MASTERARBEIT | MASTER'S THESIS

Titel | Title

Investigating the Effect of Rewarding Effort on Delayed Gratification

verfasst von | submitted by

Mag.phil. Simone Oberhollenzer BSc MA

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2026

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 840
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Psychologie
Betreut von | Supervisor:	Univ.-Prof. Dr. Veronika Job Sutnar

Abstract[1]

Delayed gratification is a key component of self-regulation, yet it remains unclear whether reinforcing effort, rather than performance, can shift intertemporal choice. This study tested whether effort-contingent rewards increase delayed gratification compared with performance-contingent rewards. In a randomized between-subjects design, 146 University of Vienna psychology students completed an n-back task under either effort-contingent monetary rewards (indexed via pre-ejection period; PEP) or performance-contingent rewards (based on hits and misses). Participants subsequently completed an unrewarded math-effort task (MET) and a shortened Monetary Choice Questionnaire (MCQ) at baseline (pre-test) and after the task battery (post-test). Controlling for baseline MCQ scores, reward contingency did not significantly predict post-task delay discounting, $F(1, 110) = 1.29$, $p = .26$, partial $\eta^2 = .01$. Likewise, the manipulation check showed no reliable group differences in physiological effort during the MET ($p = .37$). Exploratory analyses indicated a small but significant pre-post shift toward more impulsive choices across conditions (Cohen's $d = 0.25$), consistent with a fatigue-related increase in discounting. Overall, brief reward-contingency manipulations were overshadowed by stable individual differences, suggesting that stronger or longer interventions and closer alignment between training and outcome domains may be necessary to influence delay of gratification.

[1] The corresponding German abstract can be found at the end of the document.

Introduction

Imagine being a four-year-old sitting in a room, staring at a delicious marshmallow. An adult tells you that if you wait patiently and do not eat it, you will receive two marshmallows later. This scenario is famously known as the "marshmallow test". While it may seem like a simple exercise in self-control, research indicates that the ability to resist the temptation to eat a marshmallow in exchange for a larger reward has significant implications (Mischel et al., 1972). The ability to delay gratification is an important indicator of self-regulation and goal-directed behavior. Preschool children who can wait for the second marshmallow tend to perform better academically and show improved coping with frustration and stressful situations (Mischel et al., 1989). Mischel et al. (1989) reported that children, in order to increase their ability to wait patiently, either distracted themselves or consciously self-instructed themselves to resist temptation. These early findings prompted further research into the mechanisms underlying delayed gratification.

Today, the modern world offers numerous opportunities for instant gratification, including social media, fast food, gaming, and online shopping. Resisting these temptations, just as with marshmallows, requires not only willpower but also the ability to think long-term and to regulate emotions. Ainslie's (1975) theory of specious reward and hyperbolic discounting explains this struggle: it posits that the subjective value of a reward declines steeply with delay, often leading individuals to act irrationally, preferring smaller, immediate rewards to larger, distant ones. Kirby et al. (2005) set out to investigate such behavior further and found that impulsivity in decision-making may impair long-term academic success (Kirby et al., 2005). A finding that is also supported by Bembenutty's body of research (2007, 2009, 2011) and by Bembenutty and Karabenick (1998).

Thus, whether it is about resisting junk food or studying instead of watching TV, our ability to prioritize long-term rewards over immediate pleasures plays a critical role in success. This underscores the importance of a better understanding of what promotes or impedes our willingness to delay rewards. What is it precisely that empowers some individuals to wait patiently, while others give in to temptation? To what degree can resistance to temptation be learned? What role does intrinsic motivation play? And, is it possible to increase long-term reward prioritization by focusing on rewarding effort rather than performance?

Theoretical Foundations

Delayed Gratification as a Self-regulatory Capacity

Before exploring whether effort-contingent rewards can help delay gratification, it is crucial to examine the concept of delayed gratification more closely. While the ability to delay gratification is sometimes considered a stable personality trait, evidence suggests that situational factors and motivation can significantly influence it. Mischel and Ebbesen (1970) demonstrated that children could increase their waiting time for a preferred reward when they diverted their attention away from it, for example, by engaging with something else rather than focusing on the reward. Building on this, subsequent studies showed that how children think about rewards also shapes their waiting times. When children focused on the temptation, such as the sensory qualities of the reward (for example, how good it would taste), their ability to delay gratification decreased. However, representing the reward in a more abstract form or directing their attention to other enjoyable thoughts (e.g., imagining playing a fun game or thinking about a different snack) led to much longer waiting times (Mischel et al., 1989). Taken together, these findings highlight that delay of gratification is highly sensitive to attentional and cognitive strategies and may therefore not be a fixed trait but rather something that can be shaped by training and environmental structure.

Additional support for this view comes from Heckman (2006), who argues that life-cycle skill formation is dynamic. He highlights the importance of motivation and skill development, preferably at an early age, for long-term benefits (Heckman, 2006). Similarly, Duckworth et al. (2013) singled out self-control as the key ability in delayed gratification tasks. Self-control correlated significantly positively with GPA and standardized achievement test scores and negatively with body mass index (BMI) (Duckworth et al., 2013). Additionally, it showed a negative association with risky behavior among adolescents and displayed the highest predictive power for life outcomes (Duckworth et al., 2013). Moffitt and colleagues (2010) confirmed that high self-control throughout childhood is associated with improved physical health and personal finances later in life, as well as a reduced likelihood of delinquent behavior. Self-control predicts a wide range of life outcomes, often with effect sizes comparable to or larger than traditional predictors such as intelligence, and it adds predictive power beyond IQ, personality traits (except conscientiousness), reward sensitivity, and family background (Duckworth et al., 2013; Moffitt et al., 2010).

Also, Bembenutty and Karabenick (1998) examined how successful students resist immediate temptations to prioritize long-term academic goals. They defined academic delay of gratification (ADOG) as a self-regulatory act that helps students postpone immediate rewards in favor of more distant outcomes that are perceived as more valuable (Bembenutty & Karabenick, 1998). Similar to prior studies on delayed gratification, they concluded that successful students also demonstrated higher self-efficacy and used more effective cognitive strategies, such as goal-setting and self-monitoring, to maintain focus on long-term goals (Bembenutty & Karabenick, 1998).

Similarly, Weiner (1985) had already highlighted the motivational importance of effort in his attributional theory, claiming that the internal capability to control effort-investment makes it a powerful motivator. Extending this logic, attributing long-term success to sustained effort should promote behavior consistent with greater ability to delay gratification.

These studies highlight the importance of effortful control in delaying gratification and suggest that waiting time can be improved. Rather than a fixed trait, the ability to postpone gratification may be a self-regulatory capacity that is malleable, either by altering certain conditions or by employing specific strategies.

Self-Control and Intrinsic Motivation

While self-control is a key predictor of delayed gratification (Duckworth et al., 2013), it has long been conceptualized as a finite resource (Baumeister et al., 1998; Hagger et al., 2010). Such "depletion" views, however, fail to explain why individuals often persist in effortful tasks despite the demanding challenges they face. Therefore, recent research suggests a need for a more complex understanding of self-control.

Inzlicht et al. (2014) argue that self-control may primarily reflect shifts in motivation and attention over time rather than resource repletion. Accordingly, they argue that research should focus more on the mechanisms underlying this declining willingness to perform (Inzlicht et al., 2014). If the willingness to perform well decreases, so does the self-control in the task at hand.

Metcalfe and Mischel (1999) describe willpower as the ability to inhibit an impulsive response that undoes one's commitment. They propose a hot-cool system: the reflective "cool" system supports delayed gratification, while the emotionally driven "hot" system seeks instant gratification. Although the ability to use the cool system improves with age, Metcalfe and Mischel (1999) note that willpower can be improved by training individuals to obscure or reframe the stimulus.

To a large degree, willpower to self-regulate also depends on feedback and reinforcement. Mueller and Dweck (1998) showed that praising children's intelligence (ability-focused) leads them to endorse a fixed view of ability, whereas praising effort promotes beliefs that performance is improvable and under their own control (Mueller & Dweck, 1998). When children were praised for effort, they were much more likely to develop learning-oriented goals, focus more on improving their process, show greater resilience in the face of failure, and attribute outcomes to factors within their control (Mueller & Dweck, 1998). Thus, this research underscores the importance of praising the process rather than the performance when fostering a mindset that supports long-term achievement. Furthermore, these varied reactions to positive feedback regarding effort-investment may indicate that intrinsic motivation is a key factor in improving self-regulation and, in turn, in increasing the ability to delay gratification.

Also in Bembenutty's research on delayed gratification, mindset is a key factor. He found that students who reported a stronger preference for delaying gratification earned higher grades and demonstrated greater intrinsic motivation to achieve long-term goals (Bembenutty, 2011). Additionally, Bembenutty's study (2011) also emphasizes the importance of self-efficacy beliefs regarding delayed gratification. He notes that confidence in one's ability to perform well and valuing academic tasks for their own sake are positively related to ADOG (Bembenutty, 2011). Similarly, Ryan and Deci (2000) identify feelings of competence and relatedness, as well as perceived autonomy, as key ingredients of intrinsic motivation. Intrinsic motivation, they claim, fosters "performance, persistence, and well-being across life's epochs" (p. 65). This relationship between intrinsic motivation and delayed gratification can also be explained by the Self-Determination Theory (SDT) (Deci & Ryan, 1985). SDT posits that people are naturally motivated to grow, learn, and shape their behavior. As a result, people are primarily driven by internal forces and personal values rather than external control (Deci & Ryan, 1985). At its most basic level, SDT distinguishes between intrinsic and extrinsic motivation and claims that the type of motivation for action depends on the nature of the cause. Overall, when students feel competent and feel connected to the tasks, i.e., when their intrinsic motivation is high, they are also more likely to focus on long-term goals. This connection between intrinsic motivation and self-regulation explains why individuals with strong internal motivation are generally better at resisting short-term temptations and improving their performance.

Also, Job et al. (2024) posit that the right circumstances can support continuous control and thereby positively influence effort mobilization. However, in their critical review of the law of least effort, they propose that effort mobilization does not necessarily depend on self-control, i.e., the ability to inhibit an impulse in pursuit of a more distant goal. Instead, they propose the "need for effort" (p. 3) as a functional health resource and learning mechanism. While Job et al.'s (2024) position overlaps with Deci and Ryan's (1985) view that people can be driven by internal forces, they leave open the possibility that intrinsic motivation may not be mediated by self-control. Basic emotions such as interest, curiosity, and a sense of growing competence in individuals suggest a motivational system that, in some situations, has been shown to favor effortful seeking over energy conservation (Job et al., 2024).

The Effects of Rewarding Effort

Whether self-control mediates the effects of intrinsic motivation or whether people are equipped with an innate drive for effort and thereby display greater self-control, the effects of rewarding effort-investment on both self-control and intrinsic motivation are much clearer.

Lepper et al. (1973) explored the influence of extrinsic rewards on intrinsic motivation in children. Compared to a control group that did not receive a reward and a group of children who were rewarded unexpectedly, children who had been motivated by an external incentive subsequently showed diminished interest in a drawing activity, i.e., exhibited diminished intrinsic motivation following external rewards (Lepper et al., 1973). These results support the so-called overjustification hypothesis, the idea that receiving external rewards for an activity that is enjoyed can shift people's focus from intrinsic interest to the reward itself (Lepper et al., 1973).

However, prior studies have suggested a direct link between effort reinforcement and self-regulation. The aforementioned findings by Mueller and Dweck (1998) highlighted that praise for effort, rather than for outcome, fosters a growth mindset and greater resilience in children. Similarly, Eisenberger and Adornetto (1986) showed that when high effort is systematically paired with reward, individuals later persist longer on complex tasks, highlighting the ability to influence intrinsic motivation. Eisenberger (1992) conceptualised this as "learned industriousness," where effort investment itself becomes a secondary reinforcer. These findings suggest that the nature of reinforcement plays a crucial role in subsequent improvements in self-regulation, intrinsic motivation, and delay of gratification.

A meta-analysis by Deci et al. (1999) showed that expected rewards, when they are perceived as controlling, tend to reduce intrinsic motivation and undermine self-regulation. In contrast, more informational rewards that signal competence have weaker or no undermining effects. Building on this distinction, some evidence suggests that rewards are less likely to undermine intrinsic motivation when they provide competence-relevant feedback and are contingent on task engagement (e.g., Cameron et al., 2004), particularly for moderately challenging tasks.

Such findings not only underscore the importance of rewarding effort in enhancing the ability to delay gratification but also align with the conclusions of Bembenutty (2011) and Ryan and Deci (2000), indicating that self-efficacy beliefs are positively correlated with intrinsic motivation. Furthermore, these studies suggest that intrinsic motivation can transform effort-aversive behavior into effort-seeking behavior. While such theories run counter to traditional motivational theories that often portray cognitive effort as aversive, they align with studies on the secondary reward value of effort. They propose that if effort is systematically rewarded, it can acquire intrinsic value.

A large-scale study by Clay et al. (2022) found that participants in an effort-contingent reward group selected more challenging tasks in a later phase of the study, despite having been told that no additional rewards would be provided. This effect was measured not only behaviorally but also physiologically, with pre-ejection period (PEP) reactivity recorded as an index of sympathetic effort mobilization (Clay et al., 2022). Their findings further support the hypothesis that rewarding effort can increase intrinsic motivation and, in turn, lead individuals to seek cognitive effort intrinsically, with significant implications for self-regulation and delayed gratification ability.

Together, these findings support the idea that rewarding effort can influence later effort investment and task persistence, both key components of self-regulation. Delayed gratification, however, is more than a behavioral outcome. Delayed gratification also reflects underlying valuation and control processes in the brain. Therefore, the next section outlines neurobiological correlates of effort and delay with a focus on reward contingency and intertemporal choice.

Neurobiological Correlates of Effort and Delay

The ability to delay gratification has been discovered to be reflected in neural mechanisms. McClure et al. (2004) found that decisions involving immediate rewards engaged limbic regions (associated with affective valuation), whereas decisions involving delayed rewards engaged prefrontal regions (linked to cognitive control).

Furthermore, Balconi et al. (2024) examined economic decisions where greater effort yielded higher rewards. Their EEG results indicated that less effortful choices were associated with distinct patterns reflecting a perceived advantage in avoiding boredom. This supports the view that effort is subjectively evaluated in the context of perceived long-term flexibility.

Furthermore, Balconi et al. (2024) examined effort-based economic decisions where greater effort yielded higher rewards while recording EEG. In a series of trials, participants repeatedly chose between a high-effort/high-reward option, expressed through many key presses to obtain a larger monetary gain, and a low-effort/low-reward option, expressed through fewer key presses for a smaller monetary gain.Although most participants preferred to invest high effort and thereby chose the more profitable option, the EEG data showed that choosing the low-effort option was accompanied by systematic changes in oscillatory activity. Specifically, low-effort choices were associated with altered power across delta, theta, alpha, and beta bands, including stronger increases at a right temporoparietal site (TP10) and reduced delta/theta activity at the contralateral site (TP9), alongside a decrease in delta activity at a left frontal site (AF7). Balconi et al. interpret these neural differences as an indication that some individuals may experience low-effort choices as advantageous because they avoid repetitive, potentially boring or automatic responding and preserve a sense of flexibility for alternative actions.

These findings suggest a plausible neurocognitive pathway: if effort becomes subjectively more valued (via effort-contingent rewards), control-related prefrontal systems may be more strongly engaged, ultimately supporting the regulation of impulses and leading to more patient outcomes.

Summary and Research Question

Overall, the ability to delay gratification has long been linked to self-regulation skills and has been well studied as a predictor of academic achievement and overall life success. Although often described as a trait, recent evidence indicates that cognitive strategies, environmental cues, motivational framing, and neurobiological processes can influence self-control and, in turn, delay gratification. Effort-contingent rewards have proven effective in nurturing self-control and intrinsic motivation, thereby enabling individuals to persist in tedious tasks.

Theoretical frameworks such as the SDT (Deci & Ryan, 1985), attributional theory (Weiner, 1985), and learned industriousness (Eisenberger, 1992), as well as the findings from Clay et al. (2022), Eisenberger and Adornetto (1986), and Bembenutty (2011), all agree that effort, when appropriately reinforced, can become intrinsically valuable. Additionally, recent neuroscientific evidence (Balconi et al., 2024; McClure et al., 2004) indicates that the valuation of effort is reflected in increased activation of control-related brain regions, suggesting a neurocognitive pathway by which effort-contingent reward may also improve the ability to delay gratification. If effort itself can be experienced as intrinsically rewarding (Eisenberger, 1992; Clay et al., 2022), interventions that target the valuation of effort, such as rewarding effort, should enhance self-regulation and, as a result, improve the ability to wait for gratification.

While previous research has mainly examined how motivation and attention affect delay tolerance, the impact of effort-contingent versus performance-contingent rewards on delayed gratification remains largely unexplored. Therefore, this study aims to address this gap and investigate whether rewarding effort, rather than performance, leads to greater delay of gratification.

Hypotheses

H1 (Main Hypothesis): Participants in the effort-contingent reward condition will show greater delay of gratification than participants in the performance-contingent reward condition, controlling for baseline delay discounting.

H2 (Manipulation Check): Participants in the effort-contingent reward condition will exhibit greater physiological effort (shorter PEPs) than performance-rewarded participants during the transfer task (MET).

H3 (Exploratory): Exposure to effort-investment through cognitively demanding tasks will cause a change in delay of gratification compared to pre-task delay discounting rates, regardless of reward condition.

Method

Design

This study employed a between-subjects experimental design to investigate the effects of effort-based versus performance-based rewards on delayed gratification. Participants were randomly assigned to one of two conditions for the first task (n-back), where they completed the same sequence but were exposed to different reward contingencies (Independent Variable). The effort-contingent group received rewards contingent on physiological effort (PEP changes), whereas the performance-contingent group received rewards contingent on task accuracy (hits/misses). This design allowed for isolating the effects of effort reinforcement from performance outcomes.

Effort was assessed by analyzing the PEP, a physiological marker of sympathetic nervous system activation; shorter PEPs indicate greater effort. Performance was assessed by task accuracy (hits and misses). Following the n-back task, participants completed a math-effort task (MET) without external rewards to examine the transfer of internally regulated effort.

The primary outcome variable was delayed gratification, measured using a shortened version of Kirby's (1999) Delay Discounting Monetary Choice Questionnaire (MCQ)[2] administered twice: at baseline (one week pre-test) and post-task. This should allow both a between-group comparison of post-task MCQ scores while controlling for baseline, and within-subject analyses of pre-port change across the session..

Sample

Participants for this study were recruited from the University of Vienna participant pool. The final analytic sample consisted of 146 psychology students, the majority of whom were female (68.5%), with 27.4% male and 4.1% identifying as diverse. Participants had a mean age of 22.3 years (*SD* = 3.6, range: 18-40), and a total of 100 participants identified as female, 40 as male, and 6 as diverse.

To ensure the study had sufficient statistical power, an a priori power analysis was conducted with GPower to determine the required sample size. Given a power (*β-1*) of .8, an alpha (α) level of .05, and the expected effect size of Cohen's f = .25, with two groups and

[2] Delay discounting was measured using selected items of the MCQ developed by Kirby, Petry, and Bickel (1999). The full set of items was obtained from the PhenX Toolkit (https://www.phenxtoolkit.org/protocols/view/530301).

one covariate, GPower suggested a required sample size of N = 128 participants. To account for potential participant dropouts or exclusions, the goal was to recruit an additional 10% of participants beyond the required sample size. Therefore, the plan was to enroll at least N = 141 participants to ensure robust study findings.

Participants had to meet specific inclusion criteria to be eligible for the study. They had to be aged 18 and over to ensure a cognitively homogeneous sample. All participants needed to be fluent in either English or German, as these were the languages used in the study. Additionally, participants must not have had any diagnosed cardiovascular disorders to ensure that the physiological PEP measurements accurately reflected task-related effort rather than underlying health conditions. Furthermore, individuals who had previously participated in similar studies on reward-based motivation were excluded to prevent prior knowledge from influencing the study's results.

Before conducting the analyses, several exclusion criteria were decided on to ensure data quality and interpretability. Participants with incomplete or aberrant responses, such as excessive missing data or response patterns indicating disengagement, were to be removed from the analysis. In addition, any participant who demonstrated non-compliance with effort during the MET task was to be excluded based on behavioral indicators. Additionally, participants in the effort-contingent group who received less than 2€ in total rewards during the n-back task were excluded, as such minimal reinforcement was unlikely to elicit the intended motivational framing of effort-based rewards. This exclusion criteria was based on Clay et al. (2022) and Eisenberger (1992), who concluded that effort must be consistently and meaningfully rewarded to become internalized. Therefore, this exclusion criterion ensured that effort-contingent participants had been exposed to a sufficient reward signal to make experimental manipulation meaningful.

A total of 199 participants were initially recruited for the study. Based on the aforementioned exclusions criteria, eight participants were excluded manually for documented protocol violations or irregularities: one assumed the task was about frustration tolerance and did not engage meaningfully (ID 7), one had a reward value entered incorrectly by the tester (ID 29), and six participants (IDs 43, 67, 107, 153, 160 and 171) had anomalous reward outcomes or other data irregularities.

Moreover, three participants (18, 83, and 175) had problematic data after importation and were therefore excluded. Of the remaining participants, 10 were excluded due to missing values on the Reward variable. Next, 32 participants in the effort contingent group were excluded because they earned less than 2€ in the n-back task[3]. Finally, all cases were checked for completeness of key demographic variables, including gender, age, study, and parental college education, and no additional participants needed to be excluded at this stage. The final sample, therefore, consisted of 146 participants with complete data on all relevant variables.

All inferential analyses involving delay discounting required paired MCQ data. Therefore, hypotheses testing for H1 and H3 was conducted on the subsample that completed the MCQ both at baseline and post-task (N = 113), as 33 participants without post-task MCQ data were excluded from these analyses.

Procedure

Before their laboratory session, participants provided informed consent and completed baseline assessments online. One week before attending the laboratory, in order to assess baseline delay of gratification, all participants were sent the Monetary Choice Questionnaire (MCQ; Kirby et al., 1999). On the day of testing, participants were greeted upon arrival at the laboratory and fitted with a non-invasive impedance cardiography device. They were seated in front of a computer where the testing would take place and listened to relaxing music from a standardized YouTube playlist. During this period, baseline cardiovascular measurements, specifically HR and PEP, were recorded.

Participants were then randomly assigned to one of two experimental conditions: effort-contingent reward or performance-contingent reward. At this stage, the group assignment was revealed only to the experimenter, not to the participant. Subsequently, participants completed several additional questionnaires unrelated to the present study but part of a larger research project.

The participants then began the n-back task, during which they received monetary rewards contingent on either performance or effort. Depending on group assignment, either greater physiological effort mobilization or better performance led to higher monetary rewards.

Next, participants continued with Task 2, the MET. At this stage, participants were informed that no external rewards would be provided during the task, thereby allowing

[3] To assess the impact of this exclusion criterion, all analyses were also conducted without excluding participants who earned less than 2 € in rewards (inclusive sample). The corresponding results and interpretations are presented in the Appendix.

assessment of effort regulation and cognitive persistence in the absence of extrinsic incentives.

After completing both experimental tasks, participants again completed the shortened MCQ to assess post-task delay discounting, enabling a direct within-person pre–post comparison in addition to the between-condition analysis.

Before leaving, participants answered a brief set of questions to assess whether they had remained blind to the study's true purpose, and were then debriefed about the study's actual aims, including details on the reward contingencies and the rationale for the two experimental conditions. Participants received their cumulative earnings and signed to acknowledge receipt of their payment. At this stage, any remaining questions were addressed. All data were treated confidentially and in accordance with data protection guidelines.

Measures and Materials

Cardiovascular Assessment

A non-invasive impedance cardiography device recorded heart rate (HR) and PEP. PEP, the time interval between the onset of ventricular depolarization and the opening of the aortic valve, served as the primary index of sympathetic beta-adrenergic activation and cognitive effort (Kelsey, 2012). Shorter PEP indicates higher effort.

MCQ shortened (Pre- and Post-test Measurement)

To assess delay discounting, a shortened 9-item version of the original 27-item Monetary Choice Questionnaire (MCQ; Kirby et al., 1999) was administered online one week before the laboratory session (baseline/pre-test) and again immediately after the experimental tasks (post-test). The original MCQ comprises 27 items, organized in a 3 × 3 × 3 matrix reflecting three dimensions: reward magnitude (small, medium, large), delay duration (short, medium, long), and discount rate (low, medium, high). For the present study, one item from each combination of these dimensions was selected, and items with the highest test-retest reliability and discriminative validity as reported by Kirby et al. (1999) were prioritized. This approach ensured that the shortened 9-item version of the MCQ preserved full coverage of the original instrument's range of reward amounts, time delays, and impulsivity levels. The items selected were those with the highest test-retest reliability within each combination, as recommended by Kirby et al. (1999). Participants completed the shortened MCQ twice: once online approximately one week before the laboratory session (baseline/pre-test) and again in the laboratory immediately after the experimental task battery (post-test).

All items were translated into German, and monetary values were adapted to Euros to maximize contextual relevance for the study population. Participants were asked to indicate, for each item, whether they preferred a smaller immediate reward or a larger delayed reward.

Delay discounting was operationalized as the proportion of impulsive choices, with higher values indicating a preference for immediate rewards. Individual k-values were not calculated (as suggested by Kirby et al., 1999)[4] because of the abbreviated nature of the scale.

The final 9 items of the shortened MCQ were:

1. Würden Sie lieber heute 55 € oder 75 € in 61 Tagen erhalten?
   (Would you like to receive 55 € today or 75 € in 61 days?)
2. Würden Sie lieber heute 19 € oder 25 € in 53 Tagen erhalten?
   (Would you like to receive 19 € today or 25 € in 53 days?)
3. Würden Sie lieber heute 31 € oder 85 € in 7 Tagen erhalten?
   (Would you like to receive 31 € today or 85 € in 7 days?)
4. Würden Sie lieber heute 14 € oder 25 € in 19 Tagen erhalten?
   (Would you like to receive 14 € today or 25 € in 19 days?)
5. Würden Sie lieber heute 27 € oder 50 € in 21 Tagen erhalten?
   (Would you like to receive 27 € today or 50 € in 21 days?)
6. Würden Sie lieber heute 49 € oder 60 € in 89 Tagen erhalten?
   (Would you like to receive 49 € today or 60 € in 89 days?)
7. Würden Sie lieber heute 24 € oder 35 € in 29 Tagen erhalten?
   (Would you like to receive 24 € today or 35 € in 29 days?)
8. Würden Sie lieber heute 33 € oder 80 € in 14 Tagen erhalten?
   (Would you like to receive 33 € today or 80 € in 14 days?)
9. Würden Sie lieber heute 34 € oder 50 € in 30 Tagen erhalten?
   (Would you like to receive 34 € today or 50 € in 30 days?)

N-Back Task (Task 1)

Participants completed a computerized n-back task to induce cognitive effort (Kane & Conway, 2016). Participants viewed a sequence of letters and pressed a key if the current

[4] Modeling individual k-values based on this short form of the MCQ would yield unreliable and underpowered results.

letter matched the one presented n steps earlier. The task consisted of 24 blocks varying in difficulty (1-back to 4-back).

Effort-Based Condition: Rewards were contingent on block-wise changes in PEP (effort).

Performance-Based Condition: Rewards were contingent on accuracy (hits/misses).

Math-Effort Task (MET - Task 2)

The MET served as an unrewarded transfer task (Engle-Friedman et al., 2003). Participants completed 40 addition problems of varying difficulty (Level 1: 1–9; Level 2: 7–35). Before each trial, participants rated their willingness to solve the problem using a visual analog slider. No external rewards or performance feedback were provided, isolating intrinsic motivation and effort regulation.

All statistical analyses were conducted in R (Version 2024.09.0). Assumptions for parametric tests (normality and homogeneity of variance) were checked prior to analysis. Statistical significance was set at $\alpha = .05$.

Results

Results Overview

Because pre–post MCQ analyses required complete data at both time points, H1 and H3 were conducted on the subsample with paired MCQ responses ($N = 113$; missing post-test data were due to an earlier study version). Sensitivity analyses indicated that restricting the sample to these participants did not materially change the pattern of results.

Delay Discounting (MCQ) Scores by Group and Time

The mean and standard deviation of the proportion of impulsive choices on the MCQ were calculated separately for each group (effort, performance) at pre-test and post-test ($N = 113$). At pre-test, the mean proportion of impulsive choices was similar for the effort-contingent ($M = 0.21$, $SD = 0.23$) and performance-contingent ($M = 0.21$, $SD = 0.22$) reward groups. After finishing the tasks, both groups showed a higher mean proportion of impulsive choices, suggesting a slight decrease in the ability to delay gratification. This increase appears larger in the performance-contingent group ($M = 0.26$, $SD = 0.23$) compared to the effort-contingent group ($M = 0.23$, $SD = 0.22$).

Figure 1 displays the distribution of impulsive choices at pre- and post-test. For both groups, the median proportion of impulsive choices was relatively low at pre-test, with wide variability and a few outliers. The plot indicates a slight upward shift in both distributions over time. At post-test, the performance group shows higher medians and upper quartiles than the effort group. Although interquartile ranges overlap substantially across conditions and time points, indicating similar distributions overall, the performance group appears to have more participants with higher scores (i.e., lower ability to delay gratification) at post-test.

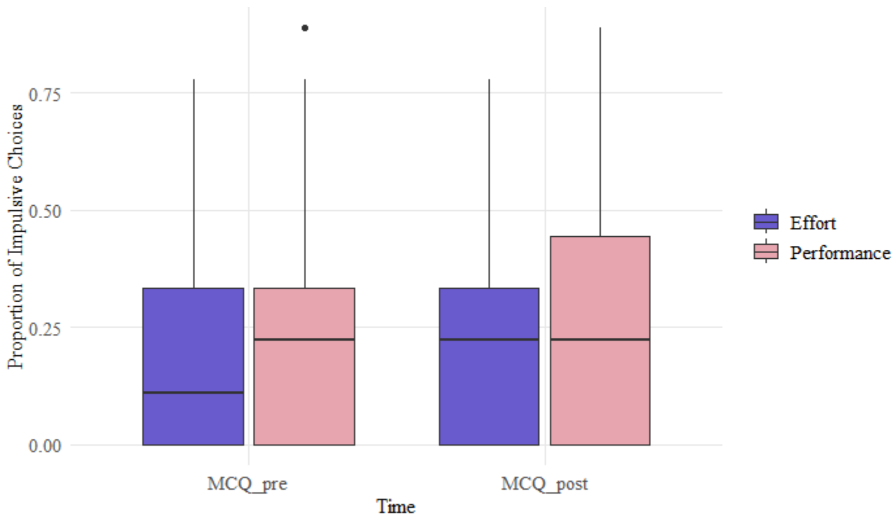

Figure 1. Distribution of MCQ scores by group (pre- vs post-test).

To interpret any group differences in intrinsic motivation or delay discounting, it is important to establish whether the two contingencies produced comparable monetary outcomes. Figure 2 therefore displays the distribution of earned monetary rewards, allowing evaluation of potential confounding differences in mean payout, variability, and skewness between conditions. In both groups, most rewards lie in the mid-range, around €4.00 to €5.00. However, the distribution shapes differed: the performance group showed a narrow, pronounced peak in this range, whereas the effort group displayed a flatter distribution with a long right tail. This suggests that, although average rewards were comparable across groups, payouts were more variable in the effort condition, with a small number of participants earning particularly high rewards. Thus, any effects of reward contingency are unlikely to be attributable to differences in average payout, although differences in payout variability (i.e., a more skewed distribution in the effort condition) could still have influenced participants' reward experience.

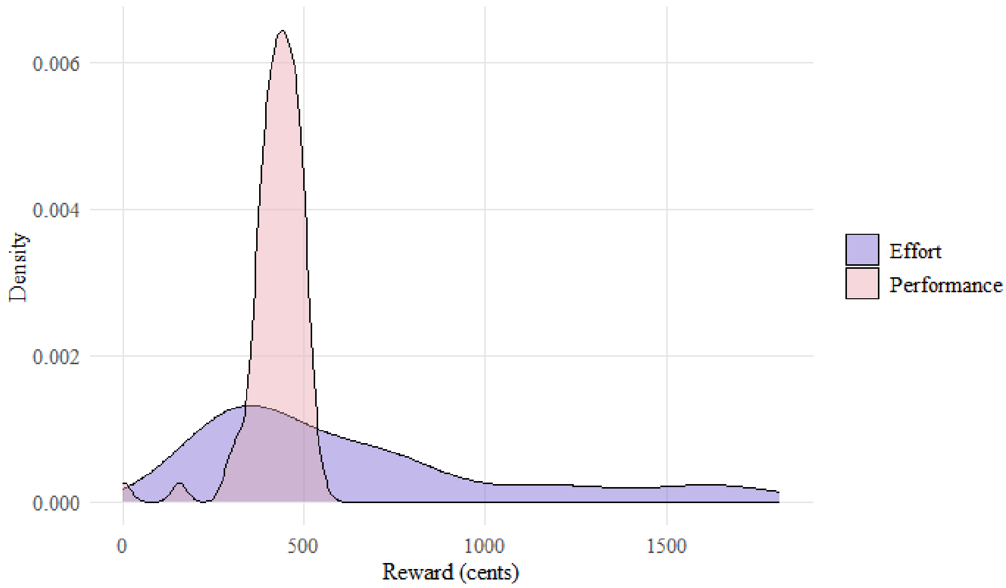

Figure 2. Reward density by group.

Main Analysis (ANCOVA)

Assumption checks

ANCOVA assumptions were examined prior to analysis. Based on the random assignment, independence of observations was assumed. Table 1 summarizes the statistical tests of assumptions. Although the assumption of normality of residuals was violated (Shapiro-Wilk $W = 0.95$, $p < .001$), ANCOVA remains robust to deviations from normality given the sample size ($N > 100$). Levene's test indicated no violation of homogeneity of variance (p = .72), and there was no significant interaction between the covariate and the independent variable, supporting the assumption of homogeneity of regression slopes (p = .77). A more detailed analysis of the ANCOVA assumption checks is provided in the Appendix.

Table 1
Assumption Checks for ANCOVA (N = 113)

Assumption	Test/Output	Result	Interpretation
Normality of residuals	Shapiro-Wilk W = 0.951, p = .0004	Violated (p < .05)	The residuals are not normally distributed, but ANCOVA is robust for N > 100.
Homogeneity of variance	Levene's $F(1, 111)$ = 0.13, p = .72	Met (p > .05)	No evidence for variance difference between groups.
Homogeneity of regression slopes	Interaction $F(1, 109)$ = 0.09, p = .77	Met (p > .05)	No group x pre-MCQ interaction.

Note. W = Shapiro–Wilk statistic; F = F test statistic. A p-value < .05 was taken to indicate a violation of the corresponding assumption. In the present data, the assumption of normality was violated (p = .0004), whereas homogeneity of variance (p = .72) and homogeneity of regression slopes (p = .77) were supported.

Figure 3 (p.19) depicts the relationship between baseline (MCQ_pre) and post-task (MCQ_post) scores. The strong positive association and the parallel linear trends across groups are consistent with the assumptions of linearity and homogeneity of regression slopes.

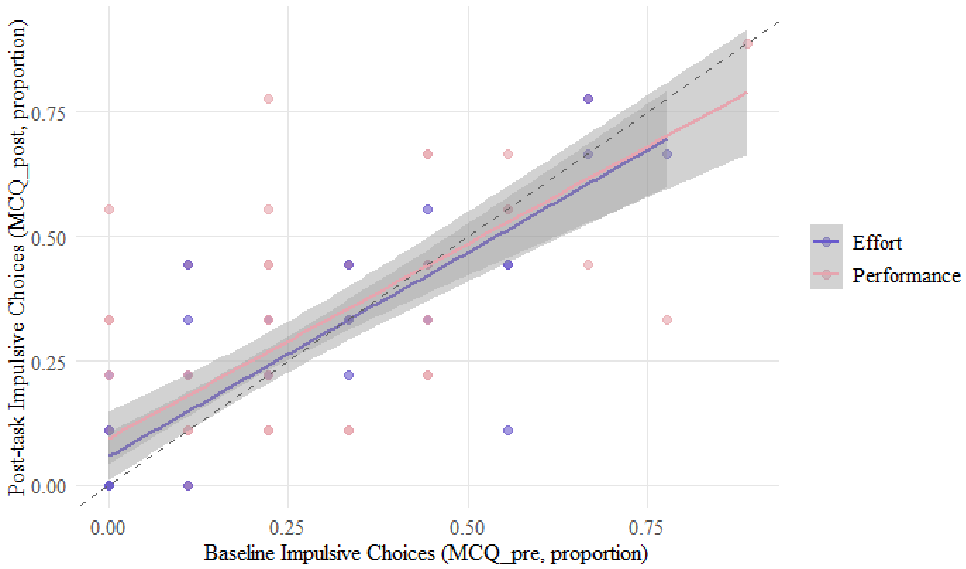

Figure 3. Pre- vs. post-task delay discounting (MCQ_pre and MCQ_post) by reward condition.
Note. Points represent individual participants; lines show separate linear regression fits for the effort and performance groups. The dashed line indicates equality (MCQ_post = MCQ_pre).

Hypothesis 1 Results

Hypothesis 1 (H1) predicted that participants in the effort-contingent reward condition would show greater delay of gratification (i.e., lower MCQ scores / fewer impulsive choices) than participants in the performance-contingent reward condition, controlling for baseline discounting. To test H1, an analysis of covariance (ANCOVA) examined whether post-task delay discounting (MCQ_post) differed between the effort (n = 42) and performance (n = 71) groups when controlling for baseline discounting (MCQ_pre). As shown in Table 2, the covariate MCQ_pre was strongly related to MCQ_post, $F(1,110) = 165.00$, $p < .001$, partial $\eta^2 = .60$. However, the main effect of group was small and non-significant, $F(1,110) = 1.29$, $p = .26$, partial $\eta^2 = .01$.

Table 2
ANCOVA Results for Post-Task MCQ Scores

Effect	SS	df	MS	*F*	*p*	Partial η^2	95% CI for η^2
group	0.027	1	0.027	1.29	.258	0.01	[0.00, 1,00]
MCQ_pre	3.510	1	3.510	165.000	<.0001***	0.60	[0.51, 1.00]
Residuals	2.340	110	0.021				

Note. SS = sum of squares; df = degrees of freedom; MS = mean square.

The model explained about 60% of the variance in MCQ_post. Adjusted mean post-task MCQ scores were 0.230 (SE = 0.023, 95% CI [0.185, 0.275]) for the effort-contingent reward group and 0.258 (SE = 0.017, 95% CI [0.224, 0.293]) for the performance-contingent reward group, and their difference was not statistically significant, estimate = −0.028, SE = 0.028, $t(110) = -0.998$, $p = .32$.

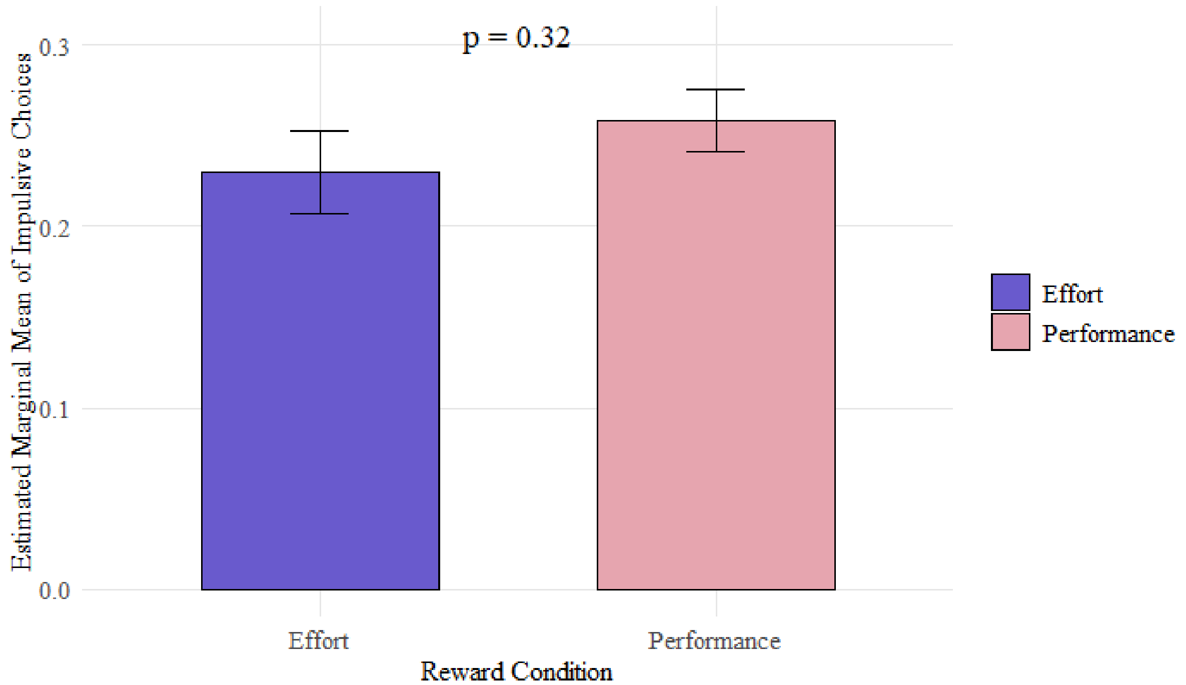

Figure 4. Barplot of adjusted mean post-task MCQ scores (± *SE*) for both groups after controlling for baseline.

H2 (Manipulation Check)

Hypothesis 2 (H2) served as a manipulation check and predicted that participants in the effort-contingent reward condition would exhibit greater physiological effort during the unrewarded transfer task (MET), indexed by shorter PEP (i.e., a more negative PEP change score) than participants in the performance-contingent reward condition. To test H2, an independent-samples *t*-test compared mean PEP reactivity during the MET (METPEPc) between the effort-contingent and performance-contingent groups using the full sample (N = 146). Negative values indicate a shortening of PEP and thus greater sympathetic effort mobilization. Descriptively, the effort group showed slightly more negative PEP change (M = -1.39 ms, SD = 7.37) than the performance group (M = -0.22 ms, SD = 7.70; see Table 3).

Table 3
PEP Change (ms) During MET Task by Group

Group	Mean	*SD*	Variance	*n*	Min	Median	Max
Effort	-1.39	7.37	54.26	51	-20.01	-0.21	16.41
Performance	-0.22	7.70	59.33	95	-16.45	0.06	17.72

Note. Negative values indicate a decrease in PEP (greater sympathetic effort mobilization).

An independent-samples Welch *t*-test comparing mean PEP change (METPEPc) between the groups revealed no statistically significant group difference, $t(106.42) = -0.89$, p = .373, 95% CI [-3.74, 1.41], Cohen's d = -0.15, 95% CI [-0.49, 0.19]. The distributions of PEP changed, as illustrated by the approximately symmetric boxplot (see Figure 5), and showed substantial overlap across conditions. No extreme outliers or pronounced group differences were evident.

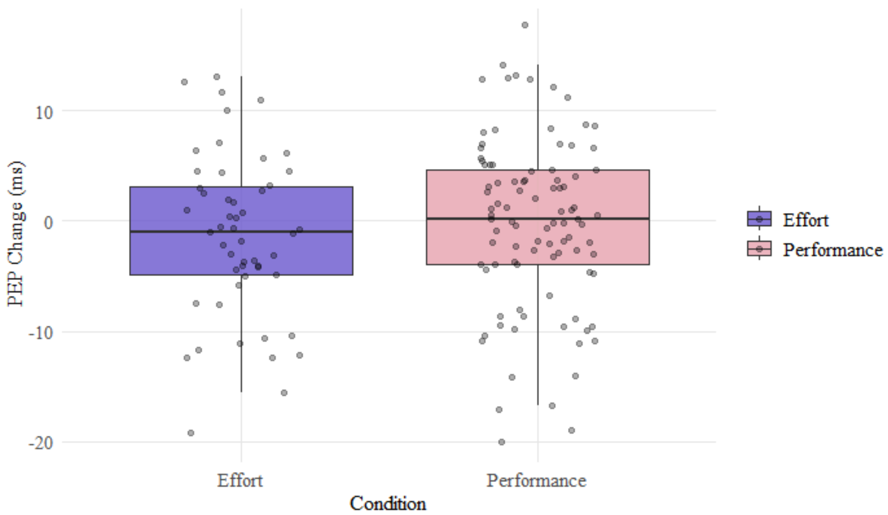

Figure 5. PEP change during MET task by condition.

H3 (Exploratory)

H3 Results

To examine whether delay discounting changed from pre- to post-test following exposure to cognitively demanding tasks, Hypothesis 3 (H3) predicted a within-subject change in MCQ scores from baseline to post-task regardless of reward condition. To test this, a paired-samples t-test was conducted on the proportion of impulsive choices in the MCQ ($N = 113$). The results supported the hypothesis, revealing a statistically significant change in impulsive choices from pre- to post-task, $t(112) = -2.68$, $p = .008$, mean difference = -0.04, 95% CI [-0.07, -0.01]. The direction of the effect indicated a shift toward greater impulsivity (i.e., lower delay of gratification) after the task battery. The effect size ($d = 0.25$, 95% CI [0.06, 0.44]) represents a small effect (Cohen, 1988).

To assess whether this change depended on reward condition, change scores were computed (ΔMCQ = MCQ_pre - MCQ_post). Both groups showed, on average, negative change scores (effort: $M = -0.02$; performance: $M = -0.05$). The difference between groups was not significant, Welch's $t(106.61) = 1.00$, $p = .32$, Cohen's $d = 0.18$, 95% CI [-0.20, 0.56]. Thus, the observed decrease in delay of gratification appeared to be a general effect of the experimental session rather than a result of the specific reward contingency. Figure 6 illustrates the individual trajectories of MCQ scores from pre- to post-test, visualizing the subtle shift toward higher impulsivity across the sample.

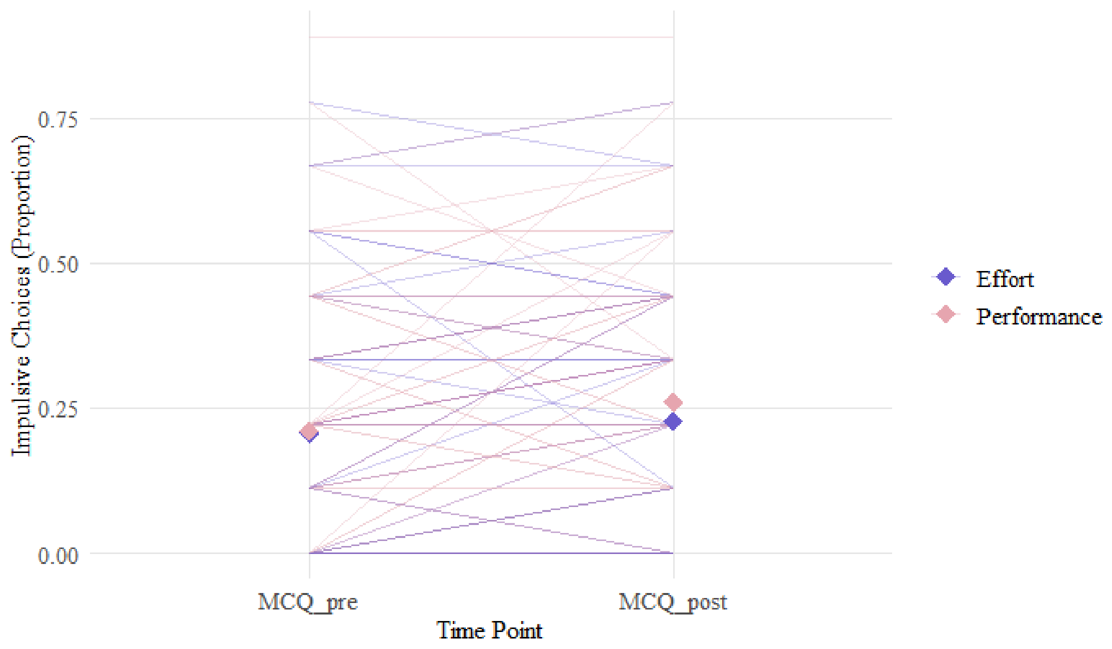

Figure 6. Individual change in gratification delay.

Exploratory Analyses of Mechanisms

Given the observed shift toward impulsivity, exploratory regression analyses were conducted to identify potential predictors of this change. First, multiple regression models tested whether motivational or physiological variables, specifically total monetary reward earned, physiological effort mobilization (PEP change), and self-reported fatigue, predicted the change in MCQ scores. None of these variables accounted for significant variance in ΔMCQ after controlling for baseline impulsivity (all *p*s > .05; see Appendix X for full regression tables). This indicates that neither earning more money nor exerting greater physiological effort was associated with changes in the ability to delay gratification.

Second, the relationship between baseline impulsivity and change was examined to test for regression-to-the-mean effects. Baseline MCQ scores (MCQ_pre) emerged as the only consistent predictor of change. Participants were divided into tertiles based on their baseline scores:

- Low and medium impulsivity: These groups showed small increases in impulsive choices (ΔMCQ ranging from −0.05 to −0.10).
- High impulsivity: This group showed little change or slight improvement (ΔMCQ = +0.03).

A quadratic regression of ΔMCQ on baseline scores confirmed a small but reliable association ($R^2 = .11$, $F(2,110) = 6.82$, $p = .002$), with a trend toward a curvilinear pattern (MCQ_pre^2: $b = 0.43$, $SE = 0.25$, $p = .095$). This U-shaped pattern suggests that individuals who were initially very impulsive tended to deteriorate less or improve slightly compared to those who already showed high patience at baseline, a pattern consistent with regression to the mean.

Discussion

Interpretations of the Results

H1: Effort-Contingent Rewards and Delay of Gratification

The present study investigated whether rewarding effort rather than performance would lead to greater delay of gratification. H1 hypothesized that effort-contingent rewards would increase participants' willingness to forgo smaller immediate rewards in favor of larger delayed ones.

The findings did not support this hypothesis. After controlling for baseline delay discounting, reward contingency did not have a statistically significant impact on post-task MCQ scores. Instead, baseline delay discounting emerged as the strongest and the only significant predictor, accounting for approximately 60% of the variance in post-task scores. In this experiment, delay of gratification appears more as a stable disposition than a malleable state. This interpretation is strengthened by the fact that delay discounting was assessed twice, once at baseline and again after the task battery. The strong baseline–post association therefore reflects substantial short-term stability, leaving limited variance for a brief experimental manipulation to explain.

Nevertheless, the descriptive trends suggest a potential benefit of rewarding effort. While both groups exhibited a slight increase in impulsivity from pre- to post-task, the effort-contingent group's increase was smaller ($M = 0.21 \rightarrow 0.23$) than that of the performance-contingent group ($M = 0.21 \rightarrow 0.26$). As suggested by the concept of learned industriousness (Eisenberger, 1992; Eisenberger & Adornetto, 1986), pairing effort with reinforcement can condition participants to experience effort investment itself as rewarding. Notably, the post-test MCQ was completed immediately after a series of demanding cognitive tasks, so the overall pre–post increase in impulsivity may partly reflect transient state factors (e.g., mental fatigue) rather than enduring preference change. Thus, the smaller drift toward impulsivity in the effort-contingent group may indicate slightly greater resistance to a task-induced, state-related shift toward immediate gratification, given that the post-test MCQ was administered directly after the cognitive task battery.

Hypothesis 2: Physiological Effort (Manipulation Check)

H2 predicted that participants rewarded for effort would exhibit greater physiological effort (shorter PEP) during the transfer task. However, the data did not support this hypothesis ($p = .37$). Although the effort-contingent group showed a numerically larger PEP response, the distributions overlapped considerably.

The lack of a significant effect suggests that the manipulation did not evoke the expected physiological persistence in the effort-contingent group. From the perspective of Ryan and Deci's Self-Determination Theory (2000), both groups may have perceived the monetary rewards as "controlling" rather than "informational." As a result, the internalization of effort values in the effort-contingent group may not have occurred, which could explain why its participants failed to show sustained motivation.

Alternatively, the overall task demands and the laboratory context may have created a ceiling effect. Since all participants were university students in a laboratory, wearing heart rate sensors and performing complex tasks, the context itself may have been sufficient to elicit similar levels of effort mobilization across both conditions. As both groups might have tried just as hard, the benefits of the reward manipulation could have simply been masked. This lack of physiological differentiation is critical for interpreting the broader study. As Balconi et al. (2024) argue, decisions to engage in effortful behaviour are shaped by the subjective valuation of effort. If the experimental manipulation failed to alter this valuation, as the results in H2 indicate, this may explain the lack of effects observed in the behavioral delay discounting task (H1).

Hypothesis 3 (Exploratory)

Changes in Delay Discounting Over Time. H3 examined whether delayed gratification changed over the course of the experiment, independent of group contingency. The results revealed a small but significant increase in impulsive choices from pre- to post-test ($d = 0.25$).

Ainslie's (1975) model of specious reward and hyperbolic discounting could explain the direction of the effect. Ainslie posits that we discount rewards in a hyperbolic curve and claims that a reward becomes subjectively more valuable as it becomes more imminent. The cognitive load of the n-back task and the MET likely steepened this discount curve. Ultimately, then, the "specious" immediate options became disproportionally more attractive than the delayed options.

The overjustification effect (Deci et al., 1999) provides an alternative explanation. Since the design lacks a control group, it is possible that the external rewards in both contingencies contributed to a decrease in intrinsic interest in the unrewarded MET.

Crucially, baseline impulsivity was the only consistent predictor of change, pointing to the relative robustness of delay discounting. Participants who were highly impulsive at baseline tended to become only slightly more impulsive, or even to exhibit slight improvements. In contrast, participants who initially chose more delayed options showed a proportionally larger shift toward impulsivity.

However, this association likely reflects the statistical phenomenon of regression to the mean, where extreme scores naturally shift toward the average when retested. In other words, participants who emerged as highly patient in the baseline MCQ simply had more 'room' to deteriorate, whereas highly impulsive participants were already near the bottom of the scale and were more likely to choose more delayed options.

In addition, repeated administration of the same MCQ items may have contributed to the observed pre–post pattern. Because participants encountered the identical choice set at baseline and again after the task battery, post-test responding could have been influenced by familiarity with the questionnaire format, anchoring to prior responses, or reduced engagement when completing the second administration under fatigue. Such retest effects are especially relevant for the shortened 9-item version, where each item has a relatively large impact on the overall score. Therefore, part of the small shift toward impulsivity may reflect a measurement-related testing effect rather than a genuine change in intertemporal preferences.

These findings suggest that relatively short but exhausting experimental interventions, such as those used in the present study, produce only modest, state-like fluctuations that could likely be driven by measurement variance or cognitive load. The core self-regulatory trait, however, appears to remain broadly stable.

Exploratory Analyses of Mechanisms. Additional exploratory regressions did not support the idea that the observed decrease in delay of gratification was driven by other motivational variables such as monetary reward amount, physiological effort mobilisation, or self-reported fatigue. None of these predictors explained changes in MCQ scores after controlling for baseline impulsivity. This suggests that the small shift toward impulsivity was a general effect of the experimental context (i.e., the sustained cognitive load) rather than being driven by individual differences in subjective fatigue or reward sensitivity. In other words, although the task battery as a whole nudged the sample toward immediacy, variations in participants' fatigue did not predict the magnitude of this shift.

Theoretical Implications

The present findings contribute to debates on the adaptability of delayed gratification and the conditions under which effort-contingent rewards can shape self-regulation. The data support the view that self-control and delay discounting have a strong trait-like component with predictive power (Duckworth et al., 2013; Kirby et al., 1999; Moffitt et al., 2011). While small fluctuations were observed, baseline MCQ scores were the strongest predictors of post-task discounting. Also, exploratory analyses identified pre-task delay discounting as the only robust predictor of change. Hence, within this relatively brief experimental episode, pre-existing differences outweighed situational manipulations. Such results suggest that the self-regulatory capacities highlighted by Mischel et al. (1989) and Bembenutty (2011) may not be readily altered by brief interventions.

At the same time, the significant increase in impulsivity from pre- to post-test aligns with models of self-control that emphasize dynamic shifts in motivation and attention as key factors in self-regulation, rather than the depletion of a fixed resource (Inzlicht et al., 2014). After a sequence of demanding tasks, rather than showing disengagement, participants displayed a modest reorientation toward immediate rewards. On the one hand, therefore, the ability to delay gratification appears as a rather stable trait; on the other, however, it also proves to be influenced by situational and motivational factors.

In regards to the effects of effort-contingent rewarding, the results of this study did not confirm increased willingness to delay gratification. The main hypothesis, derived from learned industriousness theory (Eisenberger, 1992; Eisenberger & Adornetto, 1986) and recent work by Clay et al. (2022), posited that pairing effort with reward would increase participants' valuation of effort. Similarly, SDT (Deci & Ryan, 1985; Ryan & Deci, 2000) and attributional approaches to effort (Mueller & Dweck, 1998; Weiner, 1985) suggest that effort investment can support long-term goal pursuit by fostering autonomous motivation. Thus, it was hypothesized that rewarding effort would increase willingness to delay gratification.

However, in the data of this study, reward contingency did not exert a robust effect on delay discounting. While the tendency for the effort group to show greater resilience is directionally consistent with theoretical predictions, the influence of effort-contingent rewards on delayed gratification appears subtle at best.

Similarly, the study failed to link SDT and growth mindset theories with increased long-term goal pursuit. The null finding of this study does not undermine such a connection but indicates that simply linking monetary outcomes to effort in a working memory task does not automatically translate into a preference for long-term monetary decisions. Instead, it raises the question whether effort-focused reinforcement may simply strengthen intrinsic motivation and persistence within the trained domain itself (e.g., academic tasks; Bembenutty, 2011). Generalizing increased intrinsic motivation to a different abstract task, such as monetary decision-making, may require additional mechanisms that have not yet been accounted for.

Practical Implications

Although the present study did not find an effect of effort-contingent rewards on delay discounting, the results offer insight into the important effects of self-regulation and long-term goal commitment, which are relevant in educational fields.

As a single, brief intervention does not seem to reliably alter participants' evaluations of the rewards, effort-contingent rewards might not be effective when expected to be quick fixes. However, that does not mean that they could not still be helpful as part of long-term practices.

Additionally, the lack of a transfer effect could indicate the importance of domain alignment between the rewarded behavior and the targeted outcome. In this study, effort was rewarded based on a cognitively demanding working memory task (n-back), while delay of gratification was assessed via abstract monetary choices. While the n-back task may strengthen persistence and self-regulation within the trained domain, this effect may not be detectable by the MCQ, which employs a structurally different decision-making process. Therefore, to improve academic delay of gratification, effort-based reinforcement might be better embedded in relevant academic activities, with reflection on their long-term effects. This calls into question the habit of academic grading, which is most commonly based on rewarding performance rather than effort. However, it is important to note that the study does not argue for ignoring task outcomes; rather, it advocates a more holistic approach that incorporates individual effort in grading. It may help build greater resilience to strenuous tasks, thereby prolonging the effort-investment, which could ultimately translate into improved future performance.

Furthermore, the final shift toward more impulsive choices suggests that prolonged cognitive effort may lead individuals to prefer immediate rewards. Such a mechanism is critical for settings where important decisions are made after exhausting meetings or study sessions, as it reveals moments of reduced willingness to delay gratification. The results of this study should hint that making important decisions after cognitive stress should be avoided. Instead, short recovery periods and recreational breaks should be seen as innovative alternatives that may benefit long-term planning.

Baseline impulsivity has proven itself to be the most influential predictor of change. Hence, interventions will likely be most effective when tailored either to individuals or to groups that distinguish among individuals with high, middle, or low initial impulsivity. Since participants with higher initial impulsivity benefited most from the intervention, it can be hypothesized that this group would also benefit most in applied settings. Moreso, given prior scientific insights into self-regulation, a more detailed assessment of general self-regulation capabilities is suggested to help identify the pain points of highly impulsive individuals. Ultimately, such knowledge can be used to enable more specialized effort-focused interventions.

The present study underscored the importance of combining effort-contingent rewards with other evidence-based strategies. The literature on delay of gratification suggests classic cognitive strategies such as distraction, reframing, and the abstract representation of rewards, while motivational theories emphasize autonomy support and meaningful goal setting. Although the one-time effort-contingent rewards in this study did not produce a significant effect, the findings must be understood within the broader theoretical framework. As such, they advocate for a more detailed integration of effort-focused feedback in broader programs.

Limitations of the Study

While the present study provides valuable insights into effort-contingent rewards and delay of gratification, several limitations need to be acknowledged when interpreting the findings.

In terms of statistical power, the design was not optimized to detect minimal effects. An a priori power analysis was conducted for a medium-sized effect (Cohen's $f = .25$) with a planned sample of 128 participants. Consequently, effects smaller than the targeted effect would have been difficult to detect. In addition, the effective group sizes ($n = 42$ in the effort condition; $n = 71$ in the performance condition) fell short of the planned N and, thereby, further limited sensitivity for potential subtle differences (e.g., $f^2 \approx .01$). Thus, it remains possible that small but meaningful effects of effort-contingent rewarding on delay discounting exist but remained undetected in this study.

Furthermore, the present study employed only a single-session manipulation, in which participants were not informed of the exact contingency of their rewards. First, the manipulation may have been too brief to override pre-existing individual differences in self-control. Second, because participants were not told whether rewards depended on effort or performance, it is unclear to what extent they could form - even implicitly - a stable association between effort investment and the rewarding experience. The task itself provided no feedback other than the monetary amount shown at the end of each block, and informal debriefings suggested that many participants in the effort-contingent condition experienced these rewards as partly arbitrary. While the broader project aimed to test whether effort-contingent rewards can foster intrinsic motivation without explicit instruction, in the case of this smaller study on delayed gratification, this may have been an overly ambitious expectation. Moreover, the study did not include a control group because all participants received some form of reward. The absence of a no-reward control group further limits conclusions about the absolute impact of rewarding effort versus not rewarding at all. Overall, the manipulation likely remained too weak to show significant changes in the subjective value of effort. In combination with the brief, single-session exposure, this probably contributed to the absence of a clear physiological transfer effect in the unrewarded MET and to the overall small impact of reward contingency on delay discounting.

In addition, the distribution of earned monetary rewards differed between conditions. Although average payouts were broadly comparable, the effort-contingent condition showed greater variability and a more right-skewed distribution, with a small number of participants earning particularly high rewards. This is relevant because payout variability (and occasional very high rewards) may shape perceived reward salience, fairness, and motivational impact independently of whether rewards are contingent on effort or performance. Consequently, any effects (or null effects) cannot be attributed solely to contingency structure, as differences in the experienced reward distribution may have introduced additional noise or masked subtle contingency effects.

There are also measurement constraints to consider. For feasibility reasons, the original MCQ was shortened to a 9-item version and operationalized as the proportion of impulsive (smaller–sooner) choices. While this allows for capturing broad individual differences, it also reduces sensitivity to small, state-like changes and does not allow for stable estimation of individual discount parameters (e.g., k-values). Additionally, the overall level of impulsive choices in this sample was relatively low. Resultantly, this may have created a floor effect and limited the range of observable changes. Both these factors could cause subtle shifts in delay discounting to go undetected.

As participants were predominantly young, female psychology students at the University of Vienna, the characteristics of the sample restrict the generalisability of the findings. This group is likely homogeneous with respect to cognitive ability, educational background, and familiarity with psychological experiments, and may differ from the general population. Additionally, the context of being awarded study points for their participation could have led to increased or sustained motivation in the performance-contingent group. Such behavior would further blur the effect.

Furthermore, the laboratory context also limits generalizability. The decisions on delayed gratification were made in a controlled setting using a cognitive task and hypothetical monetary choices. They are not necessarily representative of real-life pressures that shape self-regulation in daily life. As such, it remains unclear to what extent similar reactions to effort-rewarding would emerge in more diverse samples or in naturalistic settings such as workplaces or classrooms, or how real financial decisions would be affected by it.

While baseline impulsivity emerged as a robust predictor of change, other personality and motivational variables were not assessed. The design of this study did not allow for investigating other individual differences that might influence responsiveness to effort-contingent rewards, such as, for example, conscientiousness, goal orientation or stress-level. Future work could incorporate such measures to further identify for whom and under which conditions effort-based rewards are most effective.

Finally, the exploratory analyses must be interpreted with caution. Conducting multiple exploratory tests increases the risk of Type I errors. Therefore, the exploratory results regarding potential mechanisms should be viewed primarily as hypothesis-generating for future research rather than as definitive evidence.

Future Research

The present study opens several avenues for future research on effort-contingent rewards and delay of gratification. First, future studies could strengthen reward manipulation. As the current single-session design provided only a brief exposure to effort-based reinforcement, multi-session interventions could test whether effort valuation can be conditioned over time.

Second, future studies should aim for larger sample sizes to ensure adequate statistical power and recruit a more heterogeneous sample to improve generalizability and enable detailed moderator analyses. In combination with careful experimental control and statistical adjustment for potential confounds, this would help clarify whether the observed patterns hold across different populations and under varying contextual conditions.

Third, future research should focus on aligning training and the outcome domain. A domain-congruent study design, for example, would investigate the rewarding nature of effort in academic tasks and subsequently measure academic delay of gratification rather than hypothetical monetary choices. Such a design could clarify whether effort-contingent reinforcement is domain-specific or influences delayed gratification in general.

Fourth, the questionnaire itself could be further refined. Rather than using a shortened version of the questionnaire, it would be recommended to use the full MCQ. This would enable estimation of individual discount parameters (k-values) and improve differentiation of individual choices.

Fifth, future designs could reinforce the target behavior more directly by rewarding delayed choices themselves. For example, studies could implement intertemporal choice paradigms in which a subset of delay-discounting decisions is paid out for real (e.g., one randomly selected choice), or provide bonuses contingent on selecting the larger-later option. This would increase domain congruence between reinforcement and the outcome and clarify whether rewarding effort transfers to patience, or whether delay behavior requires direct reinforcement.

Finally, exploring the effects of effort-rewarding on delayed gratification would probably be best analyzed through a longitudinal design. Longitudinal designs, including follow-up assessments, would be able to analyze whether the tendency to choose the more immediate reward after a strenuous task is due to a short-lived decrease in motivation or enduring changes in self-regulation. Additionally, it could help investigate moderators such as intrinsic motivation, conscientiousness, and "need for effort." Furthermore, longitudinal studies would help identify which personality profiles respond best to effort-contingent rewards, ultimately improving not only effort-rewarding interventions but also their impact on long-term goal pursuit in individuals.

Conclusion

The present study examined whether rewarding effort rather than performance can shift delay of gratification within a single experimental session. Theoretically, the study contributes by bringing together effort-contingent reward, delay discounting, and physiological measures of effort within one experimental framework. Across this framework, observed effects were small, suggesting that brief and relatively implicit reward contingencies may be insufficient to reliably alter intertemporal choice in a high-functioning student sample. At the same time, the data are compatible with the view that demanding cognitive contexts can modestly shape short-term preferences, highlighting the importance of situational factors alongside stable individual differences. Future research should therefore test more sustained and transparent manipulations, improve alignment between the domain of reinforcement and the domain of outcome assessment, and incorporate longitudinal follow-ups to clarify whether any effects reflect transient state shifts or durable changes in self-regulation.

References

Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin, 82*(4), 463–496. https://doi.org/10.1037/h0076860

Balconi, M., Acconito, C., & Angioletti, L. (2024). Not everyone chooses profit (if it is too tiring): What behavioral and EEG data tell us. *Applied Sciences, 14*(11), 4793. https://doi.org/10.3390/app14114793

Bembenutty, H. (2007). Self-regulation of learning and academic delay of gratification: Gender and ethnic differences among college students. *Journal of Advanced Academics*, *18*(4), 586–616. https://doi.org/10.4219/jaa-2007-553

Bembenutty, H. (2009). Teaching effectiveness, course evaluation, and academic performance: The role of academic delay of gratification. *Journal of Advanced Academics, 20*(2), 326–355. https://doi.org/10.1177/1932202X0902000206

Bembenutty, H. (2011). Academic delay of gratification and academic achievement. *New Directions for Teaching and Learning, 2011*(126), 55–65. https://doi.org/10.1002/tl.444

Bembenutty, H., & Karabenick, S. A. (1998). Academic delay of gratification. *Learning and Individual Differences, 10*(4), 329–346. https://doi.org/10.1016/S1041-6080(99)80126-5

Clay, G., Mlynski, C., Korb, F. M., Goschke, T., & Job, V. (2022). Rewarding cognitive effort increases the intrinsic value of mental labor. *Proceedings of the National Academy of Sciences, 119*(5), 1–8. https://doi.org/10.1073/pnas.2111785119

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. ed.). Erlbaum.

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

Duckworth, A. L., Tsukayama, E., & Kirby, T. A. (2013). Is it really self-control? Examining the predictive power of the delay of gratification task. *Personality and Social Psychology Bulletin, 39*(7), 843–855. https://doi.org/10.1177/0146167213482589

Eisenberger, R. (1992). Learned industriousness. *Psychological Review, 99*(2), 248–267. https://doi.org/10.1037/0033-295X.99.2.248

Eisenberger, R., & Adornetto, M. (1986). Generalized self-control of delay and effort. *Journal of Personality and Social Psychology, 51*(5), 1020–1031. https://doi.org/10.1037/0022-3514.51.5.1020

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312*(5782), 1900–1902. https://doi.org/10.1126/science.1128898

Inzlicht, M., Schmeichel, B. J., & Macrae, C. N. (2014). Why self-control seems (but may not be) limited. *Trends in Cognitive Sciences, 18*(3), 127–133. https://doi.org/10.1016/j.tics.2013.12.009

Job, V., Mlynski, C., & Nikitin, J. (2024). Challenging the law of least effort. *Current Opinion in Psychology, 60*, 101881. https://doi.org/10.1016/j.copsyc.2024.101881

Kirby, K. N., Petry, N. M., & Bickel, W. K. (1999). Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *Journal of Experimental Psychology: General, 128*(1), 78–87. https://doi.org/10.1037/0096-3445.128.1.78

Mischel, W., & Ebbesen, E. B. (1970). Attention in delay of gratification. *Journal of Personality and Social Psychology, 16*(2), 329–337. https://doi.org/10.1037/h0029815

Mischel, W., Ebbesen, E. B., & Zeiss, A. R. (1972). Cognitive and attentional mechanisms in delay of gratification. *Journal of Personality and Social Psychology, 21*(2), 204–218. https://doi.org/10.1037/h0032198

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science, 244*(4907), 933–938. https://doi.org/10.1126/science.2658056

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., Houts, R., Poulton, R., Roberts, B. W., Ross, S., Sears, M. R., Thomson, W. M., & Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences, 108*(7), 2693–2698. https://doi.org/10.1073/pnas.1010076108

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

List of Figures

List of Tables

List of Abbreviations

ADOG	Academic Delay of Gratification
AI	Artificial Intelligence
ANOVA	Analysis of Variance
ANCOVA	Analysis of Covariance
APA	American Psychological Association
BMI	Body Mass Index
CI	Confidence Interval
df	Degrees of Freedom
HR	Heart Rate
MCQ	Monetary Choice Questionnaire
MET	Math-Effort Task
PEP	Pre-ejection Period
R	(Programming Language)
SE	Standard Error
SD	Standard Deviation
t	t-test statistic
p	Probability value
F	F statistic
η^2	Eta squared (effect size)

AI Disclosure

AI-based tools were used solely in a supportive capacity. All content, arguments, and analyses reflect my own critical engagement with the subject matter. I acknowledge the use of ChatGPT for assistance with R programming and for refining and clarifying my own writing - particularly in summarizing and rephrasing longer passages, such as in the data reporting section. At no point were text passages copied or inserted without revision. Grammarly was used to support language editing and stylistic polishing. All final decisions regarding content, structure, and data analysis were made independently by me.

Appendix

Detailed ANCOVA Diagnostics

Independence of observations was assumed, given the individual testing procedure and the random assignment of participants to the two reward conditions. Linearity between the covariate and the dependent variable was examined by scatterplots of baseline and post-test MCQ scores within each group, which overall suggested an approximately linear relationship. The MCQ is an established measure of the ability to delay gratification, and visual inspection of the plots (residuals vs. fitted values and residuals vs. leverage) did not reveal extreme outliers or influential cases.

The normality of residuals was evaluated using the Shapiro–Wilk test, which indicated a deviation from normality, $W = 0.95$, $p = .0004$. However, Q–Q plots showed that this deviation was mainly due to heavier tails, while the central part of the distribution fitted the normal line to a great degree. Furthermore, given the sample size ($N = 113$), ANCOVA is generally considered robust to moderate non-normality. Homogeneity of variance was supported by Levene's test, $F(1, 111) = 0.13$, $p = .72$. Homogeneity of regression slopes was examined by adding the Group × MCQ_pre interaction to the model. This interaction was not significant, $F(1, 109) = 0.09$, $p = .77$, indicating that the relationship between baseline and post-task MCQ scores did not differ across groups.

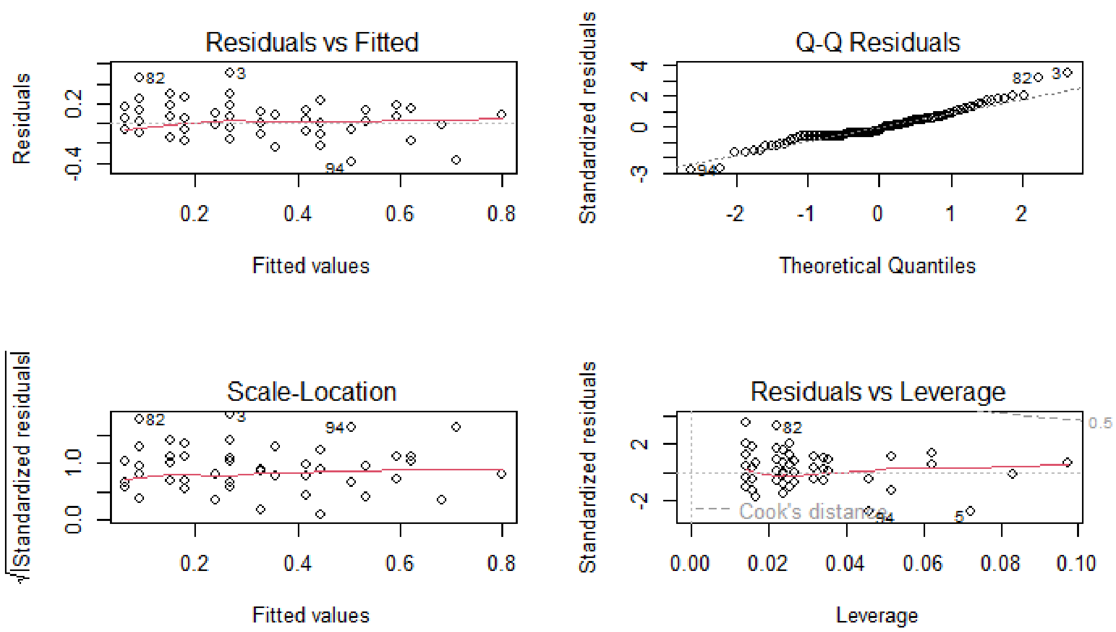

Figure A.1. Residual diagnostic plots for the ANCOVA model.

Inclusive Sample Results

To examine whether excluding participants who earned less than 2 € after task 1 affected the results, all analyses were repeated on an inclusive sample that retained these cases. This inclusive sample consisted of N = 178 participants with complete demographic data and N = 138 participants with complete MCQ data (effort n = 67, performance n = 71).

Demographic characteristics were highly similar to the primary sample. Descriptive MCQ scores showed the same pattern as in the primary analyses: both groups became more impulsive from pre- to post-test, and the performance group consistently showed slightly higher impulsive choice proportions than the effort group (see Figure A.2).

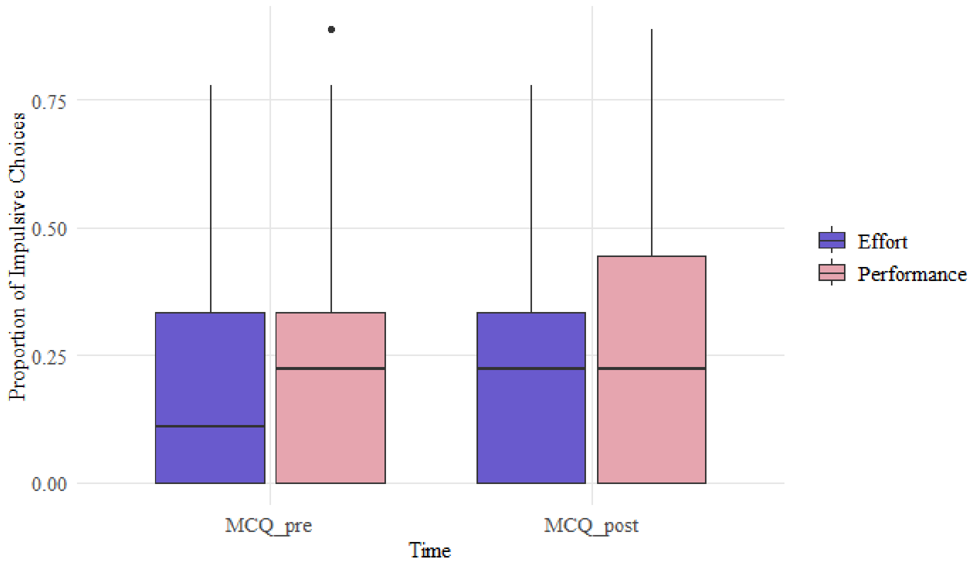

Figure A.2. MCQ pre- vs post scores by group (inclusive sample).

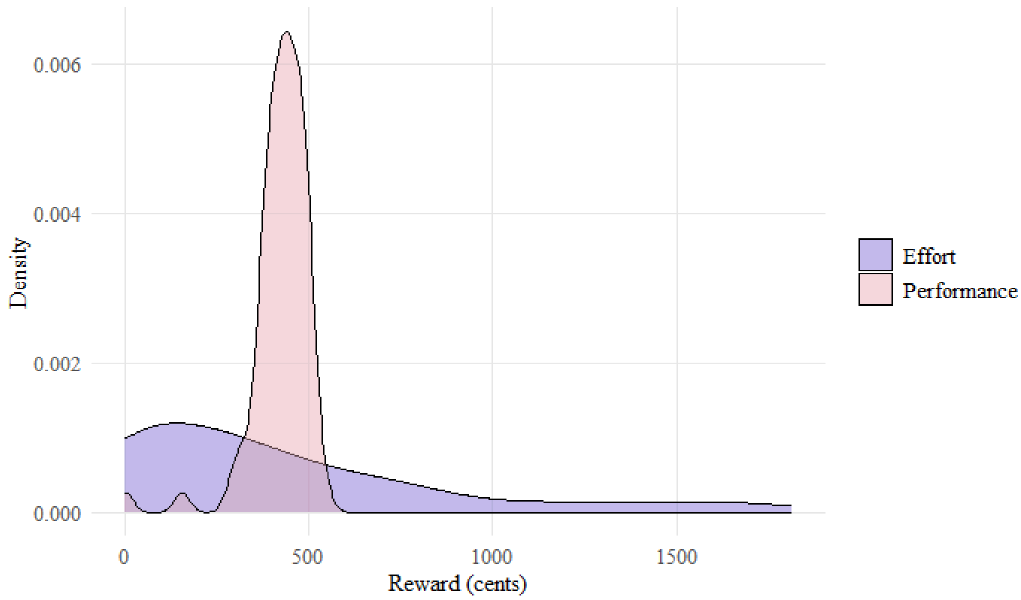

Figure A.3. Reward density by group (inclusive sample).

H1: ANCOVA Analysis (Inclusive Sample)

To test whether excluding participants who earned < 2 € after Task 1 influenced the findings, all analyses were repeated with an inclusive analytic sample in which these participants were retained (see Table A.1). Descriptive and inferential results were very similar across samples: the performance group consistently showed slightly higher adjusted post-MCQ scores than the Effort group, but the group effect remained small and statistically non-significant in both analyses (partial η^2 = .01 vs. .02). Thus, the main conclusions do not depend on the reward-based exclusion criterion.

Table A.1.

Comparison of ANCOVA Results for Main vs. Inclusive Sample

	Main Sample	**Inclusive Sample**
N (MCQ complete)	113	138
n Effort / n Performance	42/71	67/71
Adj. post-MCQ mean (Effort)	0.230	0.224
Adj. post-MCQ mean (Perf.)	0.258	0.254
Group effect F (df_1, df_2)	1.29 (1, 110)	3.111 (1, 135)
Group effect *p*	.26	.08
Partial η^2 (group)	.01	.02
Covariate *F* (MCQ_pre)	165.00	223.47
Partial η^2 (MCQ_pre)	.60	.62
R^2 / adj. R^2	.60/ .59	.63 / .62
Contrast Effort-Perf.	-0.028 (*p* = .32)	-0.030 (*p* = .22)

H2: Manipulation Check (Inclusive Sample)

To assess the robustness of the manipulation check, the analysis was repeated on the full physiological sample, excluding participants based on reward amount. As shown in Table A.2, both the restricted (N = 146) and inclusive (N = 176) samples yielded very small, non-significant group differences in PEP change. Although the direction of the mean difference flipped between samples, the effect sizes were similar in magnitude, and the confidence intervals included zero in both cases.

Table A.2.
Comparison of PEP Change (ms) Between Reward Conditions in Main vs. Inclusive Sample

	Main Sample	**Inclusive Sample**
N (MCQ complete)	146	176
n Effort / n Performance	51/95	81/95
Effort: *M* (*SD*)	-1.39 (7.37)	0.78 (6.97)
Performance: *M* (*SD*)	-0.22 (7.70)	-0.22 (7.70)
Mean difference Effort - Perf	-1.17	1.00
t(df)	-0.89 (106.42)	0.91 (173.38)
p	.37	.37
Cohen's *d*	-0.15	0.14
95% CI for *d*	[-0.49, 0.19]	[-0.16, 0.43]

Note. PEP change = METPEPc; negative values indicate a shortening of PEP (greater effort mobilization).

Although the direction of the mean group difference in PEP change flipped between the two analyses, the pattern is best interpreted as a random fluctuation rather than a meaningful effect. In the main sample (excluding participants with rewards < 2 €), the effort group showed a slightly shorter PEP than the performance group (in the hypothesized direction), whereas in the inclusive sample, the effort mean PEP was approximately 1 ms longer than that of the Performance group. As illustrated in the boxplots (Figure Z), these mean differences are very small relative to the within-group variability and their confidence intervals both include zero (restricted: $d = -0.15$; inclusive: $d = 0.14$). Therefore, the apparent "flip" simply reflects sampling variability when adding more participants, and both analyses converge on the same conclusion that the reward manipulation did not reliably affect physiological effort during the MET.

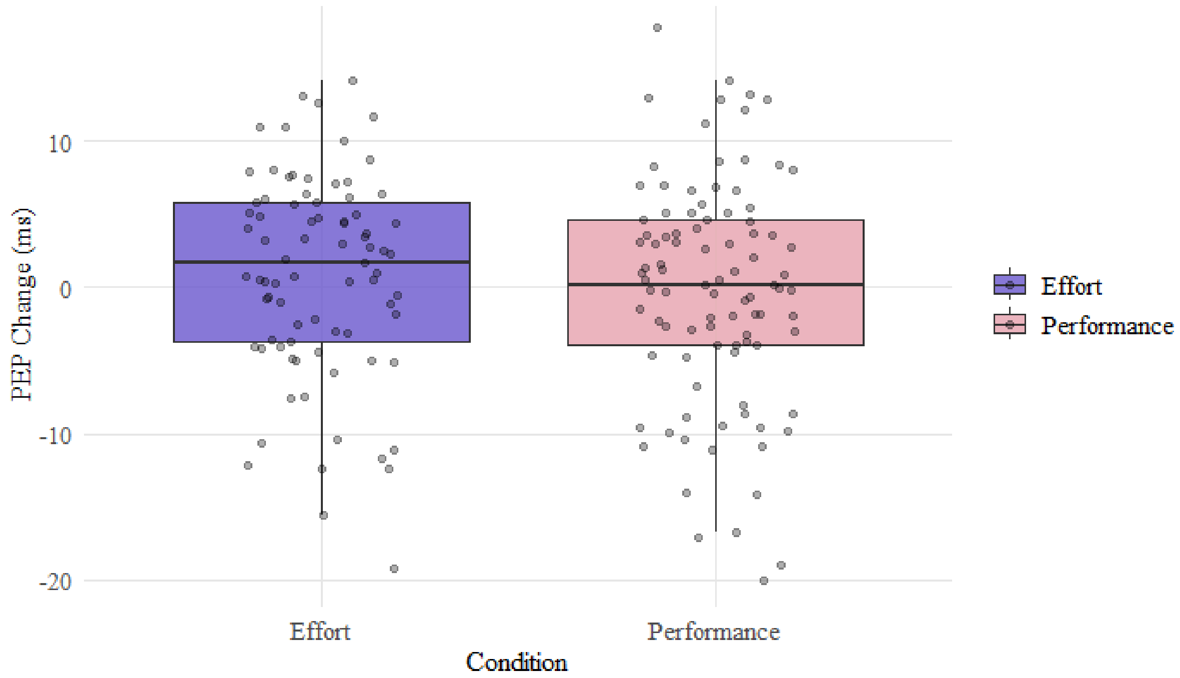

Figure A.4. PEP change during MET task by condition (inclusive sample).

H3: Inclusive Sample

Comparison of H3 Results for Restricted vs. Inclusive MCQ Samples

Table A.3.
Overall pre-post change in delay discounting (MCQ_pre - MCQ_post)

	Main Sample	**Inclusive Sample**
N	113	138
Mean diff (MCQ_pre - MCQ - post)	-0.038	-0.035
t(df)	-2.68 (112)	-2.84 (137)
p	.008	.005
Cohen's *d*	0.25	0.24
95% CI for *d*	[0.06, 0.44]	[0.07, 0.41]

Note. Negative values of MCQ_diff indicate a shift toward more impulsive choices. Cohen's *d* is computed as effort – performance; positive values therefore indicate a slightly smaller increase in impulsivity in the effort group compared to the performance group.

Table A.4.
Group differences in change scores (MCQ_diff = MCQ_pre - MCQ_post)

	Main Sample	**Inclusive Sample**
N (Effort / Performance)	42/71	67/71
Effort M_diff	-0.021	-0.022
Perf. M_diff	-0.049	-0.049
Welch's *t*(df)	1.00 (106.61)	1.09 (126.74)
p	.32	.28
Cohen's *d*	0.18	0.18
95% CI for *d*	[-0.20, 0.56]	[-0.15, 0.52]

Note. Negative values of MCQ_diff indicate a shift toward more impulsive choices. Cohen's *d* is computed as effort - performance; positive values therefore indicate a slightly smaller increase in impulsivity in the Effort group compared to the performance group.

Exploratory Analyses

Moderation

To explore whether the change in delay discounting depended on baseline impulsivity and reward condition, a linear regression with MCQ_diff (MCQ_pre – MCQ_post) as the outcome and MCQ_pre, group (Performance vs. Effort), and their interaction as predictors was conducted. The overall model was significant, $F(3, 109) = 3.87$, $p = .011$, $R^2 = .10$, adj. $R^2 = .07$. Baseline impulsivity showed a positive but only marginally significant association with MCQ_diff, suggesting that participants who were more impulsive at baseline tended to show slightly less deterioration (or even some improvement) in delay of gratification. Neither the main effect of group nor the MCQ_pre × Group interaction was significant, indicating no evidence that reward condition moderated this relationship. All results are shown in Table 4.5. Figure A.5 visualizes the interaction between baseline impulsivity and group on MCQ change.

Table A.5.

Moderation of Change in Delayed Gratification by Baseline Impulsivity And Reward Condition (Restricted Sample, N = 113)

Predictor	***b***	***SE***	***t***	***p***
Intercept	-0.06	0.03	-1.91	.058
MCQ_pre	0.18	0.10	1.82	.072
Group (Performance vs. Effort)	-0.04	0.04	-0.93	.353
MCQ_pre x Group (Performance)	0.04	0.13	0.30	.765

Note. Negative values indicate a decrease in PEP (greater sympathetic effort mobilization).

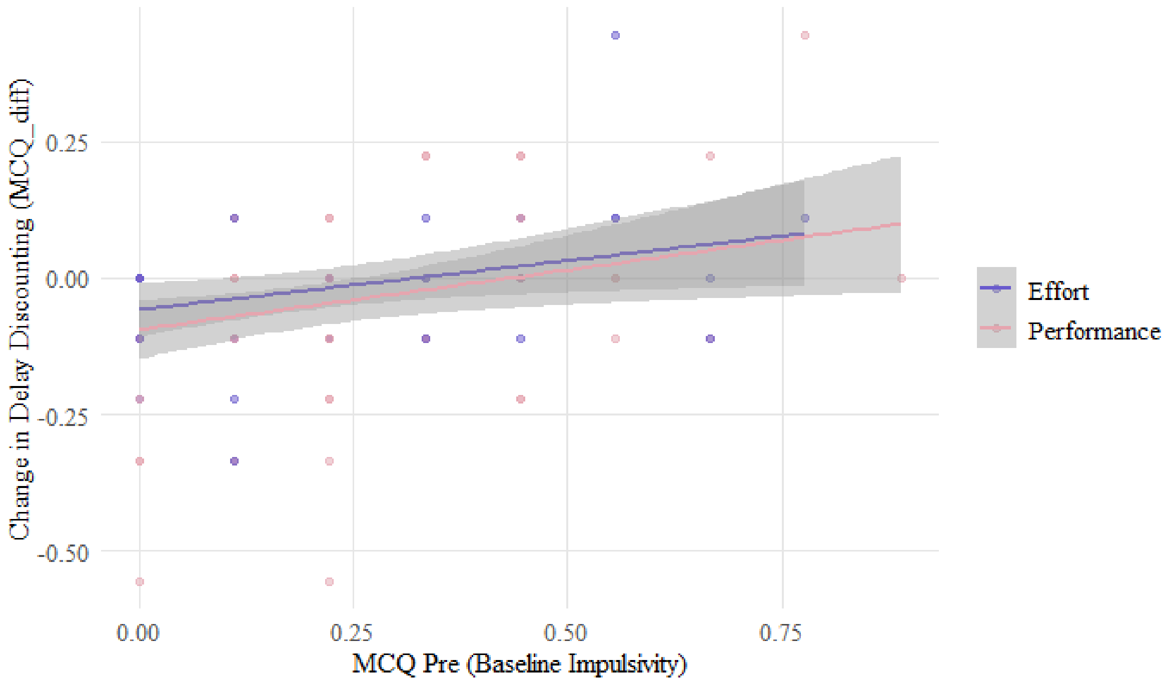

Figure A.5. Interaction between baseline impulsivity and group on MCQ change.

Regression Analysis

A further exploratory regression examined whether baseline impulsivity, reward condition, reward magnitude, and physiological effort (PEP change) jointly predicted change in delayed gratification. The model was significant, $F(4, 108) = 2.87$, $p = .026$, $R^2 = .10$, adj. $R^2 = .06$. Baseline impulsivity again emerged as the only significant predictor, indicating that participants with higher initial impulsivity tended to show less increase in impulsive choices (i.e., more positive MCQ_diff). Reward condition, reward amount, and PEP change were not significant predictors. The results are summarized in Table A.6.

Table A.6.

Regression predicting change in delay discounting (restricted sample, N = 113)

Predictor	***b***	***SE***	***t***	***p***
Intercept	-0.05	0.04	-1.25	.213
MCQ_pre	0.20	0.06	3.24	.002
Group (Performance vs. Effort)	-0.03	0.03	-1.00	.319
Reward (cents)	-0.00001	0.00005	-0.25	.807
METPEPc (PEP change, ms)	-0.0004	0.0020	-0.22	.829

Note. Positive MCQ_diff values indicate increased delay of gratification from pre- to post-test. A positive METPEPc indicates a decrease in effort-taking.

Further explorative checks in an attempt to explain the direction of the effect of H3

Cognitive Fatigue. For the subset with complete fatigue data (n = 82), participants reported higher fatigue after the task battery than at baseline, $t(81) = 9.10$, $p < .001$, with a mean difference of 0.40. However, change in fatigue was not related to change in delay discounting, $r(80) = -.04$, $p = .71$. In regression models including reward condition and baseline impulsivity, change in fatigue (or post-MET fatigue) did not significantly predict MCQ_diff, whereas higher baseline impulsivity did. Thus, although participants became more fatigued, fatigue does not appear to explain the observed decrease in delay of gratification.

Reward Amount. To examine whether the actually earned reward influenced change in delay discounting, the correlation between reward amount and MCQ_diff was calculated. Reward was not related to change in impulsivity, $r(112) = .01$, $p = .91$. In a regression model including reward condition and baseline impulsivity, reward amount did not significantly predict MCQ_diff, whereas higher baseline impulsivity did (p = .0015). Thus, within the observed range, the size of the monetary reward itself does not appear to explain the change in delay of gratification.

Physiological Effort (PEP). To examine whether physiological effort during the MET task was associated with changes in delay discounting, the correlation between PEP change (METPEPc; more negative values indicating stronger effort mobilization) and MCQ_diff was computed. The association was essentially zero, $r(112) = -.01$, $p = .93$, 95% CI [−.19, .18]. Thus, within this sample, participants who mobilized more (or less) physiological effort did not show systematically different changes in delay of gratification.

Baseline Impulsivity as Predictor of Change. To examine whether initial impulsivity predicted changes in delay discounting, MCQ_diff was regressed on baseline MCQ_pre and reward condition (and their interaction). The overall model explained a small portion of variance, $R^2 = .10$, $F(3, 109) = 3.87$, $p = .01$. There was a trend suggesting that participants with higher baseline impulsivity showed slightly larger improvements in delay of gratification, $\beta = 0.18$, $t(109) = 1.82$, $p = .07$. Neither the main effect of reward condition, $\beta = -0.04$, $p = .35$, nor the MCQ_pre × group interaction, $\beta = 0.04$, $p = .76$, were significant, indicating that this (trend-level) association did not differ between the Effort and Performance groups.

Explorative Analyses with the Inclusive Sample

The corresponding models fitted in the inclusive sample ($N = 138$) yielded a very similar pattern of results. Baseline impulsivity remained the only robust predictor of MCQ_diff, while reward condition, reward magnitude, physiological effort, and their interaction terms were not significant.

Abstract (German)

Verzögerte Belohnung (Delayed Gratification) ist ein zentraler Bestandteil der Selbstregulation. Unklar ist jedoch, ob die Verstärkung von Anstrengung, im Vergleich zu Leistung, intertemporale Entscheidungen tatsächlich verändern kann. Diese Studie prüfte daher, ob anstrengungskontingente Belohnungen die Fähigkeit zum Belohnungsaufschub im Vergleich zu leistungsabhängigen Belohnungen erhöhen. In einem randomisierten Between-Subjects-Design bearbeiteten 146 Psychologiestudierende der Universität Wien eine n-back-Aufgabe und erhielten dabei entweder anstrengungskontingente monetäre Belohnungen (operationalisiert über die Prä-Ejektionsperiode; PEP) oder leistungsabhängige Belohnungen (basierend auf Treffern und Fehlern). Anschließend absolvierten die Teilnehmenden eine Math-Effort-Aufgabe (MET), für die sie nicht belohnt wurden. Eine Woche vor dem Labortermin (Baseleine/Pre-Test), sowie unmittelbar nach der Aufgabenbatterie (Post-Test) beantworteten die Teilnehmer eine verkürzte Version des Monetary Choice Questionnaire (MCQ). Unter Kontrolle der Baseline-MCQ-Werte sagte die Belohnungskontingenz die Verzögerungsdiskontierung nach der Aufgabenbatterie nicht signifikant vorher, $F(1, 110) = 1.29$, $p = .26$, partielles $\eta^2 = .01$. Auch die Manipulationskontrolle ergab keine verlässlichen Gruppenunterschiede in der physiologischen Anstrengung während der MET ($p = .37$). Explorative Analysen zeigten über beide Bedingungen hinweg eine kleine, aber signifikante Prä-Post-Verschiebung hin zu impulsiveren Entscheidungen (Cohen's $d = 0.25$), was mit einer ermüdungsbedingten Zunahme der Diskontierung vereinbar ist. Insgesamt wurden kurzfristige Manipulationen der Belohnungskontingenz von stabilen individuellen Unterschieden überlagert, was dafür spricht, dass stärkere bzw. längerfristige Interventionen und eine engere Passung zwischen Trainings- und Outcome-Domäne notwendig sein könnten, um Belohnungsaufschub zu beeinflussen.