Spring 2015

University of Stavanger

The Impact of Feedback and Incentive Schemes on Performance

A Laboratory Experiment

Master's Thesis in Business Administration

Alexandra M. Bruun-Olsen

Camilla Engelsen Eian

Universitetet
i Stavanger

FACULTY OF SOCIAL SCIENCES,
UIS BUSINESS SCHOOL

MASTER'S THESIS

STUDY PROGRAM:	THESIS IS WRITTEN IN THE FOLLOWING SPECIALIZATION/SUBJECT:
Master in Business Administration	Economics IS THE ASSIGNMENT CONFIDENTIAL? No

TITLE:

The Impact of Feedback and Incentive Schemes on Performance

AUTHORS		ADVISOR:
Student number:	Name:	Ola Kvaløy
211890	Alexandra M. Bruun-Olsen	
223381	Camilla Engelsen Eian	

ACKNOWLEDGE RECEIPT OF 2 BOUND COPIES OF THESIS

Stavanger, ……/…… 2015 Signature administration:……

Abstract

In this thesis we analyze how different feedback (objective and subjective), incentive schemes (fixed pay and performance pay), and the interaction between them impact performance. To answer our research question we conduct a lab experiment on students from the University of Stavanger. The experiment consists of subjects performing a combined coloring- and calculation task over three periods and receiving either objective or subjective feedback after each period. To conduct our analysis we employ a 2 (feedback; objective or subjective) x 2 (incentive scheme; fixed pay or performance pay) between subjects design.

The analysis reveals one main effect; objective feedback has a greater impact on performance than subjective feedback. The effect is explained mainly by two components: (1) the quantitative nature of the task causes subjects to experience objective feedback as more understandable compared to subjective feedback and; (2) negative subjective feedback has a more damaging psychological effect compared to negative objective feedback. Our analysis also displays a trend in which subjects assigned performance pay perform better compared to subjects assigned fixed pay. However, the effect is not statistically significant. Moreover, we find that there exists a certain complementarity between objective feedback and performance pay but an interaction effect cannot be determined.

Preface

This paper is written as the final part of our master's degree program in Business Administration at the University of Stavanger.

First and foremost we would like to thank our thesis advisor professor Ola Kvaløy for engagement, quality discussions, and motivational quality advising.

We would also like to thank the University of Stavanger and the Norwegian Research Council for financially supporting our research and experiment.

In addition, a great thanks is directed to Econa for their financial support, and especially Econa's analyst and advisor Atle Kolbeinstveit for always being available to answer any questions or advise.

Furthermore, we would thank William Gilje Gjedrem, who took the time to give us invaluable advice on dealing with data analysis and the various data tools.

Finally, thanks to family and friends for contribution of valuable input.

Alexandra Marie Bruun-Olsen & Camilla Engelsen Eian

Stavanger, June 2015

Table of Contents

List of figures

List of tables

1. Introduction

In our increasingly globalized society, organizations are faced with rising levels of competition forcing them to become more cost efficient in order to survive. Herbert Hoover once said, "Competition is not only the basis of protection to the consumer, but is the incentive to progress". This has proven to be particularly true by modern day organizations progressing to put considerable emphasis on behavioral aspects and performance enhancing measures as an approach to increase efficiency. Thus, the importance of research on employee behavior and measures to modify it rises accordingly.

Performance feedback is one of the more popular interventions of Organizational Behavior Management (OBM) strategies and has long been recognized as essential for learning and motivation (Ilgen, Fisher, & Taylor, 1979). The strategy has the advantages of low economic costs, simplicity and flexibility, which allows a wide range of organizations to implement it regardless of size and scale (Prue & Fairbank, 1981). Although there exists empirical evidence that feedback improves performance (e.g. Azmat & Irriberri, 2010) it is also shown that effects are not necessarily consistent (Balcazar, 1985). This raises the question of how feedback should be utilized to generate the most positive and consistent effects.

The nature of feedback is generally based on the performance measure used to evaluate the worker. Quantitative performance measures constitute grounds for objective feedback and typically lack subjectivity. On the other hand, feedback based on qualitative performance measures require a certain level of subjectivity and is categorized accordingly. In addition, accounting for individuals' social preference like the desire to reciprocate behavior suggests that objective and subjective feedback should impact the psychological aspect of motivation differently (Fehr & Falk, 2001). To our knowledge, there exists no research comparing the effects of objective and subjective feedback. The motivation for our study draws on this shortage and our main contribution is a direct comparison of the effects of objective and subjective feedback on performance. Moreover, we assert that provision of empirical evidence on this area could contribute to increase the understanding of how to facilitate performance feedback programs and generally improve utilization.

In the context of a classic principal-agent model, employees work as agents, with the firm or a manager as their principle. Although the agents are compensated for their work the relationship may encounter certain incentive problems. Moral hazard is of particular risk and occurs if interests between the principal and the agent are not aligned. As such, firms are faced with the challenge of optimally structuring incentives in a way that motivates employees, reduces moral hazard and allows the firm to derive maximal value out of each employee. As a method of aligning interests, the principal can provide the agent with incentives that prompt extrinsic motivation or intrinsic motivation. Feedback is suggested to appeal to the agent's intrinsic motivation by supporting the agent's sense of competence and providing guidelines on how to improve performance.

The economic view is that monetary incentives such as rewards, bonuses and incentive schemes are efficient in aligning interests and improving performance. However, research has proven that effects are not always consistent. For example, Kvaløy et al. (2013) find that effect of the monetary incentive performance pay is only positive when it is accompanied by motivational talk. Amabile (1993) explains this as a motivational synergy, where extrinsic and intrinsic factors complement each other. This contradicts the psychological literature's view that extrinsic motivation works in opposition to intrinsic motivation and opens up for a line of research on incentives designed to simultaneously appeal to both extrinsic and intrinsic motivation. Balcazar et al. (1985) find that adding rewards to feedback improves consistency in its effects. Thus, we add monetary incentives (fixed pay and performance pay) to determine whether being combined with extrinsic factors can reinforce the effect of feedback. Our research question is worded as follows:

Does objective and subjective feedback generate different effects in performance, and are these effects reinforced when combined with different incentive schemes?

To collect the relevant data we conducted a laboratory experiment on students from the University in Stavanger. The experiment consisted of a combined coloring and calculation task and was conducted over three periods, each lasting 10 minutes. Subjects were informed about their assigned payment scheme at the beginning of the experiment, and after each period they were provided with either objective or subjective individual feedback on their work. Performance was equally evaluated in both the objective and subjective feedback treatment by grading after the number of completed

task sheets. The only difference was the in which feedback was expressed. Objective feedback was given by visibly counting the triangles and stating the evaluation in an objective matter. In providing subjective feedback, triangles were not visibly counted, but rather expressed as the experimenter's personal evaluation of their work. Ultimately, the feedback resulted in a grade on a scale from 1 to 6 (where 1 was the lowest and 6 was the highest possible grade), which was used as our performance variable.

Our research design consisted of a 2 x 2 between-subjects design: feedback type (objective or subjective); and incentive scheme (performance pay or fixed pay), which made a total of four treatment groups (FP-objective; FP-subjective; PP-objective; and PP-subjective). Treatments were randomly assigned over 8 sessions and thus, the session students chose to attend was decisive for which treatment group they would belong to. In the performance pay treatment pay was contingent on performance, while subjects in the fixed pay treatment received pay non-contingent on performance.

Our analysis rested primarily on the Mann Whitney U-test and regression analyses. The Mann-Whitney U test allowed us to compare groups without assuming a parametric distribution while the regression analysis was used to estimate effect sizes and examine the potential interaction effect. Average performance (mean grade) was used as our key dependent variable and consisted of an average grade deducted from period 2 and 3.

Our analysis revealed one main effect; Objective feedback had a greater impact on performance than subjective feedback. The effect was explained mainly by two components: (1) the quantitative nature of the task caused subjects to experience objective feedback as more understandable compared to subjective feedback and; (2) negative subjective feedback had a more damaging psychological effect compared to negative objective feedback. Moreover, subjects assigned performance pay had a higher average performance compared to subjects assigned fixed pay. However, this effect was not statistically significant. A comparison of average performance over treatments showed that the combination of objective feedback and performance pay yields the highest average performance, but an interaction effect could not be determined. Thus, although the results display certain tendencies, we cannot conclude that feedback and incentive scheme function as a motivational synergy.

This thesis consists of 6 chapters. Chapter 2 presents relevant theory and existing literature. This is used to highlight the essence of our research question and consists of three main topics: motivation, monetary incentives and feedback. The last section presents this thesis' contribution to literature. Chapter 5 presents the experimental design and procedure used to answer our research question. Moreover, results are reported in chapter 4 and discussed in chapter 5. Final conclusions are presented in chapter 6 along with suggestions for future research.

2. Theory and Existing Literature

2.1 Motivation

The understanding of what motivates workers is essential in the process of providing incentives. Generally, we distinguish between *intrinsic* and *extrinsic* motivation. Intrinsic motivation refers to doing something because it is inherently interesting or enjoyable, while extrinsic motivation refers to doing something because it is rewarded by incentives and leads to a separable outcome (Deci & Ryan, 2000). Research has shown that behavior and performance is very different when one is driven by intrinsic versus extrinsic reasons. Because extrinsically motivated behaviors are not inherently interesting, the economic view is that they must be externally prompted in some way. For example, a student with a lower GPA than her peers may feel extrinsically motivated to study more in order to achieve rewards in terms of higher grades and an empowering feeling. Applied to an organizational setting, extrinsic motivation is high when significant incentives are provided, these being either symbolic or monetary.

Intrinsic motivation is strongest when an opportunity to learn on the job is present. Ryan & Deci (2000) stated that behavior driven by intrinsic motivation is a critical element in cognitive, social, and physical development because of the knowledge and skill growth associated with acting on one's inherent interests. However, they further cite Skinner (1953) on his operant theory that states that all behaviors are motivated by rewards (i.e., by separable outcomes) and that the rewards for intrinsically motivated activities are in the activity itself. The relationship between intrinsic and extrinsic motivation is a complex one. Using a classic blood donation example, what would happen if an intrinsically motivated blood donor started to receive a payment for his blood? Experimental research on intrinsic motivation has rested primarily on the behavioral measure called the "free choice" measure and is used to uncover the effects of external rewards on intrinsic motivation (Deci & Ryan, 2000). Deci (1972) conducted a free-choice experiment with college students who were paid to perform an interesting SOMA Cube puzzle, and discovered that their reward for doing so made them less likely to perform the puzzle during a free-choice period. Lepper et al.'s (1973) study supported this result by finding that nursery school children who were offered a "good player award" for drawing (an activity they normally did and liked) were less likely to continue the activity

when back in their regular classrooms. Researchers argued this as an "overjustification effect" related to the self-perception theory, and suggested that intrinsic motivation would be discounted when sufficient extrinsic (rewards) and intrinsic (interest) factors were present (Sansone & Harackiewicz, 2000). Hence, the findings suggest that rewarding the blood donor for his blood would impair his intrinsic motivation and maybe even reduce the likelihood of future blood donations.

Although several researchers propose that extrinsic and intrinsic motivation are incompatible (e.g. Deci, 1971; Deci & Ryan, 1985; Deci & Ryan, 2000; Lepper & Greene, 1973), Amabile (1993) argues that you can avoid undermining intrinsic motivation by using extrinsic and intrinsic factors as complements. This constitutes what she calls a "motivational synergy" and explains the mechanism as "extrinsics in service of intrinsics". The reasoning for the mechanism is that as long as the extrinsic factor supports an individual's sense of competence without undermining the individual's sense of self-determination, it should positively impact intrinsic motivation. Hence, incentives should be designed to stimulate both extrinsic and intrinsic motivation to resolve the crowding-out problem. Applying this to an organizational setting, a firm's challenge is to design incentives that combine extrinsic and intrinsic motivational factors in a way that increases motivation and enhances performance.

Fehr and Falk (2001) argue that economists tend to ignore the psychological foundations of motivation in that they do not consider the individual's social preferences. A person is said to exhibit social preferences if the person does not only care about material rewards allocated to himself but also cares about the material rewards allocated to others. In other words, social preferences include non-pecuniary motives like the desire to reciprocate or the desire to avoid social disapproval. Moreover, Fehr and Falk (2001) suggest that such motives play a significant role in shaping human behavior and should therefore be considered in attempting to explain why economic incentives sometimes backfire. As a method of merging the psychological and economic view of motivation, this thesis attempts to consider both aspects by acknowledging that human beings are social beings in addition to rational (in an economic sense) individuals.

2.2 Monetary incentives

Traditionally, economic studies of personnel policies and workplace productivity have been concentrated on monetary incentive mechanisms. Monetary incentives serve the purpose of motivating employees to align their objectives in direction of the firm's. This does, however, require that the firm have knowledge about the incentive's effect on motivation and performance, as well as the necessary magnitude. This knowledge is not necessarily given, which may cause firms to encounter incentive problems. The basic framework for analyzing most incentive problems is called the principal-agent model. The principal-agent problem arises when an employee enters a mutually agreed contract with a principle (firm) that states the employee is working as an agent in the firm. Because the agent possesses better information about his abilities than the firm, the problem of asymmetric information arises. This presents a risk of moral hazard and conflict of interest and needs to be accounted in provision of incentives. Moral hazard is best explained by the problem that arises when imperfect contracts provide potential gain by acting contrary to the agreement. As a consequence, the agent may pursue his own agenda and ignore the best interest of the principle, which represents the essence of incentive problems (Laffont & Martimort, 2002).

The principal-agent conflict can be analyzed by modeling the objectives of both parties; the objective of the principal is to maximize firm value; and the agent's objective is to maximize own value (utility). The principle's challenge is therefore to provide sufficient incentives that maximize the agent's utility, further motivating him to exert the necessary effort to maximize firm value. One approach to mitigate the incentive problem is to provide the agent with incentives in terms of incentive schemes that sufficiently matches job design and the agent's effort and performance.

Generally, incentive schemes are distinguished between fixed pay and performance pay. Although there are several variations of the two, our study will mainly focus on *salary* as fixed pay and *piece-rate* as performance pay. The salary incentive scheme is a function of a worker's input (effort) in a current period and is often calculated as number of work hours. Piece-rate, on the other hand, is related to a worker's produced output. According to economic theory, the worker always provides the minimum possible effort under salary pay because additional effort does not accrue additional benefit. However, incentives are generated by termination contracts stating the worker will be fired if he does not perform according to the firm's pre set standard (Shearer, 2004). Under a piece-rate

incentive scheme, the worker has an incentive to increase output because it increases pay accordingly.

The firm's choice of incentive scheme is somewhat conditioned on its objective to maximize firm value. Lazear (1986) argued three main issues that affect a firm's choice of incentive scheme: sorting workers across jobs; inducing appropriate effort levels; and selecting quantity versus quality of output. Investing resources in sorting workers across jobs is important because it allows the firm to allocate their workers to where they are most valued and thus, improve efficiency. For example, a firm may (rightfully) require a different level of productivity (measured in output) from an accountant than a salesman because, naturally, the firm would like to allocate those workers with exceptional sales skills to the sales department and workers with less sales skills elsewhere. Hence, the firm needs to offer the "right" incentive scheme to attract the right workers for the job.

In the process of designing incentives, the firm also needs to consider the monitoring costs associated with the incentive scheme. If worker productivity is low, the firm will benefit from assigning workers to a piece-rate incentive scheme to boost productivity. It might also help the firm to attract more able workers because high-ability workers benefit from the opportunity to distinguish themselves from low-ability workers. However, if monitoring costs are high it may not be profitable to invest in piece-rate pay. As an alternative, the firm could offer salaries, which is a function of effort rather than output and does not require monitoring (i.e. low monitoring costs). In addition, Lazear (1986) states that salaries could increase the quality of the produced output by directing focus away from quantity. The downside of offering salaries is that it increases the possibility of attracting low ability workers. Because workers bear the cost of monitoring to distinguish themselves from their peers, it does not pay off for a low-ability worker to do so. Firms are aware of this fact and consequently offer salary workers a lower level of compensation to weigh up for the lack of ability. Thus, we summarize that piece-rate pay is most likely the preferred choice when output is easily quantified and monitoring costs are low. When the opposite is true, salaries is the better alternative.

In 2004, Shearer conducted a field experiment with the goal of providing evidence of incentive effects from piece-rates and fixed wages. The experiment took place within a tree-planting firm in

Canada and showed that average productivity increased by 20 % when workers were paid piece-rates rather than fixed wages. This finding is supported by Lazear's study from 2000, where he found that productivity of workers from a large autoglass corporation, Safelite Glass Corportion, increased by 44 % when introduced to piece-rate. The effects were documented by examining 3000 workers' behavior over a 19-month period. Lazear did, however, emphasize that the productivity increase was split in two components and that half of the increase was due to the firm's ability to attract the most productive workers and the possible reduction in quits among those.

Although there exists evidence for positive effects of piece-rate pay, its effects are disputed. Kvaløy et al. (2013) asserted that monetary rewards sometimes induce worse performance and referred to what psychologists call a "hidden cost of reward". The term refers to crowding out intrinsic motivation by undermining their confidence in own abilities. It could also be discussed in the context of social preferences. If the agent inhibits the desire to reciprocate behavior, the response to an economic incentive will be contingent on how the agent perceives the principal's behavior. More specifically, if the agent perceives the principal's behavior as kind they value the principal's payoff positively. On the other hand, if the principal's behavior is perceived as hostile, the agent values the principal's payoff negatively. Results from Dickhaut et al. (1995) imply that reduction in performance could be due to workers' perceiving performance pay as a signal of distrust. If this is the case, agents reciprocating what they perceive as hostile behavior could explain reduction in performance. Fehr and Gächter's (2000) study supports the implication as they found that subjects exerted more effort once performance pay was removed. Moreover, these results illustrate that performance pay in certain situations may work against its purpose and demotivate fair-minded workers.

To overcome the negative impact of performance pay, Kvaløy et al. (2013) suggested that motivational efforts such as motivational talk could work as a complement to, and enhance the effect of, monetary incentives. To investigate this proposition they conducted a field experiment where students were hired to enter data from ice hockey game reports into a database. Their results showed that performance pay increased performance *only* when it was accompanied by motivational talk, which implies that the monetary incentive was reinforced by appealing to other elements of motivation than material rewards.

In 1971, Deci conducted an experiment to assess the causal relationship between an individual's assigned payment scheme and motivation. Using the previously mentioned SOMA Cube puzzle, they found that students in the extrinsic incentive condition (piece-rate) spent less time with the puzzles when they no longer had instrumental value compared to students who had received pay non-contingent on performance (fixed pay). In addition, a further experiment using the same task examined the effects of "verbal rewards" on subsequent intrinsic motivation. After each puzzle solved, students would receive (false) feedback that their time solving the puzzle was "much better than average" for their peers. Rather than decreasing subsequent intrinsic motivation, as by the monetary incentive, the feedback condition *increased* later motivation (Deci, 1971).

Lepper and Greene (2000) present the proposition that verbal rewards providing evidence of one's ability and competence in performing a task will have a greater positive (or less negative) impact on intrinsic motivation than rewards without such information. Moreover, Deci & Ryan (1985) suggest that feedback stating competence and providing guidelines for performance improvement should have these positive effects, which is also in line with the psychological view of reciprocate behavior.

Based on this, we assert that also feedback should be considered as an incentive, in terms of a symbolic reward and positive behavior by the principle, to mitigate the incentive problem presented in the principal-agent model.

2.3 Feedback

Performance feedback is one of the more popular interventions of Organizational Behavior Management (OBM) strategies and is often used as an approach to organizational change (Prue & Fairbank, 1981). Compared to other intervention strategies performance feedback has the advantages of low economic costs, simplicity and flexibility, which allows a wide range of organizations, regardless of size and scale, to implement these programs. Although there exists research proving positive effects of feedback, the use of the term, its exact meaning, and the principles causing its effectiveness is highly disputed (Alvero, Bucklin, & Austin, 2001).

Ramaprasad (1983) defines feedback as information about the gap between actual level and the reference level of a system parameter (e.g. input, process or output). However, he extends the definition so that the information only serves as feedback if it is used to alter the gap. Thus, provision of feedback requires the existence of a reference level (e.g. budgets, production schedules, sales targets etc.), data on the actual level of the parameter and a mechanism for comparing the two to generate information about the gap (Ramaprasad, 1983). In terms of the principal-agent model, provision of feedback is contingent on the principal's ability to accurately measure the agent's performance. The agent's performance can be measured quantitatively, qualitatively or by a combination of the two. Quantitative metrics typically consists of outputs such as sales or production units, while qualitative performance measures usually involve measuring less concrete metrics such as customer satisfaction. Naturally, different performance measures constitute grounds for different types of feedback. Feedback based on quantitative performance measures usually lack subjectivity and is thereby categorized as objective feedback. On the other hand, qualitative performance measures *require* somewhat subjectivity and subsequent feedback is categorized accordingly.

One of the most important goals of feedback is to provide an incentive to enhance performance. Considering that feedback is a result of measured performance, the economic view is that its power as an incentive is directly related to the agent's perception of accuracy. When performance is easily quantified and the number of uncontrollable factors is low, objective feedback can work as a powerful incentive by providing the agent with specific information on how to alter the gap between expected and actual performance. When the opposite is true, the agent imposes a risk of being measured inaccurately and objective feedback loses its power as an incentive. In such cases, qualitative performance measures and subjective feedback are claimed to be a better fit (Lazear & Gibbs, 2009). Subjectivity in evaluating performance allows for several metrics to be included and reduces the agent's risk of being held accountable for uncontrollables. However, one of the potential risks of using subjective evaluations is evaluator bias. If managers do not have sufficient incentives to provide accurate employee ratings, they might lack motivation to invest time in gathering the necessary information to do so (Fox, Bizman, & Herrman, 1983). Consequently, the performance evaluation will suffer from discrepancies and the incentive effect of the feedback intervention will be (negatively) impacted accordingly.

Considering feedback in the context of social preferences characteristics may be depicted differently. Fehr and Falk (2001) state that economists tend to narrow human beings into economic rational beings and ignore the fact that they are also social in nature. In other words, the psychological aspect should also be considered in facilitating feedback. Accounting for an agent's desire to reciprocate behavior may suggest that the extent to which feedback is positive or negative is essential in consideration of when to use objective or subjective feedback. Based on the fact that subjective feedback is a function of the principal's personal evaluation, agents may perceive positive subjective feedback as a more honest and kind behavior of the principle compared to positive objective feedback. Correspondingly, the opposite may be true in the case of negative feedback. Negative subjective feedback may be perceived as a more hostile behavior compared to negative objective feedback. Thus, objective feedback may be optimal in conveying negative evaluations.

During the academic year 1990-1991, Azmat and Irriberri (2010) conducted a natural experiment at a high school where students were provided with (objective) relative performance feedback on grades in addition to the usual individual performance feedback. Relative performance feedback is often referred to as comparative feedback as it contains information about an individual's relative performance compared to other subjects in a treatment group. The study found that provision of relative performance feedback had a strong and positive effect on performance and increased overall grades by 5 %. Closer related to an organizational setting, Azmat and Irriberri (2012) wanted to examine whether the effects of relative performance feedback was contingent on incentive scheme. They argued that, under a piece-rate incentive scheme, individuals were informed about both relative performance *and* relative income, and the effects had not yet been distinguished. Thus, they conducted a study that examined the effect of relative performance feedback under a piece-rate and a flat-rate incentive scheme in a real-effort setting. The study's result under piece-rate incentives were in line with previous research; relative performance feedback had a strong and positive effect and increased individual performance by 17 %. However, under flat-rate incentives individuals reacted only when consequences in terms of relative income were provided but the effect was insignificant.

Provision of relative performance feedback is not necessarily uniformly associated with positive effects. It could lead to some distortions in incentives, especially under a piece-rate incentive scheme because of the excessive short-focus. A study in 2013 by Charness et al. showed that ranking feedback motivated subjects to put in extra effort but introducing a sabotage opportunity caused subjects to expose themselves to costs in order to improve their rank, which had a strong damaging effect on performance. This implies that relative performance feedback may work as an incentive by triggering competitive preferences, but also as an encouragement to engage in unethical measures to improve own ranking. Because quantitative measures only correspond to part of employee's tasks, such incentive distortions are of probable risk. Consequently, this may lead to the firm suffering negative consequences due to the competitive environment created by relative performance feedback (Lazear,1989; Konrad, 1999; Chen, 2003). In such cases, subjective evaluations can be used to avoid incentive distortions by mitigating the problems of excessive short-term focus (Gibbs, Merchant, Van der Stede, & Vargus, 2005).

Gibbs et al. (2005) discussed the many benefits of using subjective evaluations and state that subjectivity could indeed be used to improve incentives. The discussion was based on data collected from U.S. firms in the automobile retailing industry regarding allocation of annual bonus awards. They argued that subjectivity could improve incentives by mitigating problems due to contract incompleteness (not accounting for important but unmeasured tasks) that encouraged excessive short-term focus and performance measure manipulation. The evidence was presented in the context of Human Performance Technology (HPT) framework and showed that subjective bonuses were effective in mitigating perceived weaknesses in formula bonus awards, reducing employee risk and recalibrating incentives.

In a review of the performance feedback literature, Balcazar et al. (1985) systematically evaluated effects of performance feedback. The main purpose of the review was to determine feedback combinations and characteristics that produced the highest levels of consistent effects. Feedback was categorized as consistently effective when it uniformly produced desired mean increases or decreases in performance compared to mean baseline levels and/or levels produced by any other independent variable(s). The review stated three main results: (1) feedback does not uniformly improve performance; (2) adding rewards and/or goal setting procedures to feedback improves

consistency in its effects; and (3) some characteristics of feedback are more consistently associated with improved performance than others. In 2001, Alvero et al. replicated the review to update the findings by comparing it to later literature (1985-1998). This resulted in support of Balcazar et al.'s findings but with some deviation in arguments. For example, Alvero et al. found that feedback with antecedents (excluding goals) produced the highest levels of consistent effects (100 %) compared to feedback alone (47 %). Feedback combined with antecedents refers to subjects receiving feedback and an antecedent stimulus in some form (e.g. staff training and weekly task objectives), while feedback alone refers to subjects only receiving information about the quantity or quality of their performance. A comparison of individual performance to a standard and previous performance, which is the two most popular contents of feedback, yielded consistency effects of 50 % and 56 %, respectively. Alvero et al.'s conclusion was that feedback in combination with other procedures (e.g. antecedents or behavioral consequences) produced overall higher levels of consistent effects than feedback alone.

According to the "Law of Effect", first introduced by Edward Thorndike in 1930, any behavior followed by *pleasant* consequences is likely to be repeated, and any behavior followed by *unpleasant* consequences is likely to be stopped (Gray, 2010). Azmat & Irriberri (2012) studied the affective response of feedback and found that positive (negative) feedback affects subjects' happiness positively (negatively). This implies positive (negative) feedback to be a pleasant (unpleasant) consequence of behavior and thus, works to modify behavior towards improvement. Given that feedback is most effective when combined with other procedures such as behavioral consequences, its potential to act as an incentive may also vary with the principal's power related to anticipated rewards or sanctions (Ilgen, Fisher, & Taylor, 1979). Thus, the outcome of feedback is ultimately a function of the principal's history of consistency in provision of behavioral consequences following the feedback intervention. If previous feedback interventions have not led to changes in reinforcement and rewards or punishment it is less likely that the next intervention will have the intended effect (Prue & Fairbank, 1981). This means that, all else equal, the more power of the principal the greater likelihood the agent will try to respond in line with the feedback (Ilgen, Fisher, & Taylor, 1979).

2.4 Contribution to literature

With the increasing focus on behavioral aspects as an approach to organizational change, the importance of understanding measures that impact these elements rises accordingly. Recall that, because extrinsically motivated behaviors are not inherently interesting, the economic view is that they must be externally prompted in some way (e.g. by monetary incentives and rewards). Shearer (2004) and Lazear (1986) show that monetary incentives such as piece-rate pay could impact performance and productivity positively. On the contrary, Deci (1971) and Kvaløy et al. (2013) find that piece-rate pay could in fact be damaging on performance unless it is accompanied by intrinsic factors such as feedback or motivational talk. Fehr and Falk (2001) argue that economists have a hard time understanding why economic incentives sometimes backfire because they do not consider the psychological foundations of motivation. Moreover, they suggest that social preferences like the desire to reciprocate behavior play a substantial role in shaping human behavior and should be accounted for in the process of incentive design.

The dominant view in psychological literature is that extrinsic motivation works in opposition to intrinsic motivation (e.g. Deci, 1971; Deci & Ryan, 1985; Deci & Ryan, 2000; Lepper & Greene, 1973). However, Amabile (1993) suggests that extrinsic and intrinsic factors could complement each other and work as a motivational synergy. The argument is supported by Deci (1971) and Kvaløy et al.'s (2013) findings where the incentive effect switches from negative to positive when they simultaneously prompt intrinsic motivation by feedback or motivational talk. Performance feedback as a complementary intrinsic factor is our main focus in this thesis, and is suggested to provide a reinforcing by supporting workers' sense of competence and providing guidelines on how to improve performance (Amabile, 1993).

Although there exists empirical evidence that feedback improves performance (e.g. Azmat & Irriberri, 2010) it is shown that effects are not necessarily consistent (Balcazar, 1985). In addition, accounting for individuals' social preference like the desire to reciprocate behavior suggests that objective and subjective feedback impact the psychological aspect of motivation differently (Fehr & Falk, 2001). This raises the question of how feedback should be utilized to generate the most positive and consistent effects. The previous section (2.3) presented the theoretical aspects as to when objective and subjective feedback should be utilized. However, to our knowledge, there exists

no research examining the effects. The motivation for our study draws on this shortage and our main contribution is a direct comparison of the effects of objective and subjective feedback on performance. We assert that provision empirical evidence on this area could contribute to increase the understanding of how to facilitate performance feedback programs and generally improve utilization.

Finally, based on Amabile's (1993) argument that intrinsic and extrinsic factors complement each other, we suggest that monetary incentives and feedback work as a motivational synergy in enhancing performance. Closest related to our research is the frequently mentioned study by Kvaløy et al. (2013) that investigate the relationship between monetary incentives and simple motivational talk. The study provides certain indications of how the combination of intrinsic and extrinsic motivational factors impact performance but does not provide any implications of effects of feedback based on actual performance. Thus, in addition to the comparison of subjective and objective feedback, we attempt to test feedback and monetary incentives as a motivational synergy (in terms of an interaction effect). Our research question is worded as follows:

Does objective and subjective feedback generate different effects in performance, and are these effects reinforced when combined with different incentive schemes?

3. Experimental design and procedure

3.1 General description

In preparation of answering our research question we examined a large body of existing literature on motivation, feedback and monetary incentives. The purpose was to gather sufficient information about previous research methods to optimally design ours. The dominant research method in this line of literature is the classic experiment, which is an empirical quantitative approach. Experiments fall within the category of causal research design and are carried out to measure the effect of independent variables on some dependent variable in a controlled environment. Causal research falls within the category of conclusive research and has two main purposes; determining a cause and effect relationship; and determining the nature of the relationship and the predicted effect. This study's main goal was to determine whether there existed a causal relationship between the independent variables feedback and incentive schemes and the dependent variable performance. To extend our research, we attempted to examine the nature of the relationship by looking at the effect size, further examining whether one feedback type should be preferred over the other in a certain setting.

To collect the relevant data for our study we conducted a lab experiment. The laboratory in this context is understood broadly as all arenas that are artificial in relation to arenas in which behavior usually occurs (Fostervold, 2001). The experiment was thoroughly designed by considering the essential factors of experimental design; manipulation and control, well aware that if an experiment is conducted correctly it provides excellent data to draw conclusions about causality (Fostervold, 2001). Thus, proper design was of the essence.

An experiment assumes a large degree of control over the involved variables and it allowed us to manipulate the independent variables (feedback type and incentive schemes) while holding all other variables but the dependent variable (performance) constant. As such, great emphasis was put on the simplicity and comprehensibility of the task in the design process and preparation of instructions. This was to ensure that performance would not be impacted by external factors such as skills or knowledge. Also, the feedback manipulation was carefully formulated so that it was similar in

wording, but yet provided separate messages. By this, we minimized the risk of data errors due to poor manipulation or subjects misunderstanding the instructions. In addition, a short questionnaire was added to gain data on subjects' motivation.

The research question in this study is not limited to any certain professions or population, rather the contrary; it seeks to gain an understanding of performance enhancing measures across *all* work related situations where one might encounter the need to evaluate performance and provide incentives. Thus, the desired sample for our study was one that would represent a general population. Based on this and our available resources, students across all lines of education from the University of Stavanger were chosen as our sample. This provided a wide perspective of potential treatment effects on several different personalities of people from different backgrounds. In addition, we assert that including non-business students gave a larger insight on the psychological motivational effect of feedback compared to business students who might be more selective in exerting effort based on knowledge of utility maximization, economic benefits etc.

Students were recruited by E-mail, ItsLearning, fraternity Facebook pages and flyers distributed on campus. Participants could choose between eight sessions over two days, where each session lasted approximately 45 minutes. A limit of 18 people was set per session including a planned overbook as preparation for no-shows. Each treatment was conducted twice and had been randomly assigned to the eight sessions. Thus, the session the student attended was decisive for which treatment they would be assigned to.

Altogether 136 students volunteered to participate in the experiment. However, 23 no-shows yielded 113 subjects for our analysis (43 males and 70 females).

3.2 The Experiment

The experiment consisted of a combined coloring- and calculation task over three time periods. Each period lasted 10 minutes and subjects were informed that they would receive oral individual feedback on their performance after each period. The feedback would result in an awarded grade on a scale from 1 to 6, where 1 was the lowest and 6 was the highest possible grade. The coloring- and

calculation task for each period consisted of 3 task sheets (appendix 3) containing 45 triangles with multiplication tasks that specified the triangle's color. The color intervals were given on the bottom of each task sheet and varied from sheet to sheet to avoid biased results from memorization. With memorization being controlled for, the subjects' challenge was to calculate as many multiplication tasks as possible and give the triangles their belonging color.

Subjects were instructed to read the instructions individually before they were read out loud by the experimenter. During each period, they were notified when there were 5 minutes left and 30 seconds left. In addition, subjects were seated individually but could potentially view other subjects flipping pages. However, monitoring resulted in disciplined behavior in line with instructions.

3.2.1 Incentive Scheme

Subjects were informed about their assigned incentive scheme at the beginning of the experiment. In the fixed pay treatment the individual payoff was NOK 100, regardless of performance. In the performance pay treatment, pay was directly linked to the subjects' performance (grade). Each grade point counted for NOK 10 and total pay was based on their grade from all three periods. Thus, if a subject received grade 3 in period 1, grade 4 in period 2, and grade 5 in period 3, the subject would receive a total payoff of NOK 120 (30 + 40 + 50). Mistakes and/or potential cheats were not checked during the experiment. This fact was not brought to the subjects' attention at any point.

3.2.2 Feedback

Subjects' performance-based grade was given in two different ways. Depending on treatment, feedback was either objective or subjective. It is important to note that both objective and subjective feedback were based on a quantitative performance measure. Performance was equally evaluated by grading after the number of completed task sheets. The only difference was the way in which feedback was expressed. In the objective feedback treatment, the colored triangles were visibly counted and openly graded in line with this number. In the subjective feedback treatment, triangles were *not* visibly counted but rather expressed as the experimenter's personal evaluation of their work. To ensure consistency in the feedback manipulation, a standard sentence for each treatment was formulated as presented below.

Objective feedback – “This qualifies to a grade X”

Subjective feedback – “Based on what I see, I would like to give you a grade X”

As preparation to any objections on grade, a standard justification for each treatment was set to ensure consistency in feedback on objections as well.

Justification for objective feedback – “Your grade is based on the average of previous participants’ performance”

Justification for subjective feedback – “Your grade is based on my evaluation of how good your performance was”

To ensure consistency and avoid any confounding factors in grading, we constructed a grading system (table 1) that was easily practiced. The grading system was constructed by piloting the task on acquaintances and estimating an average based on their performance. Average performance was 1 to 1.5 task sheets, which was then set to be the average grade 3.

Table 1: *Grading system*

# of task sheets completed	Grade
< 0.5	1
> 0.5 < 1	2
> 1 < 1.5	3
> 1.5 < 2	4
> 2 < 2.5	5
> 2.5	6

The grading system was used identically in both the objective and subjective treatment. As presented in table 1, subjects obtained the grade 2 if they finished 1 task sheet, grade 4 if they finished 2 task sheets, etc. As an extra precaution to ensure consistency in grading, experimenters

switched groups of subjects to evaluate in the second period so that all subjects were evaluated by each experimenter at least once. In addition to being a precaution, it was a method of illustrating to the subjects that grading was consistent.

3.2.3 Questionnaire

The last part of the experiment consisted of a short questionnaire. The first part of the questionnaire required filling out demographic information such as age, gender and line of education. This information was demanded to categorize subjects to control for any variation in effects. The last part of the questionnaire consisted of four "to which degree" questions regarding subjects' motivation to perform in the experiment, the extent of which their payment scheme and feedback affected their motivation and their liking of receiving feedback in the work place. "To which degree" questions were chosen because of their well-fit design in mapping attitudes towards a given subject.

3.3 Treatments

Our research design consisted of a 2 x 2 between-subjects design: feedback type (objective or subjective); and incentive scheme (performance pay or fixed pay). The between-subjects design is often called an independent measures design because it provides the opportunity of comparing groups independently due to participants only being subjected into a single treatment (Shuttleworth, 2009). The four treatment groups are presented in table 2.

Table 2: *Overview treatments*

	Fixed pay	Performance pay
Subjective feedback	FP-subjective	PP-subjective
Objective feedback	FP-objective	PP-objective

3.4 Pilot

The experiment was piloted on 4 students and ourselves. Naturally, feedback was not provided in the latter case. The pilot was conducted in groups of two so that we were able to discover weaknesses and adjust them accordingly before conducting it on the next group. This allowed greater learning in the process of conducting the experiment and also in testing adjustments.

One of the main goals of the pilot was to ensure that the task completion time was not too short/long. We discovered that what we may have thought to be a relatively straightforward and "boring" task was substantially more challenging, exciting and time-consuming than expected. The time frame was initially set to be seven minutes per period, but the pilot showed a need to extend this to ten minutes. Also, some parts of the instructions needed better specification.

Participants of the pilot experiment were asked to answer some questions regarding the task and their motivation to complete it. One of the questions was whether they developed any specific strategies that increased performance. This was to ensure that the potential performance effect was in fact due to our independent variables (feedback and incentive schemes) and not strategy development. Pilot participants responded that several strategies were tested but with no conclusion on a single best strategy. When they were asked about what motivated them to perform their best they responded that, regardless of incentive scheme, they wanted to improve their grade and do better than in the previous period.

3.5 Limitations

As all studies face certain limitations, ours is no exception. In presenting the study's limitations we focus mainly on internal and external validity. *Internal validity* addresses the issue of whether we can be certain that the estimated effects are in fact due to the independent variables (X) and not extraneous factors (Fostervold, 2001). The *degree* of internal validity is thus related to research design, how well the study was run, how variables were measured, etc. On the other hand, *external validity* addresses the question of whether the estimated effects can be generalized and is often split into two types; population validity and ecological validity. Population validity considers the representativeness of the research sample, while the ecological validity addresses whether the

effects can be generalized to other settings (i.e. whether effects from a laboratory setting can be generalized to a real-world setting) (Shuttleworth, 2009). The goal of any study is to achieve a high degree of both internal and external validity. However, most often one type of validity goes on expense of the other. Thus, the research question works as an indicator of which type of validity should be emphasized (Fostervold, 2001). In this study, internal validity has received most emphasis based on the fact that external validity requires a great deal of resources, which was not available for this thesis.

The first limitation addressing the internal validity is the sample. Moreover, we assert that the small sample size is the study's greatest limitation. A component within internal validity is selection bias, which refers to any systematic differences between subjects prior to treatments. In other words, the optimal sample would be one where subjects in each group were identical and functionally equivalent (Fostervold, 2001). However, this is not easily accomplished, and the small sample size and lack of thorough randomization brings some bias to our results. In addition, the experimental task requires that subjects are familiar with calculation. Repetitive testing (as done due to subjects performing the same task in each period) can bias our results by subjects gradually refreshing their calculation skills over periods. Hence, the possibility of increased learning through the experiment may negatively impact our internal validity.

The small sample size also highly impacts external validity. Generalization requires that the research sample is perfectly randomized and representative for the whole population, which leads us to the question of population validity (Fostervold, 2001). Naturally, a sample of 113 subjects consisting of only students is not representative for the population as a whole. In addition, the sample had a skewed gender distribution. Hence, our ability to generalize our results is limited and our population validity is most likely low.

The degree of ecological validity should also be discussed. The setting in which we observed behavior was artificial and might not illustrate behavior that would occur in a real-world setting. However, subjects were compensated for their participation in the experiment, and thus, we assert that the results could indicate trends in behavior in work related settings.

Another factor that may impact the internal validity is the level of compensation offered to the participants. Sufficient incentives are essential in order for them to have any effect on performance. In the fixed pay treatment subjects were compensated with NOK 100 for 45 minutes work. This may have been too low to generate any motivation to perform well and thus, the estimated effects from this treatment may be biased. Also, in the performance pay treatment subjects were compensated with NOK 10 for each grade point they obtained. In retrospect of the experiment it was discussed whether the grading system was set too optimistically, and whether that may have had a demotivating effect. If this were the case, it would suggest that the effect of the monetary was impaired. However, quite a few participants finished all task sheets in period 2 and 3, which implies that poor performance of other subjects was due to other factors (e.g. differences in calculation and processing skills).

One common deficiency in experiments and surveys is that the participants often tend to feel that they are being watched and/or think that a correct answer or behavior is expected. Despite the fact that the instructions emphasized that the goal was to color as many triangles as possible, the subjects in the subjective feedback treatment (who did not receive visual counting) may have been confused when they received low grades. This may further have caused them to perform contrary to the instructions by focusing excessively on quality as an attempt to seek explanations as to how performance was measured. Consequently, this may bias our results by participants not acting realistically.

3.6 Hypotheses

There are three main hypotheses related to our research question:

Hypothesis 1 addresses the issue of whether objective and subjective feedback impact performance differently. Accounting for certain agents' desire to reciprocate behavior suggests that objective and subjective feedback should impact the psychological aspect of motivation differently (Fehr & Falk, 2001). The underlying assumption is that subjective feedback as a function of the principal's personal evaluation may be perceived as a more direct kind or hostile behavior to the extent feedback is positive or negative. Moreover, this would indicate that negative subjective feedback has a more damaging psychological effect on motivation compared to negative objective feedback.

Additionally, because the nature of the experimental task is highly quantitative, subjective feedback may be perceived as somewhat abstract and confusing. According to economic theory on performance evaluations and feedback, objective feedback should be experienced as more understandable on quantitative tasks and provide clearer guidelines on how to improve performance. Results from Deci & Ryan (1985) imply that feedback stating competence and providing guidelines could be perceived as kind behavior by the principle and thus, generate positive performance effects by appealing to the agents' possible desire to reciprocate behavior. Based on this, we predict that objective feedback has a stronger positive impact on performance than subjective feedback.

Hypothesis 1:

> H_0: Subjects exposed to objective feedback do not perform differently than subjects exposed to subjective feedback.
>
> H_1: Subjects exposed to objective feedback perform better than subjects exposed to subjective feedback.

Hypothesis 2 seeks to determine whether performance pay and fixed pay impact performance differently. The economic view on monetary incentives suggests that performance pay has a stronger impact on performance than fixed pay. This is because performance is directly related to compensation, which means that additional effort accrues additional benefit. Also, according to basic utility maximizing theory, individuals will not expose themselves to costs (in terms of effort) unless it leads to increased utility. Thus, we predict that subjects assigned performance pay has an overall higher performance compared to subjects assigned fixed pay.

Hypothesis 2:

> H_0: Subjects assigned performance pay do not perform differently from subjects assigned fixed pay.
>
> H_1: Subjects assigned performance pay perform better than subjects assigned fixed pay.

Hypothesis 3 addresses whether there is an interaction effect between feedback and incentive scheme. The prediction is based on Amabile's (1993) argument that extrinsic and intrinsic motivational factors could complement each other and work as a motivational synergy, which is supported by Kvaløy et al.'s (2013) findings. They find that the effect of performance pay is only

positive when it is accompanied by an intrinsic factor (motivational talk) and Amabile (1993) explains this as "intrinsics in service of extrinsics". Additionally, Balcazar et al. find that adding rewards to feedback improves consistency in its effects implying that the relationship is mutual in that extrinsics also work in service of intrinsics. Thus, based on this, and the predictions by hypothesis 1 and 2, we expect that objective feedback as an intrinsic factor and performance pay as an extrinsic factor complement each other and have a positive interaction effect.

Hypothesis 3:

H_0: There exists no interaction effect between feedback and incentive scheme.

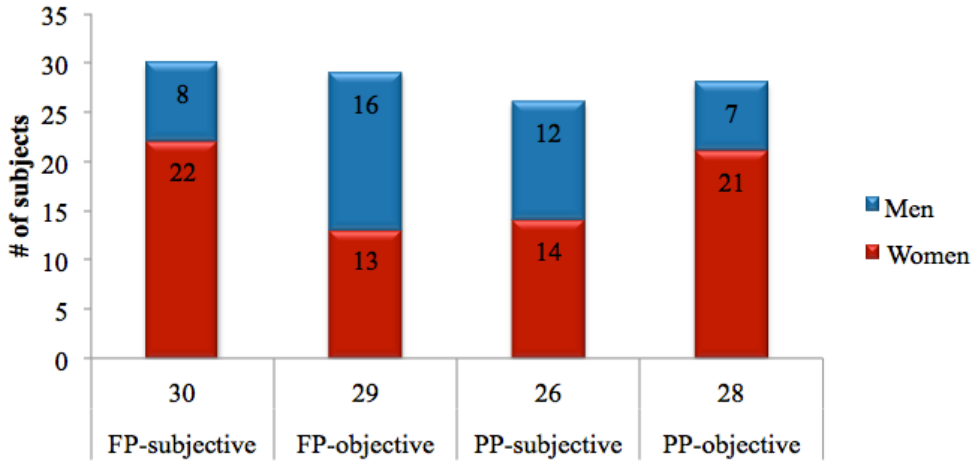
H_1: There exists a positive interaction effect between objective feedback and performance pay.

4. Results

4.1 Descriptive Statistics

Our experiment consisted of 113 participants and we were successful in acquiring a somewhat equal number of subjects in each treatment group. Figure 1 presents the number of subjects and proportions of men and women across treatments. Females are generally overrepresented in treatments due to the overweight of women in our data sample (70 females vs. 43 males).

Figure 1: *Overview gender distribution in treatment groups*



Age of subjects was similar across treatments and ranged from 19 to 54. However, due to the fact that students were used as our research sample, age ranged primarily from 19 to 25. Average age was 23.5 and 24 respectively for men and women.

Figure 2: *Overview business and non-business students over treatments*

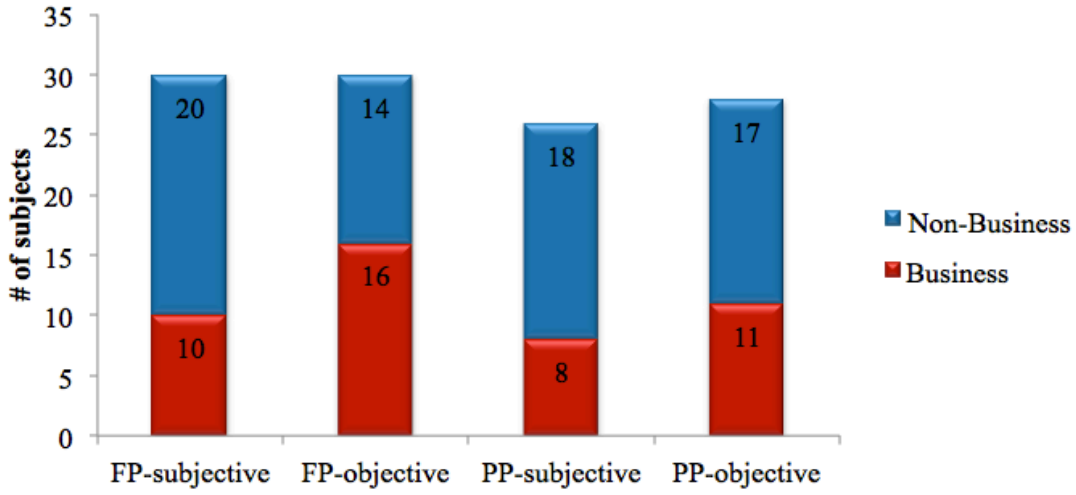


Figure 2 illustrates the proportions of business and non-business students over treatments. Of the total sample, 44 of subjects stated they were business students, while 69 stated otherwise. This shows that we acquired a sample representing students from several lines of education and thus, several different personalities.

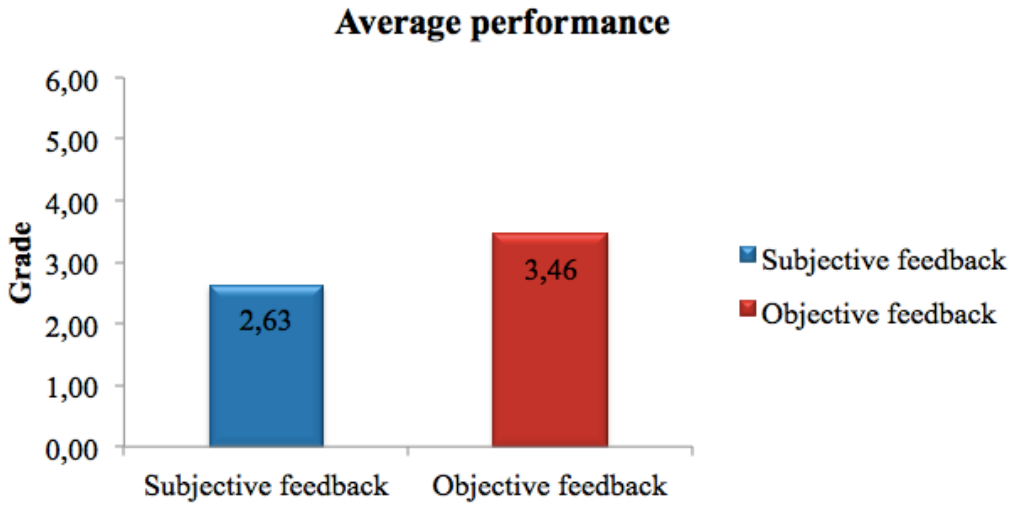
4.2 Data Analysis

Our analysis rests primarily on two different statistical tests. The first is a Mann-Whitney U-test, which similarly to a two sample t-test, allows us to compare groups but without assuming a parametric distribution. Because we have not assumed any prior distribution, a non-parametric test is wider in scope and provides more robust results. The second is a multiple regression analysis, which will be used to measure the potential interaction effect between feedback and incentive scheme. The regression analysis allows us to test several independent variables (X's) on the dependent variable (Y) and provides the possibility of adding control variables to widen our perspective.

4.2.1 Main effect testing hypothesis 1 - Feedback

Performance is defined as the subject's rewarded grade on their work after each period. Our dependent variable in testing feedback is average performance (mean grade) from period 2 and 3. Performance from period 1 is excluded because the feedback manipulation had not yet been introduced.

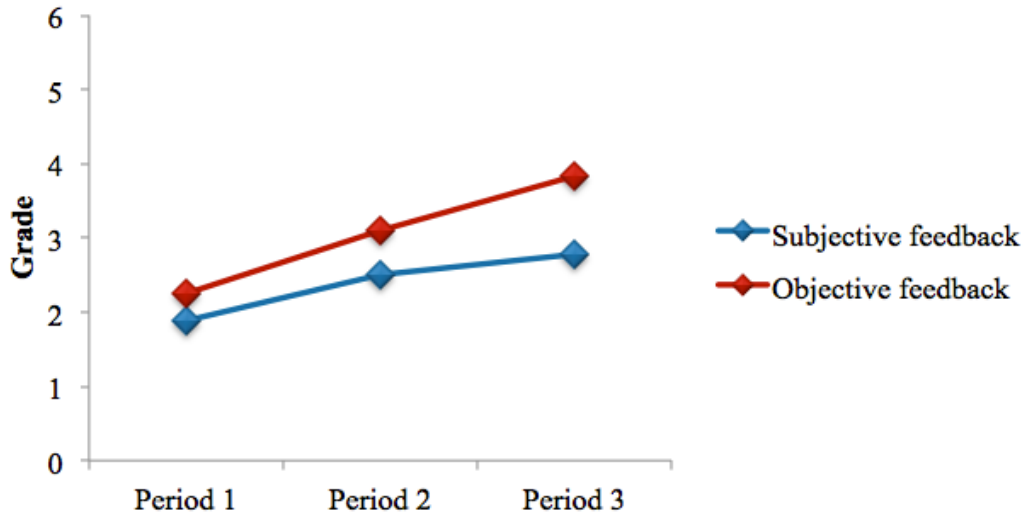
Figure 3: *Feedback – average performance (in period 2 and 3)*



As illustrated in figure 3 and predicted by hypothesis 1, subjects exposed to objective feedback perform overall better than subjects exposed to subjective feedback. Average performance for subjects in the objective treatment is 3.46 (*SD = 1.228*) while it is 2.63 (*SD = .946*) for subjects in the subjective treatment. The Mann-Whitney U-test reveals that the difference is statistically significant (*p= .000*).

To elaborate our results we test the effect of feedback in each period independently. Thus, we conduct the same test, using average performance from each period. Note that average performance in period 1 illustrates performance before the feedback manipulation had been introduced.

Figure 4: *Feedback – average performance over periods.*

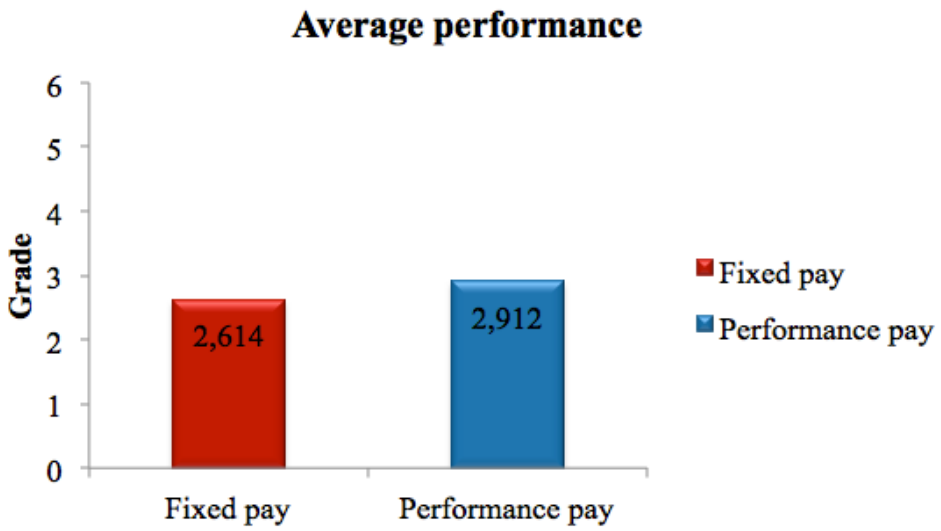


As shown in figure 4, subjects' performance is nearly equal in both treatments, which implies that there are no substantial differences in skills prior to the manipulation. Performance gradually increases over periods for subjects in both treatments. However, subjects in the objective treatment significantly perform consistently better over all periods (Mann-Whitney U test, period 2: $p = .010$, period 3: $p = .000$). Thus, we reject our null hypothesis.

4.2.2 Main effect testing hypothesis 2 – Incentive scheme

Hypothesis 2 addresses whether performance pay works as a stronger incentive to enhance performance compared to fixed pay. The dependent variable in testing the effect of incentive schemes is average performance (mean grade) from *all three periods*, as the incentive scheme manipulation is introduced prior to period 1.

Figure 5: *Incentive schemes – average performance.*



As illustrated in figure 5 and predicted by hypothesis 2, subjects assigned performance pay have an overall higher average performance compared to subjects assigned fixed pay. Average grade is 2.912 (*SD = 1.130*) and 2.614 (*SD = .965*) for subjects in the performance pay and fixed pay treatment, respectively. However, the difference is not statistically significant (Mann-Whitney U-test, $p = .123$) and thus, we fail to reject our null hypothesis.

Figure 6: *Incentive scheme – average performance over periods*

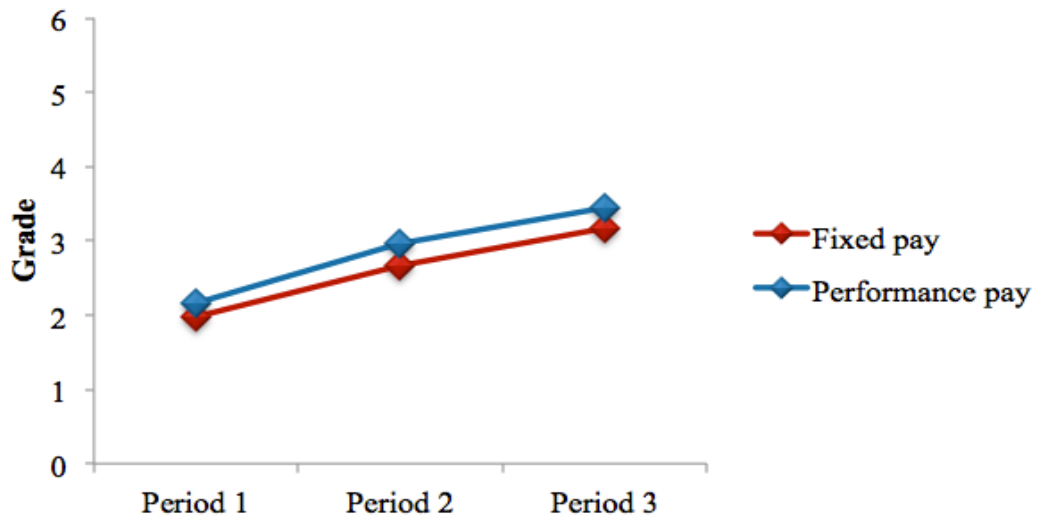


Figure 6 illustrates the average performance in each period and shows that performance gradually increases over periods but without substantial alteration depending on payment scheme. The Mann-Whitney U-test confirms this, as differences in average performance are not statistically significant for any period (period 1: $p = .201$, period 2: $p = .203$, period 3: $p = .233$).

4.2.3 Interaction effect testing hypothesis 3

The initial approach to testing hypothesis 3 is to compare treatments by conducting a Mann-Whitney U-test. Although the procedure may seem similar to the one done in previous sections, it differs in that it elaborates differences *within* the objective and subjective feedback treatments when incentive scheme is included and vice versa.

Because the feedback and incentive scheme manipulation are introduced at different points on time we conduct the Mann-Whitney U-test using average performance from period 2 and 3.

Figure 7: *Average performance (mean grade from period 2 and 3) over treatments*

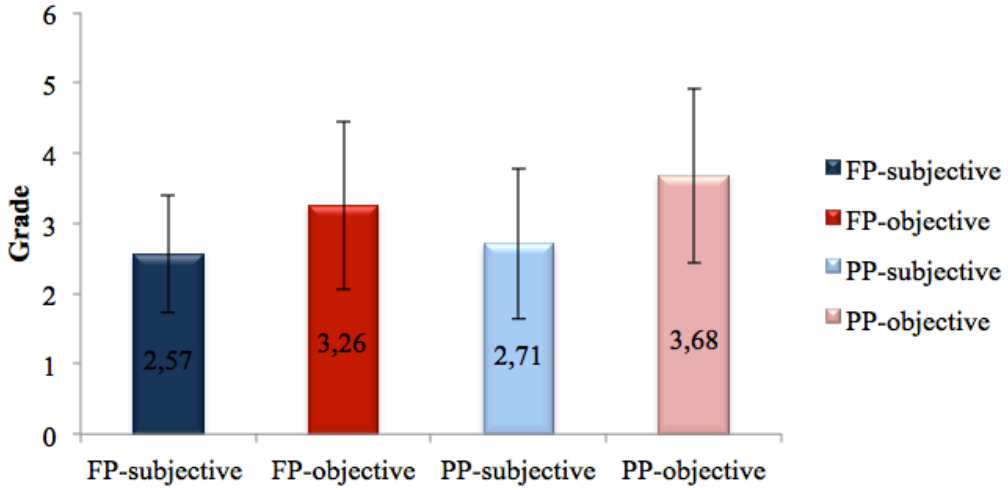


Figure 7 illustrates average performance from period 2 and 3 over treatments and shows that, depending on treatment, performance varies. The standard deviations are quite substantial, which indicates a large spread in subjects' average performance within treatments. However, small sample sizes generally creates larger deviations and thus, they are in line with our predictions.

There is a substantial difference in performance when subjects are given subjective versus objective feedback. Under performance pay (PP-subjective and PP-objective), objective feedback yields higher performance than subjective feedback and the difference is statistically significant (Mann-Whitney U-test p = .005). The same is true within fixed pay treatments (FP-subjective and FP-objective) and the effect is statistically significant on a 5 % level (Mann-Whitney U-test p = .031).

Moreover, we see that within objective feedback treatments (FP-objective and PP-objective) performance pay yields the highest average performance. However, the effect is not statistically significant (Mann-Whitney U-test p = .163). Within subjective feedback treatments (FP-subjective and PP-subjective) differences in performance are not substantial and neither significant (Mann-Whitney U-test p = .866).

Figure 8: *Average performance of treatment groups over periods*

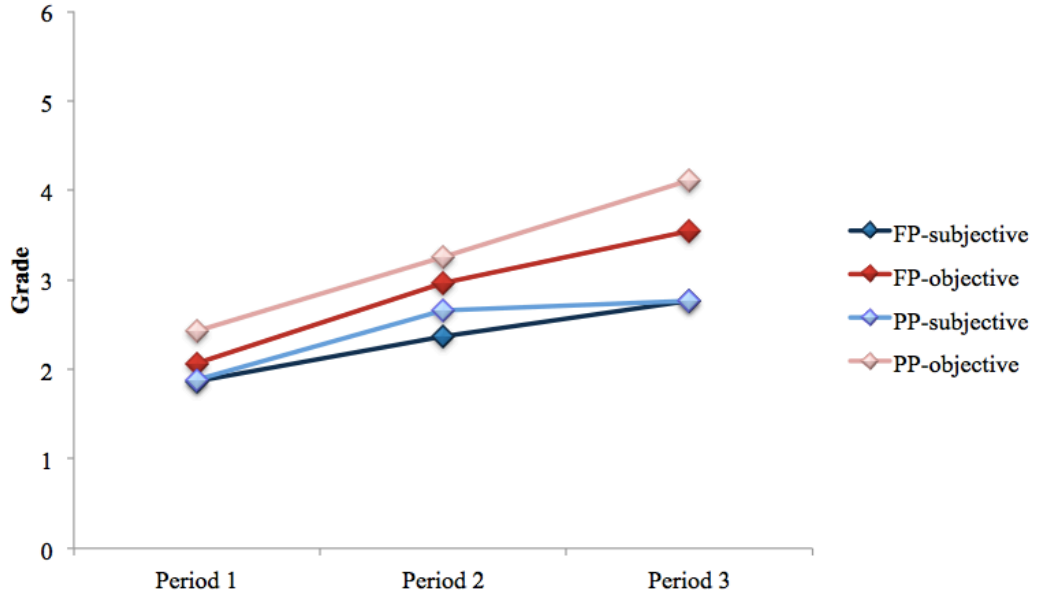


Figure 8 illustrates average performance of the four treatment groups over periods. Performance increases over all treatments from period 1 to 2. The same is true for all treatments from period 2 to 3; however, the increase of the PP-subjective treatment is quite modest. Performance increases are presented numerically in table 3.

Table 3: *Increase in performance over periods*

	FP-subjective	FP-objective	PP-subjective	PP-objective
Increase P1 to P2	.500	.897	.769	.821
Increase P2 to P3	.400	.568	.115	.857
Total increase	.900	1.483	.884	1.678

As previously reported, the objective feedback treatments experience the highest increase in performance. In addition, the combination of feedback and incentive scheme yielding the highest performance increase is objective feedback combined with performance pay.

Till now, our approach has been a simple effects procedure where we break the potential interaction effect into components and separately test them for significance. Testing hypothesis 3 requires conducting a regression analysis to estimate the potential interaction effect. A statistically significant interaction effect would mean that the impact of feedback would depend on incentive scheme and vice versa. To conduct the regression analysis we first estimate a linear regression model with an interaction term to determine the independent variables' effect on performance:

$$y_i = \beta_0 + \beta_1 o_i + \beta_2 p_i + \beta_3 o_i p_i + \varepsilon_i$$

Here, $o_i = 1$ if subject i was exposed to objective feedback (i.e., participated in FP-objective or PP-objective treatment), while $p_i = 1$ if subject i was assigned performance pay (i.e., participated in PP-subjective or PP-objective treatment). Thus, the base group for our analysis is FP-subjective. The interpretation of the beta-coefficients is that β_1 shows the effect of objective feedback without performance pay; β_2 shows the effect of performance pay without objective feedback; and β_3 shows the interaction between performance pay and objective feedback. Our key dependent variable (y) is average performance (mean grade from period 2 and 3). In addition, to elaborate our results we conduct the regression analysis with performance from each period independently.

Table 4: *Effects of feedback and incentive scheme without controls*

	Mean Grade	Period 1	Period 2	Period 3
Objective feedback	0.691**	0.202	0.599**	0.785**
	(0.285)	(0.216)	(0.294)	(0.311)
Performance pay	0.145	0.018	0.287	0.003
	(0.294)	(0.222)	(0.303)	(0.320)
Obj x PP	0.275	0.342	-0.003	0.553
	(0.413)	(0.312)	(0.426)	(0.451)
Constant	2.567***	1.867***	2.367***	2.767***
	(0.200)	(0.151)	(0.206)	(0.218)
Observations	112	112	112	112
R^2	0.146	0.071	0.083	0.187

Table 3 presents the OLS coefficient estimates (with standard deviation in parenthesis) and has the following variables:

- Dependent variable y = Performance (in terms of subjects' achieved grade)
- Independent variables
    - "Objective Feedback" – A dummy variable that took the value 1 if subjects were exposed to objective feedback and 0 if exposed to subjective feedback.
    - "Performance Pay" – A dummy variable that took the value 1 if subjects were assigned performance pay and 0 if assigned fixed pay.
    - "Obj x PP" – the interaction of both afore mentioned dummy variables.
- ***p < 0.01, **p < 0.05, *p < 0.1

With average performance (mean grade) as our dependent variable, we see that objective feedback alone has a significant effect on the performance while performance pay alone is insignificant. Moreover, there is a positive interaction effect between objective feedback and performance pay in line with predictions by hypothesis 3, but the effect is not statistically significant.

Looking at performance from each period, objective feedback alone has a significant effect in both period 2 and 3. Moreover, the increasing beta-coefficients indicate that the effect becomes stronger the more subjects are exposed to it. Performance pay alone and the interaction effect is not statistically significant for any period.

To check how predictive our model is we examine the coefficient of determination (R^2). R^2 reports how well the independent variables explain changes in the dependent variable and is always between 0 and 1 (Fostervold, 2010). The closer it is to 1, the better is the model. In our case, R^2 is low, which indicates that the predictive ability of the model is poor. As an attempt to improve R^2 we add control variables (gender, age, dummy indicating whether subjects were enrolled in business studies) to our model. Also, control variables ensure that significant effects are due to the independent variables and not systematic differences in subjects

$$y_i = \beta_0 + \beta_1 o_i + \beta_2 p_i + \beta_3 o_i p_i + \text{controls} + \varepsilon_i$$

Table 5: *Effects of feedback and incentive scheme with controls*

	Mean Grade	Period 1	Period 2	Period 3
Objective Feedback	0.527*	0.120	0.394	0.660**
	(0.296)	(0.227)	(0.302)	(0.326)
Performance Pay	0.057	-0.012	0.190	-0.076
	(0.294)	(0.226)	(0.301)	(0.324)
Obj x PP	0.520	0.439	0.281	0.760
	(0.426)	(0.327)	(0.436)	(0.469)
Gender	-0.479**	-0.186	-0.547**	-0.411*
	(0.221)	(0.169)	(0.226)	(0.243)
Age	-0.011	-0.015	-0.020	-0.002
	0.025	(0.019)	(0.026)	(0.028)
Education	0.090	0.079	0.150	0.030
	(0.216)	(0.165)	(0.220)	(0.237)
Constant	3.157***	2.332***	3.204***	3.109***
	(0.663)	(0.508)	(0.678)	(0.730)
Observations	112	112	112	112
R^2	0.184	0.088	0.138	0.209

Table 4 presents OLS coefficient estimates (standard deviations in parenthesis) and has the following variables:

- Dependent variable y = Performance (in terms of subjects' achieved grade)

- Independent variables
    - "Objective Feedback" – A dummy variable that took the value 1 if subjects were exposed to objective feedback and 0 if exposed to subjective feedback.
    - "Performance Pay" – A dummy variable that took the value 1 if subjects were assigned performance pay and 0 if assigned fixed pay.
    - "Obj x PP" – the interaction of both afore mentioned dummy variables.
- Control variables
    - "Gender" – A dummy variable that took the value 1 if subjects were female and 0 if male.
    - "Age"
    - "Education" – A dummy variable that took the value 1 if subjects were enrolled in business studies and 0 if they were enrolled in non-business studies.
- ***p < 0.01, **p < 0.05, *p < 0.1

By adding controls, objective feedback alone is significant on a 10 % level when mean grade is used as the dependent variable. Why it is only significant on a 10 % level can be explained by the skewed distribution of men and women in our sample. Women proved to be less productive than men on the experimental task and thus, the results change when the gender variable is added as a control.

Using performance of each independent period, objective feedback is significant only in period 3. The beta-coefficients show that objective feedback alone increases performance by .660 compared to subjective feedback, and that being female decreases performance by .411 compared to males (significant on a 10 % level). Performance pay alone does not have any statistically significant effects on performance in either period.

The regression analysis shows a positive interaction effect between objective feedback and performance pay, however, the effect is not statistically significant and thus, we fail to reject our null hypothesis.

4.2.4. Other testing

As a sidebar to testing our hypotheses, we examined the effect of receiving low grades in period 1 and 2 on performance in period 3. Our goal was to see whether low grades in the first periods impacted subjects' performance in the latter. In addition, we wanted to examine whether low grades could explain some of the previously presented results. More specifically if so-called "low-graders" drove the negative effects of subjective feedback. A low grade was categorized as ≤ 2, while a high grade was categorized as ≥ 3. To conduct our analysis, we added two dummy variables to our regression model (LowP1 and LowP2) along with dummies estimating the interaction effect of low grades, feedback and incentive scheme. Performance in period 3 was used as our dependent variable.

$$y_i = \beta_0 + \beta_1 \mathrm{o_i} + \beta_2 p_i + \beta_3 \mathrm{lowP1}_i + \beta_4 \mathrm{lowP2}_i + \beta_5 o_i \mathrm{lowP1}_i + \beta_6 o_i \mathrm{lowP2}_i + \beta_7 p_i \mathrm{lowP1}_i + \beta_8 p_i \mathrm{lowP2}_i \text{ controls} + \varepsilon_i$$

Here, lowP1 = 1 if subject *i* obtained a low grade in period 1, while lowP2 = 1 if subject *i* obtained a low grade in period 2. Results are reported in table 5.

Table 6: *Effects of feedback and incentive scheme with dummy variables for low grade in period 1 and 2*

	Grade Period 3
Objective feedback	0.900**
	(0.371)
Performance Pay	-0.469
	(0.335)
Low Grade P1	-1.427***
	(0.319)
Low Grade P2	-1.396***
	(0.295)
Obj x Low Grade P1	0.044
	(0.439)
Obj x Low Grade P2	0.547
	(0.345)
PP x Low Grade P1	1.105***
	(0.416)
PP x Low Grade P2	-0.554
	(0.351)
Gender	0.034
	(0.168)
Age	-0.005
	(0.020)
Education	0.068
	(0.168)
Constant	4.451***
	(0.595)
Observations	112
R^2	0.638

Table 5 presents OLS coefficient estimates (standard deviations in parenthesis) and has the following variables:

- Dependent variable y = Performance in period 3 (in terms of subjects' achieved grade)
- Independent variables
    - "Objective Feedback" – A dummy variable that took the value 1 if subjects were exposed to objective feedback and 0 if exposed to subjective feedback.

    - "Performance Pay" – A dummy variable that took the value 1 if subjects were assigned performance pay and 0 if assigned fixed pay.
    - "Low Grade P1" – A dummy variable that took the value 1 if subjects obtained a low grade in period 1 and 0 if otherwise.
    - "Low Grade P2" – A dummy variable that took the value 1 if subjects obtained a low grade in period 2 and 0 if otherwise.
    - "Obj x Low Grade P1" – A dummy variable that took the value 1 if subjects were exposed to objective feedback and obtained a low grade in period 1 and 0 if otherwise
    - "Obj x Low Grade P2" – A dummy variable that took the value 1 if subjects were exposed to objective feedback and obtained a low grade in period 2 and 0 if otherwise
    - PP x Low Grade P1" – A dummy variable that took the value 1 if subjects were assigned performance pay and obtained a low grade in period 1 and 0 if otherwise
    - PP x Low Grade P2" – A dummy variable that took the value 1 if subjects were assigned performance pay and obtained a low grade in period 2 and 0 if otherwise
- Control variables
    - "Gender" – A dummy variable that took the value 1 if subjects were female and 0 if male.
    - "Age"
    - "Education" – A dummy variable that took the value 1 if subjects were enrolled in business studies and 0 if they were enrolled in non-business studies.
- ***p < 0.01, **p < 0.05, *p < 0.1

As reported in table 5 the effect of objective feedback alone is still positive and significant. In addition, the analysis reveals that subjects who obtain a low grade in period 1 or 2 perform worse in period 3 compared to subjects who obtain a high grade.

There are no significant interaction effects between feedback and low grade, which inhibits us from determining whether "low-graders" drive the negative effect of subjective feedback. However, the interesting outcome from the analysis is the significant interaction effect between performance pay

and low grade in period 1. Moreover, the positive beta coefficient suggests that subjects who obtain a low grade in period 1 and are assigned performance pay perform *better* in period 3 compared to subjects who achieve a high grade in period 1 and are assigned fixed pay (base group). If subjects obtain a low grade in period 2 the interaction effect is negative implying that the opposite is true. However, this effect is not statistically significant.

The coefficient of determination, R^2 (.638), is considerably higher than in our previous regression models, and indicates that the added independent variables highly contribute to explain changes in performance.

Finally, we examined the relationship between the quality of subjects' work (measured in mistakes/cheats) and feedback and incentive scheme. A Mann-Whitney U test was conducted to determine whether a certain feedback type or incentive scheme encouraged more cheating and/or led to more mistakes (appendix 6). Subjects in the objective feedback treatment had an average mistake level of 5.05 % compared to 7.38 % for subjects in the subjective feedback treatment. Under performance pay, the average mistake level was 6.77 %, while it was 5.69 % under fixed pay. However, the differences are not significant (appendix 4 and 5) and given the overall low level of errors (6.21 %), we assume that they are due to miscalculations and not cheats. Thus, the results do not provide evidence that a certain feedback type or incentive scheme impair work quality.

5. Discussion

5.1 Feedback effects

The effects of feedback prove to be positive in that subjects in both the objective and subjective feedback treatments increase their performance over periods. In addition, in line with the predictions by hypothesis 1 objective feedback has a greater impact on performance than subjective feedback. The effect can be tied to the subjects' perception of accuracy, which is a crucial element for the power of the incentive. As Lazear & Gibbs (2009) argues, objective feedback can work as a powerful incentive when performance is easily quantified. Thus, the positive effect of objective feedback compared to subjective feedback could be due to the quantitative nature of the experimental task. Because the colored triangles were visibly counted in the objective feedback treatment, subjects were provided with a certain indication of how performance was measured. Consequently they were indirectly provided with guidelines on how to improve performance. This may have caused subjects to experience objective feedback as more accurate than subjective feedback. In the subjective feedback treatment, the colored triangles were not visibly counted further causing subjects to be confused about how they were measured. This confusion may have elicited subjects to explore other solutions (e.g. improving quality in terms of more accurate coloring or investing more time on calculations) shifting their focus away from the initial task.

Deci & Ryan (1985) stated that the motivational aspect of feedback is that workers obtain a sense of empowerment when they are provided with feedback stating their competence. This means that negative feedback may undermine the sense of achievement and consequently impair workers' motivation to perform. The presented intuition is supported by Azmat & Irriberi's (2012) study that found that positive (negative) feedback affected workers' happiness positively (negatively). In addition, the lack of indication as to how performance was measured in the subjective feedback treatment may cause negative subjective feedback to strike even harder. The intuition could be further elaborated by considering the individual's desire to reciprocate behavior. Fehr and Falk (2001) argued that this was an important psychological aspect of motivation and that such motives alone could cause changes in performance. If subjects perceived subjective feedback as abstract and confusing in addition to experiencing uncertainty about how performance was measured, subjects

may have felt they were treated unfairly. In the case of reciprocate behavior, an individual would respond positively to positive actions by the principal or manager and vice versa. If subjects experienced the experimenters' evaluations as unfair, this may have resulted in negative behavior by not exerting the necessary effort to increase performance.

5.2 Incentive scheme effects

The analysis in testing hypothesis 2 fails to provide any significant effects of performance pay having a greater impact on performance compared to fixed pay. However, the data displayed a trend in which subjects assigned performance pay perform better than subjects assigned fixed pay. This is in line with the majority of previous research on monetary incentives and suggests that material rewards do in fact play a part in enhancing motivation and increasing performance.

Although there exists a certain trend of performance pay being more effective in enhancing performance than fixed pay, the trend is not as clear as we may have anticipated. This could be explained by subjects' being intrinsically motivated to achieve high grades, which is supported by the pilot subjects' statement (see 3.4) that their main focus was to improve their grade, regardless of incentive scheme. Moreover, this is seemingly true for several of the subjects as they increased performance despite the fact that the fixed pay condition meant that additional effort did not accrue additional benefit.

The intrinsic motivation to improve performance could also be discussed in the context of individuals' competitive preferences. Individuals are assumed to have inherent competitive preferences, which causes them to get extra utility from performing better than others and disutility from performing worse (Azmat & Irriberri, 2012). Alvero et al. (2001) showed that the possibility of comparing performance to a previous period increased the consistency in the effects of feedback. Thus, we assume that the competitive preferences extend to wanting to compete with themselves, especially when a previous performance for comparison is available. In addition, some subjects attended the experiments with friends and although they did not communicate during the experiment, competitiveness could arise by subjects knowing they would exchange results with their friends when the experiment was over.

According to previous research, intrinsic motivation should be discounted when sufficient extrinsic factors are present. Our results thus lead to a discussion of whether the compensation level was sufficient. Recall that the limitations section presented a concern that our grading system may have been too strict. Consequently this would mean that subjects under performance pay had to work harder for their money than subjects under fixed pay. The incentive schemes and grading system were initially designed so that they were predicted to yield approximately the same payoff for both the performance pay and fixed pay treatments. I.e., we expected the average payoff for subjects in the performance pay treatment to be approximately NOK 100. However, the average payoff was NOK 86 for the performance pay treatment, while it was NOK 100 for the fixed pay treatment.

5.3 Interaction effect

Contrary to what was predicted by hypothesis 3, the regression analysis yielded no significant interaction effect between objective feedback and performance pay. However, there exists a certain trend; objective feedback combined with performance pay yields the highest average performance, while subjective feedback combined with fixed pay yields the lowest.

Although the interaction effect is insignificant, it is interesting to explore the trends yielded by the analysis. The results from the regression model with controls (table 4) show that performance pay has a negative impact on performance in period 1 and 3, while the effect of objective feedback is positive throughout the experiment. The negative effect of performance pay can be tied to what psychologists call “the hidden cost of reward”, which suggests that performance pay undermines confidence in own abilities and thus crowds out intrinsic motivation. In addition, Deci (1971) found that intrinsic motivation to perform a task was impaired when there was no longer a sufficient instrumental value to it. Thus, if the compensation level was to low to correspond to the grading system, the low instrumental value of the task may have had a detrimental effect on intrinsic motivation and impaired the potential effect of performance pay. However, in line with Kvaløy et al.’s (2013) findings, we find that the effect switches from negative to positive when accompanied by objective feedback (interaction effect). The interaction effect is positive and also greater than objective feedback alone in period 1 and 3, which indicates that objective feedback and performance

pay complement each other. This trend is in line with the predictions from hypothesis 3, but the statistically insignificant result inhibits us from rejecting the null hypothesis and supporting Amabile's (1993) suggestion of motivational synergies.

5.4 Other testing

When variables for low grade were added, the regression analysis revealed that subjects who obtained low grades in period 1 and 2 also performed worse in period 3 compared to subjects who obtained high grades. This implies that obtaining a low grade in period 1 or 2 negatively impacts performance.

Based on the suggested psychological effect of subjective feedback, the main purpose of expanding the regression model was to see whether such an effect could be precluded by so-called "low-graders" driving the negative effects. However, the positive interaction effect between objective feedback and low grade (Obj x Low Grade P1 and Obj x Low Grade P2) implies that subjective feedback compared to objective feedback has a negative effect on performance *regardless* of whether the subject is a "high-grader" or a "low-grader". But because the interaction effect is not statistically significant we cannot preclude the possibility that low-graders are the reason why subjective feedback has a negative effect compared to objective feedback.

The analysis reveals a positive and significant interaction effect between performance pay and low grade in period 1. The interaction effect shows that being assigned performance pay compared to fixed pay positively impacts subjects' performance in period 3 if he obtained a low grade in period 1 rather than a high grade. This supports previous findings in research on monetary incentives in that performance pay works as a stronger incentive to enhance performance than fixed pay. Moreover, it is interesting to see that the interaction effect between performance pay and low grade in period 2 is *negative.* This may imply that the incentive effect of performance pay is crowded out by the psychological effect of receiving low grades. However, this interaction effect is not significant.

6. Conclusion

The purpose of this thesis was to examine whether objective and subjective feedback generated different effects in performance and whether incentive schemes reinforced these potential effects. We find that subjects exposed to objective feedback have a greater and more consistently increasing performance over periods compared to subjects exposed to subjective feedback. The effect is explained mainly by two components: (1) the quantitative nature of the task caused subjects to experience objective feedback as more understandable compared to subjective feedback and; (2) negative subjective feedback had a more damaging psychological effect compared to negative objective feedback. However, both the objective and subjective feedback treatments experienced an increase in performance, further suggesting that feedback overall has a positive impact on performance.

The analysis concerning incentive schemes did not provide any significant effects. However, it did display a tendency of performance pay yielding a higher increase in performance over periods compared to fixed pay. Moreover, the treatment that yielded the highest performance increase was objective feedback combined with performance pay. The effect was not statistically significant but showed that when the performance pay alone had a negative effect, combining it with objective feedback produced a positive effect. In addition, the interaction effect was greater than the effect of objective feedback alone. Although our results do not provide empirical evidence supporting Amabile's (1993) suggestion of motivational synergies we assert that they do provide important implications for future research.

To generalize the findings in this thesis additional research is required. However, we believe that our research and experimental design is a stepping-stone toward expanding the scarce body of research examining the effects of feedback. By making small adjustments like increasing the sample size and ensuring that the compensation level is set optimally, we believe that the design could produce valid data and a sufficient conclusion to add to existing literature.

7. References

Azmat, G., & Irriberri, N. (2010, 04 14). The importance of relative performance feedback information: Evidence from a natural experiment using high school students. *Journal of Public Economics* , pp. 435-452.

Azmat, G., & Irriberri, N. (2012). *The Provision of Relative Performance Feedback Information: An Experimental Analysis of Performance and Happiness.* Centre for Economic Performance.

Alvero, A. M., Bucklin, B. R., & Austin, J. (2001). An Objective Review of the Effectiveness and Essential Characteristics of Performance Feedback in Organizational Settings. *Journal of organizational behavior management* , pp. 3-29.

Amabile, T. (1993). Motivational synergy: Toward new conceptualizations of intrinsic and extrinsic motivation in the workplace. *Human Resource Management Review , 3* (3), ss. 185-201.

Balcazar, F., Hopkins, B., & Suarez, Y. (1985). A critical, objective review of performance feedback literature. *Journal of Organizational Behavior Management , 7*, ss. 65-89.

Charness, G., Masclet, D., & Villeval, M. (2013, 08 02). The Dark Side of Competition for Status. *Management Science* , pp. 38-55.

Chen, K.-P. (2003, 02). Sabotage in Promotion Tournaments. *Journal of Law, Economics and Organization* , pp. 119-139.

Deci, E. (1971). Effects of Externally Mediated Awards on Intrinsic Motivation. *Journal of Personality and Social Psychology , 18* (1), ss. 105-115.

Deci, E. (1972, October 2). The effects of contingent and noncontingent rewards and controls on intrinsic motivation. *Organizational Behavior and Human Performance , 8* (2), ss. 217-229.

Deci, E., & Ryan, R. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology* (25), ss. 54-67.

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

Dickhaut, J., Berg, J., & McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior , 10* (1), ss. 122-142.

Fehr, E., & Falk, A. (2001). *Psychological Foundations of Incentives.* Universitas Turicensis, Institute for Empirical Research in Economics. Zürich: Institute for Empirical Research in Economics.

Fehr, E., & Gachter, S. (2002). Do Incentive Contracts Crowd Out Voluntary Cooperations. *European Economic Review , 46*, ss. 687-724.

Fox, S., Bizman, A., & Herrman, E. (1983). The Halo Effect: Is it a Unitary Concept? *Journal of Occupational Psychology* , pp. 289-296.

Fostervold, K. (2010). *Forskningsmetoder I* (Vol. 3). Oslo: NKI Forlaget.

Gibbs, M. J., Merchant, K. A., Van der Stede, W. A., & Vargus, M. E. (2005, 05/06). The Benefits of Evaluating Performance Subjectively. *Performance Improvement* , pp. 26-32.

Gray, P. (2010). Psychology. In P. Gray, *Psychology* (pp. 108-109). New York: Worth Publishers.

Ilgen, D. R., Fisher, C. D., & Taylor, M. (1979). Consequences of Individual Feedback on Behavior in Organizations. *Journal of Applied Psychology* , pp. 349-371.

Kvaløy, O., Nieken, P., & Schöttner, A. (2013). *Hidden Benefits of Reward: A Field Experiment on Motivation and Monetary Incentives.* Center for Economic Studies & Ifo Institute.

Konrad, K. A. (1999, 10 21). Sabotage in Rent-seeking Contests. *Journal of Law, Politics and Organization* , pp. 155-165.

Lazear, E. P., & Gibbs, M. (2009). Personnel Economics in Practice. New Jersey: John Wiley & Sons, Inc.

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

Lazear, E. (1986, 07). Salaries and Piece Rates. *Journal of Business , 59* (3), ss. 405-431.
Laffont, J.-J., & Martimort, D. (2002). *The Theory of Incentives.* New Jersey, USA: Princeton University Press.

Lepper, M., & Greene, D. (1973, 10). Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology , 28* (1), ss. 129-137.

Lepper, M., & Henderlong, J. (2000). Chapter 10: Turning "Play" into "Work" and "Work" into "Play": 25 years of Research on Intrinsic Versus Extrinsic Motivation. I C. Sansone, & J. Harackiewicz, *Intrinsic and Extrinsic Motivation - The Search for Optimal Motivation and Performance* (ss. 257-307). San Diego, California, USA: Academic Press.

Prue, D. M., & Fairbank, J. A. (1981). Performance Feedback in Organizational Behavior Management. *Journal of organizational behavior management* , pp. 1-16.

Sansone, C., & Harackiewicz, J. (2000). Chapter 1. Looking Beyond Rewards: The Problem and Promise on Intrinsic Motivation. I C. Sansone, & J. Harackiewcz, *Intrinsic and Extrinsic Motivation: The search for Optimal Motivation and Performance* (s. 2). San Diego, California, USA: Academic Press.

Shuttleworth, M. (2009, April 24). *Between Subjects Design*. Retrieved May 27, 2015, from Expolorable.com: https://explorable.com/between-subjects-design

Shearer, B. (2004). Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment. *Review of Economic Studies* (71), ss. 513-534.

Ramaprasad, A. (1983). On the Definition of Feedback. *Behavioral Science* , pp. 4-13.

Terstiege, S. *Objective versus Subjective Performance Evaluations.* Mannheim [u.a.]: Sonderforschungsbereich/Transregio15.

Appendix 1: Instructions Fixed Pay

Instruksjoner

Velkommen til eksperiment!
Foran dere ligger et oppgavesett bestående av tre deler. Oppgaven dere skal gjennomføre er en kombinert regne- og fargeleggingsoppgave. Hver del består av *tre ark* med trekanter som hver inneholder et regnestykke. Svaret på regnestykket angir fargen som skal fargelegges og fargeintervallene vil være angitt nedenfor bildet. *Fargene som benyttes i intervallene vil variere fra ark til ark*. Målet er å produsere så mange fargelagte trekanter som mulig. Hele trekanten må fargelegges for at den skal bli godkjent. Eksperimentet vil bestå av tre perioder à 10 minutter med opphold mellom periodene. I oppholdet vil arbeidet deres bli vurdert og dere vil motta en individuell tilbakemelding i form av en karakter. Karakterskalaen går fra 1-6, hvor 1 er laveste karakter og 6 er høyeste karakter. I de påfølgende periodene skal dere utføre samme oppgave med nye ark. Eksperimentet avsluttes med et spørreskjema.

Det er ikke lov å kommunisere med andre deltakere under eksperimentet.

Lønn
Du vil få utbetalt kr 100 for å gjennomføre dette eksperimentet.

Appendix 2: Instructions Performance Pay

Instruksjoner

Velkommen til eksperiment!
Foran dere ligger et oppgavesett bestående av tre deler. Oppgaven dere skal gjennomføre er en kombinert regne- og fargeleggingsoppgave. Hver del består av *tre ark* med trekanter som hver inneholder et regnestykke. Svaret på regnestykket angir fargen som skal fargelegges og fargeintervallene vil være angitt nedenfor bildet. *Fargene som benyttes i intervallene vil variere fra ark til ark*. Målet er å produsere så mange fargelagte trekanter som mulig. Hele trekanten må fargelegges for at den skal bli godkjent. Eksperimentet vil bestå av tre perioder à 10 minutter med opphold mellom periodene. I oppholdet vil arbeidet deres bli vurdert og dere vil motta en individuell tilbakemelding i form av en karakter. Karakterskalaen går fra 1-6, hvor 1 er lavest karakter og 6 er høyeste karakter. I de påfølgende periodene skal dere utføre samme oppgave med nye ark. Eksperimentet avsluttes med et spørreskjema.

Det er ikke lov å kommunisere med andre deltakere under eksperimentet.

Lønn
Utbetalingen vil være basert på karakterene du får etter hver av de tre periodene.
Hvert karakterpoeng utgjør kr 10.

Eksempel: Dersom du får karakteren 3 i periode 1, 4 i periode 2 og 5 etter periode 3 vil du få utbetalt kr 120 (30 + 40 + 50).

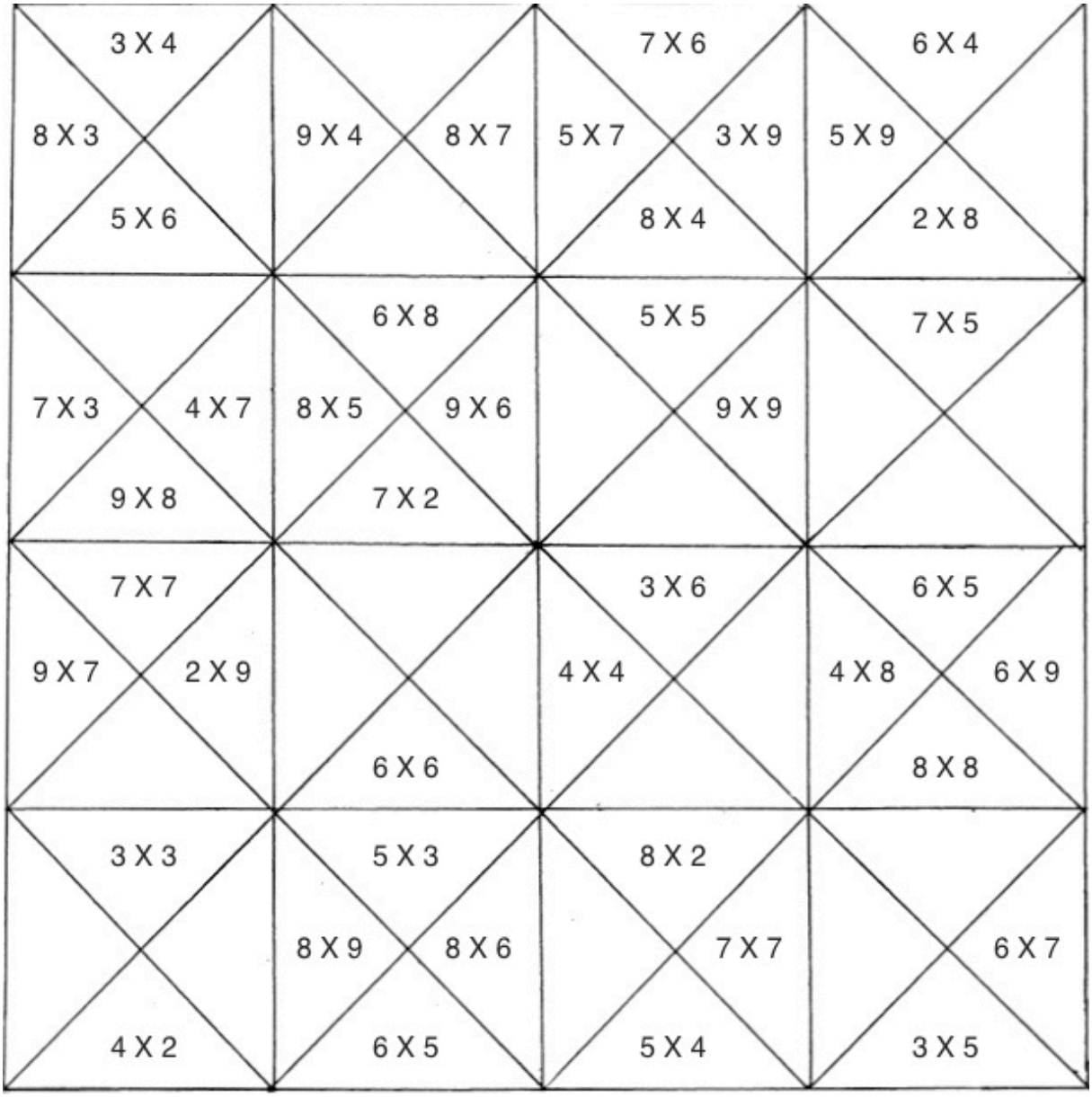
Appendix 3: Problem Set

DEL 1

VENNLIGST IKKE SNU ARKET FØR DU BLIR BEDT OM DET

Ark 1

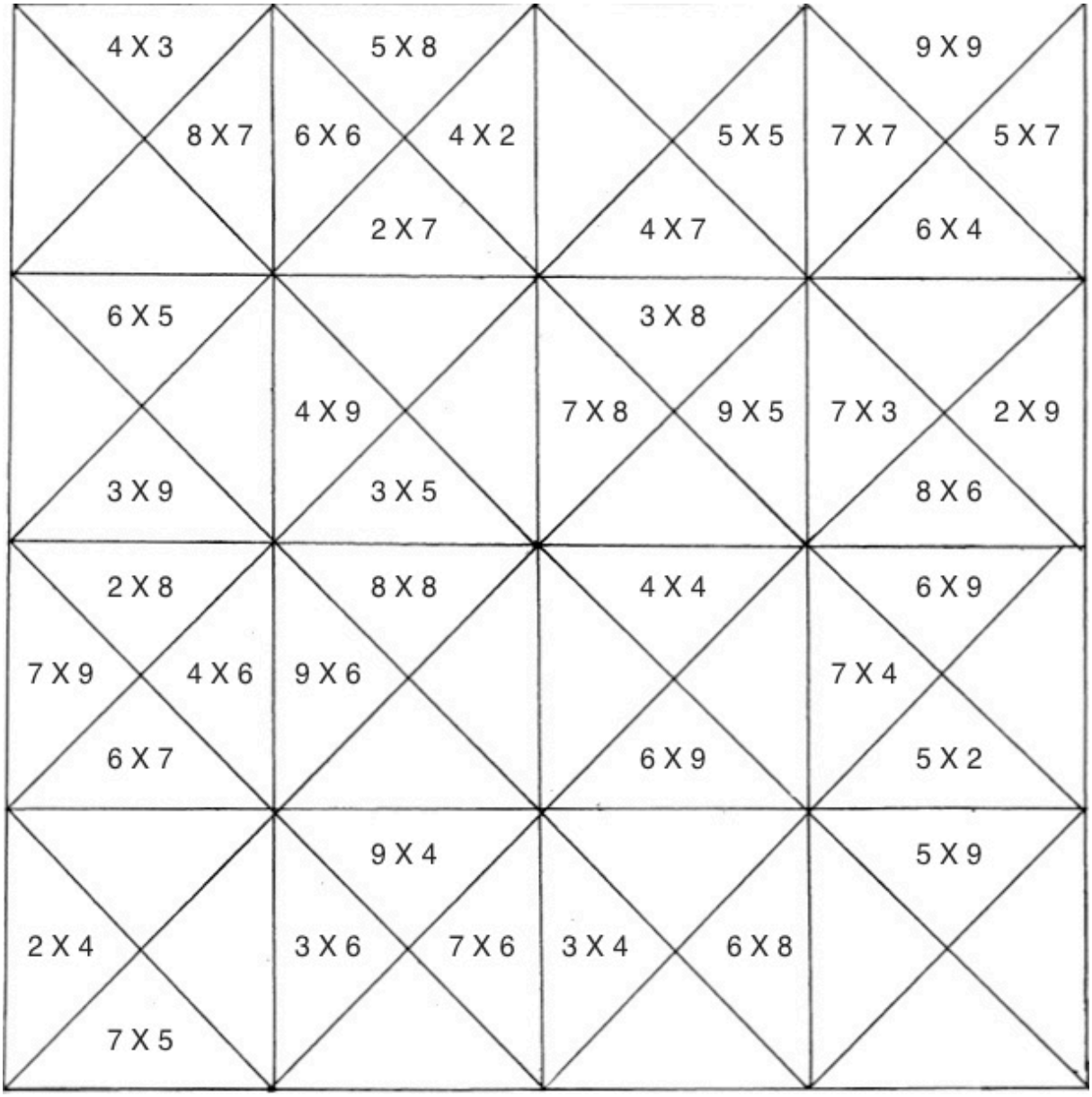
Fargelegg så mange ruter som mulig. Tallene representerer ulike farger. Fargeintervallene er angitt nedenfor bildet. Blanke felt skal ikke fargelegges.



Fargeintervall:
1-15 = Blå
16-30 = Lilla
31-45 = Grønn
46-60 = Rød
61-75 = Svart
76-90 = Rosa

Ark 2

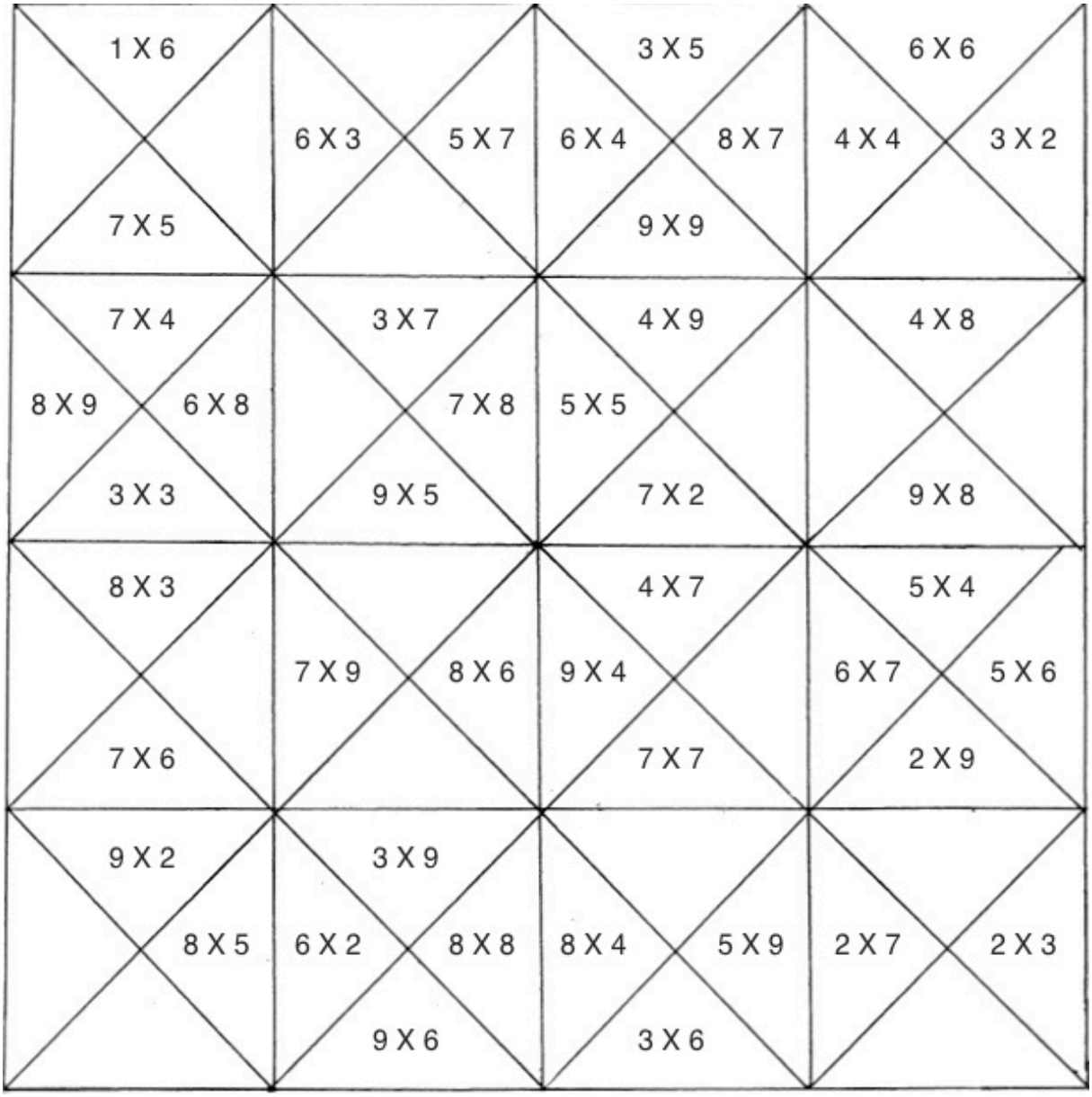
Fargelegg så mange ruter som mulig. Tallene representerer ulike farger. Fargeintervallene er angitt nedenfor bildet. Blanke felt skal ikke fargelegges.



Fargeintervall:
1-15 = Grønn
16-30 = Lilla
31-45 = Blå
46-60 = Svart
61-75 = Rød
76-90 = Rosa

Ark 3

Fargelegg så mange ruter som mulig. Tallene representerer ulike farger. Fargeintervallene er angitt nedenfor bildet. Blanke felt skal ikke fargelegges.



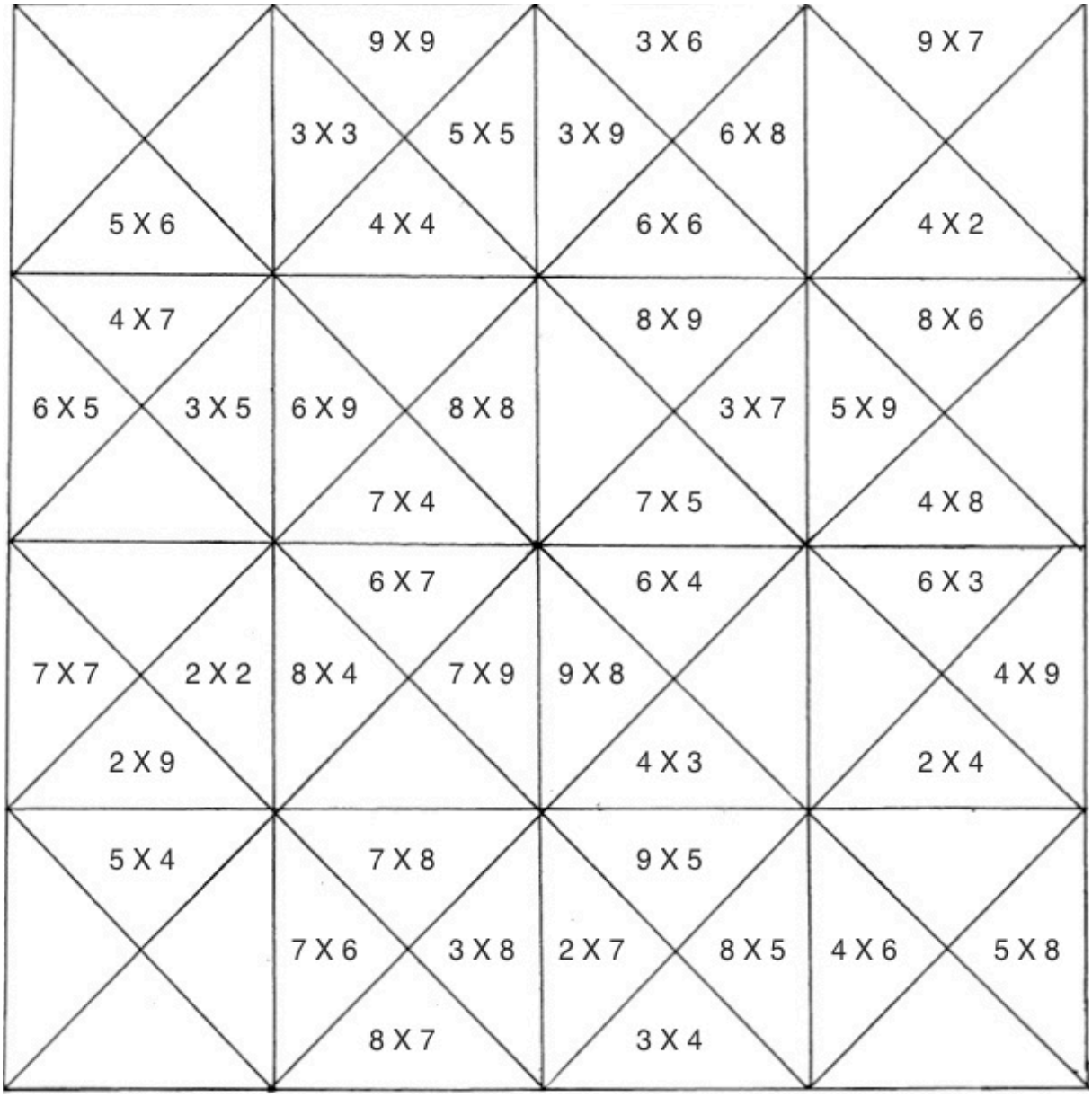
Fargeintervall:
1-15 = Brun
16-30 = Grønn
31-45 = Svart
46-60 = Blå
61-75 = Rød
76-90 = Gul

DEL 2

VENNLIGST IKKE SNU ARKET FØR DU BLIR BEDT OM DET

Ark 1

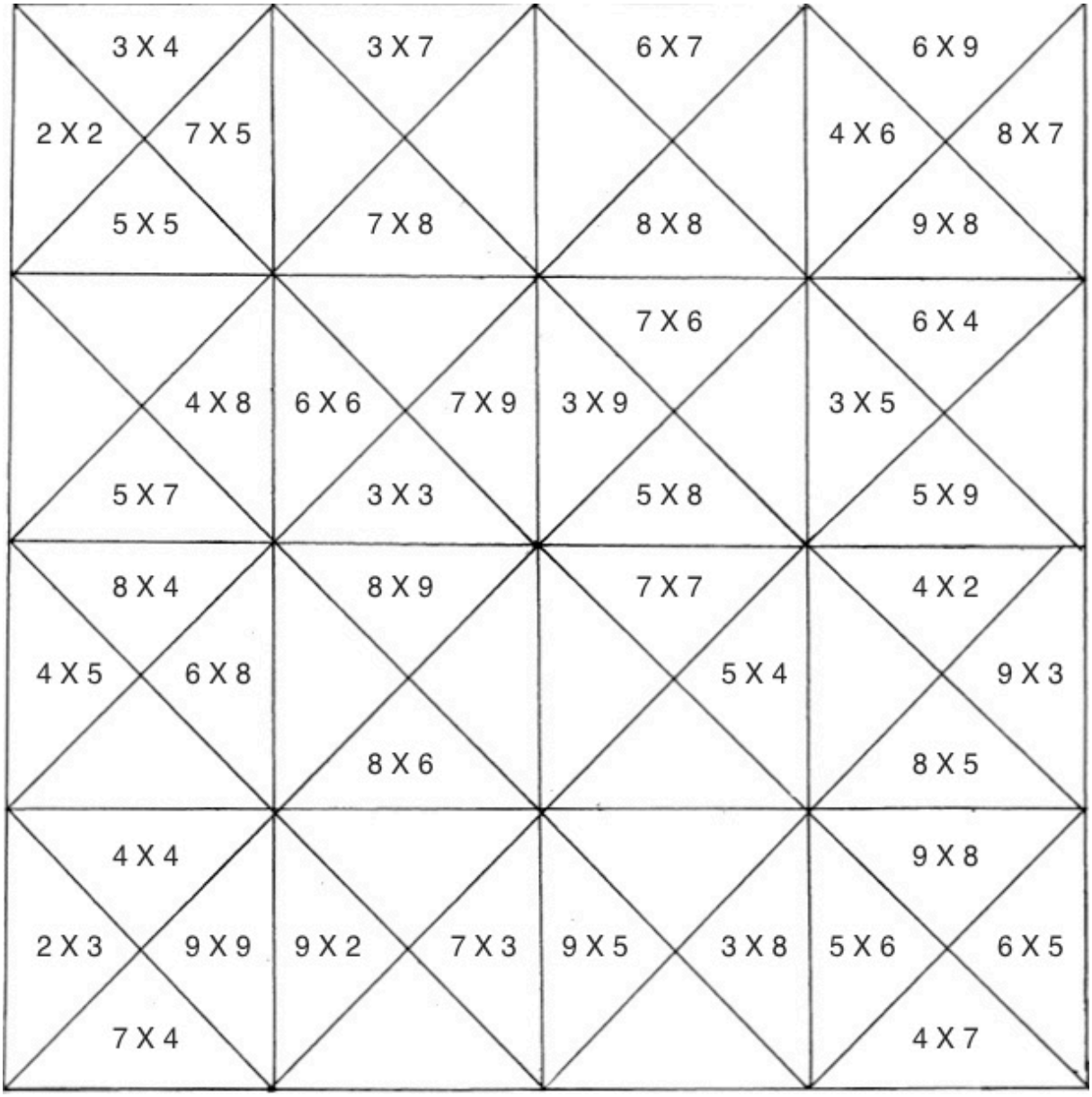
Fargelegg så mange ruter som mulig. Tallene representerer ulike farger. Fargeintervallene er angitt nedenfor bildet. Blanke felt skal ikke fargelegges.



Fargeintervall:
1-15 = Blå
16-30 = Brun
31-45 = Grønn
46-60 = Rød
61-75 = Svart
76-90 = Rosa

Ark 2

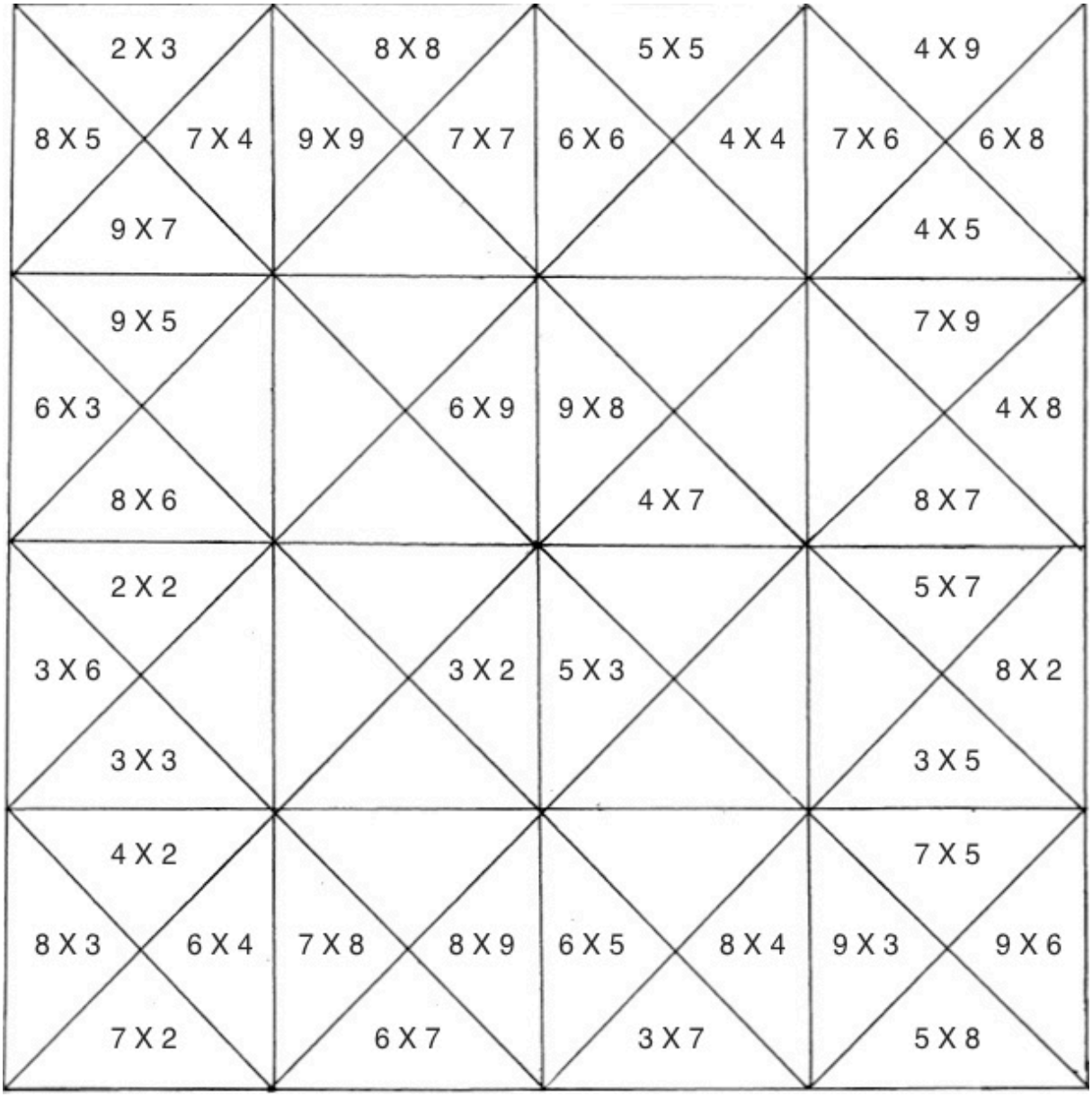
Fargelegg så mange ruter som mulig. Tallene representerer ulike farger. Fargeintervallene er angitt nedenfor bildet. Blanke felt skal ikke fargelegges.



Fargeintervall:
1-15 = Grønn
16-30 = Lilla
31-45 = Blå
46-60 = Svart
61-75 = Rød
76-90 = Rosa

Ark 3

Fargelegg så mange ruter som mulig. Tallene representerer ulike farger. Fargeintervallene er angitt nedenfor bildet. Blanke felt skal ikke fargelegges.



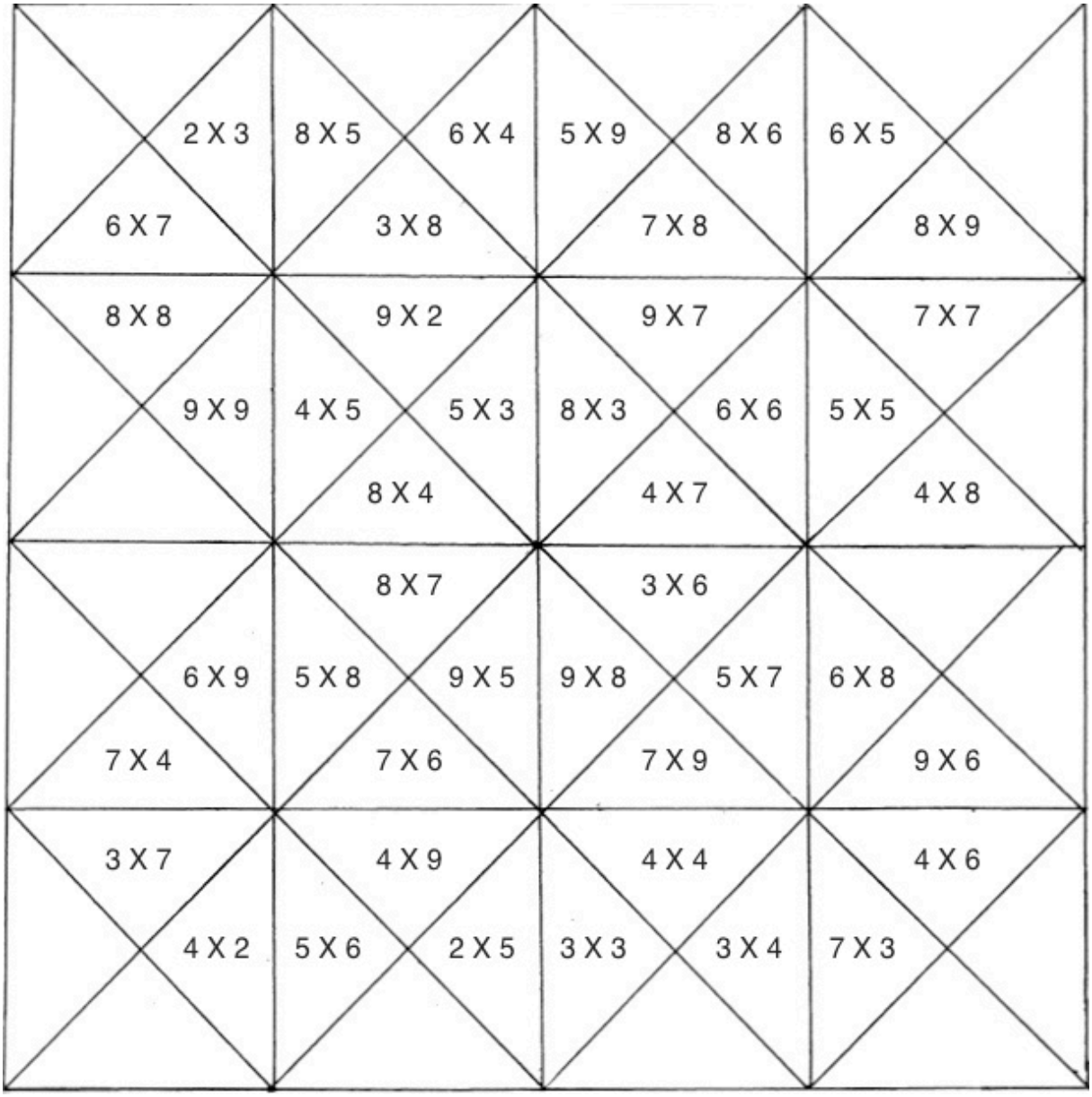
Fargeintervall:
1-15 = Lysegrønn
16-30 = Lyseblå
31-45 = Svart
46-60 = Blå
61-75 = Rød
76-90 = Gul

DEL 3

VENNLIGST IKKE SNU ARKET FØR DU BLIR BEDT OM DET

Ark 1

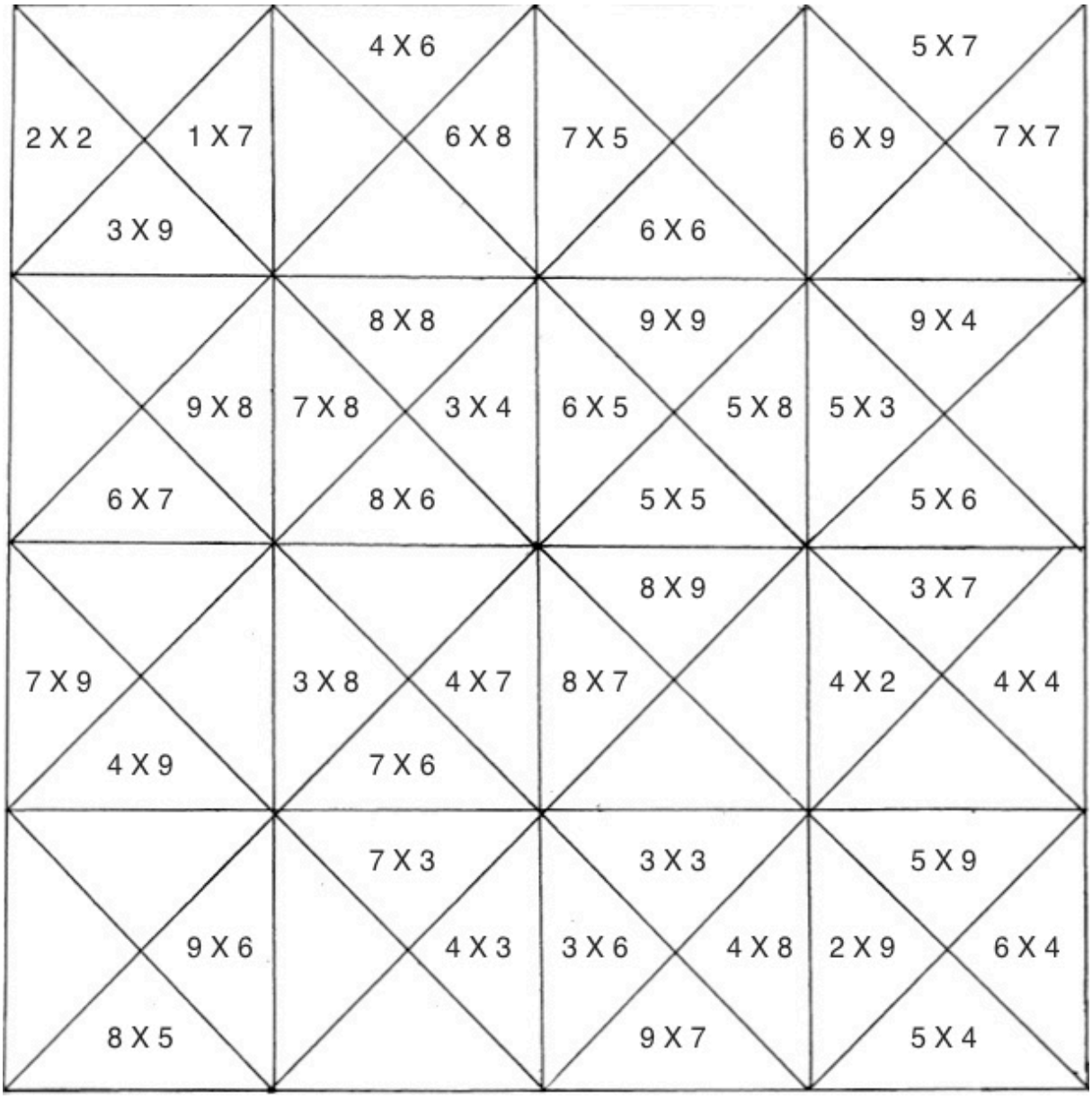
Fargelegg så mange ruter som mulig. Tallene representerer ulike farger. Fargeintervallene er angitt nedenfor bildet. Blanke felt skal ikke fargelegges.



Fargeintervall:
1-15 = Blå
16-30 = Oransje
31-45 = Grønn
46-60 = Rød
61-75 = Svart
76-90 = Gul

Ark 2

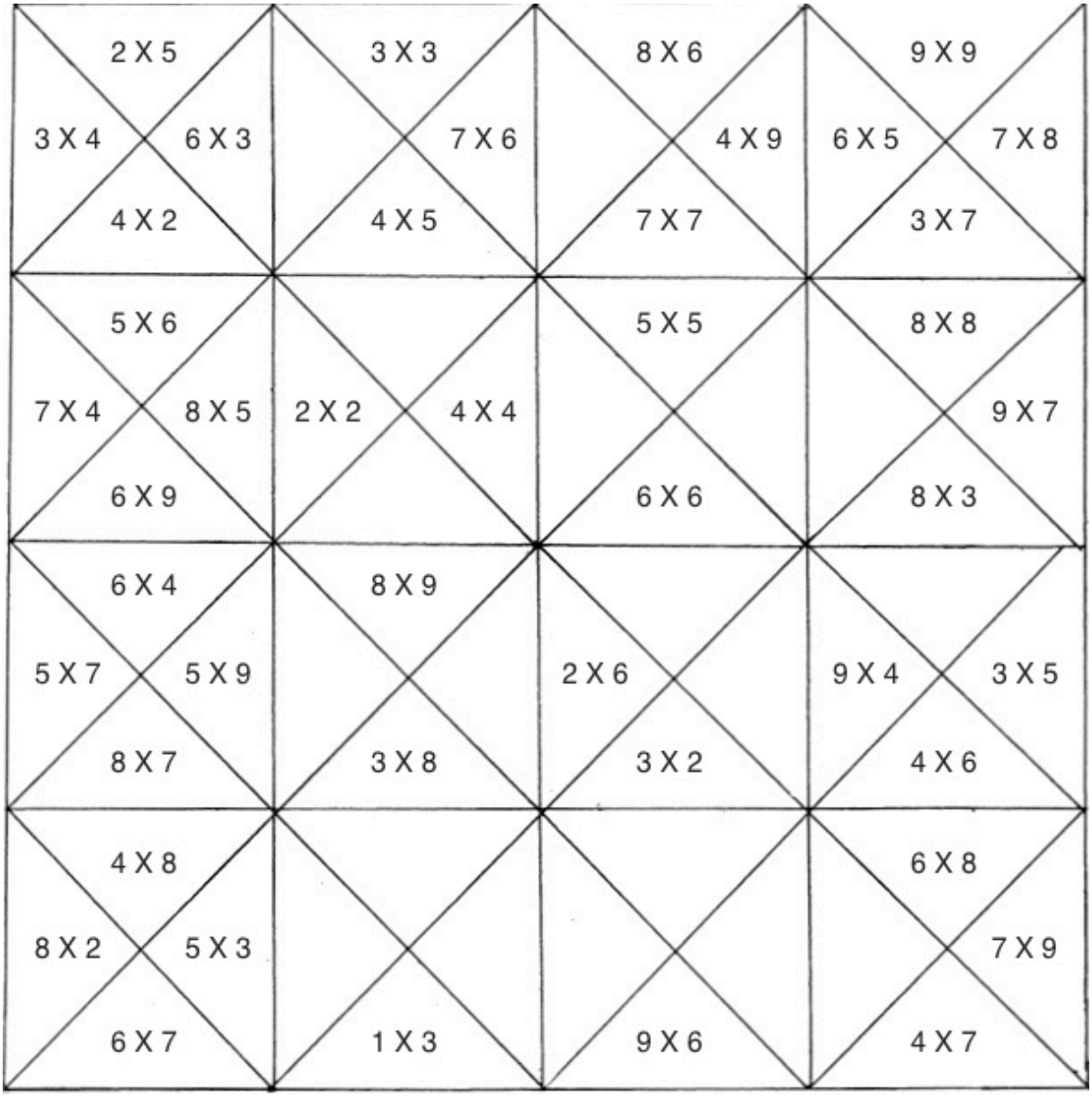
Fargelegg så mange ruter som mulig. Tallene representerer ulike farger. Fargeintervallene er angitt nedenfor bildet. Blanke felt skal ikke fargelegges.



Fargeintervall:
1-15 = Lilla
16-30 = Brun
31-45 = Lyseblå
46-60 = Lysegrønn
61-75 = Gul
76-90 = Rosa

Ark 3

Fargelegg så mange ruter som mulig. Tallene representerer ulike farger. Fargeintervallene er angitt nedenfor bildet. Blanke felt skal ikke fargelegges.



Fargeintervall:
1-15 = Blå
16-30 = Grønn
31-45 = Svart
46-60 = Lilla
61-75 = Rød
76-90 = Gul

DEL 4

SPØRRESKJEMA

VENNLIGST IKKE SNU ARKET FØR DU BLIR BEDT OM DET

Kjønn

O Kvinne

O Mann

Alder

Studieretning _____

På spørsmålene nedenfor, vennligst sett ring rundt svaret som passer best for deg.

1 = i svært liten grad, 3 = i middels grad, 5 = i svært stor grad.

---

1. I hvilken grad var du motivert til å prestere godt i dette eksperimentet? 1 2 3 4 5

---

2. I hvilken grad påvirket din kompensasjonsordning din motivasjon? 1 2 3 4 5

---

3. I hvilken grad påvirket tilbakemeldingen du fikk i pausen din motivasjon i periode 2? 1 2 3 4 5

---

4. I hvilken grad setter du pris på tilbakemelding på ditt arbeid i jobbsituasjon? 1 2 3 4 5

Takk for at du deltok i dette eksperimentet!

Appendix 4: Mann Whitney U-test Mistakes and Feedback

Ranks

	Feedback	N	Mean Rank	Sum of Ranks
Mean_mstks	Objective Feedback	57	52.36	2984.50
	Subjective Feedback	56	61.72	3456.50
	Total	113		

Test Statistics[a]

	Mean_mstks
Mann-Whitney U	1331.500
Wilcoxon W	2984.500
Z	-1.520
Asymp. Sig. (2-tailed)	.128

a. Grouping Variable: Feedback

Ranks

	Feedback	N	Mean Rank	Sum of Ranks
Prct_mstks_P1	Objective Feedback	57	56.06	3195.50
	Subjective Feedback	56	57.96	3245.50
	Total	113		
Prct_mstks_P2	Objective Feedback	57	51.89	2958.00
	Subjective Feedback	56	62.20	3483.00
	Total	113		
Prct_mstks_P3	Objective Feedback	57	53.11	3027.50
	Subjective Feedback	56	60.96	3413.50
	Total	113		

Test Statistics[a]

	Prct_mstks_P1	Prct_mstks_P2	Prct_mstks_P3
Mann-Whitney U	1542.500	1305.000	1374.500
Wilcoxon W	3195.500	2958.000	3027.500
Z	-.317	-1.709	-1.286
Asymp. Sig. (2-tailed)	.752	.087	.198

a. Grouping Variable: Feedback

Appendix 5: Mann Whitney U-test Mistakes and Incentive Schemes

Ranks

	Compensation	N	Mean Rank	Sum of Ranks
Mean_mstks	Fixed pay	59	53.71	3169.00
	Performance pay	54	60.59	3272.00
	Total	113		

Test Statistics[a]

	Mean_mstks
Mann-Whitney U	1399.000
Wilcoxon W	3169.000
Z	-1.116
Asymp. Sig. (2-tailed)	.264

a. Grouping Variable: Compensation

Ranks

	Compensation	N	Mean Rank	Sum of Ranks
Prct_mstks_P1	Fixed pay	59	57.11	3369.50
	Performance pay	54	56.88	3071.50
	Total	113		
Prct_mstks_P2	Fixed pay	59	55.06	3248.50
	Performance pay	54	59.12	3192.50
	Total	113		
Prct_mstks_P3	Fixed pay	59	54.09	3191.50
	Performance pay	54	60.18	3249.50
	Total	113		

Test Statistics[a]

	Prct_mstks_P1	Prct_mstks_P2	Prct_mstks_P3
Mann-Whitney U	1586.500	1478.500	1421.500
Wilcoxon W	3071.500	3248.500	3191.500
Z	-.038	-.673	-.997
Asymp. Sig. (2-tailed)	.969	.501	.319

a. Grouping Variable: Compensation

Appendix 6: Descriptive Statistics on Mistakes

Descriptive Statistics

Feedback		N	Minimum	Maximum	Mean	Std. Deviation
Objective Feedback	Prct_mstks_P1	57	0	16	3.72	4.065
	Prct_mstks_P2	57	0	54	4.39	8.095
	Prct_mstks_P3	57	0	81	7.05	15.694
	Mean_mstks	57	.00	46.00	5.0526	7.53215
	Feedback	57	0	0	.00	.000
	Valid N (listwise)	57				
Subjective Feedback	Prct_mstks_P1	56	0	69	6.07	10.958
	Prct_mstks_P2	56	0	89	7.13	13.050
	Prct_mstks_P3	56	0	93	8.95	18.064
	Mean_mstks	56	.00	82.00	7.3810	12.29115
	Feedback	56	1	1	1.00	.000
	Valid N (listwise)	56				

Descriptive Statistics

Compensation		N	Minimum	Maximum	Mean	Std. Deviation
Fixed pay	Compensation	59	0	0	.00	.000
	Prct_mstks_P1	59	0	31	4.19	5.396
	Prct_mstks_P2	59	0	54	4.90	8.104
	Prct_mstks_P3	59	0	93	7.98	18.651
	Mean_mstks	59	.00	46.00	5.6893	8.62343
	Valid N (listwise)	59				
Performance pay	Compensation	54	1	1	1.00	.000
	Prct_mstks_P1	54	0	69	5.65	10.583
	Prct_mstks_P2	54	0	89	6.67	13.282
	Prct_mstks_P3	54	0	88	8.00	14.836
	Mean_mstks	54	.00	82.00	6.7716	11.73052
	Valid N (listwise)	54				