# The Effect of Reward and Punishment Schemes on Performance: Risk Preference as Moderator among Accounting Students

ANDI IRFAN[1], SALMA NAURA ISMAIL[2], DESVI EMTI[3]
[1]Department of Accountancy Diploma III, Faculty of Economic and Social Sciences
Universitas Islam Negeri Sultan Syarif Kasim Riau
INDONESIA

[2]Undergraduates Student on Department of Accountancy, Faculty of Economic and Social Sciences
Universitas Islam Negeri Sultan Syarif Kasim Riau
INDONESIA

[3]Department of Management, Faculty of Economic and Social Sciences
Universitas Islam Negeri Sultan Syarif Kasim Riau
INDONESIA

*Abstract:* - This study examines the effects of reward and punishment schemes on performance and investigates whether risk preference moderates these relationships. The study used a 3 × 2 between-subjects factorial design involving undergraduate accounting students with an auditing concentration at Universitas Islam Negeri Sultan Syarif Kasim Riau. The three compensation conditions were reward, neutral, and punishment, while risk preference was classified as risk taker or risk averse. 74 students participated initially, and 52 observations remained after the manipulation check. The hypotheses were examined using analysis of variance (ANOVA). The reported results indicate a significant main effect of the reward condition on performance ($F = 5.443$, $p = .025$) and a significant main effect of risk preference ($F = 8.313$, $p = .006$). The punishment condition was not statistically significant ($F = 1.565$, $p = .221$), and neither the reward x risk-preference interaction ($F = 0.452$, $p = .506$) nor the punishment x risk-preference interaction ($F = 1.130$, $p = .297$) was statistically significant. The findings therefore provide evidence of differences associated with reward framing and risk preference in the reported sample, but not for the two interaction effects.

*Key-Words:* - Reward, Punishment, Performance, Risk Preference, Experimental Accounting.

Received: March 17, 2026. Revised: April 14, 2026. Accepted: May 18, 2026. Published: November 11, 2026.

## 1. Introduction

Performance is a central concern in organizations because incentive systems can shape how individuals allocate effort toward assigned tasks. Compensation is one mechanism through which organizations communicate expected performance and consequences. In experimental settings, the way an incentive is framed may matter even when the underlying economic value is similar. Prior experimental research has examined differences between bonus and penalty frames and has shown that contract framing can influence effort and preferences [1], [2], [3].

Risk preference is another individual characteristic relevant to decisions under uncertainty. Individuals differ in their willingness to accept uncertain outcomes, and these differences can affect responses to incentive contracts. Zaleskiewicz [4], for example, distinguishes instrumental and stimulating forms of risk taking, indicating that risk-taking behavior is not necessarily driven by a single motive. Holt and Laury [5] likewise show that risk attitudes are relevant to responses to incentives.

Previous Indonesian studies have also reported relationships between reward, punishment, and employee performance [6], [7], [8], [9]. Purnama [10] examined a compensation scheme involving fines and treated risk preference as a moderating variable. The present study extends that experimental logic by comparing reward, neutral, and punishment conditions while retaining risk preference as a potential moderator.

The research question is therefore whether compensation framing is associated with performance and whether the relationship differs according to risk

preference. The contribution of the study is empirical rather than purely theoretical: it applies a 3 × 2 factorial experiment to accounting students acting in a junior-auditor scenario. The Research Justification section also formalizes the participant's decision problem to clarify the direction of the hypothesized effects.

The study tests five hypotheses:

H1: Participants in the reward condition exhibit higher performance than participants in the neutral and punishment conditions.

H2: Participants in the punishment condition do not exhibit higher performance than participants in the neutral condition.

H3: Risk preference is associated with performance, with risk takers exhibiting higher performance than risk-averse participants.

H4: Risk preference moderates the relationship between the reward condition and performance.

H5: Risk preference moderates the relationship between the punishment condition and performance.

## 2. Research Justification

A reward is the provision of additional income to someone who has met certain requirements. In the study by Brink and Rankin, respondents tended to prefer compensation schemes that combine bonuses and penalties of equal economic value over other forms of compensation [2].

If an employee is offered a reward at the beginning of a predetermined agreement, performance tends to improve. A larger reward is associated with higher morale and stronger motivation to earn it.

According to Hannan et al., in their study "Bonus versus Penalty: Does Contract Frame Affect Employee Effort?," a penalty compensation scheme affects performance because the greater the fine, the higher the employee's performance [3]. If an employee is subject to a penalty from the beginning of a predetermined agreement, performance tends to improve, because employees are motivated to meet company targets in order to avoid the penalty.

Risk preference, which differs across individuals, influences employee performance [5]. According to Brink and Rankin, risk takers tend to perform better because the risk of failure motivates them [2]. However, risk-averse employees tend to perform worse because they dislike taking risks in any decision. Employees with a high risk appetite, in contrast, tend to perform better. If performance and risk preference are strongly related, both factors need to be considered; for example, an employee with low performance may perceive himself or herself as facing a high risk.

The study by Brink and Rankin indicates that rewards influence performance [2]. Compensation in the form of rewards improves employee performance. In this study, risk preference was added as a moderating variable to test whether it can strengthen or weaken the effect of rewards on performance. A reward contract provides salary certainty: an employee who completes the tasks or exceeds the targets is promised a reward, whereas an employee who fails to reach the targets still receives the base salary.

Individuals who are afraid of making mistakes and have a low risk appetite tend to prefer a compensation system that requires less effort to achieve, such as a scheme that includes rewards and penalties for failure [2]. In this study, risk preference was included to determine whether this variable would weaken or strengthen the effect of punishment on performance. According to research by Brink and Rankin, risk takers are investors who are willing to take opportunities with high levels of risk [2]. Risk-averse employees tend to weaken the effect of punishment on performance because they dislike taking risks in any decision, even though failing to do so results in a fine paid to the company. High risk takers, in contrast, tend to improve performance even when a penalty is imposed at the beginning of the period, whereas risk-averse employees tend to perform worse because they are reluctant to take risks.

### 2.1 Formal Model of Effort under Alternative Incentive Frames

To make the hypothesized behavioral dynamics explicit, this section formalizes the decision problem of a participant acting as a junior auditor. The model is used only to derive the direction of the effects tested in H1 to H5; it does not generate numerical predictions.

Setup. A participant chooses effort $e$ in [0, 1]. Effort determines the probability $e$ of meeting the performance target and costs $c(e) = k\, e^2/2$, with $k > 0$; intrinsic motivation adds a constant marginal benefit $m \geq 0$. Outcomes are valued relative to a reference point by the function in Eq. (1), where $\rho > 0$ captures risk preference (a larger $\rho$ describes a stronger preference for risk, so risk takers have a higher $\rho$ than risk-averse participants, following a power specification of the kind used for risk attitudes in [5]) and $\lambda > 1$ captures loss aversion [2], [3]. Monetary stakes are expressed in units such that the bonus $b > 1$ and the fine $f > 1$, and $\varphi$ in (0, 1] is the perceived credibility of the stated consequence. In the reward frame the reference point is the base pay and success yields the bonus; in the penalty frame the reference point is the full pay and failure incurs the fine; in the neutral frame there is no contingent pay.

$$v(x) = x^{\rho} \text{ for } x \geq 0; \quad v(x) = -\lambda(-x)^{\rho} \text{ for } x < 0 \quad (1)$$

$$e_N^* = m / k \quad (2)$$

$$e_R^* = (\varphi\, b^{\rho} + m) / k \quad (3)$$

$$e_P^* = (\varphi\, \lambda\, f^{\rho} + m) / k \quad (4)$$

Equations (2) to (4) give the optimal effort in the neutral, reward, and penalty frames, respectively. They are obtained by maximizing $V_N(e) = m\, e - k\, e^2/2$, $V_R(e) = e\, \varphi\, v(b) + m\, e - k\, e^2/2$, and $V_P(e) = -(1 - e)\, \varphi\, \lambda\, f^{\rho} + m\, e - k\, e^2/2$ over $e$. Each objective is strictly concave in $e$ because $k > 0$, so the first-order condition identifies a unique solution, which is interior when $\varphi\, b^{\rho} + m \leq k$ and $\varphi\, \lambda\, f^{\rho} + m \leq k$.

Proposition 1 (reward effect, H1). $e_R^* - e_N^* = \varphi\, b^{\rho} / k > 0$. Differencing (3) and (2) shows that the reward frame raises effort relative to the neutral frame. (5)

Proposition 2 (risk preference, H3). $\partial e_R^* / \partial \rho = \varphi\, b^{\rho}\, ln(b) / k > 0$ for $b > 1$, whereas $\partial e_N^* / \partial \rho = 0$. Because $d(x^{\rho})/d\rho = x^{\rho} \ln(x)$, participants with a higher ρ (risk takers) exert more effort under a reward contract. (6)

Proposition 3 (moderation, H4). $\partial(e_R^* - e_N^*)/\partial \rho = \varphi\, b^{\rho}\, ln(b) / k > 0$. The effect of the reward frame is larger for participants with a higher ρ. (7)

Proposition 4 (penalty frame, H2 and H5). $e_P^* - e_N^* = \varphi\, \lambda\, f^{\rho} / k > 0$ and $\partial(e_P^* - e_N^*)/\partial \rho = \varphi\, \lambda\, f^{\rho} \ln(f) / k > 0$ for $f > 1$. If $b = f$ (economic equivalence), $e_P^* - e_R^* = \varphi(\lambda - 1)\, b^{\rho} / k > 0$. (8)

Remark. Every effect in Eqs. (5) to (8) is proportional to φ. When the incentive is hypothetical, as in the present experiment where consequences were described in a case and participants received only a souvenir, φ may be small, so the predicted effects shrink toward zero and the interaction effects, which are of second order, become hard to detect in small samples. The model also implies a positive penalty effect when the penalty is credible, which differs from the wording of H2; H2 is therefore tested here as a conservative hypothesis that the penalty frame does not outperform the neutral frame in this setting.

## 3. Research Methods

This study employed a 3 x 2 between-subjects design. The experiment was designed using three compensation conditions crossed with two risk-preference groups. Risk preference was measured rather than manipulated, so the number of participants in each risk-preference group was not controlled by design (see Tables 3 and 4). The subjects were students from the Faculty of Economics and Social Sciences of UIN Suska Riau, Accounting Study Program, with a concentration in Accounting and Auditing. The use of accounting students as surrogate auditors was feasible because they could assume the role of junior auditors. Junior auditor positions in public accounting firms are typically filled by fresh graduates, thus considered to have adequate skills for the experiment Lee and Welker, as cited in Nahartyo [11]. This was also taken into consideration, as most undergraduate accounting students with a concentration in auditing are prospective auditors.

Table 1. 3 x 2 Between-Subjects Experimental Design

| | | Compensation Scheme | | |
|---|---|---|---|---|
| | | Reward | Neutral | Punishment |
| Risk Preference | Risk Taker | 1 | 2 | 3 |
| | Risk Averse | 4 | 5 | 6 |

*Source: Processed Data, 2022*

Note: Cells 1 to 6 are the experimental conditions formed by crossing the three compensation schemes with the two risk-preference groups.

The criteria for the subjects in this study were students with a GPA above 3.00 due to their theoretical understanding of the work of prospective auditors, students who had taken Audit 1 and 2, Management Accounting, and Strategic Management with a minimum grade of B, students who had completed a minimum of 100 Semester Credit Units, and students enrolled in the class of 2020.

The performance instrument used in this study was an adaptation of the experimental task instrument developed in Libby's research in Chow [12], where symbols were translated into letters of the alphabet using a translation key. Translating codes and numbers into letters is a simple task that can be performed by anyone [11]. The reward and punishment instrument used was an instrument adapted from Prabowo [13] and Wibisono [14], which was a development of the research instrument created by Xu and Ziegenfuss [15]. Based on this research instrument, the researcher created a modified instrument by providing case stories appropriate to the group. The risk preference instrument in this study uses an adaptation of an instrument developed by Zaleskiewicz [4], which differentiates between stimulating and instrumental risk taking based on the underlying motives for taking risks. Stimulating risk taking is characterized by behavior that is quicker, easier, and more automatic. Instrumental risk taking, on the other hand, is more achievement-oriented and relies on more complex information processing. This instrument has been tested for validity and reliability, as shown by the numerous researchers who have used it to measure risk preference.

Data and instrument originality. The experimental data reported in this paper were collected specifically

for this study, and the instruments are adapted from the sources cited above.

To test the instrument's validity, the researcher first conducted a pilot test on the instrument to be used in the experiment. The pilot test subjects were undergraduate students from the Accounting program at UIN Suska Riau, Faculty of Economics and Social Sciences. The pilot test was carried out to determine whether the cases or treatments were understandable to the subjects and to identify potential design errors. A reliability test was then conducted to determine whether the measurement instrument produces consistent results.

The research hypothesis was tested using analysis of variance (ANOVA) to determine the influence of risk preference and the influence of rewards and punishments on performance.

In this experimental study, the subject played the role of a Junior Auditor at the Ahmad Public Accounting Firm. You have been entrusted by the Public Accounting Firm to conduct a financial audit. You have a partner named Mr. Radit, who is your superior. While conducting the audit, you discover an unusual expense amounting to Rp200,000,000 (two hundred million rupiah). You then question Mr. Radit about the expense. He explains that the expense was used to purchase office supplies. Mr. Radit also states that the invoice for the expense has been lost. Unsatisfied with Mr. Radit's response, you investigate the expense.

After investigation, it was discovered that the Rp. 200,000,000 (two hundred million rupiah) was not actually used to purchase office supplies, but was instead used by Mr. Radit for his personal needs. When Mr. Radit learned of your discovery, he immediately paid Rp. 20,000,000 (twenty million rupiah) to avoid reporting his actions to the Managing Partner of KAP Ahmad. Regarding KAP Ahmad, you should be aware that KAP Ahmad Yani has a Standard Operating Procedure (SOP) for its employees that supports its anti-fraud policy. Those who assist and combat fraud will receive a commission based on the agreed-upon total services rendered. Those who refuse and choose not to intervene, believing the fraud to be non-existent, face no risk. Furthermore, those who support the fraud will be subject to penalties in the form of salary deductions or even termination of employment, as well as confiscation of the fraud proceeds.

This experimental research was conducted on undergraduate students majoring in Accounting, majoring in Auditing, at the Faculty of Economics and Social Sciences, UIN Suska Riau. The researcher provided an explanation of the case, including demographic data, action decision columns, and manipulation checks before participants completed the task. Participants were instructed to pay close attention to the instructions in the case. Participants were given approximately 30 minutes to complete each case. After completing the case, the researcher immediately checked their answers to determine whether they had passed.

In the next task, subjects were asked to read and understand the procedures provided. Before completing the actual experimental task, subjects were required to practice recognizing the case. To maintain calm in the room, subjects were asked to raise their hands if they did not understand the purpose of the experimental instrument, and the researcher would then approach them directly. After all subjects received the experimental instrument, a link to the Google Form from the first stage to the last, the researcher then read out the roles and cases in the experiment. Subjects were also asked to carefully read the information and instructions for the case together with the researcher. After completing all stages of the experimental task, the researcher asked the subjects to provide feedback and suggestions. Based on this feedback, the researcher made the necessary improvements. The researchers hoped that the questions raised by the subjects during the pilot test would not recur and that participants would pass the manipulation check, so that the experiment would run smoothly with less bias in the results. Each subject received a souvenir as a token of appreciation for their participation.

Before conducting this experiment on actual participants, a pilot test was conducted. The pilot test aimed to determine whether the experimental case is understandable to the participants. The pilot test was conducted online via Google Meet on April 3, 2023, at 11:00 a.m. WIB. The subjects were 10 undergraduate Accounting students from the Faculty of Economics and Social Sciences, class of 2018, who had taken Audit 1&2, Management Accounting, and Strategic Management and had a GPA of 3.00 or higher. The subjects in this pilot test were not real experimental subjects, but the pilot test environment was designed to mimic the actual experiment. Of the 10 subjects who participated in the pilot test, only 5 passed the manipulation check and were included in further data processing.

In the pilot test, the subjects consisted of 1 male and 9 female students. The Cronbach's alpha of 0.931 for the 11 items (Table 2 (Appendix)) indicates high internal consistency of the instrument. Cronbach's alpha does not measure comprehension; comprehension was assessed through the manipulation check, which only 5 of the 10 pilot subjects passed. During the pilot test,

some subjects still had questions about the experimental instrument, including the possibility of bonuses and fines. After the pilot test, the researcher asked the subjects for criticism and suggestions and made improvements where needed. The researchers hoped that the questions raised during the pilot test would not recur and that participants would pass the manipulation check, so that the experiment would run smoothly with less bias in the results.

# 4. Results and Discussion

## 4.1 Descriptive Analysis

The study was conducted with 74 subjects, undergraduate Accounting students majoring in Auditing at UIN Suska Riau. Of these 74 subjects, 22 failed the manipulation check, leaving 52 subjects for data processing.

Table 3. Descriptive Statistical Test

| Description | | Frequency | Min | Max |
|---|---|---|---|---|
| Gender | Male | 12 | 5 | 10 |
| | Female | 40 | 5 | 10 |
| | Total | 52 | 5 | 10 |
| Age | 20-22 Years | 51 | 5 | 10 |
| | >22 Years | 1 | 9 | 9 |
| | Total | 52 | 5 | 10 |
| Concentration | Accounting Audit | 52 | 10 | 10 |
| GPA | Below 3.00 | 3 | 9 | 10 |
| | Above 3.00 | 49 | 5 | 10 |
| | Total | 52 | 5 | 10 |
| Group | Reward | 20 | 6 | 10 |
| | Neutral | 19 | 5 | 10 |
| | Punishment | 13 | 7 | 10 |
| | Total | 52 | 5 | 10 |

*Source: Processed Data, 2023*

Note: Min and Max are the lowest and highest performance scores observed within each category; they are not demographic values.

Based on the demographic table above, it can be seen that the subjects consisted of 12 males and 40 females. In both groups, the minimum and maximum performance scores were 5 and 10, respectively. Of these subjects, the majority were aged 20 to 22 (51), while only one was over 22. All subjects were from the Accounting Audit concentration. In addition, 49 participants had a GPA above 3.00, while 3 had a GPA below 3.00. These three participants did not meet the GPA criterion stated in the Research Methods. They passed the manipulation check and therefore remained among the 52 analyzed observations; this departure from the inclusion criterion is acknowledged as a limitation. By compensation scheme, there were 20 participants in the reward group, 19 in the neutral group, and 13 in the punishment group.

Table 4. Group Statistics Risk Preference

| Risk Preference | N | Mean | Std. Deviation |
|---|---|---|---|
| Risk Averse | 42 | 7.70 | 1.337 |
| Risk Taker | 10 | 9.00 | 1.269 |
| **Total** | **52** | **8.75** | **1.370** |

*Source: Processed Data, 2023*

Table 4 shows that mean performance was 7.70 for risk-averse participants (n = 42) and 9.00 for risk-taker participants (n = 10). The groups are markedly unbalanced, with risk-averse participants outnumbering risk takers by about four to one. The imbalance arises because risk preference was measured rather than manipulated, so group sizes could not be controlled by assignment. Small and unequal cells reduce statistical power, especially for the interaction tests, and make ANOVA more sensitive to unequal variances. The risk-preference results should therefore be interpreted with caution.

Table 5. Levene's Test

| Performance | Levene's Test for Equality of Variances | |
|---|---|---|
| | F Test | Significant |
| Based on Mean | 1.839 | 0.183 |

*Source: Processed Data, 2023*

Levene's test (based on the mean) was not significant ($F = 1.839$, $p = 0.183$), so the null hypothesis of equal error variances across the experimental groups is not rejected and the homogeneity-of-variance assumption of ANOVA is met. Because the groups are unbalanced (Table 4) and some cells are small, this test has limited power, so the result indicates that no heterogeneity was detected rather than proving that the variances are equal.

## 4.2 Hypothesis Testing

The experiment used six cells in a 3 x 2 between-subjects design, and each participant received one compensation scheme. Table 6 reports the F-tests for the effect of each term on performance.

Table 6. Tests of Between-Subjects Effects

| Dependent Variable : Performance | | | | | |
|---|---|---|---|---|---|
| **Variable** | **F-Value** | **df** | **p** | **Partial $\eta^2$** | **Information** |
| *Reward* | 5.443 | 1 | 0.025 | 0.120 | H1 Accepted |
| *Punishment* | 1.565 | 1 | 0.221 | 0.038 | H2 Rejected |
| *Risk Preference* | 8.313 | 1 | 0.006 | 0.172 | H3 Accepted |
| *Reward x Risk Preference* | 0.452 | 1 | 0.506 | 0.011 | H4 Rejected |
| *Punishment x Risk Preference* | 1.130 | 1 | 0.297 | 0.027 | H5 Rejected |

*Source: Processed Data, 2023*

Note: df = degrees of freedom of the tested effect and of the error term; partial $\eta^2$ = effect size.

Based on the test results, two hypotheses were accepted and three hypotheses were rejected. The reward compensation scheme and risk preference had significant effects on performance at the 0.05 level, whereas the punishment scheme and the two interaction terms did not. The results of the hypothesis test showed that the reward compensation scheme tends to improve performance. The F-value was 5.443 with a significance level of 0.025. The mean performance in the reward compensation scheme was greater than the mean performance in the neutral and punishment compensation schemes. The bonus therefore appears to have motivated participants to perform well. This result is also consistent with Wirawan's finding that employees tend to maximize performance under reward agreements [16].

The results of the hypothesis test showed that the punishment compensation scheme tends not to improve performance. The F-value was 1.565 with a significance level of 0.221. The mean performance in the punishment compensation scheme was lower than the mean performance in the reward compensation scheme and higher than the mean performance in the neutral compensation scheme. The results of this study align with those of Brink and Rankin [2], who found that fine compensation schemes had no positive impact on performance. This was demonstrated by a significance value exceeding 0.05. These results also support research by Purnama [10], which demonstrated that fine schemes have a negligible impact on performance.

The test results indicate that H3 is accepted, as a high risk preference tends to improve performance. The calculated F-value for the test results was 8.313, with a significance value of 0.006. The significance value for the third hypothesis is less than 0.05. Performance in the high risk preference group (risk takers) had a mean value of 9.00. These results align with research by Purnama [10], who studied risk preference in a fine-based compensation scheme. Employees who enjoy challenges or prefer risks are expected to perform well, which supports H3. This interpretation does not rely on the endowment effect.

The test results indicate that H4 is rejected because a high risk preference and the promise of rewards do not improve performance compared to a low risk preference. The calculated F-value for the test results is 0.452, with a significance level of 0.506. The significance level for this fourth hypothesis is greater than 0.05. The mean performance in this reward compensation scheme is 8.90. The performance in the reward compensation scheme group for risk takers has a mean value of 9.13, while the risk averse group has a mean value of 8.00. The mean value in the reward compensation scheme group for risk takers is higher than that for risk averse groups.

The test results indicate that H5 is rejected because a high risk preference and the promise of punishment do not improve performance compared to a low risk preference. The calculated F-value for the test results is 1.130, with a significance level of 0.297. The significance level for this fifth hypothesis is greater than 0.05. The mean performance in the punishment compensation scheme group is 8.62. Performance in the punishment compensation scheme group for risk takers had a mean value of 8.78, while for risk-averse participants it was 8.25. The mean value in the punishment compensation scheme for risk takers was higher than for risk averse.

Both interaction terms test whether the effect of the compensation scheme depends on risk preference. Neither was significant. Because only 10 participants were risk takers (Table 4), the cells that contain risk takers are very small and these tests have low power, so the nonsignificant interactions should not be read as evidence that moderation is absent.

## 4.3 Discussion

### 4.3.1 Rewards Condition and Performance

The reported reward effect is statistically significant (F = 5.443, p = 0.025). The original descriptive comparison indicates that mean performance in the reward condition exceeded the means reported for the neutral and punishment conditions. Within this experiment, the result is consistent with the interpretation that a bonus frame can be associated with higher task performance.

This finding is compatible with prior incentive-contract research, although the literature is not uniform. Hannan, Hoffman, and Moser [3] reported higher effort under a penalty frame in their experiment, attributing

the result to loss aversion. Brink and Rankin [2] examined risk preference and loss aversion under bonus, penalty, and combined contract frames. These studies indicate that incentive framing and individual differences can jointly shape behavior. The present finding should therefore be interpreted as sample- and task-specific rather than as evidence that rewards universally outperform penalties.

#### 4.3.2 Punishment Condition and Performance

The punishment main effect is not statistically significant ($F = 1.565$, $p = 0.221$). The reported punishment-group mean was lower than the reward-group mean and higher than the neutral-group mean. The descriptive ordering does not by itself establish a statistically reliable difference; the inferential result indicates that the punishment condition did not produce a statistically significant main effect under the reported analysis.

This result differs from some experimental evidence. Hannan et al. [3] reported greater effort under a penalty contract than under an economically equivalent bonus contract, whereas other research has found weaker or context-dependent framing effects. The difference may reflect task design, participant population, incentive magnitude, or the way the experimental manipulation was communicated. The present study cannot identify which of these mechanisms explains the observed result.

#### 4.3.3 Risk Preference and Performance

Risk preference shows a statistically significant main effect in the reported analysis ($F = 8.313$, $p = 0.006$). The risk-taker group had a mean performance score of 9.00 compared with 7.70 for the risk-averse group. The finding is descriptively consistent with the hypothesis that participants classified as risk takers performed differently from risk-averse participants in the experimental task.

The interpretation should not rely on the phrase risk-loving endowment effect. The endowment effect is a distinct behavioral concept referring to the tendency to value an owned object more highly than an equivalent object not owned. It is not synonymous with risk seeking. Zaleskiewicz [4] instead provides a framework for distinguishing different motives for economic risk taking. Accordingly, the result should be discussed in terms of risk preference rather than an endowment effect.

#### 4.3.4 Rewards × Risk Preference

The reward × risk-preference interaction is not statistically significant ($F = .452$, $p = 0.506$). The reported means were 9.13 for risk takers and 8.00 for risk-averse participants within the reward condition. Although the descriptive means differ, the nonsignificant interaction indicates that the study does not provide statistical evidence that risk preference changes the effect of the reward condition on performance.

The difference between the two group means should therefore not be presented as proof of moderation. A moderation claim requires evidence that the difference in reward effects across risk-preference groups is statistically distinguishable from zero.

#### 4.3.5 Punishment × Risk Preference

The punishment × risk-preference interaction is not statistically significant ($F = 1.130$, $p = 0.297$). The reported mean performance under punishment was 8.78 for risk takers and 8.25 for risk-averse participants. Thus, although the descriptive mean for risk takers was higher, the interaction test does not support the claim that risk preference significantly moderates the punishment–performance relationship.

#### 4.3.6 Relation to the Formal Model

The formal model predicts positive reward, risk-preference, and moderation effects (Propositions 1 to 3). The data support the first two predictions but not the moderation prediction. For the penalty frame, the model predicts a positive effect when the penalty is credible, whereas the data show no significant punishment effect. Two explanations are consistent with the model: low perceived credibility of hypothetical incentives (a small $\varphi$) and the small number of risk takers. The present data cannot test these explanations, so they are offered as directions for future research.

## 5. Conclusion

The results of the reward compensation test showed a tendency to improve performance. The results of the punishment compensation test showed a tendency to not improve performance. The results of the high risk preference test showed a tendency to improve performance. The results of the high risk preference test with the promise of a reward did not improve performance compared to the low risk preference test. The results of the high risk preference test with the promise of a punishment did not improve performance compared to the low risk preference test.

The study has limitations that qualify these conclusions. First, risk preference was measured rather than manipulated, which produced a markedly unbalanced sample (42 risk-averse and 10 risk-taking participants) and very small cells for the interaction

tests. Second, 22 of the 74 participants (about 30 percent) failed the manipulation check and were excluded, and three of the remaining 52 participants did not meet the GPA criterion. Third, the incentives were hypothetical because participants received only a souvenir, so they may have perceived the stated consequences as weakly credible. These conditions limit statistical power and generalizability, and the results should be read as exploratory.

Based on the conclusions outlined above, the researcher offers suggestions for further research, including increasing the number of samples studied in terms of both quantity and quality and expanding the research location to avoid differences in perception that could influence the research results. Furthermore, research subjects in one experimental room should ideally be given only one compensation scheme so that researchers can convey the experimental procedure more clearly and avoid bias.

**Declaration of Generative AI and AI-assisted Technologies in the Writing Process**

During the preparation of this work, the author(s) utilized artificial intelligence (AI) technologies solely for the purpose of linguistic refinement and grammatical correction. Following the application of these tools, the author(s) comprehensively reviewed, verified, and edited the resulting text. The author(s) maintain full accountability and responsibility for the accuracy, integrity, and final content of the publication.

*References:*

[1] B. K. Church, T. Libby, and P. Zhang, "Contracting frame and individual behavior: Experimental evidence," *J. Manag. Account. Res.*, vol. 20, no. 1, pp. 153–168, 2008.

[2] A. G. Brink and F. W. Rankin, "The effects of risk preference and loss aversion on individual behavior under bonus, penalty, and combined contract frames," *Behav. Res. Account.*, vol. 25, no. 2, pp. 145–170, 2013.

[3] R. L. Hannan, V. B. Hoffman, and D. V. Moser, "Bonus versus penalty: does contract frame affect employee effort?," in *Experimental Business Research: Economic and Managerial Perspectives VOLUME II*, Springer, 2005, pp. 151–169.

[4] T. Zaleskiewicz, "Beyond risk seeking and risk aversion: Personality and the dual nature of economic risk taking," *Eur. J. Pers.*, vol. 15, no. S1, pp. S105–S122, 2001.

[5] C. A. Holt and S. K. Laury, "Risk aversion and incentive effects," *Am. Econ. Rev.*, vol. 92, no. 5, pp. 1644–1655, 2002.

[6] G. D. Koencoro, "Pengaruh Reward dan Punishment Terhadap Kinerja (Survei Pada Karyawan PT. INKA (Persero) Madiun)." Brawijaya University, 2013.

[7] R. Apriyanti, K. Bahrun, and M. Finthariasari, "Pengaruh Kepemimpinan, Reward Dan Punishment Terhadap Kinerja Karyawan (Studi Kasus Pada PT. K3/SIL Ketahun Begkulu Utara)," *Entrep. Dan Manaj. Sains*, vol. 1, no. 2, pp. 189–194, 2020.

[8] Gunanto and Ikatrinasar, "Pengaruh reward dan punishment terhadap kinerja welder studi kasus di proyek pipeline spread 4 labuan maringgai-tanjung karang," *J. Inkofar*, vol. 1, no. 1, 2018, [Online]. Available: ojs.politeknikmeta.ac.id.

[9] E. S. Dymastara, "Pengaruh Reward Dan Punishment Terhadap Kinerja Karyawan Pada PT. Sandabi Indah Lestari Bengkulu Utara," *J. Enterp. Dan Manaj. Sains*, vol. 1, no. 2, pp. 160–167, 2020.

[10] I. A. Purnama, "Pengaruh Skema Kompensasi Denda Terhadap Kinerja Dengan Risk Preference Sebagai Variabel Moderating (Studi Eksperimen pada Mahasiswa S1 Fakultas Ekonomi UNY dan Mahasiswa S2 Fakultas Ekonomika dan Bisnis UGM)," *Nominal Barom. Ris. Akunt. dan Manaj.*, vol. 4, no. 1, pp. 129–145, 2015.

[11] E. Nahartyo, "Desain and implementasi riset eksperimen," *ed UPP STIM YKPN*, 2013.

[12] C. W. Chow, J. C. Cooper, and W. S. Waller, "Participative budgeting: Effects of a truth-inducing pay scheme and information asymmetry on slack and performance," *Account. Rev.*, pp. 111–122, 1988.

[13] E. PRABOWO, "PENGARUH PERSEPSI MAHASISWA AKUNTANSI MENGENAI KOMPENSASI REWARD DAN PUNISHMENT TERHADAP KECURANGAN PELAPORAN KEUANGAN DENGAN MOTIVASI."

[14] P. W. Wibisono, "Pengaruh Reward and Punishment dan Job Rotation terhadap Fraud," *Skripsi. Univ. Negeri Yogyakarta*, 2015.

[15] Y. Xu and D. E. Ziegenfuss, "Reward systems, moral reasoning, and internal auditors' reporting wrongdoing," *J. Bus. Psychol.*, vol. 22, no. 4, pp. 323–331, 2008.

[16] A. Wirawan and I. N. Afani, "Pengaruh Reward dan punishment terhadap kinerja dan motivasi karyawan pada Cv Media Kreasi Bangsa," *J. Appl. Bus. Adm.*, vol. 2, no. 2, pp. 242–257, 2018.

**Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)**

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

**Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself**

No funding was received for conducting this study.

**Conflict of Interest**

The authors have no conflicts of interest to declare that are relevant to the content of this article.

**Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)**

This article is published under the terms of the Creative Commons Attribution License 4.0
https://creativecommons.org/licenses/by/4.0/deed.en_US

# APPENDIX

Table 2. Reliability Test Reliability Statistics

| Cronbach's Alpha | N of Items |
|---|---|
| .931 | 11 |

**Correlations**

| | | P1 | P2 | P3 | P5 | P6 | P9 | P11 | P12 | P13 | P15 | P16 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| P1 | Pearson Correlation | 1 | .667* | .606 | .456 | .375 | .272 | .700* | .645* | .749* | .458 | .583 | .717* |
| | Sig. (2-tailed) | | .035 | .063 | .185 | .286 | .447 | .024 | .044 | .013 | .183 | .077 | .020 |
| | N | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| P2 | Pearson Correlation | .667* | 1 | .531 | .456 | .458 | .068 | .758* | .645* | .656* | .583 | .667* | .729* |
| | Sig. (2-tailed) | .035 | | .115 | .185 | .183 | .852 | .011 | .044 | .040 | .077 | .035 | .017 |
| | N | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| P3 | Pearson Correlation | .606 | .531 | 1 | .830** | .796** | .371 | .557 | .587 | .596 | .531 | .417 | .778** |
| | Sig. (2-tailed) | .063 | .115 | | .003 | .006 | .291 | .094 | .074 | .069 | .115 | .231 | .008 |
| | N | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| P5 | Pearson Correlation | .456 | .456 | .830** | 1 | .913** | .745* | .639* | .707* | .641* | .571 | .342 | .844** |
| | Sig. (2-tailed) | .185 | .185 | .003 | | .000 | .013 | .047 | .022 | .046 | .085 | .333 | .002 |
| | N | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| P6 | Pearson Correlation | .375 | .458 | .796** | .913** | 1 | .612 | .554 | .484 | .398 | .354 | .167 | .701* |
| | Sig. (2-tailed) | .286 | .183 | .006 | .000 | | .060 | .097 | .156 | .255 | .315 | .645 | .024 |
| | N | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| P9 | Pearson Correlation | .272 | .068 | .371 | .745* | .612 | 1 | .429 | .527 | .459 | .408 | .102 | .780* |
| | Sig. (2-tailed) | .447 | .852 | .291 | .013 | .060 | | .217 | .117 | .182 | .242 | .779 | .009 |
| | N | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| P11 | Pearson Correlation | .700* | .758* | .557 | .639* | .554 | .429 | 1 | .678* | .852** | .758* | .846** | .898** |
| | Sig. (2-tailed) | .024 | .011 | .094 | .047 | .097 | .217 | | .031 | .002 | .011 | .002 | .000 |
| | N | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| P12 | Pearson Correlation | .645* | .645* | .587 | .707* | .484 | .527 | .678* | 1 | .907** | .807** | .645* | .877** |
| | Sig. (2-tailed) | .044 | .044 | .074 | .022 | .156 | .117 | .031 | | .000 | .005 | .044 | .001 |
| | N | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| P13 | Pearson Correlation | .749* | .656* | .596 | .641* | .398 | .459 | .852** | .907** | 1 | .890** | .866** | .926** |
| | Sig. (2-tailed) | .013 | .040 | .069 | .046 | .255 | .182 | .002 | .000 | | .001 | .001 | .000 |
| | N | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| P15 | Pearson Correlation | .458 | .583 | .531 | .571 | .354 | .408 | .758* | .807** | .890** | 1 | .875** | .855** |
| | Sig. (2-tailed) | .183 | .077 | .115 | .085 | .315 | .242 | .011 | .005 | .001 | | .001 | .002 |
| | N | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| P16 | Pearson Correlation | .583 | .667* | .417 | .342 | .167 | .102 | .846** | .645* | .866** | .875** | 1 | .764* |
| | Sig. (2-tailed) | .077 | .035 | .231 | .333 | .645 | .779 | .002 | .044 | .001 | .001 | | .010 |
| | N | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| Total | Pearson Correlation | .717* | .729* | .778** | .844** | .701* | .780* | .898** | .877** | .926** | .855** | .764* | 1 |
| | Sig. (2-tailed) | .020 | .017 | .008 | .002 | .024 | .009 | .000 | .001 | .000 | .002 | .010 | |
| | N | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

**Case Processing Summary**

| | | N | % |
|---|---|---|---|
| Cases | Valid | 10 | 100.0 |
| | Excluded[a] | 0 | .0 |
| | Total | 10 | 100.0 |

a. Listwise deletion based on all variables in the procedure.