Developing an experimental paradigm for the study of loot boxes

by

Patrick McMillen Hancock

B.A., Metropolitan State University, 2021

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Psychological Sciences
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2025

Approved by:

Major Professor
Michael Young

Copyright

Abstract

Loot boxes are lucrative random reward systems prominently featured in the modern video game industry. Psychological research on loot boxes has relied primarily on self-report correlational studies, which have found concerning similarities between loot boxes and gambling that warrant further study in an experimental paradigm. However, loot boxes may also serve as a model to better understand human operant conditioning, which often fails to replicate the results of animal learning studies. In pursuit of both of these goals, we drew from the cognitive decision-making and operant conditioning literature to develop an experimental paradigm for the study of loot boxes. In two experiments, the paradigm was used to investigate how different reinforcement schedules (fixed-ratio, variable-ratio, and random-ratio) and schedule variability (Study 1), as well as variable reinforcement magnitude (Study 2), influenced players' decisions to open loot boxes. Participants in Study 1 initially opened loot boxes more on highly variable schedules, but the effect was short-lived. Study 2 found no effects of variable magnitude on loot box opening decisions. Although we failed to find support for any of our hypotheses, we discuss a number of modifications that can be made to the paradigm in future studies.

Keywords: Loot boxes, gambling, ratio schedules of reinforcement, variable magnitude

Table of Contents

List of Figures v
List of Tables vi
Acknowledgements vii
Dedication viii
Chapter 1 - The Emergence and Study of Loot Boxes 1
Chapter 2 - Loot Boxes as Cognitive Decision-Making 10
Chapter 3 - Loot Boxes as Operant Conditioning 14
Chapter 4 - Study 1: Schedule Variability 20
Chapter 5 - Study 2: Reinforcer Magnitude 45
Chapter 6 - General Discussion 61
References 70
Appendix A - Study 1 Supplemental Bayesian Figures 80
Appendix B - Study 2 Supplemental Bayesian Figures 82

List of Figures

Figure 1*: Probability of Reinforcement per Trial Under Different Reinforcement Schedules* 17
Figure 2*: Variability by Ratio Requirement Continuum* 22
Figure 3*: Study 1 - Screenshots from Game* 26
Figure 4*: Study 1 - Decision Tree* 30
Figure 5*: Study 1 - Boxplot of Response Requirements by Schedule* 34
Figure 6: *Study 1 - Probability of Loot Box Opening by Programmed MAD, Trial, and Round* . 39
Figure 7: *Study 1 - Individual Model Fits* 40
Figure 8: *Study 1 - Correlation Between Programmed and Experienced MAD* 44
Figure 9: *Study 2 - Example Magnitude Orders by Condition* 50
Figure 10: *Study 2 - Screenshots from Game* 51
Figure 11: *Study 2 - End of Trial Choices by Condition, Round, and Trial* 56
Figure 12: *Study 2 - Posterior Mean Distributions by Condition* 57
Figure 13: *Study 2 - Individual Model Fits* 58
Appendix A Figure 14: *Study 1 - End of Trial Choice Posterior Parameter Distributions* 80
Appendix A Figure 15: *Study 1 - End of Trial Choice MCMC Chains* 81
Appendix B Figure 16: *Study 2 - End of Trial Choice Posterior Parameter Distributions* 82
Appendix B Figure 17: *Study 2 - End of Trial Choice MCMC Chains* 83

List of Tables

Table 1: *Study 1 - Parameter Estimates Modeling End of Trial Choices*..... 38
Table 2: *Study 2 - Parameter Estimates Modeling End of Trial Choices*..... 55

Acknowledgements

This thesis is many years in the making, and there are therefore many acknowledgements that need to be made. I would first like to thank the faculty and staff of Inver Hills Community College, especially Dr. Julie Luker and Emily Johnson, as well as my friends Sommer and Blake. Similarly, I would like to thank Drs. Kimberly Halvorson and Kerry Kleyman, Kaitlyn Schwieters, and Valerie Wilwert from Metropolitan State University, all of whom gave me all of the research experience and support I needed to apply to a graduate program at all. The same is true of my friends Emily, Rachel H., Lucia, Danielle, and Rachel S.

From Kansas State University, I would like to thank my undergraduate research assistants, Dalton Green and Kenzie Liby, for their help with data collection on both projects during the Fall 2024 and Spring 2025 semesters, my lab mates, Brian Howatt, Kari Payne, Robert Southern, and Lucas Watson, and my friends Miki Azuma and Shannon Ruble for keeping me sane in graduate school.

Lastly, I want to thank my mother Diane, my father Randal, my sister Laura, my husband Tony, and my friends Claire, Lizzie, Sarah, and Ezra.

Dedication

This thesis is dedicated to Laura Miller.

Chapter 1 - The Emergence and Study of Loot Boxes

Loot boxes are a feature of modern video games in which players can open 'mystery boxes' that reward them with randomly generated in-game items (Ballou et al., 2022; King & Delfabbro, 2018; Nielsen & Grabarczyk, 2019; Zendle & Cairns, 2018). Players can obtain these loot boxes through a variety of means, such as completing in-game tasks or maintaining a daily login streak, but purchasing loot boxes directly via microtransactions is typically the fastest and easiest method (Ballou et al., 2022; Larche et al., 2023). Loot box contents vary greatly across games. While loot boxes in multiplayer first-person shooter games like *Fortnite* (Epic Games., 2017) and *Overwatch* (Blizzard Entertainment., 2016) reward players with costumes, poses, and emotes for the player's character, loot boxes in the strategy card game *Hearthstone* (Blizzard Entertainment, 2014a) instead reward cards that players can use to create a custom deck for competitive play. The diversity of loot box mechanics across different games has made the psychological study of them a significant challenge (Ballou, 2023; Ballou et al., 2022; Newall, 2023).

Loot boxes are prolific in the game industry. Zendle et al. (2020) inventoried the top-listed mobile games on the Apple and Google Play stores and found that a majority of the most popular games at that time featured loot boxes. Across genres and platforms, notable titles such as *Fortnite* (Epic Games., 2017)*, Overwatch* (Blizzard Entertainment., 2016), *Hearthstone* (Blizzard Entertainment, 2014a), *Star Wars Battlefront II* (Electronic Arts Digital Illusions Creative Entertainment, 2017)*, FIFA 22* (EA Vancouver & EA Romania, 2021), *Clash of Clans* (Supercell, 2012), *League of Legends* (Riot Games, 2009), *Pokémon Go* (Niantic, 2016), *Mario Kart Tour* (Nintendo EPD, 2019), *Counterstrike: Global Offensive* (Valve & Hidden Path Entertainment, 2012)*,* and *Rocket League* (Psyonix, 2015) have included loot boxes, although

several have since reworked their loot box systems (e.g., *League of Legends*) or removed loot boxes entirely (*Mario Kart Tour, Rocket League*) in response to public criticism. Regardless of how consumers feel about them, loot boxes are still extremely lucrative, with *Counterstrike* alone making a profit of nearly $1 billion off of its loot boxes in 2023 (Niemeyer, 2024).

Psychological research on loot boxes is still in its early stages. Up to now, the study of loot boxes has largely focused on establishing similarities between loot boxes and traditional forms of gambling (e.g., Drummond & Sauer, 2018; Zendle & Cairns, 2018). Whereas determining whether loot boxes pose a public health threat as an unregulated form of gambling is an important line of research in its own right, the widespread success of loot boxes in the gaming industry also suggests that loot boxes may have an untapped potential as a paradigm to study decision-making and operant conditioning with human participants.

Informed by the classic decision-making and operant conditioning literature, we developed an experimental paradigm modeled after loot boxes. In the following sections, we introduce the nascent loot box literature and its limitations, review the cognitive decision-making and operant conditioning literatures and discuss how they informed the design of our paradigm, present the results of two studies that demonstrate the paradigm's potential contributions to both the study of loot boxes as well as operant conditioning more broadly, and discuss how the paradigm can be refined in future studies.

The Psychological Study of Loot Boxes

Loot boxes are best known for the controversy they sparked in the late 2010s due to their resemblance to gambling. Because loot box rewards are randomly generated from large pools of possible rewards, with each reward having a different probability of occurring, players purchase loot boxes with no guarantee of ever winning their desired reward (Nielsen & Grabarczyk,

2019). Similar to a slot machine, players risk exorbitant spending if they chase after rare loot box rewards. In 2017, a 19-year-old Reddit poster, Kensgold, published an open letter to game developers where he detailed his experiences with loot boxes and urged game developers to address the dangers that loot boxes pose to players (Kensgold, 2017). A follow-up report published by the gaming outlet *Kotaku* verified that Kensgold had indeed spent $17,000 on loot boxes in *Star Wars Battlefront II, Counterstrike: Global Offensive,* and other games (Gach, 2017). Psychological researchers took notice of loot boxes the following year, with Drummond and Sauer (2018) arguing that "loot boxes are psychologically akin to gambling" (p. 530) and King and Delfabbro (2018) dubbing loot boxes a "predatory monetization scheme" (p. 1967) greatly in need of regulation.

As such, prior loot box research has focused on drawing parallels between loot boxes and gambling using cross-sectional surveys and crowdsourced participants (Brooks & Clark, 2019, 2023; Li et al., 2019; Macey & Hamari, 2019; Xiao et al., 2022; Zendle & Cairns, 2018, 2019). Past research has consistently reported concerning positive correlations between scores on the Problem Gambling Severity Index (PGSI, Ferris & Wynne, 2001) and self-reported loot box spending, with a meta-analysis by Garea et al., (2021) reporting an overall correlation of .26. However, the study of loot boxes has been hindered by the diversity of loot box mechanics and prizes that make generalization across games difficult. As such, the loot box literature is inconsistent in the measurement of some key variables, such as loot box spending, and lacks clear operationalizations for others, such as loot box engagement.

Although the amount of money that players spend on loot boxes is an important outcome within the literature, prior work has measured spending inconsistently. Loot box spending measurements include categorical classifications of whether participants had purchased a loot

box in the past year or not (Li et al., 2019), estimates of past month loot box spending using bracketed (e.g., $0 to $20) (Zendle & Cairns, 2018) or free response dollar amounts (Brooks & Clark, 2019, 2023; Drummond et al., 2020; Garrett et al., 2023; Macey & Hamari, 2019; Zendle & Cairns, 2019), and free response estimates of spending within the past year (Xiao et al., 2022). Curiously, Xiao et al. is one of the few studies that did not replicate the positive correlation between PGSI scores and loot box spending found in prior work, despite the fact that their timeframe was deliberately chosen to match the timeframe of the PGSI. As such, it is unclear whether the results of prior work are valid or merely due to methodological artifacts.

Beyond measurement, spending is also analyzed inconsistently across studies. Loot box spending distributions are known to be highly skewed, with the majority of participants reporting minimal (or no) spending and a small minority, known as "whales," that report spending exorbitant amounts of money (Close et al., 2021, p. 2). If the goal is to determine whether or not loot box mechanics put players at the risk of financial and psychological harm, then it is crucial for statistical analyses to incorporate high-spenders, who are likely the represent the individuals at the highest-risk. Despite this, many studies, such as Brooks and Clark (2023), Drummond et al., (2020) Garea et al. (2023), Garrett et al. (2023), and Hall et al. (2021) report excluding outliers in loot box spending according to data-driven cutoffs (e.g., 3.29 standard deviations above the mean or an ad hoc dollar threshold). While excluding outliers of loot box spending allows researchers to use linear analyses, it may also remove information about an important subgroup.

To further illustrate why this is an issue, consider the exclusions made by Garrett et al. (2023), which used the typical 3.29 *SD* cutoff. In Study 1, this resulted in the exclusion of participants who reported spending $165.99 or more on loot boxes in the past month, while in

Study 2, spending of $179.43 or more was excluded. Neither of these cutoff values are an objectively outlandish or unreasonable value for monthly loot box spending. Data-driven cutoffs may seem appropriate in isolation, but are senseless when compared across studies. A hypothetical participant that spent $170 on loot boxes would be excluded from Study 1, yet be included in Study 2, merely as a function of sampling error. Additionally, there is no guarantee that participants in cross-sectional studies are capable of or willing to accurately report their loot box spending. A study by Auer et al. (2023) found that although online gamblers were able to accurately recall their deposits (money placed into their online account for the express purpose of gambling) within the past month, high spenders systematically underestimated their spending more than low spenders. If the loot box and gambling parallels hold, the same may be true of loot box whales, further obscuring the true distribution of spending.

Beyond excessive spending, players may engage in other thoughts and behaviors that are harmful, such as playing a game for an excessive amount to time to earn free loot boxes. Brooks and Clark (2019) developed the Risky Loot-box Index (RLI), a five-item scale designed to measure participants' levels of risky engagement with loot boxes, including excessive time investments and preoccupations with loot boxes. RLI items include five-point Likert statements such as "the thrill of opening loot boxes has encouraged me to buy more" and "I have bought more loot boxes after failing to receive valuable items" (Brooks & Clark, p. 28). Initial validity evidence for the RLI was limited, with Brooks and Clark merely reporting that overall RLI scores were correlated with both self-reported loot box spending and ratings on an individual item that stated "My loot box use has caused me problems" (p. 31). The RLI has become widely used in loot box research (e.g., Brooks & Clark, 2023; Drummond, Sauer, Ferguson, et al., 2020; Forsström et al., 2022; Garrett et al., 2023; Larche et al., 2021, 2023; Primi et al., 2024; Spicer et

al., 2022; Xiao et al., 2024) and while these subsequent studies have consistently found positive correlations between RLI scores, self-reported loot box spending, and PGSI scores, psychometric evaluations of the RLI have cast doubt on the measure's continued use (e.g., Forsström et al., 2022; Primi et al., 2024; Sidloski et al., 2022). Notably, Sidloski et al. suggested that the observed correlations between RLI and PGSI scores may actually arise from participants providing similar answers to both measures precisely because they believe that loot boxes are a form of gambling.

Other markers of risky loot box engagement include items assessing whether participants have ever profited from selling loot boxes or loot box rewards (Brooks & Clark, 2019, 2023), known as a cash-out feature (Zendle, Cairns, et al., 2020), whether they believe that purchasing loot boxes has enhanced their experience of the game (Li et al., 2019), the age at which they first opened a loot box (Brooks & Clark, 2019), and their participation in or viewing of Esports content (Macey & Hamari, 2019). The construct of loot box engagement remains poorly defined, and many of the methods of measuring it are confounded with specific game features that can lead to false comparisons when data is aggregated across players of different games. For example, games such as *Overwatch* do not allow players to cash out for profit (although players may be able to trade with others or sell/purchase loot box rewards on a black market; Ballou et al., 2022). Comparing loot box engagement between individuals is only valid if all participants play games with this option. For instance, Player A, who plays *Counterstrike* and can sell items on the Steam market, should not be considered to have higher loot box engagement than Player B, who plays *Overwatch*, just because B indicates that they have never profited from loot boxes. Other game features that dictate how players engage with loot boxes are the ability to re-roll loot box rewards (e.g., *League of Legends*), exchange duplicate rewards for in-game currency that

may be used for future loot box purchases (*Hearthstone, Overwatch*), or purchase loot box rewards directly, often at a higher cost or packaged alongside other items (*Fortnite*). To resolve these issues, some studies have limited their samples to players of a single game (e.g., Larche et al., 2021; Zendle, 2019), but an experimental approach would allow game mechanics and features to be standardized across participants and provide more objective measures of engagement, such as the raw number of loot boxes purchased.

Other than the methodological issues in loot box research, Newall (2023) has highlighted that focusing solely on how loot boxes resemble gambling overlooks their unique characteristics. For instance, loot boxes lack an objective lose condition. Unlike slot machines, where players risk betting money to win nothing, loot boxes always provide rewards, although those rewards may be duplicates or undesirable to the player. Research has also yet to establish a clear relationship between loot box use and impulsivity (Garrett et al., 2023; Spicer et al., 2022; Zendle et al., 2019), even though high impulsivity is typically linked with problem gambling behaviors (MacKillop et al., 2014). Larche et al. (2023) proposed that this may be due to the presence or absence of loot box earning mechanics (i.e., earning loot boxes is typically slower and requires self-control), and found that participants who preferred to earn loot boxes demonstrated a greater ability for planning and executive function than loot box purchasers according to scores on the Barratt Impulsiveness Scale (Patton et al., 1995). Additionally, loot box opening animations are designed to be visually and auditorily appealing to players (Kao, 2019), potentially serving as a form of reinforcement or acting as a cue to operant responding. Prior studies comparing spending on booster packs in physical trading card games and loot box spending have found little correlation between the two (Mattinen et al., 2023; Zendle et al., 2021), suggesting that loot boxes as they appear in modern games may be uniquely more

effective at incentivizing spending. Most importantly, the ability to profit from selling loot box rewards is uncommon in the games industry and players are instead motivated to open loot boxes by their desire to collect digital items and gain gameplay advantages (Zendle et al., 2019). An experimental loot box paradigm could address these issues by manipulating the amount and type of rewards that are available, the ways in which loot box rewards impact core gameplay, earning mechanics for free loot boxes, and opening animations.

Experimental Designs for Loot Boxes

An experimental paradigm for loot boxes is frequently suggested as a possible solution to the limitations of cross-sectional surveys (e.g., Ballou, 2023; Brooks & Clark, 2023; Kao, 2019; Zendle et al., 2020; Zendle & Cairns, 2018), but only a handful of studies have utilized experimental or even quasi-experimental designs. Zendle (2019) conducted a natural experiment by analyzing *Heroes of the Storm* (Blizzard Entertainment, 2014b) players' spending and PGSI scores before and after the game removed loot boxes. The study found that players with high PGSI scores spent less money after loot boxes were removed, suggesting that loot boxes did in fact incentivize spending above and beyond direct purchase microtransactions. Two studies by Larche et al. (2021, 2023) showed video clips of the opening of loot boxes to a sample of *Overwatch* players and measured participants' self-reported arousal, disappointment, and frustration after each. Although Larche et al.'s design overcame the issue of varied loot box implementation by restricting the sample only to *Overwatch* players, their results are still subject to concerns of ecological validity (i.e., whether behavior elicited by the video stimuli is representative of how players would behave in response to real loot boxes with real stakes). The only true experimental design to date was Kao (2019) which developed a custom *Unity* platform, *Infinite Loot Box,* in order to investigate the impact of different loot box opening animations (i.e.,

the presence and quality of audiovisual effects) on behavior. Despite being open-source, no study has made further use of *Infinite Loot Box*.

Although developing an experimental paradigm for loot boxes presents a serious challenge, the benefits to the study of loot boxes are clear. Unlike self-report measures, an experimental design is able to objectively track behaviors, such as the number of loot boxes opened, and model individual behavior over time. Experimental data could then be compared to data from self-report measures established in the literature (e.g., the PGSI and RLI) using a multitrait-multimethod matrix (Campbell & Fiske, 1959), thus defining the boundaries between the constructs of impulsivity and loot box engagement and providing validation for both paradigms. A custom experimental paradigm would be able to manipulate game features (e.g., how loot boxes are obtained and their impact on gameplay) and ensure that all participants play the same game, allowing for direct behavioral comparisons. Loot boxes are an opportunity to innovate the study of experiential learning in humans, but understanding and defining loot boxes requires integrating multiple areas of psychological research. As such, the following sections discuss loot boxes from the cognitive decision-making and Behaviorist operant conditioning perspectives.

Chapter 2 - Loot Boxes as Cognitive Decision-Making

In traditional judgement and decision-making (JDM) terms, loot boxes represent decisions made under uncertainty (or in certain cases discussed later, decisions under risk; see Knight, 1921). Because loot box rewards are randomly generated, it is impossible to know precisely what will be received or when. Instead, players must try to predict outcomes and make decisions based on different sources of information. When relying on the results of prior loot box purchases, players make decisions from experience (also known as experience-based decision-making; Hertwig et al., 2004; Hertwig & Pleskac, 2010; Rakow & Newell, 2010). When explicit probability information is available (and in the absence of prior experience), players make decisions from description (also known as descriptive choice, Kahneman & Tversky, 1979). Under rational principles, the prospect or choice that results in the maximum amount of gain should be chosen, but human decisions from both description and experience[1] demonstrate systematic deviations from optimality.

Prior work has found that human decision-making demonstrates an inconsistent sensitivity to a prospect's expected value (EV), the average outcome over time (Kim, 2018) in decisions under risk. Following Knight's (1921) distinction, decisions under risk (also known as descriptive choice) are those in which explicit and perfect probability information about each outcome is available. This probability information can be used to calculate the EV of each prospect by multiplying the magnitude of an outcome (V) by its probability of occurring (p).

[1] Decisions can be (and in real-life, often are) made from both descriptive information and experience. For simplicity, we focus only on results that isolate the two sources of information, but see Newell and Rakow (2007) for an example of how the two sources of learning impact each other.

Although EV represents the mathematically optimal choice, it fails as a model of human decision-making (Harless & Camerer, 1994). For example, consider the St. Petersburg paradox (Bernoulli, 1713; Hayden & Platt, 2009; Martinez & Zeilberger, 2024; Seidl, 2013). In its simplified form, the St. Petersburg paradox is a hypothetical game of chance in which the player pays an initial fee, then flips a fair coin. If the coin lands on heads, the player doubles their winnings and flips the coin again. If the coin lands on tails at any point, the player is forced to end the game and leave with their winnings up to that point.

In this scenario, optimal decision-making dictates that the player only agree to pay an initial fee that is less than what they would expect to win during the game, thus turning a profit. The paradox lies in the fact that the EV for the game is infinite, as the coin flip could, in theory, result in heads ad infinitum. The conjoint probability of each successive heads flip is asymptotic; longer and longer streaks are increasingly unlikely, but never impossible. If the EV of the game is infinite, players should therefore be willing to pay an infinite amount of money to play the game. In practice however, human participants have no use for theoretical profits, with most only willing to pay a few dollars to play the game (Hayden & Platt, 2009). Similarly, just because loot boxes do not have a lose condition (and therefore, always a positive EV) does not mean that all of the possible rewards are equally worth the investment.

Another scenario in which EV diverges from actual human decision-making is in the evaluation of risky prospects. When given the choice between a guaranteed gain and a probabilistic gain, participants often demonstrate a strong preference for the certain prospect (risk aversion), even when the probabilistic gain has a higher EV (Kahneman & Tversky, 1979). Furthermore, people typically overweight the probability of rare outcomes (that is, low

probability outcomes) in descriptive choice (Hertwig et al., 2004), meaning that they behave and make choices as if rare outcomes are far more likely to occur than they truly are.

Certainty effects and the overweighting of rare events are put into practice in two mechanics that are ancillary to loot boxes: pity timers and probability disclosures. Pity timers are a game mechanic in which players are guaranteed to receive a certain loot box reward after they have opened a specified number of boxes. Pity timers are regarded positively within the gaming community (Xiao et al., 2022), perhaps in part due to the established preference for certainty (i.e., players prefer the guarantee of getting what they want eventually, even if they have to open many loot boxes to secure that guarantee), although it is not clear whether pity timers lead to increased loot box spending or other forms of engagement. Meanwhile, probability disclosures are an intervention strategy intended to reduce loot box spending by requiring game developers to communicate the exact probabilities of obtaining each loot box reward to players. Probability disclosures have been mandated in several countries, including the UK, Belgium, and China (Xiao et al., 2022; Xiao & Newall, 2022). In theory, probability disclosures would discourage players from chasing the rarest rewards, but this strategy may backfire if small probabilities are indeed overweighted by players. There is very little research on the long-term impact of probability disclosures on spending and engagement, but initial evaluations suggest that probability disclosures alone are ineffective (Xiao & Newall, 2022).

It should be noted that risk aversion is not uniform across contexts and is typically seen when prospects are framed as gains (Kahneman & Tversky, 1979). In the context of loot boxes, prospects may be subject to alternative framings. For instance, do players perceive the choice between purchasing a loot box or not as a probabilistic gain (a potentially desirable reward from loot boxes) and a certain gain (saving money by not purchasing) or as a probabilistic loss (getting

an undesired item) and a neutral prospect (in that not opening loot boxes results in no change in the player's wealth)? Do players evaluate loot box outcomes using their current state or their expectations and aspirations as a reference point? To our knowledge, there is no peer reviewed research examining loot boxes from the lens of prospect theory, but we suspect that the latter may be true in both cases.

Decisions under risk and descriptive choice, however, are only one side of the coin. Decisions from experience (or decisions under uncertainty, according to Knight, 1921) demonstrate their own systematic deviations from optimality. As a matter of fact, experienced-based decisions have been found to be biased in the opposite direction of descriptive choice, with rare events actually being underweighted (that is, treated as less likely to occur than they actually are) rather than overweighted (Hertwig et al., 2004). This difference has been termed the description-experience gap (Hau et al., 2008; Hertwig & Erev, 2009; Nie et al., 2024) and there remains debate around the causes of this gap.

One common explanation is the role of sampling error and overgeneralization, in which humans make conclusions from a limited sample of outcomes (Fox & Hadar, 2006). In the case of loot boxes, especially in games where descriptive probabilities are not provided, the first few rewards that players encounter may heavily bias their expectations going forward (similar to the phenomenon of big wins in the gambling literature, Kassinove & Schare, 2001). While the last 20 years of decision-making research has seen an increased interest in experience-based decision-making, classic animal learning studies from the Behaviorist tradition have already made significant strides in the area of experiential learning.

Chapter 3 - Loot Boxes as Operant Conditioning

Loot boxes are perhaps best modeled by the process of operant conditioning, where outcomes influence the frequency of behaviors by altering the underlying stimulus-response (S-R) association through reinforcement (Ferster & Skinner, 1957; Thorndike, 1901). The association between loot boxes (S) and opening them (R) becomes stronger when players are rewarded with cards, costumes, and coins (reinforcers), thus increasing the likelihood that players will continue engaging in behaviors that allow them to open loot boxes. Because loot boxes deliver reinforcement when a quota of operant responses has been met (i.e., the player has opened a requisite number of loot boxes), loot boxes follow a ratio schedule of reinforcement. Ratio schedules include fixed-ratio (FR), variable-ratio (VR), and random-ratio (RR). Ratio schedules elicit different patterns of operant behavior depending on the response requirement (i.e., how many operant responses need to be made to receive reinforcement) and the schedule type.

In an FR schedule, the number of responses required for reinforcement is constant. Under an FR 10 schedule, for example, reinforcement is always delivered after the 10^{th} response. Continuous reinforcement (CRF), a special case of an FR schedule, delivers reinforcement after every response (i.e., an FR 1 schedule). Operant responses are learned quickly on CRF schedules compared to other ratio schedules (e.g., intermittent schedules, see Gottlieb, 2004 and Papini & Overmier, 1985), but said responses are extinguished just as easily (Gibbon et al., 1980). Terminal response rates may persist when transitioning from a CRF to another FR schedule, although responding may cease entirely if the new ratio requirement is too large, also known as ratio strain (Ferster & Skinner, 1957). Whereas the stability of FR and CRF schedules makes them easier to learn, intermittent schedules, including VR and RR schedules, have been shown to

have higher terminal response rates (i.e., asymptotic performance) and are highly resistant to extinction (Gibbon et al., 1980), with B. F. Skinner (1953) himself crediting the success of casinos to the power of VR schedules of reinforcement.

In a VR schedule, the number of responses required is subject to change each time reinforcement is received. VR schedules are defined by their mean number of required responses and specified range of responses. Two VR schedules may have equivalent mean response requirements but very different ranges. Consider a VR 10 schedule that uniformly ranges from nine to eleven (i.e., reinforcement may be received after the 9^{th}, 10^{th}, or 11^{th} response). The same is true of a VR 10 schedule with a range from one to nineteen. Extending the terminology of mixed reinforcement schedules (MR, schedules where two or more independent schedules of reinforcement are alternatively available), every possible ratio requirement of a VR may be thought of as a component schedule (Ferster & Skinner, 1957). Therefore, VR schedules may also differ in their number of component schedules. For instance, a VR schedule where reinforcement may be received after the 1^{st}, 9^{th}, and 20^{th} response (with each requirement being equally likely) would have three component schedules, identical to the aforementioned VR 10 (9-11). Although the number of component schedules[2] and schedule range are not fully independent (i.e., reducing or expanding the range of a VR schedule often results in changing either the range or the mean number of required responses, and vice versa), both are necessary to fully understand just how variable a VR schedule might be.

[2] In studies of MR schedules, ratio requirements often vary more substantially (e.g., requirements of 2 and 18 or 4 and 36 used in (Mazur, 1983) than those discussed here, with small differences possibly being too subtle to significantly impact behavior. We use the term component schedules here for lack of a better term in the literature.

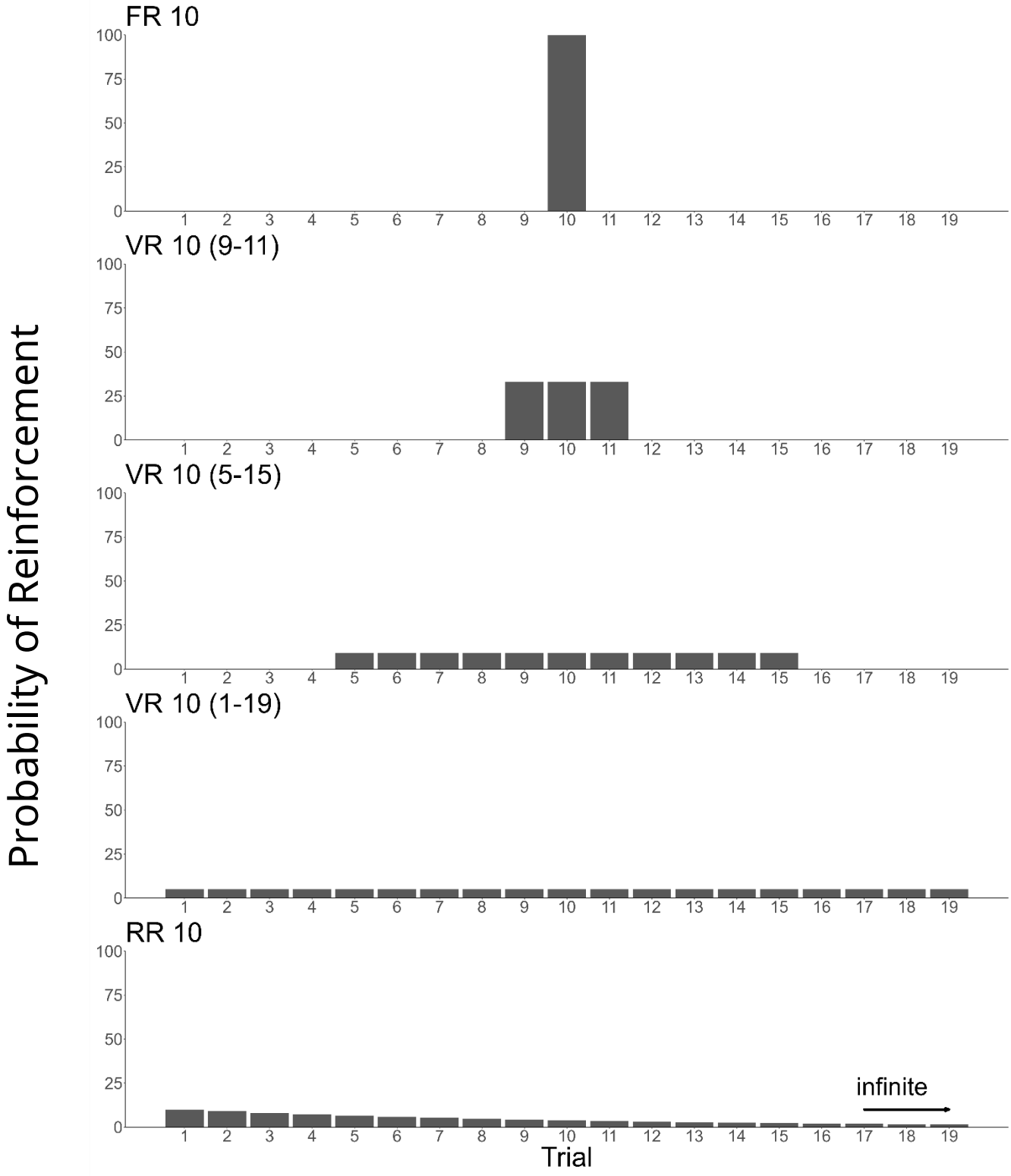
Often overlooked in the literature, a random ratio (RR) schedule maintains a constant probability of reinforcement for each response (e.g., on an RR 10, an average of one out of every 10 responses will receive reinforcement, making the constant probability of reinforcement 10%; Schoenfeld et al., 1956). An RR schedule is as truly random as machine and computer programming allows, with no defined range or number of component schedules. Reinforcement may be received after any response, from the very 1st to the 1000th . The probability of being reinforced on the n^{th} response (p_n) can be calculated using Equation 1 below, with P being the constant chance of reinforcement and n being the response number (see Figure 1 for a comparison of FR, VR, and RR distributions with equivalent mean number of required responses):

$$p_n = P * (1 - P)^{n-1}. \quad (1):$$[3]

Because the probability of receiving reinforcement after each response is conditional upon not receiving reinforcement after all previous responses, long streaks of unreinforced responses become increasingly unlikely. For example, on an RR 10 schedule, there is a 10% chance of receiving reinforcement after the first response, a 9% chance of receiving reinforcement after the second response, a 3.5% chance of receiving reinforcement after the tenth response, and so forth.

[3] This equation is an equivalent functional form of the equation to calculate the conjoint probability of successively flipping heads in the St. Petersburg paradox and therefore follows the same geometric distribution.

Figure 1: *Probability of Reinforcement per Trial Under Different Reinforcement Schedules*



The existing loot box literature is inconsistent regarding what schedule of reinforcement loot boxes supposedly use. Most studies classify loot boxes as VR schedules (e.g., Drummond & Sauer, 2018) while others define them as RR schedules (Kao, 2019). It should be noted that disagreement over which reinforcement schedules are in use also exists in the gambling literature (see Crossman, 1983). Although the exact schedule that a given game uses cannot be determined without access to the game's code, loot boxes likely operate on RR schedules because the probability of receiving a potential reward remains constant regardless of how many loot boxes are opened. However, pity timers mechanics (Xiao et al., 2022) would actually more closely resemble a VR, in that reinforcement is guaranteed to be received after the maximum number of required responses[4]. Because players always 'win' when they open loot boxes, loot boxes could also be considered to operate on CRF schedules, with reinforcer magnitude varying rather than the timing of reinforcement.

Because probability disclosures are not uniform across games, the current studies prioritized loot box purchasing behavior in a purely experienced-based decision-making scenario. As such, the design of the following studies drew primarily from the operant conditioning literature, with some cognitive undertones. Whether FR, VR and RR schedules with equivalent number of mean responses elicit different operant behavior as a function of schedule variability in a simulated loot box task were the focus of Study 1. Study 2 brought the paradigm

[4] Pity timers may also be argued to be a VR schedule that is linked to an FR 1 (i.e., if an item is not received after a certain number of attempts, the schedule switches to an FR 1 where reinforcement is guaranteed after the next attempt). This distinction is not explored for the purposes of the present proposal, but may be relevant to future studies.

closer to real loot boxes by examining the impact of different schedules of reinforcement and variable magnitude (VM).

Chapter 4 - Study 1: Schedule Variability

The primary goal of Study 1 was to determine whether the different types of ratio schedules would make participants more or less likely to open loot boxes when the mean number of required responses was equivalent across schedules. Study 1 addressed a niche, but persistent, open question in the gambling literature and was well-suited for exploration with the new simulated loot box paradigm. Prior studies have compared operant behavior under different ratio schedules using concurrent and blocked designs. In concurrent designs, two or more schedules are available simultaneously and subjects can choose where to allocate their responses (e.g., Field et al., 1996). Subjects in blocked designs experience multiple blocks of conditioning, with each block implementing a different schedule of reinforcement (e.g., Crossman et al., 1987; Mazur, 1983). In both cases, operant behavior has been shown to be influenced by schedule variability.

Participants have been shown to prefer schedules with high variability in concurrent designs, even over some schedules with lower ratio requirements. A study by Field et al. (1996) found that pigeons preferred VR 60 schedules with high variability over FR 30 schedules, despite the difference in mean ratio requirements. Fifteen VR 60 schedules were presented, each with four component schedules that varied in range (e.g., 1-30-90-120, 15-30-90-106). The VR schedules with the largest range (i.e., the highest variability) were highly preferred over the FR 30 schedule, with Field et al. noting that VR 60 schedules containing a component schedule at or near one were especially preferential (although in this design the size of the minimum component schedule is confounded with range). Hurlburt et al. (1980) used electronic slot machines to offer concurrent VR and RR schedules, but did not find evidence of preference for either schedule. However, Hurlburt et al. did not report the mean number of required responses

for either schedule nor the specified range of the VR schedule used (Haw, 2008). As such, Hurlburt et al.'s results do not provide conclusive evidence whether humans prefer VR or RR schedules when the mean number of required responses are equivalent.

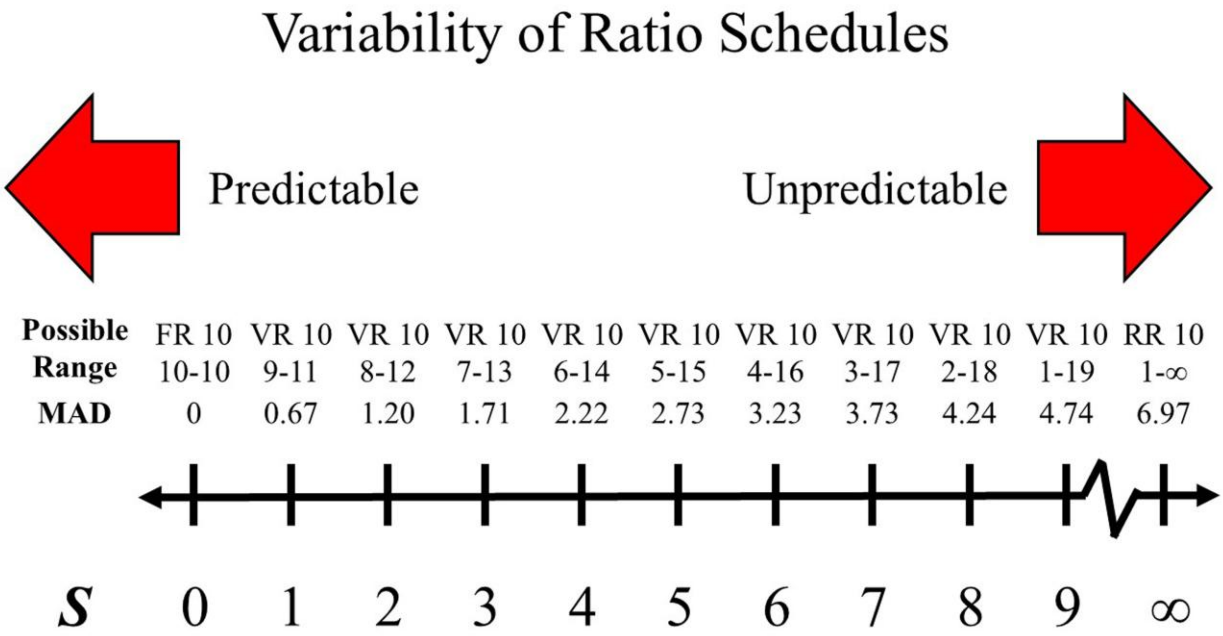
Behavioral data from studies using both concurrent and blocked designs suggest that behavior on FR, VR, and RR schedules of reinforcement is modified by schedule variability, although the overlap between schedule type, range, and number of component schedules make it difficult to precisely define variability using these measures alone. Instead, different schedules with the same mean number of required responses may be organized from least to most variable according to the mean absolute deviation (MAD) of ratio requirements. The formula for a uniformly distributed VR MAD is shown in Equation 2, where *S* is the absolute value difference of the minimum or maximum ratio requirement from the mean.

$$MAD = \frac{S(S+1)}{2 * S+1} \quad (2):$$

Based on MAD, FR schedules are the least variable because reinforcement can only be received at the exact ratio requirement (MAD = 0). Because RR schedules do not have a uniform distribution of response requirements, the RR MAD was computed using simulations. Unsurprisingly, the simulated RR MAD was much higher (MAD = 6.97 for an RR 10, with simulated ratio requirements ranging from 1 to 139) compared to other ratio schedules with equivalent mean number of required responses. VR schedules fall somewhere in between depending on their specified range. It may be the case that the less variable a VR schedule is (e.g., a VR 10 with a range of 9-11, MAD = 0.70), the more it elicits behavior characteristic of an FR schedule while VRs with higher variability may function similarly to RRs (especially those with a minimum ratio requirement of one, as in Field et al., 1996). This continuum of schedule variability and the calculated MAD for FR, VR, and RR schedules with mean number

of 10 required responses are shown in Figure 2. The potential of VRs to elicit behavior similar to both FRs or RRs depending on their variability may explain why previous work has failed to detect behavioral differences in humans under VR and RR schedules (e.g., Hurlburt et al., 1980).

Figure 2*: Variability by Ratio Requirement Continuum*



Given the unique properties of RR schedules (as shown in Figure 1), we predicted that the choice to open loot boxes would be influenced by the variability of the schedule of reinforcement. Based on Platt's (1964) method of strong inference, we proposed two competing hypotheses. Hypothesis H1 drew from the operant conditioning and experienced-based decision-making literatures and predicted that the number of loot boxes opened would increase as the variability of the schedule increases, meaning that participants would open the most loot boxes when they are exposed to the RR 10 and VR 10 (1-19) schedules. Conversely, Hypothesis H2 drew from the descriptive choice literature and predicted that human participants would prefer certain and predictable loot boxes over probabilistic ones, similar to pity timers. As such,

participants would open more loot boxes when exposed to the least variable schedules, namely the FR 10 and VR 10 (9-11).

Method

Participants

We recruited a total of 142 (77% female, M_{Age} = 19 years old) participants from Kansas State University's SONA subject pool. Participants received course credit in exchange for their participation. Given the novelty of the experimental design, prior literature was insufficient to conduct a formal power analysis, although the lack of clear differences in behavior across VR and RR schedules with equivalent mean response requirements suggested that the differences may be small, if present at all (Haw, 2008; Hurlburt et al., 1980; Mazur, 1983). Based on studies in our lab with similar designs and constraints on the number of participants that could be recruited within a semester (see Lakens, 2022 for further discussions on sample size justifications), we deemed a sample size of 100 participants to be appropriate. However, we continued data collection until the end of the semester in order to allow students to complete their required research credits.

Of the 142 total participants, 13 were ultimately excluded from the analyses. Eleven of these participants were removed because they played an alternative version of the game in which the first trial was rigged to always win. Because our statistical models would fail if there were too few instances of participants choosing to open loot boxes, we piloted this alternative version early on in data collection in order to explore whether we could leverage the established effects of big wins (e.g., Kassinove & Schare, 2001) to make opening the loot boxes more attractive if needed. This manipulation was ultimately unnecessary as our initial data checks showed that the frequency of loot box opening on the original version was sufficient. One participant was

excluded because they failed to experience any wins throughout the entire study and therefore did not have sufficient information to infer the reinforcement contingencies in effect.[5] Finally, one participant had incomplete data due to a saving error and was excluded, leaving the final sample with 129 participants.

Procedure

Participants played a modified version of Kao's (2019) *Infinite Loot Box*, a two-dimensional videogame programmed in *Unity* (Unity Technologies, 2022). The modifications made to *Infinite Loot Box* included programming new reinforcement schedules, incorporating new assets from the Unity Asset Store[6], and other minor aesthetic changes. Participants played the game on a lab computer in order to mitigate potential distractions, ensure that participants stayed on task, and control for performance issues (e.g., frame rate or operating system differences) across different machines.

Before starting the game, participants completed a written informed consent form and were given brief verbal instructions by a research assistant. Participants were told that they needed to earn a certain number of points in order to win the game and how to control the game.

---

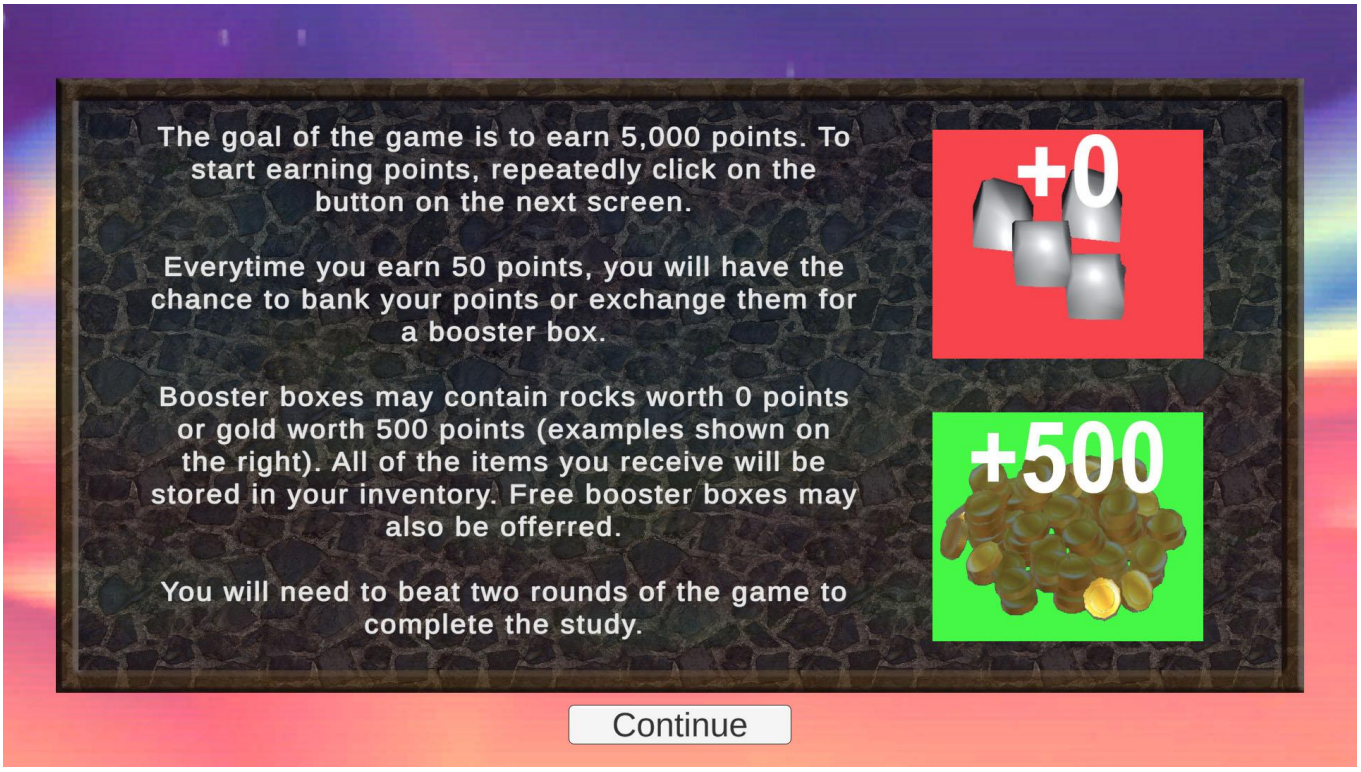
[5] Of the remaining 129 participants, 11 participants had their data partially excluded due to failing to experience a win in one of the two rounds. Data was retained from the round where they experienced at least one win.

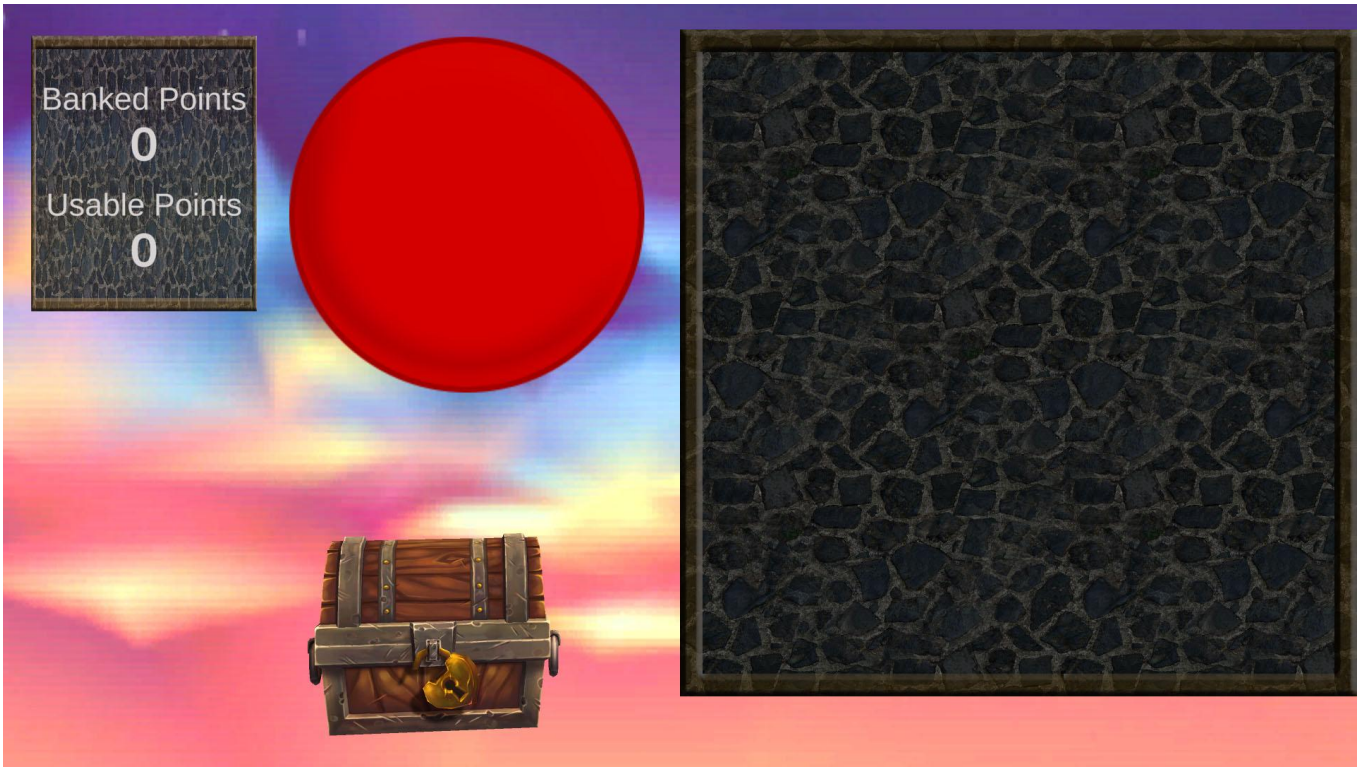
[6] Rocks were obtained from the Stylized Low Poly Rocks asset pack (https://assetstore.unity.com/packages/3d/environments/landscapes/stylized-low-poly-rocks-271334). Coins (including those used in Study 2) were obtained from the Gold Coins asset pack (https://assetstore.unity.com/packages/3d/props/gold-coins-1810). The aesthetic of the "Harvester" was obtained from the Sweet Land GUI asset pack (https://assetstore.unity.com/packages/2d/gui/sweet-land-gui-208285).

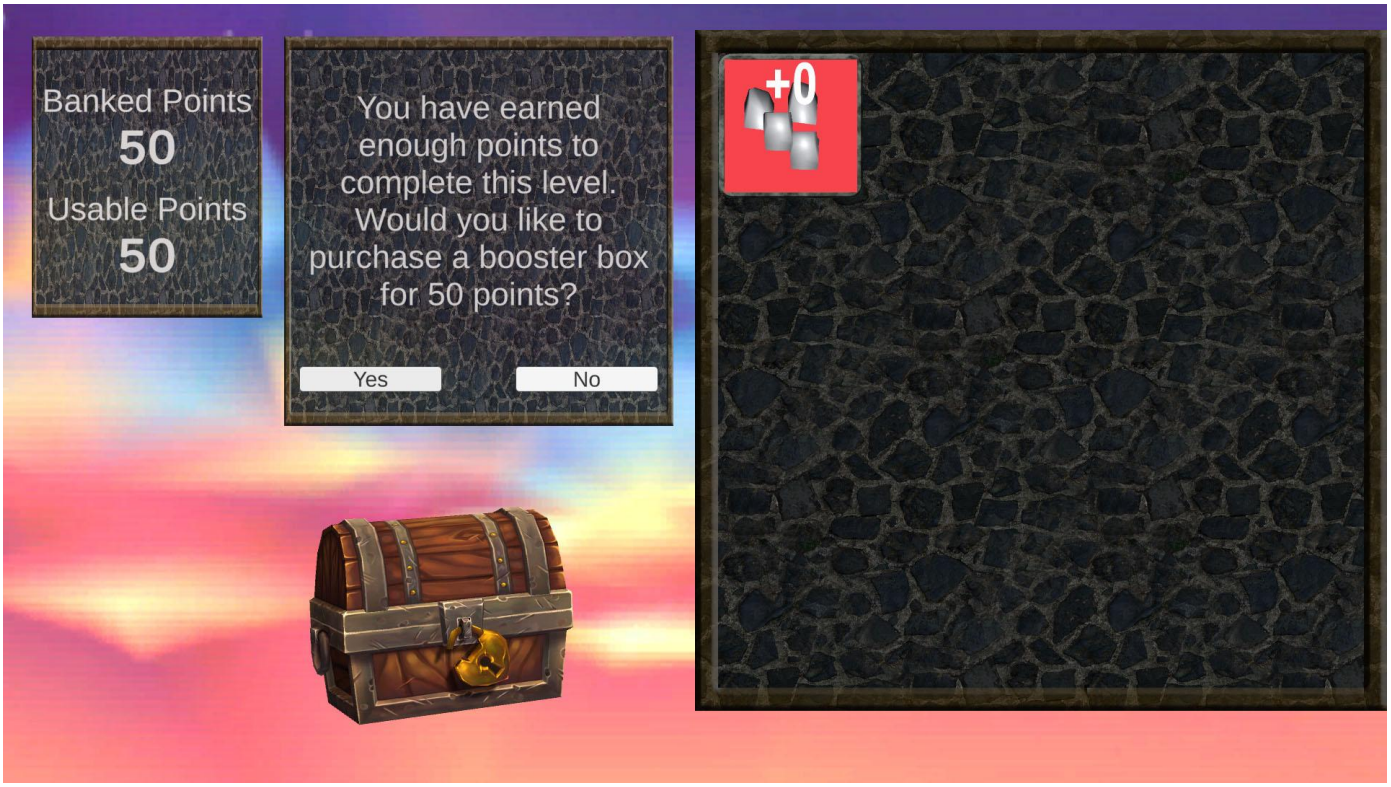
Participants were also told that they may be able to earn more points by purchasing loot boxes (referred to as 'boosters' in the game and instructions). After the instructions, participants had the opportunity to ask clarifying questions prior to starting the game. A short instructions screen before the game's main screen reiterated the goal of the game in written form.

The main screen of the game, shown in Figure 3 (nicknamed the "Harvest" screen) displayed a trial point tracker that showed how many points had been earned on the current trial (referred to as "Usable Points" in the game), and a total point tracker that showed how many points had been earned in the entire round (referred to as "Banked Points" in the game). Participants earned points by clicking on the red button in the middle of the screen at a rate of one point per click (FR 1). Points from clicking were added to the trial point tracker. Participants advanced to the next trial after they earned 50 points on the current trial (i.e., clicking the red button 50 times).

Figure 3: *Study 1 - Screenshots from Game*

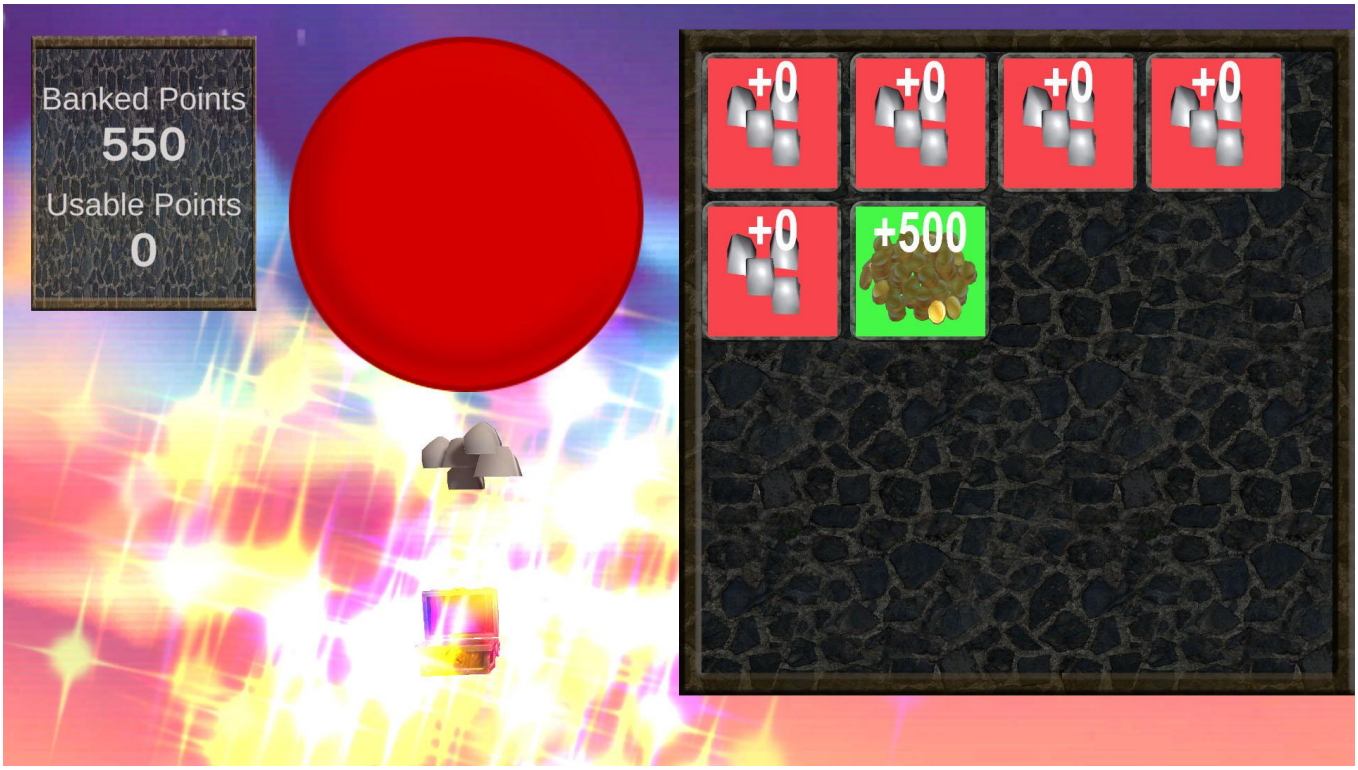





Banked Points
50
Usable Points
50
You have earned enough points to complete this level. Would you like to purchase a booster box for 50 points?
Yes
No
+0


Banked Points
0
Usable Points
50
You have earned enough points to complete this level and have earned a free booster box. Press the button below to claim your prize.
Claim Prize

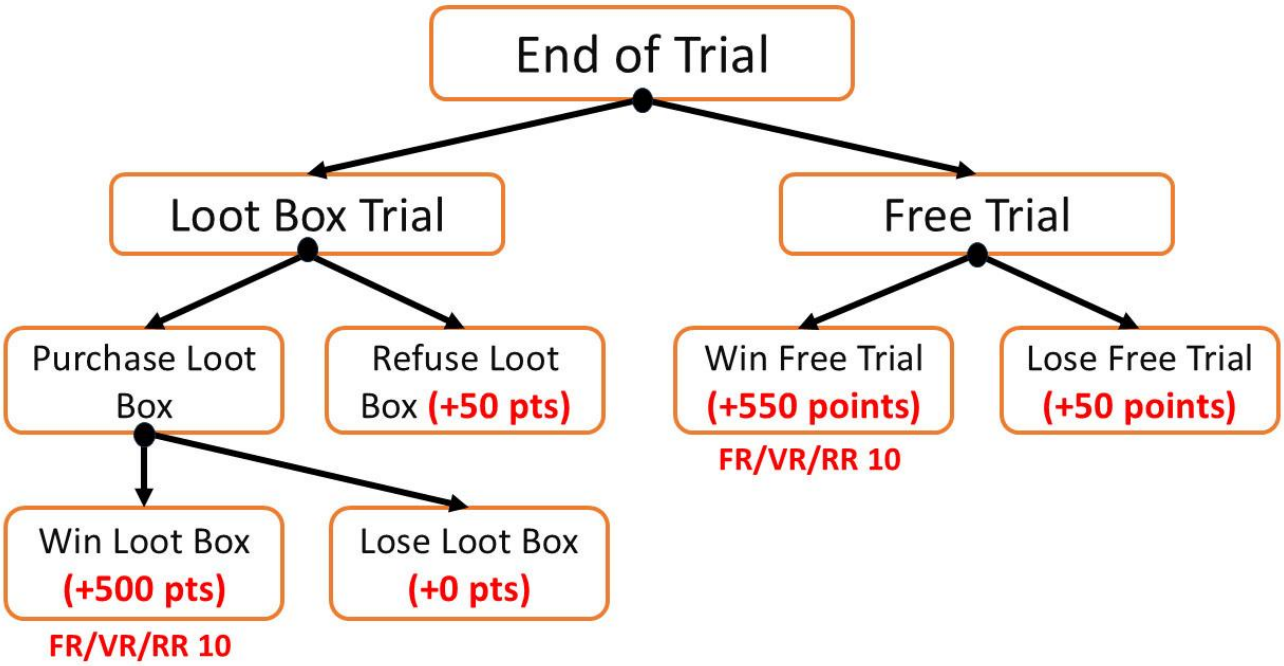

Banked Points
550
Usable Points
0
+0
+0
+0
+0
+0


Banked Points
550
Usable Points
0
+0
+0
+0
+0
+0
+500

At the end of each trial, participants were given the choice to purchase a booster with the 50 points they earned on the previous trial. Doing so would cause the chest below the red button to open, revealing its contents. Each purchased booster could reward a pile of gold worth 500 points or rocks worth zero points according to an FR 10, VR 10, or RR 10 schedule. These points were immediately added to the participant's total (i.e., Banked) points. All of a participant's obtained items were added to the inventory panel on the right side of the screen so that the participant would not need to rely on memory to infer the underlying reinforcement schedule. The goal of the game was to accrue 5,000 points. Importantly, only trial points could be used to purchase boosters. Once points had been added to the participant's total, they could no longer be exchanged for a booster. If participants chose not to open a loot box, 50 points were added to their total points while the chest at the bottom of the screen remained closed. Regardless of the participants' choice or the loot box outcome, the chest at the bottom of the screen disappeared and was replaced with another chest at the end of each trial.

In order to ensure that participants were sufficiently exposed to the schedule of reinforcement, the 1st trial and every 10th trial afterward (11th, 21st, 31st, etc.) gave the participant a free booster. This manipulation mirrors a technique used in real videogames (e.g., *Overwatch*) in which players are occasionally awarded free loot boxes (Ballou et al., 2020). Participants could not refuse this booster and were required to open it to proceed to the next trial. On free trials, 50 points were added to participants' total points in addition to the outcome of the booster. The possible choices participants could make at the end of each trial and the associated point outcomes for each are shown in Figure 4. Importantly, in order to ensure that participant choices were influenced only by the characteristics of each reinforcement schedule, all outcomes were designed to have equivalent expected values.

Figure 4: *Study 1 - Decision Tree*



In order to compare behaviors elicited by different schedules, participants played two rounds of the game with two different schedules of reinforcement. One round always used an RR 10. The schedule of reinforcement for the other round was determined by generating a random value (*S*) that determined the range of an FR/VR 10 schedule. For example, if zero was selected, the boosters in that round followed an FR 10 schedule. If nine was selected as the random number, the boosters in that round followed a VR 10 schedule with a 1-19 range (i.e., 10 – 9 and 10 + 9). In this way, behavior on RR schedules could be compared to behavior on FR/VR schedules with equivalent mean number of required responses that sampled across the continuum of schedule variability. The FR 10 condition was oversampled such that participants had a 20% probability of being assigned to it in order to increase power (see McClelland, 1997). This overweighting was to ensure that both ends of the continuum were sufficiently represented in the data. Whether participants encountered the RR or the FR/VR first was counterbalanced.

The study was designed to take 30-45 minutes to complete based upon click rates that we have observed in prior studies (250-300 clicks per minute, assuming a constant). Note that the main reinforcer, points, were used as markers of completion progress rather than any kind of currency. As such, we assumed that participants would be motivated to open loot boxes in order to complete the study as fast as possible (with the 500-point prize representing a two-minute time save). Participants were allowed to take a small break on an intermission screen between rounds if they desired and were given up to an hour to complete the study. The game automatically recorded the current schedule of reinforcement, the participant's choice at the end of each trial (including free booster trials), the current trial, the current round, the current response requirement (for FR/VR schedules), participant demographics, the time, location, and duration of each click, and the pre-ratio (or post-reinforcement) pause (Felton & Lyon, 1966; Griffiths & Thompson, 1973), the time between winning a loot box and continuing to click on the Harvester. The latter four variables were collected for exploratory analyses that will inform future studies.

Results

Manipulation Checks

Completion Time

Participants took between 27 and 57 minutes to complete the entire game, with an average completion time of about 38 minutes. In terms of trials, participants took between 42 and 127 trials to complete each round, with an average of about 92 trials.

Assignment of Conditions

Of the 129 participants, 21 (16%) played the FR 10 condition, just shy of the 20% oversampling that was implemented into the program. The remaining VR 10 conditions had sample sizes ranging from 7 (VR 10 7-13) to 19 (VR 10 8-12) participants. By round,

participants were roughly equally likely to be assigned to the counterbalanced schedule orders (i.e., an RR 10 schedule followed by one of the randomly assigned FR/VR 10 schedules or one of the randomly assigned FR/VR 10 schedules followed by an RR 10 schedule), with the exception of the VR 10 5-15 and VR 10 8-12 conditions which both appeared more frequently in the 2nd round than the 1st, with the VR 10 5-15 being assigned in the 2nd round 77% of the time and the VR 10 8-12 being assigned in the 2nd round 68% of the time. These fluctuations in schedule frequency are attributable to the random elements implemented into the game and the repeated measures nature of the data meant that there was sufficient data from each condition despite this.

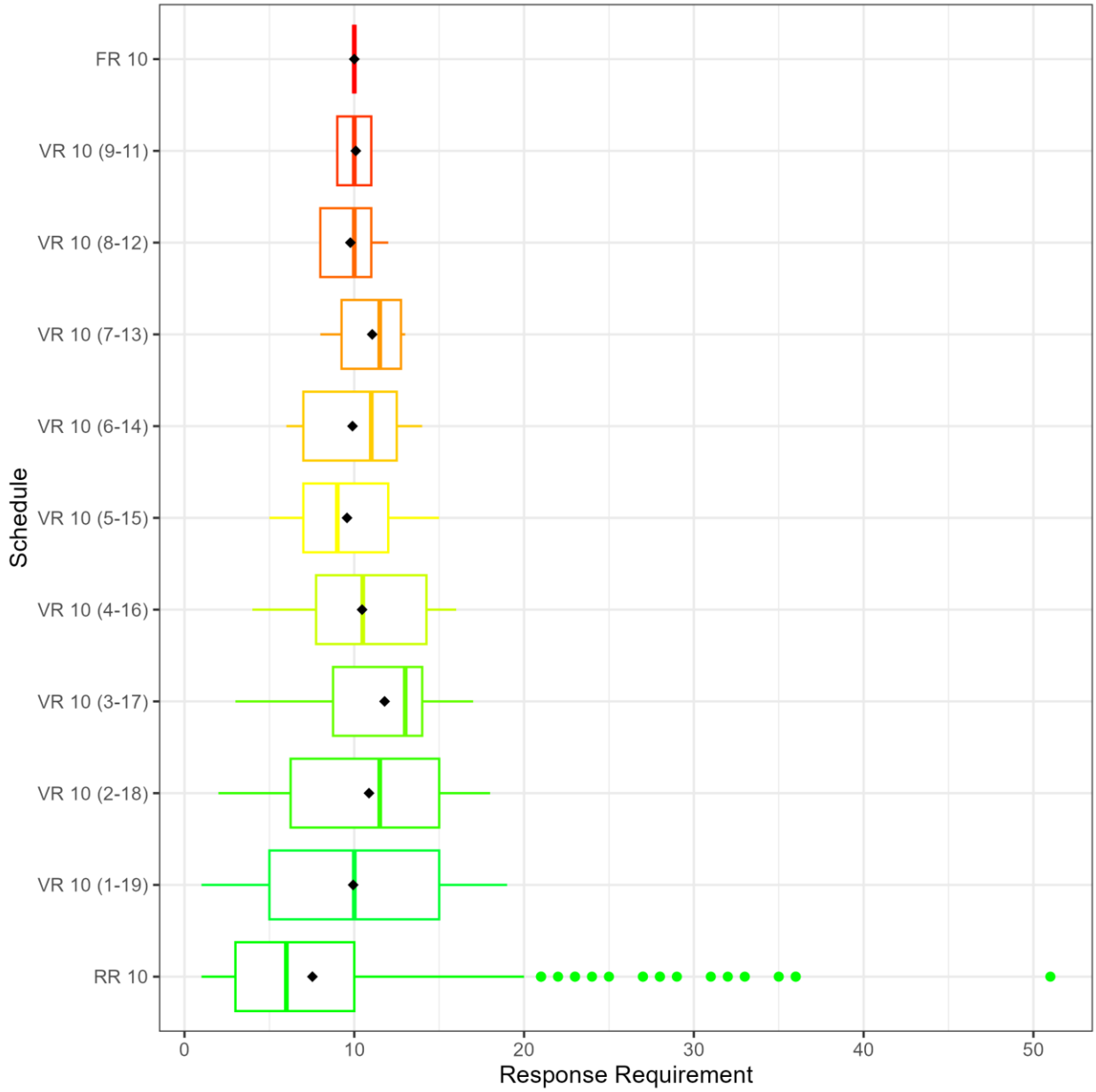
Amount and Timing of Reinforcement

There was a total of 722 wins (loot boxes that paid out 500 points) across all participants, with an average of six wins across the entire game (or three wins per round). Given the known effects of big wins (Kassinove & Schare, 2001) that may drive preferences for more variable schedules (e.g., RR 10), the data were also examined to determine when participants experienced their first win. First wins occurred anywhere from the very 1st to the 101st trial, with a mean of about 33 trials to the first win. This wide distribution was partly driven by the high variability of the RR 10 schedule, although it should be noted that repeatedly to open loot refusing boxes that were not free increased the number of trials it took to be reinforced for the first time.

The distribution of response requirements for each schedule are shown in **Figure 5**. Because RR schedules had no programmed response requirement, we calculated the number of loot boxes that were opened until reinforcement was received and used that as the response requirement, resetting the requirement after reinforcement was received (e.g., if a participant chose to open five consecutive loot boxes and was reinforced after the second and fifth, the

response requirements would be two and three, respectively). The mean response requirement for all schedules was at or near ten responses as programmed. The RR 10 condition had the lowest mean response requirement overall, but this lower mean may be due to censoring of true response requirements, as participants that experienced long streaks of loot boxes with no reinforcement may have chosen to stop opening loot boxes, thus obscuring what the true response requirement may have been.

Figure 5*: Study 1 - Boxplot of Response Requirements by Schedule*



Note. Black diamonds represent the mean response requirement.

Participant Choices

Based on the average number of trials and excluding free trials, participants made an average of 82 decisions in each round of the game. Participants chose to open loot boxes 23% of

the time (excluding free trials, in which participants were forced to open the loot boxes). In general, there seemed to be a decrease in the number of loot box openings from Round 1 to Round 2 (the effect of round is examined further in the main analyses). Eight participants never chose to open a loot box within one of the rounds and only experienced wins through free trials (with the exception of one participant that also had no wins in that round and was thus excluded from subsequent analyses). One participant chose to open a high proportion of loot boxes relative to the number of trials, choosing to open 88% of the time across 128 trials. Although these behaviors could be indicative of participants engaging in inattentive behavior or response sets (i.e., absent-mindedly and repeatedly clicking to open or not open regardless of the reinforcement contingencies), we opted to include them in the final analyses because they constituted a relatively small proportion of the total sample and may have been reflecting behavioral patterns that real videogame players have shown in regards to loot boxes. In particular, prior literature has distinguished between players that adamantly refuse to spend money on loot boxes (Larche et al., 2023) and 'whales', players that report extremely high loot box spending (Close et al., 2021).

End of Trial Choices

We conducted a series of logistic multilevel models to predict whether participants chose to open a loot box or not at the end of each trial (excluding free trials) using the trial number (within-subjects, continuous), the programmed MAD for each schedule (within-subjects, continuous), and their interaction as predictors. After data collection had completed, we identified the game round (within-subjects, categorical) as a potential unmodeled dependency in the data (Brauer & Curtin, 2018) that needed to be included in the models (including all possible

two- and three-way interactions with the trial number and programmed MAD). All continuous predictors were mean-centered and all categorical predictors were effect-coded.

Following the recommendations of Barr et al. (2013) and Scandola and Tidoni (2024), we first fit a model with a maximal random-effects structure (i.e., allowing all possible intercepts, slopes, and interactions to vary) using the lme4 package in R (Bates et al., 2015). Because trial number was on a much different scale than the other predictors, we divided trial numbers by a scalar constant (Trial/10) in order to prevent our models from encountering large eigenvalue errors. Unsurprisingly, the maximal model failed to converge using the conventional approach. Because simplifying a model's random effects structure ad hoc until finding a model that converges as suggested by Bates et al. (2018) has been shown to an increase the risk of Type 1 error (Scandola & Tidoni, 2024), we chose to use a fully Bayesian approach using the brms library for R (Bürkner, 2017).

We specified N(0,1), a normal distribution, as a weakly informed prior for all regression weights (*b*), except for the intercept. We specified the model to run four separate Markov Chain Monte Carlo (MCMC) chains for 6,000 iterations each, with an adaptive delta of .95, to ensure that model results were reliable and not influenced by starting values. All R-hats (indicators of whether the MCMC chains had converged) were below 1.05, within the cutoff for acceptable R-hat values. The model passed other diagnostics, such as the bulk and tail estimated sample sizes (ESS), although there were 27 divergent transitions after warm-up, indicating some instability in the model. Per STAN help threads (see https://mc-stan.org/learn-stan/diagnostics-warnings.html), a small number of divergent transitions may be acceptable if R-hat and ESS values do not indicate further issues. Even so, we interpreted our model with a degree of caution.

Table 1 shows the parameter estimates and 95% credible intervals for each, while the model predictions are shown in Figure 6 (see Appendix A - for plots of each posterior parameter distribution and associated MCMC chains). Figure 7 shows the model predictions for each participant.

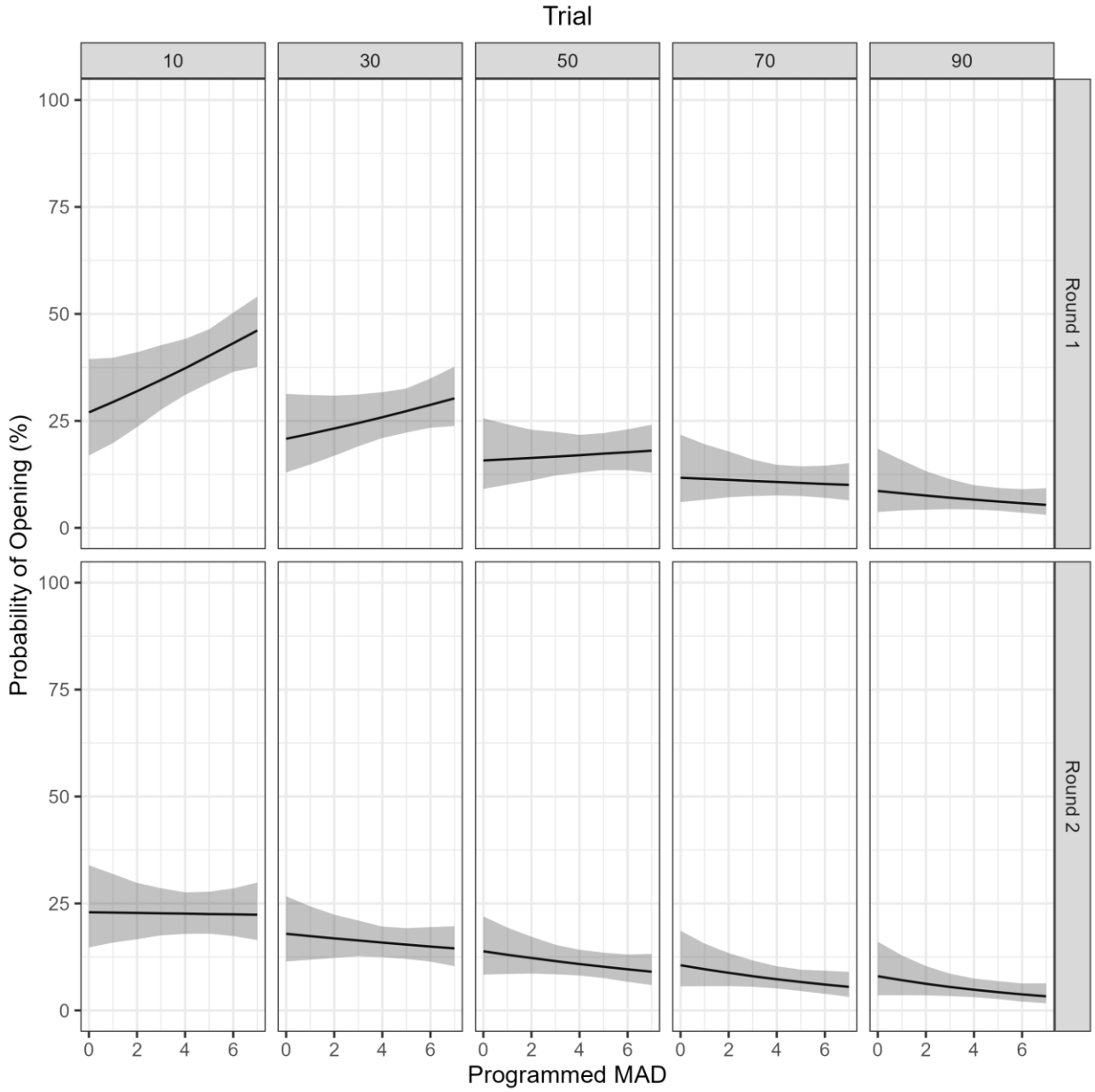
Overall, participants were strongly inclined to avoid opening loot boxes at the start of the study and became increasingly less likely to open loot boxes with each passing trial. The probability of opening loot boxes became even smaller in the 2nd round compared to the 1st, further demonstrating that as participants progressed through the entire study, the less and less likely they became to open loot boxes. While Figure 6 suggests that participants may have initially had a small preference for higher MADs, the effect was short-lived due to the negative interaction between MAD and trial that rendered the effect virtually non-existent by trial 50.

Table 1: *Study 1 - Parameter Estimates Modeling End of Trial Choices*

	Estimate (log-odds) 95% CI	Est. Error	Odds-Ratio
Intercept	**-1.79** **[-2.06, -1.53]**	0.14	0.14
Trial	**-0.26** **[-0.31, -0.22]**	0.02	0.44
MAD	-0.02 [-0.07, 0.04]	0.03	0.50
Round	**0.32** **[0.17, 0.46]**	0.08	0.58
Trial*MAD	**-0.02** **[-0.03, -0.01]**	0.01	0.50
Trial*Round	-0.03 [-0.06, 0.01]	0.02	0.49
MAD*Round	0.05 [-0.05, 0.14]	0.05	.51
Trial*MAD*Round	0.00 [-0.02, 0.01]	0.01	0.50

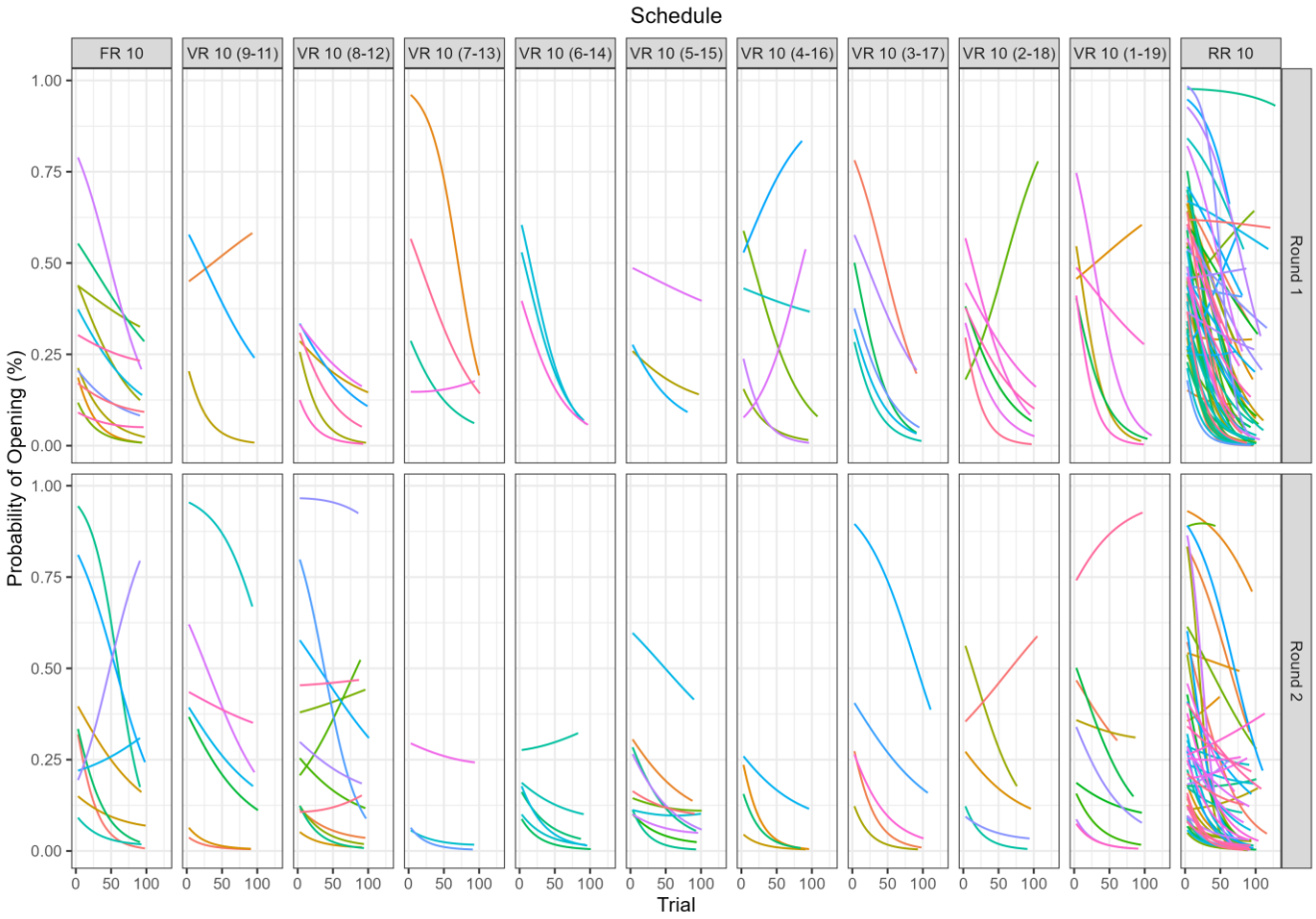
Note. Trial and MAD were mean-centered. Trial was divided by 10 in order to avoid scaling issues. Round was effect-coded with Round 1 as the reference group.

Figure 6: *Study 1 - Probability of Loot Box Opening by Programmed MAD, Trial, and Round*



Note. Error ribbons represent the 95% credible interval of the mean.

Figure 7: *Study 1 - Individual Model Fits*



Note. Lines represent model fits for each participant. Predictions were obtained by generating predictions from the model 1,000 times per data point and averaging estimates across all predictions.

Discussion

The goal of Study 1 was to examine differences in operant behavior elicited by FR, VR, and RR schedules of reinforcement with an equivalent mean number of required responses. Our competing hypotheses predicted that schedules on either of the extremes of the variability continuum could lead to more loot box purchases. Although the results failed to find unambiguous support for either of these competing hypotheses, there are several key takeaways from Study 1 with interesting implications for past literature that can help us to improve the paradigm for future studies.

The fact that participants only briefly showed a preference for opening loot boxes on higher MAD schedules is unexpected, but rational given that the EV for purchasing and not purchasing loot boxes were equivalent by design. Participants, especially those on less variable schedules, likely recognized that they would earn the same number of points in the long run regardless of their decisions. This interpretation is also consistent with the individual fits shown in Figure 7 where some participants, in contrast to the majority, were actually more likely to buy loot boxes over time. Whether participants defaulted to purchasing or not purchasing could be due to individual differences in risk aversion or preferences that are relevant to the public health aspects of loot boxes, but tangential to the current study. Despite the fact that it may have led to neither hypothesis being supported, it was necessary for the EVs to be equivalent at this stage in task development, as EV imbalances that were not carefully calibrated would have likely overpowered all other effects.

A potential confound recognized early in the design process was that of sampling error (Fox & Hadar, 2006), such that participants could potentially have made overgeneralizations about the reinforcement contingencies after opening a small number of loot boxes. Depending on

participant choices and the random mechanisms implemented in the program, the actual response requirements and MADs may have differed greatly from what had been programmed for each schedule, especially on highly variable schedules such as the RR 10 or the VR 10 1-19. To explore this issue, we calculated experienced MADs using Equation 3 below:

$$MAD_E = \frac{\Sigma|S-R|}{T} \quad (3):$$

Where S is the current experienced response requirement, R is the mean experienced response requirement for the entire round, and T is the number of times reinforcement was received[7].

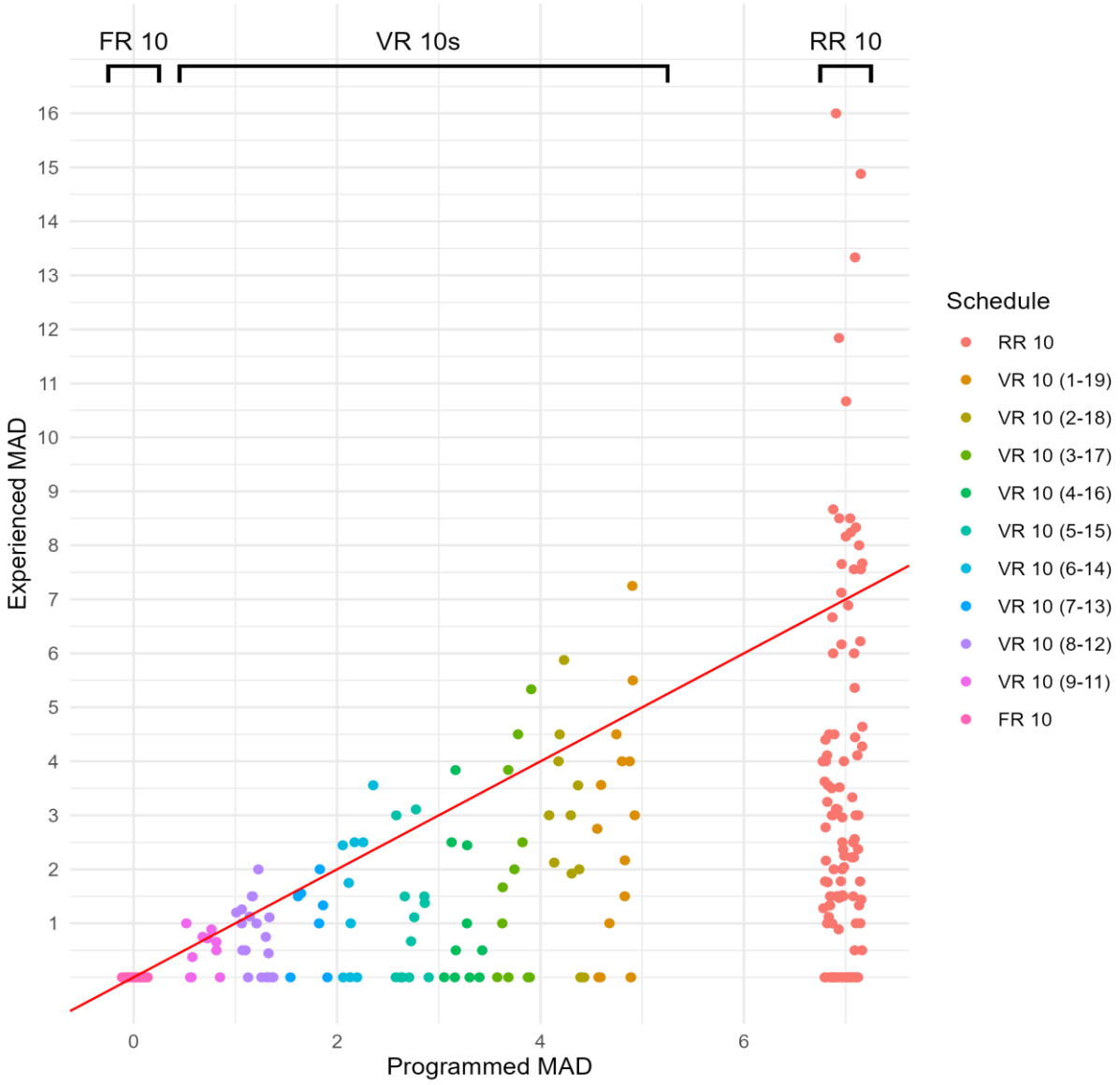
Experienced MADs ranged from zero to sixteen. Rounds with only a single win had an experienced MAD of zero by default (i.e., there can be no variation in response requirements when only a single response requirement was experienced). A scatterplot of the correlation between programmed and experienced MADs is shown in Figure 8. The correlation between the two MADs was moderately positive, $r(245) = .40$, $t(245) = 6.83$, $p < .001$. Unsurprisingly, RR 10 schedules showed the most deviance from the simulated programmed MAD of 6.67. When the RR 10 schedules were excluded, the correlation increased to $r(125) = .54$, $t(125) = 7.17$, $p < .001$. Due to the relatively small number of wins, it was not possible to recalculate our statistical models using experienced MAD instead of programmed MAD. We intend to examine experienced MAD further in future studies, but as far as Study 1 is concerned, experienced and programmed MAD appear to have performed similarly enough so as not to have major impacts on our results. While randomizing response requirements and allowing participants to dictate their own experiences (e.g., differential contact with the reinforcement contingencies depending

[7] Experienced MAD may also be calculated based on a running average response requirement, rather than calculating the average for the entire round. Given that the majority of participants only experienced a small number of wins, the latter was chosen.

on how often they chose to purchase loot boxes) may have hindered experimental control, the random elements of our design (and participants' responses to it) were reflective of loot boxes as they would appear in real video games, which is critical for creating a paradigm that will be generalizable outside of the laboratory.

Finally, the possibility of fatigue effects, as evidenced by participants opening less and less boxes as the study progressed, cannot be neglected. Unlike animal learning studies, laboratory studies with human participants cannot run for anywhere near as long. Furthermore, the task for Study 1 was repetitive, which likely led participants to become bored and disengage from trying to learn the reinforcement contingencies. Future studies will aim to either condense the task into a shorter timeframe or improve participants' subjective evaluation of engagement (i.e., whether they view the game as fun or boring). As such, Study 2 added greater variety to the paradigm by implementing further variation in outcomes, namely the incorporation of additional outcome magnitudes beyond the 500 or zero points used in Study 1. The addition of multiple outcomes also brings the paradigm closer to how loot boxes operate in real video games.

Figure 8: *Study 1 - Correlation Between Programmed and Experienced MAD*



.

Note. Diagonal red line represents theoretical perfect correlation between programmed and experienced MAD. Data points are jittered for visualization purposes.

Chapter 5 - Study 2: Reinforcer Magnitude

Loot boxes contain a wide variety of possible rewards ranging from costumes to in-game currency. As such, a paradigm that properly models loot boxes must incorporate a wider range of outcomes beyond winning and losing. Within the reinforcement literature, this is known as variable magnitude (VM) of reinforcement, although the term magnitude has been used to refer to the duration that reinforcement is available (Powell, 1969; Todorov, 1973; Todorov et al., 1984), the number of reinforcer presentations (Landon et al., 2003), or the amount of reinforcer units rewarded, such as the number of food pellets or the value of monetary rewards (Berardi et al., 2020; Lehr, 1970). Beyond these definitions, the magnitude of loot box rewards may also be related to their subjective value to the player or relative rarity of being received (Larche et al., 2021) or their expected utility (Kahneman & Tversky, 1979). The study of VM in reinforcement is further complicated due to reinforcement rate and magnitude not being independent influences on choice behavior in concurrent schedules (Elliffe et al., 2008) and the fact that VM is an intrinsic factor in all intermittent schedules of reinforcement (i.e., the absence and presence of reinforcement are themselves different magnitudes).

Broadly, the effects of reinforcer magnitude are subtler compared to reinforcement rate. In studies that use a modified version of the generalized matching law formula, rate has consistently been found to overshadow magnitude (Davison & Baum, 2003; Landon et al., 2003). High magnitude reinforcers can lead to sustained shifts in preference (preference pulses) toward that schedule (Landon et al., 2003), but VM can moderate the effect depending on the difference between magnitudes (Davison & Baum, 2003). Lehr (1970) trained rats to correctly navigate a T maze and reinforced correct responses with food pellets. A CRF/VM group received five pellets for half of their correct responses and one pellet for all other correct responses while

a VR/VM group received six pellets for half of their correct responses and no pellets for the remainder. The last two groups received either six or three pellets for every correct response (both CRFs with fixed magnitudes). Consistent with prior literature, the VR/VM schedule led to the greatest resistance to extinction (persistence of operant responses when reinforcement is no longer available), although the CRF/VM group demonstrated comparatively similar acquisition and extinction rates.

Berardi et al. (2020) used CRF and VR schedules in tandem with VM to incentivize human participants to engage in daily physical activity. Participants wore an accelerometer that awarded them with small amounts of money if they met daily physical activity goals (e.g., exercising for a specified amount of time). If participants completed 24 daily goals, the schedule of reinforcement would change, with six different schedules being used throughout the study. Initially, participants were reinforced on a CRF schedule with fixed reinforcement magnitude (i.e., the same number of points was awarded for each goal). This was followed by a CRF schedule with varying magnitudes ($0.25 to $2.50), then four VR/VM schedules with different average reinforcement rates and ranges (e.g., a VR 1.09 with magnitudes from $0.50 to $2.50). The CRF/VM schedule resulted in the greatest increase in activity above baseline. However, these results may have been due to participant dropout, with less data being available at later stages due to participants failing to progress past the first schedule.

Issues of VM have also been a major influence in studies of gambling. The gambling literature has identified multiple phenomena related to VM that lead to persistent operant responding, including big wins and near-misses. Big wins are defined as a high magnitude reinforcement that occurs early in a gambling session or career (Kassinove & Schare, 2001). Highly variable schedules, such as an RR, are likely to result in ‘early’ wins (Haw, 2008) and

make continued gambling more likely. Allowing the magnitude of the win to vary alongside schedule type better reflects the mechanics of loot boxes. Near-misses are losses that are perceived as psychologically close to wins (e.g., matching two of three symbols on a slot machine, a roulette wheel landing on a number in physical proximity to the player's chosen number; Dixon & Schreiber, 2004). Prior loot box studies have posited that receiving a reward that is similar in rarity or subjective value to a desired reward may be interpreted as a near miss (Larche et al., 2021), spurring further loot box opening. These results further extend the idea of sampling bias leading to overgeneralizations in experienced-based decision making (Fox & Hadar, 2006). In the broader domain of cognitive psychology, outcome magnitude is typically built into the framework of cognitive decision-making, with magnitude and probability jointly determining a prospect's EV (and therefore, the prospect's subjective value as well; Harless & Camerer, 1994; Kahneman & Tversky, 1979). Interestingly, a study by Konstantinidis et al. (2018) found that increasing the magnitude of outcomes can mitigate the effects of the description-experience gap.

Because VM is merely another form of schedule variability, our hypotheses for Study 2, Hypotheses H3 and H4, followed the same structure as H1 and H2. H3 predicted that the number of loot boxes opened will increase as the order of reward magnitudes becomes more variable, with participants in the large VR and RR conditions opening the most loot boxes. Conversely, Hypothesis H4 predicted that the number of loot box opens will increase as the order of reward magnitudes becomes more predictable, such that participants in the FR and small VR conditions will open the most loot boxes. As in Study 1, additional data was recorded for future exploratory analyses.

Method

Participants

A total of 116 (66% female, M_{Age} = 19 years old) participants were recruited from Kansas State University's SONA subject pool and received course credit for their participation. Based on the results from Study 1, we deemed a sample of 100 participants to still be an appropriate benchmark. Data collection continued until the end of the week in which the 100-participant goal was reached. No participants were excluded from Study 2 according to the criteria established in Study 1.

Procedure

The game used in Study 2 was a modification of the program used in Study 1. Rather than outcomes consisting of only a chance to receive 500 points or 0 points, Study 2 had four possible point rewards each time a loot box was opened: 0, 50 (depicted as a small pile of bronze coins), 100 (depicted as a large pile of silver coins), or 500 points. In order to keep expected value equivalent, we first programmed the game to generate magnitude arrays consisting of ten magnitudes; four 0s, three 50s, two 100s, and one 500 (for an overall average of 85 points). The order of magnitudes within each array dictated the corresponding magnitude of loot box rewards each time the player chose to open a loot box and the array was regenerated when all of the magnitudes within an array was exhausted. Magnitude order was shuffled to varying degrees according to four conditions: fixed-order (FO), 'low' variable order (LR), 'high' variable order (HO), and random order (RO). Manipulating the order of magnitudes allowed for us to conceptually mimic the variability of different reinforcement schedules while keeping the actual frequency of each magnitude, and therefore the EV, constant.

In the FO condition, participants experienced the outcomes exactly as shown in Figure 9 throughout the entire round. In the three other remaining conditions, the order of outcomes was shuffled using a modified version of the code posted by StackExchange user pjs (2020 - https://stackoverflow.com/a/62438368) adapted to *Unity*. Full details of the shuffling algorithm will be made available on OSF, but in short the algorithm swaps the positions of each element in the original array according to a generated Gaussian distribution and an orderliness parameter that can be modified to control the degree of randomness. Higher values of orderliness result in less shuffling of elements while lower values of orderliness result in greater shuffling of elements in the array. As such, in the LO condition, outcomes were shuffled using an orderliness value of 0.75, while in the HO condition, outcomes were shuffled using an orderliness value of 0.25. The RO condition used an orderliness value of 0.01 in order to simulate true randomness as best as possible. For all conditions except for the fixed order, the order of magnitudes was reshuffled after every ten-loot box opens. Example orderings for each condition are shown in Figure 9.

We also increased the value of declining to open loot boxes to 85 points (although the number of clicks needed to complete each trial remained at 50; this was described to participants as a small point bonus for choosing to bank their points) to match the increased expected value of the loot boxes, as well as increasing the point goal for each round to 8,500 points. The aesthetics and instructional text were also updated to accommodate the new design, namely the new models and thumbnails associated with the new magnitudes. These changes can be seen in the screenshots provided in **Figure 10**. As in Study 1, participants played two rounds of the game with each using a different condition. Participants experienced two of the four conditions, with the order in which conditions were experienced fully counterbalanced. All other aspects of the

game were identical to Study 1, including the randomized elements that allow participant experiences to vary.

Figure 9: *Study 2 - Example Magnitude Orders by Condition*

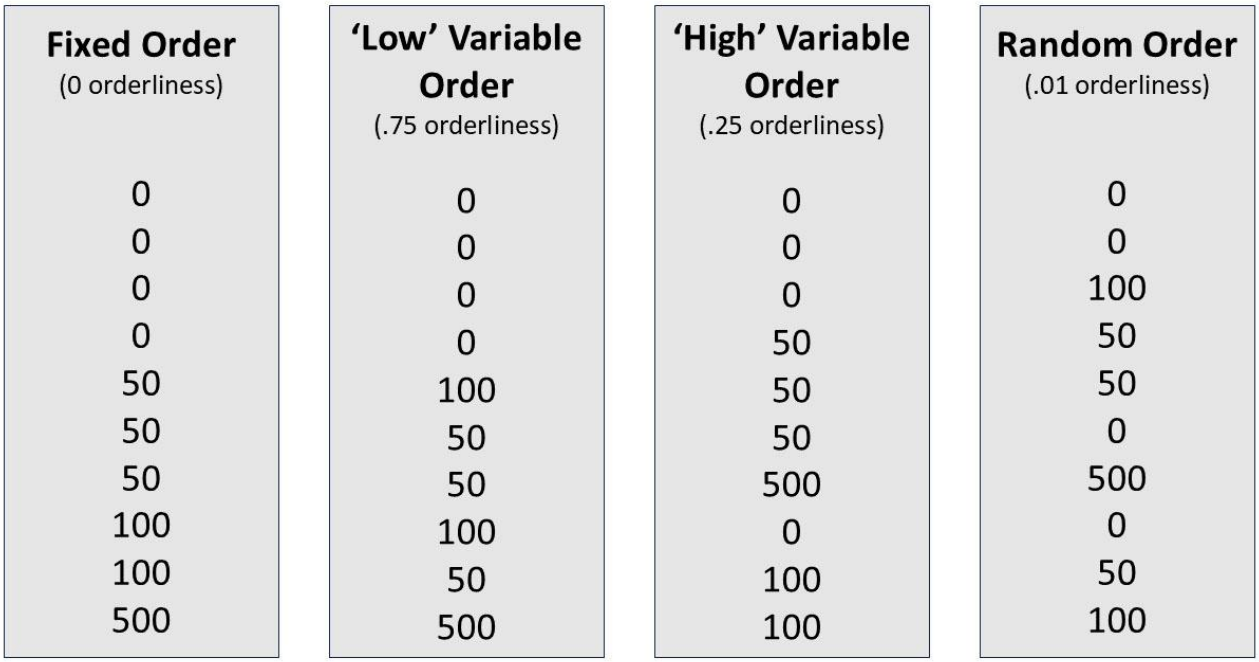
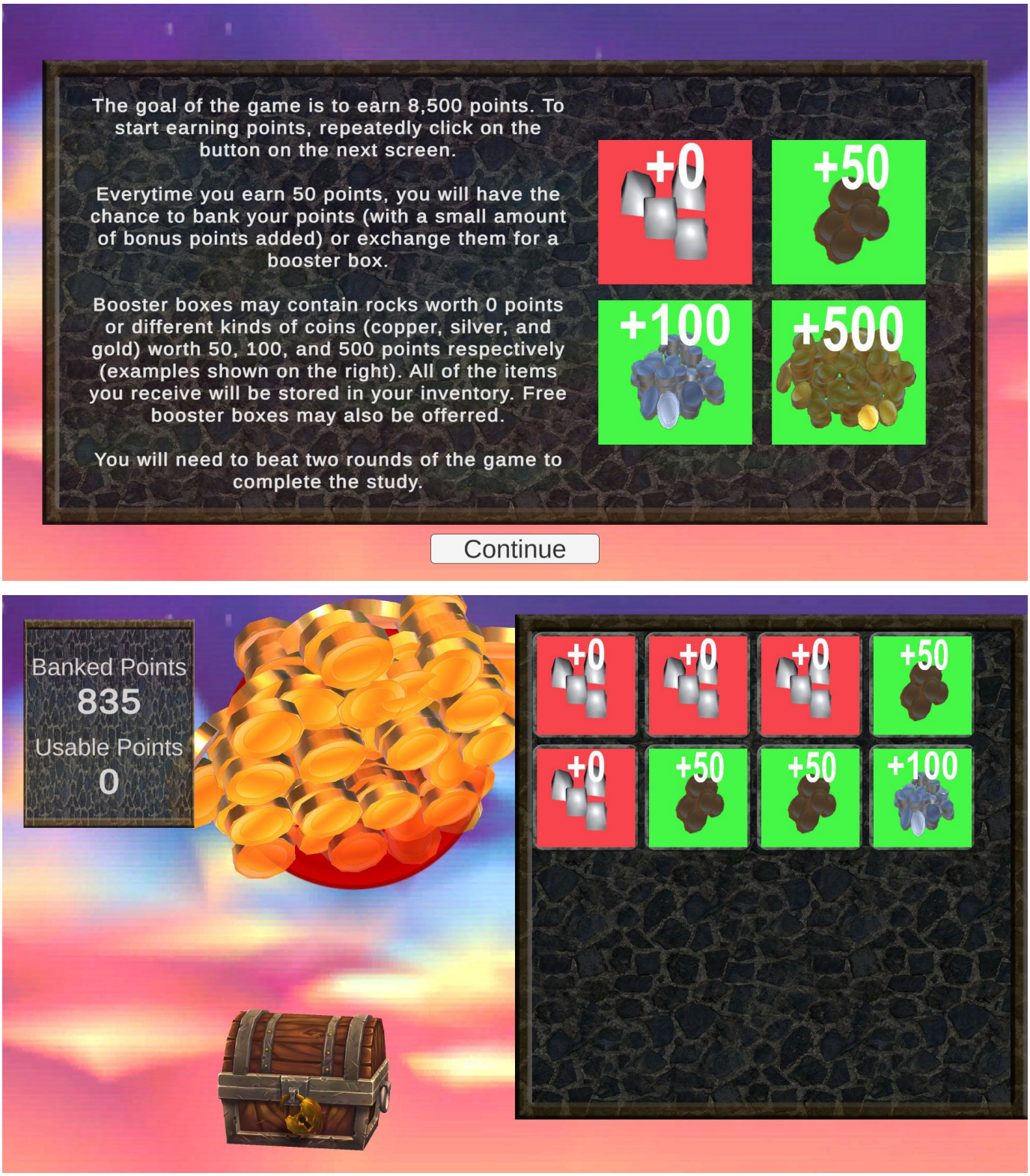


Figure 10: *Study 2 - Screenshots from Game*



Results

Manipulation Checks

Completion Time

Participants took between 28 and 48 minutes to complete the study, with an average of 37 minutes. In terms of trials, participants took between 86 and 96 trials to complete the study, with an average of 92 trials. Although these averages are similar to Study 1, the range of values was much narrower in both cases. This was likely due to the more consistent loot box payouts implemented in Study 2. Regardless of the order of magnitudes or how much they varied, participants were guaranteed to see all ten outcomes with every ten loot boxes that they opened.

Assignment of Conditions

Assignment to each of the four conditions was roughly equivalent. The most frequent condition was the random order condition with 27% of participants, and the least frequent was the high variable order condition with 22%. When factoring in the round, all four conditions appeared in both rounds 1 and 2 at virtually equal frequencies.

As an additional manipulation check, the actual orders of magnitudes that participants experienced was probed further. First, we examined the number of different orders that had been generated according to the shuffling algorithm by condition. As intended, only one possible magnitude order was generated in the fixed order condition. For the other three conditions, 71 unique orders were generated in the low variable order condition, 271 were generated in the high variable order condition, and 376 were generated in the random order condition, demonstrating that orders were indeed more variable in some conditions over others. Next, we verified that the magnitudes experienced each time a loot box was opened matched the contents within the array

(excluding magnitudes that were generated, but never actually used due to participants completing the round before opening enough loot boxes to experience them).

Amount and Timing of Reinforcement

Unsurprisingly, the design of Study 2 drastically increased the number of wins, with a total of 7,746 wins. The number of wins participants experienced per round ranged from six[8] to 54 wins, with an average of about 33 wins. The trials in which participants experienced their first win was correspondingly earlier, ranging from the very 1st trial to the 42nd trial, with a mean of about eight trials. By condition, first wins were latest in the fixed order condition with an average of 12 trials, followed by the low variable order condition which had an average of 11 trials, the high variable order condition which had an average of 7 trials, and finally, wins were earliest in the random order condition with an average of 3 trials. These results indicate the generation and shuffling of magnitude orders functioned properly and thus kept EV constant across conditions.

Participant Choices

As in Study 1, participants made an average of 82 decisions per round. Participants opened loot boxes more frequently in Study 2 than in Study 1. Overall, participants chose to open a loot box on 57% of the 19,083 choice trials. The mean number of opens was relatively similar across conditions, with participants opening an average of 49 boxes in the fixed order condition, 47 in the low variable order condition, 45 in the high variable order condition, and 46 in the random order condition.

---

[8] Six is the minimum number of wins required to complete the round. On trial 91, a player that refused all loot boxes would still have experienced all ten outcomes (six of which will be wins) through free trials.

Main Analyses – End of Trial Choices

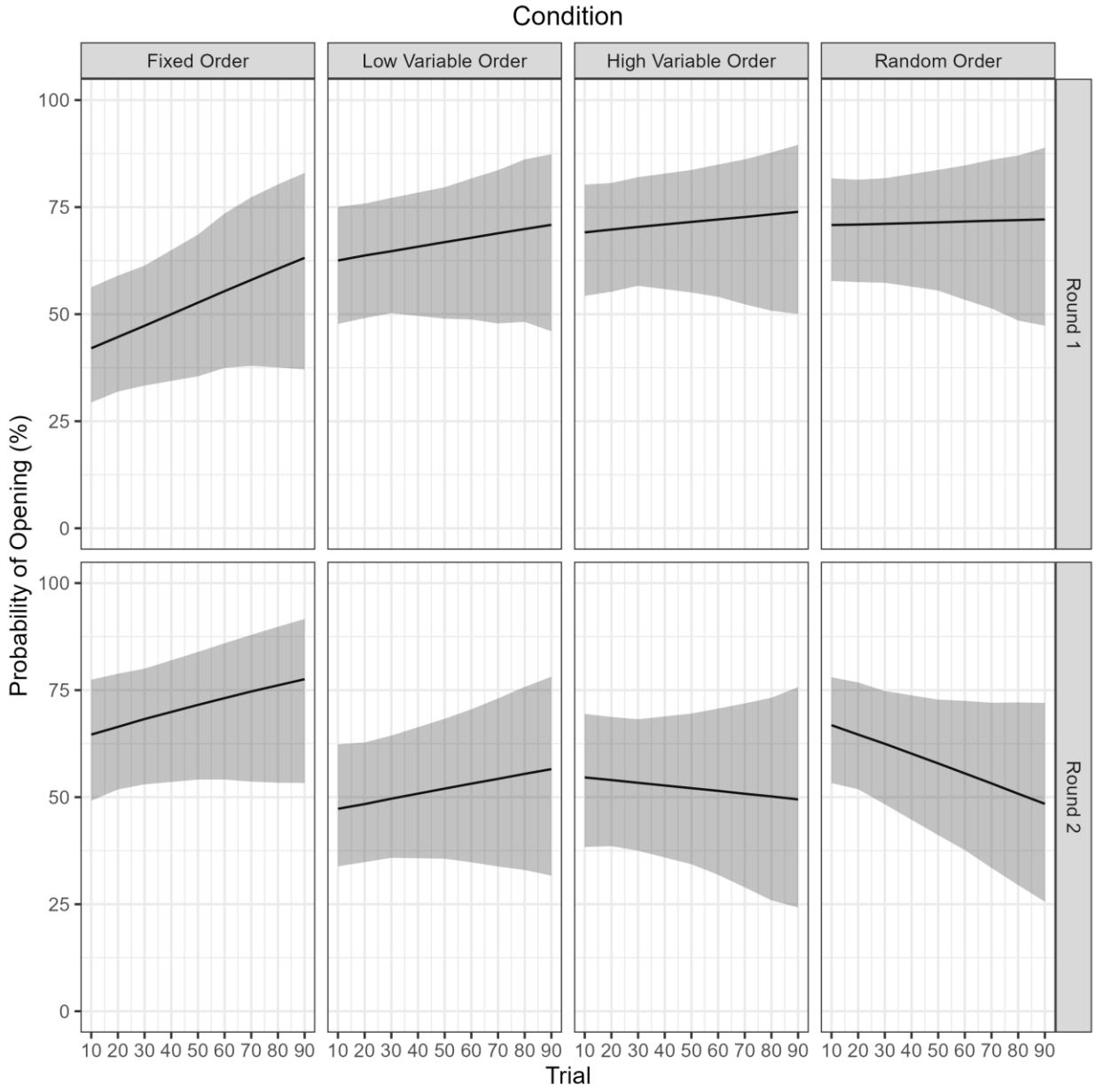
Study 2 used the same analytical strategies as Study 1. Participant choices at the end of each trial were modeled using a repeated measures logistic Bayesian regression using trial, condition (within-subjects, categorical with four levels: fixed order, low variable order, high variable order, and random order), round (categorical, within-subjects), and all possible two- and three-way interactions as predictors. Round and condition were effect coded (with round 2 coded as the reference group for round and the fixed order condition coded as the reference group for condition) and trial was mean centered. Once again following the recommendations of Barr et al. (2013) and Scandola and Tidoni (2024), the maximal random effects structure, including complex random intercepts of round and condition, were used. Conventional models were attempted, but once again failed to converge given their complexity. Instead, we used a Bayesian model with the same weakly informed priors for all regression weights (N(0, 1)). The model ran for 6,000 iterations and resulted in acceptable R-hats and MCMC chains. Parameter estimates are shown in Table 2 while posterior parameter distributions are shown in Appendix B - The model encountered an error indicating that the tail and bulk ESS were too small, indicating an issue with the sampling process that may influence the posterior distribution means. This error and the potential causes of it are addressed further in the Discussion. Model results are visualized in Figure 11. There were no effects of trial, round, condition, or their interactions, although visually there appears to be an increase in loot box purchasing later on in the fixed order condition, possibly suggesting in that participants had learned precisely when the 500 was due to appear based on prior outcomes and aligned their free trials accordingly to optimize point gains. Figure 12 shows the posterior mean distributions for each condition, while Figure 13 shows individual model fits for each participant.

Table 2: *Study 2 - Parameter Estimates Modeling End of Trial Choices*

	Estimate (log-odds) 95% CI	Est. Error	Odds-Ratio
Intercept	**0.50** **[0.11, 0.88]**	0.20	1.65
Trial	0.25 [-0.28, 0.77]	0.27	1.28
Random Order (RO)	0.13 [-0.20, 0.46]	0.17	1.14
High Variable Order (HO)	0.00 [-0.36, 0.36]	0.18	1.00
Low Variable Order (LO)	-0.12 [-0.45, 0.21]	0.17	0.89
Round	0.15 [-0.03, 0.33]	0.09	1.16
Trial*Round	0.23 [0.20, 0.67]	0.22	1.26
Trial * RO	-0.68 [-1.41, 0.05]	0.37	0.51
Trial * HO	-0.22 [-0.97, 0.52]	0.38	0.80
Trial * LO	0.22 [-0.50, 0.93]	0.36	1.25
Round * RO	0.14 [-0.18, 0.72]	0.21	1.15
Round * HO	0.15 [-0.30, 0.56]	0.22	1.16
Round * LO	0.26 [-0.17, 0.69]	0.22	1.30
Trial * Round * RO	0.28 [-0.47, 1.03]	0.38	1.32
Trial * Round * HO	0.04 [-0.75, 0.83]	0.40	1.04
Trial * Round * LO	-0.23 [-0.98, 0.50]	0.38	0.79

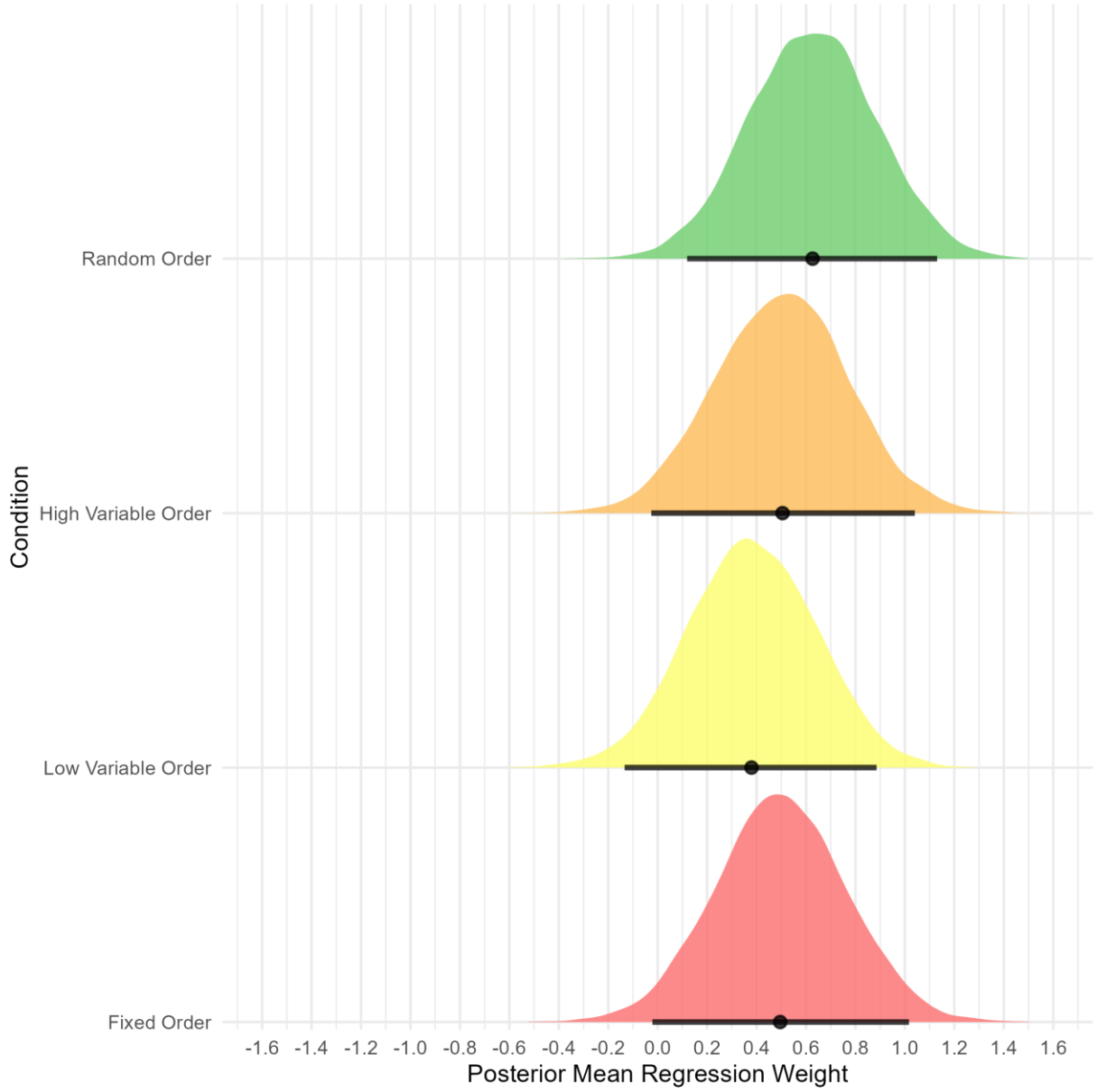
Note. Trial was mean-centered and divided by 100 to avoid scaling issues. Round was effect-coded with Round 1 as the reference group. Condition was effect coded with the fixed order condition as the reference group.

Figure 11: *Study 2 - End of Trial Choices by Condition, Round, and Trial*



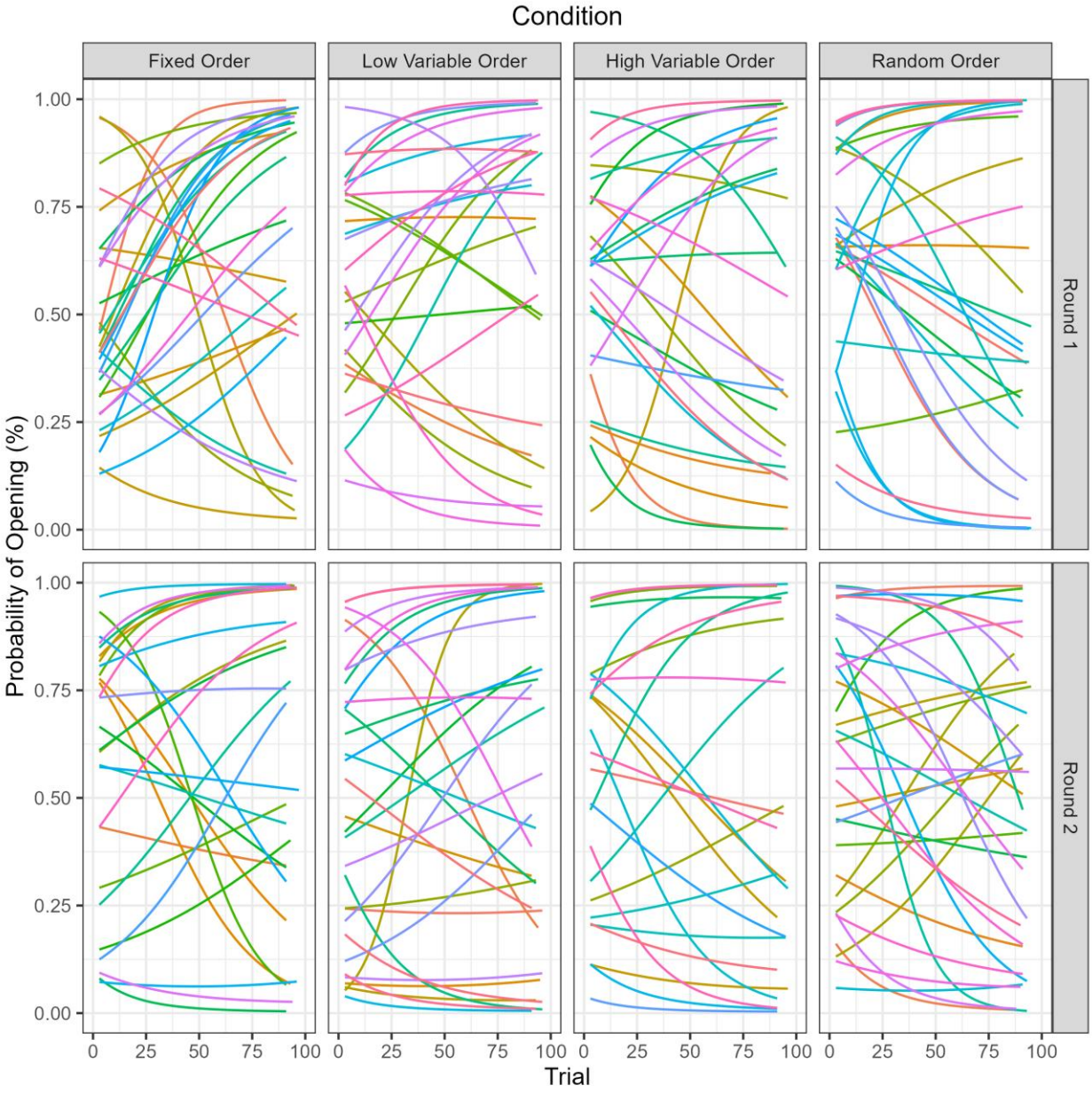
Note. Error ribbons represent 95% credible interval.

Figure 12: *Study 2 - Posterior Mean Distributions by Condition*



Note. Distributions were generated from 12,000 draws of the model fits. Black dots and bars represent the overall posterior mean regression weight of the distribution and the 95% credible interval.

Figure 13: *Study 2 - Individual Model Fits*



Note. Lines represent model fits for each participant. Predictions were obtained by generating predictions from the model 1,000 times per data point and averaging estimates across all predictions.

Discussion

The goal of Study 2 was to further replicate loot boxes by incorporating VM, thus manipulating another dimension of schedule variability. In order to keep the EV of purchasing and not purchasing loot boxes equivalent as in Study 1, only the order of outcomes varied across conditions. We once again predicted that participants would purchase the most loot boxes either when the order was highly varied (as in the RO condition) or did not vary at all (as in the FO condition). Study 2 failed to find support for either hypothesis, once again suggesting that further improvements to the paradigm are needed.

One interesting trend, however, is the increased probability of loot box openings seen exclusively in the FO condition, which may be evidence of exploitation. According to the exploitation-exploration tradeoff, exploratory behaviors are those that are used to test different relationships and contingencies in the environment (in this case, the schedule of reinforcement) while exploitative behaviors capitalize on the knowledge gained through exploration (Mehlhorn et al., 2015). As it applies to Study 2, the FO condition was the most predictable, a fact that participants likely realized and used to their advantage by using the free trials to progress the order and only opening loot. However, our model found no evidence that the slope of trial in the FO condition was different from any of the other conditions and the effect will need to be replicated in a future study.

As shown in Figure 1, one of the distinguishing features of an RR schedule is the lack of a defined range. While we manipulated the order of outcomes to be as random as possible, our manipulation also unintentionally imposed a defined range in all four conditions. It is likely that participants learned when the 500-point prize was due to appear within each set of outcomes, even in the more variable HO and RO conditions, leading to the null effects of trial and condition

reported here (and consequently, our model's inability to estimate the parameters reliably). A true random order that preserves the unique properties of an RR schedule would require that each outcome have a constant probability of occurring on each response (e.g., a 50% chance of 0 points, 30% for 50 points, a 15% chance for 100 points, and a 5% chance for 500 points). In truth, this is likely how real loot boxes are programmed in videogames, but it would have been impossible to keep the EVs equivalent if we chose this method.

Fortunately, the increased number of winning loot boxes and participants choosing to open loot boxes suggests that Study 2's design improved the subjective experience of the task. We attribute this improvement to the greater amount and variety of rewards, although participants still showed a tendency to purchase less loot boxes as the study progressed.

Chapter 6 - General Discussion

Over the past decade, loot boxes have become a source of heated controversy. Some games (e.g., *Mario Kart Tour, Rocket League*) have since chosen to remove their loot box mechanics entirely, but loot boxes remain a core mechanic of other popular titles such as *Counterstrike: Global Offensive* and *Hearthstone*, in spite of loot box regulations that have already been passed in other countries such as Belgium and China (Xiao, 2021; Xiao & Newall, 2022). Although the random reward mechanisms behind loot boxes are not new, loot boxes as they appear in modern games are remarkably effective at generating revenue (e.g., Niemeyer, 2024), which has raised serious public health concerns within the gambling literature. The study of loot boxes so far has attempted to establish links between gambling and loot boxes. However, the current loot box literature has over relied on cross-sectional surveys using inconsistent and insufficiently validated measurements that have introduced threats to validity and reliability. An experimental paradigm for loot boxes addresses these limitations and provides further validation for both methods of studying loot box. Multiple articles in the field have recognized the contributions that an experimental loot box paradigm could make (Ballou, 2023; Brooks & Clark, 2023; Zendle, Cairns, et al., 2020; Zendle & Cairns, 2018), but Kao's (2019) *Infinite Loot Box* was the only such paradigm until now.

By modifying Kao's (2019) *Infinite Loot Box,* we aimed to use loot boxes as a vehicle to study human operant learning as a function of variability in both ratio requirements (Study 1) and reinforcer magnitude (Study 2). While animal learning experiments (e.g., Crossman et al., 1987; Field et al., 1996; Mazur, 1983) have found evidence that highly variable schedules are more appealing, classic studies in cognitive decision-making from description have demonstrated an opposing preference for certainty and predictability (e.g., Kahneman & Tversky, 1979;

Rakow & Newell, 2010). We predicted that participants in both studies would choose to open loot boxes more at the two extremes of variability, meaning that participants would prefer very predictable or very unpredictable schedules. Unfortunately, we found no support for any of our hypotheses in either study, although this is perhaps unsurprising given that previous studies also found no discernible differences between VR and RR with equivalent average response requirements (Haw, 2008; Hurlburt et al., 1980). Meanwhile, studies of VM have reported that differences in reinforcer magnitude are less influential than the timing and rate of reinforcement (Davison & Baum, 2003; Lehr, 1970), with large magnitude reinforcers having fleeting impacts on operant behavior (Landon et al., 2003).

As this paradigm was ambitiously novel, it is no surprise that there were several overarching limitations to our studies. The first limitation was that participants may have had insufficient training to learn the reinforcement schedules. In Study 1 in particular, wins were rare regardless of schedule (on average, three per round), which may have been too rare for participants to discern the range of response requirements, even if they accurately tracked the response requirement for each win. The short duration of studies of human operant conditioning relative to animal learning studies is often pointed to as the potential cause of conflicting results between human and animal studies (e.g., Baron et al., 1991). The schedules of reinforcement used in the study may have not been properly calibrated to the short timeframe that we have to condition human participants (in our case, less than hour), as opposed to the days or weeks of training typically used in animal studies. Of all the limitations, this may be the easiest to address by merely increasing the odds of winning or making trials shorter (i.e., requiring less than 50 clicks) to encourage participants to open more loot boxes.

The second limitation to address was the high amount of physical effort involved in the focal task that participants had to perform in order to earn points. While designing the task, we chose to have participants click on a red button 50 times per trial because it was simple enough, that all participants would be able to complete the task regardless of prior videogame experience or skills, but also effortful enough that participants would deliberate carefully about what to do with their hard-earned points. However, the total amount physical effort involved in completing the task, amounting to thousands of mouse clicks per round, appeared to be too extreme based on participant reactions during and after the study. This may have led participants to feel bored, fatigued, or frustrated and disengage with the task, thus offering another explanation for participants' tendency to open less loot boxes the further they progressed in the game. While adjusting the probability of winning or shortening the number of clicks in each trial could also help with this issue, it may be even better to explore less strenuous tasks that could replace the current one. For instance, an autorunner, such as the Google Dinosaur game (which can be played at https://trex-runner.com/), that tasks players with jumping over obstacles in order to keep earning points, would require less physical effort of participants, as performance would be determined by the timing of their responses rather than the quantity of responses (in other words, an interval schedule instead of a ratio schedule).

The biggest limitation, however, was that we developed the paradigm with two goals in mind, but did not recognize where those goals conflicted in our design decisions. The first goal was for the paradigm to be as close to a simulation of loot boxes as possible, which could make a major contribution to the current loot box literature and inform future loot box regulations. Meanwhile, the second goal was for the paradigm to be a rich and engaging method of studying operant learning with human participants that could tap into the success of videogames to

overcome the limitations of human operant studies (i.e., boring, repetitive tasks, make the task more engaging to compensate for limited training time). Although we do not believe that these two goals are necessarily mutually exclusive, it is clear that some of our design decisions were in conflict with each other. For example, our decision to keep the EV of opening and refusing loot boxes equivalent to each other in both studies was justified in order to properly study reinforcement, but was clearly not representative of real-life loot boxes. Similarly, allowing loot box outcomes to be randomly generated accurately modeled how loot boxes would behave in a real game, but came at the cost of our ability to exert control over what participants experienced – leading some participants in Study 1 to complete a round or even the entire game without experiencing any wins. As such, we now discuss a number of modifications that could be made to the task in future studies, some of which would make the paradigm a better loot box and some that would make the paradigm a better Skinner box.

In order to make a better loot box, we first need to recognize that real loot boxes have no lose condition. Players may receive duplicate or undesired rewards from loot boxes (which may very well feel like a loss to the player), but it is impossible to receive no rewards at all (Newall, 2023). Although both Study 1 and 2 could be adjusted to follow a CRF schedule (with the zero-point rewards being replaced with a small point reward, such as a one or ten points), it would also benefit the study of human operant learning to manipulate magnitude variability and schedule variability independently. That is, reinforcer magnitude may be variable (as it was in Study 2) or fixed (as it was in Study 1), and this property of the reinforcement schedule is independent of the variability in response requirements if non-reinforcement is considered to be a special type of outcome. To further illustrate, all of the schedules in Study 1 had fixed magnitude and intermittent reinforcement. Reinforcement was not received after every response,

but when it was, the magnitude was always 500 points. A fixed magnitude, continuous reinforcement version of Study 1 would deliver reinforcement after every response, with the reinforcer magnitude being the same every time (with magnitudes calibrated such that they are not overwhelmingly beneficial to the player than the fixed magnitude, intermittent reinforcement schedules). Study 2, similarly, could be modified to include fixed magnitude conditions, with or without intermittent reinforcement.

Points were a convenient reward to use for our initial studies because their value and impact on study completion were standard across participants, but real loot boxes contain different types of rewards, sometimes even multiple reward types, such as in *Overwatch* (Ballou et al., 2022; Nielsen & Grabarczyk, 2019). Our paradigm should explore the use of rewards other than points. The idle game *Cookie Clicker* (https://cookieclicker.com/), our early inspiration for the button clicking task, features a variety of purchasable upgrades related to increasing the value of clicks, which could be implemented into our paradigm as loot box rewards. For instance, loot boxes could contain a temporary point doubler that makes clicks worth more for a short period of time. Additionally, user interface or other cosmetic upgrades could be added to the loot boxes as reinforcers, although it may be difficult to get participants invested in earning rewards for a laboratory-only game. As such, making the paradigm compatible for smartphones is a long-term goal.

Although we chose to ensure that the EV of opening or refusing the loot boxes were equivalent in both studies, this decision was based in caution rather than realism. Instead, real loot boxes are far more likely to follow traditional gambling methods and have negative EVs; obscured by extremely rare but high magnitude rewards (Rachlin, 1990; Rachlin et al., 2015). This issue is further complicated by the fact that not all loot box rewards can be assigned an

objective value (and therefore, have their EVs calculated), such as those that cannot be traded with other players whatsoever (e.g., *Hearthstone, Fortnite*). In that case, it is still possible to calculate expected values by asking participants to assign subjective values to potential rewards (Kahneman & Tversky, 1979), but these subjective values likely vary greatly across games and players. Of particular interest is the matter of duplicate rewards which, depending on the game, may be less valuable precisely because the player has already obtained it (Ballou et al., 2022).

In terms of making a better Skinner box, it's important to address the issue of experimental control as it applies to our paradigm. Although we opted to let outcomes vary randomly and be partially determined by the participant (i.e., outside of free trials, participants could always choose to avoid opening loot boxes) in order to better simulate real loot boxes, this decision may have backfired in practice. Looking at Study 1, not only was the average response requirement on the RR 10 schedule noticeably below ten (Figure 5), but the experienced MAD was generally much lower than what had been simulated (Figure 8). Naturally, it is entirely plausible that this is merely the result of random chance, but it may also be the case that the data have been censored, in that participants experiencing long streaks of loot boxes without reinforcement may have chosen to stop opening the boxes all together. Because the response requirements on the RR 10 worked differently from the other schedules (i.e., a 10% chance of reinforcement after each response rather than generating a response requirement from a specified range in advance), we do not know how many loot boxes it would have taken participants to finally win. As such, the RR 10 response requirements and experienced MAD shown in Figure 5 and Figure 8 respectively overrepresent small response requirements while large response requirements that would be present if participants had continued to open loot boxes are missing.

Furthermore, the impact of early and/or big wins on sustaining gambling behavior may be less straightforward than we initially realized. Edson et al. (2021) argued that Custer and Milt's (1985, cited from Edson et al., 2021) original definition was vague in terms of what constitutes 'big'. Weatherly et al. (2004) pointed out a similar issue with the 'early' aspect of big wins. As such, Edson et al. proposed two definitions of big wins: a monetary prize larger than a specified cutoff point (in Edson et al., a cutoff of $1,000 or more was used based on average U.S. income statistics) or a monetary prize with large prize to entry fee ratios relative to other wins that participants had experienced. Using sports betting data, Edson et al. (2023) and Edson et al. (2021) both found that large magnitude wins were predictive of future problematic gambling behaviors, although they found no evidence indicating said wins were more influential when they were experienced 'early', with the impacts of big wins instead diminishing over time. In regards to the current studies, the loot boxes in Study 1 would not be defined as a big win as there are no other prize magnitudes to which they can be compared, other than winning or not winning. In Study 2, the 500-point prize in Study 2 could be considered a big win, as it had the highest prize to fee ratio of 5.88 (500 points/85 point entry fee = 5.88) compared to the 100-point (100/85 = 1.18) and 50-point prizes (50/85 = 0.59), but the difference in ratios may not be large enough to influence behavior.

The issue is further complicated by additional evidence that big, early wins may actually make participants more likely to quit while they are ahead, as demonstrated by Weatherly et al. (2004). In the context of Studies 1 and 2, participants did not have the option to stop playing before they finished the game (outside of revoking their consent to participate in the study), but they could choose to stop opening loot boxes. As such, big wins may have actually contributed to the overall trend of participants opening less loot boxes the further they progressed in the study,

with participants refusing further loot boxes due to being satisfied with one or two wins (i.e., participants did not want to waste the effort they had saved by winning a loot box by prolonging the experiment with further non-reinforced trials).

We initially piloted a version of the Study 1 game in which the first loot box was rigged to always be a win, but did not pursue it after comparing pilot data across the two versions. Although some studies of experienced-based decision-making have taken to precisely controlling the order of outcomes (e.g., Ungemach et al., 2009), we still believe it is important to allow outcomes to vary as they would in real-life loot boxes. Based on the results from Edson et al. (2021, 2023) and Weatherly et al. (2004) and given that participants' early behavior was promising, implementing a pity timer mechanic to the RR 10 condition of Study 1 that ensures that participants receive reinforcement at predetermined times, may lead to much more stable behavior. However, while this modification would be an interesting study of pity timers, the drawback is that the unique properties of the RR 10 are no longer modeled, and the schedule would essentially become a VR schedule with a higher average response requirement.

Beyond the plethora of relatively simple modifications that could be made to the task, the paradigm also has the potential to innovate the study of outcome devaluation in humans. Devaluation, naturally, refers to a decrease in the subjective value of a reward, which influences subsequent operant responses for said reward (Adams & Dickinson, 1981). In animal learning studies, devaluation is generally introduced using one of two methods. The first method is food-aversion conditioning, in which a previous food reward becomes paired with illness by giving the subjects lithium chloride (e.g., Adams, 1982; Adams & Dickinson, 1981; Colwill & Rescorla, 1985; Dickinson et al., 1983). The second method is reward satiation, in which subjects are allowed free-access to a reinforcer, thus devaluing the reward either through

habituation or in the case of food rewards, literal satiation (e.g., Colwill & Rescorla, 1985; Jones et al., 2022). Neither of these methods translate well to experiments with human subjects; food poisoning is obviously too extreme of an aversive experience and satiation methods may be costly depending on the type of reinforcer used.

Previous devaluation tasks used with human participants include the virtual vending machine used by Morris et al. (2015), in which the main devaluation component was a short video of the snack food that participants had previously received as a reward being infested by cockroaches. Gillan et al. (2011) used a digital devaluation task in which participants learned to associate different on-screen fruits with certain point values, with devaluation operating as a reduction of points (that is, fruits became associated with less points than they were previously). Loot boxes lend themselves especially well to implementing devaluation through satiation via duplicate rewards. Depending on the game, receiving a duplicate reward may offer little to no benefits (Ballou et al., 2022), essentially sating players on that particular reward.

There is no telling whether the gaming industry will keep following the path it is currently on or shift away from loot boxes in the coming years, but we expect that loot box mechanics will evolve in response to regulations rather than be eradicated completely. While our experimental paradigm could be used to better understand how players engage with loot boxes in real-life games, it could also serve as a novel paradigm to study human operant conditioning. After all, people already choose to purchase loot boxes – perhaps the best Skinner box is one that people enter of their volition.

References

Adams, C. D. (1982). Variations in the sensitivity of instrumental responding to reinforcer devaluation. *The Quarterly Journal of Experimental Psychology Section B*, *34*(2b), 77–98. https://doi.org/10.1080/14640748208400878

Adams, C. D., & Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *The Quarterly Journal of Experimental Psychology Section B*, *33*(2b), 109–121. https://doi.org/10.1080/14640748108400816

Auer, M., Hopfgartner, N., Helic, D., & Griffiths, M. D. (2023). Self-reported deposits versus actual deposits in online gambling: An empirical study. *Journal of Gambling Studies*, *40*(2), 619–637. https://doi.org/10.1007/s10899-023-10230-1

Ballou, N. (2023). A manifesto for more productive psychological games research. *ACM Games: Research and Practice*, *1*(1), Article 2. https://doi.org/10.1145/3582929

Ballou, N., Gbadamosi, C., & Zendle, D. (2022). The hidden intricacy of loot box design: A granular description of random monetized reward features. *Proceedings of DiGRA 2022 Conference: Bringing Worlds Together.* https://doi.org/10.31234/osf.io/xeckb

Baron, A., Perone, M., & Galizio, M. (1991). Analyzing the reinforcement process at the human level: Can application and behavioristic interpretation replace laboratory research? *The Behavior Analyst*, *14*(2), 95–105. https://doi.org/10.1007/BF03392557

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). *Parsimonious Mixed Models* (No. arXiv:1506.04967). arXiv. http://arxiv.org/abs/1506.04967

Bates, D., Machler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Berardi, V., Hovell, M., Hurley, J. C., Phillips, C. B., Bellettiere, J., Todd, M., & Adams, M. A. (2020). Variable magnitude and frequency financial reinforcement is effective at increasing adults' free-living physical activity. *Perspectives on Behavior Science*, *43*(3), 515–538. https://doi.org/10.1007/s40614-019-00241-y

Bernoulli, N. (1713). *Correspondence of Nicolas Bernoulli concerning the St. Petersburg game* (R. J. Pulskamp, Trans.). https://www.probabilityandfinance.com/pulskamp/NBernoulli/correspondence_petersburg_game.pdf

Blizzard Entertainment. (2014a). *Hearthstone* [Video game].

Blizzard Entertainment. (2014b). *Heroes of the Storm* [Video game].

Blizzard Entertainment. (2016). *Overwatch* [Video game].

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, *23*(3), 389–411. https://doi.org/10.1037/met0000159

Brooks, G. A., & Clark, L. (2019). Associations between loot box use, problematic gaming and gambling, and gambling-related cognitions. *Addictive Behaviors*, *96*, 26–34. https://doi.org/10.1016/j.addbeh.2019.04.009

Brooks, G. A., & Clark, L. (2023). The gamblers of the future? Migration from loot boxes to gambling in a longitudinal study of young adults. *Computers in Human Behavior*, *141*, Article 107605. https://doi.org/10.1016/j.chb.2022.107605

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/doi:10.18637/jss.v080.i01

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105. https://psycnet.apa.org/doi/10.1037/h0046016

Close, J., Spicer, S. G., Nicklin, L. L., Uther, M., Lloyd, J., & Lloyd, H. (2021). Secondary analysis of loot box data: Are high-spending "whales" wealthy gamers or problem gamblers? *Addictive Behaviors*, *117*, Article 106851. https://doi.org/10.1016/j.addbeh.2021.106851

Colwill, R. M., & Rescorla, R. A. (1985). Postconditioning devaluation of a reinforcer affects instrumental responding. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*(1), 120–132. https://doi.org/10.1037/0097-7403.11.1.120

Crossman, E. (1983). Las Vegas knows better. *The Behavior Analyst*, *6*, 109–110. https://doi.org/10.1007/BF03391879

Crossman, E. K., Bonem, E. J., & Phelps, B. J. (1987). A comparison of response patterns on fixed-, variable-, and random-ratio schedules. *Journal of the Experimental Analysis of Behavior*, *48*(3), 395–406. https://doi.org/10.1901/jeab.1987.48-395

Davison, M., & Baum, W. M. (2003). Every reinforcer counts: Reinforcer magnitude and local preference. *Journal of the Experimental Analysis of Behavior*, *80*(1), 95–129. https://doi.org/10.1901/jeab.2003.80-95

Dickinson, A., Nicholas, D. J., & Adams, C. D. (1983). The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *The Quarterly Journal of Experimental Psychology Section B*, *35*(1b), 35–51. https://doi.org/10.1080/14640748308400912

Dixon, M. R., & Schreiber, J. E. (2004). Near-miss effects on response latencies and win estimations of slot machine players. *The Psychological Record*, *54*(3), 335–348. https://doi.org/10.1007/BF03395477

Drummond, A., & Sauer, J. D. (2018). Video game loot boxes are psychologically akin to gambling. *Nature Human Behaviour*, *2*(8), 530–532. https://doi.org/10.1038/s41562-018-0360-1

Drummond, A., Sauer, J. D., Ferguson, C. J., & Hall, L. C. (2020). The relationship between problem gambling, excessive gaming, psychological distress and spending on loot boxes in Aotearoa New Zealand, Australia, and the United States—A cross-national survey. *PLoS One*, *15*(3), e0230378.

EA Vancouver, & EA Romania. (2021). *FIFA 2022* [Video game].

Edson, T. C., Louderback, E. R., Tom, M. A., Philander, K. S., Slabczynski, J. M., Lee, T. G., & LaPlante, D. A. (2023). A large-scale prospective study of big wins and their relationship with future involvement and risk among actual online sports bettors. *Computers in Human Behavior*, *142*, 107657. https://doi.org/10.1016/j.chb.2023.107657

Edson, T. C., Tom, M. A., Philander, K. S., Louderback, E. R., & LaPlante, D. A. (2021). A large-scale prospective study of big wins and their relationship with future financial and time involvement in actual daily fantasy sports contests. *Psychology of Addictive Behaviors*, *35*(8), 948–960. https://doi.org/10.1037/adb0000715

Electronic Arts Digital Illusions Creative Entertainment. (2017). *Star Wars Battlefront II* [Video game].

Elliffe, D., Davison, M., & Landon, J. (2008). Relative reinforcer rates and magnitudes do not control concurrent choice independently. *Journal of the Experimental Analysis of Behavior*, *90*(2), 169–185.

Epic Games. (2017). *Fortnite* [Video game].

Felton, M., & Lyon, D. O. (1966). The post-reinforcement pause. *Journal of the Experimental Analysis of Behavior*, *9*(2), 131–134. https://doi.org/10.1901/jeab.1966.9-131

Ferris, J. A., & Wynne, H. J. (2001). *The Canadian problem gambling index*. Canadian Centre on substance abuse Ottawa, ON. http://jogoremoto.pt/docs/extra/TECb6h.pdf

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. Appleton-Century-Crofts.

Field, D. P., Tonneau, F., Ahearn, W., & Hineline, P. N. (1996). Preference between variable-ratio and fixed-ratio schedules: Local and extended relations. *Journal of the Experimental Analysis of Behavior*, *66*(3), 283–295. https://doi.org/10.1901/jeab.1996.66-283

Forsström, D., Chahin, G., Savander, S., Mentzoni, R. A., & Gainsbury, S. (2022). Measuring loot box consumption and negative consequences: Psychometric investigation of a

Swedish version of the Risky Loot Box Index. *Addictive Behaviors Reports*, *16*, Article 100453. https://doi.org/10.1016/j.abrep.2022.100453

Fox, C., & Hadar, L. (2006). "Description from Experience" = sampling error + prospect theory: Reconsidering Hertwig, Barron, Weber, & Erev (2004). *Judgement and Decision Making*, *1*(2), 159–161. https://doi.org/10.1017/S1930297500002370

Gach, E. (2017, November 29). Meet the 19-year-old who spent over $17,000 on microtransactions. *Kotaku Australia*. https://www.kotaku.com.au/2017/11/meet-the-19-year-old-who-spent-over-17000-on-microtransactions/

Garea, S. S., Drummond, A., Sauer, J. D., Hall, L. C., & Williams, M. N. (2021). Meta-analysis of the relationship between problem gambling, excessive gaming and loot box spending. *International Gambling Studies*, *21*(3), 460–479. https://doi.org/10.1080/14459795.2021.1914705

Garea, S. S., Sauer, J. D., Hall, L. C., Williams, M. N., & Drummond, A. (2023). The potential relationship between loot box spending, problem gambling, and obsessive-compulsive gamers. *Journal of Behavioral Addictions*, *12*(3), 733–743. https://doi.org/10.1556/2006.2023.00038

Garrett, E. P., Drummond, A., Lowe-Calverley, E., de Salas, K., Lewis, I., & Sauer, J. D. (2023). Impulsivity and loot box engagement. *Telematics and Informatics*, *78*, Article 101952. https://doi.org/10.1016/j.tele.2023.101952

Gibbon, J., Farrell, L., Locurto, C. M., Duncan, H. J., & Terrace, H. S. (1980). Partial reinforcement in autoshaping with pigeons. *Animal Learning & Behavior*, *8*(1), 45–59. https://doi.org/10.3758/BF03209729

Gillan, C. M., Papmeyer, M., Morein-Zamir, S., Sahakian, B. J., Fineberg, N. A., Robbins, T. W., & De Wit, S. (2011). Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder. *American Journal of Psychiatry*, *168*(7), 718–726. https://doi.org/10.1176/appi.ajp.2011.10071062

Gottlieb, D. A. (2004). Acquisition with partial and continuous reinforcement in pigeon autoshaping. *Learning & Behavior*, *32*(3), 321–334. https://doi.org/10.3758/BF03196031

Griffiths, R. R., & Thompson, T. (1973). The post-reinforcement pause: A misnomer. *The Psychological Record*, *23*(2), 229–235. https://doi.org/10.1007/BF03394160

Hall, L. C., Drummond, A., Sauer, J. D., & Ferguson, C. J. (2021). Effects of self-isolation and quarantine on loot box spending and excessive gaming—Results of a natural experiment. *PeerJ*, *9*, e10705.

Harless, D. W., & Camerer, C. F. (1994). The predictive utility of generalized expected utility theories. *Econometrica*, *62*(6), 1251–1289. https://doi.org/10.2307/2951749

Hau, R., Pleskac, T. J., Kiefer, J., & Hertwig, R. (2008). The description-experience gap in risky choice: The role of sample size and experienced probabilities. *Journal of Behavioral Decision Making*, *21*, 493–518.

Haw, J. (2008). Random-ratio schedules of reinforcement: The role of early wins and unreinforced trials. *Journal of Gambling Issues*, *21*(21), 56–67. https://doi.org/10.4309/jgi.2008.21.6

Hayden, B. Y., & Platt, M. L. (2009). The mean, the median, and the St. Petersburg paradox. *Judgment and Decision Making*, *4*(4), 256–272. https://doi.org/10.1017/S1930297500003831

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, *15*(8), 534–539. https://doi.org/10.1111/j.0956-7976.2004.00715.x

Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences*, *13*(12), 517–523. https://doi.org/10.1016/j.tics.2009.09.004

Hertwig, R., & Pleskac, T. J. (2010). Decisions from experience: Why small samples? *Cognition*, *115*, 225–237. https://doi.org/10.1016/j.cognition.2009.12.009

Hurlburt, R. T., Knapp, T. J., & Knowles, S. H. (1980). Simulated slot-machine play with concurrent variable ratio and random ratio schedules of reinforcement. *Psychological Reports*, *47*(2), 635–639. https://doi.org/10.2466/pr0.1980.47.2.635

Jones, B. O., Cruz, A. M., Kim, T. H., Spencer, H. F., & Smith, R. J. (2022). Discriminating goal-directed and habitual cocaine seeking in rats using a novel outcome devaluation procedure. *Learning & Memory*, *29*(12), 447–457. https://doi.org/10.1101/lm.053621.122

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291. https://doi.org/10.2307/1914185

Kao, D. (2019). Infinite loot box: A platform for simulating video game loot boxes. *IEEE Transactions on Games*, *12*(2), 219–224. https://doi.org/10.1109/TG.2019.2913320

Kassinove, J. I., & Schare, M. L. (2001). Effects of the "near miss" and the "big win" on persistence at slot machine gambling. *Psychology of Addictive Behaviors*, *15*(2), 155–158. https://doi.org/10.1037/0893-164X.15.2.155

Kim, N. (2018). *Judgement and decision-making: In the lab and the world.* Palgrave.

King, D. L., & Delfabbro, P. H. (2018). Predatory monetization schemes in video games (eg 'loot boxes') and internet gaming disorder [Editorial]. *Addiction*, *113*(11), 1967–1969. https://psycnet.apa.org/doi/10.1111/add.14286

Knight, F. H. (1921). *Risk, uncertainty and profit* (Vol. 31). Houghton Mifflin.

Konstantinidis, E., Taylor, R. T., & Newell, B. R. (2018). Magnitude and incentives: Revisiting the overweighting of extreme events in risky decisions from experience. *Psychonomic Bulletin & Review*, *25*(5), 1925–1933. https://doi.org/10.3758/s13423-017-1383-8

Lakens, D. (2022). Sample size justification. *Collabra: Psychology*, *8*(1), Article 3267. https://doi.org/10.1525/collabra.33267

Landon, J., Davison, M., & Elliffe, D. (2003). Concurrent schedules: Reinforcer magnitude effects. *Journal of the Experimental Analysis of Behavior*, *79*(3), 351–365. https://doi.org/10.1901/jeab.2003.79-351

Larche, C. J., Chini, K., Lee, C., & Dixon, M. J. (2023). To pay or just play? Examining individual differences between purchasers and earners of loot boxes in Overwatch. *Journal of Gambling Studies*, *39*(2), 625–643. https://doi.org/10.1007/s10899-022-10127-5

Larche, C. J., Chini, K., Lee, C., Dixon, M. J., & Fernandes, M. (2021). Rare loot box rewards trigger larger arousal and reward responses, and greater urge to open more loot boxes. *Journal of Gambling Studies*, *37*, 141–163. https://doi.org/10.1007/s10899-019-09913-5

Lehr, R. (1970). Partial reinforcement and variable magnitude of reward effects in rats in a T maze. *Journal of Comparative and Physiological Psychology*, *70*(2, Pt.1), 286–293. https://doi.org/10.1037/h0028739

Li, W., Mills, D., & Nower, L. (2019). The relationship of loot box purchases to problem video gaming and problem gambling. *Addictive Behaviors*, *97*, 27–34.

Macey, J., & Hamari, J. (2019). eSports, skins and loot boxes: Participants, practices and problematic behaviour associated with emergent forms of gambling. *New Media & Society*, *21*(1), 20–41. https://doi.org/10.1177/1461444818786216

MacKillop, J., Miller, J. D., Fortune, E., Maples, J., Lance, C. E., Campbell, W. K., & Goodie, A. S. (2014). Multidimensional examination of impulsivity in relation to disordered gambling. *Experimental and Clinical Psychopharmacology*, *22*(2), 176–185. https://doi.org/10.1037/a0035874

Martinez, L., & Zeilberger, D. (2024). A guide to the risk-averse gambler and resolving the St. Petersburg paradox once and for all. *The College Mathematics Journal*, *55*(3), 226–234. https://doi.org/10.1080/07468342.2024.2305103

Mattinen, T., Macey, J., & Hamari, J. (2023). A ruse by any other name: Comparing loot boxes and collectible card games using Magic Arena. *Proceedings of the ACM on Human-Computer Interaction*, *7*(CHI PLAY), 721–747. https://doi.org/https://doi.org/10.1145/3611047

Mazur, J. E. (1983). Steady-state performance on fixed-, mixed-, and random-ratio schedules. *Journal of the Experimental Analysis of Behavior*, *39*(2), 293–307. https://doi.org/10.1901/jeab.1983.39-293

McClelland, G. H. (1997). Optimal design in psychological research. *Psychological Methods*, *2*(1), 3.

Mehlhorn, K., Newell, B. R., Todd, P. M., Lee, M. D., Morgan, K., Braithwaite, V. A., Hausmann, D., Fiedler, K., & Gonzalez, C. (2015). Unpacking the exploration–exploitation tradeoff: A synthesis of human and animal literatures. *Decision*, *2*(3), 191–215. https://doi.org/10.1037/dec0000033

Morris, R. W., Quail, S., Griffiths, K. R., Green, M. J., & Balleine, B. W. (2015). Corticostriatal control of goal-directed action is impaired in schizophrenia. *Biological Psychiatry*, *77*(2), 187–195. https://doi.org/10.1016/j.biopsych.2014.06.005

Newall, P. (2023). Beyond gambling: The dangers of analogistic reasoning in addiction science, and how loot box psychology should create its own unique theory. *Addiction Research & Theory*, 1–6. https://doi.org/10.1080/16066359.2023.2279082

Newell, B. R., & Rakow, T. (2007). The role of experience in decisions from description. *Psychonomic Bulletin & Review*, *14*(6), 1133–1139. https://doi.org/10.3758/BF03193102

Niantic. (2016). *Pokémon Go* [Video game].

Nie, D., Hu, Z., Yang, J., & Zhu, D. (2024). How robust and how large is the description-experience gap? -- Evidence from an eye-tracking research. *Current Psychology*, *43*(42), 32824–32836. https://doi.org/10.1007/s12144-024-06616-y

Nielsen, R. K. L., & Grabarczyk, P. (2019). Are loot boxes gambling?: Random reward mechanisms in video games. *Transactions of the Digital Games Research Association*, *4*, 171–207.

Niemeyer, K. (2024, January 21). "Counter-Strike" raked in nearly $1 billion in 2023 from randomized loot boxes. *Business Insider*. https://www.businessinsider.com/counter-strike-1-billion-loot-boxes-gambling-esports-video-games-2024-1

Nintendo EPD. (2019). *Mario Kart Tour* [Video game].

Papini, M. R., & Overmier, J. B. (1985). Partial reinforcement and autoshaping of the pigeon's key-peck behavior. *Learning and Motivation*, *16*(1), 109–123. https://doi.org/10.1016/0023-9690(85)90007-4

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt impulsiveness scale. *Journal of Clinical Psychology*, *51*(6), 768–774. https://doi.org/10.1002/1097-4679(199511)51:6%253C768::AID-JCLP2270510607%253E3.0.CO;2-1

pjs. (2020, June 17). *Answer to "How to lightly shuffle a list in python"* [Online post]. Stack Overflow. https://stackoverflow.com/a/62438368

Platt, J. R. (1964). Strong inference: Certain systematic methods of scientific thinking may produce much more rapid progress than others. *Science*, *146*(3642), 347–353. https://doi.org/10.1126/science.146.3642.347

Powell, R. W. (1969). The effect of reinforcement magnitude upon responding under fixed-ratio schedules. *Journal of the Experimental Analysis of Behavior*, *12*(4), 605–608. https://doi.org/10.1901/jeab.1969.12-605

Primi, C., Sanson, F., & Donati, M. A. (2024). Measuring risky loot box use: An item response theory analysis of the Risky Loot Box Index among adolescents. *Psychology of Addictive Behaviors*. https://doi.org/10.1037/adb0001009

Psyonix. (2015). *Rocket League* [Video game].

Rachlin, H. (1990). Why do people gamble and keep gambling despite heavy losses? *Psychological Science*, *1*(5), 294–297. https://doi.org/10.1111/j.1467-9280.1990.tb00220.x

Rachlin, H., Safin, V., Arfer, K. B., & Yen, M. (2015). The attraction of gambling. *Journal of the Experimental Analysis of Behavior*, *103*(1), 260–266. https://doi.org/10.1002/jeab.113

Rakow, T., & Newell, B. (2010). Degrees of uncertainty: An overview and framework for future research on experience-based choice. *Journal of Behavioral Decision Making*, *23*, 1–14.

Riot Games. (2009). *League of Legends* [Video game].

Scandola, M., & Tidoni, E. (2024). Reliability and feasibility of linear mixed models in fully crossed experimental designs. *Advances in Method and Practices in Psychological Science*, *7*(1), 1–21. https://doi.org/10.1177/25152459231214454

Schoenfeld, W. N., Cumming, W. W., & Hearst, E. (1956). On the classification of reinforcement schedules. *Proceedings of the National Academy of Sciences*, *42*(8), 563–570. https://doi.org/10.1073/pnas.42.8.563

Seidl, C. (2013). The St. Petersburg paradox at 300. *Journal of Risk and Uncertainty*, *46*(3), 247–264. https://doi.org/10.1007/s11166-013-9165-9

Sidloski, B., Brooks, G. A., Zhang, K., & Clark, L. (2022). Exploring the association between loot boxes and problem gambling: Are video gamers referring to loot boxes when they complete gambling screening tools? *Addictive Behaviors*, *131*, 107318.

Skinner, B. F. (1953). *Science and human behavior*. Simon and Schuster.

Spicer, S. G., Fullwood, C., Close, J., Nicklin, L. L., Lloyd, J., & Lloyd, H. (2022). Loot boxes and problem gambling: Investigating the "gateway hypothesis." *Addictive Behaviors*, *131*, Article 107327. https://doi.org/10.1016/j.addbeh.2022.107327

Supercell. (2012). *Clash of Clans* [Video game].

Thorndike, E. (1901). *Animal intelligence: Experimental studies*. Routledge.

Todorov, J. C. (1973). Interaction of frequency and magnitude of reinforcement on concurrent performances. *Journal of the Experimental Analysis of Behavior*, *19*(3), 451–458. https://doi.org/10.1901/jeab.1973.19-451

Todorov, J. C., Hanna, E. S., & Sá, M. C. N. B. D. (1984). Frequency versus magnitude of reinforcement: New data with a new procedure. *Journal of the Experimental Analysis of Behavior*, *41*(2), 157–167. https://doi.org/10.1901/jeab.1984.41-157

Ungemach, C., Chater, N., & Stewart, N. (2009). Are probabilities overweighted or underweighted when rare outcomes are experienced (rarely)? *Psychological Science*, *20*(4), 473–479. https://doi.org/10.1111/j.1467-9280.2009.02319.x

Unity Technologies. (2022). *Unity* (Version 2022.3.15f1) [Computer software].

Valve, & Hidden Path Entertainment. (2012). *Counter Strike: Global Offensive* [Video game].

Weatherly, J. N., Sauter, J. M., & King, B. M. (2004). The "big win" and resistance to extinction when gambling. *The Journal of Psychology*, *138*(6), 495–504. https://doi.org/10.3200/JRLP.138.6.495-504

Xiao, L. Y. (2021). Regulating loot boxes as gambling? Towards a combined legal and self-regulatory consumer protection approach. *Interactive Entertainment Law Review*, *4*(1), 27–47. https://doi.org/10.4337/ielr.2021.01.02

Xiao, L. Y., Fraser, T. C., & Newall, P. W. S. (2022). Opening Pandora's loot box: Weak links between gambling and loot box expenditure in China, and player opinions on probability disclosures and pity-timers. *Journal of Gambling Studies*, *39*(2), 645–668. https://doi.org/10.1007/s10899-022-10148-0

Xiao, L. Y., Fraser, T. C., Nielsen, R. K. L., & Newall, P. W. (2024). Loot boxes, gambling-related risk factors, and mental health in Mainland China: A large-scale survey. *Addictive Behaviors*, *148*, 107860.

Xiao, L. Y., & Newall, P. (2022). Probability disclosures are not enough: Reducing loot box reward complexity as a part of ethical video game design. *Journal of Gambling Issues*, 1–18. https://doi.org/10.4309/LDOM8890

Zendle, D. (2019). Problem gamblers spend less money when loot boxes are removed from a game: A before and after study of Heroes of the Storm. *PeerJ*, *7*, Article e7700. https://doi.org/10.7717/peerj.7700

Zendle, D., & Cairns, P. (2018). Video game loot boxes are linked to problem gambling: Results of a large-scale survey. *PloS One*, *13*(11), Article e0206767. https://doi.org/10.1371/journal.pone.0206767

Zendle, D., & Cairns, P. (2019). Loot boxes are again linked to problem gambling: Results of a replication study. *PloS One*, *14*(3), Article e0213194. https://doi.org/10.1371/journal.pone.0213194

Zendle, D., Cairns, P., Barnett, H., & McCall, C. (2020). Paying for loot boxes is linked to problem gambling, regardless of specific features like cash-out and pay-to-win. *Computers in Human Behavior*, *102*, 181–191. https://doi.org/10.1016/j.chb.2019.07.003

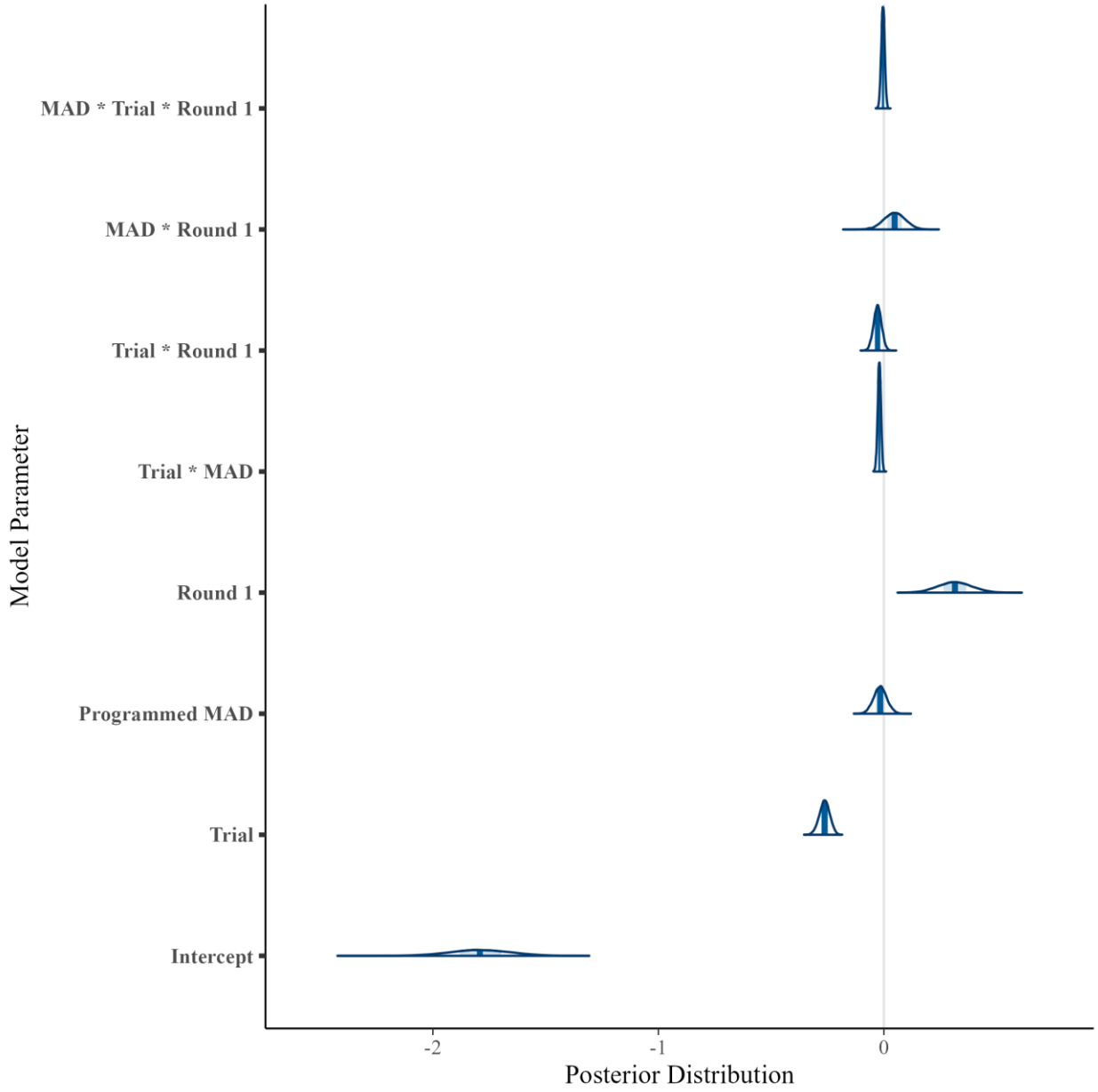
Zendle, D., Meyer, R., Cairns, P., Waters, S., & Ballou, N. (2020). The prevalence of loot boxes in mobile and desktop games. *Addiction*, *115*(9), 1768–1772. https://doi.org/10.1111/add.14973

Zendle, D., Meyer, R., & Over, H. (2019). Adolescents and loot boxes: Links with problem gambling and motivations for purchase. *Royal Society Open Science*, *6*(6), Article 190049. https://doi.org/10.1098/rsos.190049

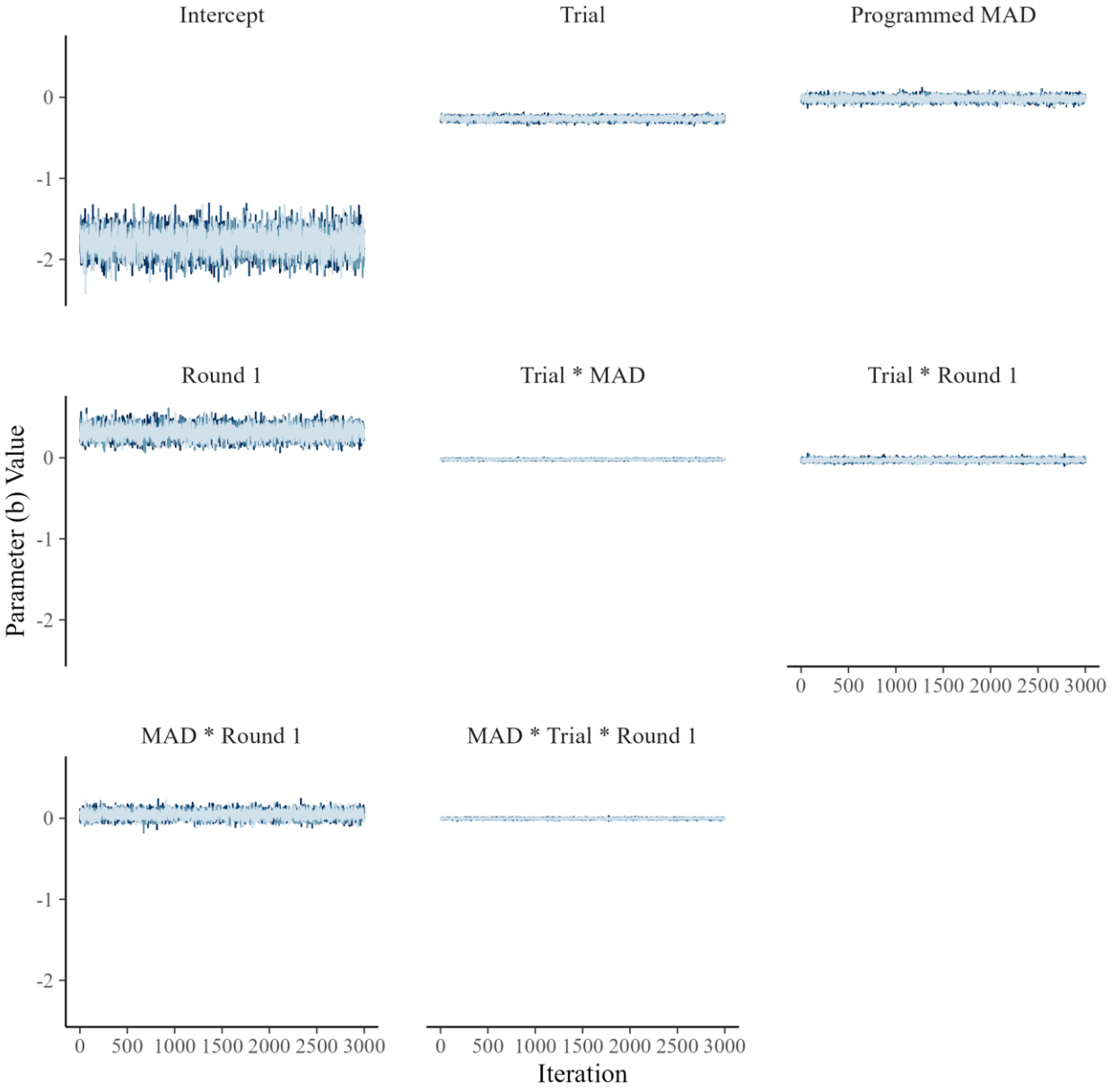
Zendle, D., Walasek, L., Cairns, P., Meyer, R., & Drummond, A. (2021). Links between problem gambling and spending on booster packs in collectible card games: A conceptual replication of research on loot boxes. *Plos One*, *16*(4), Article e0247855. https://doi.org/10.1371/journal.pone.0247855

Appendix A - Study 1 Supplemental Bayesian Figures

Appendix A Figure 14: *Study 1* **-** *End of Trial Choice Posterior Parameter Distributions*

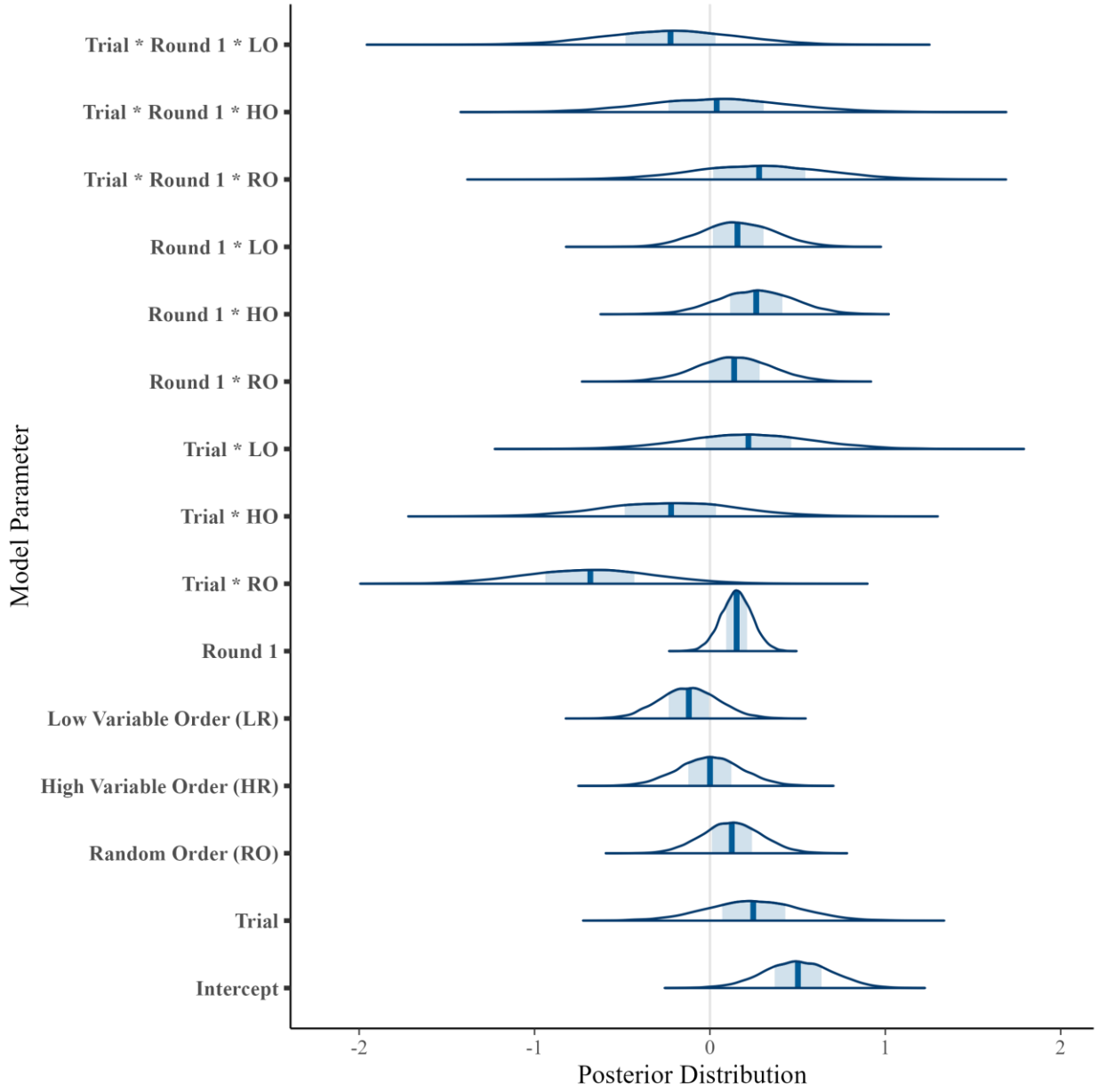


Appendix A Figure 15: *Study 1 - End of Trial Choice MCMC Chains*



Appendix B - Study 2 Supplemental Bayesian Figures

Appendix B Figure 16: *Study 2 - End of Trial Choice Posterior Parameter Distributions*



Appendix B Figure 17: *Study 2 - End of Trial Choice MCMC Chains*