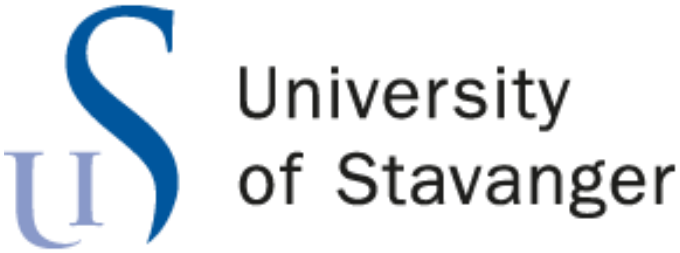


UiS Business School

Master Thesis

THE EFFECT OF PUBLIC RECOGNITION ON PERFORMANCE IN COMPETITIVE ENVIRONMENTS

June 14, 2018

Maria Kristiina Lehtinen

5021

UNIVERSITY OF STAVANGER BUSINESS SCHOOL

MASTER'S THESIS

STUDY PROGRAMME:	THIS THESIS HAS BEEN WRITTEN WITHIN THE FOLLOWING FIELD OF SPECIALISATION:
Business Administration and Economics	Economic Analysis IS THE THESIS CONFIDENTIAL? No

TITLE:

The Effect of Public Recognition on Performance in Competitive Environments

AUTHOR(S)		SUPERVISOR:
		Kristoffer Wigestrand Eriksen
Candidate number:	Name:	
5021 ……	Maria Kristiina Lehtinen ……	

Abstract

This Master thesis investigates the effects of public recognition on performance in competitive environments. Theory suggests three key mechanisms affecting motivation and thus, performance, namely conformity, altruism and reciprocity. The formulated hypotheses are based on previous theory and research and extend it by examining long-term effects. To study the effects of public recognition on performance, a real-effort laboratory experiment was conducted at the University of Stavanger. The results of previous research, that providing public recognition has a positive effect on performance for both rewarded and not rewarded subjects, could not be confirmed. However, it could be confirmed that the effect was stronger for subjects who were not rewarded than for subjects who were rewarded due to public recognition. The positive effect on performance remained for not rewarded subjects when public recognition was provided repeatedly, however the size of the effect decreased. The long-term effect on rewarded subjects could not be confirmed.

Preface

During my studies, I was fascinated by all the underlying concepts and psychological processes in the field of behavioral economics. I specifically liked the research and theory about motivation and therefore, I have chosen to conduct a laboratory experiment and study the effects of public recognition on performance in competitive environments.

I would like to thank my advisor Kristoffer Wigestrand Eriksen for good advice and helpful feedback during the whole process. I would also like to thank my fiancé Alexander Böhrer for helping to program the experiment in z-Tree and always encouraging and supporting me with this thesis. Lastly, I would like to thank the Business School at UiS for financing the experiment and the people who participated in the experiment.

Maria Lehtinen
Stavanger, June 2018

Table of Contents

Table of Figures

Table of Tables

1 Introduction

In today's working environments, it is quite common to reward people for good performance. But does a reward, that is given in front of several observants or even competitors, have a positive influence on future performance?

Employees are the most important success factor to a company. Since recognition is a crucial factor to employees' motivation, companies aim to increase motivation by recognizing employees' efforts and rewarding them, subsequently. Several studies have concluded that employee recognition and symbolic rewards increase performance (e.g. Kosfeld & Neckermann, 2011; Bradler, Dur, Neckermann & Non, 2016; Kube, Maréchal & Puppe, 2012; Stajkovic & Luthans, 2003). However, to my knowledge, previous research lacks the question if effects on performance remain, increase or decrease, when recognition is provided repeatedly. Thus, this paper attempts to confirm and even extend previous research.

Motivation is the most important mechanism which makes employees increase their performance. The underlying key mechanisms affecting motivation are conformity, altruism and reciprocity. People like to be accepted in a group and therefore try to perform and behave similarly with the group. This is called conformity (Bernheim, 1994). People are conditionally altruistic, which means that they rather care and respect an employer who also cares for them (Bradler et al., 2016). Reciprocity, the last mechanism, means that people respond to an action kindly or unkindly, depending on how well they were treated in the first place (Falk & Fischbacher, 2006). Different kinds of rewards may increase or decrease the effect of the key mechanisms presented. This paper focuses on how symbolic rewards and recognition affect motivation and subsequently, performance.

This master thesis investigates the effects of public recognition on performance in competitive environments. To examine these effects, a real-effort experiment is conducted, and the results are presented. In the experiment, the five subjects who perform best are rewarded with a symbolic reward and publicly recognized in front of other subjects, repeatedly. This paper examines the effect of this treatment on both rewarded and not rewarded subjects. In addition, this paper answers the question which of those two groups is more affected by the rewarding procedure. Since this paper aims to extend prior research, public recognition was provided repeatedly to examine the long-term effects on performance.

Both employers and employees may benefit from the results of this thesis. Employers benefit especially financially if the companies' results can be improved by recognizing employees with

no or little amount of money. In addition, recognition and rewards help employees being loyal and staying longer in the company. The employees benefit as well, since better performance may lead to higher salary or a promotion. For many employees it is also important to be recognized, be seen and feel that others see the importance of the work they do. (Fisher, 2015)

The structure of this thesis is as follows. First, relevant theory and previous research on motivation in general, the key motivating mechanisms, as well as employee recognition and rewards are expounded. Second, the methodology for this paper is introduced, including the experimental design, procedures and hypotheses. Third, the results of the experiment are presented and analyzed. Fourth, the results are critically discussed, including possible limitations of the research, and an outlook for further research is given. Last, the conclusion of this paper is presented.

2 Theoretical and empirical background

2.1 Motivation

The key mechanism, that leads to an increase in performance when the right reward is provided, is motivation.

> *To be motivated means to be moved to do something.*
> (Ryan & Deci, 2000, p. 54)

It is important for everyone's well-being at work to be able to satisfy individual needs. People experience comfort or discomfort depending if their needs are met. That is why leaders should be aware of the needs of their employees. Maslow's hierarchy of needs (Maslow, 1943), which is shown in Figure 1, is a good theory base for that. (Mikkelsen & Laudal, 2016)

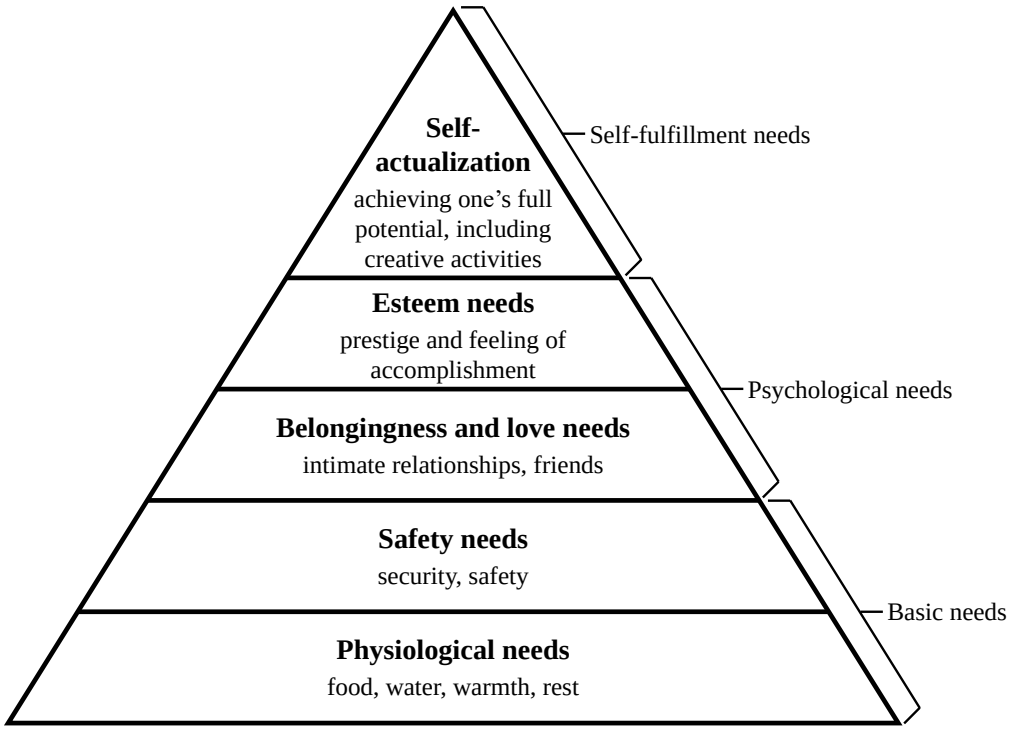


Figure 1: Maslow's hierarchy of needs (Maslow, 1943)

This study focuses on the top three levels of the Maslow's hierarchy of needs. Belongingness, esteem needs, and self-actualization can be improved by recognizing and rewarding employees.

Motivation can be divided into two parts: intrinsic and extrinsic motivation. Intrinsic motivation is "doing something for its own sake" (Reiss, 2012, p. 152). In other words, people do things and take part in activities, because they simply want to do them. At a workplace setting, intrinsic motivation can be triggered with recognition, since it makes the employees feel that they are part of something. On the Maslow's hierarchy of needs this corresponds to the belongingness

needs. By contrast, extrinsic motivation is considered as an inducement from the outside. When the leader recognizes the employees, they feel appreciated because of the work they have done. The employee can respond to the appreciation by working harder and increasing the performance. Hence, the leaders are able to affect the employees' extrinsic motivation. In terms of a work environment, employees' behavior is usually driven by rewards, either monetary or non-monetary. While monetary incentives are considered to appeal to an individual's extrinsic motivation, non-monetary incentives can appeal to both intrinsic or extrinsic motivation (Erkal, Gangadharan & Koh, 2018).

It is important to notice that there are some factors that are not causing higher motivation when present but will cause dissatisfaction if not present. Herzberg's two factor model shows that it is not enough to remove dissatisfiers to improve the performance at work, providing satisfiers is also needed. Motivator factors, satisfiers, are the factors that motivate the workers, but the lack of the factors does not result in dissatisfaction. It is important to make sure that the motivator factors exist in the organization by for example starting different reward and recognition programs. According to Herzberg's research, the different satisfiers are: achievement, recognition, work itself, responsibility, advancement and growth. Hygiene factors, dissatisfiers, are the factors that dissatisfy the workers if absent, but they do not cause higher motivation if they are present. Hygiene factors are for example: company policy and administration, supervision, work condition, salary, relationship with peers, personal life, status and security. (Fisher, 2015)

The focus on improving motivation at the workplace should be on motivator factors rather than hygiene factors, since motivator factors can increase the motivation. In this study, the focus is going to be on one motivator factor especially: recognition.

2.2 Key mechanisms affecting motivation

Since this thesis examines the effect of public recognition on performance, it is important to determine the different mechanisms that affect motivation. The following three mechanisms are especially important to this study.

2.2.1 Conformity

Conformity means that people want to match their behavior to the group (e.g. Bernheim, 1994; Bradler et al., 2016). Most of the people are willing to conform, because they are afraid that acting differently will affect their social status. Therefore, people may have preferences for behavior that deviates from the group behavior but choose to conform to a behavior that's

identical to the group behavior. (Bernheim, 1994) According to Hollander (1958), conforming to group expectations makes the members of the group to get a more positive impression of the conforming member and increases that member's status in the group.

Chen, Harper, Konstan and Li (2010) ran a field experiment on what social comparisons do to the contributions on social communities online. The experiment was carried out on a movie recommendation community online, and they found that after a user in the community got information about how many movies a median user rated, the users below the median increased their ratings by 530% in a month. The people above the median decreased their ratings by 62%. Conformity is an important mechanism here, which makes people behave similarly with the majority.

During the experiment, preferences for conformity might lead to changes in performance after the reward is introduced. The participants get to know if their performance was among the best or worst ones and those who have high preference for conformity may increase their performance, also wanting to be among the best, if they belonged to the group of people who did not receive any reward. On the contrary, participants who received a reward might even decrease their performance to match the group's behavior.

2.2.2 Altruism

Altruism is generally defined "as behavior that benefits another organism, not closely related, while being apparently detrimental to the organism performing the behavior" (Trivers, 1971, p. 35). In addition to that, people can only be altruistic if it includes the intention of helping one other (Sussman & Cloninger, 2011). According to Bradler et al. (2016, p. 4), "employees are conditionally altruistic, meaning that they care more for an employer who cares for them." Subsequently, employees increase performance when they perceive the employer's kindness.

Already in the 18th century, Adam Smith introduced his ideas of altruism; people not only being selfish but also interested of others well-being.

> *How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it, except the pleasure of seeing it.*
>
> (Smith & Hanley, 2009, p. 13)

In the experiment, the reward may work as a signal to the subjects of the experimenter's kindness. Transferred to the real work-life, altruism is considered as some kind of reciprocity, however, in regard to reciprocity, only employees who receive a reward will increase their

efforts, while all employees will increase performance according to altruism, although, just a few ones receive a reward. That is because altruism claims that all employees can perceive the employer's kindness.

2.2.3 Reciprocity

"A reciprocal action is modeled as the behavioral response to an action that is perceived as either kind or unkind" (Falk & Fischbacher, 2006, p. 294). In this paper, reciprocity means that if employees are treated well, they want to give something back. According to Berg, Dickhaut and McCabe (1995, p. 122), "a fundamental assumption in economics is that humans act in their own self-interest". On the contrary to Berg et al. (1995), Bradler et al. (2016) found out that employees receiving a reward, either monetary or non-monetary, may feel the need to reciprocate.

Typically, people think that wages trigger the reciprocal behavior, but the experiment of Kube et al. (2012) shows different results. They examined in a field experiment how workers' performance changes with monetary and non-monetary rewards. The important mechanism in their study was reciprocity. The findings suggest that it does not matter much, whether the reward is monetary or non-monetary, but the time and effort invested to the gift matters more. Thus, in their experiment, workers reciprocate by performing better when they notice that the employer has put a lot of effort in the gift they receive.

This study aims to replicate previous findings by increasing people's performance through public recognition. Thus, the focus is on non-monetary rewarding, which should, according to the literature, lead to reciprocal behavior.

2.3 Rewards

This paper examines how recognition and non-monetary rewards may affect performance. There are different types of rewards: monetary and non-monetary. Monetary rewards come in monetary form, whereas non-monetary rewards are not money, even though they can have monetary value (e.g. Fisher, 2015; Sonawane, 2008). However, some researchers narrow this definition down and exclude any involvement of money in non-monetary rewards (e.g. Mathauer & Imhoff, 2006; Manolopoulos, 2007). This paper mostly concentrates on symbolic rewards, which are rewards with no significant monetary value, but that are nevertheless valuable to the person receiving it (Kosfeld & Neckermann, 2011, p. 87). Recognition is used for acknowledging employees for their good performance. There is usually no reward without

recognition, but there may be recognition without a reward (Zeb, Jamal & Ali, 2015). Recognition itself may also be considered as a symbolic reward.

2.3.1 Monetary rewards

Monetary rewards are the incentives employees receive in monetary form, for example salaries, bonuses or stock options. Monetary rewards may boost the extrinsic motivation, but they do not have impact on intrinsic motivation (Waqas & Saleem, 2014). Often the monetary rewards, for example bonuses, are implemented to promote sales or boost the result of the company.

Early research in the field of motivation has focused on monetary rewards. Porter and Lawler (1968) developed a theory about managerial attitudes and performance behavior. The theory was tested in different organizations and verified that pay is a function of effort. Managers work harder when pay is considered both as a reward and pay for performance.

In some organizations monetary rewards have worked well, for instance Safelite Glass Corporation changed their compensation method from hourly wages to piece-rate pay. On average level, the output per worker increased by 44 percent (Lazear, 2000). Hence, piece-rate pay works best in simple jobs, where it is easy to measure the performance. Usually, the jobs which are paid by piece-rate, there are only few tasks to do and the skills required are not that high. These kinds of jobs are usually not motivating. (Lazear & Gibbs, 2014, p. 166)

Many companies in the manufacturing industry, where there are lots of piece-rate jobs, were following the Taylorism practice, which means that firms hire skilled engineers who work out the best way to organize production first and then hire people placing them on the different production tasks by their skills (Lazear & Gibbs, 2014, p. 166). Taylorism has not been popular lately, since the well-being of employees has become an important topic. However, monetary rewards can be good when they are accompanied with non-monetary rewards. Kvaløy, Nieken and Schöttner (2015) found that motivational talk improves performance, but only when it is complemented with performance pay. In their experiment, the performance pay alone even reduced the performance.

However, according to Gneezy and Rustichini (2000), performance related pay is important in all kinds of jobs, and not only in simple jobs. They found out that most of the times, offering more money led to higher performance. If the monetary incentives are low, the effect on performance may be negative, but high monetary rewards increased performance. Therefore, Gneezy and Rustichini (2000, p. 791) characterized the saying “Pay enough or don’t pay at all”.

Following Herzberg's two factor model, money is a hygiene factor, which means that it does not cause higher motivation and lead to better performance, however, if money was absent, employees would be dissatisfied (Fisher, 2015). Van Herpen, Van Praag and Cools (2005) were able to prove empirically that money only works as a boost for extrinsic motivation, but not for intrinsic motivation.

2.3.2 Non-monetary rewards

Basically, non-monetary rewards are rewards where there is no money involved directly. They affect the employee's intrinsic motivation. Theory does not state the one and only definition about non-monetary rewards. On the one hand, some sources state that non-monetary rewards are for example recognition, coaching, promotion and flexibility at work (e.g. Mathauer & Imhoff, 2006; Manolopoulos, 2007). On the other hand, there are definitions that include such intangible rewards, but also gift cards, tailored events, entertainment tickets and different tangible things (e.g. Fisher, 2015; Sonawane, 2008).

Non-monetary rewards are considered to help recognizing employees. There has been a lot of research proving that non-monetary rewards are effective at improving performance, especially because of the effect on intrinsic motivation of the workers. Many times, the non-monetary rewards lead to even better performance than monetary rewards (e.g. Asraf, Bandiera & Jack, 2014; Waqas & Saleem, 2014; Kvaløy et al., 2015) and are more effective in the long run (e.g. Lawler, 1969; Zobal, 1999).

In general, the rewards at work are something that the employees are proud of. Usually people do not talk about how much they earn, but people are not shy to mention the rewards they have gotten at work. Considering the employees above the minimum level of salary, monetary incentives show diminishing returns (Fisher, 2015).

Kosfeld and Neckermann (2011) verified that a congratulation card affects employees' performance positively. Bradler et al. (2016) further examined Kosfeld and Neckermann's (2011) findings and conducted a similar experiment. But instead of rewarding the employees in the beginning of the work period, the employees' performance was rewarded in the middle of the working period surprisingly. There were four treatment groups, which varied in how many subjects got a reward. An increase in work performance was noticed in this experiment as well, which was mostly driven by positive response from those participants that did not receive a reward.

This research focuses on non-monetary rewards and recognition, which will be further examined in the next chapter. The reward in the experiment that goes along with this thesis was a certificate of achievement (Appendix A) given to the five best performers in the experiment. Hence, the reward can be considered as symbolic.

2.3.3 Recognition

> *Recognition is first and foremost a constructive response; it is also a judgment made about a person's contribution, reflecting not just work performance but also personal dedication and engagement. Lastly, recognition is engaged in on a regular or ad hoc basis, and expressed formally or informally, individually or collectively, privately or publicly, and monetarily or non-monetarily.*
>
> (Brun & Dugas, 2008, p. 727)

Recognition is a good tool which can be applied in every workplace. It boosts the employees' motivation which improves their engagement in work and leads to better performance (Burgess & Ratto, 2003). Sometimes even a small gesture, like the leader saying "well done" to his employee, is enough to recognize the employee and keep him motivated (Fisher, 2015). In addition, by recognizing employees, the employer can signal what kind of work behavior is desirable (Frey, 2007).

According to Vroom (1964), the employees work hard if the goals are extrinsic and the employees believe that they are skilled enough to achieve the goals. The type of rewards that work the best are non-monetary such as recognition and promotion, but also higher wages motivate some people. Appelbaum and Kamal (2000) agree with Vroom that income has at least a small effect on motivation. However, employees need recognition for their accomplishments. Recognition increases productivity, fosters creativity and inspires employees. Therefore, high employee turnover is often based on a lack of recognition at the workplace.

Luthans (2000, p. 38) found out that "employees value highly personalized, instant and specific social rewards like recognition, attention, and sincere appreciation, which are based on the employees' efforts and not on how long they have served in the company". Hence, effective leadership makes use of social rewards, such as recognition, to increase motivation. Furthermore, Stajkovic and Luthans (2003) studied whether behavioral management has a positive effect on task performance. They found out that money, feedback and social recognition have a significant effect on task performance. But even more important is that the effect was even stronger when all the reinforcers were used in combination.

In the experiment, the five people who were able to decode most sequences are publicly recognized in front of everyone else in the room. They are asked to come to the front of the classroom and shake hands with the experimenter, receive the certificate of achievement (Appendix A) and are congratulated by the experimenter. Thus, the rewarding mechanism in this study is highly relatable to recognition. Therefore, the result of this study can also contribute to existing empirical work in the field.

3 Methodology

The purpose of this thesis is to investigate the effect of public recognition on performance in competitive environments. As mentioned earlier, similar studies have been made, but to my knowledge none of them have tried to find out if the effect of public recognition on performance stays the same when recognition is repeated. Own data was collected by running a real-effort experiment with students of the University of Stavanger. The task in this experiment was to decode a sequence of letters into numbers.

3.1 Method

There are two different approaches for scientific research: qualitative and quantitative. Both approaches have their positive and negative sides, hence it is important to find the approach that fits the research questions best.

Empirical evaluation involving numerical measurement is typical for quantitative business research (Zikmund, Babin, Carr & Griffin, 2013, p. 133). Quantitative methods measure the effect of one or more independent variables on a dependent variable using numerical measures (Lakshman, Sinha, Biswas, Charles & Arora, 2000). The heart of the quantitative strategy is to test if the theory works in practice (Bryman, 2012). Many times, discovering something new is easier to accept as a fact if it is something that can be quantified. When there are not that many dependent variables and the experimenter has control over them, the quantitative method works well (Lakshman et al., 2000).

Qualitative research often gathers data from interviews and focus groups. In this study, the data is gathered from an experiment with a larger sample size than what is typical for qualitative research. Qualitative research is often not generalizable because of the small sample sizes. Therefore, the data collected during the experiment is quantitative and can be analyzed by using statistical methods, which could be more challenging for qualitative types of data. (Lancaster University Management School, 2016)

This study is dedicated to the effect of public recognition on performance in competitive environments. Since the quantitative approach should be applied preferably, performance can be measured very well, and the effect size matters a lot to the overall outcome of this thesis, the quantitative approach has been chosen to examine if public recognition influences performance and how strong this effect is.

There are three different types of business research: exploratory, descriptive and causal. Exploratory research is used to clarify unstructured research problems. Descriptive research is for describing the characteristics of objects or people. It usually answers to who, what, where and how type of questions. Hence, the data is collected without environment manipulation. The suitable research type for this study is causal research, which tries to find cause-and-effect relationships. (Zikmund et al., 2013, p. 51-56) In this study, the purpose is to find out if public recognition causes changes on performance.

According to Zikmund et al. (2013, p. 57), experiments are a good method to find cause-and-effect relationships. The major advantage with an experiment is that the independent variable can be manipulated to see if there are changes in the dependent variable. In this case, the dependent variable is performance and the independent variable is public recognition.

An economic experiment can be conducted as a field experiment or a laboratory experiment. Field experiments are conducted in a naturally occurring environment, whereas laboratory experiments are conducted under controlled conditions in an artificial setting. (Zikmund et al., 2013, p. 267-268).

A laboratory experiment was chosen for this study, because the researcher has better control over the whole research setting. In a laboratory setting, it is easier to isolate other variables that could affect the performance and concentrate on what is really of interest, how public recognition affects performance. In addition, the time and budget are quite limited, hence, a field experiment would not be feasible.

Laboratory experiments in economics and psychology are similar concerning the planning and the design of the experiments, but they have different aims and interests. In economics, an induced valuation theory is commonly applied. This theory's central idea is that subjects prefer more reward medium instead of less. Economists believe that people want to maximize their utility considering the costs and benefits. The amount of the reward or if the reward is even accessible is depending on the actions of the subject. Consequently, experimental subjects are compensated according to the theory that is tested and incentives are used to be able to strengthen the validity of the experiment. This should also ensure that the experimental subjects are acting as they would in real life. Psychologists believe that using incentives may reduce the validity and the experiment may not be transferable to real life. In psychological experiments, the subjects are often deceived to be able to create a laboratory setting which is not that artificial, this is rarely done in economic experiments. (Ariely & Norton, 2007; Smith, 1976)

The specific type of laboratory experiment that is chosen in this study, is a real-effort experiment, which means that subjects work on a specific task (Brüggen & Strobel, 2007, p. 232). A real-effort experiment is chosen, because the results of such experiments should be more field-relevant and more generalizable than other types of laboratory experiments. Real-effort experiments are also less artificial than stylized-effort experiments (Dutcher, Salmon & Saral, 2015).

Major concerns with laboratory experiments are, that the data is unrealistic and that students as experimental subjects are an unrepresentative group. The real-effort experiment was chosen to counteract the first concern. The external validity, if the results are applicable beyond the experimental subjects, is claimed to be a problem in laboratory experiments. Usually the experiments are made using students, which may be a group that is not representing the whole population, hence the external validity is lowered. (Zikmund et al., 2013, p. 273) Students are also used in this experiment because of money, time and convenience issues. However, the experimental subjects being students should not be a problem for external validity in this case, since the task does not require any prior knowledge and in addition, the task is really simple.

Nevertheless, there are many positive sides of laboratory experiments as well. First, a laboratory allows tight control, which is required for causal knowledge, the type of experiment that will be conducted. Second, the experiment should be easy to replicate, because there are fewer variables the experimenter has no control of compared to field experiments. Third, laboratory experiments come usually at a lower cost than field experiments, which makes them more suitable for theses and studies where there are financial constraints. Fourth, the randomization of the experimental subjects to different treatments is easier than in the field. (Falk & Heckman, 2009; Zikmund et al., 2013)

3.2 Experimental design

The subjects worked on a task of decoding a sequence of letters into numbers. This kind of task has been used in experiments before (e.g. Clark & Friesen, 2009; Charness, Masclet & Villeval, 2014; Gjedrem, 2016; Erkal, Gangadharan & Nikiforakis, 2011) and was chosen, because it is simple, easy to understand and does not require any prior knowledge. A simple task was chosen to avoid other factors affecting the performance than the treatment subjects receive. In addition, with such a task, it is easy to measure the performance, since the indicator is the amount of correctly decoded sequences of letters.

Figure 2 shows a list of letters and corresponding random numbers, like the subjects had it during the experiment, exemplarily. The subjects had to decode a sequence of four letters. The letters were from A to Z and numbers used from 1 to 99, whereas numbers are distinct within such a list, meaning that one number can just be assigned to one letter at the same time.

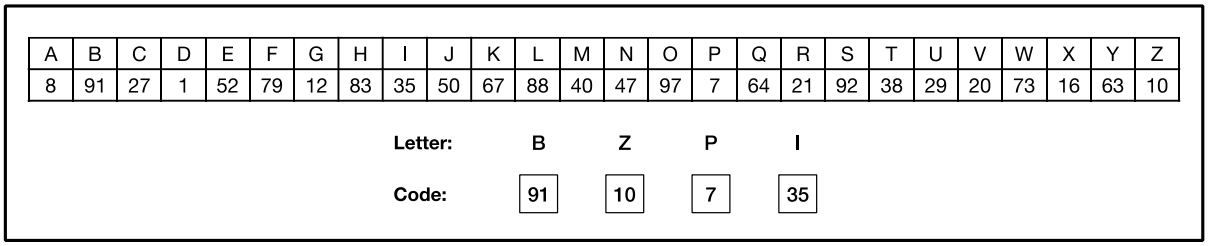


Figure 2: Example of a decoding task

The experiment consisted of four stages, each stage lasting for four minutes. The subjects were able to continue with a new sequence of letters, even if the one before was not decoded correctly. Only correctly decoded sequences were added up to measure the performance.

The following subchapters explain the experimental design from the first to the fourth stage for the control and the treatment group. Figure 3 shows the process flow of the whole experiment distinguished between both groups. The sessions for the control and treatment groups were scheduled at different times and subjects did not know that there is a treatment or control group.

3.2.1 Control group

In the control group, there was a one-minute break in between the stages and the subjects remained seated during these breaks. The decoding list was changed after every stage to prevent the subjects from memorizing the sequences.

3.2.2 Treatment group

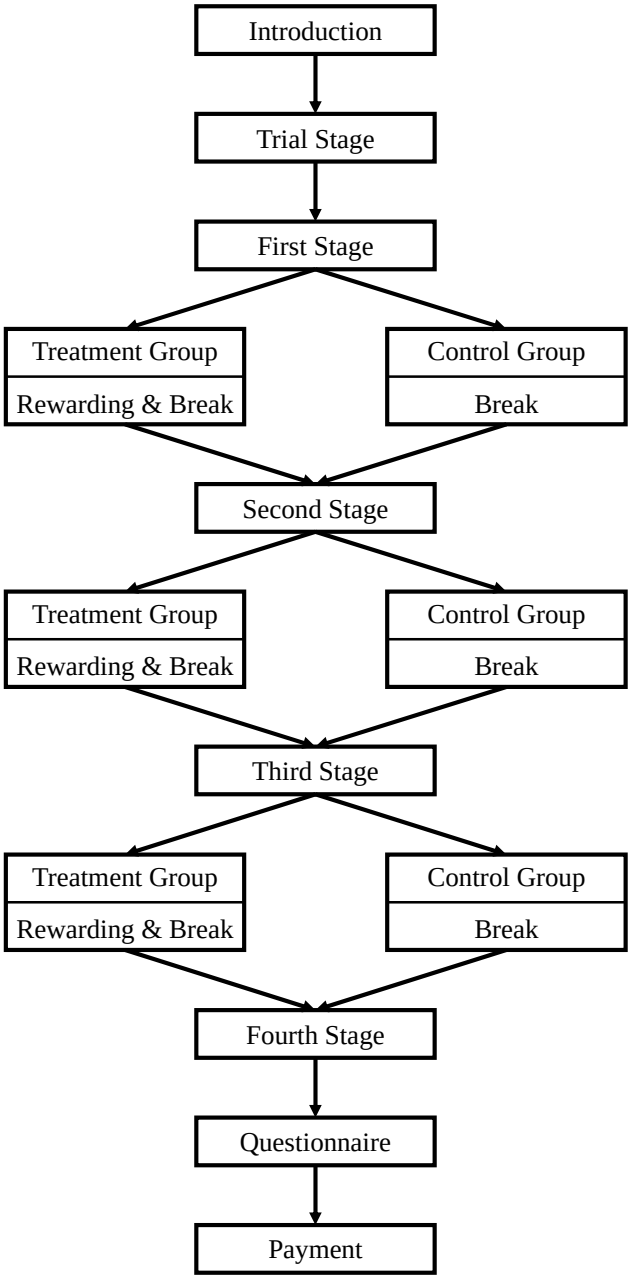


Figure 3: Process flow of the experiment

Instead of a break in between the stages, in the treatment group public recognition and a symbolic award was given to the subjects after the first, second and third stage. On the screen of the experimenter, it was shown who was among the top five of all the experimental subjects within the room. As already mentioned, the number of correctly decoded sequences was used as performance indicator. There were numbers attached on the computer screens, thus the subjects knew which number they were assigned to. After the first, second and third stage, the experimenter called the subjects that belonged to the top five to the front of the classroom to receive their reward (Appendix B). To check that just the subjects, that were called, come to the front, the same numbers that were on the front of the subjects' computer screens were also attached to the back of the screens. Subjects did not know before that point that some of them will receive a reward. The possibility of getting the reward in the treatment groups was almost equal, the size of the groups varied from 16 to 19 subjects, and the top five were rewarded in each group. The reward was a symbolic reward with no monetary value, in this case a certificate of achievement (Appendix A). When the subjects received the certificate of achievement, the experimenter also shook hands with them. This is how public recognition was ensured in the experiment. However, there was no recognition and reward after the last stage, since rewarding them would not have affected the final results, and it would have also made the experiment unnecessarily longer. As in the control group, the decoding list was changed after each stage.

3.3 Procedure

The experiment was conducted at the University of Stavanger on Wednesday, March 7th, 2018 and Thursday, March 8th, 2018. There were three sessions each day at 10:00, 11:30 and 13:00. The experiment lasted approximately 30 minutes for both treatment and control groups. On Wednesday session at 13:00 and on Thursday session at 10:00, the participants were assigned to the control group, participants of other sessions were assigned to the treatment group. The control groups were chosen to be on different days and different times to ensure that neither day nor time have an influence on the experimental results. Furthermore, the students decided themselves the time they want to show up and were not informed about which group they are assigned to. Approximately 100 subjects were needed for the experiment, 30 for the control group and 70 for the treatment group. There were 14-19 people participants in each session.

The subjects were recruited by sending an invitation (Appendix C) to students of the University of Stavanger by using their student e-mail addresses. Some students got the e-mail also in their private e-mail. The sign-up link was shared on Facebook in relevant UiS groups. The survey for signing up was created with SurveyMonkey[1]. A software, called z-Tree, was used for programming the experiment task which was carried out on computers. Z-Tree (Zurich Toolbox for Ready-made Economic Experiments) is a software used for developing and conducting economic experiments (Fischbacher, 2007).

While the previous chapter explained the process from the first to the fourth stage, the following part of this chapter focuses on the other parts of the process that are shown in Figure 3 as well. In the beginning of the experiment, the subjects were seated and informed about their task and that it is about decoding as many sequences of letters as possible within the given time. When they were done reading the instructions (Appendix D), which were already placed on the desks when subjects entered the room, the experimenter welcomed everyone and asked if there are any questions about the experiment (Appendix B). Then, subjects could familiarize themselves with the program and the task for one minute during a trial stage which was like the other stages, but shorter and without recognition. After that, the list of letters with corresponding numbers was changed to ensure that subjects did not memorize the list. The instructions and the task for both control and treatment group were the same. The experiment started, and subjects worked on their decoding tasks and had breaks in between the stages or got recognition, depending on which group they belonged to. The experimenter was sitting in front of a computer screen,

[1] https://www.surveymonkey.com

following how the experiment was going. The information about the best five performers was only relevant in the treatment group. When the experiment was done; the subjects were able to see how they performed in each stage. A questionnaire, which asked for gender, age and faculty was then filled out. After that, the subjects received their show-up payment. All subjects were given NOK 50 show-up pay. The payment was given to provide an incentive for participating and thus, recruiting more subjects for the experiment. A fixed pay was chosen to make sure that the payment is independent from the work performance. This has also been done in the field experiment by Bradler et al. (2016).

3.4 Hypotheses

There are four research questions to investigate. First, I assume that subjects who were rewarded in the first stage will perform better in the second stage. As already exposed, previous research found that symbolic rewards and non-monetary incentives increase performance (e.g. Stajkovic & Luthans, 2003; Kostfeld & Neckermann, 2011; Bradler et al., 2016). Furthermore, due to reciprocity, altruism and conformity, theory strongly supports this assumption (e.g. Bernheim, 1994; Levine, 1998; Bradler et al., 2016).

> **H_1:** *Public recognition after the first stage has a positive effect on performance in stage two for rewarded subjects.*

The second hypothesis is that subjects who were not rewarded in the first stage of the experiment will also perform better in the second stage. Reciprocity does not apply in this case; however, theory suggests that altruism and conformity lead all subjects to perform better, even the ones who did not receive a reward. Conformity encourages the subjects to follow the performance of the group and perform like the majority of the group (Bernheim, 1994). Not rewarded subjects still perceive the experimenter's kindness by rewarding for performance, and according to altruism, they are willing to increase their performance as well (Levine, 1998).

> **H_2:** *Public recognition after the first stage has a positive effect on performance in stage two for subjects who are not rewarded.*

The third research question, that arises and has to be examined, is if rewarded or not rewarded subjects perform proportionally better. I assume that subjects who were not rewarded in the first stage experience a higher increase in performance than subjects who were rewarded. Theory gives reason to assume that conformity is the strongest mechanism and leads to even better performance for subjects who do not receive a reward (Bradler et al., 2016).

H_3: *Public recognition on past performance has a stronger positive effect on future performance of subjects who were not rewarded compared to subjects who were rewarded.*

To examine if the effects on performance remain or change when recognition is provided repeatedly, the fourth research question is formulated. Note that the third and fourth stage of the experiment are not covered by any other research question, yet. Since there is a limit on how much the performance can increase and how much it can be lowered by decreasing motivation, I assume that a comparison of the performance from the second to the third stage reveals the same effect direction, but that the effect size is weaker. Consequently, I assume that a rewarded person who performs better due to recognition will perform even better in the next stage, however, the performance increase will be smaller. Furthermore, I assume the same pattern for not rewarded subjects.

H_4: *The effect of public recognition on performance remains when a subject is rewarded in every stage. The effect also remains when a subject is not rewarded in any of the stages. However, the effect size is smaller.*

4 Experimental results

In total 168 people signed up for the experiment and 101 showed up. However, one subject is excluded from further analysis, since his or her answers could distort the results. In the last round, he or she decoded 32 sequences of letters wrong, which is more than the best subject decoded correctly, and in addition, he or she decoded 17 sequences correctly and zero wrong in the round before. Thus, 100 subjects are relevant to further analysis.

4.1 Descriptive statistics

Sixty percent of the subjects were male, and forty percent were female. The subjects' age was ranging from 19 to 54 ($M = 26.80,\ SD = 5.36$) and is shown in Figure 4.

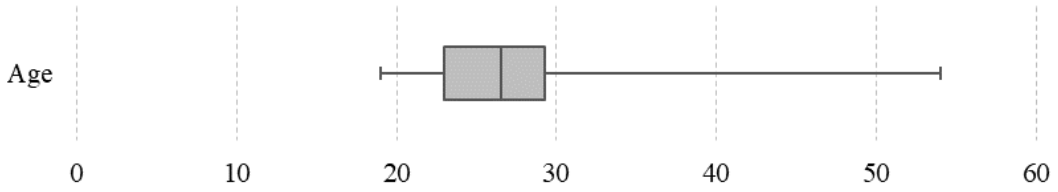


Figure 4: Boxplot with subjects' age

Since the experiment was conducted at the University of Stavanger, subjects were asked about the faculty they belong to. Thus, all six faculties were given as options, however none of the subjects studied at the Faculty of Performing Arts. One further option was "I'm not a student", which was chosen by one subject. More than half of the subjects were studying at the Faculty of Science and Technology and almost one third at the Business School as shown in Figure 5.

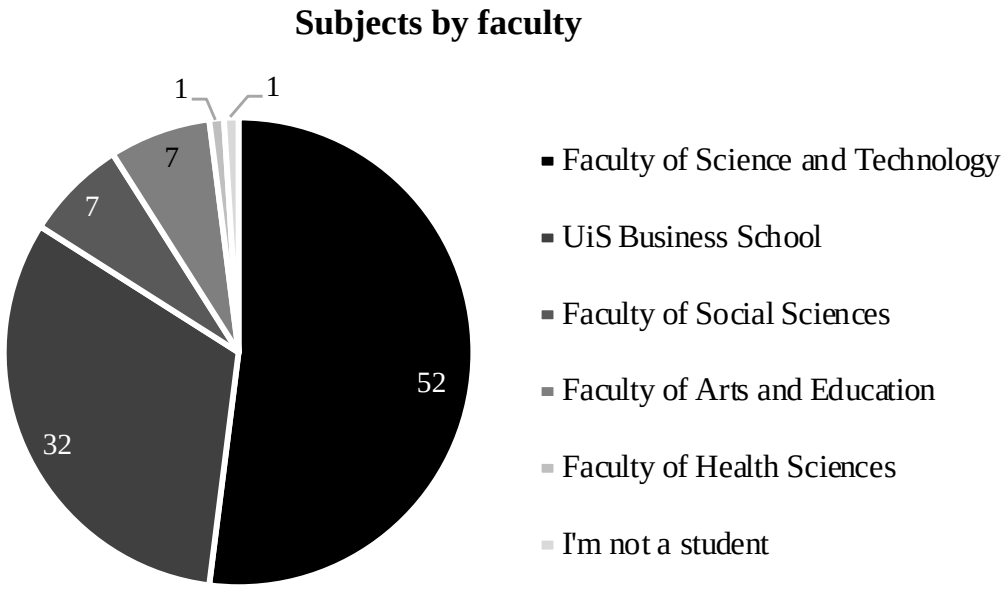


Figure 5: Pie chart with subjects per faculty

Figure 6 shows the distribution of the subjects by the time they participated in the experiment.

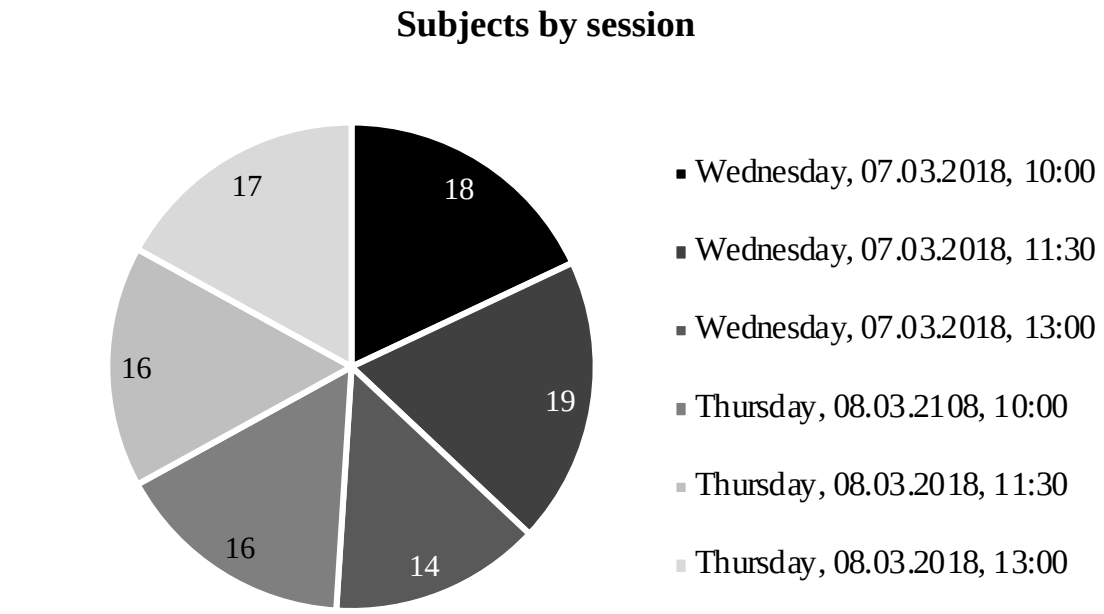


Figure 6: Pie chart with subjects per session

The subjects' performance for each stage distinguished between treatment and control group is presented as a boxplot in Figure 7. While the number of correct answers from quartile one to three is more distributed for the control group, the treatment group has stronger outliers. Nevertheless, one has to consider that there are 30 subjects in the control group, but 70 in the treatment group. The control group's median is higher in stage one and two, however, the treatment group's median is higher in stage three and four.

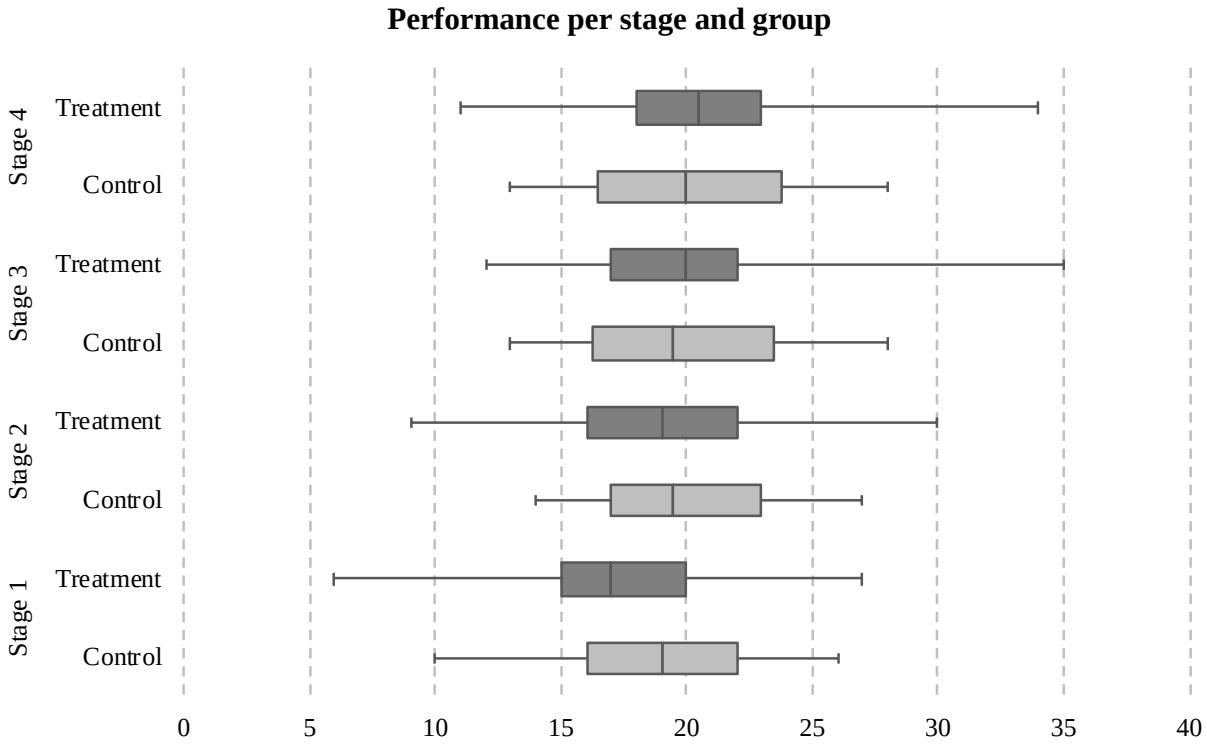


Figure 7: Boxplot with performance per stage and group

Figure 8 shows the subjects' performance in the treatment group distinguished between subjects who were rewarded in the previous stage and the ones who were not rewarded. Therefore, the first stage cannot be divided and is shown for the whole treatment group. All quartiles of the subjects who were rewarded in the previous stage have a higher value than the quartiles of the ones who were not rewarded. One should consider that in the treatment group there were 20 people who were rewarded at each stage and 50 people who were not rewarded.

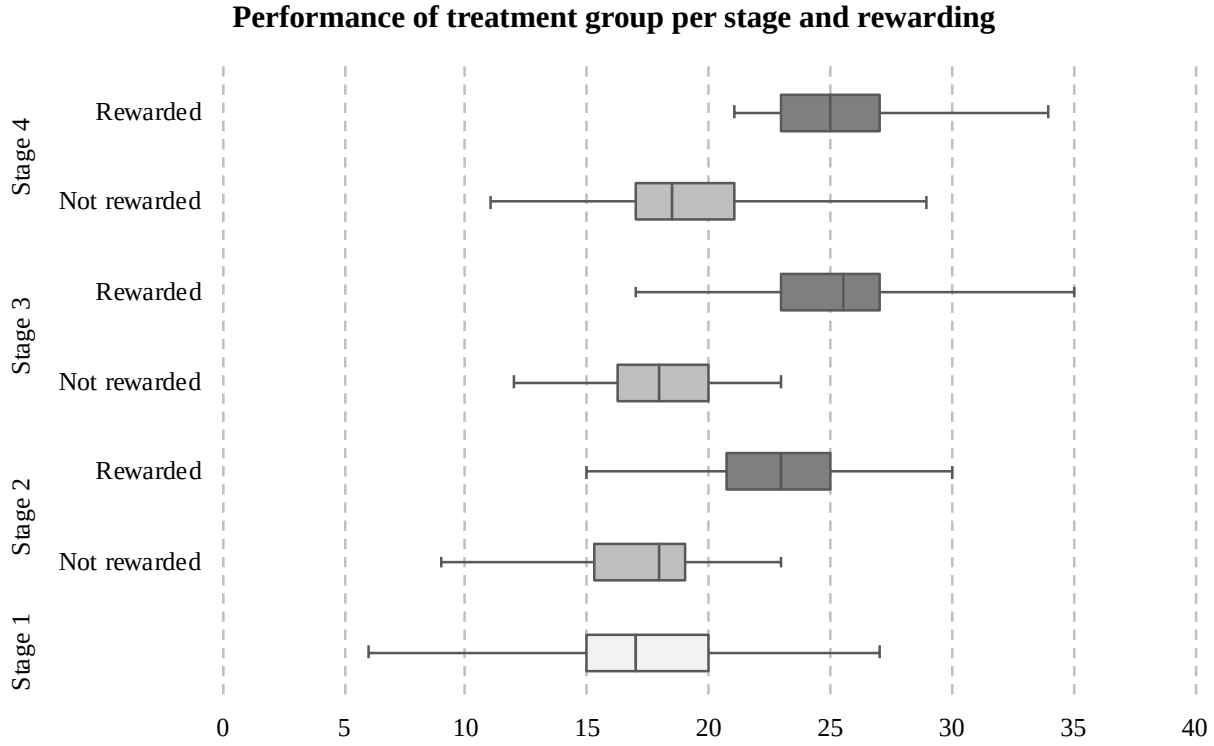


Figure 8: Boxplot with performance of treatment group per stage and rewarding

Table 1 reports means and standard deviations for correct number of answers (performance) for each stage distinguished between control and treatment group. While the treatment group started with a lower performance in average and increased the performance in each stage, the control group started with a higher average performance, however, already in stage three the performance decreased, but increased again in stage four.

	Treatment		Control		Difference between groups
	Mean	Std. dev.	Mean	Std. dev.	
Stage 1	17.09	4.60	18.77	4.04	-1.68*
Stage 2	18.97***	4.09	20.30***	3.86	-1.33
Stage 3	20.01***	4.65	19.93	4.12	0.08
Stage 4	20.57**	4.40	20.33	4.38	0.24

***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Significance within a group is reported horizontally and between groups in the last column.

Table 1: Means and standard deviations of performance

Table 2 reports, similar to Table 1, means and standard deviations for the difference of correct number of answers between stages (performance changes) distinguished between control and treatment group. This table emphasizes the increase in performance from stage to stage for the treatment group, but also reveals that this increase gets smaller and that subjects deviated less and less from the average. In the control group, there was a huge increase from stage one to two, however, after the decrease in stage three and the increase in stage four, the average performance was almost the same as in stage two.

	Treatment		Control		Difference between groups
	Mean	Std. dev.	Mean	Std. dev.	
Stage 1 to 2	1.89	3.22	1.53	1.91	0.36
Stage 2 to 3	1.04**	2.61	-0.37***	1.35	1.41***
Stage 3 to 4	0.56	2.03	0.40	1.55	0.16

***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Significance within a group is reported horizontally and between groups in the last column.

Table 2: Means and standard deviations of performance changes

Figure 9 presents the average performance of subjects for each stage distinguished between the session times. The sessions with treatment groups look very similar, however the average performance seems to be a little bit lower for the session at Thursday, 11:30. The two control groups differed more from each other, although both of them started with a high performance in stage one compared to the treatment group.

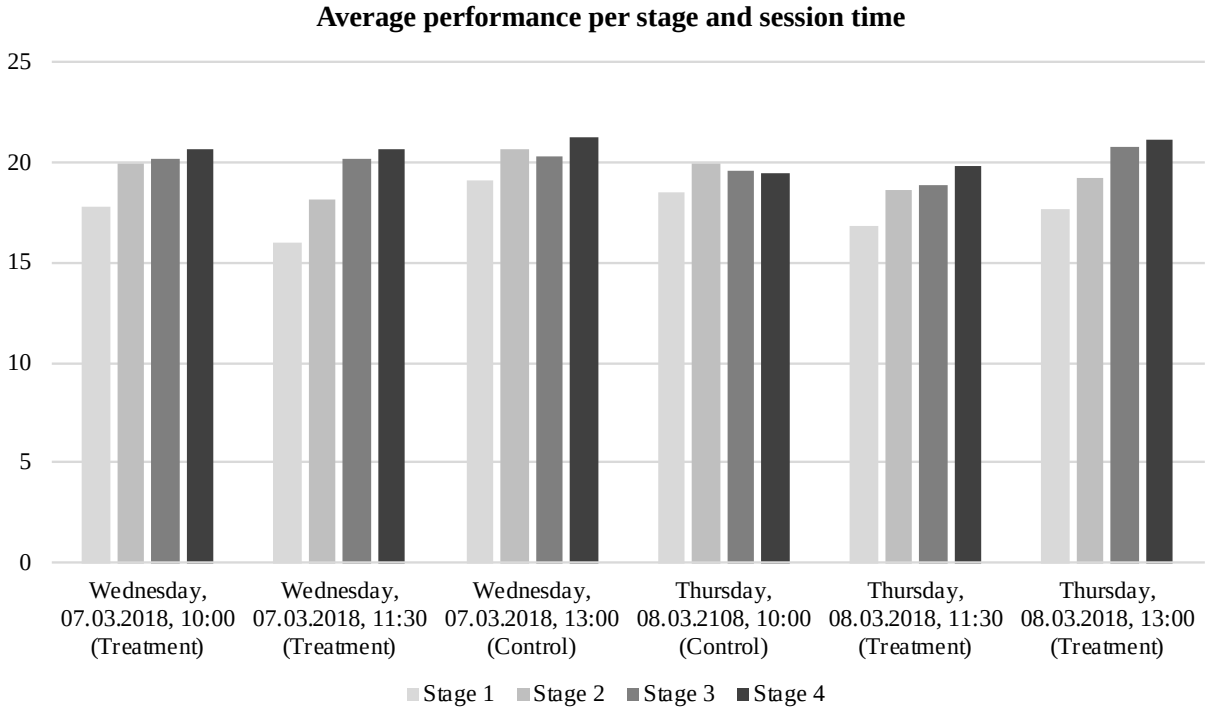


Figure 9: Bar chart with average performance per stage and session time

4.2 Hypothesis testing

This chapter examines the hypotheses from chapter 3.4. Subsequently, this chapter is divided into four parts in respect to the four hypotheses.

4.2.1 Effect on rewarded subjects

The first hypothesis suggests that public recognition for performance in stage one has a positive effect on performance in stage two for people who were rewarded. Therefore, just the rewarded subjects' performance changes from the first to the second stage are taken into account. The boxplot in Figure 10 shows the performance in the first and second stage of subjects who were rewarded after the first stage. In this case, the number of subjects is $N = 20$.

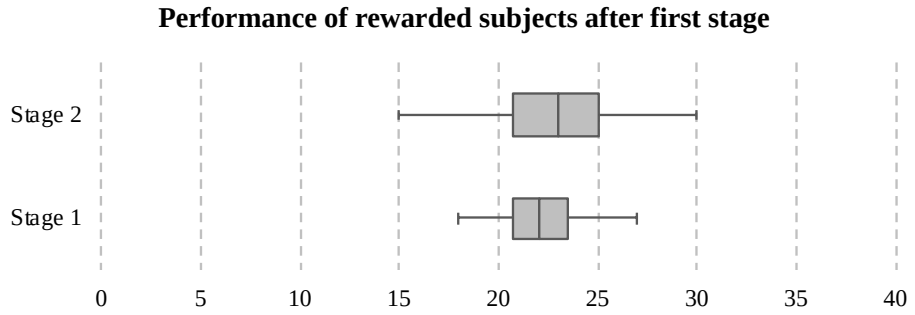


Figure 10: Boxplot with performance of rewarded subjects after first stage

There are two tests to analyze this hypothesis, because of the experimental within-subject design. The parametric paired-samples *t*-test would be chosen preferably, however it requires a normal distribution of the samples' differences (Field, 2013, p. 371), which is not given in this case, as Figure 11 reveals. In addition, the Kolmogorov-Smirnov test suggests that the performance change, $D(20) = 0.281, p < .001$, is significantly non-normal. Subsequently, the Wilcoxon signed-rank test (Wilcoxon, 1945) has to be used, which is the non-parametric equivalent to the paired-samples *t*-test (Field, 2013, p. 228).

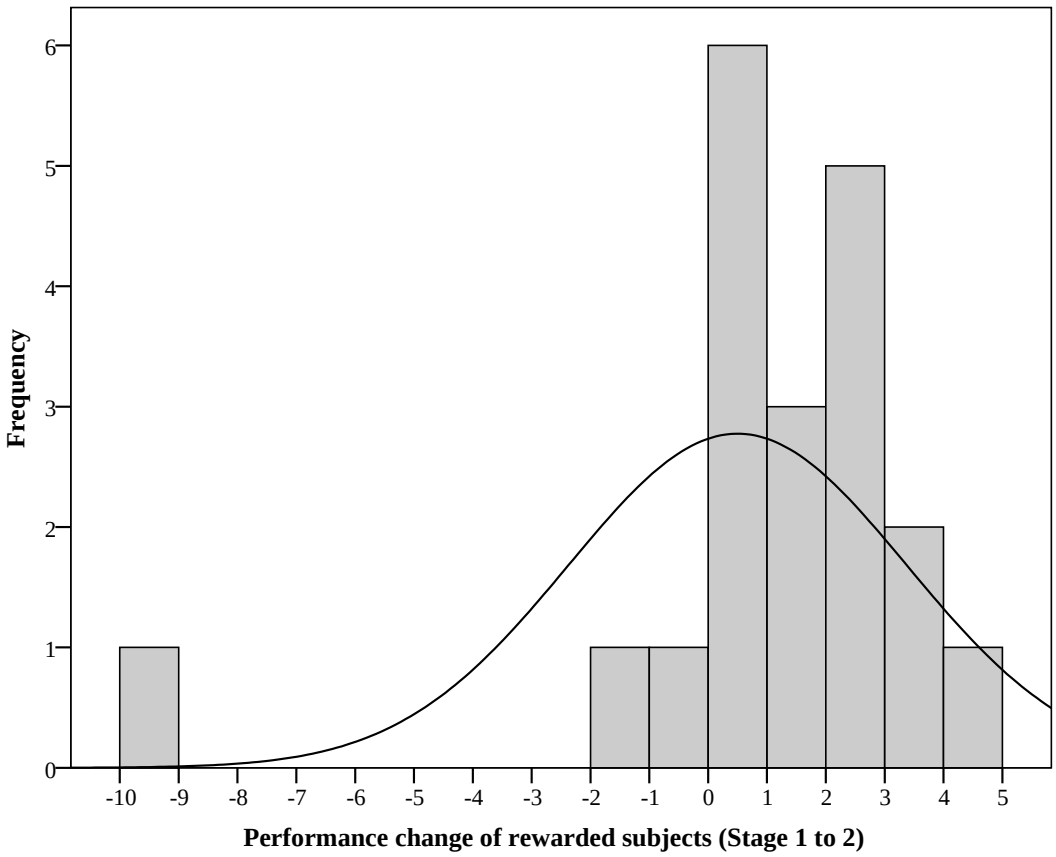


Figure 11: Histogram of rewarded subjects' performance change (Stage 1 to 2)

The results suggest that for rewarded subjects, the performance was significantly higher in the second stage ($Mdn = 23$) than in the first stage ($Mdn = 22$), $T = 81$, $z = -1.81$, $p = .07$, $r = -.40$. According to Rosenthal (1991, p. 19), the z-score can be converted into the effect size estimate, r, as follows:

$$r = \frac{z}{\sqrt{N}}$$

According to Cohen (1992), the effect size estimate $r = -.40$ represents a medium to large change in performance for subjects who were rewarded. Since the test was significant at the 10% level, the null hypothesis can be rejected and hypothesis H_1 can be accepted in the first place.

Nevertheless, it has to be examined whether the performance increase is driven by public recognition or something else, for instance learning effects. Therefore, a 2 (treatment) × 2 (top five) ANOVA was conducted on subjects' performance changes from stage one to two. Since there were no rewards in the control group, a dummy variable was calculated to indicate that the top five of the control group would have gotten a reward. Table 3 indicates that the performance change is not driven by the treatment at all ($F(1, 96) = 0.000$, $p = .994$), but by

being among the top five of the group ($F(1, 96) = 2.980$, $p = .088$). However, there is no significant effect of the two-way interaction between the treatment and the top five conditions on subjects' performance change ($F(1, 96) = 1.436$, $p = .234$).

	Performance change from stage one to two		
	Degrees of freedom	*F*-value	*p*-value
Treatment	1	0.000	0.994
Top five	1	2.980	0.088
Treatment × Top five	1	1.436	0.234

Table 3: Univariate analysis of variance for performance change from stage one to two

The results are also confirmed by testing for differences in performance change for the top five distinguished between control and treatment group. The performance increase was insignificantly higher for the control ($Mdn = 2$) than for the treatment group ($Mdn = 1$), $U = 83.5$, $z = -.75$, $p = .475$, $r = -.14$. These results diminish the explanatory power of public recognition as driver for performance of rewarded people, but rather reveal the effect of learning.

4.2.2 Effect on not rewarded subjects

The second hypothesis suggests that public recognition for performance in stage one has a positive effect on performance in stage two of people who were not rewarded. Also, in this case, just the not rewarded subjects' performance changes from the first to the second stage are taken into account. Figure 12 presents a boxplot with the performance in the first and second stage of subjects who were not rewarded after the first stage. In this case, the number of subjects is $N = 50$.

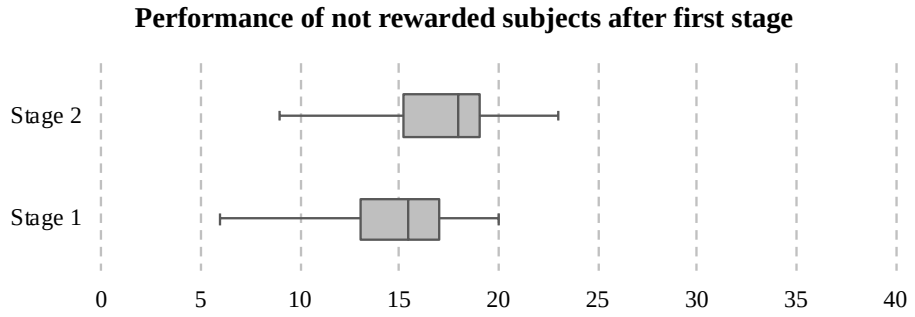


Figure 12: Boxplot with performance of not rewarded subjects after first stage

Similar to the first hypothesis, a normal distribution is neither given in this case, as Figure 13 reveals. Likewise, the Kolmogorov-Smirnov test suggests that the performance change,

$D(50) = 0.191$, $p < .001$, is significantly non-normal. Therefore, the Wilcoxon signed-rank test (Wilcoxon, 1945) is used to test this hypothesis as well.

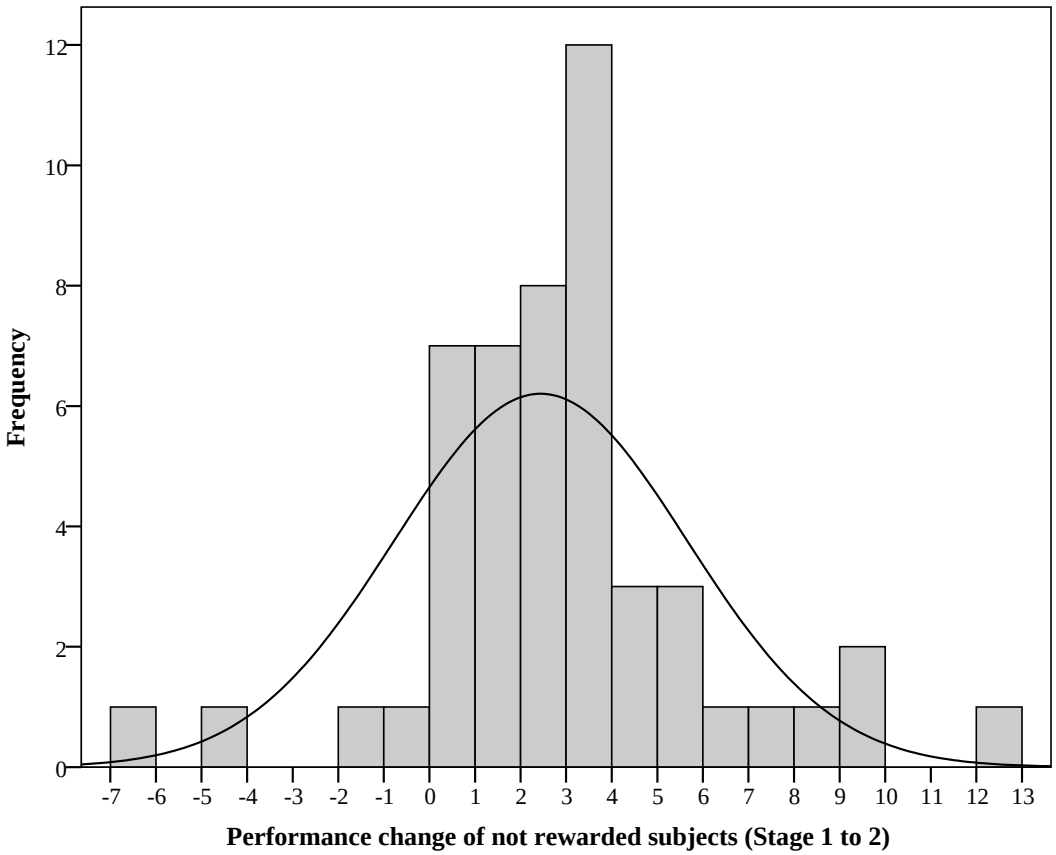


Figure 13: Histogram of not rewarded subjects' performance changes (Stage 1 to 2)

The results suggest that for not rewarded subjects, the performance was significantly higher in the second stage ($Mdn = 18$) than in the first stage ($Mdn = 15.5$), $T = 855.5$, $z = -4.64$, $p < .001$, $r = -.66$. This means that the null hypothesis can be rejected and that hypothesis H_2 can be accepted so far. Subsequently, the value of the effect size estimate r suggests a large positive effect on performance for subjects who were not rewarded (Cohen, 1992).

Just like for hypothesis H_1, it also has to be examined here whether the performance increase is driven by learning. Therefore, just the not rewarded subjects of the treatment group are included as well as those subjects of the control group that would not have been among the top five within their session. The results of the Mann-Whitney *U* test (Mann & Whitney, 1947) suggest that the performance increase was insignificantly higher for the treatment group ($Mdn = 2$) than for the control group ($Mdn = 1.5$), $U = 393.5$, $z = -1.4$, $p = .161$, $r = -.17$. These results are not as strong as for the rewarded subjects, but public recognition as a driver of performance is doubtful in this case as well.

4.2.3 Comparison of rewarded and not rewarded subjects

The third hypothesis assumes a higher increase in performance for subjects who did not receive a reward in the first stage compared to subjects who received a reward. The boxplot in Figure 14 shows the performance change from the first to the second stage, distinguished between rewarded and not rewarded subjects. Subsequently, the number of subjects in the rewarded group is $N = 20$ and in the not rewarded group is $N = 50$.

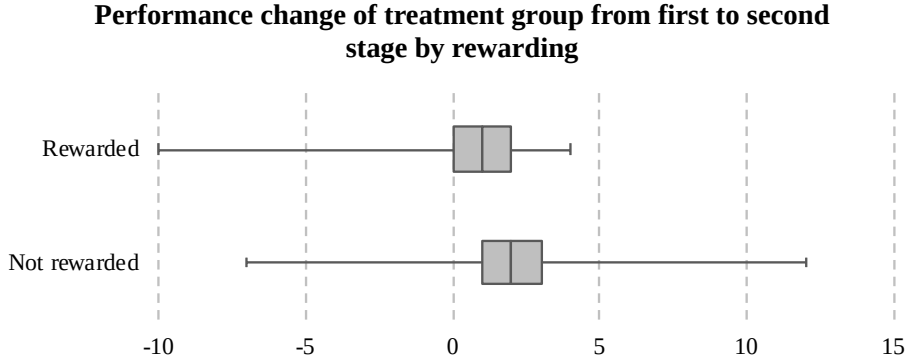


Figure 14: Boxplot with performance change from first to second stage

To test this hypothesis, the differences between subjects have to be measured. Therefore, a test for independent and not related samples is required. Since the histograms (Figure 11 & Figure 13) and the corresponding Kolmogorov-Smirnov tests did not reveal a normal distribution for either one of the groups, the Mann-Whitney *U* test (Mann & Whitney, 1947), which is equivalent to the Wilcoxon signed-rank test, is used. The results of this test suggest that for not rewarded subjects, the performance increase was significantly higher ($Mdn = 2$, $M = 2.44$) than for rewarded subjects ($Mdn = 1$, $M = 0.5$), $U = 308$, $z = -2.526$, $p = .012$, $r = -.30$. As a result, the null hypothesis can be rejected and hypothesis H_3 can be accepted. According to Cohen (1992), the effect size is at a medium level.

Also, in this case, the same test was done for the control group, assuming that there would have been a reward for the top five. Results of the Mann-Whitney *U* test (Mann & Whitney, 1947) show that the difference in performance increase was insignificant between "rewarded" ($Mdn = 2$, $M = 1.3$) and "not rewarded" subjects ($Mdn = 1.5$, $M = 1.65$) in the control group, $U = 99$, $z = -.46$, $p = .983$, $r = -.08$. Therefore, it can be assumed that the difference in performance change between rewarded and not rewarded subjects is really driven by public recognition and not the learning effect. Hypothesis H_3 is herewith accepted.

4.2.4 Remaining effects

The fourth and last hypothesis suggests a remaining effect on performance when subjects stay in the same treatment group. However, the effect size is assumed to be smaller. Recalling the results from H_1 and H_2, the performance change was significant for rewarded subjects and highly significant for subjects that were not rewarded, however, the driver was not public recognition. To test this hypothesis, the performance changes from stage two to three and three to four are taken into account. Furthermore, to test the effect on performance from stage two to three for rewarded subjects after stage two, just those subjects are included that received a reward after the first stage as well, and for not rewarded subjects vice versa. For the performance change from stage three to four, subjects are included if they were rewarded or not rewarded after the previous stages. Each of these groups of subjects was tested with the Wilcoxon signed-rank test (Wilcoxon, 1945).

The results suggest that for (1) rewarded subjects, the performance was significantly higher in the third stage ($Mdn = 26$) than in the second stage ($Mdn = 24$), $T = 66$, $z = -2.126$, $p = .033$, $r = -.57$; (2) rewarded subjects, the performance was insignificantly lower in the fourth stage ($Mdn = 25$) than in the third stage ($Mdn = 26$), $T = 22$, $z = -.998$, $p = .318$, $r = -.28$; (3) not rewarded subjects, the performance was significantly higher in the third stage ($Mdn = 18$) than in the second stage ($Mdn = 17$), $T = 541$, $z = -1.795$, $p = .073$, $r = -.27$; and (4) not rewarded subjects, the performance was significantly higher in the fourth stage ($Mdn = 18$) than in the third stage ($Mdn = 18$), $T = 303$, $z = -1.895$, $p = .061$, $r = -.29$. The results are presented in Table 4.

	Rewarded		Not rewarded	
	N	Mean	*N*	Mean
Stage 1 to 2	20	0.50	50	2.44***
Stage 2 to 3	14	1.50**	44	0.73*
Stage 3 to 4	13	-0.46	43	0.51*

This table reports Number of Subjects and Means for the performance changes distinguished between rewarded and not rewarded subjects. Note, that for performance change from stage 2 to 3, just those subjects are included that also got rewarded for stage 1, and that for performance change from stage 3 to 4, just those subjects are included that also got rewarded for stage 1 and 2.
***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 4: Means and effects sizes of performance changes

For rewarded subjects the hypothesis cannot be accepted, since the change in performance was not significant from stage one to two and three to four. However, there was a high increase in performance from stage two to three. Nevertheless, for subjects who were not rewarded, the

null hypothesis can be rejected for each stage and an increase in performance was observed in each stage. As expected, the increase in performance in terms of effect size was constantly decreasing, resulting in accepting hypothesis H_4 partly and namely for not rewarded subjects.

A multiple linear regression is run to examine effects on subjects' performance changes in different stages that are unrelated to the treatment. Since there are three dependent variables that measure a performance change – from stage one to two, from stage two to three, and from stage three to four – a regression is run for each of them. The following equation represents the basic model:

$$y = \beta_0 + \beta_1 r_1 + \beta_2 r_2 + \beta_3 r_3 + \varepsilon$$

The variable y is the observed performance change for a constant term coefficient β_0. The predictor r_1 represents the group, where $r_1 = 1$ if the subject got the treatment and $r_1 = 0$ if the subject was assigned to the control group, and its coefficient β_1, thus the effect of the treatment. Predictor r_2, which is the age of a subject, and r_3, where $r_3 = 1$ if the subject is female and $r_3 = 0$ if the subject is male, and their coefficients β_2 and β_3, respectively, extend the model by the standard control variables. The residual ε includes remaining, unexplained effects of the model. Table 5 shows the results for this model for each performance change in (1), (2), and (4).

	Performance change from stage 1 to 2	Performance change from stage 2 to 3		Performance change from stage 3 to 4
	(1)	(2)	(3)	(4)
Treatment	0.327	1.426***	0.577	0.172
	(0.638)	(0.510)	(0.564)	(0.418)
Age	0.038	-0.024	-0.024	-0.022
	(0.055)	(0.044)	(0.043)	(0.036)
Female	0.269	-0.078	-0.279	-0.098
	(0.604)	(0.483)	(0.470)	(0.396)
Session (Wed 11:30)			1.913***	
			(0.650)	
Session (Thu 13:00)			1.354**	
			(0.674)	
Constant	0.435	0.307	0.364	1.013
	(1.608)	(1.285)	(1.253)	(1.054)
Observations	100	100	100	100
R-squared	0.009	0.077	0.164	0.005

This table reports ordinary least squares (OLS) coefficient estimates. Standard errors are reported in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5: Regression effects on performance changes in respect to stages

Since faculty of a subject and session time were also used as control variables, they are introduced to the model as needed, meaning that forward selection, which is a stepwise method, was used to include session times and faculties into the model if they contribute explaining the performance change. Those variables were converted to dichotomous variables, since they were scaled from 1 to 6, originally. The probability of F to add a variable to the equation was chosen to be 0.05. However, session times were just conducive to the performance change from stage two to three and are shown in model (3). The two session times included are Wednesday, 11:30 and Thursday, 13:00.

The regression reveals several interesting findings, however the overall outcome is that neither one of the models explains any of the performance changes very well. The highest change in performance is explained by the regression model, which refers to the change from stage two to three and includes the two significant session times (4), with 16.4%.

All the regression models reveal that the predictors age and gender have almost no effect on performance change in the corresponding stages. Nevertheless, the treatment is a significant predictor at the 1% level from stage two to three (2), resulting in a significant regression equation ($F(3, 96) = 2.660,\ p = .053$) at the 10% level with an R^2 of $.077$. Although, the treatment is not a significant predictor for performance changes from stage one to two (1) and three to four (4), a slight effect of being in the treatment group was still observed.

Regression model (3) shows that two of the session times Wednesday, 11:30 and Thursday, 13:00 are significant predictors at the 1% and 5% level, respectively. As already exposed in Figure 9, the performance change from stage two to three was mainly driven by subjects who participated either in the session on Wednesday, 11:30 or Thursday, 13:00. Both of them were treatment groups. This indicates a deviation from a normal distribution of the subjects across the different sessions. However, both control groups decreased in performance and facilitated the two treatment sessions to be highly contributive and significant predictors. This results in a significant regression equation ($F(5, 94) = 3.688,\ p = .004$) at the 1% level with an R^2 of $.164$.

As a result, accepting hypothesis H_4 partly has to be qualified, since including the control group diminishes the explanatory power of public recognition in this setting. Since public recognition as driver for performance change has to be questioned, a 2 (treatment) × 3 (stage) ANOVA was conducted on subjects' performance change, where treatment is a dummy variable, coded with zero for the control group and one for the treatment group, and stage is used as an ordinally

scaled time variable, to examine if the stage is related to performance changes in respect to the group. Table 6 reveals the treatment ($F(1, 294) = 4.461,\ p = .036$) as well as the stage ($F(2, 294) = 8.276,\ p < .001$) as drivers for performance change. There is no significant effect of the two-way interaction between treatment and stage on subjects' performance change ($F(2, 294) = 1.650,\ p = .194$). The results of this *F*-test suggest that both the learning effect in terms of stage and thus time and public recognition in terms of treatment have a significant main effect on performance change. Although, stage negatively affected the performance change, the absolute performance still increased, meaning that the increase in performance was shrinking every stage and that the learning effect got smaller from stage to stage.

	Performance change		
	Degrees of freedom	*F*-value	*p*-value
Treatment	1	4.461	0.036
Stage	2	8.276	0.000
Treatment × Stage	2	1.650	0.194

Table 6: Univariate analysis of variance for treatment and stage on performance change

To further examine these results, a multiple linear regression was run on performance change and is reported in Table 7. The first model includes only the treatment group and determines the effect of being among the top five, the stage and demographics, and results in a significant regression equation ($F(4, 205) = 3.483,\ p = .009$) at the 1% level with an R^2 of $.064$. The results confirm the *F*-test, since the stage influences performance changes negatively, meaning that the effect of learning is reduced from time to time. Also, being among the top five and receiving a reward decreases the performance change significantly. However, one has to consider that in this model the control group is not included and that the subjects who did not receive a reward were identified to be the driver for performance change. Therefore, the second model includes both groups and controls for treatment. That model results in a significant regression equation ($F(5, 294) = 4.159,\ p = .001$) at the 1% level with an R^2 of $.066$. Here, treatment affects performance changes positively, suggesting that public recognition has a significant effect on performance change. In this case, the variable top five indicates which subjects got a reward (treatment group) or would have gotten a reward (control group), but is not a significant predictor to the model. The stage variable is also a significant predictor affecting performance change negatively, in this model. Nevertheless, the model does not explain much of the overall performance change and has not the explanatory power to confirm hypothesis H_4.

	Performance change (only treatment group)	Performance change
Treatment		0.620**
		(0.306)
Top five	-0.914**	-0.494
	(0.413)	(0.307)
Stage	-0.664***	-0.635***
	(0.224)	(0.171)
Age	-0.012	-0.005
	(0.034)	(0.027)
Female	-0.009	0.068
	(0.384)	(0.290)
Constant	3.082***	2.061**
	(1.070)	(0.852)
Observations	210	300
R-squared	0.064	0.066

This table reports ordinary least squares (OLS) coefficient estimates. Standard errors are reported in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 7: Regression effects on performance change

5 Discussion

The goal of this thesis was that it could be transferable to a real life setting where employees are willing to reciprocate by performing better and doing the job well if the leader is treating his employees well and recognizing them. In the experiment, some people in the room were publicly recognized by coming to the front of the class to shake hands with the experimenter and receive a certificate of achievement, with everyone in the room being able to see that. This treatment was supposed to work as a symbolic reward, which should lead to reciprocal behavior. Two other mechanisms, namely conformity and altruism, are also expected to explain the changes on performance.

The main idea with hypotheses one and two was to check if public recognition affects performance at all. The performance change for rewarded and not rewarded subjects from the first to the second stage was studied. For both groups the performance increased, and the change seemed to be significant in the first place. Other studies, for instance Kosfeld and Neckermann (2011), Bradler et al. (2016) and Asraf et al. (2014), have also shown similar results. However, further tests that included the control group, which did not receive any treatment, the learning effect seemed to have a big influence on performance improvements from stage one to two for both the rewarded and the not rewarded group. Therefore, hypothesis one and two could not be confirmed, since it could not be proven that the performance increase was driven by public recognition.

Hypothesis three tested whether the effect on performance was bigger for not rewarded subjects or for rewarded ones. The assumption of a greater effect on performance for not rewarded subjects could be proven. This result confirms the findings of Bradler et al. (2016) that the performance increase was mostly due to the subjects who did not receive the recognition and reward. Conformity seems to be an important mechanism here. If subjects have a preference for conformity and they want to follow the group norms, the performance should increase for those who were not rewarded. Altruism could partly explain this as well; the subjects perceived the kindness of the experimenter, since she recognized and rewarded the best performers without their prior knowledge. Therefore, the subjects who did not get the recognition improved even more.

The fourth and last hypothesis should test whether the effect on performance remains when public recognition is provided repeatedly. It was assumed that the effect remains, but that the effect size gets smaller. This hypothesis could be confirmed for not rewarded subjects, but not

for rewarded ones. For the rewarded subjects, the performance change decreased from stage three to four. This may be explained with conformity, since the rewarded subjects wanted to match the rest of the group and there was no need to increase the performance even further. However, the performance change was not significant. Not rewarded subjects increased their performance in every stage and the effect got smaller as hypothesized. Conformity is also working here, since not rewarded subjects tried to match their performance to the group norm. Unfortunately, the hypothesis on the effect of repeated recognition was only partly confirmed and there were still some doubts if public recognition was the driver for not rewarded subjects. More studies in this field would help clarifying the effect of repeated public recognition.

It seems that the biggest driver of the increase in performance in this study was unfortunately the learning effect and not public recognition, as I hypothesized. The limitations of the study will be discussed further in the next subchapter.

5.1 Limitations

This paper contains several limitations that should be considered when analyzing the results. Most of the subjects got better from one stage to the next one, but according to the regression model, little of that was explained by the recognition. A part of the performance increase was explained by the learning effect. I tried to minimize the learning effect by changing the decoding list in each stage, so that subjects could not memorize it. To minimize the learning effect further, it would have been even better to change the decoding list after every task that was decoded (Benndorf, Rau & Sölch, 2014).

Even though the decoding task that was chosen has been used before (e.g. Clark & Friesen, 2009; Charness et al., 2014; Erkal et al., 2011; Gjedrem, 2016)**,** it seems like it was not the best task for this study. The task was chosen, since it does not require any prior knowledge and subjects should have equal knowledge and skills when they work on the task. However, the task required typing skills and some subjects may be faster at typing on the keyboard than others, which gives those subjects better chances to master the task. Some subjects could also learn better and faster techniques to type during the experiment. Another option would have been to include a variable in the questionnaire, asking subjects about their self-estimation in terms of typing skills or how much time they spend on a computer per day or week.

According to the induced valuation theory, people want to maximize their utility (Smith, 1976). There are always costs involved and for some subjects, it may have been best to not perform that well in the first stage, since the reward was not introduced, yet. But when they got to know

about the reward, their performance might have improved, since benefits got bigger than costs. However, not everyone's performance might have been driven by the reward, because it might not be interesting to some and therefore, the reward does not have any benefit for them. That said, some of the subjects may have chosen to not try to increase their performance, because the reward was not interesting.

I believe that the sample size in the experiment was big enough, but the group of rewarded people was not that big. It could have been better for the analysis to reward a few more subjects in each treatment group. I should also consider if the symbolic reward, the certificate of achievement, fit in the laboratory experiment. Most likely it fits better in a field experiment where the subjects do not know that they are taking part in an experiment and thus, the certificate would mean more to them, since they get the feeling that their work is appreciated. In the laboratory, the subjects know that they are taking part in an experiment and the certificate from the experimenter may not be that rewarding.

The experimental subjects' morale being higher because of them knowing that they are taking part in an experiment is called the Hawthorne effect, which might be present in this experiment as well (Zikmund et al., 2013, p. 264). Therefore, the subjects might have performed differently than they would have if they had been performing the same task in a non-experimental setting. According to Chiesa and Hobbs (2008), researchers should be careful about using the term Hawthorne effect, since it often seems to refer to normal characteristics of research, which includes variables affecting the results that the experimenter is not aware of.

5.2 Further research

It seems that it would be simpler to investigate if public recognition has an effect on performance by conducting a field experiment instead of a laboratory experiment. In that case, the setting would involve a real work experience where recognition from the leader could have a more realistic effect on the participants' performance. Conducting a field experiment requires more time, resources and planning, that I had at disposal.

For further research, it would also be interesting to investigate the limit of how much a person can improve while working on such a simple decoding task. However, the tasks in real life are usually more complex and therefore, there might be room for more improvement in more complicated tasks.

6 Conclusion

This paper investigated the effect of public recognition on performance in competitive environments. I conducted a real effort experiment to examine if public recognition influences the performance of the subjects. I tried to confirm previous research and extend it by looking into the long-term effects of public recognition.

I was not able to confirm previous research in terms of public recognition having a positive effect on performance of rewarded and not rewarded subjects. However, the effect on performance was stronger for subjects who were not rewarded than for the rewarded ones. This confirms the findings of Bradler et al. (2016). Conformity is the mechanism which makes the subjects want to adjust their performance according to group norms. It could not be confirmed whether the effect of public recognition remains when recognition is provided repeatedly. Not rewarded subjects increased their performance in each stage, however, taking the control group into account diminishes the explanatory power of public recognition as a driver for performance.

There are several reasons why I was not able to find the causal relationship of public recognition and performance. The experimental design, meaning the task of decoding a list of letters, and learning effects, since the task might have been to simple, were the main weaknesses of this research. Even though, I was not able to confirm the effect of public recognition on performance, I believe that this effect still exists, and it should be studied even further. Therefore, the results may be used as management implications to adapt and improve performance related rewards. Optimally, public recognition can be used consciously as a reward to motivate both employees who receive the reward, but also the ones who did not receive one.

References

Appelbaum, S. H., & Kamal, R. (2000). An analysis of the utilization and effectiveness of non-financial incentives in small business. *Journal of Management Development, 19*(9), 733-763.

Ariely, D., & Norton, M. I. (2007). Psychology and experimental economics: A gap in abstraction. *Current Directions in Psychological Science, 16*(6), 336-339.

Benndorf, V., Rau, H., & Sölch, C. (2014). Minimizing learning behavior in experiments with repeated real-effort tasks. Retrieved June 3, 2018, from http://dx.doi.org/10.2139/ssrn.2503029

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior, 10*(1), 122-142.

Bernheim, B. D. (1994). A theory of conformity. *Journal of Political Economy, 102*(5), 841-877.

Bradler, C., Dur, R., Neckermann, S., & Non, A. (2016). Employee recognition and performance: A field experiment. *Management Science, 62*(11), 3085-3099.

Brun, J. P., & Dugas, N. (2008). An analysis of employee recognition: Perspectives on human resources practices. *The International Journal of Human Resource Management, 19*(4), 716-730.

Bryman, A. (2012). *Social Research Methods* (4th ed.). Oxford, Great Britain: Oxford University Press.

Brüggen, A., & Strobel, M. (2007). Real effort versus chosen effort in experiments. *Economic Letters, 96*(2), 232-236.

Burgess, S., & Ratto, M. (2003). The role of incentives in the public sector: Issues and evidence. *Oxford Review of Economic Policy, 19*(2), 285-300.

Charness, G., Masclet, D., & Villeval, M. C. (2014). The dark side of competition for status. *Management Science, 60*(1), 38-55.

Chen, Y., Harper, F. M., Konstan, J., & Li, S. X. (2010). Social comparisons and contributions to online communities: A field experiment on movielens. *American Economic Review, 100*(4), 1358-98.

Chiesa, M., & Hobbs, S. (2008). Making sense of social research: How useful is the Hawthorne Effect? *European Journal of Social Psychology, 38*(1), 67-74.

Clark, J., & Friesen, L. (2009). Overconfidence in forecasts of own performance: An experimental study. *The Economic Journal, 119*(534), 229-251.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159.

Dutcher, G. E., Salmon, C. T., & Saral, K. J. (2015). Is 'Real' Effort More Real? Retrieved April 20, 2018, from http://dx.doi.org/10.2139/ssrn.2701793

Erkal, N., Gangadharan, L., & Koh, B. H. (2018). Monetary and non-monetary incentives in real-effort tournaments. *European Economic Review, 101*, 528-545.

Erkal, N., Gangadharan, L., & Nikiforakis, N. (2011). Relative earnings and giving in a real-effort experiment. *American Economic Review, 101*(7), 3330-3348.

Falk, A., & Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior, 54*(2), 293-315.

Falk, A., & Heckman, J. J. (2009). Lab experiments are a major source of knowledge in the social sciences. *Science, 326*(5952), 535-538.

Field, A. P. (2013). *Discovering statistics using IBM SPSS stat*istics (4th ed.). Los Angeles, CA: Sage.

Fischbacher, U. (2007). z-Tree: Zurich Toolbox for Ready-made Economic Experiments. *Experimental Economics*, *10*(2), 171-178.

Fisher, J. G. (2015). *Strategic Reward and Recognition: Improving employee performance through non-monetary incentives*. London, Great Britain: Kogan Page.

Gjedrem, W. G. (2016). *Feedback and Performance: Experiments in Behavioral Economics* (PhD thesis UiS, no. 321). Stavanger: University of Stavanger.

Gneezy, U., & Rustichini, A. (2000). Pay enough or don't pay at all. *The Quarterly Journal of Economics*, 115(3), 791-810.

Hollander, E. P. (1958). Conformity, status, and idiosyncrasy credit. *Psychological Review, 65*(2), 117.

Kosfeld, M., & Neckermann, S. (2011). Getting more work for nothing? Symbolic awards and worker performance. *American Economic Journal: Microeconomics, 3*(3), 86-99.

Kube, S., Maréchal, M. A., & Puppe, C. (2012). The currency of reciprocity: Gift exchange in the workplace. *The American Economic Review, 102*(4), 1644-1662.

Kvaløy, O., Nieken, P., & Schöttner, A. (2015). Hidden benefits of reward: A field experiment on motivation and monetary incentives. *European Economic Review, 76*, 188-199.

Lakshman, M., Sinha, L., Biswas, M., Charles, M., & Arora, N. K. (2000). Quantitative vs qualitative research methods. *The Indian Journal of Pediatrics, 67*(5), 369-377.

Lancaster University Management School (June 26, 2016). Qualitative and quantitative research. Retrieved May 4, 2018, from https://www.lancaster.ac.uk/media/lancaster-university/content-assets/documents/learning-skills/quantitativevqualitativeresearchanswers.pdf

Lawler, E. E. (1969). 3. Job design and employee motivation. *Personnel Psychology, 22*(4), 426-435.

Lazear, E. P. (2000). Performance pay and productivity. *American Economic Review, 90*(5), 1346-1361.

Lazear, E. P., & Gibbs, M. (2014). *Personnel economics in practice* (3rd ed.). New York, NY: John Wiley & Sons.

Levine, D. K. (1998). Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics, 1*(3), 593-622.

Luthans, K. (2000). Recognition: A powerful, but often overlooked, leadership tool to improve employee performance. *Journal of Leadership Studies, 7*(1), 31-39.

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics, 18*(1), 50-60.

Manolopoulos, D. (2007). An evaluation of employee motivation in the extended public sector in Greece. *Employee Relations, 30*(1), 63-85.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review, 50*(4), 370-396.

Mathauer, I., & Imhoff, I. (2006). Health worker motivation in Africa: the role of non-financial incentives and human resource management tools. *Human resources for health, 4*(1), 1-17.

Mikkelsen, A. & Laudal, T. (ed.) (2016). *Strategisk HRM: 2: HMS, etikk og internasjonale perspektiver*. Oslo, Norway: Cappelen Damm akademisk.

Porter, L., & Lawler, E. E. (1968). *Managerial attitudes and performance*. Homewood, IL: Dorsey-Irwin.

Reiss, S. (2012). Intrinsic and extrinsic motivation. *Teaching of Psychology, 39*(2), 152-156.

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (2nd ed.). Newbury Park, CA: Sage.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54-67.

Smith, A., & Hanley, R. P. (2009). *The theory of moral sentiments*. New York, NY: Penguin Books.

Smith, V. L. (1976). Experimental economics: Induced value theory. *The American Economic Review, 66*(2), 274-279.

Sonawane, P. (2008). Non-monetary rewards: Employee choices & organizational practices. *Indian Journal of Industrial Relations, 44*(2), 256-271.

Stajkovic, A. D., & Luthans, F. (2003). Behavioral management and task performance in organizations: conceptual background, meta-analysis, and test of alternative models. *Personnel Psychology, 56*(1), 155-194.

Sussman, R. W., & Cloninger, C. R. (2011). *Origins of altruism and cooperation*. New York, NY: Springer.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology, 46*(1), 35-57.

Van Herpen, M., Van Praag, M., & Cools, K. (2005). The effects of performance measurement and compensation on motivation: An empirical study. *De Economist, 153*(3), 303-329.

Waqas, Z., & Saleem, S. (2014). The effect of monetary and non-monetary rewards on employee engagement and firm performance. *European Journal of Business and Management, 6*(31), 73-83.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin, 1*(6), 80-83.

Zeb, A., Jamal, W., & Ali, M. (2015). Reward and Recognition Priorities of Public Sector Universities' Teachers for their Motivation and Job Satisfaction. *Journal of Managerial Sciences, 9*(2), 214-224.

Zikmund, W. G., Babin, B. J., Carr, J. C., & Griffin, M. (2013). *Business Research Methods* (9th ed.). Mason, OH: Cengage learning.

Zobal, C. (1999). The "ideal" team compensation system – an overview part II. *Team Performance Management, 5*, 25-45.

Appendix

Appendix A Certificate of Achievement

Figure 15: Certificate of Achievement

Appendix B Instructions given by the experimenter during the experiment

Welcome and thank you for showing up for my experiment!

If you have any questions during the experiment, please raise your hand and I will come to you. Next to your computer screen, there is a sheet of instructions. Please read it carefully. Please decode the letters shown on the screen, not the letters shown on instructions. That is just an example.

Is someone not done reading the instructions? Do you have any questions?

Once the experiment starts, there will be a screen with tips and a trial stage before the actual task starts. We will start now.

Treatment group after stage 1: Those of you who are among the top five, meaning the people who had the highest amount of correct answers, can come to the front to receive their reward. The numbers are attached to your computer screen. These five people are:…

Treatment group after stage 2&3: The five people with highest amount of correct answers this time are:... Please come to the front.

Now we are about to start the questionnaire, while you fill it out, I will hand out the payment forms, please fill them out when you are done with the questionnaire.

About the payment form:

- ID-number/D-number is required, if you do not have either one, please write your date of birth
- Address: Norwegian address
- Tax municipality: If you do not pay taxes in Norway, please write “not available”. If you pay taxes in Norway, please write your tax municipality. Do not worry, you do not have to pay any taxes for this experiment, it is just for accounting purposes.

Appendix C Invitation e-mail

Dear student,

I would like to invite you to participate in an experiment conducted by a master student at the Business School of UiS.

The experiment is done on computers and the task for the experiment is very simple, no prior knowledge is required. The duration of the experiment is approximately 30 minutes and a fixed payment of kr 50 is given to all the participants after the experiment. The information gathered during the experiment will be confidential.

The experiment will be done in English and takes place in Ellen og Axel Lunds hus (EAL), hotel school building. The room is H-209, the computer lab on the second floor.

Dates: **Wednesday, March 7th at 10:00, 11:30 and 13:00**

Thursday, March 8th at 10:00, 11:30 and 13:00

To be able to participate, you can register via the following link at a time that fits you the best. There are limited number of places, and only the times available will be shown on the link. You can only participate once.

You will not get a confirmation e-mail, so please write down the time you signed up for.

Sign up here: https://www.surveymonkey.de/r/7WWCPL7

If you have any questions or if you have to cancel the appointment, please send an e-mail to: mk.lehtinen@stud.uis.no

Best regards,

Maria Lehtinen

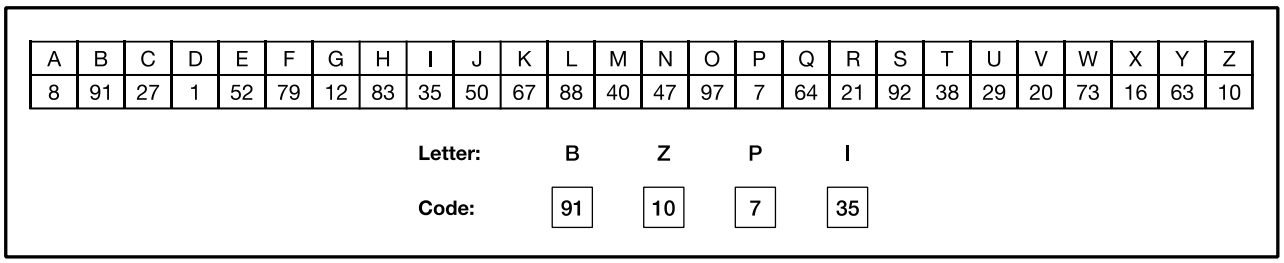
Appendix D Experimental instructions

Welcome to the experiment

In this experiment you will be working on a task of decoding letters to numbers.

Task description

You will be given a list of letters and all the letters have a corresponding number. Your task is to decode the given sequences of four letters into numbers.



A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
8	91	27	1	52	79	12	83	35	50	67	88	40	47	97	7	64	21	92	38	29	20	73	16	63	10

Stages and breaks

There are several stages with breaks in between, each stage lasts for 4 minutes. There are unlimited number of tasks on each stage. You will be given instructions on the computer screen during the experiment. When you have completed all the stages, there will be a questionnaire in the end.

Payment

You will be given a fixed payment of kr 50 in cash at the end of the experiment.

Rules

Please remain in your seat through the whole experiment. You can choose how to spend your time during the experiment, however, communicating with the other participants is not allowed. You can leave the experiment at any time, notice that your experiment will be suspended, and it is not possible to continue after leaving the room. It is not allowed to use the computer on anything else than the experiment as it may cause technical problems.

Thank you for participating!