tournament and flat-wage schemes: An experimental study. *Journal of Economic Psychology, 47*, 103-115,

De Pater, I. E., Annelies, E. M., Fischer, A. H., & Van Ginkel, W. (2009). Challenging experiences: gender differences in task choice. *Journal of Managerial Psychology, 24*(1), 4–28. http://doi.org/10.1108/02683940910922519

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). Meta-analytic review of experiments examining the effects of extrinsic reward and intrinsic motivation. *Psychological Bulletin*, *125*(6), 627-668.

Deepika, A., Oehlberg, K.A., Treadway, M.T., & Nusslock, R. (2016) Effect of failure/success feedback and the moderating influence of personality on reward motivation. *Cognition and Emotion, 30*(3), 458-471,

Diecidue, E., & van de Ven, J. (2008). Aspiration level, probability of success and failure, and expected utility. *International Economic Review*, *49*(2), 683–700.

Diggle, P. J., Liang, K, Y., & Zeger, S. L. (1994). *Analysis of Longitudinal Data*. Oxford: Oxford University Press.

Ding, Y., Wang, E., Zou, Y., Song, Y., Xiao, X., Huang, W., & Li, Y. (2017). Gender differences in reward and punishment for monetary and social feedback in children: An ERP study. *Plos ONE, 12*(3), e0174100. Available at: https://doi.org/10.1371/journal.pone.0174100

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2009). Homo Reciprocans: Survey evidence on behavioral outcomes. *Economic Journal, 119*(536), 592–612.

Drake, A. R., & Kohlmeyer, J. M. (2010). Risk-taking in new project selection: Additive effects of bonus incentives and past performance history. *Advances in Accounting, 26*(2), 207–220.

Druckman, J. N. (2001). Evaluating framing effects. *Journal of Economic Psychology, 22*(1), 91–101.

Dweck, C. S. (2000). *Self-Theories: Their role in motivation, personality, and development*. Philadelphia: Psychology Press.

Eckel, C., & Grossman, P. J. (2008). Men, women and risk aversion: experimental evidence. In C. Plott & V. Smith (Eds.), *Handbook of experimental economics Results* Vol.1 (pp. 1061–1073). New York: Elsevier.

Economist (2017). Families and work: Having it all. *Economist*, October 7th, 13.

Engelmann, J. B., Damaraju, E., Padmala S., & Pessoa L. (2009). Combined effects of attention and motivation on visual task performance: transient and sustained motivational effects. *Frontiers in Human Neuroscience, 3,* 4. doi.org/10.3389/neuro.09.004.2009.

Exadaktylos, F., Espín, A. M., & Brañas-Garza, P. (2013). Experimental subjects are not different. *Scientific Reports*, *3*, 1213. doi:10.1038/srep01213

Fagley, N. S., & Miller, P. M. (1990). The effect of framing on choice interactions with risk-taking propensity, cognitive style, and sex. *Personality and Social Psychology Bulletin, 16*(3), 496-510.

Fagley, N. S., Coleman, J. G., & Simon, A. F. (2010). Effects of framing, perspective taking, and perspective (affective focus) on choice. *Personality and Individual Differences, 48(3),* 264–269. http://doi.org/https://doi.org/10.1016/j.paid.2009.10.008.

Fehr, E., & Gächter, S. (2001). *Do incentive contracts crowd out voluntary cooperation?* Working Paper. Institute for Empirical Research in Economics, University of Zurich.

Fehr-Duda, H., De Gennaro, M., & Schubert, R. (2006). Gender, financial risk, and probability weights. *Theory and Decision. 60*(2), 283–313. http://doi.org/10.1007/s11238-005-4590-0.

Frey, B. S., & Oberholzer-Gee, F. (1997). The cost of price incentives: An empirical analysis of motivation crowding-out. *American Economic Review, 87*(4), 746-755.

Fröber, K., & Dreisbach G. (2014). The differential influences of positive affect, random reward, and performance-contingent reward on cognitive control. *Cognitive Affective & Behavioral Neuroscience, 14*(2), 530-547.

Gabriel, K.I., & Willliamson, A. (2010). Framing alters risk-taking behavior on a modified Balloon Analogue Risk Task (BART) in a sex-specific manner. *Psychological Reports, 107*(3), 699-712.

Garbarino, E., Slonim, R., & Sydnor, J. (2011). Digit ratios (2D:4D) as predictors of risky decision making. *Journal of Risk and Uncertanity, 42*(1), 1–26.

Gendolla, G. H. E., & Wright, R. A. (2005). Motivation in social settings: studies of effort-related cardiovascular arousal. In J.P. Forgas, K. Williams and W. von Hippel, (Eds.), *Social motivation* (pp. 71-90). New York: Cambridge University Press.

Gibbons, R. (1998). Incentives in organizations. *Journal of Economic Perspectives, 12*(4), 115–132.

Gill, D., & Prowse, P. (2014) Gender differences and dynamics in competition: The role of luck. *Quantitative Economics, 5*(2), 351–376.

Gjesme, T., & Nygard, R. (1970). *Achievement-related motives: Theoretical considerations and construction of a measuring instrument.* Unpublished report. Institute for Educational Research, University of Oslo.

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. *Quarterly Journal of Economics, 118*(3), 1049 – 1074.

Gomez–Mejia, L. R., & Welbourne, T. M. (1988). Compensation strategy: An overview and future steps. *Human Resource Planning, 11*(3), 173–190

Gose, K., & Sadrieh, A. (2012). Fair wages when employers face the risk of losing money. *Economics Letters, 117*(3), 687-690.

Grabner, I., & Speckbacher, G. (2016). The cost of creativity: A control perspective. *Accounting, Organizations and Society, 48*, 31-42.

Gray, J. A. (1981). A critique of Eysenck's theory of personality, In H.J. Eysenck (Ed.*), A model for personality* (pp. 246–276). Berlin: Springer Berlin Heidelberg.

Gray, J. A. (1982). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system.* Oxford: Oxford University Press.

Gray, J. A., & McNaughton, N. (2000). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system, 2nd ed*. Oxford: Oxford University Press.

Haisley, E., Mostafa, R., & Loewenstein, G. (2008). Subjective relative income and lottery ticket purchases. *Journal of Behavioral Decision Making*, *21*(3), 283–295.

Hannan, R. L., Hoffman, V. B., & Moser, D. V. (2005). Bonus versus penalty: does contract frame affect employee effort? In A. Rapoport and R. Zwick (Eds.), *Experimental business research* (pp. 151–169). Springer US.

Harackiewicz, J. M., & Manderlink, G. (1984). A process analysis of the effects of performance-contingent rewards on intrinsic motivation. *Journal of Experimental Social Psychology, 20*(6), 531-551.

Harackiewicz, J. M., Manderlink, G., & Sansone, C. (1984). Rewarding pinball wizardry: Effects of evaluation and cue value on intrinsic interest. *Journal of Personality and Social Psychology, 47*(2), 287-300.

Harackiewicz, J. M., & Elliott, A. J. (1993). Achievement goals and intrinsic motivation. *Journal of Personality and Social Psychology, 49*(6), 1753-1761.

Harshman, R. A., & Paivio, A. (1987). Paradoxical' sex differences in self-reported imagery. *Canadian Journal of Psychology, 41*(3), 287-302.

Hartmann, F., & Slapničar, S. (2012). Pay fairness and intrinsic motivation: the role of pay transparency. *The International Journal of Human Resource Management*, *23*(20), 4283–4300.

Hartmann, F., & Slapničar, S. (2014). An experimental study of the effects of negative, capped and deferred bonuses on risk taking in a multi-period setting. *Journal of Management and Governance*, 1–22. http://doi.org/10.1007/s10997-014-9297-6

Heckhausen, H. (1963). *Hoffnung und furcht in der leistungsmotivation*. Meisenheim/Glan. Germany: Hain.

Heckhausen, H. (1991). *Motivation and action.* New York: Springer.

Hickman, D. C., & Metz, N. E. (2015). The impact of pressure on performance: Evidence from the PGA TOUR. *Journal of Economic Behavior and Organization*, *116*, 319-330.

Homish, G. G., Edwards, E. P., Eiden, R. D., & Leonard, K. E. (2010). Analyzing family data: A GEE approach for substance use researchers. *Addictive Behaviors, 35*(6), 558–563.

Hosking, J. G., Floresco, S. B., & Winstanley, C. A. (2014). Dopamine antagonism decreases willingness to expend physical, but not cognitive, effort: a comparison of two rodent cost/benefit decision-making tasks. Neuropsychopharmacology, *40*(4), 1005-1015.

Hossain, T., & List, J. A. (2012). The behavioralist visits the factory: Increasing productivity using simple framing manipulations. *Management Science*, *58*(12), 2151–2167.

Hughes, A. E., Crowell, S. E., Uyeji, L., & Coan, J. A. (2012). A developmental neuroscience of borderline pathology: emotion dysregulation and social baseline theory. *Journal of Abnormal Child Psychology, 40*(1), 21–33.

Hutchinson, J. C., Sherman, T., Martinovic, N., & Tenenbaum, G. (2008). The effect of manipulated self-efficacy on perceived and sustained effort. *Journal of Applied Sport Psychology, 20*(4), 457–472.

Ilies, R., De Pater, I.E., & Judge, T. (2007). Differential affective reactions to negative and positive feedback, and the role of self-esteem. *Journal of Managerial Psychology, 22*(6), 590-609. doi.org/10.1108/02683940710778459

Jansma, J. M., Ramsey, N. F., de Zwart, J. A., van Gelderen, P., & Duyn, J. H. (2007). fMRI study of effort and information processing in a working memory task. *Human Brain Mapp*, *28*(5), 431-440.

Jentsch, J. D., Woods, J. A., Groman, S. M., & Seu, E. (2010). Behavioral characteristics and neural mechanisms mediating performance in a rodent version of the Balloon analog risk task. *Neuropsychopharmacology, 35*(8), 1797–1806.

Jimura, K., Locke, H. S., & Braver, T. S. (2010). Prefrontal cortex mediation of cognitive enhancement in rewarding motivational contexts. *Proceedings of the National Academy of Sciences, 107*(19), 8871-8876.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, New York: Prentice-Hall.

Kahneman, D., & Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, *47*(2), 263-291.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, *93*(5), 1449-1475.

Kamas, L., & Preston, A. (2012). The importance of being confident; gender, career choice, and willingness to compete. *Journal of Economic Behavior and Organization, 83*(1), 82–97.

Kilch, N. R. & Feldman, D. C. (1992). The role of approval and achievement needs in feedback seeking behavior. *Journal of Managerial Issues*, *4*(4), 554-570.
Kivikangas, J. M., Kätsyri, J., Järvelä, S., & Ravaja, N. (2014). Gender differences in emotional responses to cooperative and competitive game play. *Plos ONE, 9*(7), e100318. Retrieved from https://doi.org/10.1371/journal.pone.0100318

Kleinbaum, D. G. & Klein, M. (2010). *Logistic regression*. *A self learning text.* New York: Springer.

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology General*, *139*(4), 665–682.

Kool, W., & Botvinick, M. (2014). A labor/leisure tradeoff in cognitive control. *Journal of Experimental Psychology*, *143*(1), 131–141. doi:10.1037/a0031048.

Kreps, D. M. (1997). Intrinsic motivation and extrinsic incentives. *American Economic Review, 87*(2), 359-64.

Kruglanski, A. W., Bélanger, J. J., Chen, X., Köpetz, C., Pierro, A., & Mannetti, L. (2012). The energetics of motivated cognition: A force-field analysis. *Psychological Review, 119*(1), 1–20.

Kukla, A. (1972). Attributional determinants of achievement-related behaviour. *Journal of Personality and Psychology*, *21*(2), 166-174.

Kunz, A. H., & Pfaff, D. (2002). Agency theory, performance evaluation, and the hypothetical construct of intrinsic motivation. *Accounting, Organizations and Society, 27*(3), 275-295.

Kunz, J., & Linder, S. (2012) Organizational control and work effort – another look at the interplay of rewards and motivation. *European Accounting Review*, *21*(3), 591-621.

Lang, J. W. B., & Fries, S. (2006) A revised 10-item version of the Achievement Motives Scale: Psychometric properties in German-speaking samples. *European Journal of Psychological Assessment*, *22*(3), 216-224.

Lazear, E. P. (2000a). The Power of Incentives. *American Economic Review*, *90*(2), 410–415.

Lazear, E. P. (2000b). Performance Pay and Productivity. *American Economic Review, 90*(5), 1346–1361.

Lerner, J. S., Gonzalez, R. M., Small, D. A., & Fischhoff, B. (2003). Effects of fear and anger on perceived risks of terrorism: A aational field experiment. *Psychological Science, 14*(2), 144–50.

Li, Y., Qiao, L., Sun, J., Wei, D., Li, W., Qiu, J., Zhang, Q., & Shi, H. (2014). Gender-specific neuroanatomical basis of behavioral inhibition/approach systems (BIS/BAS) in a large sample of young adults: A voxel-based morphometric investigation. *Behavioural Brain Research, 274*(Supplement C), 400–408.

Libby, T. (2001). Referent Cognitions and Budgetary Fairness: A Research Note. *Journal of Management Accounting Research*, *13,* 91–105.

Ličen, M., Hartmann, F., Repovš, G., & Slapničar, S. (2016). The impact of social pressure and monetary incentive on cognitive control. *Frontiers in Psychology*, *7*, 93. doi: 10.3389/fpsyg.2016.00093.

Lighthall, N. R., Sakaki, M., Vasunilashorn, S., Nga, L., Somayajula, S., Chen, E.Y., Samii, N., Mather, M. (2011). Gender differences in reward-related decision processing under stress. Social Cognitive and Affective Neuroscience, *7*(4), 476-484.

Linde, J., & Sonnemans, J. (2012). Social comparison and risky choices. *Journal of Risk and Uncertainty*, *44*(1), 45–72. http://doi.org/10.1007/s11166-011-9135-z

Locke, E.A. (1991). The motivation sequence, the motivation hub, and the motivation core. *Organisational Behaviour and Human Decision Processes*, *50*(2), 288-29.

Locke, H. S., & Braver, T. S. (2008). Motivational influences on cognitive control: behavior, brain activation, and individual differences. *Cognitive, Affective, & Behavioral Neuroscience, 8*(1), 99–112.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin, 127*(2), 267-286.

Mas, A. (2006). Pay, reference points, and police performance. *The Quarterly Journal of Economics*, *121*(3), 783–821.

Massar, S. A. A., Libedinsky, C., Weiyan, C., Huettel, S. A., & Chee, M. W. L. (2015). Separate and overlapping brain areas encode subjective value during delay and effort discounting. *NeuroImage, 120,* 104-113.

Massar, S. A. A., Libedinsky, C., Weiyan, C., Huettel, S. A., & Chee, M. W. L. (2015). Separate and overlapping brain areas encode subjective value during delay and effort discounting. *NeuroImage, 120* (Supplement C), 104–113.

Maule, J., & Villejoubert, G. (2007). What lies beneath: Reframing framing effects. *Thinking and Reasoning*, *13*(1), 25–44.

Mather, M., & Lighthall, N. R. (2012). Both risk and reward are processed differently in decisions dade under stress. *Current Directions in Psychological Science, 21*(2), 36–41.

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

McClelland, D. C. (1987). *Human motivation*. Cambridge University Press.

McGuire, J., & Botvinick, M. (2010). Prefrontal cortex, cognitive control, and the registration of decision costs. *Proceedings of the National Academy of Sciences*, *107*(17), 7922-7926.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81-97.

Miller, M. A., Thomé, A., & Cowen, S. L. (2013). Intersection of effort & risk: ethological and neurobiological perspectives. *Frontiers in neuroscience*, *7*, 208. doi:10.3389/fnins.2013.00208

Möbius, M., Niederle, P., Niehaus, T. S. & Rosenblat, D. (2013). Managing self-confidence: theory and experimental evidence. NBER Working Papers 17014. National Bureau of Economic Research. Available at: http://www.nber.org/papers/w17014.pdf.

Murray, H. A. (1936). Facts which support the concept of need or drive. *Journal of Psychology*, *3*, 27-42 [Reprinted in D.C. McClelland & R.S. Steele, (Eds.), *Human motivation: A book of readings* (pp. 17-30). Morristown, NJ: General Learning.

Murayama, K., Matsumoto, M., Izuma, K., & Matsumoto, K. (2010). Neural basis of the undermining effect of monetary reward on intrinsic motivation. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(49), 20911-20916.

Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics, 122*(3), 1067-1101.

Niederle, M. & Yestrumskas, A. H. (2008). Gender differences in seeking challenges: The Role of Institutions. NBER working paper no. 13922. Cambridge. Available at: http://www.nber.org/papers/w13922.

Niederle, M., & Vesterlund, L. (2011). Gender and competition. *Annual Review in Economics*, *3*, 601–630.

Nygard, R. (1981). Toward an interactional psychology: Models from achievement motivation research. *Journal of Personality*, *49*(4), 363-387.

Oblak, K., Ličen, M., & Slapničar, S. (in press). The role of cognitive frames in combined decisions about risk and effort. Management Accounting Research. http://doi.org/10.1016/j.mar.2017.07.001

Ossewaarde, L., Qin, S., Van Marle, H. J., Van Wingen, G. A., Fernandez, G., & Hermans E. J. (2011). Stress-induced reduction in reward-related prefrontal cortex function. *Neuroimage, 55*(1), 345–352. doi10.1016/j.neuroimage.2010.11.068

Padmala, S., & Pessoa, L. (2011). Reward reduces conflict by enhancing attentional control and biasing visual cortical processing. *Journal of Cognitive Neuroscience, 23*(11), 3419-3432.

Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R. J., & Frith, C. D. (2007). How the brain translates money into force: A neuroimaging study of subliminal motivation. *Science*, *316*(5826), 904–906. doi:10.1126/science.1140459.

Pierce, W. D., Cameron, J., Banko, K. M., & So, S. (2003). Positive effects of rewards and performance standards on intrinsic motivation. *Psychological Record*, *53*(4), 561-579.

Platt, M. L., & Huettel, S. A. (2008). Risky business: the neuroeconomics of decision making under uncertainty. *Nature Neuroscience, 11*(4), 398–403.

Pochon J. B., Levy, R., Fossati, P., Lehericy, S., Poline, J. B., Pillon, B., & Dubois, B. (2002). The neural system that bridges reward and cognition in humans: An fMRI study. *Proceedings of the National Academy of Sciences of the United States of America, 99*(8), 5669-5674.

Prendergast, C. (1999). The provision of incentives in firms. *Journal of Economic Literature, 37(*1), 7–63.

Preston, S. D., Buchanan, T. W., Stansfield, R. B., & Bechara, A. (2007). Effects of anticipatory stress on decision making in a gambling task. *Behavioral Neuroscience, 121*(2). 257–263.

Robson, A. J., & Samuelson, L. (2009). The evolution of time preference with aggregate uncertainty. *American Economic Review, 99*(5), 1925–1953.

Ruchala, L. V. (1999). The influence of budget goal attainment on risk attitudes and escalation. *Behavioral Research in Accounting, 11*, 161–191.

Runeson, P. (2003). *Using students as experiment subjects–an analysis on graduate and freshmen student data.* Proceedings of the 7th International Conference on Empirical Assessment in Software Engineering. Keele University, United Kingdom, 95-102.

Rynes, S., Gerhart, B., & Parks, L. (2005). Personnel Psychology: Performance Evaluation and Pay for Performance. *Annual Review of Psychology, 56*, 571–600.

Salamone, J. D., Cousins, M. S., & Bucher, S. (1994). Anhedonia or anergia? Effects of haloperidol and nucleus accumbens dopamine depletion on instrumental response selection in a T-maze cost/benefit procedure. *Behavioural brain research, 65*(2), 221-229.

Salamone, J. D., Correa, M., Farrar, A., & Mingote, S. M. (2007). Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology, 191*(3), 461-482.

Salamone J. D., Correa M., Nunes E. J., Randall P. A., & Pardo M. (2012). The behavioral pharmacology of effort-related choice behavior: dopamine, adenosine and beyond. *Journal of the Experimental Analysis of Behaviour, 97*(1), 125-146.

Schultheiss, O. C., & Brunstein, J. C. (2005). An implicit motive perspective on competence. In A. J. Elliot & C. Dweck (Eds.), *Handbook of competence and motivation* (pp. 31-55). New York: Guilford.

Schwerter, F. (2013). *Social reference points and risk taking* (No. 11). Bonn: Bonn Econ Discussion Papers. Available at: http://www.bgse.uni-bonn.de/bonn-econ-papers-1/archive/2013/bgse11_2013

Smith, V. L. & Walker, J. (1993). Monetary rewards and decision cost in experimental economics. *Economic Inquiry, 31*(2), 245-261.

Solomon, R. L. (1948). The influence of work on behavior. *Psychological Bulletin, 45*(1), 1-40.

Spangler, W. D. (1992). Validity of questionnaire and TAT measures of need for achievement: Two meta-analyses. *Psychological Bulletin, 112*(1), 140-154.

Spielberg, J., Heller, W., & Miller, G. (2013). Hierarchical brain networks active in approach and avoidance goal pursuit. *Frontiers in Human Neuroscience*. Retrieved from http://journal.frontiersin.org/article/10.3389/fnhum.2013.00284.

Sprinkle, G. B. (2003). Perspectives on experimental research in managerial accounting. *Accounting, Organizations and Society, 28*(2), 287–318.

Sprinkle, G. B., Williamson, M. G., & Upton, D. R. (2008). The effort and risk-taking effects of budget-based contracts. *Accounting, Organizations and Society*, *33*(4–5), 436–452. http://doi.org/10.1016/j.aos.2007.11.001.

Sternberg, R. J., (2008). Increasing fluid intelligence is possible after all. *Procedures of the National Academy of Sciences, 105*(19), 6791-6792. doi: 10.1073/pnas.0803396105

Sternberg, S., (1966). High-speed scanning in human memory. *Science, 153*(3736), 652–654.

Tan, X., Van Prooijen, J.W., Proulx, T., Wu, H., Van Beest, I., & Van Lange, P. A. M. (2017). Reacting to unexpected losses in an uncertain world: High approach individuals become even more risk-seeking. *Personality and Individual Differences, 109* (Supplement C), 117–123. http://doi.org/https://doi.org/10.1016/j.paid.2016.12.051

Thaler, R. H., & Johnson, E. J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Science, 36*(6), 643–660.

Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making*, *12*(3), 183–206.

Thorndike, E. L. (1898). Animal intelligence: an experimental study of the associative processes in animals. *The Psychological Review: Monograph Supplements, 2*(4), i.

Treadway, M. T., Buckholtz, J. W., Cowan, R. L., Woodward, N. D., Li, R., Ansari, M. S., & Zald, D. H. (2012). Dopaminergic mechanisms of individual differences in human effort-based decision-making. *The Journal of Neuroscience, 32*(18), 6170–6176.

Treadway, M., Peterman, J., Zald, D., & Park, S., (2015). Impaired effort allocation in patients with schizophrenia. Schizophrenia Research, *161*, 382–5.

Valian, V. (1998). *Why so slow? The advancement of women*. Cambridge: MIT Press.

Van De Weghe, P., & Bruggeman, W. (2006). *The impact of the number of performance measures and incentive framing on performance in a multidimensional task environment* (Working Papers of Faculty of Economics and Business Administration, Ghent University, Belgium No. 06/404). Retrieved from http://econpapers.repec.org/ RePEc:rug:rugwps:06/404

Van den Bos, R., Harteveld, M., & Stoop, H. (2009). Stress and decision-making in humans: Performance is related to cortisol reactivity, albeit differently in men and women. *Psychoneuroendocrinology, 34*(10), 1449–1458

Van der Stede, W. A. (2011). Management accounting research in the wake of the crisis : Some Reflections. *European Accounting Review*, *20*(4), 605–623.

Vinkhuyzen, A. A. E., Van der Sluis, S., Boomsma, D. I., De Geus, E. J. C., & Posthuma, D. (2010). Individual differences in processing speed and working memory speed as assessed with the sternberg memory scanning task. *Behavior Genetics, 40*(3), 315–326.
Vroom, V. H., 1964. *Work and motivation*. Wiley, New York.

Walton, M. E., Kennerley, S. W., Bannerman, D. M., Phillips, P. E. M., & Rushworth, M. F. S. (2006). Weighing up the benefits of work: Behavioral and neural analyses of effort-related decision making. *Neural Networks*, *19*(8), 1302–1314.

Wang, D., Zhu, L., Maguire, P., Liu, Y., Pang, K., Li, Z., & Hu, Y., 2016. The influence of social comparison and peer group size on risky decision-making. Frontiers in Psychology, *7*, 1232.

Wang, Y., Zou, Z., Song, H., Xu, X., Wang, H., d'Oleire Uquillas, F., & Huang, X. (2016). Altered gray matter volume and white matter integrity in college students with mobile phone dependence. Frontiers in Psychology, *7*(1), 597.

Wardle, M. C., Treadway, M. T., Mayo, L. M., Zald, D. H., & De Wit, H. (2011). Amping up effort: Effects of d-Amphetamine on human effort-based decision-making. *Journal of Neuroscience, 31*(46), 16597–16602.

Wardle, M., Treadway, M., & De Wit, H. (2012). Caffeine increases psychomotor performance on the effort expenditure for rewards task. Pharmacology, *Biochemistry, and Behavior, 102*(4), 526–31.

Weiner, Y. (1982). Commitment in organizations: A normative view. *Academy of Management Review, 7*(3), 418-428.

Westbrook, A., Kester, D., & Braver, T. S. (2013). What is the subjective cost of cognitive effort? Load, trait, and aging effects revealed by economic preference. *PloS One, 8*(7), e68120.

Westbrook, A., & Braver, T. S. (2015). Cognitive effort: A neuroeconomic approach. *Cognitive, Affective & Behavioral Neuroscience*, *15*(2), 395-415. doi:10.3758/s13415-015-0334-y.

Westbrook, A., Kester, D., & Braver, T. S. (2013). What is the subjective cost of cognitive effort? Load, trait, and aging effects revealed by economic preference. *PLoS ONE, 8*(7): e68210. doi:10.1371/journal. pone.0068210.

Williams, B. A. (1988). Reinforcement, choice, and response strengthle. In L. R. . Atkinson R.C, Herrnstein R.J, Lindzey G (Ed.), *Stevens' handbook of experimental psychology* (2nd ed., pp. 167–244). New York: Wiley.

Wimperis, B. R., & Farr, J. L. (1979). The effects of task content and reward contingency upon task performance and satisfaction. *Journal of Applied Social Psychology, 9*(3), 229-249.

Wright, R. A., & Kirby, L. D. (2001). Effort determination of cardiovascular response: An integrative analysis with applications in social psychology. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 33, pp. 255-307). New York: Academic Press.

Zakrzewska, M. Z., & Brzezicka, A. (2014). Working memory capacity as a moderator of load-related frontal midline theta variability in Sternberg task. *Frontiers in Human Neuroscience, 8*. doi:10.3389/fnhum.2014.00399.

Ziegelmeyer, A., Schmelz, K., & Ploner, M. (2011). Hidden costs of control: four repetitions and an extension. *Experimental Economics, 96*(5), 1–18.

APPENDICES

LIST OF APPENDICES

Appendix A: Experimental instructions (Paper 1)

A.1: Example of the bonus/fair treatment

A sequence of letters will appear on the screen, one letter at the time until the whole sequence is shown (e.g. the letter A will appear, than the letter C, D and H until the whole sequence A C D H appears on the screen). Than the sequence of letters will disappear and you will be shown one of the letters from the sequence (e.g. the letter C).
Your task is to press the number that corresponds to the position of the letter in the sequence (if the letter shown is C, your task is to press number 2, because letter C was in the second place in the sequence A C D H). All letters in the sequence will appear in random order. In the following trials the sequence length will be prolonging until you incorrectly solve two of the three trials of a particular sequence length.
If you press the number that corresponds to the position of the letter in the sequence, a green square will appear on the screen as a feedback and if you press an incorrect number, a red square will appear.

Probe trail (2 sequences with 3 letters)
If you understood the task press ENTER.

The calibration phase
(After the calibration phase participants got feedback on their performance and payment.)
Your maximum achieved sequence length is ___.
You earned 2 experimental units (EU).

The task: The choice between an easy and a difficult task
In this phase you decide for an easy or a difficult task. The task is the same as the one in the calibration phase of the experiment. The easy task means that the sequence of letters is shorter than the number of letters you achieved in the calibration phase and the difficult task means that the sequence of letters is longer of what you achieved in the calibration phase. The easy task yields a smaller reward if correctly solved and the difficult task yields a larger reward if correctly solved.

Choice 1:
Choose between:

> An easy task that brings you 1 EU if correctly solved and 0 EU if incorrectly solved.
> A difficult task that brings you 2 EU if correctly solved and 0 EU if incorrectly solved.

The easy task means that the sequence length is two letters shorter than your maximum achieved sequence length in the calibration phase.

The difficult task means that the sequence length is as long as your maximum achieved sequence length in the calibration phase.

Your peers get equal outcome for equal performance.

You correctly incorrectly/solved the task. You earned ___ EU.

Choice 2:
Choose between:

An easy task that brings you 1 EU if correctly solved and 0 EU if incorrectly solved.

A difficult task that brings you 3 EU if correctly solved and 0 EU if incorrectly solved.

The easy task means that the sequence length is two letters shorter than your maximum achieved sequence length in the calibration phase.
The difficult task means that the sequence length is one letter longer than your maximum achieved sequence length in the first phase.

Your peers get equal outcome for equal performance.

You correctly incorrectly/solved the task. You earned ___ EU.

Choice 3:
Choose between:

An easy task that brings you 1 EU if correctly solved and 0 EU if incorrectly solved.

A difficult task that brings you 4 EU if correctly solved and 0 EU if incorrectly solved.

The easy task means that the sequence length is two letters shorter than your maximum achieved sequence length in the calibration phase.
The difficult task means that the sequence length is two letters longer than your maximum achieved sequence length in the first phase.

Your peers get equal outcome for equal performance.
You correctly incorrectly/solved the task. You earned ___ EU.

A.2: Example of the penalty/unfair treatment

A sequence of letters will appear on the screen, one letter at the time until the whole sequence is shown (e.g. the letter A will appear, than the letter C, D and H until the whole sequence A C D H appears on the screen). Than the sequence of letters will disappear and you will be shown one of the letters from the sequence (e.g. the letter C).

Your task is to press the number that corresponds to the position of the letter in the sequence (if the letter shown is C, your task is to press number 2, because letter C was in the second place in the sequence A C D H). All letters in the sequence will appear in random order. In the following trials the sequence length will be prolonging until you incorrectly solve two of the three trials of a particular sequence length.

If you press the number that corresponds to the position of the letter in the sequence, a green square will appear on the screen as a feedback and if you press an incorrect number, a red square will appear.

Probe trail (2 sequences with 3 letters)

If you understood the task press ENTER.

The calibration phase

(After the calibration phase participants got feedback on their performance and payment.)

Your maximum achieved sequence length is ___.

You earned 8 experimental units (EU).

The experiment: The choice between an easy and a difficult task

In this phase you decide for an easy or a difficult task. The task is the same as the one in the calibration phase of the experiment. The easy task means that the sequence of letters is shorter than the number of letters you achieved in the calibration phase and the difficult task means that the sequence of letters is longer of what you achieved in the calibration phase. The easy task yields a smaller reward if correctly solved and the difficult task yields a larger reward if correctly solved.

Choice 1:

Choose between:

An easy task in which you lose 1 EU if correctly solved and 2 EU if incorrectly solved.

A difficult task in which you lose 0 EU if correctly solved and 2 EU if incorrectly solved.

The easy task means that the sequence length is two letters shorter than your maximum achieved sequence length in the calibration phase.

The difficult task means that the sequence length is as long as your maximum achieved sequence length in the calibration phase.

Your peers get higher outcome for equal performance.
You correctly incorrectly/solved the task. You lost ___ EU.

Choice 2:
Choose between:

An easy task in which you lose 1 EU if correctly solved and 2 EU if incorrectly solved.

A difficult task in which you earn 1 EU if correctly solved and lose 2 EU if incorrectly solved.

The easy task means that the sequence length is two letters shorter than your maximum achieved sequence length in the calibration phase.
The difficult task means that the sequence length is one letter longer than your maximum achieved sequence length in the first phase.

Your peers get higher outcome for equal performance.
You correctly incorrectly/solved the task. You earned (or lost) ___ EU.

Choice 3:
Choose between:

An easy task in which you lose 1 EU if correctly solved and lose 2 EU if incorrectly solved.

A difficult task in which you earn 2 EU if correctly solved and lose 2 EU if incorrectly solved.

The easy task means that the sequence length is two letters shorter than your maximum achieved sequence length in the calibration phase.

The difficult task means that the sequence length is two letters longer than your maximum achieved sequence length in the first phase.

Your peers get higher outcome for equal performance.
You correctly incorrectly/solved the task. You earned (or lost) ___ EU.

Appendix B: Experimental instructions and design for Paper 2 and Paper 3 (an example of the reward treatment)[16]

The calibration phase

A sequence of letters will appear on the screen, one letter at a time until the whole sequence is shown (eg. the letter A will appear, then the letters C, D and H until the whole sequence A C D H appears on the screen). Then the sequence of letters will disappear and a number will appear on the screen (e.g. number 2).

Your task is to press the letter that corresponds to the position (the number that appeared) of the letter in the sequence (if the number shown is 2, your task is to press the letter C because that letter was in second place in the sequence A C D H). All letters in the sequence will appear in random order. In the following trials the sequence length will be extended until you incorrectly solve two of the three trials of a particular sequence length.

If you press the letter that corresponds to the position of the number that appeared on the screen, a green square will appear on the screen by way of feedback, and if you press an incorrect number a red square will appear.

(After the calibration phase the participants were given feedback on their performance)

Your maximum achieved sequence length is ___.

The experiment

The choice between an easy and a difficult task in a non-incentivised setting

In this phase, you decide for an easy or a difficult task. The task is the same as the one in the first phase of the experiment. The easy task means that the sequence of letters is shorter and the difficult task means that the sequence of letters is the same as or longer than your achievement in the first phase.

Choice 1:

Choose between:

- The easy task
- The difficult task

The easy task means that the sequence length is two letters shorter than your maximum sequence length achieved in the first phase.

The difficult task means that the sequence length is as long as your maximusequence length achieved in the first phase.

[16] Experimental instruction for reward treatment in Paper 2 and experimental instruction for incentivized treatement for Paper 3 are same.

Choice 2:
Choose between:

- The easy task
- The difficult task

The easy task means that the sequence length is two letters shorter than your maximum sequence length achieved in the first phase.
The difficult task means that the sequence length is one letter longer than your maximum sequence length achieved in the first phase.

Choice 3:
Choose between:

- The easy task
- The difficult task

The easy task means that the sequence length is two letters shorter than your maximum sequence length achieved in the first phase.
The difficult task means that the sequence length is two letters longer than your maximum sequence length achieved in the first phase.

After each choice, feedback indicating (in)correctness appears on the screen.

The choice between an easy and a difficult task in an incentivised setting (the reward treatment)
In this phase, you decide for an easy or a difficult task. The task is the same as the one in the first phase of the experiment. The easy task means that the sequence of letters is shorter and the difficult task means that the sequence of letters is the same as or longer than your achievement in the first phase. The easy task yields a smaller reward if correctly solved and the difficult task yields a larger reward if correctly solved.
You receive EU 36 of initial endowment.

Choice 1:
Choose between:

The easy task that brings you EU 4 if correctly solved and EU 0 if incorrectly solved.
The difficult task that brings you EU 8 if correctly solved and EU 0 if incorrectly solved.

The easy task means that the sequence length is two letters shorter than your maximum sequence length achieved in the first phase.

The difficult task means that the sequence length is as long as your maximum sequence length achieved in the first phase.

Choice 2:
Choose between:

> The easy task that brings you EU 4 if correctly solved and EU 0 if incorrectly solved.
> The difficult task that brings you EU 12 if correctly solved and EU 0 if incorrectly solved.

The easy task means that the sequence length is two letters shorter than your maximum sequence length achieved in the first phase.
The difficult task means that the sequence length is one letter longer than your maximum sequence length achieved in the first phase.

Choice 3:
Choose between:

> The easy task that brings you EU 4 if correctly solved and EU 0 if incorrectly solved.
> The difficult task that brings you EU 16 if correctly solved and EU 0 if incorrectly solved.

The easy task means that the sequence length is two letters shorter than your maximum sequence length achieved in the first phase.
The difficult task means that the sequence length is two letters longer than your maximum sequence length achieved in the first phase.

**After each choice, feedback indicating (in)correctness and about the reward achieved appears on the screen:*

You correctly (incorrectly) solved the task. You earned EU ___.

Appendix C: Measurement of achievement motivation

	Motivation to achieve (MTA)
1.	I like situations in which I can find out how capable I am.
2.	When I am confronted with a problem which I can possibly solve, I am excited to start working on it immediately.
3.	I enjoy situations in which I can make use of my abilities.
4.	I am attracted by situations that allow me to test my abilities.
5.	I am attracted by tasks in which I can test my abilities.
	Motivation to avoid failure (MTF)
1.	I am afraid of failing in somewhat difficult situations, when a lot depends on me.
2.	I feel uneasy doing something if I am unsure of succeeding.
3.	Even if nobody were to notice my failure, I'm afraid of tasks which I'm not able to solve.
4.	Even if nobody is watching, I feel quite anxious in new situations.
5.	If I do not understand a problem immediately, I start feeling anxious.

Note: The 10 items of the revised AMS version by Lang and Fries (2006) were rated on 4-point scales ranging from strongly disagree (1) to strongly agree (4). Correlation between the two variables is negative: $r = -0.436$, $p = 0.000$. Cronbach's alpha for MTA is 0.82 and for MTF 0.80.

Appendix D.:Measurment of Behavioural Approach and Behavioural Inhibition Scales

Behavioural approach scale (BAS) (13 items)
1. I often act on the spur of the moment. 2. When I want something I usually go all-out to get it. 3. I go out of my way to get things I want. 4. I will often do things for no other reason than that they might be fun 5. If I see a chance to get something I want I move on it right away. 6. When good things happen to me, it affects me strongly. 7. I crave excitement and new sensations. 8. It would excite me to win a contest. 9. When I go after something I use a "no holds barred" approach. 10. When I see an opportunity for something I like I get excited right away. 11. I'm always willing to try something new if I think it will be fun. 12. When I'm doing well at something I love to keep at it. 13. When I get something I want, I feel excited and energized.
Behavioural inhibition scale (BIS) (7 items)
1. Even if something bad is about to happen to me, I rarely experience fear or nervousness. (*Reversed scored*) 2. I have very few fears compared to my friends. (*Reversed scored*) 3. If I think something unpleasant is going to happen I usually get pretty "worked up." 4. I feel worried when I think I have done poorly at something important. 5. I worry about making mistakes. 6. I feel pretty worried or upset when I think or know somebody is angry at me. 7. Criticism or scolding hurts me quite a bit.

Note: The 20 items of the BIS/BAS Scales by Carver and White (1994) were rated on 4-point scales ranging from strongly disagree (1) to strongly agree (4). Correlation between the two variables is negative: $r = -0.020$, $p = 0.810$. Cronbach's alpha for BAS is 0.742 and for BIS 0.732.

Appendix E. Consent to participate in research (English translation)

University of Ljubljana
Faculty of Economics

Research Title: The role of cognitive frames in combined decisions about risk and effort

1. You are invited to participate in a research experimental study on the role of cognitive frames in combined decisions about risk and effort. The study is conducted in the context of a broader research project analysing the effects of incentive schemes on decision making conducted at the Faculty of Economics, University of Ljubljana under the supervision of Assoc. prof. dr. Sergeja Slapničar.
2. The purpose of the research is to determine how different remuneration schemes affect the combined decision about risk and effort.
3. You will be asked to solve simple cognitive tasks and make three decisions.
4. The study will last approximately 20 minutes. For the participation in the research you will receive financial compensation.
5. There are no anticipated risks to your participation.
6. The decision to participate in this study is entirely up to you. You may refuse to take part in the study *at any time* without affecting your relationship with the investigators of this study or University of Ljubljana. Your decision will not result in any loss or benefits to which you are otherwise entitled. You have the right not to answer any single question, as well as to withdraw completely.
7. This study is anonymous. We will do everything to protect your privacy. The records of this study will be kept strictly confidential. The information collected about you will be coded using numbers (ID).

Consent

Your signature below indicates that you have decided to volunteer as a research participant for this study, that the records obtained will be used in scientific purposes, and that you have read and understood the information provided above.

Date: ________________________________

Subject's Name (print): ________________________________

Subject's Signature: ________________________________

Appendix F: Summary in Slovenian language / Daljši povzetek disertacije v slovenskem jeziku

Pomen shem nagrajevanja za oblikovanje referenčnih vrednosti in vpliv na odločitve o tveganjih in naporu

Namen doktorske disertacije je raziskati vpliv shem nagrajevanja na odločitve o kognitivnem naporu v več obdobjih, pri čemer se raven zahtevanega kognitivnega napora povečuje. Raziskava vključuje teorije s področja ekonomije, poslovodnega računovodstva, socialne psihologije in uvide s področja nevroznanosti, ki omogočajo vpogled v kompleksen proces sprejemanja odločitev o kognitivnem naporu. Učinek finančnih spodbud na človekovo vedenje je eno temeljnih vprašanj, ki se poraja v ekonomiji (npr. Bénabou in Tirole, 2003; Gibbons, 1998; Kreps, 1997; Lazear, 1986, 2000; Prendergast, 1999), psihologiji in poslovodenju (npr. Gerhart in Rynes, 2003; Gomez-Mejia in Welbourne, 1988; Rynes et al., 2005; Vroom, 1964), saj je ekonomija kljub vsem formulam, modelom in teorijam namenjena raziskovanju človeškega vedenja.

Doktorska disertacija je sestavljena iz treh poglavij, ki proučujejo vpliv shem nagrajevanja na kognitivni napor s treh vidikov, s skupnim ciljem ugotoviti, ali sheme nagrajevanja s kognitivnim okvirom, ki ga ustvarijo, vplivajo na odločanje o kognitivnem naporu in kako na učinek le-teh vplivajo osebnostne značilnosti, težavnost naloge in predhodna uspešnost.

V prvem poglavju proučujemo, kako kognitivni okviri, ustvarjeni z oblikovanjem shem nagrajevanja in pravičnostjo le-teh, vplivajo na odločitve o tveganju in naporu. V vsakdanjem življenju so odločitve o tveganju in naporu medsebojno povezane v teoriji, pa ni tako, saj so teorije, ki preučujejo sprejemanje tveganj in motivacijo za napor, nepovezane. Pomembne psihološke teorije poudarjajo, da je odločanje odvisno od posameznikovih kognitivnih okvirov ali mentalnih predstavitev problema odločanja (Birnberg in drugi, 2007). Zasnova shem nagrajevanja ima pomemben vpliv na kognitivne okvire, ki vplivajo na posameznikovo zaznavanje pravičnosti, raven prizadevanja in zaznavanje rezultata kot nagrado ali izgubo. Dve vodilni psihološki teoriji – teorija organizacijske pravičnosti (Adams, 1963, 1966) in teorija obetov (Kahneman in Tversky, 1979; Kahneman, 2003) – predlagata, da se kognitivni okviri pojavijo s primerjavo izidov z referenčno točko. Pri teoriji organizacijske pravičnosti referenčna točka predstavlja primerjavo z ustrezno drugo točko, medtem ko v teoriji obetov referenčna točka v bistvu predstavlja nespremenjeno stanje *(status quo)* (Kahneman, 2003) in se lahko sproži z različnimi značilnostmi sistema spodbud. Ideja, ki je skupna obema teorijama, je, da referenčne točke oblikujejo kognitivne okvire in da odstopanje od njih povzroča notranje konflikte, ki se jim posamezniki poskušajo izogniti (Birnberg in drugi, 2007). V bolj zapletenih situacijah odločanja se posamezniki hkrati srečujejo s številnimi kognitivnimi okviri, pri čemer se postavlja vprašanje, kateri igra osrednjo vlogo pri odločanju in kako med seboj vzajemno delujejo. Čeprav si obe teoriji delita globoko povezani koncept, je zanimivo, da ostajata ločeni, saj teorija organizacijske pravičnosti uporablja referenčne točke za odločitve o naporu brez izrecnega upoštevanja tveganj izida, teorija obetov pa za napovedovanje vedenja pri sprejemanju tveganj. Vendar se v praksi odločanje o tveganju in naporu prepletata; posamezniki se spoprijemajo z možnostjo, ki zahteva veliko napora in ki lahko prinese veliko plačilo, vendar je verjetnost, da bi to plačilo dosegli, odvisna od uspeha pri dokončanju naloge. Druga možnost pa je izbira lahke naloge, ki zahteva malo napora in vključuje veliko verjetnost za uspeh, a posledično ponuja majhno plačilo.

V raziskavi želimo z uporabo obeh teorij ugotoviti, kateri kognitivni okviri prevladujejo pri sočasnih odločitvah o tveganju in naporu. Spoznanja na področju nevroznanosti, psihologije (Hughes in drugi, 2015; Salamone in drugi, 1994; Treadway in drugi, 2015; Walton in drugi, 2006; Wardle in drugi, 2012) in vedenja živali (Cocker in drugi, 2012; Hosking in drugi, 2014 a, b) ugotavljajo, da se pri odločanju o tveganju in naporu nevronske mreže, ki se aktivirajo pri obeh vrstah odločitev, prekrivajo (Salamone in drugi, 2012; Miller in drugi, 2013), kar krepi domnevo, da so odločitve o tveganju in naporu povezane.

V raziskavi smo uporabili eksperimentalno metodo, s katero smo v treh obdobjih merili odločitev o tveganju in naporu. Eksperimentalno smo manipulirali s shemo nagrajevanja (shema kaznovanja ali shema nagrajevanja) in pravičnostjo sheme nagrajevanja (pravična ali nepravična porazdelitev nagrad). S tem smo ustvarili dva okvira za oblikovanje referenčne točke posameznika. Ker je osnovno/fiksno plačilo običajno upoštevano kot referenčna točka, shema nagrajevanja v obliki nagrade (*bonus*) ustvarja dojemanje pridobitve, shema nagrajevanja v obliki kazni (*penalty*) pa ustvarja dojemanje izgube. Podobno pravična in nepravična porazdelitev nagrad/kazni glede na primerljivega posameznika ustvarja dojemanje pridobitve ali izgube. V poglavju ugotavljamo, kako sta kognitivna okvira, ustvarjena s shemo nagrajevanja in pravičnostjo, povezana; ali je njun vpliv vzajemen, nasproten, kumulativen ali morda en okvir prevlada nad drugim. V eksperimentu smo uporabili prilagojeno Sternbergovo nalogo (Sternberg, 1966). Sternbergova naloga se v psihologiji uporablja za merjenje kognitivnega napora (Burrows in Okada, 1973; D'Esposito in drugi, 2000; Jansma in drugi, 2007; Zakrzewska in Brzezicka, 2014). Udeleženci se morajo trikrat odločiti med lahko in zato manj tvegano nalogo, ki prinaša majhno nagrado/majhno kazen in težko in zato bolj tvegano nalogo, ki prinaša večjo nagrado, če je rešena pravilno. Tveganje je povezano s pravilnostjo naloge in s tem povezano nagrado/kaznijo. Težja ko je naloga, večja je nagrada oziroma manjša je kazen in večje je tveganje, da naloga ne bo pravilno rešena. Tretji kognitivni okvir pa ustvarja časovna komponenta, saj pozitivni ali negativni predhodni rezultati drugače vplivajo na trenutno odločitev (Thaler in Johnson, 1990).

Ugotavljamo, da je pogostost odločitev za večje tveganje in napor najnižja pri shemah nagrajevanja s pravičnim bonusom in najvišja pri pravičnih shemah nagrajevanja s kaznijo. Če je plačilo za enak napor in uspešnost nepravično, je manj pomembno, ali je shema nagrajevanja oblikovana kot bonus ali kazen. Če pa je plačilo pravično, večji napor in tveganje spodbudimo s shemo kaznovanja. Zdi se, da pravična shema nagrajevanja z bonusom predstavlja območje ugodja, ki najmanj pogosto sproži odločitve za visoko vloženo tveganje in napor. Ugotavljamo, da v drugem in tretjem krogu postane najpomembnejši dejavnik odločitve predhodna uspešnost in okviri, ustvarjeni s shemo nagrajevanja, niso več pomembni za posameznikovo odločitev. Ta učinek nakazuje, da je ocena verjetnosti o tem, da nekdo uspešno lahko opravi nalogo, ki temelji na predhodni uspešnosti in predhodnih odločitvah, postala pomembnejša od sheme nagrajevanja ali pravičnosti rezultatov in sama po sebi tvori referenčno točko.

Prvi prispevek raziskave je proučevanje hkratnih odločitev o tveganju in naporu, saj gre za zelo nepopolno raziskano področje v literaturi (Sprinkle in drugi, 2008). Brez upoštevanja takih odločitev je nemogoče popolnoma razumeti učinkovitost shem nagrajevanja. Večina študij na področju poslovodnega računovodstva temelji na predpostavki teorije pričakovanj o razmerju med tveganjem in naporom, pri čemer večja ko je verjetnost, da bodo napori prispevali k večji uspešnosti, bolj bo oseba motivirana za vlaganje naporov (Vroom, 1964). Drugi prispevek te znanstvene raziskave je analiziranje, kako posamezniki upoštevajo več kot en kognitivni okvir hkrati. Naše ugotovitve kažejo, da sheme z bonusi in kaznimi ustvarjajo kognitivne okvire v skladu s teorijo obetov, kar dodatno dokazuje, kako različne prakse nagrajevanja oblikujejo kognitivne okvire. Pokazali smo, da pri interakcijah več okvirov slednji stimulirajo vedenje, ki je različno od tistega, ki ga povzroči le en kognitivni okvir. Tretjič, s preučevanjem odločanja v okolju v več obdobjih smo pokazali, da učinek shem nagrajevanja sčasoma zbledi; ko se pojavi nova pomembna informacija (tj. predhodna uspešnost), ki pomaga ponovno oceniti verjetnost izida. Četrtič, naše ugotovitve vsebujejo praktične implikacije za oblikovanje učinkovitih shem nagrajevanja. Ugotovljeno je bilo, da shema kaznovanja spodbuja odločitve za velik napor in tveganja. Ker sheme kaznovanja postajajo vse bolj priljubljene, naše ugotovitve kažejo, da jih je treba izvajati z jasnim zavedanjem o njihovih učinkih. Raziskava združuje dve vplivni psihološki teoriji z literaturo in prakso o poslovodnem računovodstvu.

V drugem poglavju raziskujemo, kako se posamezniki z različnimi stopnjami motivacije, usmerjene k uspehu (*Motivation to achieve – MTA*), in motivacije, usmerjene k izogibanju neuspeha (*Motivation to avoid failure – MTF*), odzivajo na sheme nagrajevanja, oblikovane kot kazni, in nagrade pri odločanju o kognitivnem naporu. Prav tako ocenjujemo, kako na to vrsto odločanja vpliva zahtevnost naloge, saj je uspešnost pri višji težavnosti manj verjetna in je s tem povezano večje tveganje. Ker je odziv posameznikov na negotovost odvisen od kognitivnega okvira, ki je ustvarjen s shemo nagrajevanja, ti dve vrsti spodbud preučujemo ločeno.

Uspeh sodobnih organizacij je močno odvisen od motivacije njihovih zaposlenih, da se vključijo v kognitivno zahtevne naloge, tj. razvijanje novih poslovnih modelov, tehnologij, izdelkov ali storitev, ki so lahko veliko bolj negotove od običajnih nalog. Čeprav so sheme nagrajevanja pomembna determinanta za izbiro kognitivnih nalog, so rezultati prejšnjih študij o kognitivnem naporu, izzvanem s shemami nagrajevanja, izjemno nedosledni.

Bonner, Hastie, Sprinkle in Young (2000) v svojem preglednem delu poročajo, da pri približno 50 % eksperimentov denarne spodbude niso vplivale na uspešnost pri opravljanju kognitivnih nalog. Kot poudarjajo avtorji, denarne spodbude lahko postanejo manj učinkovite, če obstaja vedno večja vrzel med težavnostno stopnjo naloge in sposobnostmi, potrebnimi za njeno uspešno dokončanje. Kako se posamezniki odzovejo na shemo nagrajevanja, je odvisno tudi od motivacije, usmerjene k uspehu. Nekateri posamezniki so močno motivirani za samo kognitivno nalogo, ker so prirojeni uspešneži, ki stremijo k doseganju rezultatov, medtem ko drugi delujejo tako, da bi se izognili neuspehu.

Da bi odgovorili na naša raziskovalna vprašanja, smo opravili eksperiment, pri katerem smo merili pripravljenost za opravljanje zahtevne kognitivne naloge, tj. prilagojene Sternbergove naloge (Sternberg, 1966). Udeleženci so se morali pri pogoju, kjer smo manipulirali s shemo nagrajevanja (oblikovana kot nagrada ali kazen), in nevtralnem pogoju odločiti med težko in lahko kognitivno nalogo. V vsakem od pogojev so opravili tri odločitve, pri čemer se je težavnost težke naloge iz kroga v krog povečevala, težavnost lahke naloge pa je bila konstanta.

Uporabili smo paradigmo kognitivnega napora (Botvinick in drugi, 2009; Kool in drugi, 2010; Westbrook in drugi, 2013), pri kateri so se udeleženci odločali med zahtevno nalogo za večjo denarno nagrado ali nezahtevno nalogo za manjšo denarno nagrado v ekonomsko enakovrednih spodbudnih okoljih. Motivacijo, usmerjeno k uspehu, smo merili z revidirano lestvico motivacije, usmerjene k uspehu (*Achievement motivation scale*) (Lang in Fries, 2006).

Naši opisni rezultati izražajo to, da imajo pri vedno zahtevnejših kognitivnih nalogah denarne spodbude manjši učinek na izbiro zahtevne naloge. Medtem ko motivacija, usmerjena k uspehu, v nevtralnem pogoju jasno razloži izbiro zahtevnih nalog, je razmerje v okolju s spodbudami manj jasno. Pri pogoju brez sheme nagrajevanja ugotavljamo, da pri relativno nizki težavnosti nalog izbire težavne naloge ni mogoče razložiti z motivacijo, usmerjeno k uspehu, saj večina ljudi izbere zahtevno nalogo. Ko se težavnost naloge poveča na zmerno raven, pa opažamo pozitiven učinek motivacije, usmerjene k uspehu (MTA). Učinek MTA preneha biti pomemben pri visoki stopnji težavnosti. Na tej ravni težavnosti samo motivacija, usmerjena k izogibanju neuspeha (MTF), pojasni izbire, kar sovpada s predhodnimi raziskavami (Gendolla in Wright, 2005; Wright in Kirby, 2001; Heckhausen, 1963, 1991; McClelland in drugi, 1953). Za zahtevne naloge se odločijo samo tisti posamezniki z visoko motivacijo MTF, medtem ko se posamezniki z visoko motivacijo MTF najmanj verjetno odločijo za opravljanje težkih nalog.

Pri pogoju s shemo nagrajevanja ugotavljamo, da nagrade niso pozitivno vplivale na izbiro težke naloge pri nobeni stopnji težavnosti. V nasprotju z našimi pričakovanji tudi niso imele pozitivnega učinka pri nalogah z nizko stopnjo težavnosti. S povečevanjem težavnosti je negativni učinek še večji, kar je mogoče pripisati večji negotovosti glede uspeha pri nalogi, saj se ta s težavnostjo naloge povečuje. Pri povečanju težavnosti naloge se negotovost nagrad primerja s kognitivnim naporom, ki ga je treba vložiti za doseganje rezultatov, pri čemer ljudje nočejo sprejemati tveganj (Kahneman in Tversky, 1979). Naše ugotovitve delno podpirajo teorijo kognitivne evalvacije (Deci in drugi, 1999), v skladu s katero zunanje nagrade zmanjšujejo notranjo motivacijo. Prav tako potrjujejo nevroznanstvene dokaze, da se prednostna izbira lažje kognitivne naloge ne spremeni, četudi je za zahtevnejšo nalogo na voljo večja nagrada (Massar in drugi, 2015; Westbrook in drugi, 2013). To velja za zelo zahtevne naloge in za posameznike z visoko stopnjo MTA, medtem ko to ni razvidno pri nižjih stopnjah težavnosti.

Pri pogoju sheme nagrajevanja, ki je oblikovana kot kazen, opažamo bistveno višjo stopnjo angažiranosti za težke naloge samo pri stopnji, pri kateri je uspeh razumno dosegljiv (lahka in srednje težka naloga). Izbira težavne naloge v prvem krogu je prevladujoča, saj ponuja možnost izognitve določeni kazni. Medtem ko v okolju brez spodbud MTF ne kaže nobenega učinka, ima nasprotno interakcija med kaznijo in MTF v spodbudnem okolju znaten negativen predznak, kar nakazuje na to, da so se posamezniki z visoko stopnjo MTF redkeje odločili za težavno nalogo kot posamezniki z nizko stopnjo MTF. To pomeni, da so posamezniki z visoko stopnjo MTF manj občutljivi za denarne kazni kot posamezniki z nizko stopnjo MTF in se ne lotijo težavnejših nalog, četudi na račun določene manjše izgube. Rezultat je opazen tudi v drugem krogu pri nalogah z zmerno stopnjo težavnosti. V tem krogu prav tako poročamo o znatni interakciji med MTA in kaznijo. Negativni interaktivni učinek je tako močan, da kazni v resnici nimajo veliko učinka niti za posameznike z visoko stopnjo MTA niti za posameznike z visoko stopnjo MTF. Pri najvišji stopnji težavnosti, pri kateri je večina udeležencev ocenila, da se z izbiro težavne naloge ne morejo izogniti izgubi, se jih je večina odločila za lažjo nalogo. Kazen ni imela učinka niti samostojno niti v interakciji z MTF. Tako kot v okolju brez spodbud je ponovno pomemben dejavnik MTF. V splošnem so naši rezultati potrdili domnevo, da interakcija med vrsto spodbude, zahtevnostjo naloge in motivacijo, usmerjeno k uspehu, skupaj vpliva na pripravljenost za izbiro kognitivno zahtevne naloge.

Najpomembnejši prispevek naše študije je prispevek k dokazanemu vplivu sheme nagrajevanja na motivacijo, usmerjeno k uspehu. Enake spodbude različno vplivajo na ljudi z drugačno motivacijo, usmerjeno k uspehu, in motivacijo, usmerjeno k izogibanju neuspeha. Nagrade in kazni ustvarjajo različne kognitivne okvire, v katerih posamezniki drugače ocenijo izide tveganj in ustrezno izberejo svoje vedenje glede na tveganje. Te težnje temeljijo na motivacijskih nagnjenjih posameznikov, ki določajo, kako se ti odzivajo na morebitne nagrade in kazni. Nadalje v enem eksperimentu analiziramo kompleksen medsebojni vpliv spodbud in motivacije, usmerjene k uspehu, pri kognitivno zahtevnih nalogah. Čeprav je eksperiment kompleksen, je samo taka analiza podobna resničnim življenjskim situacijam, v katerih se sprejemajo take odločitve. Bivariatne analize teh razmerij so privedle do protislovnih rezultatov. Ugotovitve o učinkih spodbud na kognitivno odločanje so polemične in ne ponujajo nedvoumnega odgovora na to, ali je z denarnimi spodbudami mogoče spodbuditi kognicijo (npr. boljši kratkoročni spomin, boljše razumevanje, boljše kognitivno načrtovanje, prilagodljivo prilagajanje vedenja na prihajajoče znake itd.). Predhodne raziskave so redko analizirale učinke shem nagrajevanja za opravljanje kognitivnih nalog, katerih težavnost se povečuje. Na vsaki stopnji težavnosti nalog odkrijemo drugačen učinek shem nagrajevanja. Študije, ki ne upoštevajo različnih stopenj težavnosti nalog in učinek preprosto merijo na eni točki, ne da bi ga umerile glede na vsakega posameznika, odnos med shemo nagrajevanja in prizadevanjem opazujejo le v omejenem intervalu. Z naraščajočo težavnostjo nalog se verjetnost za uspeh manjša, s tem pa pri odločitvah prihaja do novega vidika odnosa med tveganjem in vedenjem. Prepričani smo, da z znanstveno razpravo prispevamo k interdisciplinarnemu pogledu, ki spodbuja napredovanje našega razumevanja delovanja shem nagrajevanja (Merchant, Van der Stede in Zheng, 2003). Praktični prispevek raziskave je v priporočilih o tem, kako naj torej organizacije spodbujajo sprejemanje kognitivno zahtevnih nalog. Pogodba, ki bi lahko pritegnila visoko motivirane kandidate, mora zagotavljati konkurenčno fiksno plačilo in stimulativno okolje.

V tretjem poglavju obravnavamo razlike v učinku denarnih spodbud glede na spol pri zahtevnih kognitivnih nalogah. Zanima nas, kako nagrade motivirajo moške in kako ženske pri različnih težavnostnih stopnjah. Pregled literature kaže, da moški in ženske drugače sprejemajo odločitve o tveganjih. Ženske so previdnejše in drugače občutijo tveganje (Croson in Gneezy, 2009), načeloma manj uživajo v izzivih, so manj tekmovalne (Niederle in Vesterlund, 2007; Datta, Gupta, Poulsen in Villeval, 2013) in so bolj negotove (Carson in Gneezy, 2009). Vendar pa pri odločanju za opravljanje lahke ali zahtevne kognitivne naloge ni pomembno samo tveganje, ampak tudi kognitivni napor. Pomembno, a premalo raziskano področje procesa odločanja je vloga osebnostnih profilov pri odzivih na spodbude glede na različno zahtevnost naloge. Vedenjski sistem umika (BIS) in vedenjski sistem približevanja (BAS) sta dva nevropsihološka sistema, ki usmerjata posameznikovo vedenje (Gray, 1981). Posameznikova nagnjenja BIS in BAS lahko vodita proces odločanja, saj sta občutljiva za nagrajevalne oziroma kaznovalne dražljaje. Pri preučevanju razlik med spoloma glede občutljivosti za spodbudo je pomembno upoštevati osnovne razlike med spoloma v motivih za motivacijo k uspehu in za izognitev neuspehu, o čemer pa imamo le omejene dokaze (De Pater et al., 2007).

Za analizo raziskovalnega vprašanja razlik med spoloma, nagnjenja BIS/BAS in vpliv shem nagrajevanja na kognitivni napor smo zasnovali eksperiment, ki je potekal pri dveh pogojih in kjer je vsak od pogojev vseboval tri obdobja. Opazovali smo izbiro kognitivnih nalog, ki odražajo preference za tveganje in napor. Udeleženci so izbirali med nalogo z nizkim tveganjem in naporom (lahka naloga) in nalogo z visokim tveganjem in naporom (težka naloga). Težka naloga ponuja večji pričakovani donos kot lahka naloga, vendar zahteva več napora. Z elementom tveganja manipuliramo s povečevanjem nagrade za vsako povečano zahtevnost naloge (lahka, zmerna in visoka težavnost). Udeleženci so izbire opravili pri pogoju brez spodbud, drugič pa pri pogoju s shemo nagrajevanja. Uporabili smo prilagojeno različico Sternbergove naloge. Z lestvico BIS/BAS smo izmerili nagnjenja BIS in BAS (Carver in White, 1994). Lestvica je sestavljena iz 20 postavk, uporablja pa Likertovo štiritočkovno lestvico.

Opisni rezultati kažejo, da so pri pogoju brez denarnih spodbud pri katerikoli stopnji težavnosti naloge moški izbirali težje naloge kot ženske. Pri pogoju nagrajevanja pa so ženske presegle moške pri izboru v drugem krogu, ko ima težka naloge zmerno stopnjo težavnosti. Pri zelo visoki stopnji težavnosti je bila raven izbire težke naloge najnižja pri ženskah, izbira pri moških pa precej višja kot pri ženskah. Za analizo vpliva sheme nagrajevanja na vsaki od stopenj težavnosti smo uporabili metodo logistične regresije z večkratnimi meritvami. Pri nalogi z nizko težavnostjo je bila pripravljenost za kognitivni napor pri ženskah in moških podobna, nagrada ni imela vpliva. Pri nalogi z zmerno težavnostjo so nagrade dejansko zmanjšale pripravljenost za težko nalogo pri moških, medtem ko se je pripravljenost pri ženskah povečala.

Ta učinek je nepredviden in ga delno lahko pojasnimo s posrednim vplivom sistema BIS, ki je višji pri ženskah. Ko je stopnja težavnosti najvišja, nagrada ne stimulira ne moških ne žensk. Negativni vpliv je zelo močan pri ženskah, ki so načeloma bolj nenaklonjene tveganjem in ki manj zaupajo v svoje sposobnosti. Pri zmerno in visoko zahtevnih nalogah poročamo o pozitivnem vplivu predhodne uspešnosti, kar pomeni, da je ovrednotenje verjetnosti, ali lahko posameznik uspešno opravi zahtevno nalogo, močno odvisno od preteklih izkušenj. Ta ugotovitev je skladna z ugotovitvami v naših drugih dveh raziskavah. Pri visoko zahtevnih nalogah sistem BIS ublaži učinek nagrade in zmanjša kognitivni napor. Tveganje, ki je povezano z izgubo nagrade, je najvišje pri zelo zahtevni nalogi. Posamezniki z visokim BIS se bojijo tvegati. Na tej stopnji ženske izbirajo manj težkih nalog kot moški.

Ugotovitve te raziskave prispevajo k literaturi, ki preučuje razlike med spoloma v različnih kontekstih shem nagrajevanja in razlike med spoloma pri občutljivosti za sheme nagrajevanja. Medtem ko so bile razlike med spoloma pri odločanju o tveganjih v literaturi že temeljito raziskane, ne poznamo obstoječih študij, ki bi raziskovale razlike med spoloma pri izbiri, ki vsebujejo odločitev o tveganju in naporu hkrati in v več obdobjih. Pretekle študije o spolih so zlasti raziskovale izbiro tveganj (Barber in Odean, 2001; Eckel in Grossman, 2008; Croson in Gneezy, 2009; Sila et al., 2016) ali napora (Gneezy, Niederle in Rustichini, 2003; Gill in Prowse, 2010) ali pa so analizirale izbiro nalog z le eno stopnjo težavnosti, ne da bi upoštevale, kakšen vpliv ima lahko spodbuda pri visoki ali nizki verjetnosti. Pri odločitvah v resničnem življenju verjetnosti izida niso znane in so odvisne od posameznikove ocene lastnih zmožnosti in preteklih izkušenj. Predhodno uspešnost lahko upoštevamo, če ocenjujemo več obdobij, v katerih je mogoče nadzorovati, kako pretekle odločitve in izkušnje vplivajo na trenutne izbire. Nazadnje raziskava kaže, da moški, ki pogosteje kot ženske izberejo nalogo z višjo stopnjo težavnosti in ki sprejmejo večje tveganje, niso dosegli splošnega višjega finančnega izida.

Ker so bile ugotovitve pridobljene v eksperimentu, jih je treba upoštevati v luči metodoloških omejitev. Denarne spodbude so bile resda precej nizke, vendar so bile v razponu prejšnjih raziskav, pri katerih je bilo ugotovljeno, da so bile učinkovite.

Rezultatov eksperimenta ne smemo razumeti pri nominalni vrednosti, saj so ti odvisni od specifične zasnove eksperimenta. Če bi spremenili relativno vrednost naših spodbud ali težavnostno stopnjo nalog, bi morda opazili drugačen vzorec odločanja. Kljub nizkim zneskom nagrad verjamemo, da smo eksperiment opravili v skladu s teorijo in najvišjimi raziskovalnimi standardi in da so rezultati logično konsistentni. Raziskava na študentih pri kognitivnih nalogah ne predstavlja omejitve, saj je rezultate mogoče posplošiti na celotno populacijo (Runeson, 2003; Exadaktylos, Espín in Brañas-Garza, 2013).

Da bi povečali veljavnost naših ugotovitev, bi lahko prihodnje raziskave proučile kognitivne naloge, ki zajemajo druga področja kognicije (npr. presojanje, sklepanje, inhibitorni nadzor, načrtovanje itd.). Neke vrste preslikava kognitivnih funkcij bi bila najbolj utemeljena, saj bi zagotovila odgovore na to, kako se ljudje odzivajo na spodbude pri različnih kognitivnih nalogah. Druga zanimiva raziskovalna tema bi bila preučitev medsebojnih vplivov shem nagrajevanja in drugih osebnih lastnosti, kot so na primer vedenjski sistem približevanja in umika (Corr in McNaughton, 2012), izogibanje tveganjem in sposobnosti za kognitivni napor. Na splošno menimo, da vprašanje občutljivosti osebnostnih lastnosti za sheme nagrajevanja še ostaja področje, ki bo navdihovalo prihodnje raziskave.