

A REINFORCEMENT-LEARNING APPROACH TO UNDERSTANDING LOSS-CHASING
BEHAVIOR IN RATS

by

ANDREW THOMAS MARSHALL

B.S., Brown University, 2010
M.S., Kansas State University, 2013

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Psychological Sciences
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2016

Abstract

Risky decisions are inherently characterized by the potential to receive gains and losses from these choices, and gains and losses have distinct effects on global risky choice behavior and the likelihoods of making risky choices depending on the outcome of the previous choice. One translationally-relevant phenomenon of risky choice is loss-chasing, in which individuals make risky choices following losses. However, the mechanisms of loss-chasing are poorly understood. The goal of two experiments was to illuminate the mechanisms governing individual differences in loss-chasing and risky choice behaviors. In two experiments, rats chose between a certain outcome that always delivered reward and a risky outcome that probabilistically delivered reward. In Experiment 1, loss processing and loss-chasing behavior were assessed in the context of losses-disguised-as-wins (LDWs), or loss outcomes presented along with gain-related stimuli. The rats presented with LDWs were riskier and less sensitive to differential losses. In Experiment 2, these behaviors were assessed relative to the number of risky losses that could be experienced. Here, the addition of reward omission or a small non-zero loss to the possible risky outcomes elicited substantial individual differences in risky choice, with some rats increasing, decreasing, or maintaining their previous risky choice preferences. Several reinforcement learning (RL) models were fit to individual rats' data to elucidate the possible psychological mechanisms that best accounted for individual differences in risky choice and loss-chasing behaviors. The RL analyses indicated that the critical predictors of risky choice and loss-chasing behavior were the different rates that individuals updated value estimates with newly experienced gains and losses. Thus, learning deficits may predict individual differences in maladaptive risky decision making. Accordingly, targeted interventions to alleviate learning deficits may ultimately increase the likelihood of making more optimal and informed choices.

A REINFORCEMENT-LEARNING APPROACH TO UNDERSTANDING LOSS-CHASING
BEHAVIOR IN RATS

by

ANDREW THOMAS MARSHALL

B.S., Brown University, 2010
M.S., Kansas State University, 2013

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Psychological Sciences
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2016

Approved by:

Major Professor
Kimberly Kirkpatrick, Ph.D.

Copyright

ANDREW THOMAS MARSHALL

2016

Abstract

Risky decisions are inherently characterized by the potential to receive gains and losses from these choices, and gains and losses have distinct effects on global risky choice behavior and the likelihoods of making risky choices depending on the outcome of the previous choice. One translationally-relevant phenomenon of risky choice is loss-chasing, in which individuals make risky choices following losses. However, the mechanisms of loss-chasing are poorly understood. The goal of two experiments was to illuminate the mechanisms governing individual differences in loss-chasing and risky choice behaviors. In two experiments, rats chose between a certain outcome that always delivered reward and a risky outcome that probabilistically delivered reward. In Experiment 1, loss processing and loss-chasing behavior were assessed in the context of losses-disguised-as-wins (LDWs), or loss outcomes presented along with gain-related stimuli. The rats presented with LDWs were riskier and less sensitive to differential losses. In Experiment 2, these behaviors were assessed relative to the number of risky losses that could be experienced. Here, the addition of reward omission or a small non-zero loss to the possible risky outcomes elicited substantial individual differences in risky choice, with some rats increasing, decreasing, or maintaining their previous risky choice preferences. Several reinforcement learning (RL) models were fit to individual rats' data to elucidate the possible psychological mechanisms that best accounted for individual differences in risky choice and loss-chasing behaviors. The RL analyses indicated that the critical predictors of risky choice and loss-chasing behavior were the different rates that individuals updated value estimates with newly experienced gains and losses. Thus, learning deficits may predict individual differences in maladaptive risky decision making. Accordingly, targeted interventions to alleviate learning deficits may ultimately increase the likelihood of making more optimal and informed choices.

Table of Contents

List of Figures	ix
List of Tables	xiv
Acknowledgements	xvi
Chapter 1 - Introduction	1
Gains and Losses	1
All Gains and Losses are Not Created Equally	4
Mechanisms of Loss-Chasing	6
Chapter 2 - Experiment 1	13
Method	16
Animals	16
Apparatus	17
Procedure	17
Magazine and lever-press training	17
Risky choice task	18
Data analysis	20
Choice behavior	21
Molar analysis	21
Molecular analysis	22
Goal-tracking behavior	24
Reinforcement learning (RL) models	25
Simple RL	25
Asymmetric RL	26
Valence-Attentive RL	27
Weighted-Reference-Point RL	27
Decision rule	28
Model fitting and selection	29
Results and Discussion	30
Choice behavior	30
Molar analysis	30

Overall effects.....	30
Individual differences	34
Molecular analysis	36
Effect of the previous outcome	36
Outcome and choice history.....	41
Goal-tracking behavior	44
Overall effects.....	44
Individual differences	47
RL models	47
Parameter recovery.	48
Model selection.....	49
Model fit (Asymmetric RL).....	49
Interim summary	54
Chapter 3 - Experiment 2.....	80
Method	83
Animals	83
Apparatus	83
Procedure	83
Magazine and lever-press training.....	83
Risky choice task.	83
Data analysis	85
Choice behavior.	85
Reinforcement learning (RL) models.	86
Results and Discussion	86
Choice behavior	86
Molar analysis.....	86
Overall effects.....	86
Individual differences	90
Molecular analysis	92
Effect of the previous outcome	92
Outcome and choice history.....	99

RL models	101
Parameter recovery.	101
Model selection.	101
Model fit (Asymmetric RL)	103
Chapter 4 - General Discussion	136

List of Figures

Figure 2.1 External receptacle for receiving food pellets delivered from the alternative feeder for Group Extra-Feedback. 65

Figure 2.2 Proportion of choices for the risky outcome for Groups Extra-Feedback and Normal-Feedback as a function of session, phase (P[0], P[1]), and the probability of the R-0 and R-1 outcomes in the P[0] and P[1] phases (i.e., sub-phase), respectively. The $P[0] | P[1] = .10$ data are shown separately for Phase 1 and Phase 4. 66

Figure 2.3 Proportion of choices for the risky outcome as a function of session and sub-phase (i.e., probability of manipulated outcome) for individual rats in the P[0] phase. The dot-dash ($P[0] = .10$ (Data: P1)), dashed ($P[0] = .90$ (Data: P2)), dotted ($P[0] = .50$ (Data: P3)), and solid lines ($P[0] = .10$ (Data: P4)) represent the individual subject-fits from the best-fitting generalized linear mixed effects model. The rats in Group Normal-Feedback are labeled as “N.” and the rats in Group Extra-Feedback are labeled as “E.”. 67

Figure 2.4 Proportion of choices for the risky outcome as a function of session and sub-phase (i.e., probability of manipulated outcome) for individual rats in the P[1] phase. The dot-dash ($P[1] = .10$ (Data: P1)), dashed ($P[1] = .90$ (Data: P2)), dotted ($P[1] = .50$ (Data: P3)), and solid lines ($P[1] = .10$ (Data: P4)) represent the individual subject-fits from the best-fitting generalized linear mixed effects model. The rats in Group Normal-Feedback are labeled as “N.” and the rats in Group Extra-Feedback are labeled as “E.”. 68

Figure 2.5 Left: Proportion of choices for the risky outcome for Groups Extra-Feedback and Normal-Feedback as a function of phase (P[0], P[1]) and the outcome of the previous choice. Right: Proportion of choices for the risky outcome for Groups Extra-Feedback and Normal-Feedback as a function of probability of the manipulated outcome and the outcome of the previous choice. The dotted lines connecting the data points are intended to demonstrate the relationship between risky choice behavior following differential outcomes rather than to imply a continuous relationship. C-2 = certain choice – 2 pellets; C-4 = certain choice – 4 pellets; R-0 = risky choice – 0 pellets; R-1 = risky choice – 1 pellet; R-11 = risky choice – 11 pellets. 69

Figure 2.6 Proportion of choices for the risky outcome as a function of the outcome of the previous choice and phase for individual rats. The dashed (P[0]) and dotted lines (P[1])

represent the individual subject-fits from the best-fitting generalized linear mixed effects model. The rats in Group Normal-Feedback are labeled as “N.” and the rats in Group Extra-Feedback are labeled as “E.”. 70

Figure 2.7 Regression coefficients from the mixed-effects analysis evaluating the effect of the previous 9 outcomes on subsequent risky choice. N = Group Normal-Feedback; E = Group Extra-Feedback; Hyp. = hyperbolic model fit; Exp. = exponential model fit. 71

Figure 2.8 Normalized regression coefficients from the mixed-effects analysis evaluating the effect of the previous 10 choices on subsequent risky choice. N = Group Normal-Feedback; E = Group Extra-Feedback; Hyp. = hyperbolic model fit; Exp. = exponential model fit. ... 72

Figure 2.9 Mean number of head entries into the food magazine during the 5-s temporal window following the onset of outcome delivery as a function of risky outcome magnitude and group. 73

Figure 2.10 Number of head entries into the food magazine during the 5-s temporal window following the onset of outcome delivery as a function of risky outcome magnitude and phase for individual rats. The dashed (P[0]) and dotted lines (P[1]) represent the individual subject-fits from the best-fitting generalized linear mixed effects model. The rats in Group Normal-Feedback are labeled as “N.” and the rats in Group Extra-Feedback are labeled as “E.” 74

Figure 2.11 Parameter recovery for the Simple, Asymmetric, Valence-Attentive, and Weighted-Reference-Point models in Experiment 1. The thick dark lines are the unit diagonals. 75

Figure 2.12 Akaike Information Criterion (AIC) for the Asymmetric Reinforcement Learning (RL) model plotted against the AICs of the Simple and Valence-Attentive RL models. Each data point represents the fit to an individual rat’s data. Model fits that did not converge were omitted. 76

Figure 2.13 Choice data (Data) and the predicted choice data from the Asymmetric RL model (Model) for each rat, ordered from top to bottom by how well the model fit the data in terms of omega-squared (ω^2). “Min” and “Max” refer to the data with the lowest and highest ω^2 , respectively. “25%ile”, “Mdn”, and “75%ile” refer to the 25th, 50th, and 75th percentile of ω^2 , respectively. The alternating gray and white boxes within each panel represent different phases and sub-phases (see Table 2.1). The differential alignments of these phase markers reflect individual differences in trials completed within each phase. 77

Figure 2.14 Log-transformed mean risky choice plotted against the log-transformed α_G (left), α_L (center), and β parameters (right) as derived from the fitting of the Asymmetric RL model. Each data point represents an individual rat. The goodness-of-fit index of the best fitting regression line for the overall model is shown (R^2). 78

Figure 2.15 Win-Stay (top panels) and Loss Sensitivity indices (bottom panels) plotted against the log-transformed α_G (left panels), α_L (middle panels), and α_L to α_G ratio (right panels) as derived from the fitting of the Asymmetric RL model. Each data point represents an individual rat. The rats in Group Normal-Feedback and Group Extra-Feedback are distinguished. 79

Figure 3.1 Proportion of choices for the risky outcome for each group (Equal-Risk, Unequal-Risk) across risk paradigms (Risk-Omission, Risk-Variability) as a function of loss condition (One-Loss, Two-Loss). 119

Figure 3.2 Proportion of choices for the risky outcome for each group (Equal-Risk, Unequal-Risk) across risk paradigms (Risk-Omission, Risk-Variability) as a function of the session in the One-Loss and Two-Loss conditions. The dotted line designates the data corresponding to the One-Loss (left) and Two-Loss conditions (right). In the figure, the abscissa is labeled as beginning at “1” and ending at “20” for the purpose of interpretability, but sessions were nested within each phase and sub-phase for analyses. 120

Figure 3.3 Proportion of choices for the risky outcome as a function of session and loss condition for individual rats within the Risk-Omission phases. The solid (Omission: One-Loss) and dashed lines (Omission: Two-Loss) represent the individual subject-fits from the best-fitting generalized linear mixed effects model. The rats in Group Equal-Risk are labeled as “E.” and the rats in Group Unequal-Risk are labeled as “U.”. 121

Figure 3.4 Proportion of choices for the risky outcome as a function of session and loss condition for individual rats within the Risk-Variability phases. The solid (Variability: One-Loss) and dashed lines (Variability: Two-Loss) represent the individual subject-fits from the best-fitting generalized linear mixed effects model. The rats in Group Equal-Risk are labeled as “E.” and the rats in Group Unequal-Risk are labeled as “U.”. 122

Figure 3.5 Proportion of choices for the risky outcome for Groups Equal-Risk and Unequal-Risk as a function of loss condition (One-Loss, Two-Loss) and the outcome of the previous choice. The dotted lines connecting the data points are intended to demonstrate the

relationship between risky choice behavior following differential outcomes rather than to imply a continuous relationship. C-2 = certain choice – 2 pellets; C-4 = certain choice – 4 pellets; R-0 = risky choice – 0 pellets; R-1 = risky choice – 1 pellet; R-11 = risky choice – 11 pellets. 123

Figure 3.6 Proportion of choices for the risky outcome for Groups Equal-Risk and Unequal-Risk as a function of loss condition (One-Loss, Two-Loss) and the outcome of the previous choice in the Risk-Omission (left) and Risk-Variability phases (right). The dotted lines connecting the data points are intended to demonstrate the relationship between risky choice behavior following differential outcomes rather than to imply a continuous relationship. C-2 = certain choice – 2 pellets; C-4 = certain choice – 4 pellets; R-0 = risky choice – 0 pellets; R-1 = risky choice – 1 pellet; R-11 = risky choice – 11 pellets. 124

Figure 3.7 Proportion of choices for the risky outcome as a function of the outcome of the previous choice and loss condition for individual rats in the Risk-Omission phases. The dotted (Omission: One-Loss) and dashed lines (Omission: Two-Loss) represent the individual subject-fits from the best-fitting generalized linear mixed effects model. The rats in Group Equal-Risk are labeled as “E.” and the rats in Group Unequal-Risk are labeled as “U.”. 125

Figure 3.8 Proportion of choices for the risky outcome as a function of the outcome of the previous choice and loss condition for individual rats in the Risk-Variability phases. The dotted (Variability: One-Loss) and dashed lines (Variability: Two-Loss) represent the individual subject-fits from the best-fitting generalized linear mixed effects model. The rats in Group Equal-Risk are labeled as “E.” and the rats in Group Unequal-Risk are labeled as “U.”. 126

Figure 3.9 Regression coefficients from the mixed-effects analysis evaluating the effect of the previous 10 outcomes on subsequent risky choice. E = Group Equal-Risk; U = Group Unequal-Risk; Hyp. = hyperbolic model fit; Exp. = exponential model fit. 127

Figure 3.10 Normalized regression coefficients from the mixed-effects analysis evaluating the effect of the previous 10 choices on subsequent risky choice. E = Group Equal-Risk; U = Group Unequal-Risk; Hyp. = hyperbolic model fit; Exp. = exponential model fit. 128

Figure 3.11 Parameter recovery for the Simple, Asymmetric, Valence-Attentive, and Weighted-Reference-Point models in Experiment 2. The thick dark lines are the unit diagonals. 129

Figure 3.12 Akaike Information Criterion (AIC) for the Asymmetric Reinforcement Learning (RL) model plotted against the AICs of the Simple and Valence-Attentive RL models. Each data point represents the fit to an individual rat’s data. Model fits that did not converge were omitted. 130

Figure 3.13 Choice data (Data) and the predicted choice data from the Asymmetric Reinforcement Learning (RL) model (Model) for each rat, ordered from top to bottom by how well the model fit the data in terms of omega-squared (ω^2). “Min” and “Max” refer to the data with the lowest and highest ω^2 , respectively. “25%ile”, “Mdn”, and “75%ile” refer to the 25th, 50th, and 75th percentile of ω^2 , respectively. The alternating gray and white boxes within each panel represent different phases and sub-phases (see Table 3.1). The differential alignments of these phase markers reflect individual differences in trials completed within each phase. 131

Figure 3.14 Mean risky choice plotted against the log-transformed α_G (left), α_L (center), and β parameters (right) as derived from the fitting of the Asymmetric Reinforcement Learning (RL) model. Each data point represents an individual rat. The goodness-of-fit index of the best fitting regression line for the overall model is shown (R^2). 132

Figure 3.15 Win-Stay Index plotted against the log-transformed α_G (left panels), α_L (middle panels), and α_L to α_G ratio (right panels) as derived from the fitting of the Asymmetric Reinforcement Learning (RL) model. Each data point represents an individual rat. The rats in Group Unequal-Risk and Group Equal-Risk are distinguished. 133

Figure 3.16 Loss Sensitivity Index from the Risk-Omission phases plotted against the log-transformed α_G (left panels), α_L (middle panels), and α_L to α_G ratio (right panels) as derived from the fitting of the Asymmetric Reinforcement Learning (RL) model. Each data point represents an individual rat. The rats in Group Unequal-Risk and Group Equal-Risk are distinguished. 134

Figure 3.17 Loss Sensitivity Index from the Risk-Variability phases plotted against the log-transformed α_G (left panels), α_L (middle panels), and α_L to α_G ratio (right panels) as derived from the fitting of the Asymmetric Reinforcement Learning (RL) model. Each data point represents an individual rat. The rats in Group Unequal-Risk and Group Equal-Risk are distinguished. 135

List of Tables

Table 2.1 Probabilities of the risky-zero (R-0), R-1, and R-11 outcomes in the P[0] and P[1] conditions in each phase of Experiment 1 for Groups Extra-Feedback and Normal-Feedback. Half of each group experienced one of two probability conditions across the two phases of Experiment 1.	57
Table 2.2 Mixed-effects model output for the molar analyses of Experiment 1.	58
Table 2.3 Mixed-effects model output for the molecular analyses of Experiment 1 evaluating the effect of the previous outcome on subsequent choice behavior. C-4 = certain choice – 4-pellet outcome; R-0 = risky choice – 0-pellet outcome; R-1 = risky choice – 1-pellet outcome; R-11 = risky choice – 11-pellet outcome	59
Table 2.4 Mixed-effects model output for the molecular analyses of Experiment 1 evaluating the effect of the previous 9 outcomes on subsequent risky choice behavior.	61
Table 2.5 Mixed-effects model output for the molecular analyses of Experiment 1 evaluating the effect of the previous 10 choices on subsequent risky choice behavior.	62
Table 2.6 Mixed-effects model output for the goal-tracking analysis of Experiment 1 evaluating the effect of phase, feedback group, and risky outcome magnitude on goal-tracking behavior following the onset of risky outcome delivery.....	63
Table 2.7 Summary statistics of the best-fitting parameter estimates and goodness-of-fit indices of the Asymmetric Reinforcement Learning (RL) model. AIC = Akaike Information Criterion; ω^2 = omega-squared.	64
Table 3.1 Probability of the risky-zero (R-0), R-1, and R-11 outcomes and the expected value of the risky choice in One-Loss and Two-Loss sub-phases of the Risk-Omission and Risk-Variability phases of the risky choice task of Experiment 2 for Groups Equal-Risk and Unequal-Risk.	112
Table 3.2 Mixed-effects model output for the molar analyses of Experiment 2.	113
Table 3.3 Mixed-effects model output for the molecular analyses of Experiment 2 evaluating the effect of the previous outcome on subsequent choice behavior. C-4 = certain choice – 4-pellet outcome; R-0 = risky choice – 0-pellet outcome; R-1 = risky choice – 1-pellet outcome; R-11 = risky choice – 11-pellet outcome	114

Table 3.4 Mixed-effects model output for the molecular analyses of Experiment 2 evaluating the effect of the previous 10 outcomes on subsequent risky choice behavior.	116
Table 3.5 Mixed-effects model output for the molecular analyses of Experiment 2 evaluating the effect of the previous 10 choices on subsequent risky choice behavior.	117
Table 3.6 Summary statistics of the best-fitting parameter estimates and goodness-of-fit indices of the Asymmetric Reinforcement Learning (RL) model. AIC = Akaike Information Criterion; ω^2 = omega-squared.	118

Acknowledgements

First and foremost, I would like to thank Dr. Kimberly Kirkpatrick, for her guidance, advice, patience, and leadership as an advisor and as a mentor. I must also recognize other current and former members of Dr. Kirkpatrick's Reward, Timing, & Decision (RTD) Laboratory, such as Dr. Jen Peterson, Catherine Hill, Sarah Stuebing, Christian Davis, Jeremy Lott, Carrie Turpen, Sydney Edmisten, Jesseca Pirkle, and Ashton Triplett, for their help with animal care and experimentation. Additionally, my committee members, Drs. Mary Cain, Michael Young, and Elliot Ludvig, for their feedback on my proposal. Emily Reinhardt, for her constant support and encouragement. Last, but not least, my 48 rats, for their data.

Chapter 1 - Introduction

Gains and Losses

A decision-maker will often face the choice between a guaranteed safer outcome (e.g., saving one's money) and a larger yet riskier payoff (e.g., gambling). In daily life, commuters have the option of taking the same roads to work or perusing an alternative route that may result in either an early workplace arrival or a tardiness-inducing traffic jam. Those given to addiction must weigh abstinence against the decision to risk their long-term health in order to achieve the temporary benefit of substance-induced euphoria. Accordingly, risky decision making has become a critical research topic in psychological science, neuroeconomics, behavioral ecology, and judgment and decision making (Doya, 2008; Heilbrunner & Hayden, 2013; Levin et al., 2012; Madden, Ewan, & Lagorio, 2007; Mazur, 2004; Myerson, Green, Hanson, Holt, & Estle, 2003; H. Peters, Hunt, & Harper, 2010; Platt & Huettel, 2008; Potenza, 2009; Rushworth & Behrens, 2008; Weatherly & Derenne, 2007; Weber, Shafir, & Blais, 2004; Winstanley, 2011). Traditional analyses of risky decision making have employed tasks that involve a tradeoff between reward certainty and reward magnitude (e.g., Rachlin, Raineri, & Cross, 1991). A person may be offered a choice between a guaranteed \$10 and a 50% chance of receiving \$40, whereas a rat may be offered a choice between a guaranteed one food pellet and a 50% chance of receiving four food pellets. Theoretically, individuals of all species should learn to optimize the tradeoff between reward magnitude and reward certainty, ultimately maximizing the amount of reward received. However, individuals tend to behave suboptimally in many choice paradigms (e.g., Marshall & Kirkpatrick, 2015; Stagner & Zentall, 2010; Zentall & Stagner, 2010). Indeed, individual differences in risky-choice behavior are related to gambling (Holt, Green, & Myerson, 2003), cigarette smoking (Reynolds, Richards, Horn, & Karraker, 2004), and percent body fat

(Rasmussen, Lawyer, & Reilly, 2010). As too much or too little risky-choice behavior can inhibit reward maximization, these individuals may exhibit relative deficits in the ability to weigh reward magnitude against reward certainty. Accordingly, the prevalence of risky (and often maladaptive) behaviors (Compton, Thomas, Stinson, & Grant, 2007; Lorains, Cowlshaw, & Thomas, 2011; Shaffer, Hall, & Vander Bilt, 1999; Shaffer & Korn, 2002) has encouraged research aimed at identifying the psychological and neurobiological mechanisms of individual differences in risky choice (e.g., Potenza, 2009).

An inherent component of decision making in risky environments is the potential to receive both gains and losses from the corresponding decisions. Theoretically, outcomes are characterized as gains and losses relative to an internal expectancy criterion known as a reference point (Kahneman & Tversky, 1979; Wang & Johnson, 2012). Losses refer to outcomes of a smaller magnitude relative to the amount that was expected (i.e., the reference point), thereby producing a negatively-signed deviation between the outcome that was expected and the one received (i.e., a prediction error). Alternatively, gains reflect outcomes of a larger magnitude than the expected amount, producing a positively-signed prediction error. Prediction error mechanisms, primarily localized to the brain's midbrain dopamine system (e.g., Schultz, Dayan, & Montague, 1997), have been assumed to reflect a teaching signal for learning the value of a particular action (see Hollerman & Schultz, 1998), thereby resulting in adjustments in subsequent behavior to optimize the tradeoff between energy intake and energy expenditure (i.e., maximizing gains while minimizing losses). Assuming that gains and losses are encoded relative to a subjective reference point, loss-induced negatively-signed prediction errors result in reductions in the subjective value attributed to the corresponding action (i.e., the choice). This decrement in value relative to the values of alternative choices in the environment is assumed to

reduce the likelihood of repeating the previous behavior (i.e., lose-shift), while gain-induced positively-signed prediction errors, which increase the choice's value, are assumed to increase the likelihood of repeating the same behavior (i.e., win-stay). Indeed, previous risky losses tend to discourage subsequent risky-choice behavior while previous risky gains encourage subsequent risky choice in both humans and non-human animals (Evenden & Robbins, 1984; Larkin, Jenni, & Floresco, 2016; Marshall & Kirkpatrick, 2013, 2015; Stopper & Floresco, 2011; Thaler & Johnson, 1990; also see Brevers et al., 2015). Accordingly, win-stay/lose-shift behavior has become well-encompassed within multiple theories of learning and valuation that incorporate prediction error mechanisms (e.g., Bush & Mosteller, 1951; Glimcher, 2011; Rescorla & Wagner, 1972; Sutton & Barto, 1998).

While win-stay/lose-shift behavior may be assumed to be a theoretically optimal strategy in ecologically valid environments (see Wilke & Barrett, 2009), some individuals exhibit the potentially suboptimal tendency to chase losses with increasingly risky behaviors, possibly to compensate for a previous loss or accumulation of losses (i.e., lose-stay; see Linnet, Røjskjær, Nygaard, & Maher, 2006). For example, after an individual loses a wager on a horse race, s/he may be more likely to bet larger amounts on subsequent races to make up for the previous loss (see McGlothlin, 1956). Additionally, after losses, individuals tend to wager significantly more than they had originally planned, while actual versus planned wagering does not considerably differ after gains (Andrade & Iyer, 2009). Indeed, “loss-chasing” behavior is a key characteristic of pathological gambling (e.g., Breen & Zuckerman, 1999; Lesieur, 1979). As losses have been suggested to drive risky decisions more than gains do (see Anselme, 2013; Anselme & Robinson, 2013) and as both human and non-human animals have been suggested to be more sensitive to losses than to gains (Kahneman & Tversky, 1979; Marshall & Kirkpatrick, 2015), one key factor

that may dictate individual differences in loss-chasing and risky choice behaviors is the subjective sensitivity to losses relative to gains. Individuals who are relatively *more* sensitive to differential losses may exhibit *less* loss-chasing behavior compared to individuals who are less sensitive. Indeed, more severe gamblers exhibit greater sensitivity to gains versus losses (Brevers, Koritzky, Bechara, & Noël, 2014). Also, individuals with other behavioral addictions (i.e., internet addiction disorder) exhibit a reduced sensitivity to losses compared to control subjects (Dong, Hu, & Lin, 2013), and severe gamblers are not as sensitive to differential losses as less severe gamblers (see Kreussel et al., 2013). Thus, as pathological gamblers (PGs), relative to normal subjects, engage in several maladaptive behaviors that produce a lower quality of life (Black, Shaw, McCormick, & Allen, 2013), further elucidation of loss-chasing mechanisms and the differential processing of gains versus losses will permit greater understanding of the mechanisms that drive some individuals to continue making risky choices despite the experience of repeated adverse consequences (see Rachlin, 1990).

All Gains and Losses are Not Created Equally

Behavior in response to gains and losses provides critical insight into the mechanisms governing individual differences in risky choice and loss-chasing behaviors. However, the dichotomization of outcomes into either gains or losses fails to capture the continuum of behavioral responses to differential gains and losses. Several analyses in the gambling literature have investigated the impact of risky outcomes that, while objectively dichotomized as losses, may be more ambiguous in terms of their effects on behavior. For example, a loss-disguised-as-a-win (LDW) is an outcome that is less than the amount wagered but delivered with the same positive-feedback stimuli as a win (e.g., wagering five dollars and the subsequent three-dollar return is accompanied by sensory stimuli that also accompany the returns that exceed the amount

wagered, otherwise known as wins). When exposed to wins, losses, and LDWs, individuals' heart-rate changes were more similar after losses and LDWs while individuals' skin-conductance responses were more similar after wins and LDWs (M. J. Dixon, Harrigan, Sandhu, Collins, & Fugelsang, 2010), reflecting a complex physiological response to LDWs (i.e., small losses). Thus, the analysis of LDWs is critical, as such outcomes make up a considerable proportion of the possible outcomes of a gamble (M. J. Dixon et al., 2010), and could be subjectively encoded as gains or losses depending on the individual's reference point (see Linnet et al., 2006). For example, one possible reference point is the individual's wager (e.g., winning more than \$5 from a \$5 bet equates to a gain). However, that same individual may be relatively insensitive to the wagered amount or have a poor memory for the wagered amount, such that this amount is essentially an ineffective reference point. Furthermore, the ultimate receipt/outcome from the wager may be more salient than the wagered amount, thus discouraging the use of the wagered amount as the reference point. Accordingly, such subjective deficits may elicit use of differential reference points (e.g., current holdings before versus after the wager), which may then alter subsequent decision making by changing how the wagered outcome is encoded (i.e., gain or loss).

Collectively, these results suggest that individuals with heightened risk-taking tendencies may exhibit impairments in loss processing mechanisms (see Worhunsky, Malison, Rogers, & Potenza, 2014). Indeed, given the prevalence of losses in gambling environments, greater risk-taking in PGs may be driven by loss processing deficits. Indeed, losses involve more than the actual loss of money (e.g., wagering five dollars and losing all of it), but also the total net deficit (e.g., wagering five dollars and receiving four dollars in return; LDWs). Potential impairments in loss processing (i.e., deficits in processing loss magnitude) relative to gain processing may

explain individual differences in response to more ambiguous outcomes (e.g., LDWs). While reward omission following a gamble may be viewed as aversive and punishing, these outcomes do not effectively inhibit gambling behavior (see Horsley, Osborne, Norman, & Wells, 2012), possibly due to the absence of tangible feedback in reward omission. Accordingly, the analysis of behavior in response to losses of differential magnitudes, rather than only following reward omission (Rogers, Wong, McKinnon, & Winstanley, 2013; Tremblay et al., 2014), may provide the best insight into the core mechanisms dictating individual motivations to persist in making risky decisions despite the experience of negative consequences. Furthermore, while research has analyzed various pharmacological effects on loss-chasing in non-human animals (Rogers et al., 2013), our understanding of the core psychological processes governing individual differences in differential loss-chasing in non-human animals remains in its infancy. Further elucidation of loss-chasing mechanisms and the neurobiology of loss-chasing and reward processing would considerably solidify the future development of neurocognitive and/or pharmacological interventions aimed at alleviating subjective deficits in such behavior.

Mechanisms of Loss-Chasing

Previous reports have considered potential correlates of loss-chasing behavior. Breen and Zuckerman (1999) reported that loss-chasers were significantly more impulsive, based on self-reports of impulsivity. Indeed, both impulsivity and risk-taking are associated with individual differences in reward processing mechanisms (e.g., Kirkpatrick, Marshall, & Smith, 2015). Campbell-Meiklejohn, Woolrich, Passingham, and Rogers (2008) showed that loss-chasing decisions were associated with elevated activity in the ventromedial prefrontal cortex, which has been implicated in processing reward value (e.g., J. Peters & Büchel, 2010). It has also been suggested that deficits in the mesolimbic dopamine pathway may result in heightened loss-

chasing behavior (Campbell-Meiklejohn et al., 2008). Campbell-Meiklejohn et al. (2012) reported that while gambling behavior was reduced with increasing stakes of the gamble (i.e., the amount potentially won or lost), participants who were administered methylphenidate did not show the corresponding decrease in gambling with increasing stakes, suggesting that increasing dopamine levels via methylphenidate administration resulted in elevated risk-taking behavior. Indeed, administration of a dopamine D₂ receptor antagonist reduced loss-chasing in rats (Rogers et al., 2013). Given the well-established relationship between dopaminergic function and reward processing (e.g., Schultz, 2007; Schultz et al., 1997), the dopaminergic involvement in loss-chasing behavior suggests that individual differences in loss-chasing and risky decision making may be primarily driven by individual differences in reward processing/valuation mechanisms.

The concepts related to and the definitions of “value” are complex and multifaceted (O’Doherty, 2014; J. Peters & Büchel, 2010), but reward value in the present context is regarded as the value attributed to a choice as determined by repeated experiences with the corresponding outcomes (i.e., prediction of future reward expectation; Sutton & Barto, 1998). Accordingly, reward value is dynamic, subject to learning, environmental changes, and time since the corresponding outcomes were experienced (e.g., Bayer & Glimcher, 2005; Devenport, Hill, Wilson, & Ogden, 1997; Dunlap & Stephens, 2012; Sutton & Barto, 1998). For example, to a naïve foraging animal, the absence of experience with a foraging patch may theoretically predict that the patch has no subjective value. Subsequent exposures (and lack thereof) to food items in the patch cause adjustments in patch value. Reward presentation that exceeds expectations elicits a positively-signed prediction error, theoretically increasing the subjective value of the patch, while smaller amounts of food contrary to expectation elicit negatively-signed prediction errors that decrease subjective patch value.

Given two different foraging animals with identical experiences with the same foraging patch, one assumption is that these animals will attribute identical values to this patch. However, as prediction errors serve as teaching signals to update expectations (e.g., Hollerman & Schultz, 1998), an addendum to this assumption is that these animals will only attribute identical values to this patch if they learn about this patch and update their expected patch values identically; value estimates will be the same if the animals learn about the patches in the same way. This implies that individuals with discrepant rates of value-updating/learning will value the patch differently despite identical experiences with the same patch. Accordingly, the choice between two patches (or the choice between a certain and a risky choice) would therefore depend on how often and how much the value of a patch/choice is updated as a function of experience. Indeed, previous research has described how individual differences in risky choice behavior may be accounted for by those in learning rates (March, 1996; Marshall & Kirkpatrick, 2013). Given the necessity of learning rate to valuation computations and choice behavior (e.g., Glimcher, 2011; Sutton & Barto, 1998), as well as the aforementioned hypothesis that loss-chasing may reflect subjective deficits in processing reward value, individual differences in loss-chasing behavior may in fact be driven by individual differences in value updating (i.e., learning the value of a choice).

The proposed mechanism by which a choice's value is learned given repeated outcomes of the choice is dependent on the theory of learning/valuation. Several theories employ a linear-operator mechanism to update expectation (e.g., Bush & Mosteller, 1951), such that updated estimates of subjective value are a function of the previous value estimates and the prediction error between the previous estimate and the most recent outcome delivered for having made the choice corresponding to that estimate. Prediction error is traditionally scaled by a learning-rate parameter (e.g., α), in which steeper learning (i.e., larger values of α up to 1) reflects more rapid

adjustments in value. Assuming that α can be treated as a free parameter (e.g., Barraclough, Conroy, & Lee, 2004), an individual's learning rate can be estimated by fitting the linear-operator model to individual subject data. Indeed, faster learning rates (i.e., larger α values) have been suggested to reflect less overall risk-taking (see March, 1996; Marshall & Kirkpatrick, 2013), suggesting that individuals prone to loss-chasing may not update value given the previous outcome as readily as individuals who learn from the environment more quickly.

A learning-rate hypothesis of loss-chasing behavior is intriguing, as the derivation of an individual's learning rate could predict one's corresponding propensity to engage in suboptimal and maladaptive behaviors (i.e., pathological gambling via heightened loss-chasing). Individuals who fail to readily incorporate a negatively-signed prediction error (i.e., a loss) into their subjective computations of value may continue making risky choices as the corresponding estimation of value has not considerably decreased (see Clark, Liu, et al., 2013). Consecutive risky losses may do little to discourage an individual in making further risky decisions (cf., Rachlin, 1990). However, assuming a constant learning rate regardless of the sign of the prediction error, this would also suggest that large gains would not be readily incorporated into subjective estimations of value, so that the value of the choice would not increase to the extent that would reasonably produce subsequent risky choices (i.e., win-stay behavior). In contrast to the latter, previous research has suggested that larger rewards exert a considerable influence over risky choice behavior (Hayden, Heilbronner, Nair, & Platt, 2008; Hayden & Platt, 2007; but see Marshall & Kirkpatrick, 2015). If individual differences in learning can account for individual differences in loss-chasing, then this may not be achievable assuming identical gain- and loss-based learning rates.

While the linear-operator mechanism described above has served as a critical component of various models of classical conditioning, machine learning, interval timing, and decision making (e.g., Bush & Mosteller, 1951; Glimcher, 2011; Guilhardi, Yi, & Church, 2007; Kirkpatrick, 2002; Rescorla & Wagner, 1972; Sutton & Barto, 1998), subsequent adjustments to the model have been employed to better understand the corresponding mechanisms of such phenomena. For example, recent research has suggested that prediction error encoding of losses may be distinct from gains (see Barraclough et al., 2004; Bayer & Glimcher, 2005; Frank, Moustafa, Haughey, Curran, & Hutchison, 2007; Niv, Edlund, Dayan, & O'Doherty, 2012). Individuals who gamble despite gains and losses may exhibit asymmetric learning rates from these outcomes. Loss-chasers may rapidly incorporate gains via positively-signed prediction errors into their expected values of gambling outcomes, producing substantial increases in expected value. However, these same individuals may less readily incorporate losses via negatively-signed prediction errors, such that the corresponding expected values do not reliably decrease, leading to repeated choices with the same risky outcome despite previous losses (see Clark, Liu, et al., 2013). Indeed, in accordance with prediction error encoding in midbrain dopamine neurons (Schultz et al., 1997), elevated or reduced dopaminergic activity induced by the administration or depletion of dopamine precursors (i.e., levodopa, tyrosine, and phenylalanine) may separately impact appetitive/gain- and aversive/loss-based learning rates, respectively (see Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006; Robinson, Standing, DeVito, Cools, & Sahakian, 2010). Bayer and Glimcher (2005) suggested that the proposed neurobiological correlates of prediction error (i.e., the midbrain dopamine system) may be more attuned to encoding positively- than negatively-signed prediction error. Indeed, PGs show greater increases in dopaminergic activity following losses than control participants (Linnet,

Peterson, Doudet, Gjedde, & Møller, 2010). Alternatively, relative to positively-signed prediction errors, negatively-signed prediction errors may be processed by distinct brain regions (see, e.g., Hong, Jhou, Smith, Saleem, & Hikosaka, 2011; Matsumoto & Hikosaka, 2007) or moderated by different neurotransmitters (see den Ouden et al., 2013). Thus, one critical determinant of enhanced risk-taking via loss-chasing may involve asymmetric gain- and loss-based learning rates, such that individuals who are more likely to chase losses are those that exhibit greater adjustments in subjective value following gains than following losses.

Alternatively, loss-chasers may exhibit such behavior given differences in multiple other mechanisms. For example, loss-chasers may indeed incorporate losses, or a series of losses, into valuation computations, but only after receipt of a gain. This hypothesis is reminiscent of the string theory of gambling (Rachlin, 1990; Rachlin, Safin, Arfer, & Yen, 2015), such that gambling sequences are bounded by gains and that the value of the string is only computed after a gain is received. This explanation would suggest that individuals treat gain-to-gain cycles as individual gambling episodes, and that previous losses are less readily integrated into estimates of value given the diminished recency of the loss. Here, only the string (gain-to-gain cycle), not the individual gamble, has value. Interestingly, loss-chasing is a direct prediction of string theory, as riskier choices with higher potential payoffs provide a greater likelihood of increasing the value of the string following consecutive losses (Rachlin, 1990; Rachlin et al., 2015).

An additional hypothesis is that loss-chasers may be more attentive to gains relative to losses; indeed, severe gamblers exhibit greater sensitivity to gains than they do to losses (Brevers et al., 2014). Furthermore, loss-chasers may exhibit differential memories for previous gains and losses; that is, loss-chasers may continue to make risky choices as their memories for gains surpasses those for losses. Indeed, stronger memories for extreme gains is correlated with

greater likelihoods to make risky decisions (Madan, Ludvig, & Spetch, 2014). Alternatively, loss-chasers may have a different reference point dichotomizing outcomes as gains and losses relative to non-loss-chasers (see Linnet et al., 2006). Thus, loss-chasing may be explained by individual differences in a multitude of mechanisms. Given the prevalence of maladaptive risk-taking behaviors that inherently involve potential gains and losses (e.g., substance abuse, pathological gambling, risky sexual behaviors, etc.), greater understanding of the relationship between loss-chasing and the mechanisms by which individuals incorporate losses and gains into estimates of value has critical implications for understanding subjective tendencies to persist in making risky choices despite the experience of negative consequences.

Chapter 2 - Experiment 1

By necessity, reward omission counters previous expectation, as reward can only be omitted when reward is expected. In these situations, reward omission is, by default, a loss. In accordance with reinforcement learning (RL), it is expected that estimates of value continue to decrease with more instances of reward omission, so that the corresponding action (i.e., choice) is less regularly exhibited in the future. However, while reward omission may eventually reduce risk-taking behaviors, repeated occurrences of reward omission may not effectively and efficiently do so (see Horsley et al., 2012). For example, probabilistic (i.e., partial) reinforcement discourages the cessation of the corresponding behavior when reward is no longer delivered (Domjan, 2010), ultimately producing problematic and maladaptive risky decision-making behaviors (e.g., slot-machine gambling, which operates via random-ratio schedules of reinforcement; see Crossman, 1983; Madden et al., 2007).

One possible explanation for the ineffectiveness of reward omission relates to both the nature of the corresponding feedback given the behavior that resulted in reward omission, as well as the structure of the environment. When a risky choice probabilistically results in 0 or 11 pellets, the delivery of either outcome provides information regarding both reward rate and the frequencies of each outcome. For example, if one of ten risky choices results in 11 pellets, then $p(0) = .90$ and $p(11) = .10$. Accordingly, $p(0)$ and $p(11)$ are anti-correlated; new information suggesting that $p(0)$ has increased may automatically imply that $p(11)$ has decreased, even though the trial-by-trial outcomes are typically independent. Win-stay/lose-shift behavior would be predicted by an individual's sensitivity to such an environment; reward omission would theoretically discourage staying behavior given the possibility that reward rate has decreased

while reward presentation would promote staying behavior given the possibility that reward rate has increased.

While win-stay/lose-shift behavior is mostly observed in risky choice paradigms in which the probabilities of reward omission and a larger reward are anti-correlated, many experiments have shown that individuals will continue to make risky choices following reward omission, such that a loss does not completely “shift” behavior but only decreases the rate of “stay” behavior (e.g., Larkin et al., 2016; Marshall & Kirkpatrick, 2013, 2015; St. Onge, Stopper, Zahm, & Floresco, 2012; Stopper & Floresco, 2011; Stopper, Khayambashi, & Floresco, 2013). That is, a loss is also followed by subsequent responding for the operandum that resulted in reward omission. Interestingly, in a non-choice context, Kello (1972) showed that reward omission coupled with other stimuli typically presented with reward (e.g., houselight offset) systematically reduced the post-omission elevations in behavior with each added stimulus (also see Mellon, Leak, Fairhurst, & Gibbon, 1995). Ultimately, in experimental contexts with anti-correlated probabilities of a larger reward and reward omission, the effects of reward omission on subsequent behavior may partially depend on the reception of explicit feedback from the corresponding decision.

This effect of anti-correlated reward probabilities may not hold when the probabilities of reward omission and a larger reward are correlated. Marshall and Kirkpatrick (2015) presented rats with the choice between a relatively more certain outcome (2 or 4 pellets) and a relatively more risky outcome (0, 1, or 11 pellets). In one condition, $p(1) = p(11)$ and the probability of reward omission was manipulated, such that $p(0)$ and $p(11)$ were anti-correlated (i.e., the P[0] condition). In a second condition, $p(0) = p(11)$, and the probability of the one-pellet reward was manipulated, such that $p(0)$ and $p(11)$ were correlated (i.e., the P[1] condition). Across

conditions, both the 0- and 1-pellet rewards were losses. In the P[0] condition, the rats demonstrated typical win-stay/lose-shift behavior, exhibiting a greater likelihood of making a risky choice following the 1- and 11-pellet outcomes than following the 0-pellet outcome. However, in the P[1] condition, the rats were more likely to make a risky choice following the 0-pellet outcome than following the 1-pellet outcome, in contrast to typical win-stay/lose-shift behavior. This latter pattern of behavior may reflect subjective responses to probabilistic feedback. In the P[1] condition, exposure to reward omission may result in the subjective prediction that $p(11)$ has increased, while exposure to the 1-pellet outcome would result in the opposite prediction, thereby decreasing subsequent risky choice behavior. Alternatively, rather than reward omission providing loss- and probability-based information of the frequency of 11-pellet reward delivery, reward omission may simply reflect the absence of information, such that the relatively high lose-shift behavior following reward omission is in fact a continuation of previous behavior given the absence of feedback to promote alternative actions. Thus, in addition to the effectiveness of the 1-pellet outcome as a loss in reducing staying behavior, the decrease in staying behavior following 1-pellet outcomes in the P[1] condition compared to that following 0-pellet outcomes may be explained in terms of the anti-correlation between $p(1)$ and $p(11)$ and/or the reception of explicit feedback given the loss. These possibilities have implications for how individuals encode and respond to the ambiguous differential losses (i.e., LDWs). LDWs may not just be a special case of loss (i.e., numerically greater than reward omission), but a unique psychological event driven by the differential integration of loss- and gain-based feedback, which may potentially distinguish these outcomes from the traditional continuum of differentially-sized gains and losses.

Accordingly, the goal of the present experiment was to determine how loss-chasing behavior is affected by the exposure to explicit feedback following differential losses. While it may be argued that reward omission and corresponding lever retraction in traditional risky choice paradigms serve as adequate feedback, the absence of explicit feedback from the corresponding choice may be psychologically distinct from the mechanisms of loss presentation in experiments of human risky decision making, such as the explicit removal of a resource (e.g., depleted monetary funds) or a small presentation of a resource that fails to meet expectations (i.e., LDWs). Accordingly, the goal of the present experiment was to determine whether rats' loss-chasing behavior was altered depending on whether the 0- and 1-pellet losses are accompanied by the explicit feedback that is typically only presented when 11-pellet outcomes are presented, similar to the presentation of win-related stimuli that accompany LDWs.

Method

Animals

Twenty-four experimentally-naive male Sprague-Dawley rats (Charles River; Kingston, NY) were used in the experiment. They arrived at the facility (Kansas State University; Manhattan, KS) at approximately 21 days of age. The rats were pair-housed in a red-illuminated colony room that is set to a reverse 12:12 hr light:dark schedule (lights off at approximately 7:30 am). The rats were tested during the dark phase. There was ad libitum access to water in the home cages and in the experimental chambers. The rats were maintained at approximately 85% of their projected ad libitum weight during the experiment, based on growth-curve charts obtained from the supplier. When supplementary feeding was required following an experimental session, the rats were fed in their home cages approximately 1 hr after being returned to the colony room (see Bacotti, 1976; Smethells, Fox, Andrews, & Reilly, 2012).

Apparatus

The experiment was conducted in 24 operant chambers (Med-Associates; St. Albans, VT) each housed within sound-attenuating, ventilated boxes ($74 \times 38 \times 60$ cm). Each operant chamber ($25 \times 30 \times 30$ cm) was equipped with a stainless steel grid floor, two stainless steel walls (front and back), and a transparent polycarbonate side wall, ceiling, and door. Two pellet dispensers (ENV-203), mounted on the outside of the operant chamber, were equipped to deliver 45-mg food pellets (Bio-Serv; Flemington, NJ) to a food cup (ENV-200R7) that was centered on the lower section of the front wall. For the rats in Group Extra-Feedback (see below), the tubing connecting one of the pellet dispensers to the food cup was disconnected and the corresponding pellet deliveries were rerouted to a receptacle outside of the operant chamber (Figure 2.1; see Freestone, MacInnis, & Church, 2013). Head entries into the food magazine were transduced by an infrared photobeam (ENV-254). Two retractable levers (ENV-112CM) were located on opposite sides of the food cup. The chamber was also equipped with a house light (ENV-215) that was centered at the top of the front wall, as well as two nosepoke keys (ENV-119M-1) that were located above the levers. Water was always available from a sipper tube that protruded through the back wall of the chamber. Experimental events were controlled and recorded with 2-ms resolution by the software program MED-PC IV (Tatham & Zurn, 1989).

Procedure

Magazine and lever-press training. The rats experienced a random-time 60-s schedule of food deliveries for magazine training, earning approximately 120 pellets in one 2-hr session. The rats then experienced lever-press training with a fixed ratio (FR) 1 schedule of reinforcement followed by a random ratio (RR) 3 schedule and then an RR 5; each of these schedules lasted until the rats earned 20 pellets on each lever. For 22 of the 24 rats, lever-press training lasted for

two sessions. For the other two rats, a third session of lever-press training was administered, which involved only the RR 3 and RR 5 schedules of reinforcement.

Risky choice task. Each pair of rats was randomly partitioned into one of two groups: Group Extra-Feedback and Group Normal-Feedback. Typically, food deliveries are separated by 0.2 s and accompanied by the sound and vibration of the pellet dispenser; thus, reward omission is not accompanied by external stimuli. The rats in Group Normal-Feedback experienced such an environment, in which the amount of exposure to external stimuli (i.e., sound and vibration of the pellet dispenser) was directly related to the amount of food delivered following each choice. Additionally, for Group Normal-Feedback, each operation of the pellet dispenser was accompanied by a 0.1-s illumination of the nosepoke key light above the corresponding lever, such that these rats experienced synchronized auditory, tactile, and visual stimuli veridically related to the number of pellets received. For Group Extra-Feedback, every risky outcome was also accompanied by multiple operations of the alternative pellet dispenser without the rats' receiving the pellets from this dispenser (i.e., food deliveries into an external receptacle; Figure 2.1) and flashing illuminations of the nosepoke key light above the risky outcome lever, such that all risky outcomes included explicit multimodal feedback (i.e., light, sound and vibration of the pellet dispenser). As in Group Normal-Feedback, every 0.1-s flash of the nosepoke key light presented to Group Extra-Feedback was time-locked to the operation of the pellet dispenser(s).

Both groups experienced the same risky choice task. Certain choices always resulted in either 2 or 4 pellets ($p_s = .50$; certain-two, C-2, and certain-four, C-4), and risky choices resulted in either no food being delivered (risky-zero, R-0), 1 food pellet (R-1), or 11 food pellets (R-11). Both choices involved an element of risk (i.e., outcome variability), as outcome variability provides greater ecological validity (see Searcy & Pietras, 2011), and is thus an essential

component of risky choice tasks (see Marshall & Kirkpatrick, 2013). Moreover, rats' preferences for riskiness/uncertainty over certainty may override true preference for an outcome (cf., Battalio, Kagel, & MacDonald, 1985). Thus, the certain outcome referred to choices in which food was always delivered, while the risky outcome referred to choices in which food was probabilistically delivered.

Each session began with 8 forced-choice trials followed by a maximum of 100 free-choice trials (Marshall & Kirkpatrick, 2013, 2015). On forced-choice trials, one lever was inserted into the chamber. Each lever corresponded to one of two choices (i.e., certain and risky), and lever assignments were counterbalanced within pairs of rats. When the lever was pressed on forced-choice trials, a fixed interval 20-s schedule began; the first lever press after 20 s resulted in lever retraction and food delivery. If the lever corresponded to the certain outcome, then the C-2 or C-4 outcome was delivered ($p = .50$). If the lever corresponded to the risky outcome, then the R-11 outcome was delivered ($p = .50$) or either the R-0 or R-1 outcome was delivered ($p = .50$), the latter of which depending on the condition, as specified below. Each of these magnitudes for the certain and risky choices was presented twice in the forced-choice trials in a random order. A 10-s inter-trial interval (ITI) intervened between successive trials.

For Group Extra-Feedback, all risky outcome food deliveries were accompanied by 11 operations of the alternative pellet dispenser and 11 0.1-s flashes of the nosepoke key light, produced in synchrony with the alternative pellet dispenser. The ITI began following the eleventh operation of the alternative pellet dispenser and eleventh flash of the nosepoke key light. Accordingly, for Group Extra-Feedback, the R-0 and R-1 outcomes were designed to resemble exposure to LDWs. Free-choice trials were identical to forced-choice trials with the following exceptions: (1) both levers were inserted into the chamber; (2) a choice on one of the

levers caused the other lever to retract; and, (3) all risky outcomes (R-0, R-1, and R-11) were possible outcomes of risky choices.

The task was divided into two conditions (P[0] and P[1]), in which either $p(\text{R-0})$ or $p(\text{R-1})$ was manipulated, respectively. Half of each group experienced the P[0] condition first, while the other half experienced the P[1] condition first (Table 2.1). In Phase 1 of the P[0] condition, $p(\text{R-0})$ was .10 (i.e., $p(\text{R-1}) = p(\text{R-11}) = .45$). In Phases 2-4, $p(\text{R-0})$ was equal to .90, .50, and .10, respectively, such that $p(\text{R-1})$ and $p(\text{R-11})$ were equal to .05, .25, and .45, respectively. For the P[1] condition, $p(\text{R-1})$ was equal to .10, .90, .50, and .10 in Phases 1-4, respectively, such that $p(\text{R-0})$ and $p(\text{R-11})$ were equal to .45, .05, .25, and .45 in Phases 1-4, respectively.

Following exposure to either the P[0] or P[1] conditions, the rats then experienced the P[1] and P[0] conditions, respectively. Here, the P[0] and P[1] conditions mirrored Phases 2-4 of the prior condition, such that secondary exposure to P[0] involved $p(\text{R-0})$ equal to .90, .50, and .10 in Phases 1-3, respectively; secondary exposure to P[1] involved $p(\text{R-1})$ equal to .90, .50, and .10 in Phases 1-3 respectively (see Table 2.1). In the P[0] condition, risky forced-choice trials involved the R-1 and R-11 outcomes; in the P[1] condition, risky forced-choice trials involved the R-0 and R-11 outcomes. Each of the differential feedback conditions lasted for 10 sessions. Each session lasted until all free-choice trials were completed or for approximately 2 hr.

Data analysis

With the following exceptions, all rats' data from all sessions were used in analyses (unless otherwise specified below). Due to equipment error, 10 sessions across multiple individual rats were removed from all data analyses (i.e., three sessions for one rat, two sessions for a second rat, one session for a third rat, and four sessions for a fourth rat). Cursory analyses

of mean choice behavior across rats before and after each of these sessions suggested that these equipment errors did not considerably impact subsequent choice behavior.

Choice behavior

Molar analysis. All summary measures were obtained from the raw data using MATLAB (The MathWorks; Natick, MA). Analyses were conducted at the molar and molecular levels. The dependent measure of the molar analyses was choice for the risky or certain outcome. Generalized linear mixed-effects models (Wright & London, 2009) employing binomial logistic regression with a logit link function were conducted with choice as the criterion (certain = 0, risky = 1). Generalized linear mixed-effects models are comparable to repeated-measures logistic regression analyses, but allow for parameter estimation as a function of manipulation condition (e.g., LL magnitude) and the individual subject (Young, Webb, Rung, & Jacobs, 2013). Thus, such models permit inclusion of both fixed and random effects, respectively. For the molar analyses, potential fixed effects included feedback group (Extra-Feedback/Normal-Feedback), probability condition (P[0]/P[1]), probability of the manipulated outcome (i.e., P[0]/P[1] = .10, .50, .90), session, and their interactions. The factors that were potential random effects included the aforementioned fixed effects except for the main effect of and interactions that included feedback group.

Model fitting occurred in two stages. Analyses first determined the model with the best fitting random-effects structure, and then the model with the best fitting fixed-effects structure that incorporated the best fitting random-effects structure. Given the current design, all potential random effects were also potential fixed effects, so the factor(s) within the best-fitting random effects structure were automatically included as fixed-effects (Young et al., 2013). Factors that did not vary as a function of subject (i.e., between-subjects factors) could be entered into the

model in this second stage. Model selection involved determining the model that minimized the Akaike information criterion (AIC; Akaike, 1973), in which the doubled negative log likelihood of the regression model is penalized by twice the number of estimated parameters. The AIC involves determining the best approximating model of the data, such that the model with the minimum AIC reflects the best fit but with a penalty for the expected improvement in fit due to added parameters (see Burnham & Anderson, 1998). Continuous predictors were mean-centered to reduce multicollinearity, and categorical predictors were effect-coded (i.e., codes summed to 0).

Molecular analysis. The first molecular analysis was identical to the molar analysis except that the identity of the previous outcome was included in the analysis as a potential fixed and/or random effect. Given the previously-reported non-monotonic relationship between previous outcome and subsequent risky choice (Marshall & Kirkpatrick, 2015), the factor corresponding to previous outcomes was set as a categorical predictor in this molecular analysis. Because this analysis relied on the effect of the previous outcome, the first trial of each session was excluded from analysis. Lastly, as the purpose of this analysis concerned the effects of the feedback, probability, and phase manipulations on trial-by-trial risky choice behavior, the predictor of session was not included in these analyses.

The second molecular analysis was similar to the modelling approach employed by Lau and Glimcher (2005), in which choice on trial N was modelled as a function of outcome and choice history. The model initially took the following form (see Lau & Glimcher, 2005):

$$\gamma + \sum_{i=1}^{10} \alpha_i (O_{Risky, N-i} - O_{Certain, N-i}) + \sum_{i=1}^{10} \beta_i (C_{Risky, N-i} - C_{Certain, N-i}), \quad (1)$$

in which γ was the model's intercept, N was the current trial, i was trial lag, α_i was the coefficient estimating the extent to which the i th previous outcome influenced choice on trial N ,

β_i was the coefficient estimating the extent to which the i th previous choice influenced choice on trial N , $O_{Risky,N-i}$ was the reward received on trial $N-i$ for making a risky choice, $O_{Certain,N-i}$ was the reward received on trial $N-i$ for making a certain choice, $C_{Risky,N-i}$ was a binary predictor indicating that a risky choice was made on trial $N-i$, and $C_{Certain,N-i}$ was a binary predictor indicating that a certain choice was made on trial $N-i$. The outcomes and choices of the previous ten trials were analyzed. For practical purposes, the model's predictors were a vector of previous outcomes and choices rather than a difference score (Eq. 1). Accordingly, the certain outcomes (C-2, C-4) were coded as -2 and -4, respectively, and certain choices were coded as -1.

Due to highly significant correlations between the i th previous outcomes and i th previous choices of similar lags, $r_s > .76$, $p_s < .001$ (e.g., the most recent outcome and most recent choice), the previous outcome and previous choice terms in Equation 1 were separated into two sets of models. Accordingly, rather than a single intercept (γ) for the combined equation, each of the two models incorporated an independent intercept. For both analyses, the most recent outcome (or choice) was entered into the random effects first, followed by the second most recent outcome (or choice), and so on. The predictors in the best-fitting random effects model were automatically included as fixed effects. As the purpose of this analysis was to determine the influence of outcome and choice histories on subsequent choice in the absence of other predictors, analyses only included the previous outcomes or choices. Model selection involved determining the model that minimized the AIC. As the AIC is dependent on sample size, the same data were used for all models. Thus, the first ten trials of every session were excluded from analysis so that there were equivalent numbers of all lagged outcomes and choices.

The second stage of this molecular analysis involved determining the functional form of the change in regression weights as a function of outcome and choice history. Each rat's

individual regression weights were computed from the results of the previous analysis (i.e., fixed effect coefficients + random effect deviations for each rat). These regression weights were subjected to a nonlinear mixed-effects model in R (Pinheiro, Bates, DebRoy, & Sarkar, 2016) to determine the extent to which previous outcomes and choices affected subsequent choice across groups, and to determine whether the decaying effect of previous choices and outcomes was better accounted for by a model that assumed exponential decay (e.g., Glimcher, 2011) or hyperbolic decay (e.g., Devenport et al., 1997). The following models were fit within the nonlinear mixed effects analyses for the exponential (Eq. 2) and hyperbolic functions (Eq. 3):

$$A * (1 - k)^{E-1}, \quad (2)$$

$$\frac{A}{1 + k * (E - 1)}, \quad (3)$$

in which A is the function's intercept, k is the decay rate of the influence of previous outcomes (or choices), and E is the lagged event (outcome or choice). Due to model convergence issues for the choice history analysis when both free parameters, A and k , were included, the regression coefficients in the choice history analysis were normalized for each rat relative to that rat's maximum regression coefficient in the choice history analysis. Thus, for the choice history analysis, the fitted exponential (Eq. 4) and hyperbolic (Eq. 5) models were as follows:

$$1 * (1 - k)^{E-1}, \quad (4)$$

$$\frac{1}{1 + k * (E - 1)}, \quad (5)$$

This normalization technique was not necessary in the outcome history analysis, and so was not employed to avoid loss of information when different rats' data were set to similar scales.

Goal-tracking behavior. Goal-tracking behavior was measured to determine whether the rats in Group Extra-Feedback elicited greater goal-tracking behavior given risky losses, given

that risky losses were cued as if R-11 outcomes were delivered. Goal-tracking behavior was operationally defined as the number of head entries into the food magazine. While rats typically exhibit head entries into the food cup at any given moment during the session, the primary time window of interest was that encompassing risky outcome delivery. For both groups, the R-11 outcome took approximately 2.5 s to deliver. For Group Extra-Feedback, the R-0 and R-1 outcomes involved the same-duration presentation of stimuli (i.e., feeder operation, flashing nosepoke key light). Accordingly, the reception of R-0 and R-1 outcomes for Group Extra-Feedback may have elicited an increase in head-entry response rate past the 2.5-s outcome delivery; moreover, it was expected that the rats in both groups continued to consume the reward corresponding to the R-11 outcome past the duration of the 2.5-s outcome delivery. Thus, this analysis involved the 5-s temporal window following the onset of risky outcome delivery. The number of head entries during this interval was compared across groups as a function of risky outcome magnitude and phase (P[0], P[1]). This analysis was identical to the mixed-effects analysis of choice behavior except that the current analysis assumed a Poisson response distribution given the count-type nature of the data.

Reinforcement learning (RL) models. A modular approach was employed in terms of fitting the RL models. Each model incorporated one of four valuation rules (Simple RL, Asymmetric RL, Valence-Attentive RL, Weighted-Reference-Point RL) and a single common softmax decision rule (Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006).

Simple RL. A Simple RL algorithm implies that both gains and losses of similar magnitudes have equivalent effects on computations of estimated value:

$$\delta_{N,T} = R_{N,T} - Q_{N,T-1}, \quad (6)$$

$$Q_{N,T} = Q_{N,T-1} + \alpha \delta_{N,T}, \quad (7)$$

in which the updated value ($Q_{N,T}$) on trial T for choice N is a function of both the previous estimate of value for that choice ($Q_{N,T-1}$) as well as the prediction error ($\delta_{N,T}$) between the previous estimate of value ($Q_{N,T-1}$) and the outcome that was received on trial T for making choice N ($R_{N,T}$). When $R_{N,T}$ exceeds $Q_{N,T-1}$, $R_{N,T}$ generates a positively-signed prediction error (i.e., gains). When $R_{N,T}$ is less than $Q_{N,T-1}$, $R_{N,T}$ generates a negatively-signed prediction error (i.e., losses). If $R_{N,T}$ equals $Q_{N,T-1}$, then $R_{N,T}$ does not generate a prediction error (i.e., $R_{N,T}$ matches the expectation of $Q_{N,T-1}$; $\delta_{N,T} = 0$). The prediction error, $\delta_{N,T}$, is scaled by a learning-rate parameter, $\alpha \in [0,1]$. Larger α values reflect faster learning and more rapid adjustments in value. The larger the value of α , the more closely $Q_{N,T}$ approximates $R_{N,T}$, such that $Q_{N,T}$ is less influenced by past outcomes and more influenced by recent outcomes.

Asymmetric RL. The Asymmetric RL algorithm, also referred to as risk-sensitive RL (Niv et al., 2012), involves differential value-updating rates following gains and losses (see Frank et al., 2007):

$$\alpha = \begin{cases} \alpha_G, \delta_{N,T} > 0 \\ \alpha_L, \delta_{N,T} \leq 0 \end{cases} \quad (8)$$

$$Q_{N,T} = Q_{N,T-1} + \alpha \delta_{N,T}, \quad (9)$$

in which α_G is the value-updating rate parameter for gains, and α_L is the value-updating rate parameter for losses. Here, α equals α_G and α_L when gains and losses are experienced, respectively. Assuming individual discrepancies in the value-updating parameters, the same prediction error following a gain or loss would elicit differentially-scaled updates in value.

Asymmetric RL (Eq. 9) reduces to Simple RL (Eq. 7) when $\alpha_G = \alpha_L$.

Valence-Attentive RL. The Valence-Attentive RL algorithm implies differences in relative attention to gains versus losses (i.e., expectancy valence model; see Busemeyer & Stout, 2002; also see Bishara et al., 2009):

$$G = \begin{cases} 1, R_{N,T} > Q_{N,T-1} \\ 0, R_{N,T} \leq Q_{N,T-1} \end{cases} \quad (10)$$

$$\delta_{N,T} = [G \cdot \omega \cdot R_{N,T} + (1 - G) \cdot (2 - \omega) \cdot R_{N,T}] - Q_{N,T-1} \quad (11)$$

$$Q_{N,T} = Q_{N,T-1} + \alpha \delta_{N,T}, \quad (12)$$

in which G is a logical operator for whether the previous outcome was greater than the previous value estimate for choice N , and $\omega \in [0,2]$ is the attention given to gains relative to losses, such that greater values of ω reflect more attention to gains versus losses. Accordingly, setting $\omega = 1$ and $G = 1$ reduces the model to Simple RL (Eq. 7). Here, the value-updating rate (α) is identical for gains and losses. In contrast to Asymmetric RL, which involves differential weighting of positively- versus negatively-signed prediction errors, Valence-Attentive RL involves differential weighting of individual outcomes. In contrast to the RL models above, in which $\delta_{N,T} = 0$ when $R_{N,T} = Q_{N,T-1}$, Valence-Attentive RL weights the individual outcomes, such that identical values for $R_{N,T}$ and $Q_{N,T-1}$ will only result in $\delta_{N,T} = 0$ if $\omega = 1$.

Weighted-Reference-Point RL. While traditional RL models assume that the reference point is the expected value of the choice that was made (e.g., Sutton & Barto, 1998), previous research has suggested that gains and losses are encoded relative to the expected value of the alternative outcome (Marshall & Kirkpatrick, 2015). Indeed, individuals may use multiple reference points when making decisions (Wang & Johnson, 2012). Weighted-Reference-Point RL implies that gains and losses may be encoded relative to zero (zero-based reference point), the expected value of the choice that was made (current-choice-based reference point), and/or the

expected value of the alternative choice (alternative-choice reference point; see Marshall & Kirkpatrick, 2015), and that individuals may update value at different rates depending on the elicited prediction error from each of these reference points:

$$\delta_{Z,T} = R_{N,T} - 0, \quad (13)$$

$$\delta_{N,T} = R_{N,T} - Q_{N,T-1} \quad (14)$$

$$\delta_{-N,T} = R_{N,T} - Q_{-N,T-1} \quad (15)$$

$$Q_{N,T} = Q_{N,T-1} + \alpha_Z \delta_{Z,T} + \alpha_N \delta_{N,T} + \alpha_{-N} \delta_{-N,T}, \quad (16)$$

in which $Q_{-N,T-1}$ is the value of the alternative choice (i.e., not choice N), $\delta_{Z,T}$ is the prediction error relative to a zero-based reference point, $\delta_{N,T}$ is the prediction error relative to a current-choice-based reference point (i.e., the same prediction error in the previous models), $\delta_{-N,T}$ is the prediction error relative to an alternative-choice reference point, and α_Z , α_N , and α_{-N} are the value-updating parameters corresponding to each of these prediction errors. Weighted-Reference-Point RL reduces to Simple RL (Eq. 7) when α_Z and α_{-N} are set to 0.

Decision rule. While each RL model had different valuation rules, all models involved the same softmax decision rule (Cohen, McClure, & Yu, 2007; Daw, 2011; Daw et al., 2006; Luce, 1959):

$$P(N_T = Risky | Q_{Certain,T}, Q_{Risky,T}) = \frac{1}{1 + \exp(-\beta(Q_{Risky,T} - Q_{Certain,T}))}, \quad (17)$$

in which $P(N_T = Risky | Q_{Certain,T}, Q_{Risky,T})$ is the probability of a risky choice given the derived value of the certain choice ($Q_{Certain,T}$) and risky choice ($Q_{Risky,T}$), and β is the inverse-temperature parameter of RL models that captures the stochasticity of choice behavior (i.e., larger values of β reflect greater exploitation of the higher-valued choice).

Model fitting and selection. The *fminsearch* optimization algorithm in MATLAB was used to fit each RL model to the trial-by-trial choice data for individual rats via maximum likelihood estimation (Daw, 2011). The parameters of each model were bounded within the following limits: $\alpha \in (0,1)$, $\omega \in (0,2)$, $\beta \in (0,10)$. Model fitting involved the random generation of 300 sets of uniformly-distributed initial parameters (i.e., 300 runs) across the boundary conditions for each parameter. Individual iterations that returned parameters outside the aforementioned parameter boundaries were excluded from analyses. The model's choices in terms of probability [$P(N_T=Risky | Q_{Certain,T}, Q_{Risky,T})$] was compared to each rat's actual choice on each trial (certain = 0, risky = 1), as the same reinforcement history experienced by the rat was provided to the model. These probabilities were then transformed given the rat's choice: For a risky choice, $P = P(N_T=Risky | Q_{Certain,T}, Q_{Risky,T})$; for a certain choice, $P = 1 - P(N_T=Risky | Q_{Certain,T}, Q_{Risky,T})$. Model fit was evaluated in terms of log likelihood, which was calculated as the sum of the log-transformed probabilities. The best-fitting model was the one with the minimum AIC. The selected model was subjected to secondary evaluation of goodness-of-fit via a pseudo- R^2 measure (i.e., omega-squared, ω^2) to aid interpretation of the adequacy of model fit. Here, individual rats' data were smoothed over a moving nine-trial window (observed data), and compared to simulated data that were generated using the selected model's fitted parameters.

A parameter recovery technique was performed to confirm that the RL models were detecting the data generative processes and parameters. For each condition order (Table 2.1), 1000 simulations for each of the four valuation rules were conducted. The simulations incorporated the same task structure that the rats experienced (see Table 2.1). For each simulation, model parameters were randomly sampled from uniform distributions [i.e., $\alpha \in (0,1)$, $\omega \in (0,2)$, $\beta \in (0,10)$] and these sampled parameters were used in the valuation and decision

rules that determined the simulated choices. The simulated data were then fit with the RL models described above to determine whether the fitted parameter matched the corresponding simulated parameter. For parameter recovery, each simulation was fit over 50 iterations.

Results and Discussion

Choice behavior

Molar analysis

Overall effects. The molar analysis of risky choice was conducted to examine the effects of the phase and probability manipulations in both feedback groups. Here, analysis involved 158,024 observations from the risky choice task, and included the overall intercept, a categorical predictor of feedback group (Normal-Feedback, Extra-Feedback), and continuous predictors of phase (P[0], P[1]), probability (.10, .50, .90), and session. Feedback group was effect coded with Normal-Feedback/Extra-Feedback as -1/+1.

For this molar analysis, the probability predictor was recoded as sub-phase. Specifically, in terms of probability, the experimental design can be viewed as an ACBA|CBA design (i.e., .10, .90, .50, .10, .90, .50, .10; Table 2.1), with the pipe indicating the phase separation. Thus, there was more exposure to the .10 (A) probability sub-phase in the first phase (P[0] or P[1]) than in the second phase (P[1] or P[0]). Also, in the first .10 (A) probability sub-phase, the rats were beginning to learn the dynamics of the task. Because the current molar analysis involved analyzing learning as a function of session within each phase/sub-phase, the .10 probability sub-phase data were separated, as the first .10 probability sub-phase may be viewed as qualitatively unique relative to the other .10 (A) probability sub-phases. Thus, to discourage averaging effects between the two .10 probability sub-phases in the first phase, the probability data were recoded such that the .90 probability sub-phases were set to “1”, the .50 probability sub-phases to “2”, the

terminal .10 probability sub-phases to “3”, and the initial .10 probability sub-phase in Phase 1 to “4”. The decision to code the terminal .10 sub-phases as “3” and the initial .10 sub-phase as “4” was data-driven, as the rats made more risky choices in the initial .10 sub-phase of Phase 1 than in the terminal .10 sub-phases (Figure 2.2). Accordingly, for these molar analyses, the predictor of probability was referred to as sub-phase. Due to this recoding, only half of the rats had data in the initial .10 sub-phase for each of the P[0] and P[1] phases. However, sub-phase was treated as a continuous predictor, such that the current regression analyses attenuated potential problems of missing data that would be detrimental to an analysis-of-variance approach to analysis.

For this and subsequent results sections, the in-text reporting of results focused on the theoretically-relevant results to the corresponding analyses, but statistical tables are provided throughout the manuscript that include the full model analyses.

Table 2.2 shows the full model output from the mixed-effects analysis. The fixed-effects structure included the full factorial model of Feedback Group \times Phase \times Sub-Phase \times Session. Intercept, phase, sub-phase, session, and Phase \times Sub-Phase served as random effects. While there was no main effect of feedback group on overall risky, there was a Feedback Group \times Phase \times Sub-Phase \times Session interaction, $t(158008) = 10.87, p < .001$.

Figure 2.2 shows the proportion of choices for the risky outcome as a function of session for Groups Extra-Feedback and Normal-Feedback, separated by phase (P[0], P[1]) and sub-phase (i.e., the probability of the manipulated outcome; R-0 in P[0], R-1 in P[1]). The far left panel of Figure 2.2 shows risky choice behavior as a function of session across groups and phases in the first sub-phase of Phase 1. Here, in the first sub-phase of the entire task, both groups increased their risky choice behavior as a function of session. Group Normal-Feedback exhibited a steeper slope than Group Extra-Feedback in the P[0] phase (Group Normal-Feedback: slope = 0.22;

Group Extra-Feedback: slope = 0.17), but Group Extra-Feedback exhibited a much steeper slope than Group Normal-Feedback in the P[1] phase (Group Normal-Feedback: slope = 0.08; Group Extra-Feedback: slope = 0.33). In the .10 probability condition of the P[1] phase, the frequency of non-zero risky reward was at its largest. Accordingly, it would be expected that both groups would exhibit comparably high levels of risky choice behavior during this condition, but this did not occur. The increased riskiness in Group Extra-Feedback in this condition may be explained by the elevated frequency with which the extra multimodal stimuli accompanied the actual delivery of the R-11 outcome (i.e., greater contingency of the multimodal stimuli with R-11 delivery). The multimodal stimuli would serve as a strong conditioned reinforcer in this condition. The more positive contingency between the multimodal feedback and outcome magnitude experienced by Group Extra-Feedback may have promoted increased risky choice, given the greater value of the conditioned reinforcer (Zentall & Stagner, 2010; also see Seo & Lee, 2009; but see Barrus & Winstanley, 2016).

Following the .10 probability sub-phase(s), the probability of the manipulated outcome (i.e., 0 pellets in P[0]; 1 pellet in P[1]) increased to .90. Unsurprisingly, all groups showed large decreases in risky choice behavior with continued progression in this sub-phase (Figure 2.2, second panel from the left). Here, while both groups showed relatively comparable decreases in the P[0] phase (Group Normal-Feedback: slope = -0.24; Group Extra-Feedback: slope = -0.23), Group Extra-Feedback exhibited a steeper decline than Group Normal-Feedback in the P[1] phase (Group Normal-Feedback: slope = -0.27; Group Extra-Feedback: slope = -0.35), likely due to their exhibiting greater risky choice at the onset of this sub-phase.

When the probability of the manipulated outcome was .50, Group Extra-Feedback exhibited slightly greater mean risky choice behavior than Group Normal-Feedback, despite the

similarities between the groups at the end of the .90 probability sub-phase. Additionally, both groups exhibited minimal change or slight increases in risky choice as a function of session (Figure 2.2, second panel from the right). However, computations of simple slopes from the statistical output equated to small decreases in risky choice as a function of session (Group Normal-Feedback, P[0]: slope = -0.08; Group Extra-Feedback, P[0]: slope = -0.10; Group Normal-Feedback, P[1]: slope = -0.15; Group Extra-Feedback, P[1]: slope = -0.12). These discrepancies between the mean data and fitted slopes were likely due to the unweighted effects coding of the mixed effects analysis and the differential weighting of group means (i.e., by the different number of choices across phases, sub-phases, and sessions for each rat within both groups). Alternatively, such discrepancies may have been caused by the even linear spacing of the recoded probability predictor as sub-phase, such that the slopes at probability = .50 would result in values between the slopes at probabilities = .90 and .10, rather than an actual decrease in mean risky choice as a function of session.

In the final sub-phase of each phase, the probability of the manipulated outcome decreased to .10, and, given the increase in expected value of the risky choice, there was a general increase in risky choice as a function of session in this sub-phase (Figure 2.2, far right panel). Group differences diverged, as Group Extra-Feedback also exhibited greater mean risky choice behavior than Group Normal-Feedback in this sub-phase. Simple slopes computations revealed that both groups exhibited an increase in risky choice as a function of session in the P[0] phase (Group Normal-Feedback: slope = 0.07; Group Extra-Feedback: slope = 0.03). Group Extra-Feedback showed a similar increase in risky choice behavior in the P[1] phase (slope = 0.10). While the simple slopes calculations indicated that Group Normal-Feedback exhibited a decrease in risky choice as a function of session in the P[0] phase (slope = -0.04), the

mean data suggest that there was a slight increase in risky choice as a function of session in the terminal .10 sub-phase of the P[1] phase.

Overall, Figure 2.2 indicates that both groups exhibited a strong decrease in risky choice as a function of session when the R-11 outcome was relatively unlikely ($P[0] | P[1] = .90$) and relative increases in risky choice as a function of session when the R-11 outcome was more likely ($P[0] | P[1] = .10$), corroborating research that has reported sensitivity to changing risk and probability in rats (Cardinal & Howes, 2005; Larkin et al., 2016; Marshall & Kirkpatrick, 2013, 2015; Mazur, 1988; Mobini et al., 2002; Mobini, Chiang, Ho, Bradshaw, & Szabadi, 2000; Stopper & Floresco, 2011, 2014; Stopper, Tse, Montes, Wiedman, & Floresco, 2014).

Individual differences. As described above, the model's random effects included intercept, phase, sub-phase, session, and Phase \times Sub-Phase. There were considerable individual differences in risky choice as a function of these factors (Figures 2.3 and 2.4). Figure 2.3 shows the individual rat data during the P[0] phase. For the six rats in each group that experienced P[0] as the first phase (Rats N.1-N.6, E.1-E.6), risky choice behavior tended to increase as a function of session in the initial $P[0] = .10$ sub-phase. When the probability of R-0 increased to .90 ($P[0] = .90$), the rats decreased their risky choice behavior as a function of session (but see, e.g., Rat N.9, in which choice behavior was relatively constant as a function of session, and Rat N.10, in which there was a slight increase in risky choice as a function of session). The probability of R-0 then decreased to .50 ($P[0] = .50$). While the model fits tended to produce slightly negative slopes, which matched some rats' data (e.g., Rat E.2), other rats showed constant risky choice (e.g., Rat N.7) or increased risky choice as a function of session in this sub-phase (e.g., Rat E.7). Moreover, when the probability of R-0 decreased again to .10 ($P[0] = .10$) in the terminal sub-phase of the P[0] phase, rats showed a general increase in risky choice as a function of session

(but see Rats E.8, E.9, and E.11, in which risky choice behavior was relatively constant and approximated zero).

For the six rats in each group that experienced the P[1] phase first (Rats N.7-N.12, E.7-E.12), risky choice behavior tended to increase as a function of session in the initial P[1] = .10 sub-phase. The probability of the R-1 outcome then increased to .90, and the rats tended to exhibit decreases in risky choice behavior as a function of session in this sub-phase (e.g., Rats N.8 and E.10) unless their risky choice behavior was relatively constant with very few risky choices made (e.g., Rat N.5). When the probability of the R-1 outcome decreased to .50, many rats showed only minimal change in risky choice as a function of session (e.g., Rats N.9 and E.6), while other rats showed small increases in risky choice (e.g., Rat E.2). Finally, after the probability of the R-1 outcome had decreased again to .10, some rats showed an increase in risky choice with session (e.g., Rats E.2 and E.4), while other rats were seemingly unaffected by the change in R-11 probability (e.g., Rats N.5 and E.1).

As seen across Figures 2.3 and 2.4, the model fits approximated the corresponding patterns in some rats' behavior (e.g., Rats E.7 and E.10 in Figure 2.4). However, for other rats, the model fits poorly approximated the rats' data (e.g., Rat N.11 in Figure 2.4). Many of these poorer fits occurred in rats that exhibited considerable differences in behavior between the initial and terminal .10 probability sub-phases of the first phase experienced, in that many of these rats showed increased risky choice behavior in the initial than in the terminal .10 probability sub-phases (e.g. Rat N.2 in Figure 2.3). These large differences in individual risky choice behavior paralleled the differences in group means in Figure 2.2, as there were large decreases in risky choice behavior from the initial .10 probability sub-phase in Phase 1 to the subsequent sub-phases. Accordingly, these data suggest that the rats' behaviors in the two .10 probability sub-

phases was dependent on the rats' procedural history, as exposure to high rates of risky loss in the .90 and .50 probability sub-phases seemingly drove some rats to general risk aversion even when the probability of the R-11 outcome increased to .45 in the terminal .10 probability sub-phases. Moreover, as the expected value of the risky choice was greater than 5 pellets in the .10 probability sub-phases, the rats who were risk averse in the terminal .10 probability sub-phase were behaving suboptimally in terms of reward rate, as the expected value of the certain choice was 3 pellets across all conditions. Thus, despite the potential to earn considerably more long-term reward given exclusive risky choice behavior, these data suggest that the history and frequency of risky losses had stronger effects on choice behavior than choices' expected values.

Molecular analysis

Effect of the previous outcome

Overall effects. A molecular analysis was conducted to determine how differential feedback, phase, and probability manipulations affected the likelihood that rats made risky choices as a function of the outcome of the previous choice. Analysis involved 156,354 observations, and included the overall intercept, a categorical predictor of feedback group and previous outcome (C-2, C-4, R-0, R-1, R-11), and continuous predictors of phase and the probability of the manipulated outcome in each phase. Probability was not recoded as sub-phase. Feedback group was effect coded with Normal-Feedback/Extra-Feedback as -1/+1. The reference level of previous outcome was the C-2 outcome. The fixed-effects structure included all main effects and interactions within the full factorial model of Feedback Group \times Previous Outcome \times Phase \times Probability except for the four-way interaction. Intercept, phase, and previous outcome served as random effects.

Table 2.3 shows the full model output from the mixed-effects analyses. A critical question of the current experiment was whether differential feedback influenced rats' choice behavior following differential outcomes. Overall, Group Extra-Feedback exhibited a significantly greater likelihood of making risky choices compared to Group Normal-Feedback, $t(156318) = 2.18, p = .030$ (see Figure 2.5). While the molar analysis did not reveal a significant difference in risky choice between the feedback groups, the mean data in Figure 2.2 indicated a trend in this direction. With the added factor of previous outcome, this molecular analysis suggested that the addition of multimodal gain-related (R-11) feedback to the delivery of risky losses (R-0, R-1) resulted in increased risky choice in Group Extra-Feedback. Assuming that such a combination of gain-related feedback with the delivery of losses serves to transform the R-0 and R-1 losses into LDWs, then the current results are potentially consistent with research indicating that LDWs drive humans to overestimate the frequency of "winning" (M. J. Dixon, Collins, Harrigan, Graydon, & Fugelsang, 2015; Jensen et al., 2012), which would theoretically lead to elevated risky choice following these perceived but nonexistent "wins".

The molecular analysis also revealed a Feedback Group \times Phase \times Previous Outcome interaction. The left panel of Figure 2.5 shows the proportion of choices for the risky outcome as a function of phase (P[0], P[1]) and outcome of the previous choice for both groups. Following C-2 and C-4 outcomes in the P[0] and P[1] phases, Groups Extra-Feedback and Normal-Feedback behaved comparably, exhibiting a high mean tendency to make certain choices after certain outcomes. In contrast, the groups diverged in their risky choice behavior following risky outcomes. Specifically, in both the P[0] and P[1] phases, Group Extra-Feedback made more risky choices than Group Normal-Feedback following risky outcomes, and this difference increased with risky outcome magnitude. Thus, the extra multimodal feedback produced a

general increase in risky choice in Group Extra-Feedback following the reception of risky outcomes, and any group differences were considerably smaller following certain outcomes. In other words, the added feedback presented to Group Extra-Feedback following risky choices (and not certain choices) resulted in greater riskiness after risky outcomes (and not certain outcomes). Additionally, as seen in Figure 2.5 (left panel), there was a distinct pattern of results following R-0 and R-1 outcomes that depended on phase. Specifically, in the P[0] phase, rats made more risky choices after R-1 than after R-0 outcomes, and the opposite pattern was exhibited in the P[1] phase (Marshall & Kirkpatrick, 2015). As described below, these differences may indicate that the rats were tracking the probability of the R-11 outcome and how often the R-11 outcome occurred relative to one of the risky losses.

Molecular analysis also revealed a Feedback Group \times Probability \times Previous Outcome interaction (Figure 2.5, right panel). As in the Feedback Group \times Phase \times Previous Outcome interaction, there were relatively small differences between feedback groups in risky choice following C-2 outcomes and C-4 outcomes, but considerably larger differences between feedback groups following risky outcomes. Interestingly, these mean group differences were smaller following all risky outcomes when the probability of the manipulated outcome was .10 than when it was .50 or .90. This suggests that the added multimodal feedback stimuli had the largest mean effects when the probability of risky losses was at its greatest (i.e., when the probability of the R-11 outcome was smallest). Thus, assuming that the added multimodal feedback transformed the R-0 and R-1 outcomes into LDWs, then these data indicate that the more often the rats experienced LDWs, the more often they made risky choices after risky outcomes. Indeed, Jensen et al. (2012) showed that humans are more likely to estimate how often they “won” when they are exposed to more LDWs. Accordingly, these data provide

support for the LDW model in risky choice in rats and how responses to such ambiguous losses are moderated by the probability of their occurrence.

Overall, these molecular analyses suggest that Group Extra-Feedback made more risky choices than Group Normal-Feedback, and that these differences were primarily observed following risky outcomes. These results are reasonable in the sense that certain outcomes were presented in the same fashion to both groups. Furthermore, the Feedback Group \times Probability \times Previous Outcome indicated that win-stay/lose-shift behavior is not absolute, but moderated by outcome probability (Marshall & Kirkpatrick, 2013). While previous experiments have analyzed win-stay/lose-shift behavior in rats in risky environments, some of these studies have collapsed across risky-outcome probabilities (see, e.g., Larkin et al., 2016; Stopper & Floresco, 2011), thus masking the effect of outcome probability on subsequent risky choice. However, the analysis of probability moderation on local risky choice behavior may be crucial to future research.

Individual differences. The random effects included intercept, phase, and previous outcome. Figure 2.6 portrays the individual differences in risky choice as random effects. For many rats, the model fits did well to approximate the individual rats' mean choice behavior following differential outcomes. Cases in which the model had a poorer fit may have been driven by the discrepancy in overall risky choice behavior between phases. For example, Rat E.6 made 578 risky choices in the P[0] phase, but only 46 risky choices in the P[1] phase. Accordingly, the increased risk aversion in the P[1] phase seemingly drove the model fit below the actual rat data given the discrepancy in number of risky choices that were made between phases (and, thus, the number of choices that could follow the corresponding risky outcomes).

Regardless of the individual rat biases, all rats were more likely to make a risky choice after a risky outcome than after a certain outcome, consistent with previously-reported

perseverative/stay biases in risky choice paradigms (see Worthy, Pang, & Byrne, 2013). Some rats exhibited greater risky choice after risky outcomes in the P[0] phase than in the P[1] phase (e.g., Rats N.1 and E.1), while some rats exhibited the opposite pattern (e.g., Rats N.12 and E.8). Thus, the effect of manipulating differential losses is seemingly dependent on the individual rat, indicating individual differences in loss sensitivity (Marshall & Kirkpatrick, 2015).

In the P[0] phase, the majority of the rats exhibited greater risky choice following larger risky outcome magnitudes (but see, e.g., Rats N.5 and E.11), which is consistent with general win-stay/lose-shift behavior in rats (Larkin et al., 2016; Marshall & Kirkpatrick, 2013; Stopper & Floresco, 2011). For the rats that showed a contrasting pattern (e.g., Rats N.5 and E.11), the elevated risky choice following risky losses is indicative of elevated loss-chasing, potentially reflecting frustration-induced choice perseverance to compensate for the previous losses (see Breen & Zuckerman, 1999; Rogers et al., 2013; Tremblay et al., 2014). Furthermore, many of the rats in both groups exhibited the averaged pattern shown in Figure 2.5, in which more risky choices were made after R-0 than after R-1 outcomes in the P[1] phase (but see, e.g., Rats N.3 and E.3). As this latter result has been demonstrated previously (Marshall & Kirkpatrick, 2015), these results indicated that trial-by-trial risky choice behavior was strongly impacted by the relationship between the frequency of large risky gains (R-11) and that of differential losses. Specifically, in the P[0] condition, the probabilities of the R-0 and R-11 outcome were anti-correlated, such that increases in $p(R-0)$ were associated with decreases in $p(R-11)$. In contrast, the probabilities of R-0 and R-11 were correlated in the P[1] phase. The greater risky choice following R-0 outcomes than R-1 outcomes in the P[1] phase may have potentially reflected the rats' tracking $p(R-11)$; the more that they receive R-0 outcomes, the more likely they were to receive subsequent R-11 outcomes. Accordingly, the differential effects of phase on risky choice

following R-0 and R-1 outcomes may indicate that the rats were using a learned model of their environment to make sequential decisions (Huh, Jo, Kim, Sul, & Jung, 2009; also see Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Doll, Simon, & Daw, 2012).

Outcome and choice history. The first molecular analysis evaluated the effect of the most recent outcome on subsequent choice. Previous research has indicated that the influence of a given previous outcome on subsequent choice decays with each subsequent outcome (e.g., Devenport et al., 1997; Glimcher, 2011). This decay has been proposed to be exponential (Bayer & Glimcher, 2005; Daw & Tobler, 2013; Glimcher, 2011), while other reports have suggested a hyperbolic decay of past events (Devenport et al., 1997; also see Vasconcelos, Monteiro, Aw, & Kacelnik, 2010). To our knowledge, while there have been several previous analyses of outcome and/or choice decay (Kennerley, Walton, Behrens, Buckley, & Rushworth, 2006; Kim, Lee, & Jung, 2013; Kim, Sul, Huh, Lee, & Jung, 2009; Kwak et al., 2014; Kumaran, Warren, & Tranel, 2015; Lau & Glimcher, 2005; Lee, Conroy, McGreevy, & Barraclough, 2004; McCoy & Platt, 2005; Sul, Kim, Huh, Lee, & Jung, 2010; Walton, Behrens, Buckley, Rudebeck, & Rushworth, 2010), none of them have reported the functional form of such decay. Moreover, Rutledge et al. (2009) only fit an exponential function to their decay data (also see Beeler, Daw, Frazier, & Zhuang, 2010). Thus, a second molecular analysis was conducted to determine whether differential risky-outcome feedback affected the decaying influence of past events on subsequent choice, as well as to determine the function that best characterized the nature of decay. As described above, separate analyses were conducted for outcome history and choice history.

The first component of this analysis involved determining the regression coefficients of previous outcomes and choices as a function of outcome lag and choice lag, respectively. Analyses involved 141,324 observations, and included the overall intercept, a random intercept,

and fixed and random effects of the previous 9 outcomes in the outcome history analysis (i.e., the random-model including 10 previous outcomes did not converge on a solution) and previous 10 choices in the choice history analysis. Tables 2.4 and 2.5 include the mixed-effects output for the outcome history and choice history analyses, respectively. Intuitively, this analysis is a win-stay/lose-shift analysis, in which more positive slopes reflect a greater likelihood of making risky choices after risky choices (or greater risky outcome magnitudes) and of making certain choices after certain choices (or greater certain outcome magnitudes). For both histories, there was a decay in regression coefficients as a function of previous outcome and choice: More recent outcomes and choices had greater influences over subsequent choices than more temporally distant outcomes and choices. All regression coefficients for outcomes and choices were significantly greater than 0, $t_s \geq 3.85$, $p_s < .001$, suggesting that the past series of outcomes and choices had partial influence over subsequent choice behavior (see Landon, Davison, & Elliffe, 2002).

The random coefficients were analyzed via nonlinear mixed-effects models to determine if the decays of the outcome and choice coefficients were better characterized by exponential or hyperbolic functions (Eq. 2-5). Figures 2.7 and 2.8 show the individual rat hyperbolic and exponential model fits for outcome history and choice history, respectively. The decay rate parameter, k , was set as a fixed and random effect in the outcome and choice history analyses. The intercept parameter, A , was set as a fixed and random effect in the outcome history analysis.

All rats showed a general decrease in regression coefficients in the outcome history analysis (Figure 2.7). The hyperbolic model analysis of outcome history revealed a significant difference between feedback groups in decay rate (k), $t(189) = 4.32$, $p < .001$. Group Extra-Feedback ($k = 1.10$) showed a significantly greater decay in the influence of previous outcomes

compared to Group Normal-Feedback ($k = 0.61$). However, there were no statistically significant differences in intercept (A) in the hyperbolic analysis, $t(189) = 1.82, p = .070$ (Group Extra-Feedback: $A = 0.25$; Group Normal-Feedback: $A = 0.19$). The exponential outcome history analysis also revealed a significantly greater decay rate in Group Extra-Feedback ($k = 0.36$) than in Group Normal-Feedback ($k = 0.23$), $t(189) = 4.53, p < .001$. Moreover, Group Extra-Feedback showed a significantly higher intercept ($A = 0.23$) than Group Normal-Feedback ($A = 0.17$) in the exponential analysis, $t(189) = 2.01, p = .045$. The hyperbolic model accounted for a better fit of the data than the exponential model (hyperbolic AIC = -974.25, exponential AIC = -834.97).

The choice history analysis revealed comparable results as the outcome history analysis. All rats showed a general decrease in normalized regression coefficients in the choice history analysis (Figure 2.8). Specifically, there was a significantly greater decay rate in Group Extra-Feedback ($k = 2.67$) than in Group Normal-Feedback ($k = 1.77$), $t(215) = 2.34, p = .020$. The exponential analysis also indicated that Group Extra-Feedback exhibited a significantly greater decay rate ($k = 0.68$) than Group Normal-Feedback ($k = 0.55$), $t(215) = 2.52, p = .012$. As in the analysis of outcome history, the hyperbolic model accounted for a better fit of the data in the choice history analysis (hyperbolic AIC = -541.79, exponential AIC = -330.89).

The current analysis revealed two key results: (1) Group Extra-Feedback exhibited a steeper decay rate of outcome and choice histories, and (2) the hyperbolic model accounted for the outcome and choice decay rate coefficients better than the exponential model did. Steeper decay rates have been suggested to be analogous to greater learning rates (see Glimcher, 2011). Interestingly, previous research has suggested that steeper learning rates are related to greater risk aversion in the presence of gains (March, 1996). Accordingly, as Group Extra-Feedback

exhibited greater risk proneness and steeper decay rate coefficients than Group Normal-Feedback, these results may suggest that the idiosyncratic rates of decaying coefficients of outcome and choice history are not necessarily indicators of individual differences in learning, but perhaps those in the decision to persevere on and/or exploit the previously made choice. Similarly, as risky foraging decisions have been suggested to be made in bouts (Krebs, Kacelnik, & Taylor, 1978), individual differences in outcome and choice history decay rates may reflect individual differences in sampling behavior and bout length.

As revealed by the analysis, the decays of outcome and choice coefficients were better described by a hyperbolic than an exponential function. Exponential decay is assumed within traditional reinforcement learning algorithms (Sutton & Barto, 1998), in which the influence of past events and the value of future events decays in an exponential fashion (see Bayer & Glimcher, 2005; Daw & Tobler, 2013; Glimcher, 2011). However, given the considerable support for hyperbolic discounting of future rewards (e.g., Myerson & Green, 1995), recent efforts have been made to implement a hyperbolic version of classic reinforcement learning principles (Alexander & Brown, 2010). Accordingly, trial-by-trial models of decision making may be improved by employing a hyperbolic decay of previous events, especially if future research corroborates the current results (see Maia, 2009).

Goal-tracking behavior

Overall effects. Goal-tracking behavior was analyzed to determine whether the differential feedback conditions affected the rats' head-entry behavior into the food trough following reinforcer delivery. Specifically, as Group Extra-Feedback received the same feedback across risky outcomes, it was hypothesized that these rats would continue to exhibit head entries to retrieve food even when food was not delivered (i.e., R-0 outcomes). This

analysis involved observations from trials in which rats made a risky choice, totaling 30,116 observations, and included the overall intercept, a categorical predictor of feedback group (Normal-Feedback, Extra-Feedback), and continuous predictors of phase (P[0], P[1]) and risky outcome magnitude (R-0, R-1, R-11). Here, risky outcome magnitude was treated as continuous because the behavior following only one choice's outcomes was analyzed, such that the different choice's outcomes did not have to be placed along the same continuum; furthermore, there was no theoretical backing to assume a non-monotonic relationship between goal-tracking behavior and risky outcome magnitude in the current paradigm, as was done for the molecular analysis above (Marshall & Kirkpatrick, 2015). Feedback group was effect coded with Normal-Feedback/Extra-Feedback as -1/+1. The fixed-effects structure included the full factorial model of Feedback Group \times Risky Outcome Magnitude + Phase \times Risky Outcome Magnitude. Intercept, phase, risky outcome magnitude, and Phase \times Risky Outcome Magnitude were random effects.

Table 2.6 includes the full model output from the mixed-effects analyses, and Figure 2.9 shows the arithmetic mean number of head entries during the 5-s temporal window following the onset of risky outcome delivery as a function of risky outcome magnitude and feedback group. The arithmetic mean of these count data were shown due to the frequency of zeros in the data, and as the link function of the generalized linear mixed effects model acts on the arithmetic mean rather than the geometric mean of the data. Group Extra-Feedback exhibited a significantly greater number of head entries compared to Group Normal-Feedback, $t(30110) = 3.68, p < .001$. The number of head entries significantly increased as a function of risky outcome magnitude, $t(30110) = 19.12, p < .001$. Moreover there was a significant Feedback Group \times Risky Outcome Magnitude interaction, $t(30110) = -3.31, p = .001$. Specifically, Group Extra-Feedback exhibited

a shallower slope as a function of risky outcome magnitude (slope = 0.05) compared to Group Normal-Feedback (slope = 0.07).

Overall, these goal-tracking results provide potential face validity to the current paradigm. Group Extra-Feedback received the same added multimodal stimuli following all risky choices, such that the additional tactile, auditory, and visual stimulation were always presented regardless of risky outcome magnitude. If Group Extra-Feedback was only attentive to the food outcome that was delivered, then the rats in this group would be expected to exhibit similar goal-tracking behavior as Group Normal-Feedback. However, not only did Group Extra-Feedback exhibit more head entries, but a shallower relationship between risky outcome magnitude and head-entry behavior. Accordingly, Group Extra-Feedback's goal-tracking behavior was seemingly governed by both the actual risky outcome magnitude, as well as the additional feedback that unreliably predicted the delivered risky outcome.

These goal-tracking results suggest that Group Extra-Feedback exhibited a reduced sensitivity to risky outcome magnitude. This reduced sensitivity may potentially be driven by the reduced salience of individual risky outcome magnitudes due to their presentations being more similar than they were for Group Normal-Feedback. Furthermore, this reduced sensitivity and reduced salience of differential risky outcomes may have contributed to the steeper outcome and choice history decay rates in Group Extra-Feedback; that is, the similarities in risky outcome presentation may have disrupted Group Extra-Feedback's memory for past outcomes, thereby resulting in the rats in Group Extra-Feedback not using the past history of outcomes and choices as determinants of future behavior to the same extent that the rats in Group Normal-Feedback did. Accordingly, a reduced sensitivity to differential risky outcomes, reduced salience of differential risky outcomes (especially risky losses), and the steeper outcome and choice history

decay rates may collectively predict the heightened risky choice behavior in Group Extra-Feedback following risky outcomes. If so, then the presentation of LDWs may result in increased risk taking due to reward hyposensitivity during gambling (see Lole, Gonsalvez, & Barry, 2015; but also see Lole, Gonsalvez, Barry, & Blaszczynski, 2014). Therefore, in conjunction with those from the molar and molecular choice analyses, these results suggest that the differential feedback had substantial influence over Group Extra-Feedback's behavior.

Individual differences. The random effects model of goal-tracking behavior included intercept, phase, risky outcome magnitude, and Phase \times Risky Outcome Magnitude. Indeed, there were substantial individual differences in head-entry/goal-tracking behavior as a function of these random effects (Figure 2.10). In both the P[0] and P[1] phases, the majority of rats showed increased head-entry behavior as a function of risky outcome magnitude (but see, e.g., Rats E.8 and N.12). Additionally, while many rats behaved relatively similarly across phases as a function of risky outcome magnitude (e.g., Rats E.1 and N.3), other rats exhibited differential head-entry behavior across risky outcome magnitudes in the different phases (e.g. Rats E.3 and N.9). As seen in Figure 2.10, Rat N.5 exhibited zero head entries across risky outcome magnitudes in the P[1] phase, which is likely due to an undiagnosed equipment issue. Accordingly, the mixed-effects model described above (Table 2.6) was also run without Rat N.5 to determine if the patterns in Figure 2.9 were driven by this rat's performance. However, the significant effects and interactions identified in Table 2.6 were maintained.

RL models

In addition to the formal statistical analyses described above, the data were modelled with different RL algorithms. Interestingly, the results from the outcome and choice history analyses suggested that the decay rates were better characterized by a hyperbolic than exponential

function. Even though exponential decay is assumed with traditional RL models (e.g., Glimcher, 2011), such that the current data may be better characterized by modifications of RL algorithms, there has been widespread application of RL models to risky decision making data (e.g., Daw et al., 2011; Niv et al., 2012). Moreover, modifications to simple RL models (as described above) and the support for such modifications present the opportunity to test competing hypotheses as to the mechanisms that best account for individual differences in risky choice and loss-chasing behaviors. Accordingly, the RL modelling approach described above was employed.

The first step of RL modelling was parameter recovery, in which data were simulated with uniformly-distributed parameters and then fit with the RL models to gauge how well the models identified the data generative mechanisms and parameters. The second step of the model fitting was fitting the RL models to the observed rat data.

Parameter recovery. Figure 2.11 shows the parameter recovery results of the Simple, Asymmetric, Valence-Attentive, and Weighted-Reference-Point RL models. In each panel, the thick line represents the unit diagonal and each data point represents the parameter that was used in its simulation, and the corresponding fitted parameter. Each panel of Figure 2.11 includes all simulations except those in which the fitted parameters exceeded the boundaries of the distribution from which the simulated parameters were sampled. This occurred in 4.3% of the simulations of Simple RL, 7.8% of the simulations of Asymmetric RL, 5.2% of the simulations of Valence-Attentive RL, and 27.8% of the simulations of Weighted-Reference-Point RL. The Simple, Asymmetric, and Valence-Attentive RL models did well to recover the parameters from the simulated data. However, the Weighted-Reference-Point RL model did poorly to recover the simulated parameters, even when one of the learning-rate parameters (α_z) was removed (Figure 2.11). This was likely due to over-parameterization and instability of the Weighted-Reference-

Point RL model (e.g., both learning-rate parameters acted simultaneously on the different prediction errors that were relative to the same reward outcome). For example, given similarly sized prediction errors relative to different reference points, concurrent increases and decreases in the two corresponding learning rates would produce similar adjustments in value, such that these parameters would become unidentifiable. Thus, the Simple, Asymmetric, and Valence-Attentive RL models were fit to the data.

Model selection. Figure 2.12 shows the AICs for the best-fitting Simple, Asymmetric, and Valence-Attentive RL models for each rat's data. The thick line in Figure 2.12 represents the unit diagonal, such that data points corresponding to the Simple or Valence-Attentive RL models below the diagonal reflect support for the Simple or Valence-Attentive RL models, respectively, and data points above the diagonal represent support for the Asymmetric RL model. Even through 300 iterations, the Simple RL model did not converge on a viable solution for 14 rats. In contrast, the Valence-Attentive and Asymmetric RL models converged on a viable solution for all 24 rats. The mean (median) AIC of the viable Asymmetric RL fits was 4409.5 (4171.1), while those of the Valence-Attentive RL and viable Simple RL fits were 7783.8 (7763.9) and 5972.2 (6438.4), respectively. For two rats, the Valence-Attentive RL model provided a better account of the data than the Asymmetric RL model ($\max \Delta AIC = 36.7$); however, for the remaining rats, the mean ΔAIC in support of the Asymmetric RL model was 3682.8. Thus, Asymmetric RL provided the best account of the risky choice data.

Model fit (Asymmetric RL). Figure 2.13 shows the simulations of the fitted Asymmetric RL model parameters to individual trial data of five different rats. The alternating gray and white boxes in Figure 2.13 indicate the different phases and sub-phases (see Table 2.1). The observed data in Figure 2.13 were smoothed over a moving nine-trial window (see Lau &

Glimcher, 2005). The panels represent the data and model fit corresponding to the minimum, 25th percentile, median, 75th percentile, and maximum of the ω^2 values (i.e., the top panel includes the model with the lowest ω^2 value; the bottom panel, the largest ω^2 value). Even for the data that were most poorly fit by the model in terms of ω^2 (Figure 2.13, top panel), the model approximated the rat's behavior.

Table 2.7 shows the summary statistics of the best-fitting parameter estimates of the Asymmetric RL model, as well as the corresponding fit indices. The gain-based learning rates / value-updating rate (α_G) were significantly smaller than those of losses (α_L), $z = 4.29$, $p < .001$ (Wilcoxon signed-ranks test). As the rats learned more readily from losses than they did gains, these data suggest that the rats were seemingly loss averse (Bhatti, Jang, Kralik, & Jeong, 2014), and more sensitive to losses than they were to gains (Marshall & Kirkpatrick, 2015). In contrast, there were no differences between Groups Extra-Feedback and Normal-Feedback in α_G , $z = 0.72$, $p = .470$, α_L , $z = -0.14$, $p = .885$, and β , $z = -0.84$, $p = .403$ (Wilcoxon rank-sum tests), suggesting that the group differences in choice behavior cannot be fully explained in terms of significantly different gain-based learning rates, loss-based learning rates, and choice stochasticity.

To determine whether individual differences in local risky choice (see Figure 2.6) could be elucidated by individual differences in model parameters and whether the model parameters differentiated behavioral differences across feedback groups, simple and multiple linear regression models were performed. In the first series of models, mean risky choice across all phases and probabilities was regressed on α_G , α_L , and β . The data and the model parameters were log-transformed to correct for positive skewness. Figure 2.14 (left panel) shows the significant positive relationship between α_G and mean risky choice, $b = 0.44$, $t(22) = 5.26$, $p < .001$, model $R^2 = .56$. In contrast, there was no relationship between mean risky choice and α_L

(Figure 2.14, middle panel), $b = -0.18$, $t(22) = -1.17$, $p = .254$, model $R^2 = .06$. Lastly, there was a significant negative correlation between β and mean risky choice (Figure 2.14, right panel), $b = -0.51$, $t(22) = -2.39$, $p = .026$, model $R^2 = .21$.

A second series of regression models involved two derived measures of molecular choice behavior (Figure 2.15). First, the regression criterion was a measure of win-stay/lose-shift behavior following risky choice outcomes. In all conditions, the R-0 outcome was always an objective loss and the R-11 outcome was always an objective gain. Accordingly, a Win-Stay Index was derived as dividing the proportion of risky choices following R-11 outcomes by that following R-0 outcomes. Larger values indicated a greater proportion of risky choice following gains than losses. Accordingly, the inverse of the Win-Stay Index can be interpreted as an index of loss-chasing behavior, such that smaller values of the Win-Stay Index were indicative of relatively more risky choices following risky losses than gains.

The second criterion was a measure of sensitivity to loss magnitude. A Loss Sensitivity Index was derived as dividing the proportion of risky choices following R-1 outcomes by that following R-0 outcomes (i.e., R-1 / R-0). Here, larger values indicated greater sensitivity to loss magnitude. Accordingly, larger values were also indicative of increasingly different rates of loss-chasing behavior following differential losses. In these models, the indices were separately regressed on α_G , α_L , and the $\alpha_L:\alpha_G$ ratio to determine how win-stay behavior and loss sensitivity were related to gain- and loss-based learning rates, as well as to the relative rate of learning from gains versus losses. For each index, the models involved one of the three aforementioned model parameter measures, feedback group, and the corresponding interaction to determine whether feedback group moderated the relationship between the modelled parameters and behavioral indices of local risky choice behavior. The main effects and interaction were entered

simultaneously into the model. Feedback group was effect coded with Normal-Feedback/Extra-Feedback as -1/+1, and the model parameters were mean-centered.

The top-left panel of Figure 2.15 shows the relationship between α_G and win-stay behavior. There was no main effect of feedback group, $b = 0.03$, $t(20) = 0.64$, $p = .529$, no main effect of α_G , $b = 0.06$, $t(20) = 0.89$, $p = .382$, and no Feedback Group \times α_G interaction, $b = -0.13$, $t(20) = -1.92$, $p = .069$. The top-middle panel of Figure 2.15 shows the relationship between α_L and win-stay behavior. Win-stay behavior significantly increased with α_L , $b = 0.18$, $t(20) = 2.26$, $p = .035$, but there was no main effect of Group, $b = 0.03$, $t(20) = 0.66$, $p = .516$, and no Group \times α_L interaction, $b = -0.08$, $t(20) = -1.01$, $p = .326$. The significant effect of α_L indicates that the more the rats learned from losses, the more likely they were to make risky choices after R-11 outcomes than to make risky choices after R-0 outcomes. Accordingly, the less that the rats learned from losses, the smaller the ratio of risky choice behavior after R-11 gains to risky choice behavior after R-0 losses. This latter statement implies that the lower loss-based learning rates were associated with a lower likelihood of following R-11 gains than R-0 losses with risky choices; the lower the likelihood, the more that the rats chased R-0 losses relative to how often they stayed on the risky choice after receiving R-11 gains. Lastly, the top-right panel of Figure 2.15 shows the relationship between $\alpha_L:\alpha_G$ and win-stay behavior. There was no main effect of feedback group, $b = 0.03$, $t(20) = 0.71$, $p = .483$, no main effect of $\alpha_L:\alpha_G$, $b = 0.03$, $t(20) = 0.53$, $p = .605$, and no Feedback Group \times $\alpha_L:\alpha_G$ interaction, $b = 0.03$, $t(20) = 0.51$, $p = .617$.

The bottom-left panel of Figure 2.15 shows the relationship between α_G and loss sensitivity. There was no main effect of feedback group, $b = 0.01$, $t(20) = 0.20$, $p = .844$, no main effect of α_G , $b = -0.003$, $t(20) = -0.06$, $p = .952$, and no Feedback Group \times α_G interaction, $b = -0.04$, $t(20) = -0.83$, $p = .418$. The bottom-middle panel of Figure 2.15 shows the relationship

between α_L and loss sensitivity. There was no main effect of feedback group, $b = 0.002$, $t(20) = 0.09$, $p = .926$, but loss sensitivity significantly increased with α_L , $b = 0.22$, $t(20) = 6.67$, $p < .001$, and there was a significant Feedback Group \times α_L interaction, $b = -0.10$, $t(20) = -3.03$, $p = .007$, with Group Extra-Feedback exhibiting a shallower slope (slope = 0.12) than Group Normal-Feedback (slope = 0.32). This result suggests that loss-based learning rates significantly predicted loss sensitivity, and that the strength of this relationship was lower in Group Extra-Feedback, who may have exhibited loss-processing or loss-based-learning deficits due to the more similar presentations of all risky outcomes compared to Group Normal-Feedback. Indeed, this reduced sensitivity to losses is paralleled by the shallower relationship between risky outcome magnitude and goal-tracking described above (Figures 2.9 and 2.10). The bottom-right panel of Figure 2.15 shows the relationship between $\alpha_L:\alpha_G$ and loss sensitivity. Loss sensitivity significantly increased with $\alpha_L:\alpha_G$, $b = 0.09$, $t(20) = 2.72$, $p = .013$, but there was no main effect of feedback group, $b = 0.01$, $t(20) = 0.35$, $p = .733$, and no Feedback Group \times $\alpha_L:\alpha_G$ interaction, $b = -0.04$, $t(20) = -1.15$, $p = .262$.

The Asymmetric RL results provide key insight into the mechanisms governing individual differences in risky choice and loss-chasing behavior. Specifically, individual differences in gain- and loss-based learning rates seem to primarily predict those in risky choice and loss-chasing behavior. As revealed in Figure 2.14 (left panel), individuals who made more risky choices were also those that learned more readily from outcome gains, α_G ; theoretically, for the individuals with greater gain-based learning rates, the expected value of the risky choice would thus increase at a greater rate following R-11 outcomes, therefore increasing the likelihood of future risky choices. Furthermore, comparisons of the fitted RL parameters with local risky choice behavior indicated that individual differences in α_L were related to relative

staying behaviors following gains versus losses (Figure 2.15, top-middle panel). Specifically, rats that were more likely to follow R-0 than R-11 outcomes with risky choices were those rats that less readily learned from losses, suggesting that deficits in loss-based learning may contribute to elevated loss-chasing behaviors in pathological gamblers (Clark, Liu, et al., 2013). Similarly, rats that more readily learned from losses also showed greater discrimination in behavior following differential risky losses (Figure 2.15, bottom-middle panel). Notably, the Feedback Group \times α_L interaction on the Loss-Sensitivity Index complemented the results from the choice and goal-tracking analyses. Specifically, Group Extra-Feedback exhibited a shallower relationship between α_L and the Loss-Sensitivity Index, suggesting that the differential feedback may have impaired rats' abilities to adapt their risky behaviors following differential losses. In other words, the augmented feedback experienced by Group Extra-Feedback may have partially prohibited the rats' abilities to translate their learning from losses to corresponding subsequent behaviors. Thus, the differentiability of R-0 and R-1 losses by Group Extra-Feedback was seemingly limited by the extra multimodal feedback that was experienced with risky losses.

Interim summary

The present experiment investigated the effect of differential risky outcome feedback on molar choice, molecular choice, the influence of past outcomes and choices on subsequent choice, and goal-tracking behavior. Overall, the addition of extra multimodal stimuli to risky losses in Group Extra-Feedback increased risky choice behavior, especially following previous risky outcomes, suggesting that the accompanying multimodal stimuli with risky outcomes specifically affected choice behavior following these outcomes (Figure 2.5). Moreover, choice behavior in Group Extra-Feedback, compared to Group Normal-Feedback, was more strongly affected by more recent outcomes/choices than by more temporally distant outcomes/choices,

suggesting that the extra multimodal stimuli decreased the extent to which past gains and losses affected future risky choice (Figures 2.7 and 2.8). Indeed, the increased goal-tracking behavior in Group Extra-Feedback (Figure 2.9) may be explained by this shortened temporal window of previous outcome/choice influence. The steeper decay rate of outcome and choice histories in Group Extra-Feedback may imply that the stimuli associated with these choice and outcome events did little to affect future behavior after only a few more trials. Similarly, the reduced salience of losses given the consistent multimodal feedback across risky outcomes may have produced both the greater decay in outcome/choice influence, as well as the shallower sensitivity of goal-tracking behavior to risky outcome magnitude. Thus, the extra multimodal stimuli may have increased risky choice and goal-tracking behaviors by constraining the effects of past losses on future choice behavior, such that the individual rats in Group Extra-Feedback were more sensitive to the frequency of gain-related stimuli than risky outcome magnitude relative to Group Normal-Feedback.

The empirical results were paralleled by the fitting of multiple RL algorithms to the rats' individual trial data to evaluate the psychological mechanisms governing individual differences in risky choice behavior across the two feedback groups. The Asymmetric RL algorithm emerged as the best model, suggesting that individual differences in risky choice in the current task were most effectively captured by assuming differential gain- and loss-based learning rates. The rats exhibited greater loss-based than gain-based learning rates, but there were no overall group differences in the fitted Asymmetric RL parameters. However, in conjunction with the proposed reduced effect of past risky losses on risky choice and goal-tracking behavior in Group Extra-Feedback, there was a significant interaction between feedback group and loss-based learning rates on loss sensitivity. Specifically, Group Extra-Feedback exhibited a shallower

relationship between loss-based learning and loss sensitivity, suggesting that the consistent multimodal stimuli across all risky outcomes produced a deficit in processing risky losses. The employment of such complex multimodal stimuli by casinos may encourage risky decision making by impairing individuals' ability to incorporate aversive loss-type outcomes into their estimated value of differential actions/behaviors in their environments. Indeed, individuals tend to overestimate how often they "won" when LDWs are present in the environment (M. J. Dixon et al., 2015), and problem gamblers have been shown to overestimate "win" frequency relative to non-problem gamblers (M. J. Dixon et al., 2014). Individuals who become pathological gamblers may be more susceptible to continue gambling when LDWs are present, thereby leading to overestimates of "winning" and consequent continuations of gambling behavior.

Table 2.1 Probabilities of the risky-zero (R-0), R-1, and R-11 outcomes in the P[0] and P[1] conditions in each phase of Experiment 1 for Groups Extra-Feedback and Normal-Feedback. Half of each group experienced one of two probability conditions across the two phases of Experiment 1. Across all phases, the certain-two (C-2) and C-4 outcomes were delivered following certain choices with probabilities of .50.

Group	Order	Phase 1	Phase 2
Extra-Feedback (n=12)	1 (n=6)	P[0] R-0 = .10, .90, .50, .10 R-1 = .45, .05, .25, .45 R-11 = .45, .05, .25, .45	P[1] R-0 = .05, .25, .45 R-1 = .90, .50, .10 R-11 = .05, .25, .45
		P[1] R-0 = .45, .05, .25, .45 R-1 = .10, .90, .50, .10 R-11 = .45, .05, .25, .45	P[0] R-0 = .90, .50, .10 R-1 = .05, .25, .45 R-11 = .05, .25, .45
	1 (n=6)	P[0] R-0 = .10, .90, .50, .10 R-1 = .45, .05, .25, .45 R-11 = .45, .05, .25, .45	P[1] R-0 = .05, .25, .45 R-1 = .90, .50, .10 R-11 = .05, .25, .45
		P[1] R-0 = .45, .05, .25, .45 R-1 = .10, .90, .50, .10 R-11 = .45, .05, .25, .45	P[0] R-0 = .90, .50, .10 R-1 = .05, .25, .45 R-11 = .05, .25, .45

Table 2.2 Mixed-effects model output for the molar analyses of Experiment 1. Note: continuous variables were mean-centered, and categorical variables were effect coded with Group Normal-Feedback/Extra-Feedback (Feedback Group) as -1/+1.

Factor	Estimate	T (158008)	P	95% CI
Intercept	-2.18	-11.95	< .001	[-2.54, -1.83]
Feedback Group	0.28	1.55	.122	[-0.08, 0.64]
Phase	-0.12	-0.43	.671	[-0.67, 0.43]
Sub-Phase	0.67	8.61	< .001	[0.51, 0.82]
Session	-0.07	-4.94	< .001	[-0.10, -0.04]
Feedback Group × Phase	0.25	0.86	.376	[-0.30, 0.79]
Feedback Group × Sub-Phase	0.04	0.52	.604	[-0.11, 0.19]
Phase × Sub-Phase	-0.33	-1.52	.130	[-0.75, 0.10]
Feedback Group × Session	0.01	0.64	.524	[-0.02, 0.04]
Phase × Session	-0.04	-5.92	< .001	[-0.05, -0.03]
Sub-Phase × Session	0.16	55.85	< .001	[0.15, 0.16]
Feedback Group × Phase × Sub-Phase	0.19	0.88	.378	[-0.23, 0.61]
Feedback Group × Phase × Session	0.04	6.18	< .001	[0.03, 0.05]
Feedback Group × Sub-Phase × Session	0.02	8.00	< .001	[0.02, 0.03]
Phase × Sub-Phase × Session	0.03	4.60	< .001	[0.02, 0.04]
Feedback Group × Phase × Sub-Phase × Session	0.06	10.87	< .001	[0.05, 0.08]

Table 2.3 Mixed-effects model output for the molecular analyses of Experiment 1 evaluating the effect of the previous outcome on subsequent choice behavior. Note: continuous variables were mean-centered, and categorical variables were effect coded with Group Normal-Feedback/Extra-Feedback (Feedback Group) as -1/+1. The reference level of previous outcome was the certain-two (C-2) outcome. C-4 = certain choice – 4-pellet outcome; R-0 = risky choice – 0-pellet outcome; R-1 = risky choice – 1-pellet outcome; R-11 = risky choice – 11-pellet outcome

Factor	Estimate	T (156318)	P	95% CI
Intercept	-0.89	-8.90	< .001	[-1.09, -0.69]
Feedback Group	0.22	2.18	.030	[0.02, 0.41]
Phase	-0.05	-0.25	.806	[-0.47, 0.36]
Probability	-1.05	-22.95	< .001	[-1.14, -0.96]
C-4	-2.22	-15.41	< .001	[-2.50, -1.94]
R-0	1.29	13.43	< .001	[1.10, 1.48]
R-1	1.30	12.77	< .001	[1.10, 1.50]
R-11	1.66	12.48	< .001	[1.40, 1.92]
Feedback Group × Phase	0.18	0.83	.407	[-0.24, 0.59]
Feedback Group × Probability	0.10	2.42	.016	[0.02, 0.19]
Phase × Probability	0.30	3.29	.001	[0.12, 0.48]
Feedback Group × C-4	-0.06	-0.44	.657	[-0.35, 0.22]
Feedback Group × R-0	-0.03	-0.28	.779	[-0.21, 0.16]
Feedback Group × R-1	0.05	0.49	.627	[-0.15, 0.24]
Feedback Group × R-11	0.23	1.74	.083	[-0.03, 0.49]
Phase × C-4	-0.04	-0.92	.356	[-0.13, 0.05]
Phase × R-0	0.06	0.89	.374	[-0.07, 0.19]
Phase × R-1	0.10	1.47	.143	[-0.03, 0.23]
Phase × R-11	-0.25	-2.98	.003	[-0.41, -0.08]
Probability × C-4	-0.05	-0.74	.461	[-0.17, 0.08]
Probability × R-0	-0.26	-2.75	.006	[-0.45, -0.08]
Probability × R-1	-0.22	-2.27	.023	[-0.41, -0.03]
Probability × R-11	0.33	2.74	.006	[0.09, 0.57]
Feedback Group × Phase × Probability	-0.23	-3.32	.001	[-0.36, -0.09]

Factor	Estimate	T (156318)	P	95% CI
Feedback Group × Phase × C-4	-0.01	-0.32	.749	[-0.09, 0.07]
Feedback Group × Phase × R-0	-0.13	-2.15	.031	[-0.25, -0.01]
Feedback Group × Phase × R-1	-0.14	-2.24	.025	[-0.25, -0.02]
Feedback Group × Phase × R-11	0.22	3.75	< .001	[0.11, 0.34]
Feedback Group × Probability × C-4	-0.09	-1.46	.144	[-0.21, 0.03]
Feedback Group × Probability × R-0	-0.03	-0.33	.745	[-0.20, 0.14]
Feedback Group × Probability × R-1	0.18	2.06	.039	[0.01, 0.34]
Feedback Group × Probability × R-11	0.18	1.49	.137	[-0.06, 0.41]
Phase × Probability × C-4	-0.21	-1.64	.102	[-0.46, 0.04]
Phase × Probability × R-0	-0.75	-3.89	< .001	[-1.13, -0.37]
Phase × Probability × R-1	0.84	4.27	< .001	[0.45, 1.22]
Phase × Probability × R-11	0.18	0.72	.469	[-0.30, 0.66]

Table 2.4 Mixed-effects model output for the molecular analyses of Experiment 1 evaluating the effect of the previous 9 outcomes on subsequent risky choice behavior.

Factor	Estimate	T (141314)	P	95% CI
Intercept	-0.89	-12.51	< .001	[-1.03, -0.75]
Outcome – Lag 1	0.23	12.37	< .001	[0.19, 0.26]
Outcome – Lag 2	0.11	15.27	< .001	[0.10, 0.13]
Outcome – Lag 3	0.07	11.88	< .001	[0.06, 0.08]
Outcome – Lag 4	0.06	10.89	< .001	[0.05, 0.07]
Outcome – Lag 5	0.04	9.29	< .001	[0.03, 0.05]
Outcome – Lag 6	0.04	8.53	< .001	[0.03, 0.05]
Outcome – Lag 7	0.04	7.37	< .001	[0.03, 0.05]
Outcome – Lag 8	0.04	8.04	< .001	[0.03, 0.04]
Outcome – Lag 9	0.05	9.47	< .001	[0.04, 0.06]

Table 2.5 Mixed-effects model output for the molecular analyses of Experiment 1 evaluating the effect of the previous 10 choices on subsequent risky choice behavior.

Factor	Estimate	T (141313)	P	95% CI
Intercept	-0.48	-13.62	< .001	[-0.54, -0.41]
Choice – Lag 1	1.28	20.24	< .001	[1.16, 1.41]
Choice – Lag 2	0.37	11.81	< .001	[0.30, 0.43]
Choice – Lag 3	0.17	8.02	< .001	[0.13, 0.21]
Choice – Lag 4	0.17	6.76	< .001	[0.12, 0.22]
Choice – Lag 5	0.11	4.67	< .001	[0.06, 0.15]
Choice – Lag 6	0.15	7.00	< .001	[0.11, 0.19]
Choice – Lag 7	0.14	5.16	< .001	[0.09, 0.20]
Choice – Lag 8	0.09	3.85	< .001	[0.05, 0.14]
Choice – Lag 9	0.17	7.57	< .001	[0.13, 0.22]
Choice – Lag 10	0.17	10.02	< .001	[0.14, 0.21]

Table 2.6 Mixed-effects model output for the goal-tracking analysis of Experiment 1 evaluating the effect of phase, feedback group, and risky outcome magnitude on goal-tracking behavior following the onset of risky outcome delivery. Note: continuous variables were mean-centered, and categorical variables were effect coded with Group Normal-Feedback/Extra-Feedback (Feedback Group) as -1/+1.

Factor	Estimate	T (30110)	P	95% CI
Intercept	0.37	8.11	< .001	[0.28, 0.46]
Feedback Group	0.12	3.68	< .001	[0.06, 0.19]
Phase	-0.03	-0.54	.591	[-0.16, 0.09]
Risky Outcome Magnitude	0.06	19.12	< .001	[0.05, 0.06]
Feedback Group × Risky Outcome Magnitude	-0.01	-3.31	.001	[-0.02, -0.004]
Phase × Risky Outcome Magnitude	0.01	2.25	.025	[0.001, 0.02]

Table 2.7 Summary statistics of the best-fitting parameter estimates and goodness-of-fit indices of the Asymmetric Reinforcement Learning (RL) model. AIC = Akaike Information Criterion; ω^2 = omega-squared.

	α_G	α_L	β	AIC	Pseudo R^2 (ω^2)
Minimum	.0003	.0076	0.9636	2491.1	.07
25 th Percentile	.0010	.0099	1.6041	3351.3	.32
Median	.0017	.0142	2.1505	4171.1	.49
75 th Percentile	.0027	.0255	2.6080	5465.4	.70
Maximum	.0073	.0795	4.4161	7157.4	.87
Mean	.0022	.0209	2.2833	4409.5	.50
Standard Deviation	.0017	.0169	0.9489	1336.4	.24

Figure 2.1 External receptacle for receiving food pellets delivered from the alternative feeder for Group Extra-Feedback.



Figure 2.2 Proportion of choices for the risky outcome for Groups Extra-Feedback and Normal-Feedback as a function of session, phase (P[0], P[1]), and the probability of the R-0 and R-1 outcomes in the P[0] and P[1] phases (i.e., sub-phase), respectively. The $P[0] | P[1] = .10$ data are shown separately for Phase 1 and Phase 4.

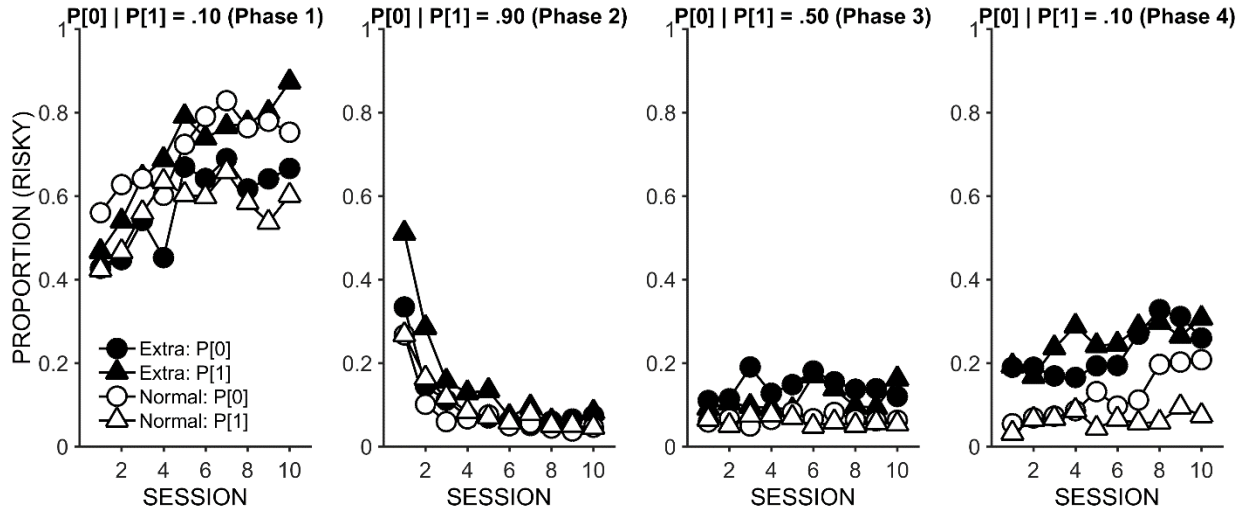


Figure 2.3 Proportion of choices for the risky outcome as a function of session and sub-phase (i.e., probability of manipulated outcome) for individual rats in the P[0] phase. The first six rats of Groups Normal-Feedback and Extra-Feedback experienced the P[0] phase prior to the P[1] phase, while the remaining rats experienced the P[1] phase prior to the P[0] phase. The dot-dash (P[0] = .10 (Data: P1)), dashed (P[0] = .90 (Data: P2)), dotted (P[0] = .50 (Data: P3)), and solid lines (P[0] = .10 (Data: P4)) represent the individual subject-fits from the best-fitting generalized linear mixed effects model. The rats in Group Normal-Feedback are labeled as “N.” and the rats in Group Extra-Feedback are labeled as “E.”.

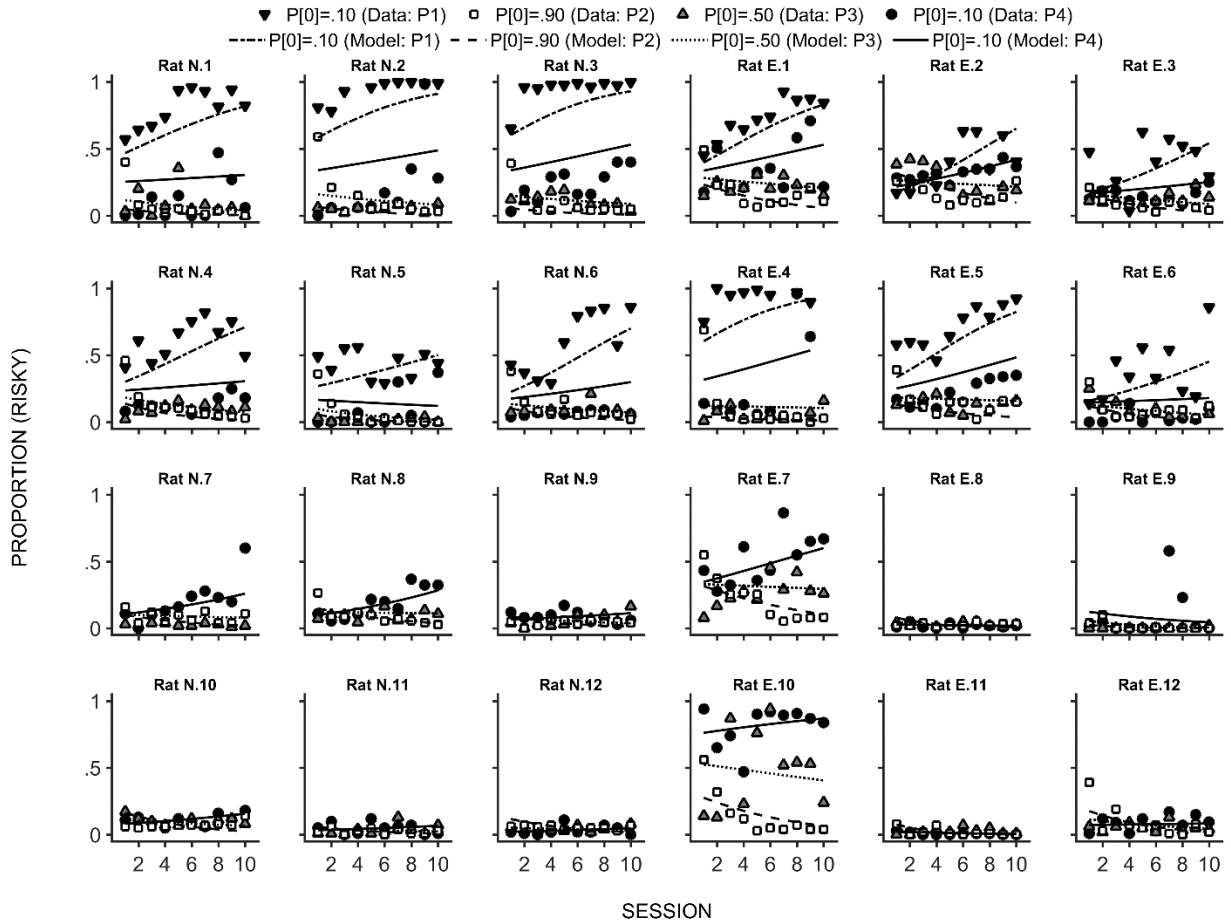


Figure 2.4 Proportion of choices for the risky outcome as a function of session and sub-phase (i.e., probability of manipulated outcome) for individual rats in the P[1] phase. The first six rats of Groups Normal-Feedback and Extra-Feedback experienced the P[0] phase prior to the P[1] phase, while the remaining rats experienced the P[1] phase prior to the P[0] phase. The dot-dash ($P[1] = .10$ (Data: P1)), dashed ($P[1] = .90$ (Data: P2)), dotted ($P[1] = .50$ (Data: P3)), and solid lines ($P[1] = .10$ (Data: P4)) represent the individual subject-fits from the best-fitting generalized linear mixed effects model. The rats in Group Normal-Feedback are labeled as “N.” and the rats in Group Extra-Feedback are labeled as “E.”.

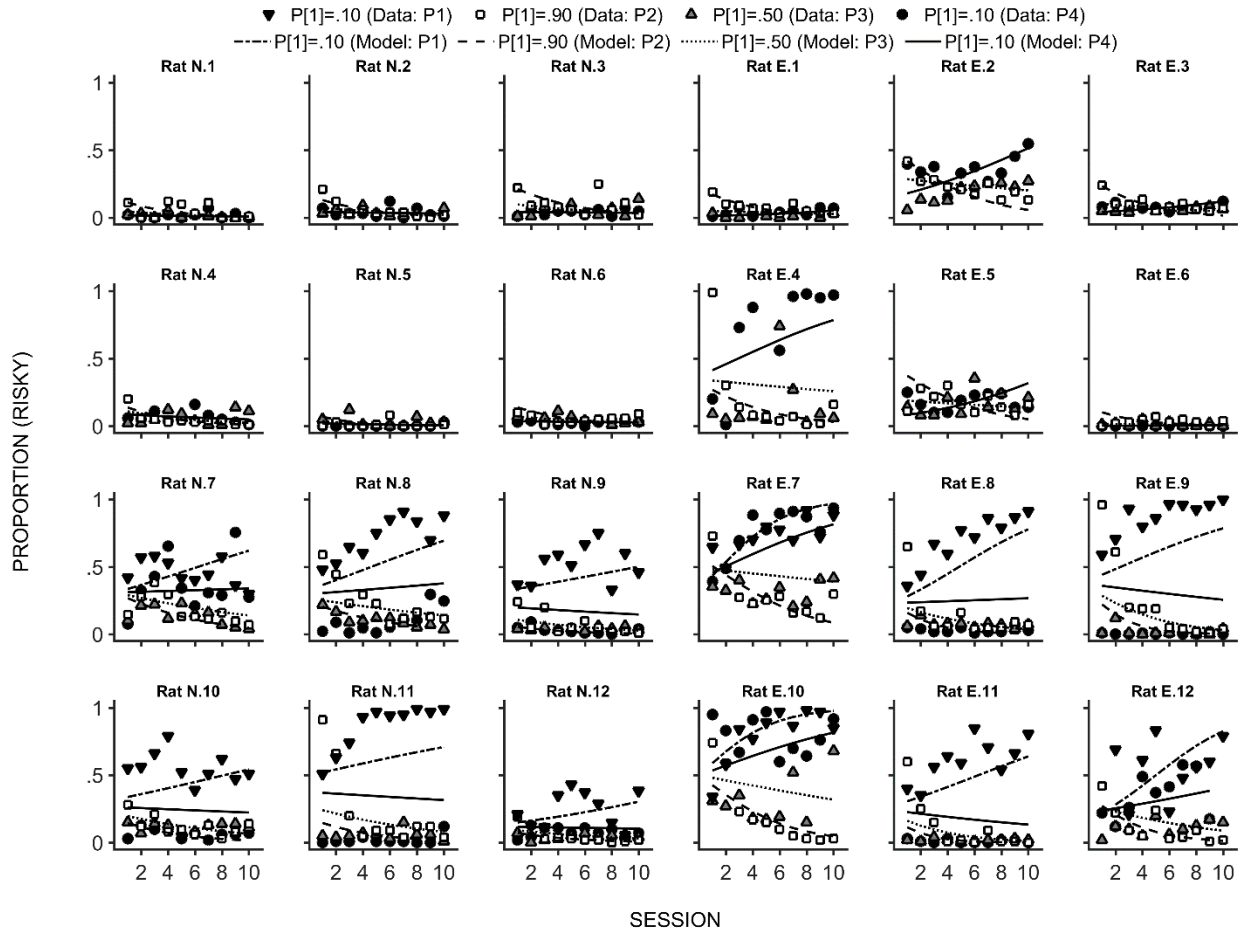


Figure 2.5 Left: Proportion of choices for the risky outcome for Groups Extra-Feedback and Normal-Feedback as a function of phase (P[0], P[1]) and the outcome of the previous choice. Right: Proportion of choices for the risky outcome for Groups Extra-Feedback and Normal-Feedback as a function of probability of the manipulated outcome and the outcome of the previous choice. The dotted lines connecting the data points are intended to demonstrate the relationship between risky choice behavior following differential outcomes rather than to imply a continuous relationship. C-2 = certain choice – 2 pellets; C-4 = certain choice – 4 pellets; R-0 = risky choice – 0 pellets; R-1 = risky choice – 1 pellet; R-11 = risky choice – 11 pellets.

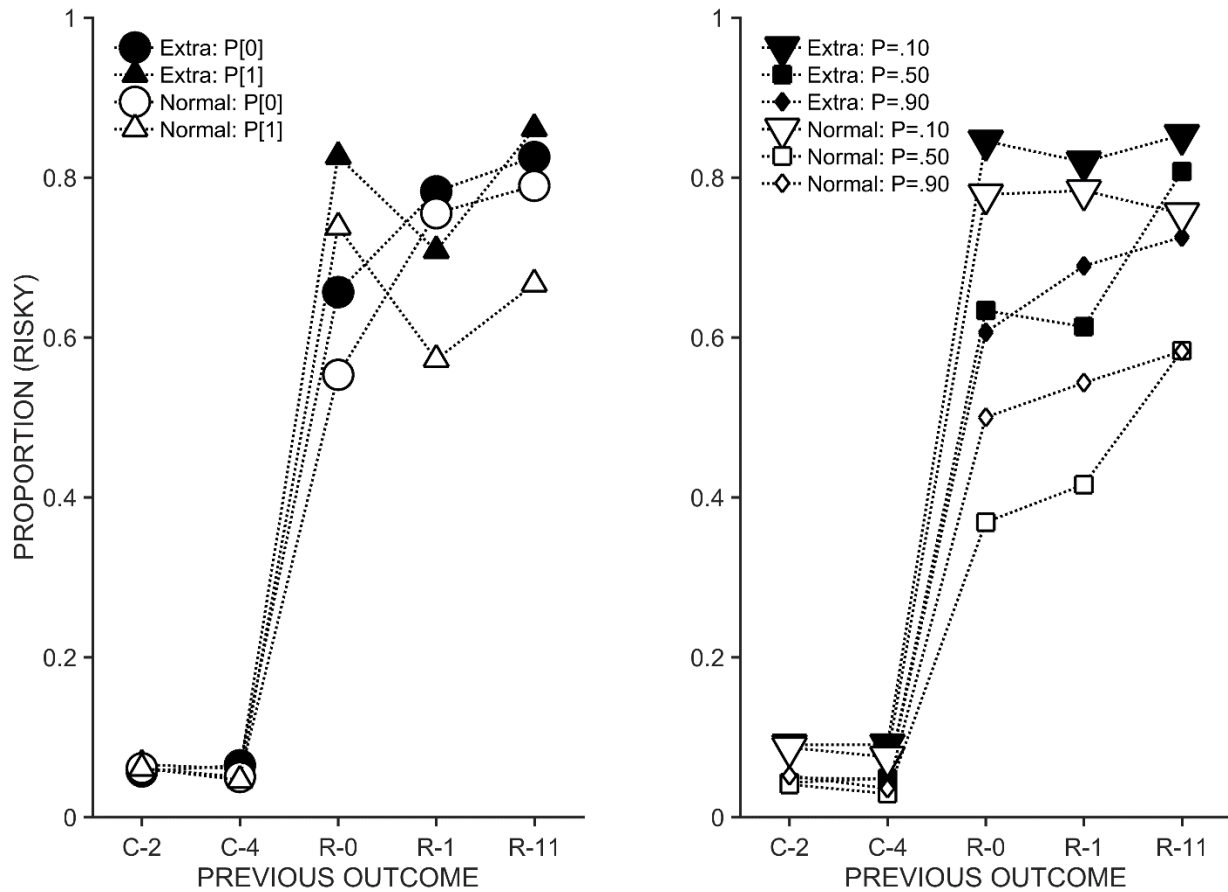


Figure 2.6 Proportion of choices for the risky outcome as a function of the outcome of the previous choice and phase for individual rats. The dashed (P[0]) and dotted lines (P[1]) represent the individual subject-fits from the best-fitting generalized linear mixed effects model. The rats in Group Normal-Feedback are labeled as “N.” and the rats in Group Extra-Feedback are labeled as “E.”.

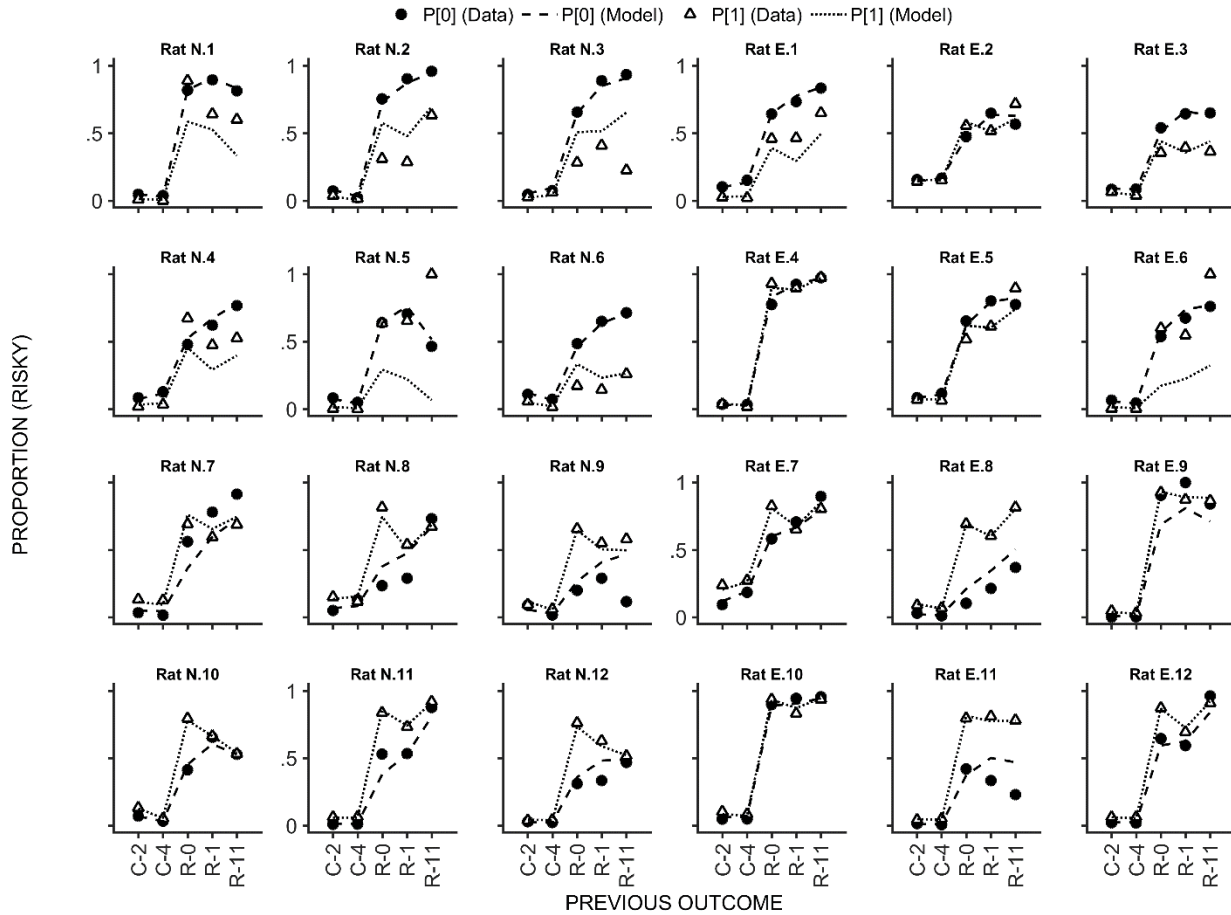


Figure 2.7 Regression coefficients from the mixed-effects analysis evaluating the effect of the previous 9 outcomes on subsequent risky choice. N = Group Normal-Feedback; E = Group Extra-Feedback; Hyp. = hyperbolic model fit; Exp. = exponential model fit.

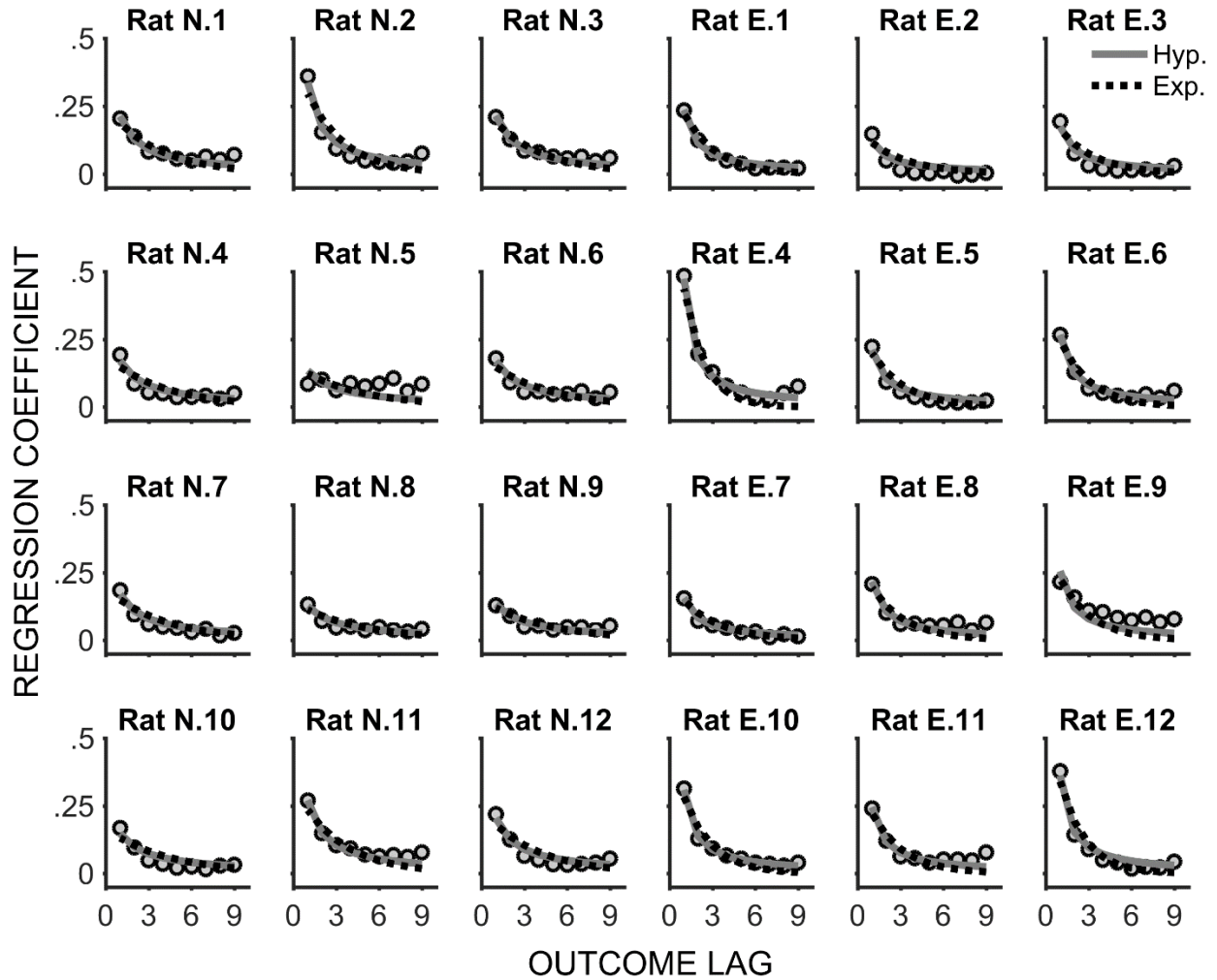


Figure 2.8 Normalized regression coefficients from the mixed-effects analysis evaluating the effect of the previous 10 choices on subsequent risky choice. N = Group Normal-Feedback; E = Group Extra-Feedback; Hyp. = hyperbolic model fit; Exp. = exponential model fit.

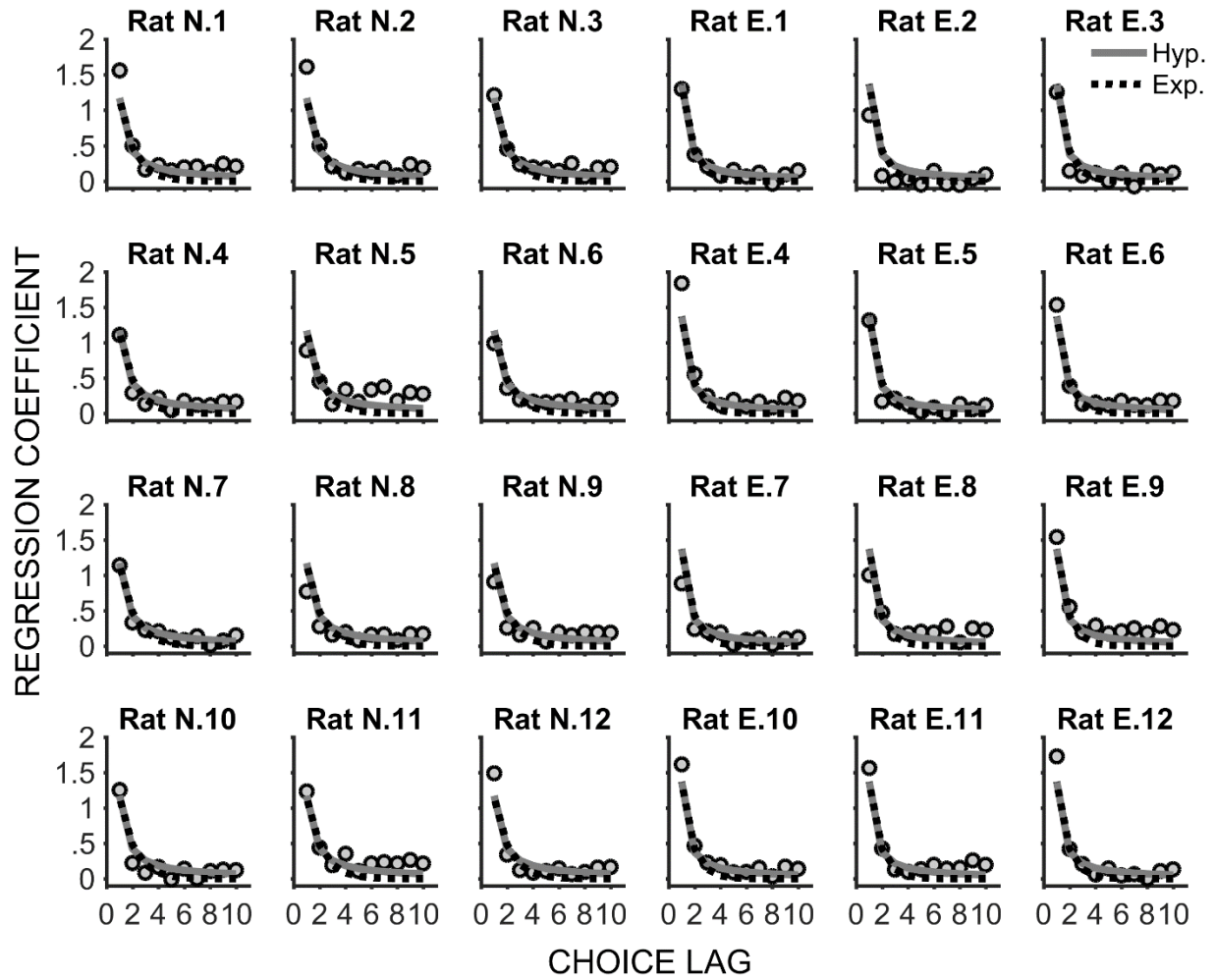


Figure 2.9 Mean number of head entries into the food magazine during the 5-s temporal window following the onset of outcome delivery as a function of risky outcome magnitude and group.

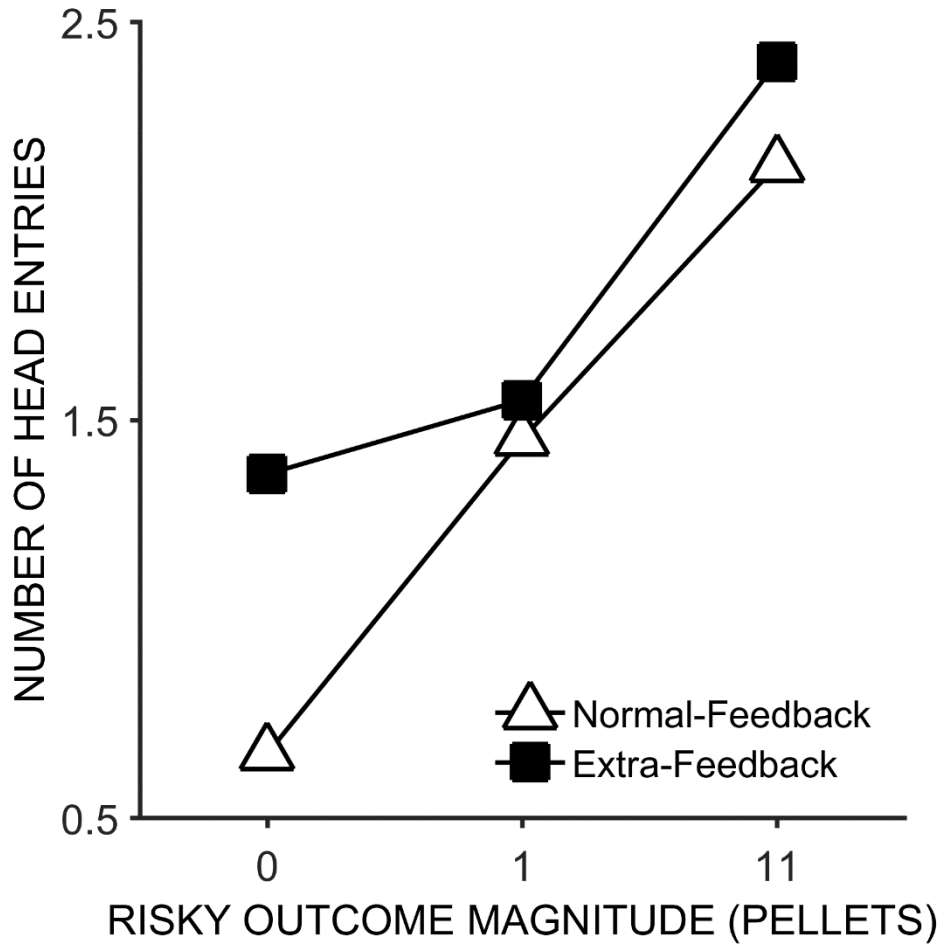


Figure 2.10 Mean number of head entries into the food magazine during the 5-s temporal window following the onset of outcome delivery as a function of risky outcome magnitude and phase for individual rats. The dashed (P[0]) and dotted lines (P[1]) represent the individual subject-fits from the best-fitting generalized linear mixed effects model. The rats in Group Normal-Feedback are labeled as “N.” and the rats in Group Extra-Feedback are labeled as “E.”.

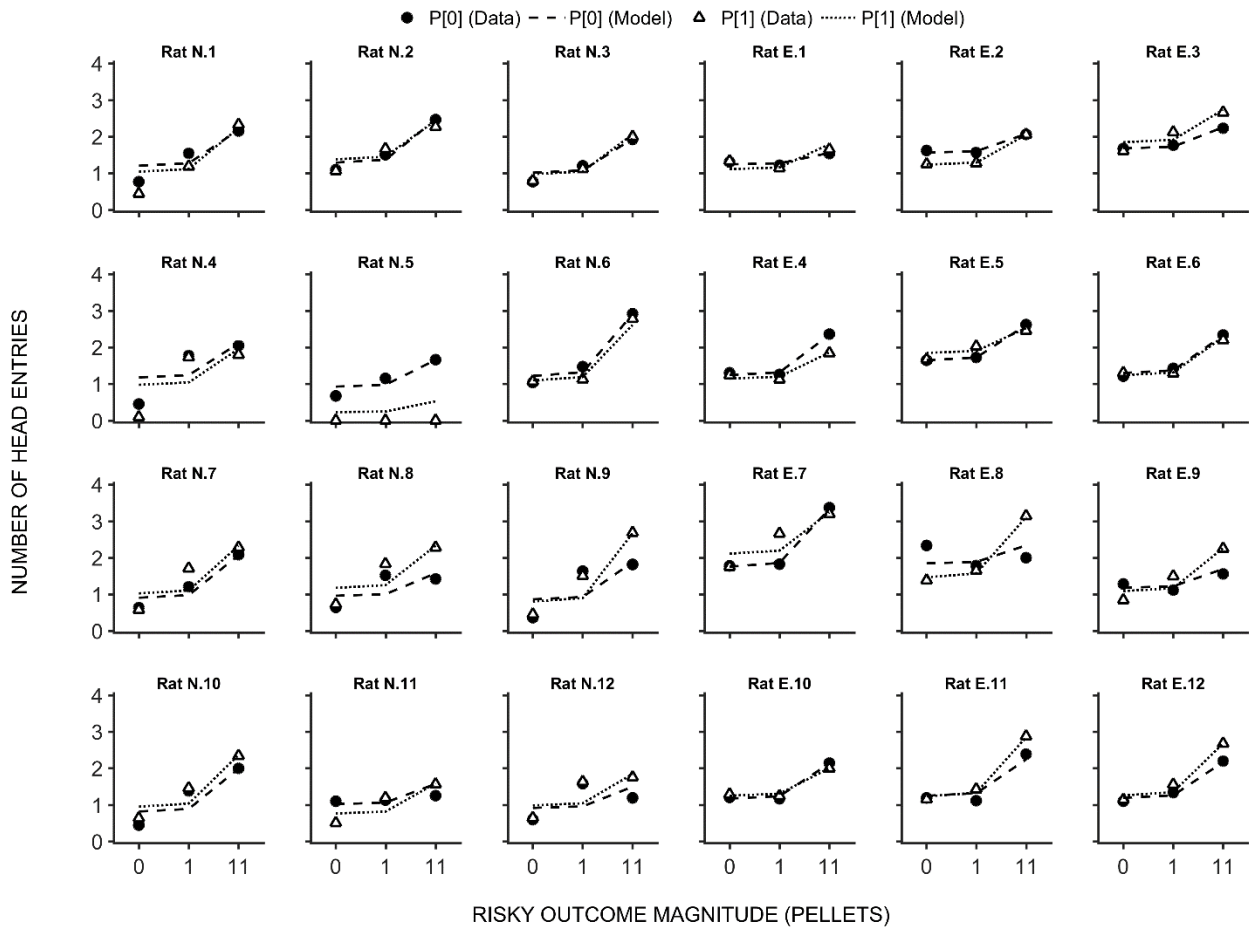


Figure 2.11 Parameter recovery for the Simple, Asymmetric, Valence-Attentive, and Weighted-Reference-Point models in Experiment 1. The thick dark lines are the unit diagonals.

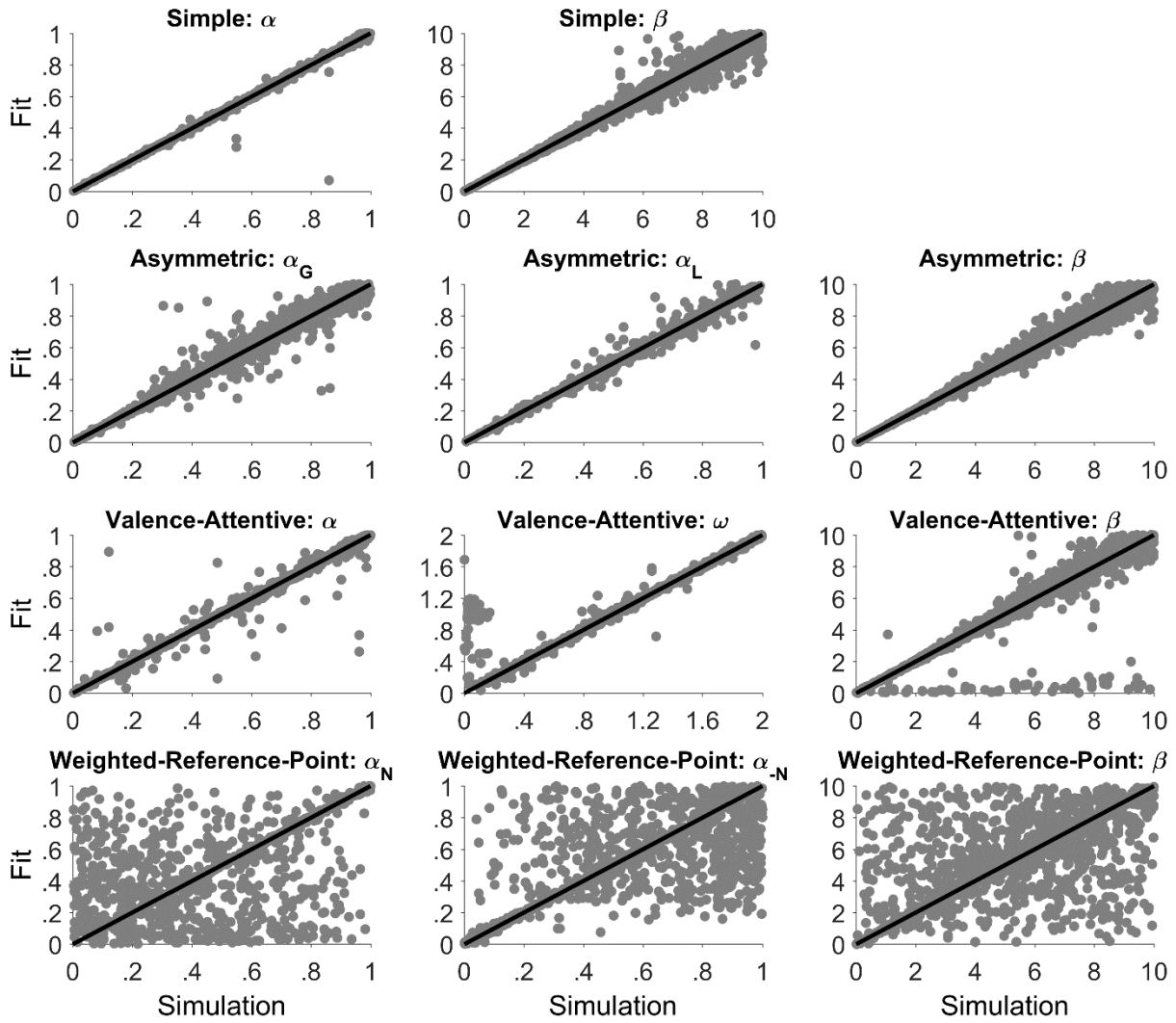


Figure 2.12 Akaike Information Criterion (AIC) for the Asymmetric Reinforcement Learning (RL) model plotted against the AICs of the Simple and Valence-Attentive RL models. Each data point represents the fit to an individual rat's data. Model fits that did not converge were omitted.

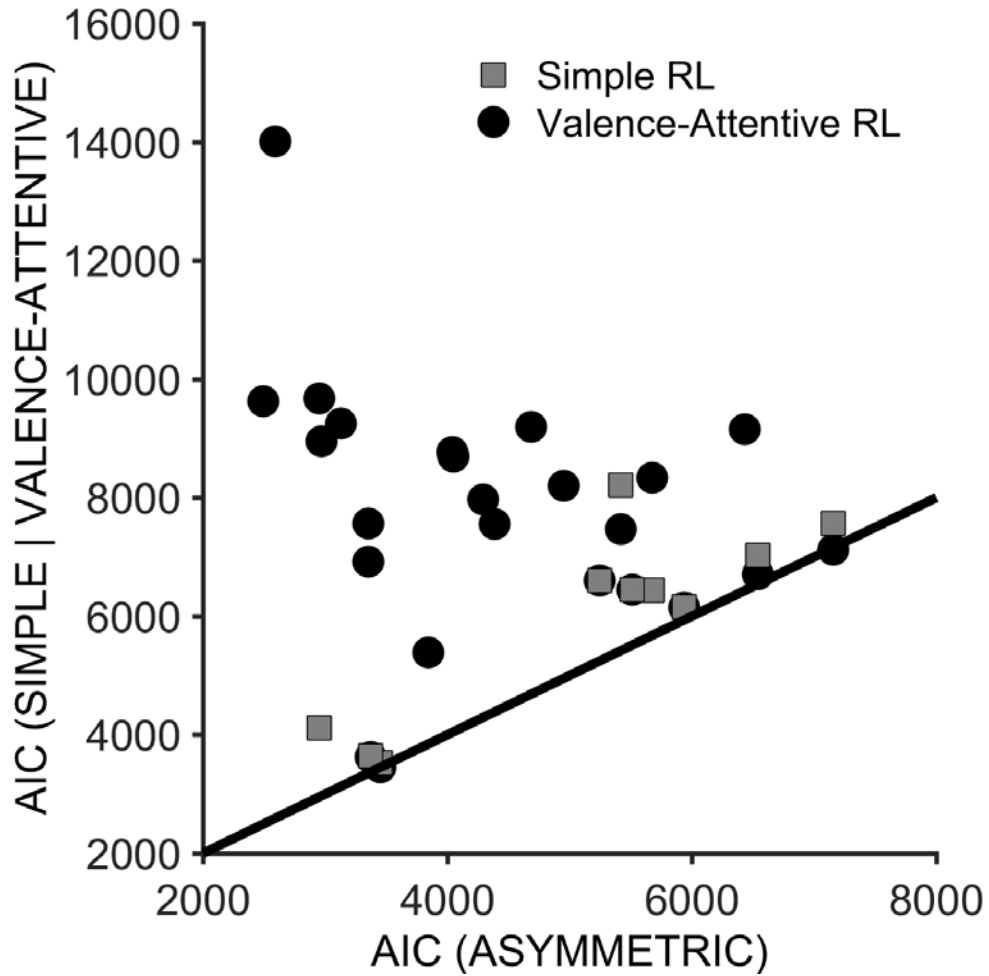


Figure 2.13 Choice data (Data) and the predicted choice data from the Asymmetric RL model (Model) for each rat, ordered from top to bottom by how well the model fit the data in terms of omega-squared (ω^2). “Min” and “Max” refer to the data with the lowest and highest ω^2 , respectively. “25%ile”, “Mdn”, and “75%ile” refer to the 25th, 50th, and 75th percentile of ω^2 , respectively. The alternating gray and white boxes within each panel represent different phases and sub-phases (see Table 2.1). The differential alignments of these phase markers reflect individual differences in trials completed within each phase.

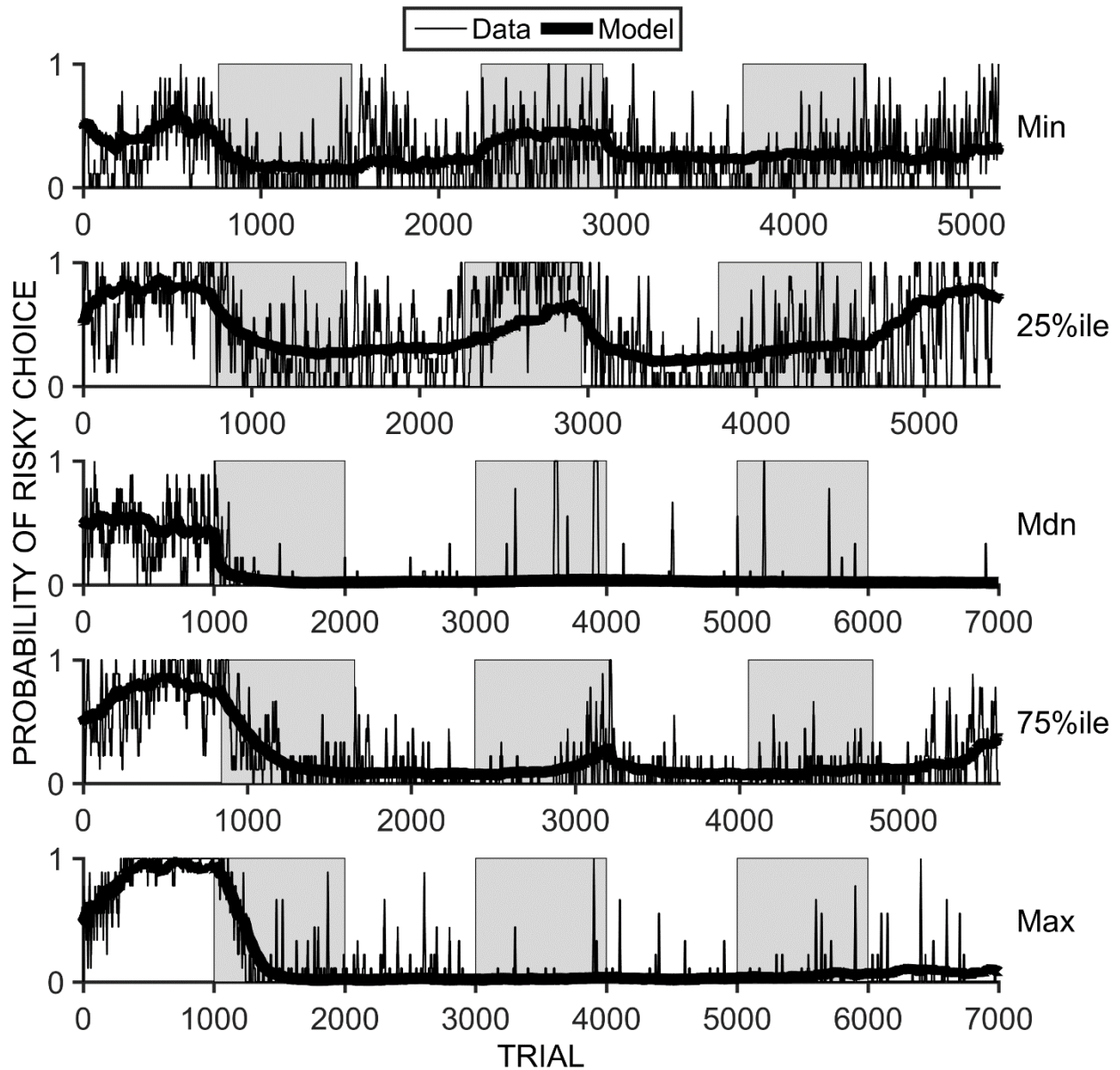


Figure 2.14 Log-transformed mean risky choice plotted against the log-transformed α_G (left), α_L (center), and β parameters (right) as derived from the fitting of the Asymmetric RL model. Each data point represents an individual rat. The goodness-of-fit index of the best fitting regression line for the overall model is shown (R^2).

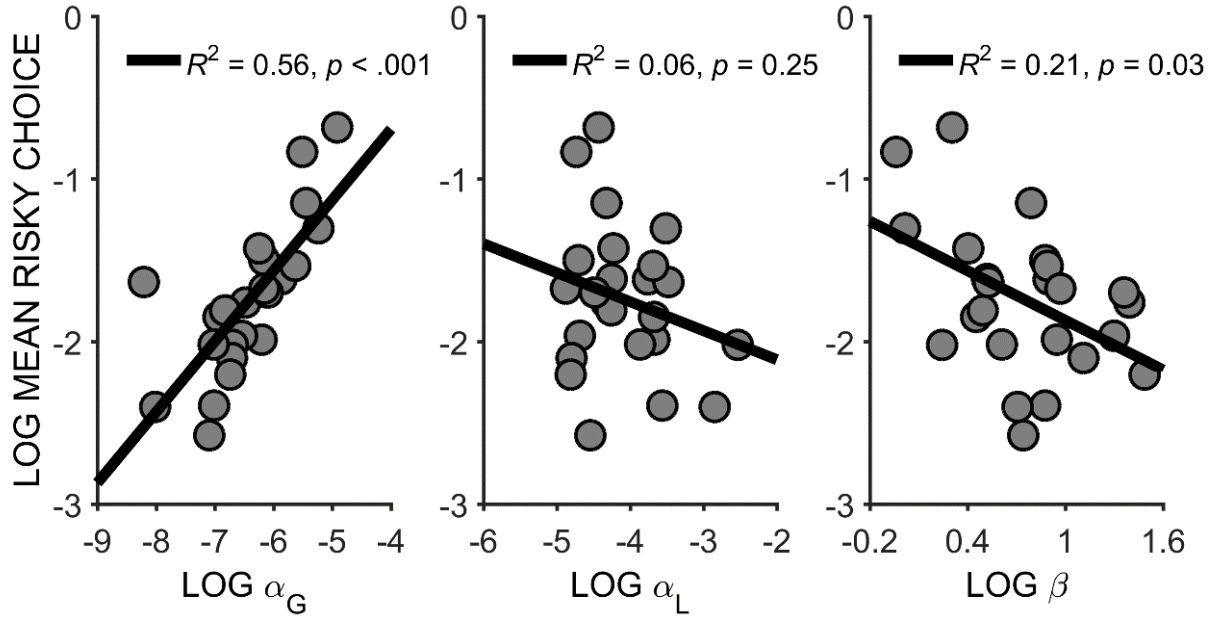
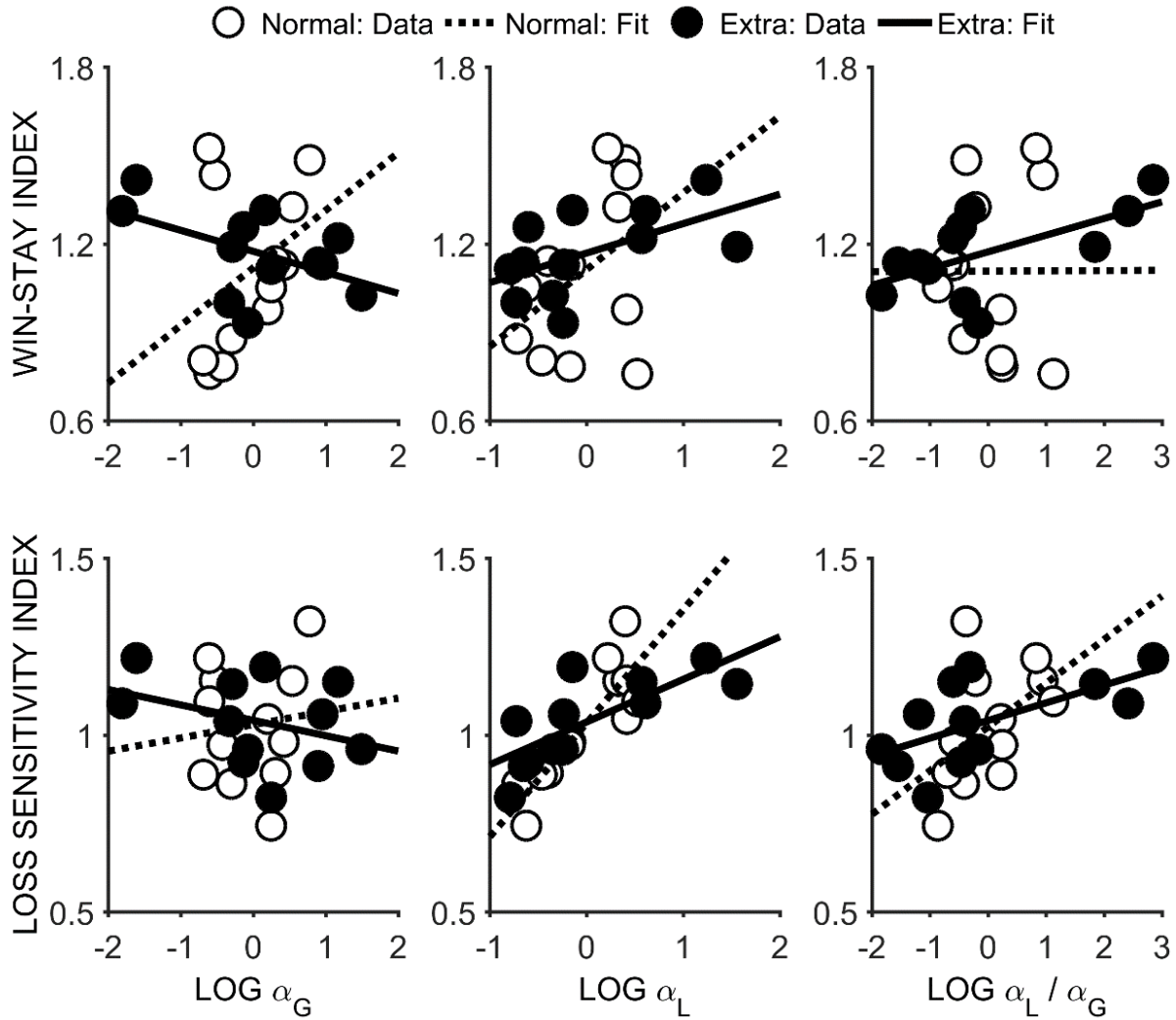


Figure 2.15 Win-Stay (top panels) and Loss Sensitivity indices (bottom panels) plotted against the log-transformed α_G (left panels), α_L (middle panels), and α_L to α_G ratio (right panels) as derived from the fitting of the Asymmetric RL model. Each data point represents an individual rat. The rats in Group Normal-Feedback and Group Extra-Feedback are distinguished.



Chapter 3 - Experiment 2

The results of Experiment 1 revealed that manipulation of feedback and information in terms of miscued risky gains produced elevated rates of risky choice. Ultimately, the general risky choice paradigm in Experiment 1 is parametrically unique relative to traditional analyses of risky choice in rats. In Experiment 1, risky choices probabilistically resulted in two risky losses and one risky gain. However, a more common approach in the literature has been to provide animals (and humans) with the choice between a smaller certain outcome (e.g., 1 pellet) and a single probabilistic larger outcome (e.g., 4 pellets; e.g., Cardinal & Howes, 2005; Floresco, St. Onge, Ghods-Sharifi, & Winstanley, 2008; Stopper & Floresco, 2011), in contrast to the variable non-zero reward magnitudes for both certain and risky choices described above (also see Kirkpatrick, Marshall, Smith, Koci, & Park, 2014; Marshall & Kirkpatrick, 2013, 2015). Alternatively, instead of probabilistic reward omission following a risky choice, other experimental procedures have included risky choice outcomes of variable reward magnitudes without the possibility of reward omission (e.g., Bateson & Kacelnik, 1995; Reboreda & Kacelnik, 1991). Thus, risky choice procedures typically incorporate risk in terms of potential reward omission (Risk-Omission) or non-zero reward variability (Risk-Variability), but not both (but see Kirkpatrick et al., 2014; Marshall & Kirkpatrick, 2013, 2015). Yet, as seen in Experiment 1, the manipulation of the risky loss probability (R-0, R-1) can differentially impact risky decision making on both molar and molecular levels depending on the magnitude of the risky loss that is manipulated.

Despite the similarities between the Risk-Omission and Risk-Variability procedures (i.e., a certain outcome versus a probabilistic larger outcome), the patterns of win-stay/lose-shift behavior are inconsistent across these procedures. In Risk-Omission procedures, animals exhibit

traditional win-stay/lose-shift behavior, exhibiting greater risk-taking behavior following risky wins (or gains) and a relatively reduced likelihood of risk-taking behavior following losses (e.g., Stopper & Floresco, 2011). However, in Risk-Variability procedures, McCoy and Platt (2005) and Hayden and Platt (2009) reported that macaques and humans, respectively, exhibited greater risk-taking behavior following risky outcomes as the previous risky outcome deviated more from the expected value of the risky choice (also see O'Neill & Schultz, 2010); in other words, greater risky losses and greater risky gains were followed by an increased likelihood of subsequent risky choices. Thus, loss-chasing behavior may depend on not just the previous outcome, but also on the parametric structure of the environment (i.e., reward omission, reward variability). Despite repeated experiences, behavior within a Risk-Omission or Risk-Variability paradigm could be fundamentally altered by the addition of a non-zero variable-reward amount or reward omission, respectively. For example, if a Risk-Variability paradigm with risky outcomes of 1 and 11 pellets were modified to include reward omission, the 1-pellet outcome would now be bounded by outcomes of differential feedback (i.e., the absence of feedback in reward omission and the presence of feedback with the 11-pellet outcome). The experience of reward omission may affect the encoding of the 1-pellet outcome, as the feedback of a 1-pellet outcome is now a mixture of loss- and gain-based information from the 0- and 11-pellet outcomes, respectively. The simple inclusion of reward omission may transform the 1-pellet outcome into an LDW as it is a loss relative to expectation, but also accompanied by gain-related stimuli that do not co-occur with reward omission. Such dynamic risk-taking environments may be likened to ticketing and point/demerit systems in regards to driving. For example, the issuance of a warning instead of a citation (i.e., relative win) by a police officer for a moving violation may be encoded

differently depending on whether such warnings in fact appear on the individual's driving record (i.e., potential loss).

Accordingly, there were two goals of the present experiment. The first goal was to determine potential differences in loss-chasing behavior in both Risk-Omission and Risk-Variability procedures. After exposure to each of these environments, the Risk-Variability procedure was modified to include reward omission, and the Risk-Omission procedure was modified to include a smaller non-zero 1-pellet reward. The addition of a second loss (i.e., 1-pellet outcome) to the Risk-Omission procedure introduced a potential LDW to the experimental environment. As described above, the addition of a second loss (i.e., reward omission) to the Risk-Variability procedure may also transform the 1-pellet outcome into an LDW. Accordingly, a 1-pellet outcome in the Risk-Variability procedure may be encoded considerably differently upon the introduction of probabilistic reward omission. Indeed, the mechanisms by which individuals process the combination of loss-based stimuli (e.g., the magnitude of the 1-pellet reward) and gain-based stimuli (e.g., reward delivery being accompanied by feeder operation) has implications for how LDWs are processed in human participants. If the prospect of potential reward omission is in fact necessary for the 1-pellet outcome to be treated as an LDW, then this experiment may differentiate the possible impact of LDWs on loss-chasing behavior when the 1-pellet outcome had already been experienced (i.e., Risk-Variability) or is newly introduced within the environment (i.e., Risk-Omission). Thus, the present experiment manipulated the feedback and information provided by risky losses in terms of the addition of alternative risky outcomes, as Experiment 1 did with the added presentation of multimodal feedback.

As in Experiment 1, multiple reinforcement-learning (RL) models were fit to the data. The parameters of these models were correlated with the rats' behaviors to determine the overall

involvement of the corresponding psychological mechanisms in loss-chasing behavior, providing insight into the explanatory factors related to individual susceptibilities to exhibit loss-chasing and risk-taking behaviors. Translationally, the results are applicable to the development of subject-specific therapies to treat various maladaptive behaviors given the characteristics of the corresponding outcome history of such risky choices.

Method

Animals

Twenty-four experimentally-naive male Sprague-Dawley rats, approximately 21 days of age on arrival, served as subjects. The housing and husbandry conditions were identical to those described in Experiment 1, with the exception that the reverse 12:12 hr light:dark schedule was set such that the lights turned off at approximately 7:00 am.

Apparatus

The experimental apparatus was identical to that described in Experiment 1, except that the external receptacle (Figure 2.1) was not used in this experiment.

Procedure

Magazine and lever-press training. Magazine and lever-press training procedures were identical to those described in Experiment 1. There was one session of magazine training and two sessions of lever-press training.

Risky choice task. The risky choice task was identical to that experienced by Group Normal-Feedback in Experiment 1 with the following exceptions. Table 3.1 depicts the design of this experiment. Each pair of rats was randomly partitioned into one of two groups (Equal-Risk and Unequal-Risk), and the task was divided into two phases of different risk paradigms (Risk-Omission, Risk-Variability), each composed of two sub-phases of different loss conditions

(One-Loss, Two-Loss). Here, One-Loss and Two-Loss refer to the number of risky outcome magnitudes regarded as losses relative to the expected value of both the risky and certain choices.

The One-Loss sub-phase of each risk paradigm occurred before the corresponding Two-Loss sub-phase. In the Risk-Omission/One-Loss sub-phase, risky choices resulted in the probabilistic delivery of 0 (R-0) or 11 food pellets (R-11). In the Risk-Variability/One-Loss sub-phase, risky choices resulted in the probabilistic delivery of 1 (R-1) or 11 food pellets (R-11). For Group Equal-Risk, the probability of each of these outcomes following a risky choice was .50. For Group Unequal-Risk, the probability of the smaller outcome (0 pellets in Risk-Omission, 1 pellet in Risk-Variability) was equal to .67, such that $p(\text{R-11}) = .33$. The four risky forced-choice trials twice resulted in each of the corresponding risky outcome magnitudes. In the Risk-Omission and Risk-Variability/Two-Loss sub-phases for Groups Equal-Risk and Unequal-Risk, risky choices probabilistically resulted in the delivery of R-0, R-1, and R-11 outcomes ($p_s = 0.33$). For the Two-Loss sub-phases, the four risky forced-choice trials twice resulted in each of the non-zero risky outcome magnitudes (R-1, R-11). Thus, while the initial probability conditions were different for Groups Equal-Risk and Unequal-Risk in the One-Loss sub-phases, they were equated in the Two-Loss sub-phases. Half of the rats in each group experienced the Risk-Omission phases (One-Loss \rightarrow Two-Loss) followed by the Risk-Variability phases (One-Loss \rightarrow Two-Loss), and vice versa for the other half of the rats. Each sub-phase lasted for 10 sessions. Each session lasted until all free-choice trials were completed or for approximately 2 hr.

Data analysis

Choice behavior. The choice behavior analyses were identical to those described in Experiment 1 with the following exceptions. One session for two different rats was removed from data analysis due to equipment error; cursory analyses of mean choice behavior across rats before and after each of these sessions suggested that the equipment errors did not considerably impact subsequent choice behavior. Regarding the mixed-effects models, potential fixed effects for the molar analyses included group (Equal-Risk/Unequal-Risk), risk paradigm (Risk-Omission/Risk-Variability), loss condition (One-Loss/Two-Loss), session, and their interactions. Potential random effects included the aforementioned fixed effects except for group. Previous outcome was also a potential fixed and random effect in the molecular analyses. As in Experiment 1, because the purpose of this analysis concerned the effects of the group, risk-paradigm, and loss-condition manipulations on molecular risky choice behavior, the predictor of session was not included in these analyses.

In regards to the analysis of outcome and choice histories, there was a significant correlation between previous outcome and previous choice coding as in Experiment 1, $r_s > .70$, $p_s < .001$. Accordingly, the individual rats' outcome and choice histories were analyzed separately. Equations 2 and 3, which included free parameters for the intercept (A) and decay rate (k), were employed for the outcome history analysis. As in Experiment 1, the individual rats' regression coefficients in the choice history analysis were normalized by each rat's maximum regression coefficient, and the hyperbolic and exponential equations with only the decay rate parameter were fit to these data (Eq. 4 and 5). Lastly, in contrast to the 9 previous outcomes that were used in the outcome history analysis in Experiment 1, the outcome history analysis in the current experiment included the 10 previous outcomes.

Reinforcement learning (RL) models. RL model analyses were identical to those described in Experiment 1.

Results and Discussion

Choice behavior

Molar analysis.

Overall effects. The molar analysis of risky choice was conducted to examine the effects of the loss-condition and risk-paradigm manipulations across groups. Analysis involved 89,923 observations, and included the overall intercept, categorical predictors of group (Unequal-Risk, Equal-Risk) and risk paradigm (Risk-Variability, Risk-Omission), and continuous predictors of loss condition (One-Loss, Two-Loss), and session. The categorical predictors were effect coded with Unequal-Risk/Equal-Risk as -1/+1 and Risk-Variability/Risk-Omission -1/+1. The fixed-effects structure included the full factorial model of Group \times Risk Paradigm \times Loss Condition \times Session. Intercept, risk paradigm, loss condition, session, and Risk Paradigm \times Loss Condition were random effects.

Table 3.2 shows the full model output from the mixed-effects analysis. Group Equal-Risk was significantly more likely to make risky choices relative to Group Unequal-Risk, $t(89907) = 3.38, p = .001$. This main effect was expected, as $p(R-11)$ was greater for Group Equal-Risk than for Group Unequal-Risk across conditions. Accordingly, this effect confirms that the mean group differences were sensitive to the between-groups manipulation.

Analysis also revealed a Risk Paradigm \times Loss Condition interaction, $t(89907) = -2.16, p = .031$. In the Risk-Omission phases, there was a decrease in risky choice from the One-Loss to Two-Loss conditions (slope = -0.21); thus, the addition of a 1-pellet loss increased risk aversion. In contrast, there was an increase in risky choice from the One-Loss to Two-Loss conditions in

the Risk-Variability condition (slope = 0.32), in that the addition of a 0-pellet loss *reduced* risk aversion. These results are counterintuitive, as they imply that risk aversion is decreased when greater losses are added as potential risky outcomes. However, as described below, these results are ably explained in terms of how risky choice behavior changed as a function of session across the One- and Two-Loss sub-phases.

The analysis also indicated that there was a significant Group \times Risk Paradigm \times Loss Condition interaction (Figure 3.1), $t(89907) = 3.15$, $p = .002$ (Table 3.2). Here, there were group differences in how risky choice behavior was affected in the Two-Loss sub-phases depending on whether the rats were in the Risk-Omission or Risk-Variability phases. With the exception of Group Unequal-Risk in the Risk-Omission phases, there were increases in risky choice when the second risky loss was added in the Two-Loss phase. Simple slopes computations revealed positive slopes for Group Equal-Risk in the Risk-Omission (slope = 0.59) and Risk-Variability phases (slope = 0.34), and Group Unequal-Risk in the Risk-Variability phase (slope = 0.30). However, simple slopes computations revealed a decrease in risky choice from the One- to Two-Loss conditions in the Risk-Omission phase (slope = -1.00). As the parametric structure of the environment was identical across groups and risk paradigms in the Two-Loss sub-phases, these results suggest that the comparable risky choice parameters in the Two-Loss sub-phases were not the only determinant of risky choice in those sub-phases.

Figure 3.2 shows the Group \times Risk Paradigm \times Loss Condition \times Session interaction. This interaction was not statistically significant, $t(89907) = -1.89$, $p = .058$, but the patterns in Figure 3.2 do well to illustrate the rats' behavior across groups, risk paradigms, and loss conditions. A mixed-effects model without the Group \times Risk Paradigm \times Loss Condition \times Session interaction had a larger AIC ($\Delta\text{AIC} = 1.57$), but produced similar estimates and effect

sizes as to those produced by the current model. While the similarities in estimates, effect sizes, and AIC supported the simpler model, the current patterns in the data warranted inclusion of the full factorial model to facilitate understanding of the lower-order interactions.

Specifically, in the One-Loss condition, Group Equal-Risk exhibited a greater increase in risky choice as a function of session in the Risk-Variability phase (slope = 0.21) than in the Risk-Omission phase (slope = 0.07). This result can be explained in terms of the expected value of the risky choice, as the risky choice in the Risk-Variability sub-phase had a greater expected value than that in the Risk-Omission sub-phase (see Table 3.1). Similarly, Group Unequal-Risk showed an increase in risky choice with session in the Risk-Variability sub-phase (slope = 0.06), and a decrease in risky choice as a function of session in the Risk-Omission sub-phase (slope = -0.16). This result is not surprising because the expected value of the risky choice for Group Unequal-Risk in the Risk-Omission phase was less than that of the risky choice in the Risk-Variability phase (see Table 3.1). Indeed, the terminal group risky choice behavior was ordered with respect to the expected value of the risky choice, indicating that expected value at least partially contributed to risky choice behavior in the One-Loss sub-phase.

In contrast to the relatively large changes in risky choice as a function of session across groups and risk paradigms in the One-Loss sub-phases, mean risky choice changed relatively little as a function of session in the Two-Loss sub-phases. In the Risk-Variability phase, Group Equal-Risk showed a small increase in risky choice as a function of session (slope = 0.01). However, Group Equal-Risk showed relatively little change in risky choice as a function of session in the Risk-Omission sub-phase (slope = -0.0001). Group Unequal-Risk also showed small decreases in the Risk-Omission (slope = -0.09) and Risk-Variability conditions (slope = -0.05).

These results indicate that while the groups were sensitive to the parameters of the risky choice in the One-Loss sub-phases, the group differences in the Two-Loss sub-phases were primarily driven by terminal behavior in the preceding One-Loss sub-phase. Accordingly, the significant lower-order interactions were seemingly driven by between-groups behavioral differences in the One-Loss sub-phases. Interestingly, as described above, risky choice behavior increased in the Risk-Variability/Two-Loss sub-phase across sessions for Group Equal-Risk, even though this set of conditions accounted for the largest decrease in expected value of the risky choice between the One- and Two-Loss sub-phases. The only set of conditions in which an increase in risky choice with session may have been expected was for Group Unequal-Risk in the Risk-Omission paradigm; here, there was an increase in expected value of the risky choice from the One- to Two-Loss sub-phases. However, this set of conditions accounted for the largest *decrease* in risky choice as a function of session (Figure 3.2). Thus, while the expected value of the risky choice seemed to predict risky choice behavior in the One-Loss sub-phases, the inclusion of a second risky loss in the Two-Loss sub-phases potentially discouraged the use of expected value to dictate overall risky choice (Marshall & Kirkpatrick, 2015). Therefore, rats' behavior in the Two-Loss sub-phases was likely driven by the risky choice parameters (i.e., outcome magnitudes and probabilities) prior to the paradigmatic shift, seemingly indicative of sensitivity to past environmental conditions (also see McNamara, Green, & Olsson, 2006; Stewart, Chater, Stott, & Reimers, 2003). Accordingly, as the expected value of the risky choice was identical in the Two-Loss sub-phases across all conditions, these results may suggest that changes in behavior within and across the One- and Two-Loss conditions were not strictly governed by the corresponding changes in expected value of the risky choice. Alternatively, these data may suggest that the rats were potentially more sensitive to the specific outcomes of

the choices (and their probabilities of occurrence) rather than to the expected value of the choices (Marshall & Kirkpatrick, 2015).

Individual differences. As described above, the mixed-effects model's random-effects structure included intercept, risk paradigm, loss condition, session, and Risk Paradigm \times Loss Condition. Figures 3.3 and 3.4 depict the substantial individual differences in risky choice as a function of risk paradigm, loss condition, and session. The Risk-Omission data are shown in Figure 3.3. During the One-Loss sub-phase, in which risky choices probabilistically resulted in R-0 and R-11 outcomes, some rats showed an increase in risky choice as a function of session (e.g., Rats E.2 and E.4), while other rats showed a decrease in risky choice as a function of session (e.g., Rats U.4 and U.11). Indeed, only three rats in Group Equal-Risk [i.e., $p(\text{R-0}) = p(\text{R-11})$] showed a decrease in risky choice in this phase as a function of session, while all but one rat in Group Unequal-Risk [i.e., $p(\text{R-0}) > p(\text{R-11})$] showed a decrease in risky choice as a function of session. In the Two-Loss phase, in which the probabilities of delivery of R-0, R-1, and R-11 outcomes were equal, the majority of rats in Group Equal-Risk maintained relatively high rates of risky choice behavior, while the majority of rats in Group Unequal-Risk maintained relatively low rates. Two rats in Group Equal-Risk (i.e., Rats E.1 and E.6) decreased their risky choice as a function of session in the Two-Loss phase. Accordingly, when the probability of risky losses increased in the Two-Loss phase, some rats decreased their risky choice behavior, suggesting that they were sensitive to the increase in risky loss frequency. In contrast, the other rats in Group Equal-Risk may have been less attentive to the increase in frequency of risky losses, in that they may have primarily attended to the consistent yet less frequent delivery of R-11 gains. Similarly, two rats in Group Unequal-Risk (i.e., Rats U.4 and U.6) showed decreases in risky choice across sessions in the Two-Loss phase. Here, the decrease in risky choice as a

function of session may have reflected an aversion to R-1 outcomes (see Marshall & Kirkpatrick, 2015). Generally, the rats exhibited greater changes in behavior during the One-Loss than during the Two-Loss sub-phases, similar to that seen in the group data (Figure 3.2), suggesting that some rats' behaviors were less sensitive to the addition of the R-1 outcome in the Two-Loss sub-phase. As other rats behaved differently between the One- and Two-Loss sub-phases, these data strongly suggest considerably heterogeneity in loss processing and loss sensitivity across rats.

There was a generally distinct pattern of individual differences in the Risk-Variability phases across groups (Figure 3.4). In the One-Loss sub-phase, all rats in Group Equal-Risk showed an increase in risky choice as a function of session. There was less consensus in Group Unequal-Risk, in which some rats showed relatively little change in behavior with session (e.g., Rats U.4 and U.12), and some rats showed an increase in risky choice (e.g., Rats U.2 and U.3), suggesting greater individual variability in treatment of R-1 and R-11 outcomes when the outcomes were unequally presented (Group Unequal-Risk) than when they were equally presented (Group Equal-Risk). Subsequently, when risky choices probabilistically resulted in R-0, R-1, and R-11 outcomes in the Two-Loss sub-phase, many of the rats in Group Equal-Risk and Unequal-Risk exhibited relatively little sensitivity to the probabilistic addition of R-0 outcomes (see, e.g., Rats E.4 and U.11). Indeed, of the rats that were apparently sensitive to the change in task conditions, some rats showed an increase in risky choice (e.g., Rats E.10 and U.3), and other rats showed a respective decrease (e.g., Rats E.1 and U.10). Accordingly, there was substantial individual variability in treatment of added R-0 losses in the Risk-Variability phases, with some rats becoming more risk prone and others becoming more risk averse. Therefore, as in the Risk-Omission phases (Figure 3.3), different rats were differentially affected by the change in risky outcome parameters in the Risk-Variability phases, but there seemed to be

greater individual variability in the Risk-Variability phases than in the Risk-Omission phases. In other words, there appeared to be more substantial individual differences in the Two-Loss sub-phases given the addition of R-0 outcomes in the Risk-Variability phases than given the addition of R-1 outcomes in the Risk-Omission phases. Accordingly, the extension of the range of risky outcomes in adding the R-0 outcome (Risk-Variability) may have elicited greater individual differences in risky choice in terms of risk and loss sensitivity compared to the addition of the R-1 outcome to a consistent outcome range (Risk-Omission).

Molecular analysis

Effect of the previous outcome

Overall effects. A molecular analysis of risky choice was conducted to determine how the outcome of the previous choice and the loss-condition and risk-paradigm manipulations impacted subsequent risky choice in both groups. Analysis involved 88,965 observations, and included the overall intercept, categorical predictors of group (Unequal-Risk, Equal-Risk), risk paradigm (Risk-Variability, Risk-Omission), and previous outcome (C-2, C-4, R-0, R-1, R-11), and a continuous predictor of loss condition (One-Loss, Two-Loss). The categorical predictors were effect coded with Unequal-Risk/Equal-Risk as -1/+1 and Risk-Variability/Risk-Omission -1/+1. The reference level of previous outcome was the C-2 outcome. The fixed-effects structure included the full factorial model of Group \times Loss Condition \times Previous Outcome + Group \times Risk Paradigm \times Loss Condition + Risk Paradigm \times Previous Outcome. Intercept, risk paradigm, loss condition, previous outcome, and Risk Paradigm \times Loss Condition were random effects.

The full model output from the mixed-effects analyses is included in Table 3.3. Figure 3.5 shows the proportion of choices for the risky outcome as a function of the outcome of the

previous choice separated by group and loss condition. There was a general tendency for certain choices to be made after certain outcomes and risky choices to be made after risky outcomes. In addition, Group Equal-Risk made more risky choices than Group Unequal-Risk, regardless of the previous outcome. Interestingly, the shift from the One- to Two-Loss condition seemed to have induced greater mean staying behavior. That is, in the Two-Loss condition, there was a greater tendency to make certain choices after certain outcomes and risky choices after risky outcomes. Alternatively, the behavior in the Two-Loss condition may partially be an effect of training, such that the behavior that was continuing to develop in the One-Loss condition was further solidified in the Two-Loss condition.

Accordingly, analysis revealed a significant Group \times Loss Condition \times Previous Outcome interaction. Prior to discussion of this interaction, it should be noted that this significant interaction collapsed across risk paradigms (Risk-Omission, Risk-Variability). Due to the treatment of previous outcome and risk paradigm as categorical predictors, there was an incomplete design matrix (i.e., empty cells); that is, there were no R-0 outcomes in the Risk-Variability/One-Loss sub-phase and no R-1 outcomes in the Risk-Omission/One-Loss sub-phase, so that an analysis that included a Loss Condition \times Risk Paradigm \times Previous Outcome interaction was not permitted. Thus, an interpretation of how risky choice following R-0 and R-1 outcomes was affected following addition of a second risky loss to the Risk-Omission and Risk-Variability paradigms in the corresponding Two-Loss sub-phases should be interpreted with caution. However, Figure 3.6 was included to portray the risky choice data split by group, risk paradigm, loss condition, and previous outcome to better inform the reader.

As described above, after shifting from the One- to Two-Loss sub-phases, Group Equal-Risk showed a decrease in risky choice following C-2 (slope = -0.002) and C-4 outcomes (slope

= -0.01), as seen in Figure 3.5. This behavior is indicative of greater staying behavior after certain outcomes in the Two-Loss sub-phase (i.e., reduced risky choice behavior after certain choices). Group Equal-Risk also showed an increase in risky choice following R-0 (slope = 0.82), R-1 (slope = 0.55), and R-11 outcomes (slope = 0.96), again reflecting greater staying behavior after risky outcomes in the Two-Loss sub-phase (Figure 3.5). In the Two-Loss sub-phase relative to the One-Loss sub-phase, Group Equal-Risk had the potential to receive more loss outcomes following risky choices; also, the probabilities of the R-11 and risky loss outcomes decreased, but there were no changes in certain choice parameters. Accordingly, while the certain and risky choices for Group Equal-Risk resulted in outcomes greater or less than the expected value of that choice 50% of the time in the One-Loss sub-phases, there was an asymmetric shift in gain/loss probabilities between choices in the Two-Loss sub-phases. These results may suggest that the increase in the variability of outcome magnitudes and the decrease in probability of receiving each of the previous outcomes from the One-Loss sub-phase induced a greater overall likelihood of exploiting the previously-made choice.

In contrast to the increased likelihood of staying behavior in Group Equal-Risk, Group Unequal-Risk showed a general decline in risky choice from the One- to Two-Loss sub-phases. The largest decreases were in regards to risky choice following C-2 (slope = -0.21) and C-4 outcomes (slope = -0.58). Additionally, there was a relatively smaller decrease in risky choice from the One- to Two-Loss sub-phases following R-0 (slope = -0.16), R-1 (slope = -0.02) and R-11 outcomes (slope = -0.06). The discrepancies between the differences in group means in Figure 3.5 and the simple slopes are likely driven by the unweighted effects coding of the mixed effects analysis and the differential weighting of group means (i.e., by the different number of choices following each previous outcome across groups and rats within groups). Accordingly,

for Group Unequal-Risk, following the shift to the Two-Loss sub-phase (i.e., in which half of the One-Loss risky losses were replaced with an additional risky loss; see Table 3.1), there was a general reduction in likelihood to make risky choices following each of the previous outcomes, and this decrease was more accentuated following certain outcomes. This suggests that the relative consistency of risky loss frequencies did not considerably impact mean risky choice behavior following risky outcomes (see Yechiam & Telpaz, 2013). Moreover, as the probability of risky losses did not change from the One- to Two-Loss sub-phases, the general decrease in risky choice may have been driven by the increase in the number of outcomes that were equivalent to risky losses (i.e., R-0 and R-1).

In comparing groups across One- and Two-Loss sub-phases, Group Equal-Risk exhibited greater changes in behavior following risky outcomes and Group Unequal-Risk exhibited greater changes in behavior following certain outcomes. Interestingly, these changes occurred despite the fact that (1) the parameters of the certain choice were consistent, and (2) both groups experienced the same paradigm in the Two-Loss sub-phases. As Group Equal-Risk made more risky choices than Group Unequal-Risk, these results may be a function of simple exposure; more choices for the risky outcome in Group Equal-Risk would permit faster detection of changes in the risky choice parameters. Moreover, the relatively minimal change in behavior following risky outcomes in Group Unequal-Risk, in contrast to the larger increases in risky choice following the risky outcomes in Group Equal-Risk, may have been a function of the smaller changes in risky choice parameters experienced by Group Unequal-Risk than Group Equal-Risk. Indeed, Payne (2005) reported that humans are particularly sensitive to task parameters that increase and decrease the probabilities of gains and losses, respectively, and Horstmann, Villringer, and Neumann (2012) suggested that individuals may be more sensitive to

frequencies of gains and losses than to expected values in the Iowa Gambling Task (Bechara, Damasio, Tranel, & Damasio, 1997). The greater parametric changes in the Two-Loss sub-phases experienced by Group Equal-Risk may have promoted the greater change in risky choice behavior.

Overall, these results suggest that loss-chasing behavior following risky outcomes was differentially affected by group, as well as the parameters of the risky losses. For Group Unequal-Risk, risky loss probability consistently exceeded that of risky gain probability, suggesting that the addition of another risky loss magnitude without changing loss probability produced a small decline in risky choice. Accordingly, these results corroborate previous research which has suggested that the exposure to differential losses may have a greater impact over risky decision making than simply the expected value of the choices (Marshall & Kirkpatrick, 2015; also see Nygren, Isen, Taylor, & Dulin, 1996). Indeed, as seen in Figure 3.2, the group's mean choice behavior was relatively constant across sessions in the Two-Loss sub-phases. Thus, as the rats in Group Unequal-Risk were trained to essentially expect the delivery of more losses than gains, the increase in number of outcomes regarded as losses may have induced greater risk aversion.

In contrast, when the probability of gains was identical to the probabilities of each loss in the environment (i.e., Group Equal-Risk), the addition of a second risky loss in the Two-Loss phase seemed to have *increased* risky choice following risky outcomes. One potential explanation for these results regards averaging artifacts. As seen in Figure 3.2, mean risky choice behavior was relatively constant across sessions in the Two-Loss sub-phases, despite its being relatively dynamic in the One-Loss sub-phases. Accordingly, the behavioral differences between the One- and Two-Loss sub-phases in terms of loss-chasing behavior may reflect the

rats' developing their risk-taking preferences during the One-Loss sub-phase, which were then maintained in the Two-Loss sub-phase. For example, a gradual increase in risky choice following risky outcomes in the One-Loss sub-phases followed by relative maintenance of such behavior in the Two-Loss sub-phases would have produced greater post-risky-choice risk-taking in the Two-Loss than in the One-Loss sub-phases simply by the mechanism of averaging rats' behavior across sessions within the sub-phases. However, the considerable individual differences in risky choice behavior (Figures 3.3 and 3.4) warrant further investigation into the subjective preferences and psychological mechanisms governing risky decision making in dynamic environments.

Indeed, a second potential explanation for these behavioral patterns in Group Equal-Risk may be in terms of LDWs. In the Two-Loss sub-phase, one third of the R-11 outcomes were replaced by losses and one third of the initial risky loss outcomes (R-0 or R-1) were replaced by the new risky loss magnitude (R-1 or R-0), such that risky gains and losses were partially replaced with a second loss outcome. This second risky loss may have assumed gain- and loss-related characteristics via integration of gain- and loss-related information, as the second risky loss was essentially introduced by decreasing the probability of both risky gains (R-11) and losses (R-0 or R-1). If these probability manipulations ultimately produced some form of LDWs, then the increase in risky choice following risky outcomes in Group Equal-Risk in the Two-Loss condition may be explained by the introduction of these LDWs and the generalized association between LDW outcomes and the other risky outcomes in the environment. Accordingly, these potential LDWs may have conditionally promoted greater risky choice following risky outcomes (see Experiment 1). Moreover, this increase in risky choice may reflect the rats' tracking the probability of each risky outcome. Specifically, deliveries of the added risky loss were

ultimately predictive of fewer deliveries of the initial risky loss that was experienced in the One-Loss condition, which may have transformed the new risky loss into an indirect gain. However, the receipt of the new risky loss in the Two-Loss condition was also predictive of fewer larger R-11 gains (i.e., an indirect loss). Indeed, these specific changes in probability relationships did not occur for Group Unequal-Risk, such that the same behavioral differences would not be expected. Accordingly, the probabilistic receptions of a second risky loss may have indeed served to produce LDW-type events, thereby producing increased risky choice behavior following risky outcomes in Group Equal-Risk.

Individual differences. The random effects of the aforementioned molecular analysis were intercept, risk condition, loss condition, session, previous outcome, and Risk Condition \times Loss Condition. There were considerable individual differences in risky choice at the molecular level (Figure 3.7 and 3.8). Primarily, the rats exhibited post-outcome staying behavior, in that they were relatively more likely to make the same choice that was made in the previous trial. In the Risk-Omission condition (Figure 3.7), when shifting from the One-Loss to Two-Loss phases, some rats relatively maintained the same rate of risky choices following each of the previous outcomes even though the R-1 outcome was then available to be delivered (e.g., Rats E.9 and U.3). Other rats exhibited greater changes in behavior, and some of these changes occurred only after certain or risky outcomes. For example, Rat E.12 made fewer risky choices after certain outcomes in the Two-Loss phase, while Rat U.4 made fewer risky choices following R-0 and R-11 outcomes in the Two-Loss phase than in the One-Loss phase. Moreover, in the Two-Loss sub-phase, following all previous outcomes, Rat U.11 exhibited near-exclusive preference for the certain choice, while Rat E.7 exhibited near-exclusive preference for the risky choice.

In the Risk-Variability condition (Figure 3.8), one relatively common pattern was an increase in risky choice following all previous outcomes when shifting from the One-Loss to Two-Loss sub-phases (e.g., Rats E.3 and U.6). Other rats showed greater post-outcome staying behavior after this shift (i.e., Rats E.10, and U.10). Accordingly, there were substantial individual differences in the likelihood of risky choice following each previous outcome, as well as the sensitivity to the changes in reinforcement contingencies in the Two-Loss phases of each risk condition. These individual differences may be indicative of distinct strategies that different rats are using to decipher possible changes in the choice environment, as well as differential sensitivities and motivations to risky choices and outcomes.

Outcome and choice history. As in Experiment 1, analyses were conducted to determine whether the differential frequencies of losses and gains between groups modified the decaying influence of past outcomes on choice, and to determine the functional form of this decay. Separate analyses were conducted for outcome history and choice history.

The first component of analysis involved 80,343 observations, and included the overall intercept, a random intercept, and the fixed and random effects of either the previous 10 outcomes or choices in the corresponding analyses. Tables 3.4 and 3.5 include the mixed-effects model output for the outcome history and choice history analyses, respectively. Regression coefficients decayed as a function of previous outcome and choice. All regression coefficients for the previous outcomes and choices were significantly greater than 0, $t_s \geq 3.52$, $p_s < .001$, suggesting that, as in Experiment 1, the rats' history of outcomes and choices had partial influence over subsequent choice behavior (see Landon et al., 2002).

Figure 3.9 shows the individual rat hyperbolic and exponential model fits for the outcome history analysis. For all rats, there was a general decrease in regression coefficients. In the

outcome history analysis, there were no differences between groups in the intercept (A) of the hyperbolic model, $t(213) = 1.41, p = .161$ (Group Equal-Risk, $A = 0.13$; Group Unequal-Risk, $A = 0.18$), or the decay rate (k) of the hyperbolic model, $t(213) = -0.27, p = .791$ (Group Equal-Risk, $k = 0.80$; Group Unequal-Risk, $k = 0.74$). Similarly, the exponential analysis of outcome history indicated that there were no group differences in intercept, $t(213) = 1.41, p = .160$ (Group Equal-Risk, $A = 0.12$; Group Unequal-Risk, $A = 0.16$), or decay rate, $t(213) = -0.33, p = .742$ (Group Equal-Risk, $k = 0.28$; Group Unequal-Risk, $k = 0.25$). The hyperbolic model accounted for the data better than the exponential model did (hyperbolic AIC = -1237.10, exponential AIC = -1148.70).

Figure 3.10 shows the individual rat hyperbolic and exponential model fits for the choice history analysis. For all rats, there was a general decrease in normalized regression coefficients (but see Rat E.7). The hyperbolic choice history analysis indicated that there were no significant group differences in decay rate (k), $t(215) = 1.22, p = .223$ (Group Equal-Risk, $k = 1.79$; Group Unequal-Risk, $k = 2.37$). The exponential analysis also indicated no group differences in decay rate (k), $t(215) = 1.36, p = .175$ (Group Equal-Risk, $k = 0.52$; Group Unequal-Risk, $k = 0.64$). The hyperbolic model accounted for a better fit of the data than the exponential model in the choice history analysis (hyperbolic AIC = -421.37, exponential AIC = -256.46). Thus, in conjunction with the results in Experiment 1, the superiority of the hyperbolic model in accounting for the decaying influences (weights) of past outcomes and choices on subsequent choice suggests that traditional assumptions of exponential decay of these events require modification and reconceptualization (also see Alexander & Brown, 2010; Maia, 2009).

RL models

As in Experiment 1, multiple RL models were fit to both simulated and observed rat data to gain insight into the psychological processes and mechanisms that govern individual differences in risky choice in rats.

Parameter recovery. Figure 3.11 show the parameter recovery results of the Simple, Asymmetric, Valence-Attentive, and Weighted-Reference-Point RL models. In each panel, the dark line represents the unit diagonal and each data point represents the parameter that was used in its simulation, and the corresponding fitted parameter. Figure 3.11 does not include simulations in which the fitted parameters exceeded the boundaries of the distribution from which the simulated parameters were sampled. This occurred in 5.4% of the simulations of Simple RL, 15.9% of the simulations of Asymmetric RL, 6.8% of the simulations of Valence-Attentive RL, and 30.6% of the simulations of Weighted-Reference-Point RL. As seen in Figure 3.11, the Simple, Asymmetric, and Valence-Attentive RL models did well to recover the parameters from the simulated data. In contrast, the Weighted-Reference-Point RL model did poorly to recover the simulated parameters. Therefore, the Simple, Asymmetric, and Valence-Attentive RL models were fit to the data.

Model selection. Figure 3.12 shows the AICs for the best-fitting Simple, Asymmetric, and Valence-Attentive RL models. The dark line represents the unit diagonal, such that data points below the diagonal reflect support for the Simple or Valence-Attentive RL models and data points above the diagonal represent support for the Asymmetric RL model. Even through 300 iterations, the Simple RL model did not converge on a viable solution for six rats. In contrast, the Valence-Attentive and Asymmetric RL models converged on a viable solution for all 24 rats. Whereas the majority of data points in Figure 3.12 were proximal to the unit

diagonal, there were multiple AICs for the Valence-Attentive RL model that exceeded those of the Simple and Asymmetric RL models, indicative of a poorer fit by the Valence-Attentive RL model. The mean (median) AIC of the Asymmetric RL fits was 3141.9 (3284.3), while mean AICs of the Valence-Attentive RL and viable Simple RL fits were 3478.9 (3659.1) and 3039.7 (2914.6), respectively. In regards to individual rats, the Simple RL model was the best model of the data for 3 rats, the Valence-Attentive RL model was the best model for 9 rats, and the Asymmetric RL model provided the best account of the data for 12 rats. For the rats in which the Valence-Attentive RL model had a lower AIC than the Asymmetric RL model, the mean Δ AIC was 51.5; for the rats in which the Asymmetric RL model had the lower AIC, the mean Δ AIC was 787.8. While there were instances in which the Simple or Valence-Attentive RL models were superior to the Asymmetric RL and while the Simple RL model had a lower mean AIC, the Simple RL model did not converge for all rats. Thus, these data suggest that the Asymmetric RL model provided the best account of the risky choice data set as a whole. However, it is noteworthy that the Simple and Valence-Attentive RL models were more comparable in AIC to that of the Asymmetric RL models in the current experiment than in Experiment 2. Given the substantial individual differences in both molar and molecular risky choice behavior (Figures 3.3, 3.4, 3.7, and 3.8), these similarities in model performance may reflect the aforementioned possibility that different rats are using different strategies to interact with the changing environments. For example, instances in which the Valence-Attentive RL model outperformed the Asymmetric RL model may indicate that these rats' behaviors were driven by differences in attention given to gains versus losses rather than differences in gain- and loss-based learning. Given the general support for the Asymmetric RL model over the other RL models, the model fits of Asymmetric RL were further evaluated.

Model fit (Asymmetric RL). Figure 3.13 shows the fits of the Asymmetric RL model to individual trial data of five different rats. The alternating gray and white boxes in Figure 3.13 indicate the different phases and sub-phases (see Table 3.1). As in Experiment 1, the observed data were smoothed over a moving nine-trial window (see Lau & Glimcher, 2005). The panels represent the data and model fit corresponding to the minimum, 25th percentile, median, 75th percentile, and maximum of the ω^2 values (i.e., the top panel includes the model with the lowest ω^2 value; the bottom panel, the largest ω^2 value). Table 3.6 shows the summary statistics of the best-fitting parameter estimates of the Asymmetric RL model, as well as the corresponding fit indices. The lowest ω^2 value was negative, which was primarily due to the rat exhibiting little change in choice behavior, such that the mean of the data provided a better account of the data than the model (Figure 3.13, top panel). Despite the noisiness of the rats' data, the model approximated the rats' average local choice preferences.

In regards to the fitting parameters, there was no significant overall difference between α_L and α_G across rats, $z = 1.00$, $p = .317$ (Wilcoxon signed-ranks test). Indeed, the instances in which the Simple RL model provided a better account of the data would likewise suggest similar gain- and loss-based learning rates. Interestingly, analysis revealed significantly greater gain-based learning / value-updating rates (α_G) in Group Equal-Risk relative to Group Unequal-Risk, $z = 2.22$, $p = .026$, as well as significantly smaller β values in Group Equal-Risk, $z = -2.57$, $p = .010$; there were no group differences in α_L , $z = -0.78$, $p = .436$ (Wilcoxon rank-sum tests). The group differences in α_G may be due to the increased exposure to R-11 gains in Group Equal-Risk relative to Group Unequal-Risk, or the necessity for Group Equal-Risk to track changing R-11 probabilities, which Group Unequal-Risk did not experience. Furthermore, the smaller β values

in Group Equal-Risk reflect greater exploration (i.e., less exploitation of the choice with the larger subjective value).

Simple and multiple linear regression models were performed to determine how individual differences in model parameters corresponded to individual differences in the rats' behavior (Figure 3.14). In the first series of models, mean risky choice data across all conditions were regressed on α_G , α_L , and β . The model parameters were log-transformed to correct for positive skewness. The data did not require transformation.

There was a significant positive relationship between mean risky choice and α_G , $t(22) = 2.58$, $p = .017$, in that greater risky choice behavior was related to greater gain-based learning (Figure 3.14, left panel). Also, there was a significant negative relationship between mean risky choice and α_L , $t(22) = -4.52$, $p < .001$, such that greater loss-based learning was associated with less overall risky choice (Figure 3.14, middle panel). Lastly, there was a significant negative relationship between mean risky choice and β , $t(22) = -3.75$, $p = .001$, such that the more risk averse rats were those that were more likely to exploit the higher-valued choice (Figure 3.14, right panel). Here, the expected value of the risky choice exceeded that of the certain choice (see Table 3.1), suggesting that greater exploitation of the higher-valued choice in risk-averse rats was due to the *subjective* value of the risky choice being less than that of the certain choice. This discrepancy was likely driven by subjective asymmetry in α_G versus α_L . As the majority of risky outcomes were losses for Groups Equal-Risk and Unequal-Risk, the prevalence of losses in conjunction with greater loss-based learning rates would therefore reduce the value of the risky choice below that of the certain choice. The rats who were more likely to exploit the higher-valued option were indeed those that made more certain choices. Similarly, Addicott, Pearson, Kaiser, Platt, and McClernon (2015) reported greater exploration in more regular gamblers.

A second series of models involved regressing two measures of molecular choice behavior on the model parameters. Due to the paradigmatic differences between Experiments 1 and 2, the indices of molecular choice were computed differently from those of Experiment 1, but approximated the computations in Experiment 1 to permit comparable interpretations. Similar to Experiment 1, the two behavioral indices were the Win-Stay and Loss Sensitivity indices, but, as stated, their computations differed slightly.

In Experiment 1, the Win-Stay Index evaluated risky choice behavior following R-11 outcomes versus that following R-0 outcomes. Here, because R-0 outcomes were not presented in the Risk-Variability/One-Loss phase, the amount of exposure to R-0 outcomes was less than that of R-11 outcomes. Accordingly, a global difference in risky choice behavior following these outcomes may potentially be confounded by the frequency of exposure to these outcomes. Thus, the Win-Stay Index was computed separately for the Two-Loss sub-phase within each risk paradigm, as all risky outcomes were possible in the Two-Loss sub-phase and the opportunity for exposure to the outcomes was equivalent. Here, the Win-Stay Index was the proportion of risky choices following R-11 outcomes minus the proportion of risky choices following the risky loss that was consistently presented throughout both loss conditions of that risk type (i.e., R-0 in the Risk-Omission phases; R-1 in the Risk-Variability phases). For example, the proportion of risky choices following R-0 outcomes in the Risk-Omission/Two-Loss sub-phase was subtracted from the proportion of risky choices following R-11 outcomes in the Risk-Omission/Two-Loss sub-phase. This ensured that the derived index of win-stay behavior was not affected by the probabilistic addition of the new risky loss magnitude that was presented in the Two-Loss sub-phases of each risk paradigm (i.e., R-1 in the Risk-Omission/Two-Loss sub-phase; R-0 in the Risk-Variability/Two-Loss sub-phase). The Win-Stay Index was averaged across risk paradigms

to produce a general measure of staying behavior following R-11 outcomes. Accordingly, the larger the value of the Win-Stay Index, the more likely the rats were to make risky choices after R-11 gains than after risky (R-0 or R-1) losses. Relatedly, the smaller the value of the Win-Stay Index, the more likely the rats were to chase risky losses than they were to follow R-11 gains with risky choices.

As in Experiment 1, the Win-Stay Index was separately regressed on α_G , α_L , and the $\alpha_L:\alpha_G$ ratio. Each of the analyses involved one of the three aforementioned model parameter measures, group, and the corresponding interaction to determine whether group moderated the relationship between the modelled parameters and local risky choice indices. The main effects and interaction were entered simultaneously into the model. To correct for negative skewness, the Win-Stay Index was exponentially transformed. Group was effect coded with Unequal-Risk/Equal-Risk as -1/+1, and the model parameters were mean-centered.

The left panel of Figure 3.15 shows the relationship between α_G and the Win-Stay Index. Here, there was no main effect of group, $b = 0.01$, $t(20) = 0.43$, $p = .675$, but win-stay behavior following risky outcomes significantly increased with α_G , $b = 0.03$, $t(20) = 2.63$, $p = .016$, and there was a significant Group \times α_G interaction, $b = -0.03$, $t(20) = -2.18$, $p = .042$. Here, Group Equal-Risk exhibited a shallower relationship between α_G and win-stay behavior (slope = 0.01) than Group Unequal-Risk (slope = 0.06). For the relationship between α_L and the Win-Stay Index (Figure 3.15, middle panel), there was no effect of group, $b = 0.04$, $t(20) = 1.72$, $p = .101$, no effect of α_L , $b = -0.002$, $t(20) = -0.25$, $p = .806$, and no Group \times α_L interaction, $b = 0.01$, $t(20) = 0.66$, $p = .517$. Similarly, there was no effect of group, $b = 0.03$, $t(20) = 1.23$, $p = .232$, no effect of $\alpha_L:\alpha_G$, $b = -0.01$, $t(20) = -0.99$, $p = .332$, and no Group \times $\alpha_L:\alpha_G$ interaction, $b = 0.01$, $t(20) = 1.00$, $p = .330$, on win-stay behavior following risky outcomes (Figure 3.15, right panel).

These results indicate that win-stay behavior in the current experiment was primarily driven by gain-based learning. The significant Group \times α_G interaction suggests that win-stay behavior was more strongly related to gain-based learning in Group Unequal-Risk than in Group Equal-Risk. This result may reflect general biases in Group Equal-Risk regardless of learning rate, or the greater homogeneity of gain-based learning rates in Group Equal-Risk (Figure 3.15, left panel). Alternatively, this interaction may be driven by the parametric structure of the environment; specifically, for Group Unequal-Risk, the probability of the R-11 outcome was constant at .33 across all phases, but changed every 10 sessions for Group Equal-Risk. Therefore, the between-groups differences in the stability of such probability information may have weakened the strength of this relationship in Group Equal-Risk (e.g., win-stay propensities may have been driven by gain-based learning *and* learning the changes in reinforcement contingencies, among other task dynamics).

The Loss Sensitivity Index of Experiment 1 accounted for the difference in risky choice behavior following R-1 and R-0 outcomes. Here, the R-1 and R-0 outcomes were only presented together in the Two-Loss sub-phases, such that a metric of an overall difference in risky choice behavior following these outcomes may be confounded by their deliveries with and without the probabilistic delivery of the other risky loss. Accordingly, in conjunction with the goals of the present experiment, the Loss Sensitivity Index in this experiment measured how risky choice following risky losses changed between the One- and Two-Loss sub-phases. Specifically, for each risk type, the proportion of risky choices following the risky loss in the One-Loss sub-phase (i.e., R-0 in the Risk-Omission/One-Loss sub-phase; R-1 in the Risk-Variability/One-Loss sub-phase) was subtracted from the mean proportion of risky choices following both risky losses in the Two-Loss sub-phase. The Loss Sensitivity Index for each risk type was analyzed separately,

and, as done for the analyses of win-stay behavior, the Loss Sensitivity Index was separately regressed on α_G , α_L , and $\alpha_L:\alpha_G$. The Loss Sensitivity Index was exponentially transformed to correct for negative skewness. Here, larger values of the Loss Sensitivity Index can be interpreted as greater increases (or smaller decreases) in risky choice behavior following risky losses when there were two potential risky losses in the environment (R-0 and R-1) than when there was one potential risky loss in the environment (R-0 or R-1).

Figure 3.16 shows the relationships between the model parameters from the Asymmetric RL fits and the Loss Sensitivity Index for each group within the Risk-Omission phases. The left panel of Figure 3.16 shows this relationship between the Loss Sensitivity Index and α_G . There was a main effect of group, $b = 0.07$, $t(20) = 2.23$, $p = .037$, but no effect of α_G , $b = 0.04$, $t(20) = 2.05$, $p = .053$, and no Group \times α_G interaction, $b = -0.02$, $t(20) = -1.28$, $p = .214$. The middle panel of Figure 3.16 shows the relationship between the Loss Sensitivity Index and α_L . Here, there was a main effect of group, $b = 0.09$, $t(20) = 3.45$, $p = .003$, and a main effect of α_L , $b = -0.03$, $t(20) = -2.54$, $p = .019$, such that greater learning from losses (α_L) was associated with smaller increases (or larger decreases) in risky choice behavior following risky losses when the number of potential risky losses doubled in the Risk-Omission/Two-Loss sub-phase. In contrast, there was no Group \times α_L interaction, $b = 0.01$, $t(20) = 1.04$, $p = .310$. The right panel of Figure 3.16 shows the relationship between the Loss Sensitivity Index and $\alpha_L:\alpha_G$. There was a main effect of group, $b = 0.07$, $t(20) = 2.44$, $p = .024$, a main effect of $\alpha_L:\alpha_G$, $b = -0.03$, $t(20) = -3.17$, $p = .005$, and no Group \times $\alpha_L:\alpha_G$ interaction, $b = 0.003$, $t(20) = 0.34$, $p = .737$. Similar to the main effect of α_L , the main effect of $\alpha_L:\alpha_G$ indicates that greater learning from losses relative to gains was also associated with smaller increases (or larger decreases) in risky choice behavior following risky losses from the Risk-Omission/One-Loss to Risk-Omission/Two-Loss sub-phase.

This pattern of results indicates that greater loss-based learning rates predicted the extent to which rats decreased their risky choice behavior following risky losses when the quantity of risky loss outcomes increased. Accordingly, when the R-1 outcome was included as a potential risky outcome in the Two-Loss sub-phase, the rats who were less likely to chase risky losses (relative to their own behavior when the only risky loss was the R-0 outcome) were those that exhibited greater loss-based learning rates relative to gain-based learning rates. Thus, subjective deficits in risky decision making may reflect a reduced ability to learn from experienced losses (see Clark, Liu, et al., 2013).

Figure 3.17 shows the relationships between the best-fitting Asymmetric RL parameters and the Loss Sensitivity Index in the Risk-Variability phase. Regarding the relationship between α_G and the Loss Sensitivity Index (Figure 3.17, left panel), there was no main effect of group, $b = -0.001$, $t(20) = -0.05$, $p = .964$, no effect of α_G , $b = -0.01$, $t(20) = -0.62$, $p = .541$, and no Group $\times \alpha_G$ interaction, $b = 0.01$, $t(20) = 0.75$, $p = .464$. The middle panel of Figure 3.17 shows the relationship between the Loss Sensitivity Index and α_L . Here, there was no main effect of group, $b = -0.01$, $t(20) = -0.34$, $p = .736$, or α_L , $b = 0.01$, $t(20) = 0.60$, $p = .553$, but there was a Group $\times \alpha_L$ interaction, $b = -0.02$, $t(20) = -2.22$, $p = .038$. Here, Group Equal-Risk exhibited a negative relationship between α_L and the Loss Sensitivity Index (slope = -0.02), while Group Unequal-Risk exhibited a positive relationship between these two variables (slope = 0.03). The right panel of Figure 3.17 shows the relationship between the Loss Sensitivity Index and $\alpha_L:\alpha_G$, in which there was no main effect of group, $b = -0.01$, $t(20) = -0.38$, $p = .706$, or $\alpha_L:\alpha_G$, $b = 0.001$, $t(20) = 0.09$, $p = .930$, and no Group $\times \alpha_L:\alpha_G$ interaction, $b = -0.02$, $t(20) = 2.05$, $p = .054$. Therefore, the rats in Group Equal-Risk who exhibited greater loss-based learning rates showed smaller increases (or larger decreases) in loss-chasing behavior when there was a probabilistic

addition of R-0 losses in the Two-Loss sub-phase of the Risk-Variability condition, which parallels the patterns in the Risk-Omission condition (see Figure 3.16, middle panel). However, the rats in Group Unequal-Risk who exhibited greater loss-based learning rates showed larger increases (or smaller decreases) in loss-chasing when R-0 became a potential risky outcome in the Risk-Variability/Two-Loss sub-phase, suggesting that substituting half of the R-1 outcomes with R-0 outcomes increased loss-chasing behavior. Accordingly, individual differences in risky choice following losses seemingly depends on the frequency and magnitude of risky losses experienced in past conditions.

One explanation for this pattern of results is that loss-based learning is independent of post-loss decision making, but the results from the Risk-Omission phases do not support this hypothesis (Figure 3.16). Alternatively, these results may in fact be driven by structure of the environment. Group Equal-Risk exhibited negative relationships between α_L and the Loss Sensitivity Index in both the Risk-Omission and Risk-Variability conditions (see Figures 3.16 and 3.17). For Group Equal-Risk, the probabilities of both risky gains and losses changed from the One- to Two-Loss sub-phases in both risk paradigm conditions, suggesting that the relationship between α_L and the Loss Sensitivity Index may be related to the overall increase in number of losses experienced; the rats that exhibited greater loss-based learning rates were those that were less likely to chase losses in the Two-Loss sub-phase, possibly because they more readily learned from the increased frequency of losses. Alternatively, the environmental changes may have been more salient and, thus, learned more quickly by the rats that exhibited greater loss-based learning rates (see Behrens, Woolrich, Walton, & Rushworth, 2007). In contrast, the rats in Group Unequal-Risk with comparable loss-based learning rates as rats in Group Equal-Risk exhibited greater decreases in loss-chasing when the R-1 outcome was added as a potential

risky choice outcome in the Risk-Omission/Two-Loss sub-phase, but greater increases in loss-chasing when the R-0 outcome was added as a potential risky choice outcome in the Risk-Variability/Two-Loss sub-phase. Thus, the treatment of R-0 and R-1 losses seemingly depended on the probabilities of each of these outcomes relative to the probabilities of other risky choice outcomes. Moreover, the differential relationships between α_L and loss sensitivity in Group Unequal-Risk suggest that the partial substitution of some risky losses with other risky losses is influenced by the loss magnitude; the addition of R-1 outcomes (Risk-Omission/Two-Loss) decreased relative loss-chasing while the addition of R-0 outcomes (Risk-Variability/Two-Loss) increased it. Thus, these results suggest that rats exhibited distinct sensitivities to differential losses, such that the R-1 outcome may in fact be viewed as generally more aversive than the R-0 outcome under particular conditions (see Marshall & Kirkpatrick, 2015).

Table 3.1 Probability of the risky-zero (R-0), R-1, and R-11 outcomes and the expected value of the risky choice in One-Loss and Two-Loss sub-phases of the Risk-Omission and Risk-Variability phases of the risky choice task of Experiment 2 for Groups Equal-Risk and Unequal-Risk.

Group	Risk Paradigm	Risky Choice Task Sub-Phase			
		One-Loss		Two-Loss	
		Outcome (Probability)	Expected Value	Outcome (Probability)	Expected Value
Equal-Risk (n=12)	Risk-Omission	R-0 (.50) R-11 (.50)	5.5		
	Risk-Variability	R-1 (.50) R-11 (.50)	6	R-0 (.33) R-1 (.33) R-11 (.33)	4
Unequal-Risk (n=12)	Risk-Omission	R-0 (.67) R-11 (.33)	3.67		
	Risk-Variability	R-1 (.67) R-11 (.33)	4.33		

Table 3.2 Mixed-effects model output for the molar analyses of Experiment 2. Note: continuous variables were mean-centered, and categorical variables were effect coded with Group Unequal-Risk/Equal-Risk (Group) as -1/+1 and Risk-Variability/Risk-Omission (Risk Paradigm) as -1/+1.

Factor	Estimate	T (89907)	P	95% CI
Intercept	0.11	0.31	.757	[-0.57, 0.79]
Group	1.17	3.38	.001	[0.49, 1.85]
Risk Paradigm	-0.12	-0.66	.510	[-0.46, 0.23]
Loss Condition	0.06	0.34	.732	[-0.27, 0.38]
Session	0.01	0.28	.773	[-0.03, 0.04]
Group × Risk Paradigm	0.24	1.36	.175	[-0.11, 0.59]
Group × Loss Condition	0.41	2.47	.014	[0.08, 0.73]
Risk Paradigm × Loss Condition	-0.27	-2.16	.031	[-0.51, -0.02]
Group × Session	0.06	3.44	.001	[0.03, 0.10]
Risk Paradigm × Session	-0.05	-14.30	< .001	[-0.06, -0.04]
Loss Condition × Session	-0.08	-11.55	< .001	[-0.09, -0.06]
Group × Risk Paradigm × Loss Condition	0.39	3.15	.002	[0.15, 0.63]
Group × Risk Paradigm × Session	0.01	3.98	< .001	[0.01, 0.02]
Group × Loss Condition × Session	-0.06	-8.57	< .001	[-0.07, -0.04]
Risk Paradigm × Loss Condition × Session	0.08	11.51	< .001	[0.06, 0.09]
Group × Risk Paradigm × Loss Condition × Session	-0.01	-1.89	.058	[-0.03, 0.0004]

Table 3.3 Mixed-effects model output for the molecular analyses of Experiment 2 evaluating the effect of the previous outcome on subsequent choice behavior. Note: continuous variables were mean-centered, and categorical variables were effect coded with Group Unequal-Risk/Equal-Risk (Group) as -1/+1 and Risk-Variability/Risk-Omission (Risk Paradigm) as -1/+1. The reference level of previous outcome was the certain-two (C-2) outcome. C-4 = certain choice – 4-pellet outcome; R-0 = risky choice – 0-pellet outcome; R-1 = risky choice – 1-pellet outcome; R-11 = risky choice – 11-pellet outcome

Factor	Estimate	T (88937)	P	95% CI
Intercept	0.38	1.63	.104	[-0.08, 0.84]
Group	0.79	3.40	.001	[0.34, 1.25]
Risk Paradigm	0.01	0.10	.918	[-0.23, 0.26]
Loss Condition	0.13	1.07	.284	[-0.11, 0.37]
C-4	-1.68	-14.37	< .001	[-1.91, -1.45]
R-0	0.98	7.43	< .001	[0.72, 1.24]
R-1	1.33	13.13	< .001	[1.13, 1.53]
R-11	1.04	8.64	< .001	[0.80, 1.28]
Group × Risk Paradigm	0.17	1.35	.177	[-0.08, 0.42]
Group × Loss Condition	0.33	2.78	.005	[0.10, 0.57]
Risk Paradigm × Loss Condition	-0.27	-3.10	.002	[-0.44, -0.10]
Group × C-4	0.22	1.90	.057	[-0.01, 0.45]
Group × R-0	-0.23	-1.74	.082	[-0.48, 0.03]
Group × R-1	-0.14	-1.43	.153	[-0.33, 0.05]
Group × R-11	0.15	1.23	.219	[-0.09, 0.38]
Risk Paradigm × C-4	-0.04	-1.53	.127	[-0.09, 0.01]
Risk Paradigm × R-0	0.19	5.00	< .001	[0.11, 0.26]
Risk Paradigm × R-1	0.05	1.24	.214	[-0.03, 0.13]
Risk Paradigm × R-11	-0.07	-2.45	.014	[-0.12, -0.01]
Loss Condition × C-4	-0.42	-8.91	< .001	[-0.52, -0.33]
Loss Condition × R-0	0.20	2.79	.005	[0.06, 0.34]
Loss Condition × R-1	0.14	2.07	.038	[0.01, 0.26]
Loss Condition × R-11	0.32	6.38	< .001	[0.22, 0.42]

Factor	Estimate	T (88937)	P	95% CI
Group × Risk Paradigm × Loss Condition	0.26	3.04	.002	[0.09, 0.43]
Group × Loss Condition × C-4	-0.05	-1.13	.258	[-0.14, 0.04]
Group × Loss Condition × R-0	0.16	2.98	.003	[0.05, 0.26]
Group × Loss Condition × R-1	-0.05	-0.87	.386	[-0.15, 0.06]
Group × Loss Condition × R-11	0.17	3.57	< .001	[0.08, 0.27]

Table 3.4 Mixed-effects model output for the molecular analyses of Experiment 2 evaluating the effect of the previous 10 outcomes on subsequent risky choice behavior.

Factor	Estimate	T (80332)	P	95% CI
(Intercept)	-0.07	-0.33	.739	[-0.46, 0.33]
Outcome – Lag 1	0.16	9.03	< .001	[0.12, 0.19]
Outcome – Lag 2	0.08	13.29	< .001	[0.07, 0.10]
Outcome – Lag 3	0.06	12.28	< .001	[0.05, 0.07]
Outcome – Lag 4	0.05	9.34	< .001	[0.04, 0.06]
Outcome – Lag 5	0.03	7.25	< .001	[0.02, 0.04]
Outcome – Lag 6	0.03	6.93	< .001	[0.02, 0.04]
Outcome – Lag 7	0.03	7.19	< .001	[0.02, 0.04]
Outcome – Lag 8	0.03	6.76	< .001	[0.02, 0.03]
Outcome – Lag 9	0.03	8.60	< .001	[0.02, 0.04]
Outcome – Lag 10	0.03	6.38	< .001	[0.02, 0.04]

Table 3.5 Mixed-effects model output for the molecular analyses of Experiment 2 evaluating the effect of the previous 10 choices on subsequent risky choice behavior.

Factor	Estimate	T (80332)	P	95% CI
(Intercept)	-0.04	-0.42	.677	[-0.24, 0.16]
Choice – Lag 1	1.12	14.05	< .001	[0.96, 1.27]
Choice – Lag 2	0.29	14.32	< .001	[0.25, 0.33]
Choice – Lag 3	0.18	9.61	< .001	[0.15, 0.22]
Choice – Lag 4	0.15	6.39	< .001	[0.11, 0.20]
Choice – Lag 5	0.12	5.23	< .001	[0.08, 0.17]
Choice – Lag 6	0.09	4.49	< .001	[0.05, 0.14]
Choice – Lag 7	0.13	5.08	< .001	[0.08, 0.18]
Choice – Lag 8	0.08	3.52	< .001	[0.04, 0.13]
Choice – Lag 9	0.15	6.04	< .001	[0.10, 0.21]
Choice – Lag 10	0.19	7.01	< .001	[0.14, 0.24]

Table 3.6 Summary statistics of the best-fitting parameter estimates and goodness-of-fit indices of the Asymmetric Reinforcement Learning (RL) model. AIC = Akaike Information Criterion; ω^2 = omega-squared.

	α_G	α_L	β	AIC	Pseudo R^2 (ω^2)
Minimum	.00006	.00005	0.0913	908.8	-0.24
25 th Percentile	.0007	.0015	0.3620	2322.1	0.10
Median	.0012	.0050	0.9835	3284.3	0.20
75 th Percentile	.0065	.0546	2.5389	3998.5	0.47
Maximum	.5888	.7372	6.1410	5367.5	0.76
Mean	.0443	.0630	1.6833	3141.9	0.28
Standard Deviation	.1386	.1583	1.5827	1208.1	0.25

Figure 3.1 Proportion of choices for the risky outcome for each group (Equal-Risk, Unequal-Risk) across risk paradigms (Risk-Omission, Risk-Variability) as a function of loss condition (One-Loss, Two-Loss).

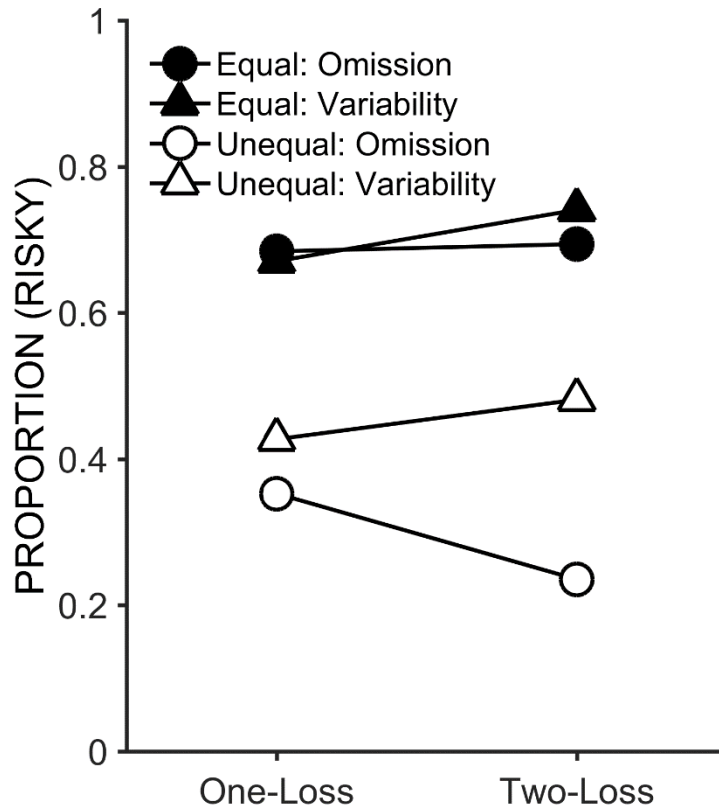


Figure 3.2 Proportion of choices for the risky outcome for each group (Equal-Risk, Unequal-Risk) across risk paradigms (Risk-Omission, Risk-Variability) as a function of the session in the One-Loss and Two-Loss conditions. The dotted line designates the data corresponding to the One-Loss (left) and Two-Loss conditions (right). In the figure, the abscissa is labeled as beginning at “1” and ending at “20” for the purpose of interpretability, but sessions were nested within each phase and sub-phase for analyses.

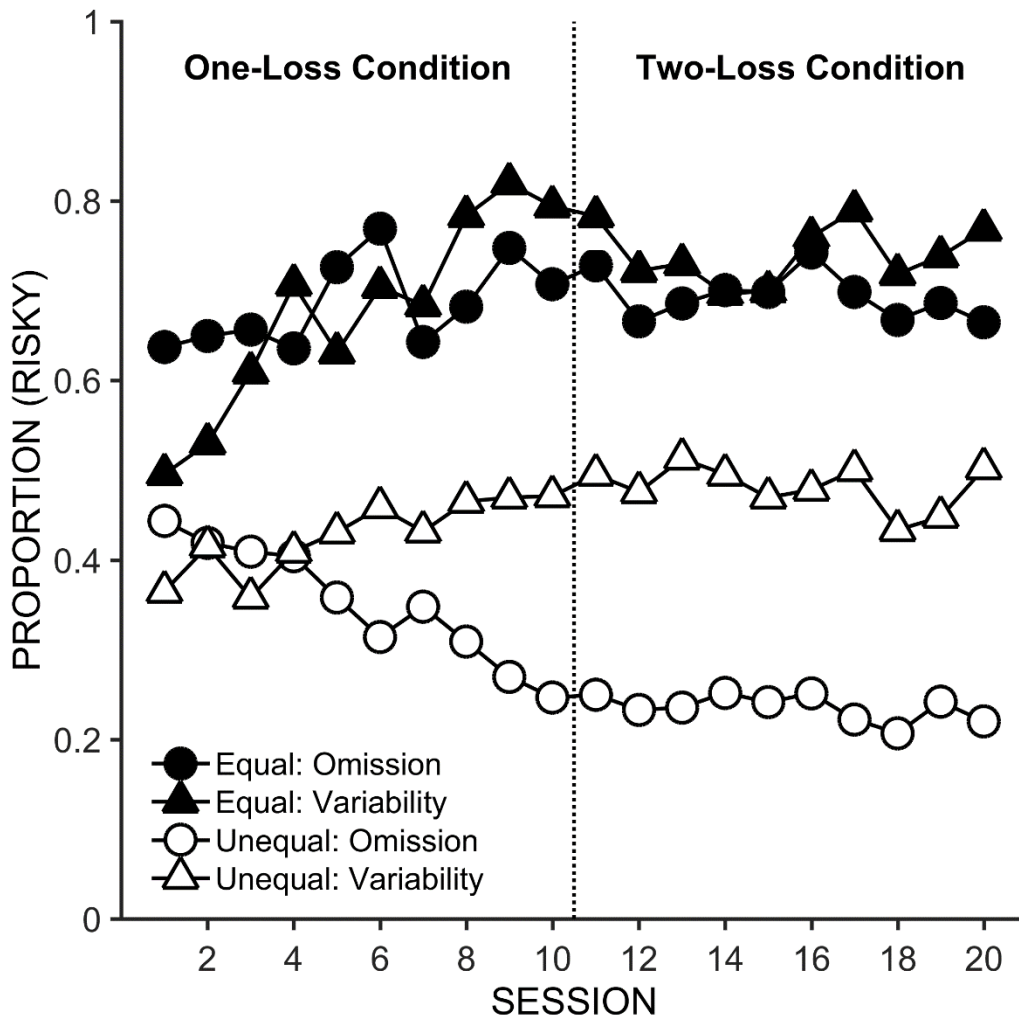


Figure 3.3 Proportion of choices for the risky outcome as a function of session and loss condition for individual rats within the Risk-Omission phases. The solid (Omission: One-Loss) and dashed lines (Omission: Two-Loss) represent the individual subject-fits from the best-fitting generalized linear mixed effects model. The rats in Group Equal-Risk are labeled as “E.” and the rats in Group Unequal-Risk are labeled as “U.”.

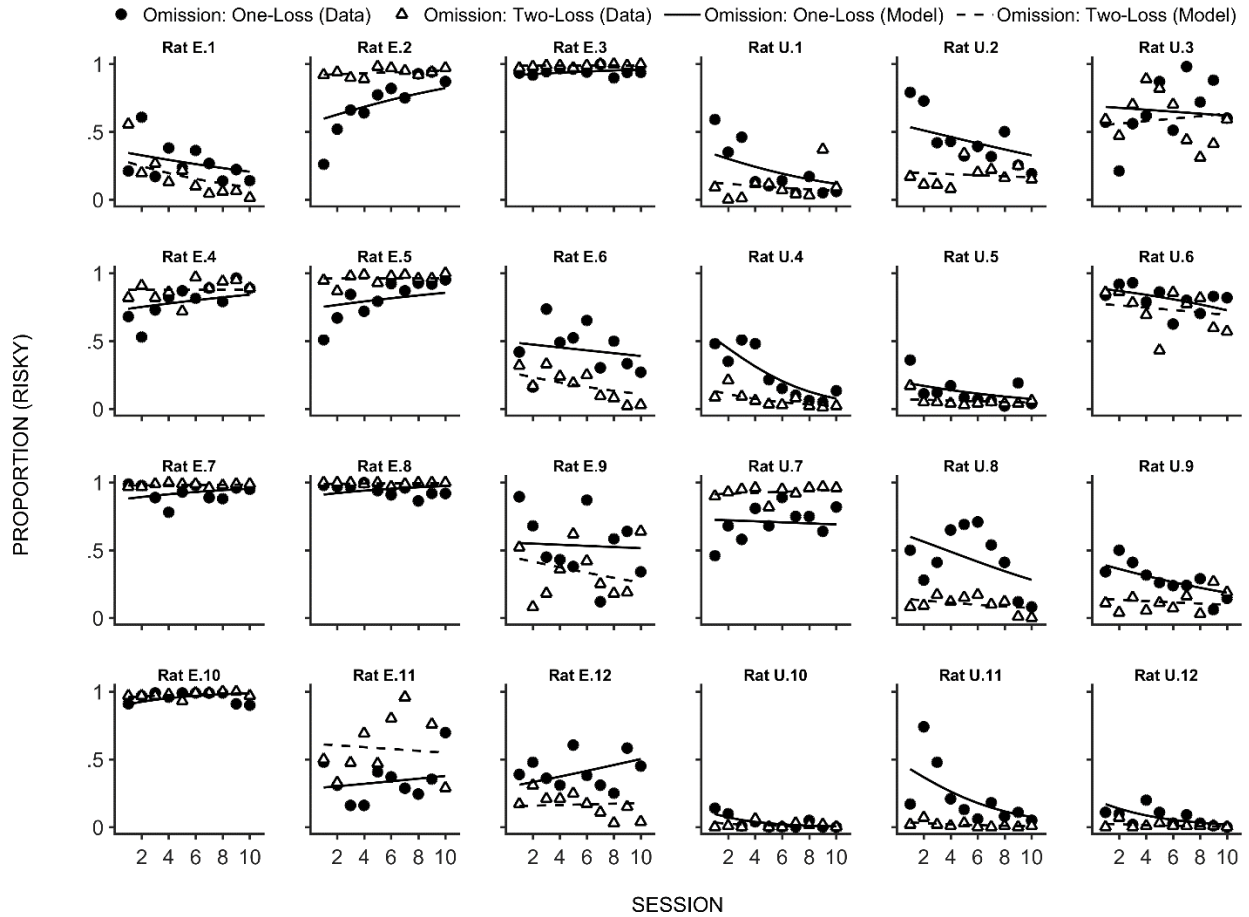


Figure 3.4 Proportion of choices for the risky outcome as a function of session and loss condition for individual rats within the Risk-Variability phases. The solid (Variability: One-Loss) and dashed lines (Variability: Two-Loss) represent the individual subject-fits from the best-fitting generalized linear mixed effects model. The rats in Group Equal-Risk are labeled as “E.” and the rats in Group Unequal-Risk are labeled as “U.”

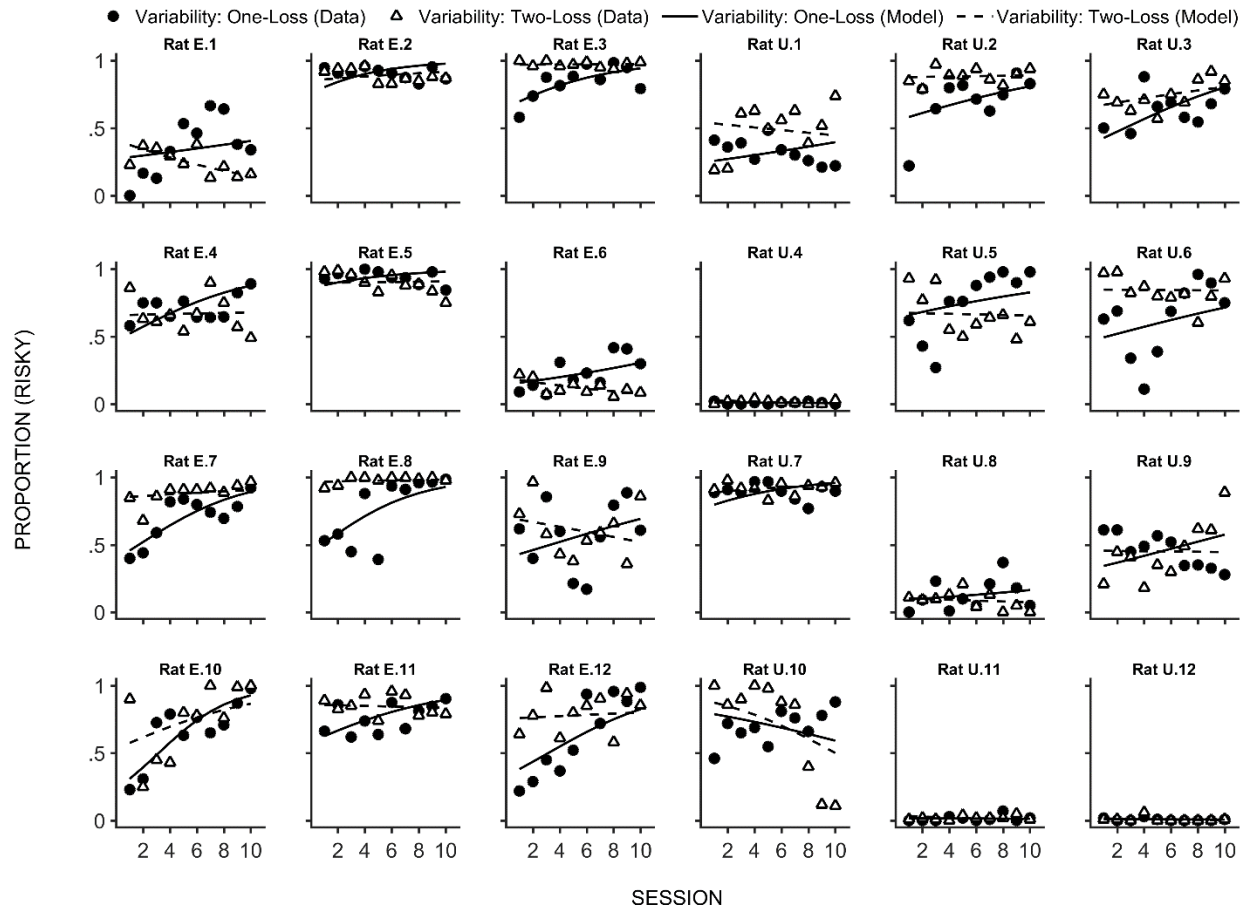


Figure 3.5 Proportion of choices for the risky outcome for Groups Equal-Risk and Unequal-Risk as a function of loss condition (One-Loss, Two-Loss) and the outcome of the previous choice. The dotted lines connecting the data points are intended to demonstrate the relationship between risky choice behavior following differential outcomes rather than to imply a continuous relationship. C-2 = certain choice – 2 pellets; C-4 = certain choice – 4 pellets; R-0 = risky choice – 0 pellets; R-1 = risky choice – 1 pellet; R-11 = risky choice – 11 pellets.

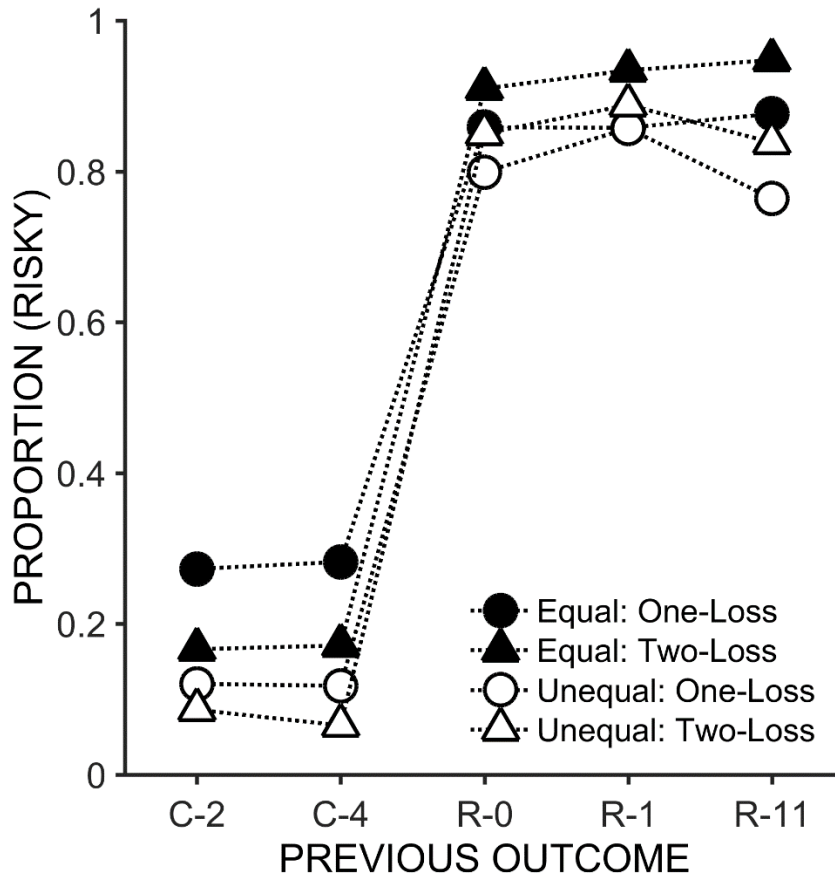


Figure 3.6 Proportion of choices for the risky outcome for Groups Equal-Risk and Unequal-Risk as a function of loss condition (One-Loss, Two-Loss) and the outcome of the previous choice in the Risk-Omission (left) and Risk-Variability phases (right). The dotted lines connecting the data points are intended to demonstrate the relationship between risky choice behavior following differential outcomes rather than to imply a continuous relationship. C-2 = certain choice – 2 pellets; C-4 = certain choice – 4 pellets; R-0 = risky choice – 0 pellets; R-1 = risky choice – 1 pellet; R-11 = risky choice – 11 pellets.

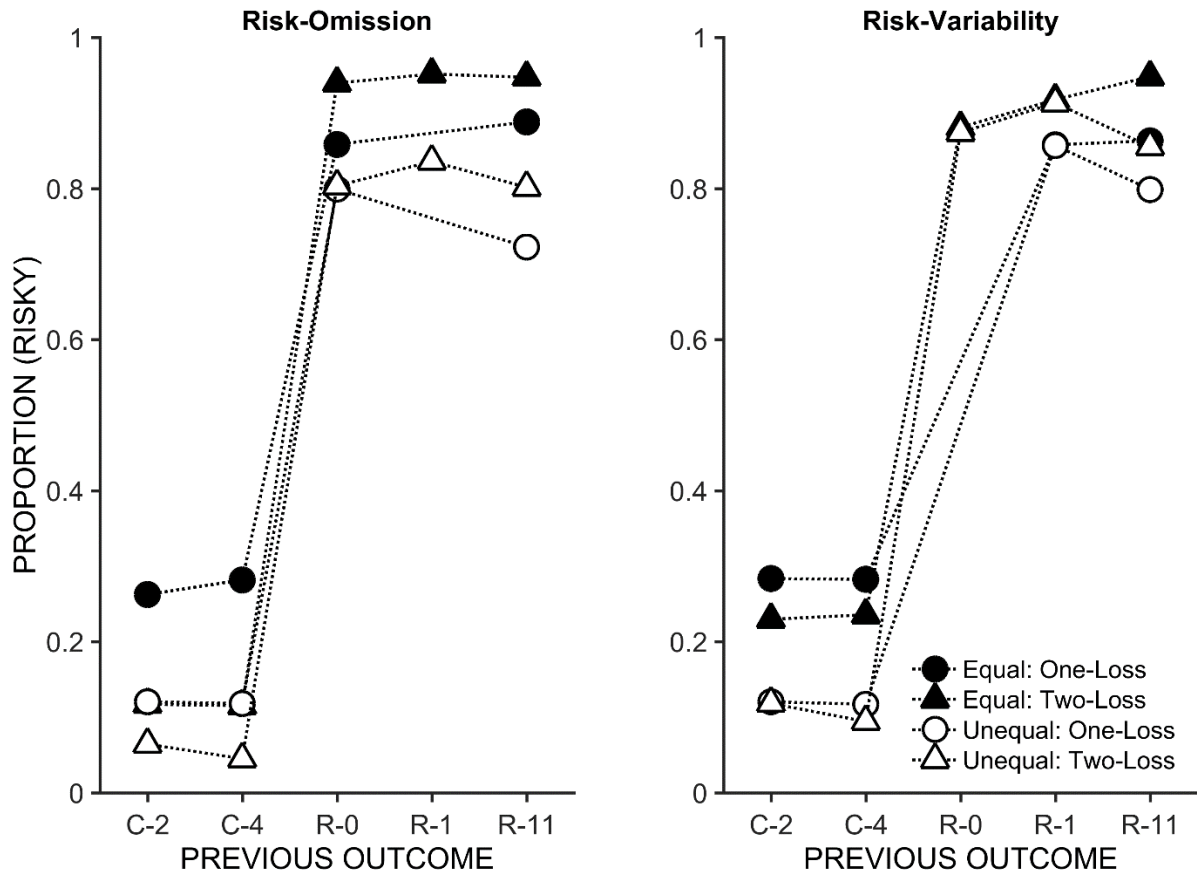


Figure 3.7 Proportion of choices for the risky outcome as a function of the outcome of the previous choice and loss condition for individual rats in the Risk-Omission phases. The dotted (Omission: One-Loss) and dashed lines (Omission: Two-Loss) represent the individual subject-fits from the best-fitting generalized linear mixed effects model. The rats in Group Equal-Risk are labeled as “E.” and the rats in Group Unequal-Risk are labeled as “U.”.

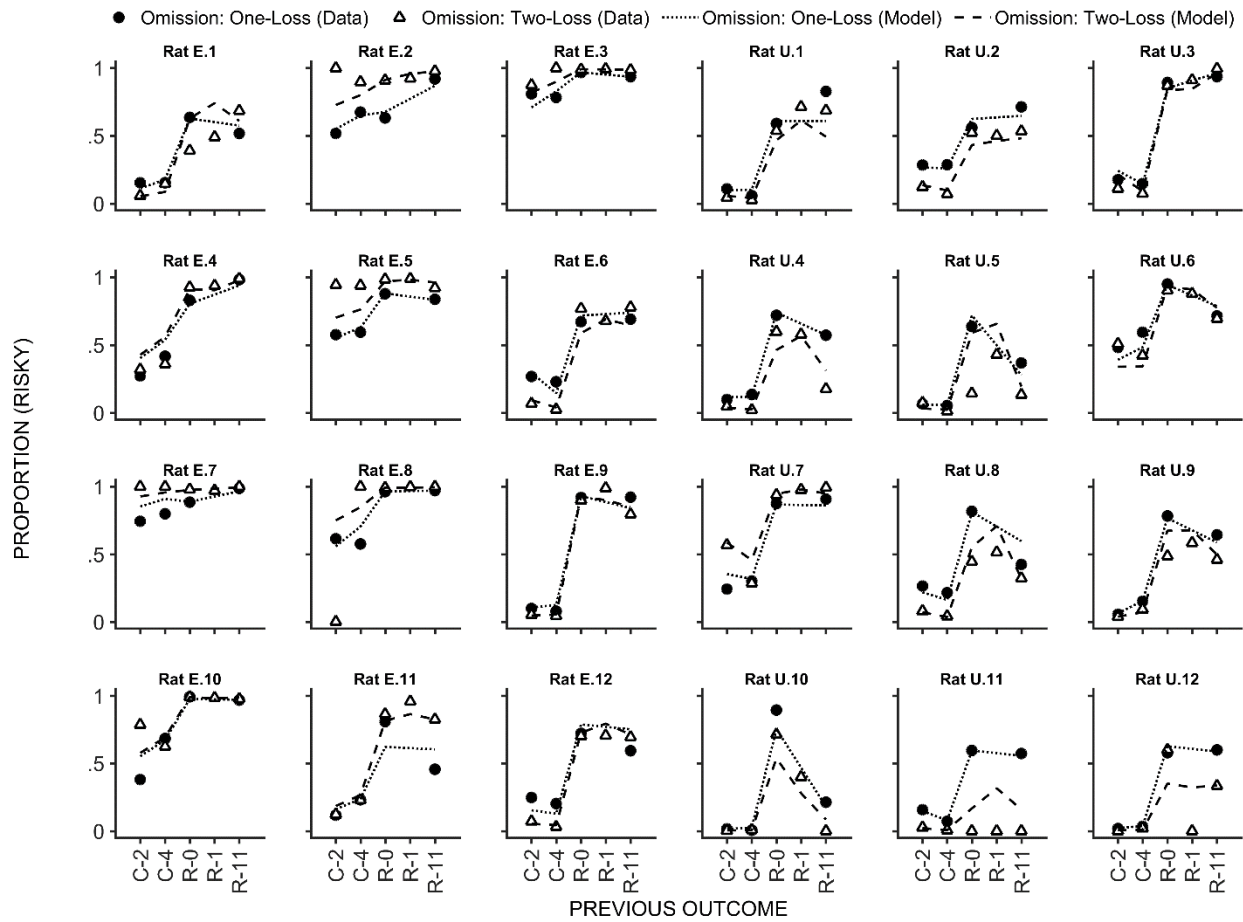


Figure 3.8 Proportion of choices for the risky outcome as a function of the outcome of the previous choice and loss condition for individual rats in the Risk-Variability phases. The dotted (Variability: One-Loss) and dashed lines (Variability: Two-Loss) represent the individual subject-fits from the best-fitting generalized linear mixed effects model. The rats in Group Equal-Risk are labeled as “E.” and the rats in Group Unequal-Risk are labeled as “U.”.

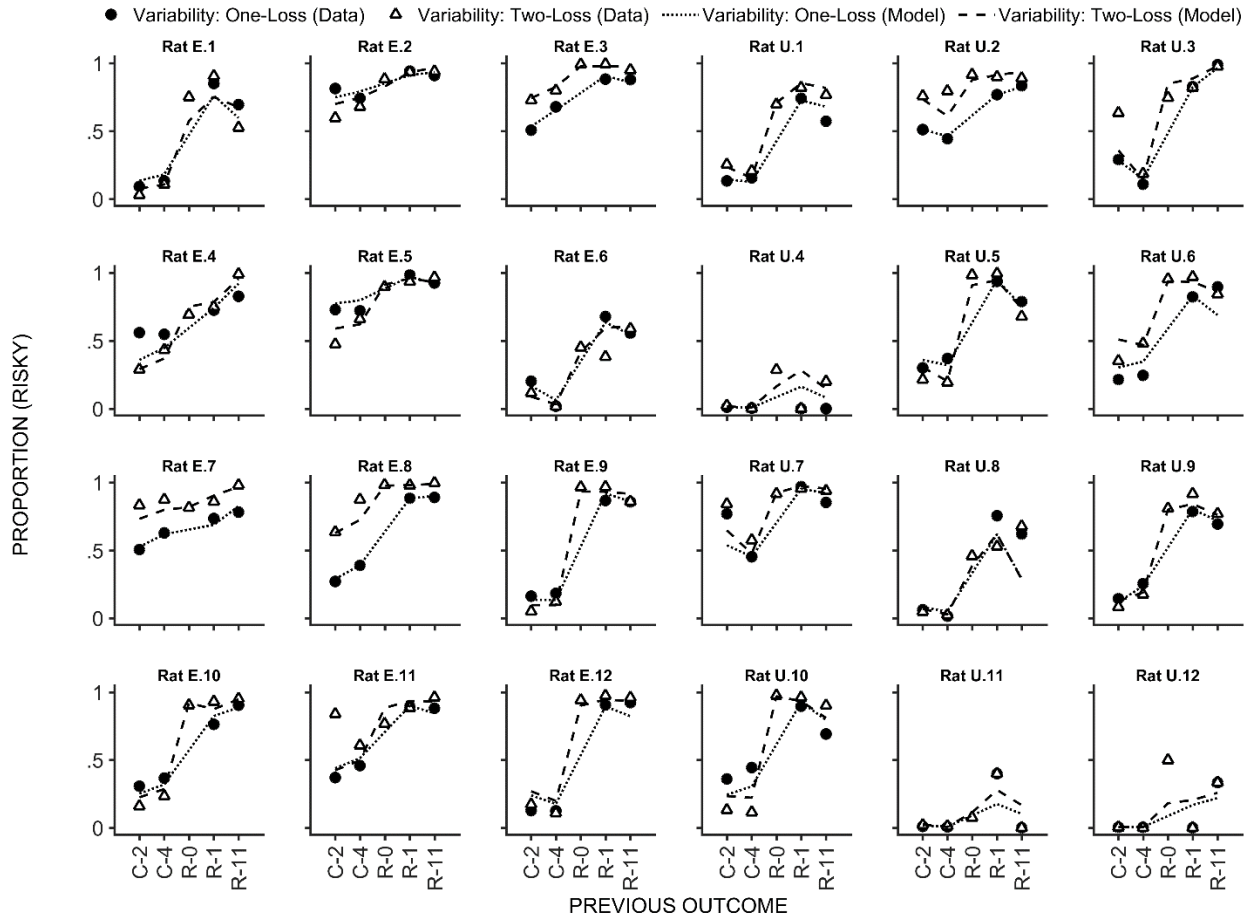


Figure 3.9 Regression coefficients from the mixed-effects analysis evaluating the effect of the previous 10 outcomes on subsequent risky choice. E = Group Equal-Risk; U = Group Unequal-Risk; Hyp. = hyperbolic model fit; Exp. = exponential model fit.

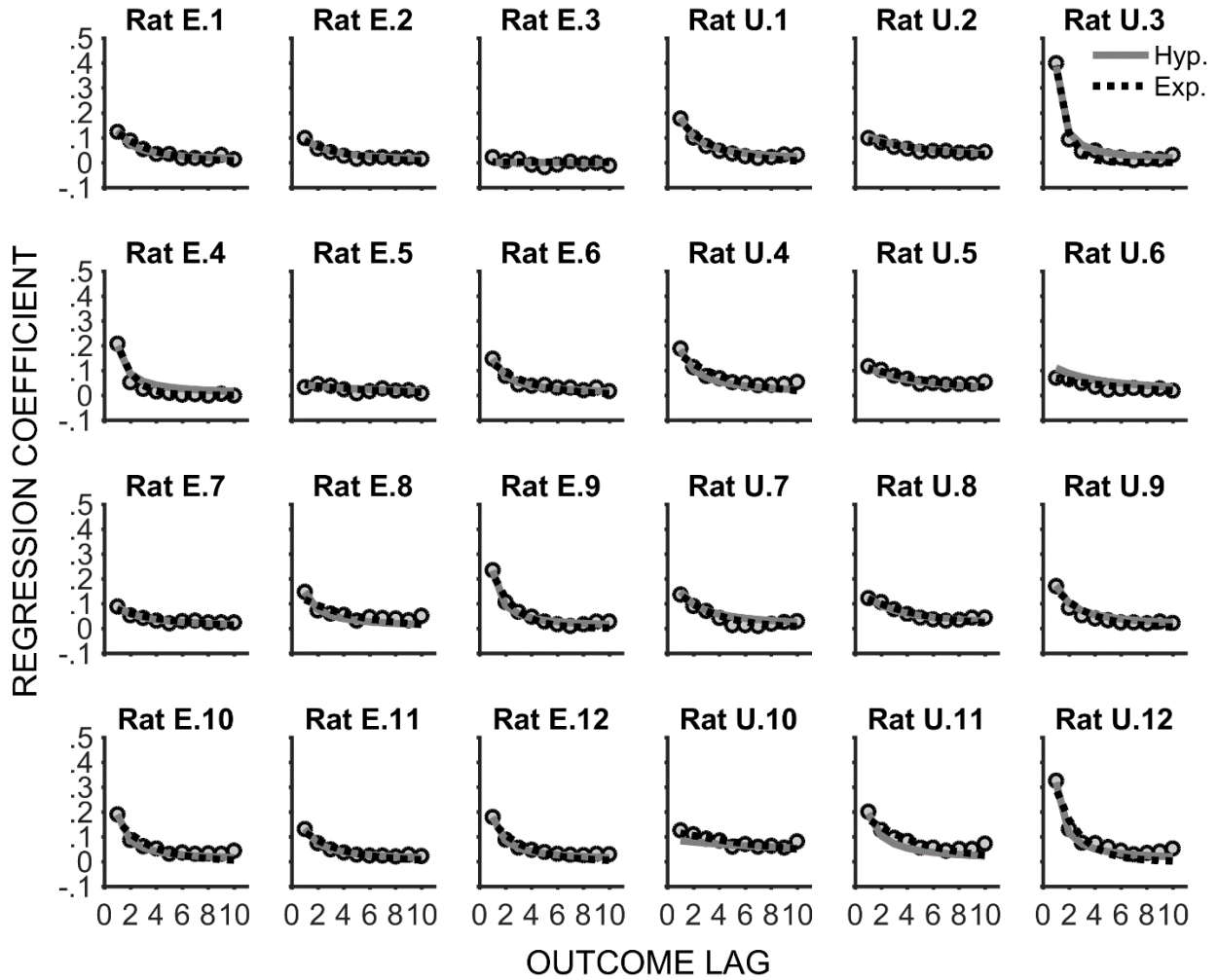


Figure 3.10 Normalized regression coefficients from the mixed-effects analysis evaluating the effect of the previous 10 choices on subsequent risky choice. E = Group Equal-Risk; U = Group Unequal-Risk; Hyp. = hyperbolic model fit; Exp. = exponential model fit.

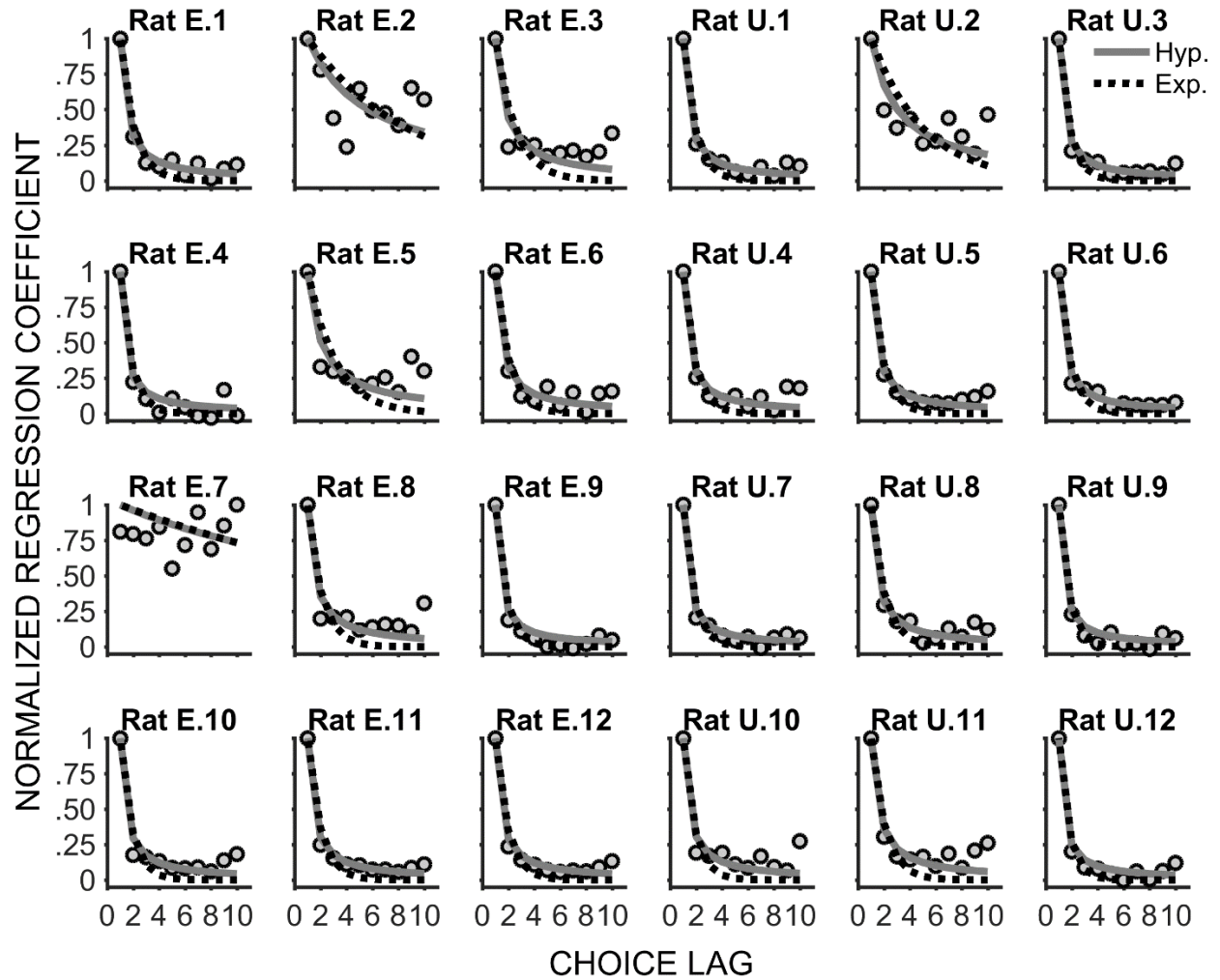


Figure 3.11 Parameter recovery for the Simple, Asymmetric, Valence-Attentive, and Weighted-Reference-Point models in Experiment 2. The thick dark lines are the unit diagonals.

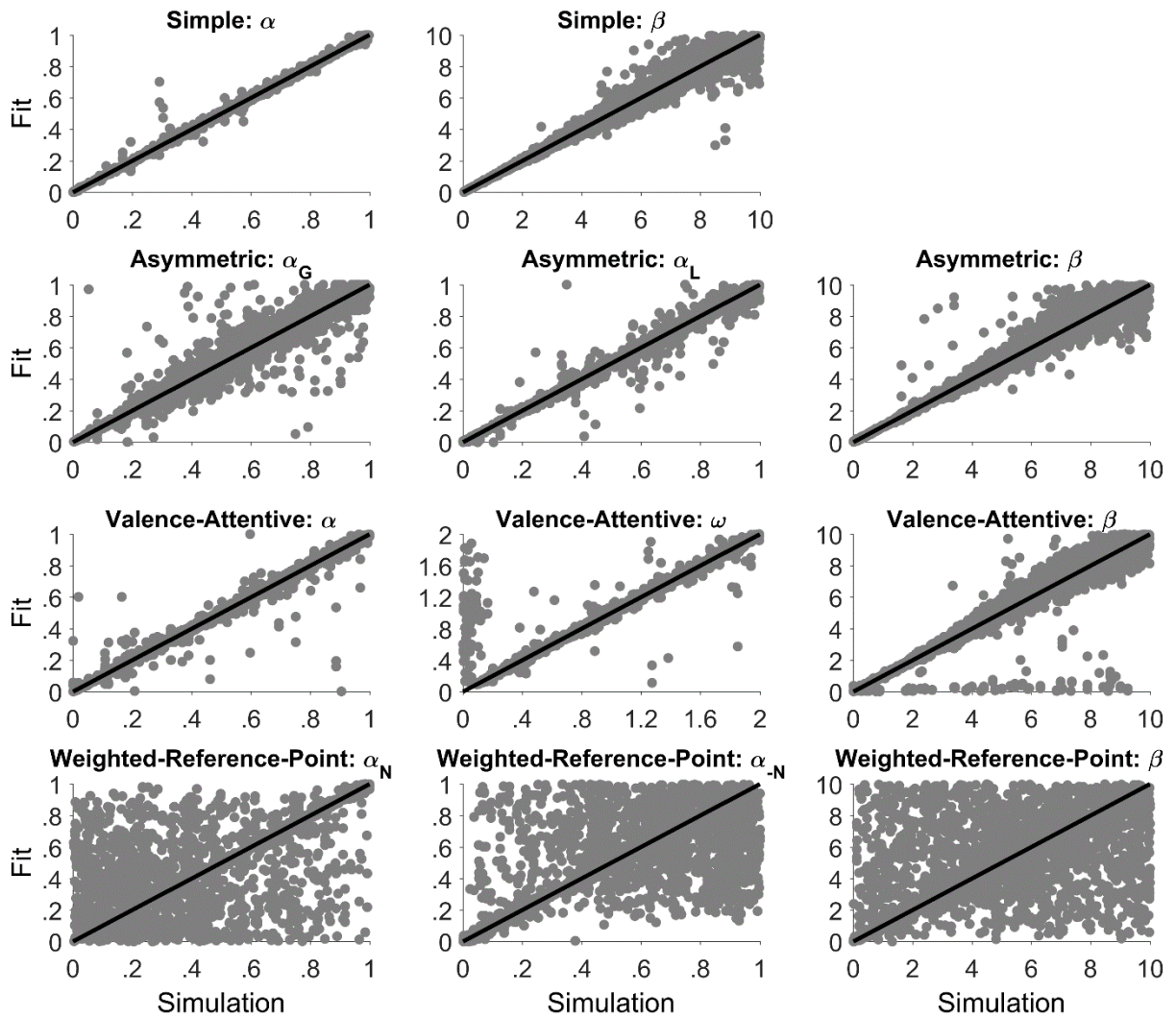


Figure 3.12 Akaike Information Criterion (AIC) for the Asymmetric Reinforcement Learning (RL) model plotted against the AICs of the Simple and Valence-Attentive RL models. Each data point represents the fit to an individual rat's data. Model fits that did not converge were omitted.

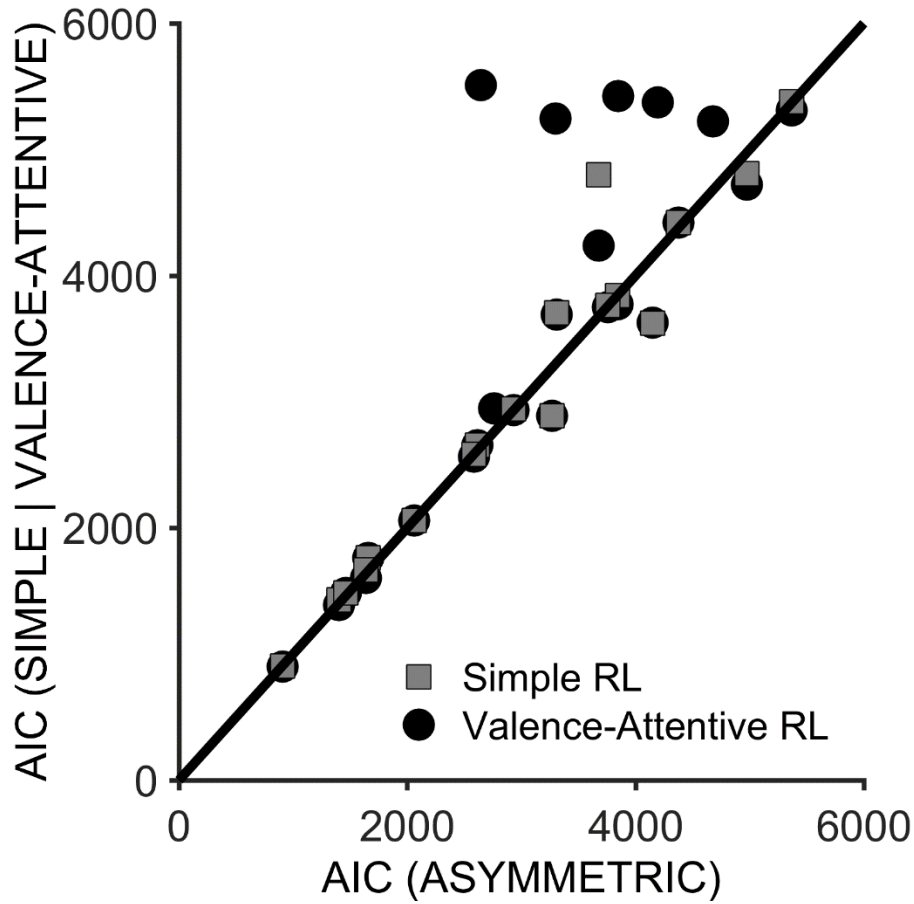


Figure 3.13 Choice data (Data) and the predicted choice data from the Asymmetric Reinforcement Learning (RL) model (Model) for each rat, ordered from top to bottom by how well the model fit the data in terms of omega-squared (ω^2). “Min” and “Max” refer to the data with the lowest and highest ω^2 , respectively. “25%ile”, “Mdn”, and “75%ile” refer to the 25th, 50th, and 75th percentile of ω^2 , respectively. The alternating gray and white boxes within each panel represent different phases and sub-phases (see Table 3.1). The differential alignments of these phase markers reflect individual differences in trials completed within each phase.

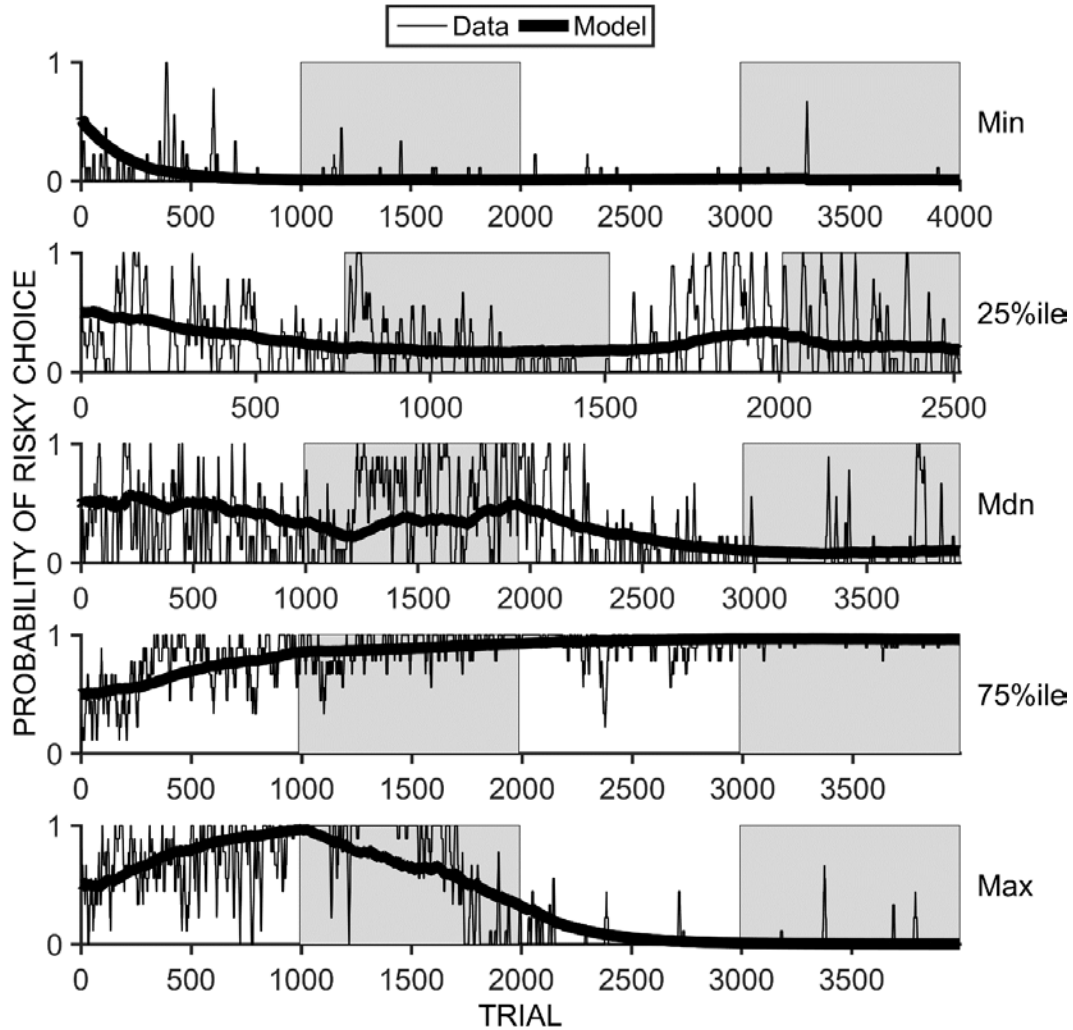


Figure 3.14 Mean risky choice plotted against the log-transformed α_G (left), α_L (center), and β parameters (right) as derived from the fitting of the Asymmetric Reinforcement Learning (RL) model. Each data point represents an individual rat. The goodness-of-fit index of the best fitting regression line for the overall model is shown (R^2).

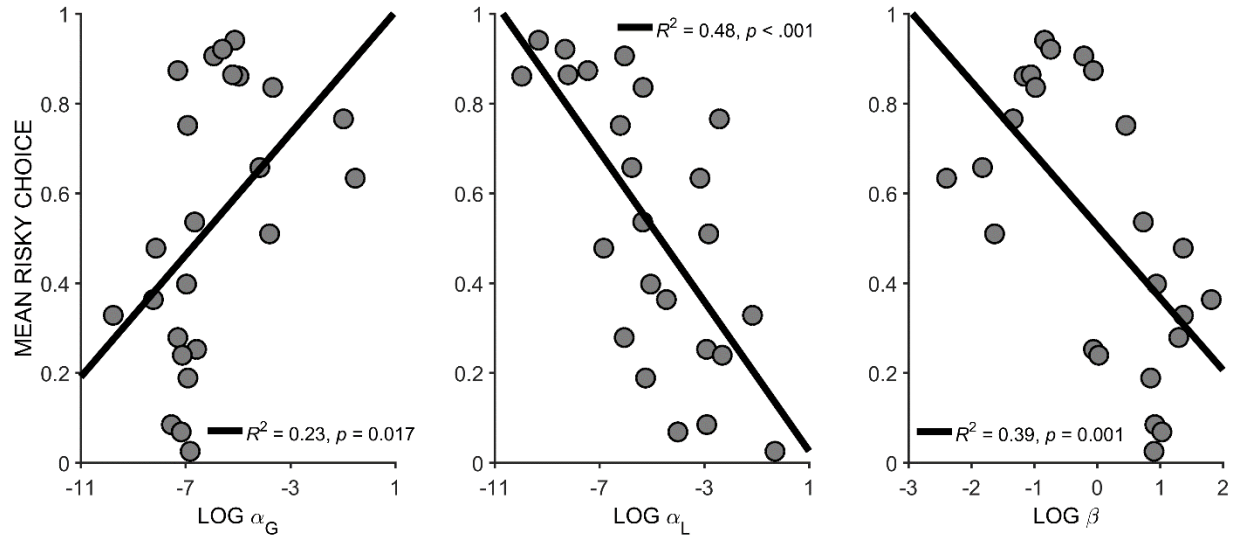


Figure 3.15 Win-Stay Index plotted against the log-transformed α_G (left panels), α_L (middle panels), and α_L to α_G ratio (right panels) as derived from the fitting of the Asymmetric Reinforcement Learning (RL) model. Each data point represents an individual rat. The rats in Group Unequal-Risk and Group Equal-Risk are distinguished.

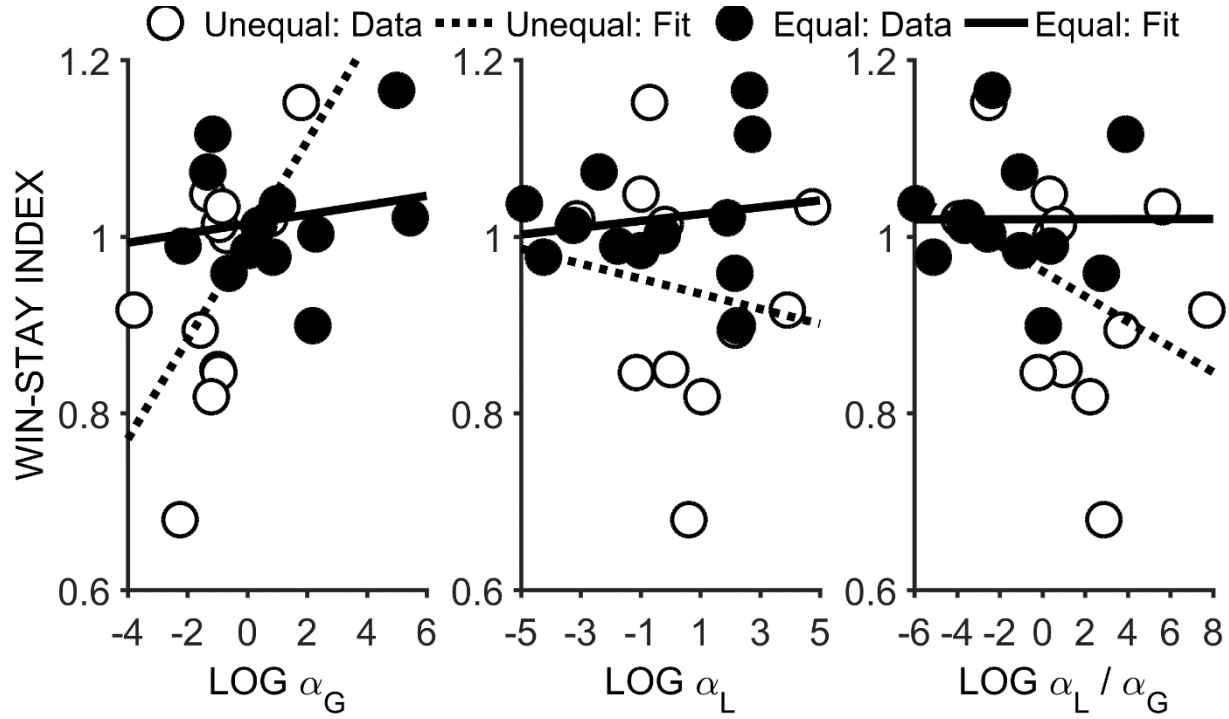


Figure 3.16 Loss Sensitivity Index from the Risk-Omission phases plotted against the log-transformed α_G (left panels), α_L (middle panels), and α_L to α_G ratio (right panels) as derived from the fitting of the Asymmetric Reinforcement Learning (RL) model. Each data point represents an individual rat. The rats in Group Unequal-Risk and Group Equal-Risk are distinguished.

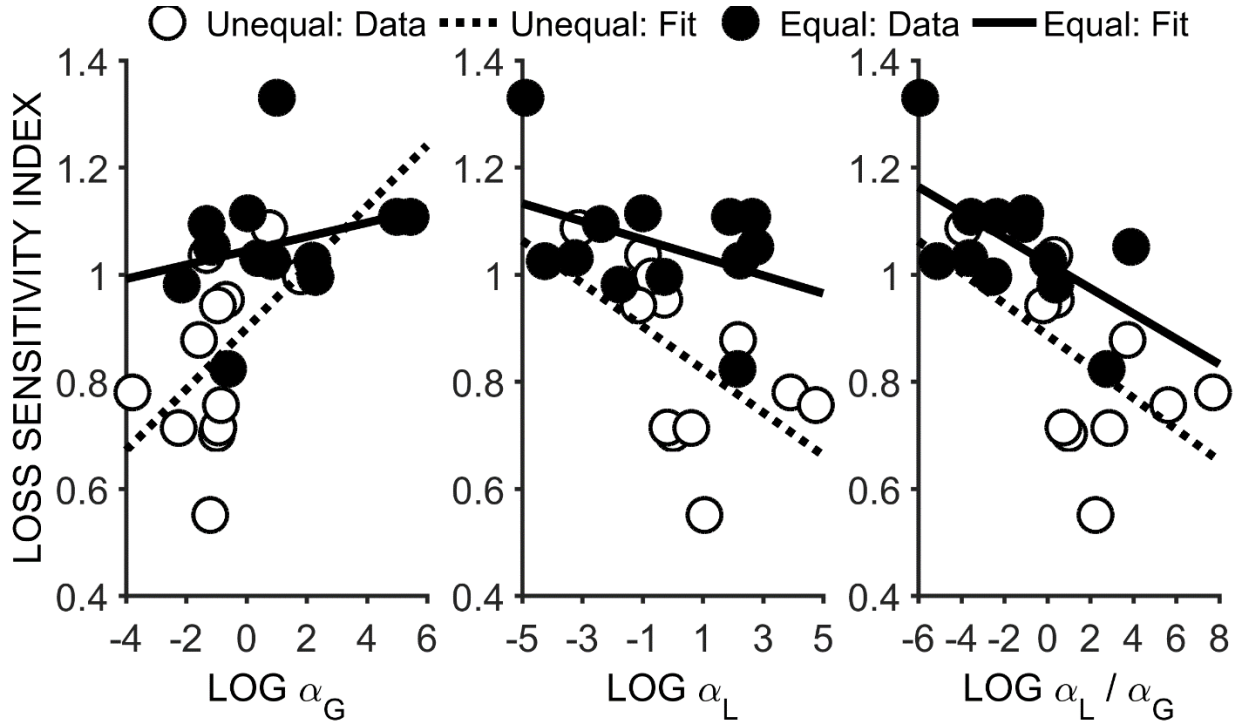
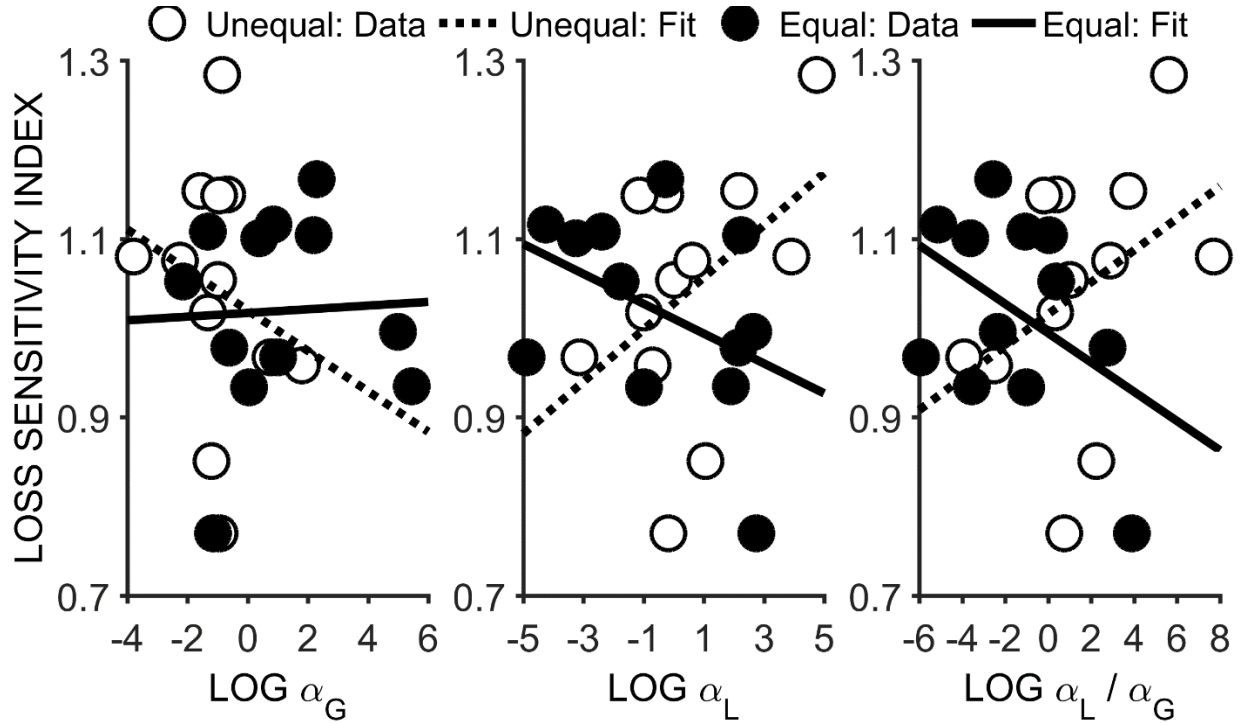


Figure 3.17 Loss Sensitivity Index from the Risk-Variability phases plotted against the log-transformed α_G (left panels), α_L (middle panels), and α_L to α_G ratio (right panels) as derived from the fitting of the Asymmetric Reinforcement Learning (RL) model. Each data point represents an individual rat. The rats in Group Unequal-Risk and Group Equal-Risk are distinguished.



Chapter 4 - General Discussion

Even though risky decision making has been proposed to be more affected by losses than gains (Anselme & Robinson, 2013), the psychological and neurobiological mechanisms of loss-processing have yet to become fully understood (Seymour, Maruyama, & De Martino, 2015). Moreover, there has been considerable difficulty in evaluating loss processing in animal models of risky decision making (see Clark, Averbach, et al., 2013; Cocker & Winstanley, 2015). Specifically, the operationalization of losses is limited with animal models, as food reward is the primary outcome type in the corresponding paradigms. As the consumption of reward cannot be readily undone in the same fashion that monetary rewards can be gained and then lost in human experiments, it has become necessary for animal research to establish a manipulation of “loss” that will have implications for choice behavior in humans. Accordingly, the primary goal across the present experiments was to investigate the mechanisms of loss processing and the corresponding behaviors in the context of a risky choice environment. These mechanisms were explored through determination of the effects of manipulating the feedback and information presented to rats within these environments.

Experiment 1 was designed to explicitly model LDWs in rats in terms of differential feedback from risky choices. LDWs have been shown to impact human behavior within risky choice contexts (M. J. Dixon et al., 2015; Jensen et al., 2012). As rats are strong models of human risky decision making, the modelling of such phenomena in rats provides future and unique opportunities to investigate loss-chasing behavior in rats and develop various behavioral and/or pharmacological therapies to reduce risky behaviors in individuals prone to make such decisions. Moreover, while there has been previous research investigating the potential neural correlates of gain- and loss-processing (den Ouden et al., 2013; Gehring & Willoughby, 2002;

Levin et al., 2012; Tom, Fox, Trepel, & Poldrack, 2007; van Holst, van den Brink, Veltman, & Goudriaan, 2010; Xue, Lu, Levin, & Bechara, 2011; Yacubian et al., 2006; Zhong et al., 2009), as well as those of more ambiguous losses (Clark, Lawrence, Astley-Jones, & Gray, 2009; Dymond et al., 2014; Habib & Dixon, 2010; Qi, Ding, Song, & Yang, 2011; van Holst, Chase, & Clark, 2014; Worhunsky et al., 2014), the implementation of LDWs in neurobiological analyses of risky choice behavior would considerably advance our understanding of loss-processing and value computation. Recently, Barrus and Winstanley (2016) reported that the accompaniment of stimulus cues with risky gains (cued task) resulted in more suboptimal choice behavior relative to behavior within a condition in which there were no such stimulus cues (uncued task); interestingly, in the cued task, but not in the uncued task, choice behavior was modulated by agonism and antagonism of D₃ receptors, which have been linked to both risk-taking behavior and addiction (Kreek, Nielsen, Butelman, & LaForge, 2005). Accordingly, the pairing of gain-related stimuli with not only gains (Barrus & Winstanley, 2016), but with losses (LDWs) may provide the most crucial insight into the mechanisms of behavioral addiction (e.g., pathological gambling). However, to our knowledge, there has yet to be a published report investigating the effects of explicit LDWs in rats. In Experiment 1 of the current report, LDWs were modelled by presenting gain-related stimuli to Group Extra-Feedback following all risky losses. Thus, these risky losses were assumed to be LDWs.

In Experiment 1, Group Extra-Feedback made more risky choices than Group Normal-Feedback, such that the provision of extra feedback following risky losses (R-0, R-1) encouraged greater risk taking at the molar and molecular levels (Figures 2.2-2.6). These results were paralleled by the increased goal-tracking behavior following risky outcome delivery in Group Extra-Feedback, which suggested that the rats in Group Extra-Feedback expected more food to

be delivered than the amount that actually was delivered. Furthermore, as revealed in the outcome and choice history analyses, Group Extra-Feedback's trial-by-trial choice behavior was relatively less affected by past reinforcement relative to more recent reinforcement in comparison to Group Normal-Feedback. Thus, differential risky outcome feedback, such as that presented by gambling machines in casinos, may encourage greater risk taking by discouraging the influence of past gambles (and losses) on future behavior (e.g., poorer memories for past outcomes). Interestingly, previous research has also suggested that better memories for past gains is predictive of greater likelihoods to make risky choices (Madan et al., 2014). Therefore, risky choice may be partially driven by an interaction between the memories for gains and losses, as well as the separate influences of gains and losses on subsequent risky choice. Accordingly, future research should continue exploring the cognitive and neurobiological mechanisms driving risky decision making behaviors in the absence and presence of environmental stimuli that have been suggested to increase risk taking (i.e., LDWs). Moreover, many previous risky choice experiments have not disambiguated differential risky losses, in that different risky outcomes are presented following trial termination without differential external cues. These experimental paradigms may be closer approximations to the environment experienced by Group Extra-Feedback rather than that experienced by Group Normal-Feedback. Accordingly, given the differences reported here between Groups Extra-Feedback and Normal-Feedback, future research may need to reconsider the paradigmatic conventions, as win-stay/lose-shift results in these previous reports may have been affected by the inconsistencies between risky outcome magnitude and the external cues that accompany the delivery of these outcomes.

In contrast to Experiment 1, Experiment 2 sought to determine how behavior within two traditional risky choice paradigms, Risk-Omission and Risk-Variability, was influenced by the

probabilistic addition of a second risky loss in the context of equivalent or differential probabilities of risky gains and losses. Here, risky choice was primarily affected by the conditions in the One-Loss sub-phases (i.e., when a risky choice probabilistically delivered a large R-11 gain or a smaller risky loss), which were well accounted for by the expected value of the risky choice. In the Two-Loss sub-phases (i.e., when the R-0, R-1, and R-11 outcomes were probabilistically delivered following risky choices), there was relatively little change in mean risky choice as a function of session, suggesting that the addition of a second loss did little to affect overall risky choice (Figure 3.2). However, there were substantial individual differences in behavior across the One- to Two-Loss sub-phases (see Figures 3.3 and 3.4), which may suggest that the minimal changes in mean behavior in the Two-Loss sub-phases were artifacts of averaging across individual rats that exhibited noticeably different patterns in behavior. Accordingly, individual differences analyses of risky choice may provide greater insight into the mechanisms of risky decision making behaviors, as averaged individual differences may hide the true psychological processes governing such behavior (see Young, 2016).

While group-level overall risky choice was unchanged as a function of session in the Two-Loss phases, there were systematic differences in molecular risky choice behavior as a function of the previous outcome between the One- and Two-Loss sub-phases. Group Equal-Risk exhibited greater staying behavior in the Two-Loss than in the One-Loss sub-phases following both certain and risky choices. In contrast, Group Unequal-Risk showed general decreases in risky choice behavior, and these decreases were accentuated following certain choices. As described above, such behavioral differences may be attributed to the experience of LDW-type outcomes. That is, while Experiment 1 involved potential LDWs in terms of the pairing of gain-related auditory, visual, and tactile stimulation with risky losses, Experiment 2

included potential LDWs in terms of the integration of loss- and gain-related outcome information and the association of this information with the probability of receiving the large R-11 gain. Therefore, while both experiments have different applications, they mutually imply that differential manipulation of outcome feedback and the information associated with these outcomes can have considerable impact on risky choice. Whether the presentation of losses is accompanied by win-related stimuli or probabilistically substituted for risky gains or other risky losses, the current experiments have demonstrated that gains and losses may not be strictly coded as objective gains and losses, but that the encoding and response to such outcomes is dependent on individual differences in gain and loss processing (Barracough et al., 2004; Cox et al., 2015; Marshall & Kirkpatrick, 2015; Van den Berg, Franken, & Muris, 2011), as well as environmental manipulations that alter the perception and treatment of these gain and loss outcomes (see M. J. Dixon et al., 2015; Engelmann & Hein, 2013; Jensen et al., 2012). Therefore, the present experiments provide key insight into the behavioral mechanisms governing individual differences in risky choice and loss-chasing behavior in rats.

A key research question of the present report involved determining which of several established RL models best accounted for rats' behaviors within different risky choice contexts, as well as how the parameters of the best fitting RL models related to individual rats' behaviors. The RL models served to complement the mixed-effects models, providing a theory-driven process model (RL) in addition to classical statistical theory-free approaches (mixed-effects). Accordingly, there are strengths to both approaches: RL can account for individuals' real-time experiences (i.e., choice and outcome history) without discretizing experimental manipulations (e.g., probability of food, phase), while the mixed-effects models permit analysis of group-level and individual-level effects of various manipulations on risky choice. Accordingly, the use of

only one of these approaches in dynamic environments may ultimately fail to capture the critical pieces of information obtained by employing the other approach. The integration of both approaches to analyze behavior in dynamic environments may serve as the most direct way to unveil the critical mechanism-manipulation interactions that govern individual differences in risky decision making. Here, several RL models were fit to the rats' trial-by-trial data. Of these models, the Asymmetric RL algorithm emerged as the superior model, particularly in Experiment 1, consistent with empirical support for this algorithm in other paradigms (Donahue & Lee, 2015; Frank et al., 2007; Niv et al., 2012; but see den Ouden et al., 2013). Across both experiments, there were strong relationships between the fitted parameters of the Asymmetric RL model and the rats' mean risky choice behavior. Individual differences in the patterns of gain-based learning, loss-based learning, and exploration-exploitation tendencies strongly predicted individual differences in risky decision making in rats. Future research should continue to investigate the neurobiological mechanisms of such psychological phenomena, as doing so will strengthen our understanding of the determinants driving risky decision making.

In Experiment 1, there were greater loss-based than gain-based learning rates, suggesting that rats are prone to be loss averse (Bhatti et al., 2014; Marshall & Kirkpatrick, 2015), similar to humans (e.g., Kahneman & Tversky, 1979). While Donahue and Lee (2015) reported greater gain-based than loss-based learning rates in primates, suggestive of possible species differences, loss-based learning in their experiment was restricted to reward omission (also see Costa, Tran, Turchi, & Averbeck, 2015). Interestingly, loss aversion has been suggested to be a product of the distribution of gains and losses (Walasek & Stewart, 2015). Accordingly, as in Experiment 1, the presence of non-zero-valued losses, the greater likelihood of receiving losses, and/or the explicit manipulation of loss-related information may have promoted greater loss-based learning.

Furthermore, individual differences in loss-based learning rates predicted both win-stay behaviors as well as rats' sensitivities to losses of different magnitudes (i.e., in terms of risky choice behavior following differential losses). As gain-based learning rates did not significantly predict either of these molecular risky choice behaviors, the results from Experiment 1 suggest that individual differences in risky choice in the presence of differential external feedback from risky losses are primarily governed by how well these individuals learn from experienced losses; the less effectively that the rats learned from losses, the less sensitive they were to differential losses and the more likely they were to chase risky losses (relative to how often they followed risky gains with risky choices). Furthermore, the inverse of the Win-Stay Index in Experiment 1 is a metric of loss-chasing behavior, and the Loss Sensitivity Index in Experiment 1 is indicative of rats' differential loss-chasing behaviors given differentially sized losses. Accordingly, the rats that did not learn as readily from losses were more likely to chase losses with risky choices and were less sensitive to differences in risky loss magnitudes. Therefore, extreme risk-taking behaviors in the context of drug addiction or gambling may be explained in terms of learning deficits (see Clark, Liu, et al., 2013); individuals who do not learn as readily from losses may be less sensitive to loss and, therefore, more likely to follow risky outcomes with additional risky choices (also see Rachlin, 1990).

As in Experiment 1, the Asymmetric RL model provided the best account of the rats' data in Experiment 2, relative to the other models tested, but there were considerable individual differences in the model fitting. Here, the metric of win-stay behavior was better predicted by individual differences in gain-based learning rates, again suggesting that subjective learning/value-updating rates are critical determinants of risky choice behavior in rats. The difference between experiments in terms of the model parameter that best predicted win-stay

rates is likely due to procedural differences, as the parameters of risky losses in Experiment 1 were seemingly more varied across conditions than they were in Experiment 2. Thus, loss-based learning rates may emerge as a better predictor of choice behavior following risky gains versus losses when the contingencies surrounding the delivery of risky losses are more dynamic. However, as in Experiment 1, the loss-sensitivity metric of loss-chasing behavior in Experiment 2 was similarly predicted by individual differences in loss-based learning rates. Therefore, the results from the RL analyses in both experiments collectively suggest that individual differences in differential gain- and loss-based learning rates can predict individual differences in molar and molecular measures of risky choice in rats. Theoretically, elevated risky decision making in pathological gamblers in the context of LDWs may be driven by a reduced sensitivity to differential losses and deficits in loss-based learning. Thus, future modelling approaches would do well to consider Asymmetric RL-type models to most effectively capture individual differences in risky choice behavior across various conditions and contexts.

While the current results supported the existence of separate gain- and loss-based learning rates, there were instances in which the Simple and Valence-Attentive RL models provided better accounts of individual rat's data; this was more common in Experiment 2 than in Experiment 1. Furthermore, the Asymmetric RL model fit the data in Experiment 2 less convincingly than it did in Experiment 1. Accordingly, while the Asymmetric RL model was the dominant model of those tested, it is likely that the currently tested RL models will need to be expanded to more effectively account for the trial-by-trial risky choice patterns across rats. Indeed, as seen in Figures 2.13 and 3.13, there was considerable noise in the rats' trial-by-trial choice behavior that was not captured by the model. Whether such noise is in fact noise or systematic patterns that can be precisely modelled with further modifications of RL models is

left to be determined. Alternatively, the poorer performances of the other models, as well as the inability for these models to converge on viable solutions, may reflect the fact that the starting values for the model fits were sampled from a uniform distribution. Indeed, Gershman (2016) argued that empirically-determined non-uniform distributions of such starting values (i.e., priors) can improve the fitting of RL models to data, such that the inability to converge on viable solutions can be avoided. Thus, the distribution of fitted parameters from the present experiment provides a key starting point for the determination of the distribution of such empirical priors.

In addition to the elucidation of risky choice mechanisms by the RL analyses, the outcome and choice history analyses provided critical awareness into the effects of past experiences on subsequent behavior. Previous reports have employed comparable analyses to investigate how certain experimental manipulations (e.g., brain lesions) impact the number of past choices and/or outcomes that influence subsequent choice behavior (e.g., Kennerley et al., 2006). As described above, there have been several other previous analyses of outcome and choice history, but, to our knowledge, there have been no reports describing the functional relationship of these regression coefficients. Here, the functional decays of the regression coefficients were better characterized by hyperbolic than exponential functions (Figures 2.7, 2.8, 3.9, and 3.10). Lau and Glimcher (2005) analyzed the decaying regression coefficients from two primates and noted that such decays could take a hyperbolic or exponential form; however, they noted that the use of only two subjects prevented their ability to statistically determine whether the hyperbolic or exponential form provided better fits to the data. Accordingly, with added power in the current analyses, these results suggest that the weights of previous outcomes and choices decay hyperbolically as these events recede into the past. Interestingly, support for the hyperbolic model stands in contrast to traditional assumptions of RL algorithms, which assume

exponentially decaying weights of past outcomes on current estimates of value of the corresponding choice. As noted above, there has been recent consideration of hyperbolic RL models (Alexander & Brown, 2010; Maia, 2009). The use of the traditional exponential RL models in the current paradigm advanced our understanding of the mechanisms of risky choice, but future implementation of simple hyperbolic RL models, as well as modified hyperbolic RL models that account for differential psychological phenomena (e.g., attention), may provide the best account of the patterns in trial-by-trial risky decision making across different individuals.

In conclusion, the results of the present experiments has provided novel insight into the mechanisms of loss processing in individual rats in risky choice environments. The differential feedback manipulation of Experiment 1 represents a novel approach to understanding the processing of LDWs in non-human animals, an approach that can be further studied to understand the psychopharmacological effects and neurobiological underpinnings of loss processing, loss-chasing, LDWs, and risky choice. Such further experimentation may permit the development of subject-specific pharmacological and/or neurocognitive therapies to alleviate maladaptive risky decision making tendencies (also see M. R. Dixon, Wilson, & Habib, 2014). Similarly, the differential risk treatments in Experiment 2 provided critical insight into the effects of additional risky losses in standard risky choice paradigms. While Experiment 1 may have more translational potential, Experiment 2 may serve to more greatly advance our understanding of the cognitive processes of risky choice behavior (e.g., how differential losses are processed relative to other losses). However, both experiments have served to strengthen our understanding of the mechanisms of loss-based phenomena. Specifically, these results have suggested that individual differences in such behaviors are primarily governed by individual differences in gain- and loss-based learning. Indeed, previous research has suggested potentially

distinct neurobiological correlates of gain- and loss-based learning and valuation (den Ouden et al., 2013; Doya, 2008; Levin et al., 2012; Yacubian et al., 2006; Zhong et al., 2009). Therefore, future pharmacological and neurobiological manipulations of risky choice behavior must consider the possibly orthogonal effects of such manipulations on gain- and loss-based changes in behavior (see Kubanek, Snyder, & Abrams, 2015). Otherwise, current understanding of the psychological, neurological, and neurocomputational mechanisms of choice behavior will remain stagnant. Proper treatment and consideration of such learning differences in future experiments and analyses should ultimately unveil the neural circuits that promote informed and adaptive decision making in various contexts.

References

- Addicott, M. A., Pearson, J. M., Kaiser, N., Platt, M. L., & McClernon, F. J. (2015). Suboptimal foraging behavior: a new perspective on gambling. *Behavioral Neuroscience, 129*, 656-665. doi: 10.1037/bne0000082
- Akaike, H. (1973). Information theory as an extension of the maximum likelihood principle. In B. N. Petrov & F. Csaki (Eds.), *Second International Symposium on Information Theory* (pp. 267-281). Budapest: Akademiai Kiado.
- Alexander, W. H., & Brown, J. W. (2010). Hyperbolically discounted temporal difference learning. *Neural Computation, 22*, 1511-1527. doi: 10.1162/neco.2010.08-09-1080
- Andrade, E. B., & Iyer, G. (2009). Planned versus actual betting in sequential gambles. *Journal of Marketing Research, 46*, 372-383. doi: 10.1509/jmkr.46.3.372
- Anselme, P. (2013). Dopamine, motivation, and the evolutionary significance of gambling-like behavior. *Behavioural Brain Research, 256*, 1-4. doi: 10.1016/j.bbr.2013.07.039
- Anselme, P., & Robinson, M. J. F. (2013). What motivates gambling behaviour? Insight into dopamine's role. *Frontiers in Behavioral Neuroscience, 7*. doi: 10.3389/fnbeh.2013.00182
- Bacotti, A. V. (1976). Home cage feeding time controls responding under multiple schedules. *Animal Learning & Behavior, 4*, 41-44. doi: 10.3758/BF03211983
- Barracough, D. J., Conroy, M. L., & Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nature Neuroscience, 7*, 404-410. doi: 10.1038/nn1209
- Barrus, M. M., & Winstanley, C. A. (2016). Dopamine D₃ receptors modulate the ability of win-paired cues to increase risky choice in a rat gambling task. *The Journal of Neuroscience, 36*, 785-794. doi: 10.1523/JNEUROSCI.2225-15.2016

- Bateson, M., & Kacelnik, A. (1995). Preferences for fixed and variable food sources: variability in amount and delay. *Journal of the Experimental Analysis of Behavior*, *63*, 313-329. doi: 10.1901/jeab.1995.63-313
- Battalio, R. C., Kagel, J. H., & MacDonald, D. N. (1985). Animals' choices over uncertain outcomes: some initial experimental results. *The American Economic Review*, *75*, 597-613.
- Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, *47*, 129-141. doi: 10.1016/j.neuron.2005.05.020
- Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, *275*, 1293-1295. doi: 10.1126/science.275.5304.1293
- Beeler, J. A., Daw, N., Frazier, C. R. M., & Zhuang, X. (2010). Tonic dopamine modulates exploitation of reward learning. *Frontiers in Behavioral Neuroscience*, *4*. doi: 10.3389/fnbeh.2010.00170
- Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*, 1214-1221. doi: 10.1038/nn1954
- Bhatti, M., Jang, H., Kralik, J. D., & Jeong, J. (2014). Rats exhibit reference-dependent choice behavior. *Behavioural Brain Research*, *267*, 26-32. doi: 10.1016/j.bbr.2014.03.012
- Bishara, A. J., Pleskac, T. J., Fridberg, D. J., Yechiam, E., Lucas, J., Busemeyer, J. R., . . . Stout, J. C. (2009). Similar processes despite divergent behavior in two commonly used measures of risky decision making. *Journal of Behavioral Decision Making*, *22*, 435-454. doi: 10.1002/bdm.641

- Black, D. W., Shaw, M., McCormick, B., & Allen, J. (2013). Pathological gambling: relationship to obesity, self-reported chronic medical conditions, poor lifestyle choices, and impaired quality of life. *Comprehensive Psychiatry*, *54*, 97-104. doi: 10.1016/j.comppsy.2012.07.001
- Breen, R. B., & Zuckerman, M. (1999). 'Chasing' in gambling behavior: personality and cognitive determinants. *Personality and Individual Differences*, *27*, 1097-1111.
- Brevers, D., Koritzky, G., Bechara, A., & Noël, X. (2014). Cognitive processes underlying impaired decision-making under uncertainty in gambling disorder. *Addictive Behaviors*, *39*, 1533-1536. doi: 10.1016/j.addbeh.2014.06.004
- Brevers, D., Noël, X., Bechara, A., Vanavermaete, N., Verbanck, P., & Kornreich, C. (2015). Effect of casino-related sound, red light and pairs on decision-making during the Iowa Gambling Task. *Journal of Gambling Studies*, *31*, 409-421. doi: 10.1007/s10899-013-9441-2
- Burnham, K. P., & Anderson, D. R. (1998). *Model Selection and Inference: A Practical Information-Theoretic Approach*. New York, NY: Springer.
- Busemeyer, J. R., & Stout, J. C. (2002). A contribution of cognitive decision models to clinical assessment: decomposing performance on the Bechara gambling task. *Psychological Assessment*, *14*, 253-262. doi: 10.1037//1040-3590.14.3.253
- Bush, R. R., & Mosteller, F. (1951). A mathematical model for simple learning. *Psychological Review*, *58*, 313-323. doi: 10.1037/h0054388
- Campbell-Meiklejohn, D., Simonsen, A., Scheel-Krüger, J., Wohlert, V., Gjerløff, T., Frith, C. D., . . . Møller, A. (2012). In for a penny, in for a pound: methylphenidate reduces the

- inhibitory effect of high stakes on persistent risky choice. *The Journal of Neuroscience*, 32, 13032-13038.
- Campbell-Meiklejohn, D., Woolrich, M. W., Passingham, R. E., & Rogers, R. D. (2008). Knowing when to stop: the brain mechanisms of chasing losses. *Biological Psychiatry*, 63, 293-300. doi: 10.1016/j.biopsych.2007.05.014
- Cardinal, R. N., & Howes, N. J. (2005). Effects of lesions of the nucleus accumbens core on choice between small certain rewards and large uncertain rewards. *BMC Neuroscience*, 6: 37. doi: 10.1186/1471-2202-6-37
- Clark, L., Averbeck, B., Payer, D., Sescousse, G., Winstanley, C. A., & Xue, G. (2013). Pathological choice: the neuroscience of gambling and gambling addiction. *The Journal of Neuroscience*, 33, 17617-17623. doi: 10.1523/JNEUROSCI.3231-13.2013
- Clark, L., Lawrence, A. J., Astley-Jones, F., & Gray, N. (2009). Gambling near-misses enhance motivation to gamble and recruit win-related brain circuitry. *Neuron*, 61, 481-490. doi: 10.1016/j.neuron.2008.12.031
- Clark, L., Liu, R., McKavanagh, R., Garrett, A., Dunn, B. D., & Aitken, M. R. F. (2013). Learning and affect following near-miss outcomes in simulated gambling. *Journal of Behavioral Decision Making*, 26, 442-450. doi: 10.1002/bdm.1774
- Cocker, P. J., & Winstanley, C. A. (2015). Irrational beliefs, biases and gambling: exploring the role of animal models in elucidating vulnerabilities for the development of pathological gambling. *Behavioural Brain Research*, 279, 259-273. doi: 10.1016/j.bbr.2014.10.043
- Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362, 933-942.

- Compton, W. M., Thomas, Y. F., Stinson, F. S., & Grant, B. F. (2007). Prevalence, correlates, disability, and comorbidity of DSM-IV drug abuse and dependence in the United States. *Archives of General Psychiatry*, *64*, 566-576. doi: 10.1001/archpsyc.64.5.566
- Costa, V. D., Tran, V. L., Turchi, J., & Averbeck, B. B. (2015). Reversal learning and dopamine: a Bayesian perspective. *The Journal of Neuroscience*, *35*, 2407-2416. doi: 10.1523/JNEUROSCI.1989-14.2015
- Cox, S. M. L., Frank, M. J., Larcher, K., Fellows, L. K., Clark, C. A., Leyton, M., & Dagher, A. (2015). Striatal D1 and D2 signaling differentially predict learning from positive and negative outcomes. *NeuroImage*, *109*, 95-101. doi: 10.1016/j.neuroimage.2014.12.070
- Crossman, E. (1983). Las Vegas knows better. *The Behavior Analyst*, *6*, 109-110.
- Daw, N. D. (2011). Trial-by-trial data analysis using computational models. In M. R. Delgado, E. A. Phelps & T. W. Robbins (Eds.), *Decision Making, Affect, and Learning: Attention and Performance* (Vol. 23, pp. 3-38).
- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*, 1204-1215. doi: 10.1016/j.neuron.2011.02.027
- Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*, 876-879. doi: 10.1038/nature04766
- Daw, N. D., & Tobler, P. N. (2013). Value learning through reinforcement: the basics of dopamine and reinforcement learning. In P. W. Glimcher & E. Fehr (Eds.), *Neuroeconomics* (2nd ed., pp. 283-298). London, UK: Academic Press.

- den Ouden, H. E. M., Daw, N. D., Fernandez, G., Elshout, J. A., Rijpkema, M., Hoogman, M., . . . Cools, R. (2013). Dissociable effects of dopamine and serotonin on reversal learning. *Neuron, 80*, 1090-1100. doi: 10.1016/j.neuron.2013.08.030
- Devenport, L., Hill, T., Wilson, M., & Ogden, E. (1997). Tracking and averaging in variable environments: a transition rule. *Journal of Experimental Psychology: Animal Behavior Processes, 23*, 450-460. doi: 10.1037/0097-7403.23.4.450
- Dixon, M. J., Collins, K., Harrigan, K. A., Graydon, C., & Fugelsang, J. A. (2015). Using sound to unmask losses disguised as wins in multiline slot machines. *Journal of Gambling Studies, 31*, 183-196. doi: 10.1007/s10899-013-9411-8
- Dixon, M. J., Harrigan, K. A., Sandhu, R., Collins, K., & Fugelsang, J. A. (2010). Losses disguised as wins in modern multi-line video slot machines. *Addiction, 105*, 1819-1824. doi: 10.1111/j.1360-0443.2010.03050.x
- Dixon, M. J., Harrigan, K. A., Santesso, D. L., Graydon, C., Fugelsang, J. A., & Collins, K. (2014). The impact of sound in modern multiline video slot machine play. *Journal of Gambling Studies, 30*, 913-929. doi: 10.1007/s10899-013-9391-8
- Dixon, M. R., Wilson, A., & Habib, R. (2014). Neurological correlates of slot machine win size in pathological gamblers. *Behavioural Processes, 104*, 108-113. doi: 10.1016/j.beproc.2014.02.007
- Doll, B. B., Simon, D. A., & Daw, N. D. (2012). The ubiquity of model-based reinforcement learning. *Current Opinion in Neurobiology, 22*, 1075-1081. doi: 10.1016/j.conb.2012.08.003
- Domjan, M. (2010). *The Principles of Learning and Behavior* (6th ed.). Belmont, CA: Wadsworth.

- Donahue, C. H., & Lee, D. (2015). Dynamic routing of task-relevant signals for decision making in dorsolateral prefrontal cortex. *Nature Neuroscience*, *18*, 295-301. doi: 10.1038/nn.3918
- Dong, G., Hu, Y., & Lin, X. (2013). Reward/punishment sensitivities among internet addicts: implications for their addictive behaviors. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *46*, 139-145. doi: 10.1016/j.pnpbp.2013.07.007
- Doya, K. (2008). Modulators of decision making. *Nature Neuroscience*, *11*, 410-416. doi: 10.1038/nn2077
- Dunlap, A. S., & Stephens, D. W. (2012). Tracking a changing environment: optimal sampling, adaptive memory and overnight effects. *Behavioural Processes*, *89*, 86-94.
- Dymond, S., Lawrence, N. S., Dunkley, B. T., Yuen, K. S. L., Hinton, E. C., Dixon, M. R., . . . Singh, K. D. (2014). Almost winning: induced MEG theta power in insula and orbitofrontal cortex increases during gambling near-misses and is associated with BOLD signal and gambling severity. *NeuroImage*, *91*, 210-219. doi: 10.1016/j.neuroimage.2014.01.019
- Engelmann, J. B., & Hein, G. (2013). Contextual and social influences on valuation and choice. In V. S. C. Pammi & N. Srinivasan (Eds.), *Progress in Brain Research* (Vol. 202, pp. 215-237). Amsterdam, The Netherlands: Elsevier.
- Evenden, J. L., & Robbins, T. W. (1984). Win-stay behaviour in the rat. *Quarterly Journal of Experimental Psychology*, *36B*, 1-26. doi: 10.1080/14640748408402190
- Floresco, S. B., St. Onge, J. R., Ghods-Sharifi, S., & Winstanley, C. A. (2008). Cortico-limbic-striatal circuits subserving different forms of cost-benefit decision making. *Cognitive, Affective, & Behavioral Neuroscience*, *8*, 375-389.

- Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences, 104*, 16311-16316. doi: 10.1073/pnas.0706111104
- Freestone, D. M., MacInnis, M. L. M., & Church, R. M. (2013). Response rates are governed more by time cues than contingency. *Timing & Time Perception, 1*, 3-20. doi: 10.1163/22134468-00002006
- Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science, 295*, 2279-2282. doi: 10.1126/science.1066893
- Gershman, S. J. (2016). Empirical priors for reinforcement learning models. *Journal of Mathematical Psychology, 71*, 1-6. doi: 10.1016/j.jmp.2016.01.006
- Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences, 108*, 15647-15654. doi: 10.1073/pnas.1014269108
- Guilhardi, P., Yi, L., & Church, R. M. (2007). A modular theory of learning and performance. *Psychonomic Bulletin & Review, 14*, 543-559.
- Habib, R., & Dixon, M. R. (2010). Neurobehavioral evidence for the "near-miss" effect in pathological gamblers. *Journal of the Experimental Analysis of Behavior, 93*, 313-328. doi: 10.1901/jeab.2010.93-313
- Hayden, B. Y., Heilbronner, S. R., Nair, A. C., & Platt, M. L. (2008). Cognitive influences on risk-seeking by rhesus macaques. *Judgment and Decision Making, 3*, 389-395.
- Hayden, B. Y., & Platt, M. L. (2007). Temporal discounting predicts risk sensitivity in rhesus macaques. *Current Biology, 17*, 49-53.

- Hayden, B. Y., & Platt, M. L. (2009). Gambling for Gatorade: risk-sensitive decision making for fluid rewards in humans. *Animal Cognition*, *12*, 201-207. doi: 10.1007/s10071-008-0186-8
- Heilbronner, S. R., & Hayden, B. Y. (2013). Contextual factors explain risk-seeking preferences in rhesus monkeys. *Frontiers in Neuroscience*, *7*, 1-7. doi: 10.3389/fnins.2013.00007
- Hollerman, J. R., & Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience*, *1*, 304-309. doi: 10.1038/1124
- Holt, D. D., Green, L., & Myerson, J. (2003). Is discounting impulsive? Evidence from temporal and probability discounting in gambling and non-gambling college students. *Behavioural Processes*, *64*, 355-367. doi: 10.1016/S0376-6357(03)00141-4
- Hong, S., Jhou, T. C., Smith, M., Saleem, K. S., & Hikosaka, O. (2011). Negative reward signals from the lateral habenula to dopamine neurons are mediated by rostromedial tegmental nucleus in primates. *The Journal of Neuroscience*, *31*, 11457-11471. doi: 10.1523/JNEUROSCI.1384-11.2011
- Horsley, R. R., Osborne, M., Norman, C., & Wells, T. (2012). High-frequency gamblers show increased resistance to extinction following partial reinforcement. *Behavioural Brain Research*, *229*, 438-442. doi: 10.1016/j.bbr.2012.01.024
- Horstmann, A., Villringer, A., & Neumann, J. (2012). Iowa Gambling Task: there is more to consider than long term outcome. Using a linear equation model to disentangle the impact of outcome and frequency of gains and losses. *Frontiers in Neuroscience*, *6*, 61. doi: 10.3389/fnins.2012.00061

- Huh, N., Jo, S., Kim, H., Sul, J. H., & Jung, M. W. (2009). Model-based reinforcement learning under concurrent schedules of reinforcement in rodents. *Learning & Memory*, *16*, 315-323. doi: 10.1101/lm.1295509
- Jensen, C., Dixon, M. J., Harrigan, K. A., Sheepy, E., Fugelsang, J. A., & Jarick, M. (2012). Misinterpreting 'winning' in multiline slot machine games. *International Gambling Studies*, *13*, 112-126. doi: 10.1080/14459795.2012.717635
- Kahneman, D., & Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, *47*, 263-291. doi: 10.2307/1914185
- Kello, J. E. (1972). The reinforcement-omission effect on fixed-interval schedules: frustration or inhibition? *Learning and Motivation*, *3*, 138-147.
- Kennerley, S. W., Walton, M. E., Behrens, T. E. J., Buckley, M. J., & Rushworth, M. F. S. (2006). Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*, *9*, 940-947. doi: 10.1038/nn1724
- Kim, H., Lee, D., & Jung, M. W. (2013). Signals for previous goal choice persist in the dorsomedial, but not dorsolateral striatum of rats. *The Journal of Neuroscience*, *33*, 52-63. doi: 10.1523/JNEUROSCI.2422-12.2013
- Kim, H., Sul, J. H., Huh, N., Lee, D., & Jung, M. W. (2009). Role of striatum in updating values of chosen actions. *The Journal of Neuroscience*, *29*, 14701-14712. doi: 10.1523/JNEUROSCI.2728-09.2009
- Kirkpatrick, K. (2002). Packet theory of conditioning and timing. *Behavioural Processes*, *57*, 89-106.

- Kirkpatrick, K., Marshall, A. T., & Smith, A. P. (2015). Mechanisms of individual differences in impulsive and risky choice in rats. *Comparative Cognition & Behavior Reviews*, *10*, 45-72. doi: 10.3819/CCBR.2015.100003
- Kirkpatrick, K., Marshall, A. T., Smith, A. P., Koci, J., & Park, Y. (2014). Individual differences in impulsive and risky choice: effects of environmental rearing conditions. *Behavioural Brain Research*, *269*, 115-127. doi: 10.1016/j.bbr.2014.04.024
- Krebs, J. R., Kacelnik, A., & Taylor, P. (1978). Test of optimal sampling by foraging great tits. *Nature*, *275*, 27-31.
- Kreek, M. J., Nielsen, D. A., Butelman, E. R., & LaForge, K. S. (2005). Genetic influences on impulsivity, risk taking, stress responsivity and vulnerability to drug abuse and addiction. *Nature Neuroscience*, *8*, 1450-1457.
- Kreussel, L., Hewig, J., Kretschmer, N., Hecht, H., Coles, M. G. H., & Miltner, W. H. R. (2013). How bad was it? Differences in the time course of sensitivity to the magnitude of loss in problem gamblers and controls. *Behavioural Brain Research*, *247*, 140-145. doi: 10.1016/j.bbr.2013.03.024
- Kubanek, J., Snyder, L. H., & Abrams, R. A. (2015). Reward and punishment act as distinct factors in guiding behavior. *Cognition*, *139*, 154-167. doi: 10.1016/j.cognition.2015.03.005
- Kumaran, D., Warren, D. E., & Tranel, D. (2015). Damage to the ventromedial prefrontal cortex impairs learning from observed outcomes. *Cerebral Cortex*, *25*, 4504-4518. doi: 10.1093/cercor/bhv080

- Kwak, S., Huh, N., Seo, J.-S., Lee, J.-E., Han, P.-L., & Jung, M. W. (2014). Role of dopamine D2 receptors in optimizing choice strategy in a dynamic and uncertain environment. *Frontiers in Behavioral Neuroscience*, 8. doi: 10.3389/fnbeh.2014.00368
- Landon, J., Davison, M., & Elliffe, D. (2002). Concurrent schedules: short- and long-term effects of reinforcers. *Journal of the Experimental Analysis of Behavior*, 77, 257-271. doi: 10.1901/jeab.2002.77-257
- Larkin, J. D., Jenni, N. L., & Floresco, S. B. (2016). Modulation of risk/reward decision making by dopaminergic transmission within the basolateral amygdala. *Psychopharmacology*, 233, 121-136. doi: 10.1007/s00213-015-4094-8
- Lau, B., & Glimcher, P. W. (2005). Dynamic response-by-response models of matching behavior in rhesus monkeys. *Journal of the Experimental Analysis of Behavior*, 84, 555-579. doi: 10.1901/jeab.2005.110-04
- Lee, D., Conroy, M. L., McGreevy, B. P., & Barraclough, D. J. (2004). Reinforcement learning and decision making in monkeys during a competitive game. *Cognitive Brain Research*, 22, 45-58. doi: 10.1016/j.cogbrainres.2004.07.007
- Lesieur, H. R. (1979). The compulsive gambler's spiral of options and involvement. *Psychiatry: Journal for the Study of Interpersonal Processes*, 42, 79-87.
- Levin, I. P., Xue, G., Weller, J. A., Reimann, M., Lauriola, M., & Bechara, A. (2012). A neuropsychological approach to understanding risk-taking for potential gains and losses. *Frontiers in Neuroscience*, 6, 1-11.
- Linnet, J., Peterson, E., Doudet, D. J., Gjedde, A., & Møller, A. (2010). Dopamine release in ventral striatum of pathological gamblers losing money. *Acta Psychiatrica Scandinavica*, 122, 326-333. doi: 10.1111/j.1600-0447.2010.01591.x

- Linnet, J., Røjskjær, S., Nygaard, J., & Maher, B. A. (2006). Episodic chasing in pathological gamblers using the Iowa gambling task. *Scandinavian Journal of Psychology*, *47*, 43-49. doi: 10.1111/j.1467-9450.2006.00491.x
- Lole, L., Gonsalvez, C. J., & Barry, R. J. (2015). Reward and punishment hyposensitivity in problem gamblers: a study of event-related potentials using a principal components analysis. *Clinical Neurophysiology*, *126*, 1295-1309. doi: 10.1016/j.clinph.2014.10.011
- Lole, L., Gonsalvez, C. J., Barry, R. J., & Blaszczynski, A. (2014). Problem gamblers are hyposensitive to wins: an analysis of skin conductance responses during actual gambling on electronic gaming machines. *Psychophysiology*, *51*, 556-564. doi: 10.1111/psyp.12198
- Lorains, F. K., Cowlishaw, S., & Thomas, S. A. (2011). Prevalance of comorbid disorders in problem and pathological gambling: systematic review and meta-analysis of population surveys. *Addiction*, *106*, 490-498. doi: 10.1111/j.1360-0443.2010.03300.x
- Luce, R. D. (1959). *Individual choice behavior: a theoretical analysis*. New York: John Wiley & Sons.
- Madan, C. R., Ludvig, E. A., & Spetch, M. L. (2014). Remembering the best and worst of times: memories for extreme outcomes bias risky decisions. *Psychonomic Bulletin & Review*, *21*, 629-636. doi: 10.3758/s13423-013-0542-9
- Madden, G. J., Ewan, E. E., & Lagorio, C. H. (2007). Toward an animal model of gambling: delay discounting and the allure of unpredictable outcomes. *Journal of Gambling Studies*, *23*, 63-83. doi: 10.1007/s10899-006-9041-5
- Maia, T. V. (2009). Reinforcement learning, conditioning, and the brain: successes and challenges. *Cognitive, Affective, & Behavioral Neuroscience*, *9*, 343-364. doi: 10.3758/CABN.9.4.343

- March, J. G. (1996). Learning to be risk averse. *Psychological Review*, *103*, 309-319.
- Marshall, A. T., & Kirkpatrick, K. (2013). The effects of the previous outcome on probabilistic choice in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, *39*, 24-38. doi: 10.1037/a0030765
- Marshall, A. T., & Kirkpatrick, K. (2015). Relative gains, losses, and reference points in probabilistic choice in rats. *PLoS ONE*, *10*, e0117697. doi: 10.1371/journal.pone.0117697
- Matsumoto, M., & Hikosaka, O. (2007). Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature*, *447*, 1111-1115.
- Mazur, J. E. (1988). Choice between small certain and large uncertain reinforcers. *Animal Learning & Behavior*, *16*, 199-205. doi: 10.3758/BF03209066
- Mazur, J. E. (2004). Risky choice: selecting between certain and uncertain outcomes. *The Behavior Analyst Today*, *5*, 190-203.
- McCoy, A. N., & Platt, M. L. (2005). Risk-sensitive neurons in macaque posterior cingulate cortex. *Nature Neuroscience*, *8*, 1220-1227. doi: 10.1038/nn1523
- McGlothlin, W. H. (1956). Stability of choices among uncertain alternatives. *The American Journal of Psychology*, *69*, 604-615. doi: 10.2307/1419083
- McNamara, J. M., Green, R. F., & Olsson, O. (2006). Bayes' theorem and its applications in animal behaviour. *Oikos*, *112*, 243-251.
- Mellon, R. C., Leak, T. M., Fairhurst, S., & Gibbon, J. (1995). Timing processes in the reinforcement-omission effect. *Animal Learning & Behavior*, *23*, 286-296.
- Mobini, S., Body, S., Ho, M.-Y., Bradshaw, C. M., Szabadi, E., Deakin, J. F. W., & Anderson, I. M. (2002). Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and

- probabilistic reinforcement. *Psychopharmacology*, *160*, 290-298. doi: 10.1007/s00213-001-0983-0
- Mobini, S., Chiang, T.-J., Ho, M.-Y., Bradshaw, C. M., & Szabadi, E. (2000). Effects of central 5-hydroxytryptamine depletion on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology*, *152*, 390-397. doi: 10.1007/s002130000542
- Myerson, J., & Green, L. (1995). Discounting of delayed rewards: models of individual choice. *Journal of the Experimental Analysis of Behavior*, *64*, 263-276. doi: 10.1901/jeab.1995.64-263
- Myerson, J., Green, L., Hanson, J. S., Holt, D. D., & Estle, S. J. (2003). Discounting delayed and probabilistic rewards: processes and traits. *Journal of Economic Psychology*, *24*, 619-635. doi: 10.1016/S0167-4870(03)00005-9
- Niv, Y., Edlund, J. A., Dayan, P., & O'Doherty, J. P. (2012). Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *The Journal of Neuroscience*, *32*, 551-562. doi: 10.1523/JNEUROSCI.5498-10.2012
- Nygren, T. E., Isen, A. M., Taylor, P. J., & Dulin, J. (1996). The influence of positive affect on the decision rule in risk situations: focus on outcome (and especially avoidance of loss) rather than probability. *Organizational Behavior and Human Decision Processes*, *66*, 59-72. doi: 10.1006/obhd.1996.0038
- O'Doherty, J. P. (2014). The problem with value. *Neuroscience & Biobehavioral Reviews*, *43*, 259-268. doi: 10.1016/j.neubiorev.2014.03.027
- O'Neill, M., & Schultz, W. (2010). Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. *Neuron*, *68*, 789-800. doi: 10.1016/j.neuron.2010.09.031

- Payne, J. W. (2005). It is whether you win or lose: the importance of the overall probabilities of winning or losing in risky choice. *The Journal of Risk and Uncertainty*, *30*, 5-19. doi: 10.1007/s11166-005-5831-x
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*, 1042-1045. doi: 10.1038/nature05051
- Peters, H., Hunt, M., & Harper, D. (2010). An animal model of slot machine gambling: the effect of structural characteristics on response latency and persistence. *Journal of Gambling Studies*, *26*, 521-531.
- Peters, J., & Büchel, C. (2010). Neural representations of subjective reward value. *Behavioural Brain Research*, *213*, 135-141. doi: 10.1016/j.bbr.2010.04.031
- Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D. (2016). nlme: Linear and Nonlinear Mixed Effects Models. Retrieved from <http://CRAN.R-project.org/package=nlme>
- Platt, M. L., & Huettel, S. A. (2008). Risky business: the neuroeconomics of decision making under uncertainty. *Nature Neuroscience*, *11*, 398-403. doi: 10.1038/nn2062
- Potenza, M. N. (2009). The importance of animal models of decision making, gambling, and related behaviors: implications for translational research in addiction. *Neuropsychopharmacology*, *34*, 2623-2624. doi: 10.1038/npp.2009.152
- Qi, S., Ding, C., Song, Y., & Yang, D. (2011). Neural correlates of near-misses effect in gambling. *Neuroscience Letters*, *493*, 80-85. doi: 10.1016/j.neulet.2011.01.059
- Rachlin, H. (1990). Why do people gamble and keep gambling despite heavy losses? *Psychological Science*, *1*, 294-297. doi: 10.1111/j.1467-9280.1990.tb00220.x

- Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective probability and delay. *Journal of the Experimental Analysis of Behavior*, 55, 233-244. doi: 10.1901/jeab.1991.55-233
- Rachlin, H., Safin, V., Arfer, K. B., & Yen, M. (2015). The attraction of gambling. *Journal of the Experimental Analysis of Behavior*, 103, 260-266. doi: 10.1002/jeab.113
- Rasmussen, E. B., Lawyer, S. R., & Reilly, W. (2010). Percent body fat is related to delay and probability discounting for food in humans. *Behavioural Processes*, 83, 23-30. doi: 10.1016/j.beproc.2009.09.001
- Reboreda, J. C., & Kacelnik, A. (1991). Risk sensitivity in starlings: variability in food amount and food delay. *Behavioral Ecology*, 2, 301-308. doi: 10.1093/beheco/2.4.301
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical Conditioning II: Current Research and Theory* (pp. 64-99). New York: Appleton-Century-Crofts.
- Reynolds, B., Richards, J. B., Horn, K., & Karraker, K. (2004). Delay discounting and probability discounting as related to cigarette smoking status in adults. *Behavioural Processes*, 65, 35-42. doi: 10.1016/S0376-6357(03)00109-8
- Robinson, O. J., Standing, H. R., DeVito, E. E., Cools, R., & Sahakian, B. J. (2010). Dopamine precursor depletion improves punishment prediction during reversal learning in healthy females but not males. *Psychopharmacology*, 211, 187-195. doi: 10.1007/s00213-010-1880-1
- Rogers, R. D., Wong, A., McKinnon, C., & Winstanley, C. A. (2013). Systemic administration of 8-OH-DPAT and eticlopride, but not SCH23390, alters loss-chasing behavior in the rat. *Neuropsychopharmacology*, 38, 1094-1104. doi: 10.1038/npp.2013.8

- Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience*, *11*, 389-397. doi: 10.1038/nn2066
- Rutledge, R. B., Lazzaro, S. C., Lau, B., Myers, C. E., Gluck, M. A., & Glimcher, P. W. (2009). Dopaminergic drugs modulate learning rates and perseveration in Parkinson's patients in a dynamic foraging task. *The Journal of Neuroscience*, *29*, 15104-15114. doi: 10.1523/JNEUROSCI.3524-09.2009
- Schultz, W. (2007). Behavioral dopamine signals. *Trends in Neurosciences*, *30*, 203-210. doi: 10.1016/j.tins.2007.03.007
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593-1599. doi: 10.1126/science.275.5306.1593
- Searcy, G. D., & Pietras, C. J. (2011). Optimal risky choice in humans: effects of amount of variability. *Behavioural Processes*, *87*, 88-99.
- Seo, H., & Lee, D. (2009). Behavioral and neural changes after gains and losses of conditioned reinforcers. *The Journal of Neuroscience*, *29*, 3627-3641. doi: 10.1523/JNEUROSCI.4726-08.2009
- Seymour, B., Maruyama, M., & De Martino, B. (2015). When is a loss a loss? Excitatory and inhibitory processes in loss-related decision-making. *Current Opinion in Behavioral Sciences*, *5*, 122-127. doi: 10.1016/j.cobeha.2015.09.003
- Shaffer, H. J., Hall, M. N., & Vander Bilt, J. (1999). Estimating the prevalence of disordered gambling behavior in the United States and Canada: a research synthesis. *American Journal of Public Health*, *89*, 1369-1376. doi: 10.2105/AJPH.89.9.1369

- Shaffer, H. J., & Korn, D. A. (2002). Gambling and related mental disorders: a public health analysis. *Annual Review of Public Health, 23*, 171-212. doi: 10.1146/annurev.publhealth.23.100901.140532
- Smethells, J. R., Fox, A. T., Andrews, J. J., & Reilly, M. P. (2012). Immediate postsession feeding reduces operant responding in rats. *Journal of the Experimental Analysis of Behavior, 97*, 203-214. doi: 10.1901/jeab.2012.97-203
- St. Onge, J. R., Stopper, C. M., Zahm, D. S., & Floresco, S. B. (2012). Separate prefrontal-subcortical circuits mediate different components of risk-based decision making. *The Journal of Neuroscience, 32*, 2886-2899. doi: 10.1523/JNEUROSCI.5625-11.2012
- Stagner, J. P., & Zentall, T. R. (2010). Suboptimal choice behavior by pigeons. *Psychonomic Bulletin & Review, 17*, 412-416.
- Stewart, N., Chater, N., Stott, H. P., & Reimers, S. (2003). Prospect relativity: how choice options influence decision under risk. *Journal of Experimental Psychology: General, 132*, 23-46. doi: 10.1037/0096-3445.132.1.23
- Stopper, C. M., & Floresco, S. B. (2011). Contributions of the nucleus accumbens and its subregions to different aspects of risk-based decision making. *Cognitive, Affective, and Behavioral Neuroscience, 11*, 97-112. doi: 10.3758/s13415-010-0015-9
- Stopper, C. M., & Floresco, S. B. (2014). What's better for me? Fundamental role for lateral habenula in promoting subjective decision biases. *Nature Neuroscience, 7*, 33-37. doi: 10.1038/nn.3587
- Stopper, C. M., Khayambashi, S., & Floresco, S. B. (2013). Receptor-specific modulation of risk-based decision making by nucleus accumbens dopamine. *Neuropsychopharmacology, 38*, 715-728. doi: 10.1038/npp.2012.240

- Stopper, C. M., Tse, M. T. L., Montes, D. R., Wiedman, C. R., & Floresco, S. B. (2014).
Overriding phasic dopamine signals redirects action selection during risk/reward decision
making. *Neuron*, *84*, 177-189. doi: 10.1016/j.neuron.2014.08.033
- Sul, J. H., Kim, H., Huh, N., Lee, D., & Jung, M. W. (2010). Distinct roles of rodent
orbitofrontal and medial prefrontal cortex in decision making. *Neuron*, *66*, 449-460. doi:
10.1016/j.neuron.2010.03.033
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA:
MIT Press.
- Tatham, T. A., & Zurn, K. R. (1989). The MED-PC experimental apparatus programming
system. *Behavior Research Methods Instruments & Computers*, *21*, 294-302. doi:
10.3758/BF03205598
- Thaler, R. H., & Johnson, E. J. (1990). Gambling with the house money and trying to break
even: the effects of prior outcomes on risky choice. *Management Science*, *36*, 643-660.
doi: 10.1287/mnsc.36.6.643
- Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in
decision-making under risk. *Science*, *315*, 515-518. doi: 10.1126/science.1134239
- Tremblay, M., Cocker, P. J., Hosking, J. G., Zeeb, F. D., Rogers, R. D., & Winstanley, C. A.
(2014). Dissociable effects of basolateral amygdala lesions on decision making biases in
rats when loss or gain is emphasized. *Cognitive, Affective & Behavioral Neuroscience*,
14, 1184-1195. doi: 10.3758/s13415-014-0271-1
- Van den Berg, I., Franken, I. H. A., & Muris, P. (2011). Individual differences in sensitivity to
reward: association with electrophysiological responses to monetary gains and losses.
Journal of Psychophysiology, *25*, 81-86. doi: 10.1027/0269-8803/a000032

- van Holst, R. J., Chase, H. W., & Clark, L. (2014). Striatal connectivity changes following gambling wins and near-misses: associations with gambling severity. *NeuroImage: Clinical*, 5, 232-239. doi: 10.1016/j.nicl.2014.06.008
- van Holst, R. J., van den Brink, W., Veltman, D. J., & Goudriaan, A. E. (2010). Why gamblers fail to win: a review of cognitive and neuroimaging findings in pathological gambling. *Neuroscience and Biobehavioral Reviews*, 34, 87-107. doi: 10.1016/j.neubiorev.2009.07.007
- Vasconcelos, M., Monteiro, T., Aw, J., & Kacelnik, A. (2010). Choice in multi-alternative environments: a trial-by-trial implementation of the Sequential Choice Model. *Behavioural Processes*, 84, 435-439. doi: 10.1016/j.beproc.2009.11.010
- Walasek, L., & Stewart, N. (2015). How to make loss aversion disappear and reverse: tests of the decision by sampling origin of loss aversion. *Journal of Experimental Psychology: General*, 144, 7-11. doi: 10.1037/xge0000039
- Walton, M. E., Behrens, T. E. J., Buckley, M. J., Rudebeck, P. H., & Rushworth, M. F. S. (2010). Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron*, 65, 927-939. doi: 10.1016/j.neuron.2010.02.027
- Wang, X. T., & Johnson, J. G. (2012). A tri-reference point theory of decision making under risk. *Journal of Experimental Psychology: General*, 141, 743-756. doi: 10.1037/a0027415
- Weatherly, J. N., & Derenne, A. (2007). Rats playing a slot machine: a preliminary attempt at an animal gambling model. *Analysis of Gambling Behavior*, 1, 79-89.

- Weber, E. U., Shafir, S., & Blais, A.-R. (2004). Predicting risk sensitivity in humans and lower animals: risk as variance or coefficient of variation. *Psychological Review*, *111*, 430-445. doi: 10.1037/0033-295X.111.2.430
- Wilke, A., & Barrett, H. C. (2009). The hot hand phenomenon as a cognitive adaptation to clumped resources. *Evolution and Human Behavior*, *30*, 161-169. doi: 10.1016/j.evolhumbehav.2008.11.004
- Winstanley, C. A. (2011). Gambling rats: insight into impulsive and addictive behavior. *Neuropsychopharmacology*, *36*, 359. doi: 10.1038/npp.2010.136
- Worhunsky, P. D., Malison, R. T., Rogers, R. D., & Potenza, M. N. (2014). Altered neural correlates of reward and loss processing during simulated slot-matching fMRI in pathological gambling and cocaine dependence. *Drug and Alcohol Dependence*, *145*, 77-86. doi: 10.1016/j.drugalcdep.2014.09.013
- Worthy, D. A., Pang, B., & Byrne, K. A. (2013). Decomposing the roles of perseveration and expected value representation in models of the Iowa Gambling Task. *Frontiers in Psychology*, *4*. doi: 10:3389/fpsyg.2013.00640
- Wright, D. B., & London, K. (2009). Multilevel modelling: beyond the basic applications. *British Journal of Mathematical and Statistical Psychology*, *62*, 439-456. doi: 10.1348/000711008X327632
- Xue, G., Lu, Z., Levin, I. P., & Bechara, A. (2011). An fMRI study of risk-taking following wins and losses: implications for the gambler's fallacy? *Human Brain Mapping*, *32*, 271-281. doi: 10.1002/hbm.21015
- Yacubian, J., Gläscher, J., Schroeder, K., Sommer, T., Braus, D. F., & Büchel, C. (2006). Dissociable systems for gain- and loss-related value predictions and errors of prediction

- in the human brain. *The Journal of Neuroscience*, 26, 9530-9537. doi: 10.1523/JNEUROSCI.2915-06.2006
- Yechiam, E., & Telpaz, A. (2013). Losses induce consistency in risk taking even without loss aversion. *Journal of Behavioral Decision Making*, 26, 31-40. doi: 10.1002/bdm.758
- Young, M. E. (2016). The problem with categorical thinking by psychologists. *Behavioural Processes*, 123, 43-53. doi: 10.1016/j.beproc.2015.09.009
- Young, M. E., Webb, T. L., Rung, J. M., & Jacobs, E. A. (2013). Sensitivity to changing contingencies in an impulsivity task. *Journal of the Experimental Analysis of Behavior*, 99, 335-345. doi: 10.1002/jeab.24
- Zentall, T. R., & Stagner, J. (2010). Maladaptive choice behaviour by pigeons: an animal analogue and possible mechanism for gambling (sub-optimal human decision-making behaviour). *Proceedings of the Royal Society B*.
- Zhong, S., Israel, S., Xue, H., Sham, P. C., Ebstein, R. P., & Chew, S. H. (2009). A neurochemical approach to valuation sensitivity over gains and losses. *Proceedings of the Royal Society B*, 276, 4181-4188. doi: 10.1098/rspb.2009.1312