Title: Bull, bear, or rat markets: Rat ‘stock market’ task reveals human-like behavioral biases

Word count: ~10,000 words

Figure count: 6 figures, 4 tables

Abstract

Investors often exhibit behavioral biases (e.g. loss aversion) that are putatively underpinned by mechanisms supporting reinforcement learning in the brain, which are largely evolutionarily conserved across mammalian species. While prior research has demonstrated that rats, like humans, exhibit behavioral economic biases in certain contexts, asset market contingencies have gone largely unexplored. Thus, we developed an experimental ‘stock market’ task in which cohorts of four rats drove asset prices up and down by selecting and subsequently buying, selling, or holding ‘stocks’ to earn sweet liquid reward. Profits and losses were operationalized as reward volumes larger than and smaller than a reference volume of reward, respectively. Following a loss, rats moved more slowly to collect reward and spent less time licking at the reward spigot, indicative of lower motivation to approach and ‘savor’ a loss reward. Rats also tended to respond suboptimally following a loss, which corresponded to an increase in risk-seeking behavior characterized by a bias against the optimal ‘hold’ option in that context. Rats’ choice of the sell option demonstrated a robust tendency towards realizing gains more quickly than losses, which is characteristic of the ‘disposition effect’ in human stock markets. Our results indicate that rats exhibit behavioral biases similar to human investors, emphasizing the suitability of the rat stock market model to future work into the behavioral neuroscience of suboptimal financial decision making.

1. Introduction

Financial decisions have a big impact on our lives – especially when those decisions are suboptimal. The neuroeconomic investigation of behavioral biases that lead to suboptimal financial decisions has gained increasing attention in recent decades (Frydman & Camerer, 2016). Financial decision-making biases, such as loss aversion or the disposition effect, are ubiquitous across cultures (Grinblatt & Keloharju, 2001), professional money managers (Shapira & Venezia, 2001), and even non-human primates (Chen et al., 2006; Lakshminaryanan et al., 2008). There is a rich history of studies demonstrating that laboratory rats, too, share our economic biases (Bhatti et al., 2014; Kagel et al., 1986; Marshall & Kirkpatrick, 2015). However, to the authors' knowledge, no previous study to-date has characterized rat behavior in a task with reinforcement contingencies that approximate those found in a stock market.

Given the inherent unpredictability of future returns of stocks, the stock market represents a particularly interesting context in which to study how reinforcement mechanisms drive presumed 'myopic' valuation of uncertain gains and losses (Choi et al., 2009; Frydman et al., 2014; Fuster et al., 2010; Gutiérrez-Roig et al., 2016; Kuhnen & Knutson, 2005; Strahilevitz et al., 2011; Thaler et al., 1997). To the reward system, the stock market contains: 1) reward signals (i.e. current prices) that do not necessarily reflect an investor's experienced value (i.e. profit/loss), and 2) price stochasticity that undermines the system's ability to rely on reward prediction errors to improve predictions of future payoffs (Fama, 1965; Schultz et al., 1997; Sutton & Barto, 1998). It has been hypothesized that these reinforcement contingencies inappropriately drive reinforcement learning mechanisms, which in turn interacts with an individual's ability to choose in a

way that maximizes expected returns (Charness & Levin, 2005; Daw et al., 2005; Gutiérrez-Roig et al., 2016).

Kahneman and Tversky's (1979) Prospect Theory provides a detailed account of behavioral biases that arise in such contexts. Three prominent biases are: 1) loss aversion, where an individual demonstrates greater sensitivity to potential losses than to gains, 2) risk-aversion, or a preference for less uncertainty, and 3) anchoring, where an individual defines gains and losses according to some reference point other than zero. A fourth behavioral bias is referred to as the 'disposition effect'. Although not originally put forth by Kahneman and Tversky (1979), the disposition effect has been postulated to stem from Prospect Theory preferences (c.f. Kaustia, 2010; Shefrin & Statman, 1985) and is characterized by investors' reluctance to sell losing stocks while selling winning stocks too quickly.

Standing in contrast to Prospect Theory, the Efficient Markets Hypothesis asserts that the behavioral biases of individual investors are irrelevant to prices because they are quickly corrected by arbitrage forces (Samuelson, 1965; Shleifer & Summers, 1990). However, a number of empirical studies have identified suboptimal patterns of investor behavior that putatively reflect an inability to inhibit simple reinforcement-based strategies (Choi et al., 2009; Erev & Roth, 1998; Kaustia & Knüpfer, 2008; Payzan-LeNestour & Bossaerts, 2015; Strahilevitz et al., 2011). This suggests that an investor's motivation, subjective reward value, and affective state are each factors that play a larger role than traditionally assumed in normative financial theory (Fama, 1970; Frydman & Camerer, 2016; Hirshleifer, 2001; Shiller, 2003; Strahilevitz et al., 2011).

For example, Kaustia and Knüpfer (2008) evaluated five years of individual investor behavior in a Finnish investor public offering (IPO) market, and found that investors displayed a strong tendency to repeat an investment that had previously led to positive returns and avoid an investment that had previously resulted in negative returns. Thorndike (1911) first described such win-stay/lose-shift behavior as the "Law of Effect"; a simple strategy that explains how animals' responses typically increase after rewarding outcomes and decrease after undesirable outcomes. When many investors respond to a market event (e.g. a highly anticipated IPO) with similarly-biased behavior, the cumulative effect can lead to autocorrelation of prices, potentially escalating into bubble markets (Daniel et al., 1998; Poterba & Summers, 1988; Shiller, 2000). In sum, the sentiment and reward history of individual investors can align on certain market events, driving subtle yet impactful deviations from fundamental value.

Here, we investigated the possibility that rats may also exhibit suboptimal investment behaviors in reward contexts that simulate the outcomes of trading decisions in the stock market. Research in non-human primates provides some evidence that this may be the case. Chen and colleagues (2006) showed that non-human primates exhibit suboptimal behaviors that are similar to humans in an experimental marketplace. In the experiment, capuchins were given 'tokens' that could be exchanged for fruit, thus creating a primitive token economy. The experimenters systematically varied either the number of tokens to represent 'wealth', or the number of the fruit pieces a monkey earned in exchange for one token, i.e. 'price'. In a pair of separate experiments, the experimenters either added or subtracted a piece of fruit to the initial exchange offer in

order to elicit putative effects of 'reference-point' and 'loss aversion'. While capuchin monkeys responded rationally to changes in wealth and price, they exhibited both reference-dependence and loss aversion. In the 'reference point' experiment, monkeys preferred gambles with a 50% chance that an experimenter would add a second piece of fruit to a single fruit offer over a 50% chance that a piece of fruit would be taken away from a two-piece offer. Monkeys were then given the option between two trades; one of which was initially two pieces of fruit from which one piece was removed, while the second trade always started as and delivered just the single piece of fruit. Monkeys demonstrated a strong preference for the option that did not involve a piece of fruit being removed, which is indicative of loss aversion. In an extension of the task, capuchins also exhibited framing effects – becoming risk-seeking when gambles were presented as a loss and risk-averse when the gamble was presented as a gain (Lakshminaryanan et al., 2008).

Like human and non-human primates, research has also shown that, while rats respond rationally to changes in wealth and price in certain simulated economic experiments (Kagel et al., 1981; Kagel et al., 1975; van Wingerden et al., 2015), certain contexts also reliably give rise to behavioral biases such as loss aversion, risk aversion, and anchoring (Bhatti et al., 2014; Constantinople, Piet, & Brody, 2019; Kagel et al., 1986; Kirkpatrick et al., 2015; Koot et al., 2009; Marshall & Kirkpatrick, 2015, 2017; Rivalan et al., 2009). This has allowed for the investigation of the neuronal mechanisms underlying such suboptimal behaviors (Ferland et al., 2018; St Onge & Floresco, 2009; Tremblay et al., 2014; Zalocusky et al., 2016; Zeeb et al., 2009). Yet, to the authors' best knowledge, no previous studies have used a simulated 'stock market' environment to explore suboptimal investment behaviors in the rat. Establishing a rat stock market model

would provide a valuable opportunity in the future to investigate suboptimal investment behavior at the neural level using techniques not available in human research while also avoiding many of the 'human' confounds such as education and numeracy (Kalenscher & van Wingerden, 2011).

In order to explore the use of rats as a model of investor biases, we designed a task that obeyed the key reinforcement contingencies of human stock markets. Our primary goals in paradigm ideation were to simulate: 1) repeated trinary choices between buy, sell and hold options of 3 virtual 'stocks', 2) subject-driven 'demand' for stocks that resulted in corresponding price fluctuations and potential boom-bust cycles, 3) current reward as a net profit/loss based on 'price' changes from previous trials, and 4) opportunity costs associated with both action and inaction in the 'market.'

By poking into one of five different nosepoke-holes, cohorts of four rats could choose to buy, hold, or sell one of three simulated 'stocks' in order to receive liquid reward. The resulting profit or loss on a given trial resulted in delivery of sodium saccharin reward at a volume greater than or less than a reference volume, respectively. Stock prices shifted dynamically according to the buy and sell decisions of all four rats within a trading cohort. Although prices in real asset markets are commonly believed to follow a random walk (Daniel et al., 1998; Fama, 1965; Malkiel, 1973; Samuelson, 1965) and could have been simulated pseudo-randomly *a priori*, we chose instead to determine stock prices at any given trial based on the current cumulative 'demand' from the four

rats in a cohort. This mechanism allowed for naturally-occurring "boom and bust" cycles and serial autocorrelation, which have been associated with investor biases in real asset markets (Barberis & Thaler, 2003; Bondt & Thaler, 1985; Hirshleifer, 2001; Shiller, 2003; Xiong, 2013).

2. Materials and Methods

2.1 Subjects

Subjects were 24 adult male Lister hooded rats, 8 of which were bred in house and 16 of which were bred by a commercial supplier (Harlan U.K.), with initial weights between 300g and 475g. Animals were housed in groups of two or three on a 12-hour light: 12-hour dark cycle (7PM lights off). All testing was carried out on weekdays during the light part of the cycle. Rats were habituated to human handling for two weeks and then placed on restricted water access. Rats received water *ad libitum* from Friday afternoon to Sunday afternoon and for one hour each weekday after testing. Rats' weights were monitored daily before testing so that no animal was allowed to drop below 85% of its free-drinking body weight. All animals maintained growth during this experiment..

2.2 Apparatus

Testing was carried out in four 34cm × 29cm × 25cm Perspex inner chambers (Med Associates Inc., St Albans, VT). The right wall of the inner testing chamber contained five square nosepoke holes, each accommodating a recessed green LED light (luminosity ~ 4.5 mcd per LED) as well as an infrared sensor to record nosepokes. The right wall

contained a recessed custom-built liquid reward magazine with a white LED (approximately 2072 mcd luminosity) delivering 0.3% weight/volume sodium saccharin solution at a constant rate of 0.05ml/sec. Capacitive contact lickometers were used to determine when the rats licked the reward spigot. piezoelectric buzzers signaled reward availability for 'gain' (2900Hz, 85dB) and 'loss' (4400Hz, 70db) payoffs during testing. The reader is referred to Wilson et al. (2006) for details of the operant chambers and reward delivery system.

Behavioral testing was interfaced by the MED-PC® IV data experimental control system (Med Associates Inc., St Albans, VT) with an HP® computer running Windows7™. Behavioral events were time-stamped (2 msec resolution) and recorded for offline data analysis and session reconstruction using *tsch* to batch-process data files using a program in the AWK programming language (Apple Computer Inc., OS X operating system). Subsequent data analysis was carried out using: Microsoft® Excel for Mac 2011, SPSS® version 24 for Mac, and R version 3.2.2 for Mac.

2.3 Testing

Six cohorts of four adult male rats (N=24) operated a virtual stock market by nosepoking in holes to select, and subsequently buy, sell, or hold, virtual assets in order to receive sweet liquid reward. The reader is referred to the SI for a comprehensive description of pre-training and shaping procedures prior to behavioral testing. Rats were placed into individual operant chambers that were networked to dynamically update trading information across all four chambers in real time. Thus, the buy and sell decisions of one rat in the cohort affect the share prices faced by all rats on subsequent trials. Rats made

choices in the task by nosepoking into an illuminated nosepoke hole of a 5-hole array. Each session comprised two blocks. For each trial, rats performed a sequence of two nosepokes in order to earn a fixed (block 1) or variable (block 2) volume of reward.

2.3.1 Block 1: Reference Point Establishment

In the initial 15 trials of each testing session (block 1), rats performed two forced-choice nosepokes into lit (but unblinking) holes, which always resulted in 0.15ml of reward, consistent with the pre-training reward volume and irrespective of the nosepoke choices made by the rat. This served as a reference reward volume (hereon, reference point) from which 'wins' and 'losses' would later deviate. Once all four rats from a cohort completed block 1 (mean completion time for block 1 = 279 secs), the group began trading virtual stocks in block 2 (Fig. 1). In total, rats underwent 8-10 days of testing, of which the final 7 days were used for analysis.

2.3.2 Block 2: Task Structure

Block 2 lasted 45 minutes, in which a mean of 100 trials were completed. Free and forced-choice trials for stock selection were pseudorandomly interleaved at a ratio of 3:1, respectively. The trial sequence within block 2 was similar to block 1, with the exception of the first discriminative stimulus (i.e. nosepoke hole light) and the payoff structure. Where the available holes were lit but unblinking in block 1, in block 2 the lights blinked to signify the 'price' of each stock (see Fig 1). Each of the three center nosepoke holes was randomly assigned to one of three stocks (stocks 1, 2, and 3) at the beginning of each testing session. Stock location remained fixed within a session but

varied between sessions, and stock price was signified by the temporal frequency of the flashing LED (blink rate) in the associated nosepoke hole on the first nosepoke response only. The volume of reward rats received upon successful completion of a nosepoke sequence varied (min = 0 ml, max = 0.450 ml, mean = 0.165 ml).

A rat began a free-choice trial by nosepoking into one of the three blinking center holes (holes 2, 3, or 4), where blink rate signified the price of each option. This was analogous to 'selecting a stock.' All lights were immediately extinguished for 2 seconds following an initial nosepoke. After the 2-sec pause, the chosen nosepoke hole (e.g. hole 2) and the immediately adjacent nosepoke holes (e.g. holes 1 and 3) were illuminated (but not blinking). The rat was then free to nosepoke a second time into one of the three lit holes, which represented an 'option selection' ('buy', 'hold', or 'sell'). Following the second nosepoke response, the hole lights were immediately extinguished and one of two tones (either a gain tone or a loss tone) was paired with the conditioned light stimulus at the reward spigot to indicate the reward outcome of their choice sequence. We enforced a minimum trial length of 7 sec. As long as the minimum 7 sec had elapsed, the trial ended once the rat ceased licking at the reward spigot (interlick interval >300 msec). The trial was followed by an intertrial interval with a mean of 2 sec. In order to allow for the natural development of behavioral strategies within the task, rats were given no training on the association between changing blink rate and profit/loss of a buy/sell option.

2.3.3 Stock Price & Blink Rate

A stock's price depended on its total number of shares held across all four rats, with greater cumulative shares leading to higher prices, and lower cumulative shares

resulting in lower prices. The variability in stock prices based on 'demand' was the primary motivation for incorporating 4 rats into each trading cohort. This allowed us to achieve some likeness to the stochastic share price fluctuations observed in human markets.

The price of a given share at any moment was equal to:

Share Price = (Initial Price × Cumulative # of Shares)/400

The initial price of stocks 1, 2, and 3 was fixed across sessions and was arbitrarily set at 80, 140, and 220 arbitrary units respectively. A stock's price increased with the demand for the share as indicated by the number of shares 'bought' by the rats, and conversely decreased with the number of 'sales' of that stock. By varying the initial starting price across stocks, we were able to manipulate the 'step' in reward volume a rat would receive for choosing the buy/sell option while keeping blink rate constant. Rats held a maximum of 830 shares and a minimum 40 shares.

The blink rate that signaled a stock's price operated on a 50/50 on/off cycle so that the time between each flash was equal to each flash length. The on/off time period is the reciprocal of blink rate (Hz):

On/Off Period = [2/(Cumulative # of Shares/100)] × 1 second

Each rat (4 rats per testing cohort) was initially endowed with 100 shares of each of the three stocks, which meant that at the start of a session, the blink rate was the same for each of the three stocks: [2/((100 x 4 rats)/100)] x 1 sec = 0.5 sec on/off = 2 Hz.

2.3.4 The 'Hold' Option

By choosing to nosepoke in the same hole as the initial nosepoke (e.g. hole 2), the rat selected an option equivalent to a 'hold' in the task. A conditioned light stimulus paired with a gain tone immediately indicated the availability of reward at the reward spigot. Choice of the 'hold' option always resulted in the reference point volume of reward (0.15ml) plus a smaller 'dividend' reward (mean = 0.015 ml). Choice of the 'hold' option neither incremented nor decremented the number of shares held in a rat's 'portfolio.' Thus, this action did not affect the subsequent blink rates representing the price of the chosen stock.

The dividend amount was based on the individual subject's current share holdings of that stock, and had a 2/3 probability of being low (e.g. 2% of holdings of the given stock) and a 1/3 probability of being high (e.g. 6% of holdings of the given stock).

2.3.5 The 'Sell' Option

By choosing to nosepoke in the hole to the right (e.g. hole 3) of the initial nosepoke hole, the rat would be selecting an option equivalent to 'sell' in the task. Note that 'buy' and 'sell' were counterbalanced to the left and right across rats. After a nosepoke, a conditioned light stimulus immediately indicated the availability of reward at the reward spigot, and either a gain or a loss tone indicated whether the reward volume was greater than or less than the reference point of 0.15ml, respectively. The 'sell' action decremented the number of shares held in the rat's virtual 'portfolio' by 10, which was

also reflected in the total number of shares held cumulatively across all rats. Akin to demand, the price of the chosen stock subsequently fell for all four rats in each of the testing chambers at that time.

The volume of reward a rat received on a given 'sell' trial was proportional to the amount gained or lost in the 'transaction', added to or subtracted from the reference point (0.15 ml), respectively.

$$\text{Profit/Loss} = \text{Price}_t - \text{Price}_{t-1}$$

$$\text{Reward} = 0.15\text{ ml} + (\text{profit or loss})*0.01\text{ ml}$$

If the selected stock's share price had increased (i.e. a greater cumulative number of shares now held by all four rats) from the previously selected trial, $t-1$, relative to the current trial, t, then the 'sell' option resulted in a gain of that magnitude. Conversely, if the share price had decreased since the previous trial, then the rat experienced a loss of that magnitude. This is intended to be akin to an investor that buys a stock at e.g. $100 and subsequently sells it at $110 for a $10 profit. Had the share price fallen to $90 in this example, the sale would have resulted in a $10 loss.

2.3.6 The 'Buy' Option

By choosing to nosepoke in the hole to the left (e.g. hole 1) of the initial nosepoke hole, the rat would be selecting an option equivalent to a 'buy' in the task. The conditioned light stimulus immediately indicated the availability of reward at the reward spigot, and either the gain or the loss tone indicated whether the reward volume was greater than or less than the reference point of 0.15ml, respectively. This action incremented the number of shares held in the rat's virtual 'portfolio' by 10, which was also reflected in the total number of shares held cumulatively across all rats. Akin to

demand, the price of the chosen stock subsequently rose for all four rats in each of the testing chambers at that time.

The volume of reward a rat received on a given 'buy' trial was proportional to the amount gained or lost in the 'transaction', added to or subtracted from the reference point (0.15 m), respectively.

$$\text{Profit/Loss} = \text{Price}_{t\text{-}1} - \text{Price}_{t}$$

$$\text{Reward} = 0.15\ \text{ml} + (\text{profit or loss})*0.01\ \text{ml}$$

If the selected stock's share price had decreased (i.e. fewer cumulative number of shares now held by all four rats) from the previously selected trial, *t-1*, relative to the current trial, *t*, then the 'buy' option resulted in a gain of that magnitude. Conversely, if the share price had increased since the previous trial, then then the rat experienced a loss of that magnitude. This is intended to represent the missed opportunity cost experienced by an investor that, for example, fails to invest on day 1 when a stock is $50 a share, only to find that the share price had increased to $60 per share on day 2.

2.4 Data Analysis

Offline data analysis was performed on behavior from the final 7 testing sessions. Summary measures were aggregated per subject across the 7 sessions and analyzed in SPSS. Regressions and hazard models were conducted on non-aggregated data in *R*. *R*'s **nlme** package was used in the generalized linear mixed modeling of MT and PPL with reward (Pinheiro et al., 2017). Relationships of task predictors to choices were evaluated by multinomial (option choice) or binomial (choice optimality) logistic regression using the **mlogit** package in R (Croissant, 2013). In order to facilitate comparison of rat and

human disposition effect behavior, we also conducted a Cox proportional hazard model, a methodology often employed in the behavioral finance literature (Barber & Odean, 2011; Odean, 1998), to evaluate rats' hazard ratios at different rates of return. Cox proportional hazard models were estimated using the **survival** package in R (Therneau, 2014). Effect sizes for paired-sample *t*-tests are reported as Cohen's *d* calculated from pooled sample variance.

2.5 Behavioral Measures

Behavioral measures were: stock choice; option (buy/hold/sell) choice; (sub)optimality of option choice; movement time to collect reward (MT); and post-pump licking (PPL), which is defined as the amount of time spent licking at the reward spigot after mechanical cessation of reward delivery. Thus, PPL can be interpreted as a measure of post-consumption behavior, and MT a measure of pre-consumption behavior. MT durations > 15 secs and PPL > 10 secs were omitted from analysis, as they were indicative of the rat taking a break or perseverating at the reward spigot, respectively.

2.6 Proportion of Realized Gains & Losses

An investor who holds a winning stock in her portfolio without selling it holds a 'paper gain.' Once sold, the paper gain becomes a 'realized gain.' Adapted from Odean (Odean), we calculated the proportion of gains realized (PGR) and the proportion of losses realized (PLR) in order to establish whether rats exhibited the disposition effect (when PGR > PLR) in our stock market task. The disposition effect represents a pattern in which stocks with a winning history are sold and those with a losing history are held,

resulting in suboptimal performance given the temporal autocorrelation in stock price. Odean (Odean) computed PGR and PLR as:

$$PGR = \frac{\#\ of\ realized\ gains}{\#\ of\ realized\ gains + \#\ of\ paper\ gains}$$

$$PLR = \frac{\#\ of\ realized\ losses}{\#\ of\ realised\ losses + \#\ of\ paper\ losses}$$

Above, the denominator represents the number of opportunities to realize a gain (loss). In the current experiment this is slightly less straightforward, as interleaving of forced-choice and paired-choice trials precluded the opportunity to realize a gain (loss) on every stock on every trial. Therefore, the PGR (PLR) denominator was calculated on a stock-by-stock basis as any trial on which the rat had the opportunity to select *and* sell a stock that had gone up (down) in price since the previous purchase.

2.7 Cox Proportional-Hazard Modeling

The Cox proportional-hazard model is a semi-parametric analysis that makes no assumption about the shape (e.g. linear) of the baseline hazard rate. This model has been employed in a number of behavioral finance studies (Barber & Odean, 2011; Feng & Seasholes, 2005; Shumway & Wu, 2005; Strahilevitz et al., 2011) to characterize the likelihood of selling a stock in a time-series conditional on some factor (e.g. return magnitude or valence). The estimated model takes the following form:

$$h(t, x(t)) = h_0(t) exp\ (\beta_1 x_1 + \cdots + \beta_p x_p)$$

where the hazard rate, $h(t,x(t))$, on trial t is conditional on p predictors. The β coefficients are estimated from the data. The main assumption of the model is that the hazards are

proportionally dispersed at all time-points, but this model can be extended to include time-varying covariates (e.g. blink-rate). The model can also be stratified to incorporate repeated-measures designs, as in the current study. From this model, one can predict the hazard ratio of a subject choosing to sell a given stock at time *t* for each covariate *k* as:

$$exp\ (\beta_k) = \frac{h_0(t)exp\ (\beta_1 x_1 + \cdots + \beta_k(x_k + 1) + \cdots + \beta_p x_p)}{h_0(t)exp\ (\beta_1 x_1 + \cdots + \beta_k x_k + \cdots + \beta_p x_p)}$$

Here, the hazard ratio, $exp(\beta_k)$, is the ratio of two stocks with the same *k* covariates and where the numerator stock has an x_k that is one unit greater than the denominator (Barber & Odean, 2011). To maximize the potential of the model, a continuous variable (such as return on sale) can be transformed into dummy variables that represent 4% wide bins, set to 1 on trials within that range, or 0 otherwise. This allows the model to isolate the marginal hazard contributed by each bin when all other bins are zero. The reader is referred to Cox and Oakes (Cox & Oakes) for further details on the Cox proportional-hazard analysis. The analysis and subsequent Chi-squared tests were carried out using the **survival** package in R statistical software (Therneau, 2014). For any sale trial *x*, return on sale was calculated as:

$$Return(x) = \frac{Price_{sale} - Price_{purchase}}{Price_{purchase}}$$

where the difference in current sale price and previous purchase price was averaged relative to the previous purchase price. Trials were included in the analysis only if the selected stock had been purchased at least once previously. The model was stratified over subject, stock, and session to account for the effects of the repeated-measures design. Blink rate, counterfactual reward, sales count and return on sale were included as time-

varying factors as appropriate.

3. Results

3.1 Task Performance

Rats (N=24) completed two blocks of trials over the course of 7 sessions. On average, rats completed 100.0 (SD = 31.76) trials per 45-minute session in Block 2. Fig 2 depicts the resulting price fluctuations of the three stocks in example trading sessions. On free choice trials, rats showed no preference for either of three stocks across all sessions. The average proportion of trials in which subjects chose Stock 1, 2, and 3 were distributed tightly around chance at .33, .35, and .32, respectively. A repeated-measures ANOVA indicated that there was no significant effect of Stock identity *per se* (within-subjects, 3 levels: Stock 1, Stock 2, Stock 3) on choice ($F_{(2,46)} = 0.47$, $p = NS$).

Within the task, a rat could incur losses and gains based on its stock and option selection. The profit (loss) of a trade was translated into a liquid equivalent and added to (subtracted from) a 0.15ml reference point. Rats received a mean of 0.165ml (*SD* = 0.044ml) and median of 0.170 ml (*IQR* = 0.036 ml) per trial. Therefore, the set reference point (0.15ml) was within about 0.01ml of the experienced measures of central tendency.

During trading (i.e. when a buy or sell option was selected), rats received a profit on nearly 2/3 of trials (63.7%) and a loss on 36.4% of trials. Although rats profited on a greater proportion of trials, rats lost a mean of 0.052 (SD = 0.005) ml per trial while profiting only 0.026 (SD = 0.006) ml of reward on average ($M_{Difference} = 0.027$ml, *SEM* =

0.001ml), paired sample *t*-test: $t(23) = 21.04$, Cohen's $d = -4.30$, $p < .001$. Rats that lost more reward on average did not also gain more reward on average, Pearson's $r = .37$, $p = .08$. The average loss trial resulted in 0.10 (SD = 0.03) ml of reward, and only a small number of trials (0.7%) were associated with a very large loss that resulted in a payoff of no reward.

3.2.1 The effect of outcome tone on reward expectation and approach motivation

We hypothesized that rats' expectations of reward on the current trial would be differently affected by the gain or loss tone, which signaled whether the upcoming reward would be greater than or equal the RP (0.15 ml) or less than the RP, respectively. Given that humans experience losses roughly twice as strongly as equivalent gains (Kahneman & Tversky, 1979), we expected loss tones to elicit a measurable effect on behavior if rats had formed a reference point at 0.15 ml. We evaluated rats' movement times (MT) to collect reward after the tone onset in order to determine whether the loss tone may have shaped rats' motivation for reward, whereby slower MT's would indicate decreased motivation (Rivalan et al., 2013). A generalized linear mixed model was fitted to evaluate the relationship of current reward volume (ml) and outcome (0=Gain, 1=Loss), as well as their interaction, as fixed predictors of MT with random intercepts for subjects and random slopes for session. The GLMM results indicated that reward was not a significant predictor of MT at a loss ($b = -0.78$, $t = -1.05$, $p > .05$) or at a gain ($b = 0.73$, $t = 1.70$, $p > .05$). The interaction between reward and outcome was not significant ($b = -1.51$, $t = -1.76$, $p > .05$). However, we found a robust effect of loss on movement time ($b = 1.68$, $t = 14.95$, $p < .001$). This behavior suggests that rats learned to frame their expectations

about reward in terms of the gain or loss tone.

In order to further quantify the differences in expectations of reward at the reference point vs. a loss or a gain, we evaluated subject-wise variance in mean MT separately with a repeated measures ANOVA (3 levels = RP, Gain, Loss) and post-hoc tests with the Bonferroni correction on aggregated data. This analysis revealed that there was a significant main effect of outcome on MT, $F(1.16, 35.69) = 133.29$, $\eta_p^2 = .85$, $p < .001$. Fig 3a depicts the significant jump in mean MT that occurs at the 0.15 ml reference point, whereby rats moved 1.56 (*SEM* = 0.13) secs more quickly to collect a gain reward compared to a loss reward on average ($p < .001$). Rats approached a reference point (RP) reward 1.42 (*SEM* = 0.12) secs more quickly than a loss ($p < .001$), but only 0.14 (*SEM* = .04) secs more slowly than a gain on average ($p < .01$). Despite the continuity of the reward volume between a loss and a gain, rats demonstrate a significant discontinuity in approach behavior based on the outcome tone and volume relative to the RP. Rats were substantially slower to approach the reward at RP and at a loss.

3.2.2 The effects of the gain vs. loss tone on expected and experienced reward

We next sought to determine whether rewards were experienced differently depending on whether the volume was at, above or below the RP. Here, we used post-pump licking (PPL) at the reward spigot after offset of delivery (where reward was delivered at a constant rate, see section 2.2 for details) as a measure of experienced satisfaction and 'savoring' of reward (Wilson et al., 2006), as a measure of experienced reward satisfaction. A generalized linear mixed model was fitted to evaluate the effects of

current outcome and reward on PPL. The GLMM revealed a significant interaction effect between reward and current outcome ($b = 2.29$, $t = 4.64$, $p < .001$). As is shown in Fig 3b, PPL increased linearly with gain reward volume ($b = 2.88$, $t = 13.18$, $p < .001$), whereas there was no significant relationship between PPL and loss reward volume ($b = 0.59$, $t = 1.34$, $p > .05$). The linearly increasing trend demonstrates a significant discontinuity at the RP and at very high and very low reward volumes.

In order to quantify the discontinuity in licking behavior at the RP vs. a gain or loss, repeated-measures ANOVA (3 levels = RP, Gain, Loss) and post-hoc tests with Bonferroni correction were carried out on aggregated data. The repeated measures ANOVA revealed a robust main effect of outcome on PPL ($F_{(1.28, 29.53)} = 21.82$, $\eta_p^2 = .49$, $p < .001$). Rats spent on average 0.31 ($SEM = .06$) secs longer licking at the reward spigot after dispensing ceased on a gain trial compared to a loss trial ($p < .001$). In contrast to (pre-consumption) MT behavior, (post-consumption) PPL at the RP was not significantly different than loss trials ($M_{DIFF} = 0.13$, $SEM = .05$ secs, $p > .05$), yet it was 0.18 ($SEM = .02$) secs shorter than gain trials ($p <.001$). Thus, PPL at the RP more closely reflected loss trials than gain trials. Based on the pattern of PPL, rats treated the RP as being similar to a loss.

3.3.1 Stock choice

Although rats did not exhibit a group-level preference for any particular stock across sessions, visual inspection of individual rats' stock choices did indicate that preferences may have developed within sessions (see SI). In order to determine which

factors may have contributed to the choice of a stock on a given free-choice trial, we performed a multinomial logistic regression with average previous reward (ml), stock hole location, perseveration at the previous stock choice (0 = switch, 1 = stay), and blink rate (Hz) as predictors of the trinary choice. Here, we refer to average previous reward as the running average of reward volume from the 5 most recent trials on which that stock was selected. Blink rate is defined as the Hz of the on/off cycle at the selected stock hole at the time of choice. Neither the outcome (gain or loss) nor the location of response 2 (buy/hold/sell) on the immediately preceding trial were found to be significant predictors of stock choice and were excluded from the model. Regression coefficients are displayed in Table 1. In contrast with our analysis of stock choices pooled across rats and sessions (see section 3.1), the logistic regression indicated that rats demonstrated a slight preference for stock 3 relative to stock 2 (but not stock 1), $b = -0.12$, $t = -2.11$, $p < .05$.

3.3.2 The effect of average previous reward volume on stock choice

Average previous reward (of a stock) was added as a predictor to the model in order to ascertain whether rats returned more often to stock choices that were previously rewarding and less often to stock choices that were previously less rewarding. The average reward of the previous 5 trials on which that stock had been selected was the strongest predictor of stock choice ($b = 2.98$, $t = 7.15$, $p < .001$). This pattern suggests that rats were tracking the average reward earned from each stock option, and were more likely to repeat those choices that had recently resulted in the highest average reward volume.

492

493 **3.4.1 Option choice**

494 In order to characterize rats' choice of trading options (buy, hold, sell), we carried
495 out a multinomial logistic regression with average reward, previous loss, and change in
496 blink rate as predictors. Here, the we refer to previous trial as the immediately preceding
497 trial and not necessarily the trial on which the stock had been previously selected. Stock
498 choice was not entered as a significant predictor of option choice.

499 We also evaluated the experienced risk and profit associated with the buy, sell,
500 and hold options. Using the variability of outcome volume as a measure of an option's
501 risk, we found that the sell option was the riskiest ($\sigma^2 = .0028ml^2$) option, although not
502 significantly more so than the buy option ($\sigma^2 = .0026ml^2$, $p>.05$). The hold option ensured
503 a much more certain (or 'safe') outcome ($\sigma^2 = .0006ml^2$), leading to the average reward
504 more than four times more reliably than either of the trade options.

505 **3.4.2 The effect of change in blink rate on option choice**

506 Change in blink rate, which indicated the change in stock price from the previous
507 trial, appeared to have a clear effect on rats' choices (see Fig 4). If rats did indeed decode
508 the functional meaning of blink rate, we would expect to find an effect of change in blink
509 rate on option choice within each bin. If rats also learned to discriminate between the
510 direction of blink rate change (or null change) and the optimality of the buy/sell option,
511 we would expect that the null change bin would be predictive of the 'hold' option, the
512 negative change bin of the 'buy' option, and the positive change bins of the 'sell' option.
513 Indeed, we found that given no change in blink rate from the previous trial, rats were

514 79% (t = 39.00, p < .001) less likely to choose the buy option and 81% (t = 38.56, p <
515 .001) less likely to choose the sell option relative to the hold option (see Table 2 for
516 regression coefficients). However, rats were more likely to choose the hold option than
517 either of the trade options until the absolute difference in blink rate from the previous
518 trial was greater than 0.15 Hz (i.e. the equivalent of either 3 buy or 3 sell decisions).
519 Once blink rate had increased by more than 0.15 Hz relative to the previous trial, rats
520 were most likely to choose the sell option (t = 8.99, p < .001). When blink rate had
521 decreased by more than 0.15 Hz relative to the previous trial, rats were most likely to
522 choose the buy option (t = 6.95, p < .001).

523

524 **3.4.3 The effect of previous reward volume on option choice**

525 It stands to reason that rats chose the option hole based on the rate of reward
526 earned from selecting that option on prior trials within a session. In order to evaluate
527 whether rats based their choices on the previous reward rate of a given option, we
528 included the average previous reward of an option as predictor. Overall, the average of
529 the reward earned from selecting an option (buy, hold, sell) on previous trials was
530 negatively correlated with option choice on the current trial (b = -2.34, t = -3.46, p <
531 .001). Since the hold option was consistently associated with the lowest rate of reward,
532 this suggests that a rat was less likely to repeat a choice of the buy or sell option despite
533 the higher average reward rate.

534 **3.4.4 The effect of previous loss on option choice**

535 The outcome of the immediately preceding trial was a significant predictor of

option choice, which suggests that gains and losses affected subsequent trial behavior. Rats selected the hold option ~10% less after a previous loss compared to a previous gain. The likelihood of choosing both the 'buy' ($t = 3.30$, $p < .001$) and the 'sell' option ($t = 2.68$, $p < .01$) increased significantly after a loss. In order to determine whether this reflected perseveration, we performed paired sample *t*-tests of % stay choices of the 'buy' and 'sell' options after a gain vs. loss ('hold' conditions always resulted in a gain). We found no significant difference between perseveration at the 'buy' ($M_{Gain} = 25.6\%$, $M_{Loss} = 28.4\%$, $t(23) = -1.29$, $p > .05$) or 'sell' ($M_{Gain} = 27.4\%$, $M_{Loss} = 26.7\%$, $t(23) = 0.45$, $p > .05$) option after a loss vs. after a gain. Thus, this change in behavior reflected a decreased likelihood of choosing of the safer 'hold' option after a previous loss, and not perseveration at either the 'buy' or 'sell' options after a loss.

3.5.1 Optimal responses

In order to characterize the optimality of rats' responses (0=suboptimal, 1=optimal), we performed a binomial logistic regression (see Table 3 for regression coefficients) with the following predictors: previous outcome (0=gain, 1=loss), current trade option (0=hold, 1=trade), session quartile (1-4), and the absolute value of change in blink rate (0.05-0.30), as well as a previous outcome by current option interaction and previous outcome by change magnitude interaction. The absolute value of the change in blink rate was divided into bins separated by 1 unit change in price (i.e. 1 bin = $\Delta \pm 0.05$ Hz = Price Δ of ± 1 share). The two trade options (buy and sell) were collapsed into a bivariate choice measure representing rats' choice between trading or holding, or 'current trade option.' However, the optimality of an option was assessed *before* collapsing the measure, thereby preserving the correct proportion of optimal choices overall.

3.5.2 The effect of current option on optimal responses

In general, rats were less effective at choosing the trade options optimally relative to the hold option. Rats were 71% less likely to choose an optimal trade response than an optimal hold response (b = -1.25, t = -30.66, p < .001). Further analysis suggests that although rats had learned to trade vs. hold optimally, they frequently chose the wrong trade option at small changes in absolute blink rate. This effect was stronger for the 'sell' condition than for the 'buy' condition (see Figs 4a-b). This pattern of behavior suggests that rats learned to respond optimally in the task, but had difficulty discriminating small changes in the blink rate.

3.5.3 The effect of absolute change in blink rate change on optimal responses

We were also interested in comparing behavior at small changes in blink rate vs. large changes in blink rate. An absolute value of 0.05 Hz was associated with the lowest proportion of optimal responses (see Fig 4b), which likely reflects greater difficulty in differentiating between a change of ±0.05 Hz *vs.* no change. The effect of absolute blink change was highly significant at each level relative to a of 0.05 Hz (all p's < .001, see Table 3). As the magnitude of blink rate change increased from 0.05 Hz, the likelihood of responding optimally rose. At the highest absolute blink rate change, rats chose the optimal trade outcome in ~60% of trials and were 11.5 times more likely to choose the optimal outcome relative to ±0.05 Hz (b = 2.45, t = 5.59, p < .001).

At zero change, rats were 5.7 times more likely to choose the optimal hold outcome relative to an absolute change of 0.05 Hz (b = 1.74, t = 40.49, p < .001). An optimal hold response at zero was significantly more likely than an optimal trade

response at magnitudes of 0.05 Hz (t = -40.49, p <.001), 0.10 Hz (t = -12.15, p <.001), and 0.15 Hz (t = -5.50, p < .001), but *not* more likely than 0.20 Hz (t = -0.91, p >.05), 0.25 Hz (t = 0.76, p>.05), or 0.30 Hz (t = 1.62, p >.05). This pattern of behavior suggests that rats were just as good at responding optimally in the hold condition at zero change in blink rate as they were at the correct trade option at large changes in blink rate.

3.5.4 The effect of session quartile on optimal responses

We also examined whether learning occurred over the course of the sessions. Although session number (1-7) was not entered as a significant predictor into the model, we did find that session quartile was a predictor of optimal responding over the course of a session. In any given session, rats made more (~15%) optimal choices in the last quartile relative to the first (t = 2.95, p < .01). There were no differences between the 4th quartile and either the 2nd (t = -0.26, p > .05) or 3rd quartile (t = -0.35, p >.05).

3.5.5 The effect of previous outcome valence on optimal responses

We next sought to further explore how a loss on the immediately preceding trial affected choice behavior on the current trial. We found a robust effect of previous loss on optimal responses. As illustrated in Fig. 5a, a previous loss reduced the overall likelihood of an optimal response by nearly 10% (t = -3.90, p < .001). The interaction term between a previous loss and a trade (vs. hold) option was also significant (b = 0.84, t = 2.30, p < .001). Rats were ~15% less likely to choose the hold option optimally after a loss trial (Fig. 5b, t(23) = -7.56, p < .001), yet there was no difference in optimal choice of the trade option after a loss trial (t(23) = 0.79, p > .05). We also found a significant relationship between the absolute value of blink rate change and the outcome of the

previous trial. At an absolute change of 0.10 Hz (i.e. at a price differential of ± 2 shares from the previous trial), rats were 45% less likely to choose optimally after a loss (b = -0.60, t = -4.08, p < .001) relative to a loss at an absolute change of 0.05 Hz. We did not find significant previous outcome by absolute change interactions at any other levels. This pattern was as if rats became more risk seeking after a loss, resulting in a higher likelihood of responding suboptimally.

3.6.1 Win-Stay/Lose-Shift vs. Optimal Strategy

We next investigated whether simple win-stay/lose-shift (WSLS) strategy at a nosepoke hole location interacted with choice of the optimal strategy based on option choice (e.g. choice of the left 'sell' option after a blink rate increase). In order to establish whether there was an interaction between the model-free WSLS strategy and the model-based optimal strategy, we compared the percentage of optimal choices on congruent trials where WSLS was also the optimal option strategy vs. incongruent trials where WSLS was not the optimal strategy. On average across rats, the WSLS strategy was also the optimal strategy on 44.5% of all trials. Of those trials, WSLS was the optimal strategy on 38.2% of trials after a gain and 78.0% of trials after a loss.

We carried out a repeated measures ANOVA on rats' average proportion of optimal choices with strategy congruence (congruent: WSLS = optimal; incongruent: WSLS = suboptimal) and previous outcome (Win-Stay vs. Lose-Shift) as within-subject factors. We found no main effect either congruence ($F_{(1,23)}$ =1.14, p = *NS*) or previous outcome ($F_{(1,23)}$ = 0.20, p = *NS*) on the proportion of optimal choices. However, one-sample ANOVA's revealed that the overall proportion of optimal win-stay trials (M = 40.5%, SD = 4.3%) was significantly higher than chance (33.3%), t(23) = 8.21, p < .001).

Conversely, the overall proportion of optimal lose-shift trials (M = 37.3%, SD =8.4%) was significantly lower than chance (66.6%), $t(23) = -17.14$, $p < .001$. Thus, rats' optimal WSLS behavior was better than expected after a gain, but worse than expected after a loss.

We found a significant interaction effect of previous outcome on congruent trials compared to incongruent trials ($F_{(1,23)} = 9.15$, $\eta_p^2 = .29$, $p < .01$). As illustrated in Figure 5c, rats were more likely to choose the optimal strategy if it aligned with the win-stay strategy compared to trials on which they conflicted, although post-hoc paired t-tests revealed that this difference did not reach significance ($M_{\text{Diff}} = 4.7\%$, $t(23) = 1.89$, $p > .05$). Conversely, rats were less likely to choose the optimal lose-shift strategy when the two strategies aligned (35.5%) compared to when they conflicted (44.3%), and this difference was significant: $M_{\text{Diff}} = 8.8\%$, $t(23) = -2.66$, $p < .05$. This behavior suggests that a previous loss differentially engaged model-free and model-based strategies in the task.

3.7.1 The effect of returns on choice of the sell option

Over the 7 testing sessions, returns from selling a stock ranged from -57.1% to 26.5% (SD = 8.5%), with an average return of 0.3%. In order to determine whether this was indicative of the disposition effect, we determined the proportion of gains realized (PGR) and the proportion of losses realized (PLR) for each rat across each stock and session (see Methods). PGR and PLR represent the number of trials a rat sold at a gain/loss with respect to the opportunities it had to sell at a gain/loss. Empirical studies of

648 the disposition effect in humans have found that PGR is consistently above PLR (Odean,
649 1998), which indicates that investors are selling at a gain on a greater proportion of
650 opportunities than at a loss. We found that rats had a mean PGR of .12 (*SEM* = .02) and
651 mean PLR of .09 (*SEM* = .01). The results of a paired-sample t-test revealed that rats
652 realized gains more often than they realized losses, $t(23) = 2.22$, Cohen's $d = 0.80$, $p <$
653 $.05$, with no correlation between PGR and PLR (Pearson's $r = .33$, $p = .12$).

654

655 Our analysis of the disposition effect proceeded by fitting a Cox Proportional
656 Hazards model (see Methods). Stratifying over subject, stock, and session, we also
657 included two factors as time-varying covariates in the model: counterfactual reward and
658 the number of times the rat had chosen the 'sell' option previous to that trial.
659 Counterfactual rewards represented the volume of reward a rat *could* have earned had it
660 chosen a different option. We found that the addition of counterfactual reward
661 significantly increased the model's R^2 from .57 to .65 ($\chi^2(1) = 635.23$, $p < .001$). This
662 pattern is as though rats were considering the outcomes of the alternative options when
663 making a response. Fig 6 illustrates the 'hazard' of an animal selling a stock on a given
664 trial based on the potential returns (see Table 4 for model coefficients). In line with the
665 PGR-PLR analysis, this revealed that rats were more likely to sell a stock at a gain than at
666 a loss, i.e. rats exhibited the disposition effect.

667

668 **4. Discussion**

669 We found evidence that rats exhibited a number of investor biases that are well-

established in humans (and to some extent in non-human primates as well, see Santos and Platt (Santos & Platt) for a review), but that remained largely unexplored in rats. Within our task, rats nosepoked into blinking holes in order to choose and subsequently buy, hold or sell a virtual stock. In order to perform optimally, rats were required to integrate previous reinforcement experience with current changes in blink rates. Suboptimal behavior resulted in a 'loss', which was equal to the reference point (RP = 0.15 ml) less the liquid equivalent of the trading loss. Although loss rewards were less than the reference point, loss trials still resulted in an average of 0.10 ml reward.

By explicitly signaling trial gains and losses with respect to a reference point of 0.15 ml, we were able to infer how expectations about lower or higher reward volumes affected rats' motivation to collect reward. After a gain tone, MT to collect reward became substantially quicker (see Fig 3a). Before even experiencing the amount of reward, rats' languid approach after hearing the loss tone implies that the stimulus reshaped expectations about the desirability of the reward volume. Furthermore, although reward outcomes at the 0.15 ml reference point were signaled as a gain, rats approached the RP of reward significantly more slowly than other gain outcomes. This discontinuity suggests that rats had some prior expectations of the reward volume based on the outcome of their choices that changed approach behavior at the RP. Thus, rats' behavior is consistent with the idea that the loss tone conceptually reframed e.g. a 0.10 ml reward as a 0.05ml loss from the 0.15 ml reference point, and a 0.15 ml reward as a zero gain.

692 Post-pump licking (PPL) was measured as a behavioral indicator of post-
693 consumption 'savoring' or 'satisfaction' (Wilson et al., 2006). Rats spent less time
694 licking at the reward spigot (PPL) after a loss outcome compared to a gain outcome (see
695 Fig 3b) but not an RP outcome, which indicates that rats savored losses less so than they
696 did gains, but not less than outcomes at the RP. After a small jump from the RP, PPL
697 increased monotonically with reward volume after a gain. We found no significant linear
698 trend between reward volume and PPL for losses. Although the custom-designed reward-
699 delivery system was built to deliver the liquid reward at a slow, constant rate, it is
700 possible that this effect can be explained purely based on greater reward volumes.
701 However, this is unlikely to explain the discontinuity in PPL at the RP (which was paired
702 with a gain tone). This behavior suggests that, post-consumption, rats experienced gain
703 rewards differently than a reference point reward or a loss reward, whereby RP and loss
704 rewards were less worth savoring than gain rewards.

705

706 The reframing of gains as losses based on a reference point (as opposed to
707 objective zero) can be defined as 'anchoring' (Kahneman & Tversky, 1979). Both MT
708 and PPL behavior suggest that rats formed an anchor at 0.15 ml of reward rather than at 0
709 ml. This finding supports and extends recent work exploring reference-dependent
710 behavior in rats (Bhatti et al., 2014; Constantinople, Piet, & Brody, 2019)}. Bhatti and
711 colleagues (Bhatti et al., 2014) found that rats in a T-maze preferred an arm that both
712 contained and delivered one pellet (gain frame), over an arm that contained four visible
713 pellets but only delivered one (loss frame). The researchers found that rats preferred the
714 1-pellet arm to the 4-pellet arm, putatively because receiving only one of the four pellets

715 framed the outcome as a loss.

716

717 The task design by Bhatti and colleagues (Bhatti et al., 2014) assumes that rats'
718 reference point for losses and gains was based on the difference between visible pellets
719 and delivered pellets, rather than on previous experience or expectations. Only one pellet
720 had been consistently delivered across all trials, which implies that a rat's reference point
721 should never have been anything other than one. Thus, it is unclear whether the observed
722 results reflected a learned avoidance due to the perceived loss of three pellets, or rather
723 more simply to the punishing effect of inaccessible food (Amsel, 1958; Lawson & Marx,
724 1958). Our own task design differs in that the reference point of reward here was
725 explicitly established in block 1, and closely approximated the mean and median volume
726 of reward delivery in block 2. In other words, the behavioral effects of gains and losses in
727 our task were categorized relative to rats' average expected reward (or the '*status quo*')
728 rather than to zero.

729

730 We show that rats exhibited both optimal and suboptimal performance within the
731 task, and that rats appeared to be utilizing a mixture of simple reinforcement and more
732 sophisticated model-based strategies. Although classical views of serial pattern learning
733 (e.g. Capaldi et al., 1980; Hulse, 1973) tended to view strategy use as mutually exclusive,
734 recent work supports the perspective that the concurrent mechanisms, such as the
735 combination observed here, are more likely to drive the encoding and reproduction of
736 complex sequential behavior (Muller & Fountain, 2010, 2016). This is exemplified here

in that task performance was highly dependent on both the magnitude of the change in blink rate and the outcome of the previous trial. Overall, rats' optimal response rate was highest when choosing the hold option, although this was not significantly different than optimal responses in each of the trade options at greater absolute changes in blink rate (Fig 4). Suboptimal behavior at lower magnitudes of change (i.e. at differences of ±1-3 shares) could primarily be explained by rats' perceptual uncertainty in discriminating blink rate changes, and the consequent uncertainty about the optimal response. The weak price change signal led to a greater likelihood of choosing the suboptimal hold option than the optimal trade option at low blink rate changes. This indicates that rats were risk averse under uncertainty about price changes: they preferred to select the hold option for a smaller, safer reward over a riskier trade option for a larger, riskier reward when the optimal outcome was unclear. This is a well-characterized bias, referred to as the 'equity premium puzzle' in human stock markets, whereby investors continue to purchase 'risk-free' bonds despite their considerable underperformance with respect to stocks (Benartzi & Thaler, 1995). These results are supported by numerous studies demonstrating that rats, like humans, tend to be risk averse under uncertainty, and furthermore that such behavior is modulated by the mesolimbic dopamine system, which is largely conserved across species (Constantinople, Piet, Bibawi, et al., 2019; Kagel et al., 1986; Pais-Vieira et al., 2007; Simon et al., 2009; St Onge & Floresco, 2009)}.

We observed a robust decrease in optimal performance on trials immediately preceded by a loss (Fig 3a). Behavioral research in humans (Charness & Levin, 2005) has found that performance drops by nearly 50% when reinforcement and utility

maximization rules conflict. Thus, rats may have found it difficult to repeat an option that previously resulted in a loss, but that represents the optimal choice on the current trial. Similarly, a previous loss may have interfered with a rat's ability to update expectations (based on prior reinforcement) with new information from the current trial. This is supported by the significant interaction effect (Fig 5c), whereby optimal performance increased when the win-stay strategy was also the optimal response, yet decreased on congruent trials with a prior loss. Neurophysiological and behavioral evidence by Steiner and Redish (Steiner & Redish, 2014) support the possibility that rats experience emotions similar to human disappointment and regret after making mistakes in a neuroeconomic decision-making task. A potential source of interference may therefore have been the animal's emotional reaction to losing the previous trial (Mendl et al., 2009). This is consistent with human research suggesting that one's immediate emotions can influence a number of decision factors, such as: the perceived likelihood of positive and negative outcomes (Johnson & Tversky, 1983), selective attention to decision attributes (Bower, 1981; Forgas, 1989), and the depth of subsequent information processing, such as whether or not one relies upon heuristics (Bodenhausen et al., 1994; Tiedens & Linton, 2001).

The interaction between a previous loss and current choice of a trade or hold option was a significant predictor of optimal choices. Prospect Theory (Kahneman & Tversky, 1979) predicts that individuals will be more likely to take risks when in a loss frame than in a gain frame. This increase in suboptimal behavior is often interpreted as an attempt to 'break even,' and has been demonstrated in both humans and monkeys

783 (Lakshminarayanan et al., 2011; Tversky & Kahneman, 1981). We found that rats, too,
784 were far more likely to choose the riskier buy/sell options than the safer hold option on
785 trials immediately preceded by a loss. Interestingly, this led to slightly (although not
786 significantly) *increased* optimal choice of the trade options, yet *decreased* optimal choice
787 of the hold option (Fig 5b). This suggests that loss aversion may have promoted better
788 trading choices by mitigating the negative impact of risk aversion at small price changes
789 (i.e. high uncertainty regarding the optimal outcome). To adopt the language of
790 Kahneman and Tversky (Kahneman & Tversky, 1979), it is as though a rat that "has not
791 made peace with his losses is likely to accept gambles that would be unacceptable to him
792 otherwise" (p. 287). Taken from an evolutionary perspective, these loss-averse behaviors
793 can be adaptive. For example, when food turns out to be scarcer than predicted, it
794 becomes necessary to take on additional risks in order to ensure one gathers enough
795 resources to survive (Li et al., 2012; Stephens & Krebs, 1986). However, it appears that
796 in the context of the rat stock market, these naturally adaptive strategies become
797 deleterious.

798

799 Within the task, rats were given trial-by-trial feedback about the quantity gained
800 or lost on individual trading decisions, but they did not receive any direct feedback
801 regarding long-term profits or losses (i.e. returns) from the sale of a stock relative to its
802 original purchase value. In the current study, rats realized a larger proportion of gains
803 than losses. In other words, rats were generally reluctant to sell at negative returns (see
804 Fig 6), despite not receiving direct feedback about any given return. This is indicative of
805 the disposition effect characterized in humans by traders' reluctance to sell losing stocks,

while selling winning stock too quickly (Shefrin & Statman, 1985). This is further supported by the results of a Cox proportional hazards model (see Fig 6), which illustrates rats' general reluctance to sell at negative returns. An interesting deviation from human behavior however, is the observation that rats were far less willing to sell at a loss compared to humans. It is possible that this reflects a species-level difference in the general willingness or ability to incur a smaller immediate loss in order to avoid a larger future loss. However, it may also be the case that the rats in our task simply did not have enough information/experience to reach such a conclusion, and that the disposition effect is a natural byproduct of 'myopic loss aversion.'

The nature of the conflict one experiences when selling at a short-term loss in order to avoid greater long-term losses resembles that of reverse-reward experiments, in which subjects' initial choice between a small reward and a large reward (e.g. 3 candies or 6 candies) results in the non-chosen reward amount (i.e. choice of 6 candies results in 3 candies) (Boysen & Berntson, 1995; Carlson et al., 2005; Hershberger, 1986). Reverse-reward tasks require inhibitory control in order to suppress a response that is not useful in attaining a goal (Christ et al., 2001) – an undertaking that is immensely difficult for pre-school children (Carlson et al., 2005), chimpanzees (Boysen & Berntson, 1995), and cockerel chicks (Hershberger, 1986). Higher levels of inhibitory control in children are predictive of better test scores (Shoda et al., 1990) and wealth (Moffitt et al., 2011) later on. Similarly, previous studies of the disposition effect have found that the effect is diminished in individuals with higher IQ's (Grinblatt et al., 2012) and with greater trading experience (Chen et al., 2007), which presumably represents greater inhibitory

control in these groups. Taken together, these results suggest that the behavioral biases observed in this task need not arise from mechanisms supporting higher levels of cognition, but may require higher levels of cognitive control to overcome.

It is important to address a number of limitations within the current paradigm. We developed this task to explore the concept that rats exhibit similar suboptimal behavioral patterns to those of humans within a simulated 'stock market' environment. Given the complexity of the task, we remained agnostic to the content of rats' learning and to any assumptions about the specific strategies that rats might develop at the outset. Our intention was to first evaluate the behavioral output of the paradigm, so that future studies might then systematically manipulate individual task components in order to identify specific neural and computational contributions to the established behavior.

However, the complex sequential choices required from the current task should be considered in the light of their complexity from a serial learning perspective (Constantinople, Piet, Bibawi, et al., 2019; Garlick et al., 2017; Muller & Fountain, 2016)}. While we show that some of the rats' behavior reflects a structural understanding of the optimal task structure (i.e. go right to 'sell' on the second response, after observing a positive change in blink rate at the first response hole), it is also possible that uninvestigated simple reinforcement-based strategies better explain performance. We refrained from training rats to criterion on the optimal strategies prior to testing in order to allow for strategies to develop spontaneously, but this also precludes us from excluding this alternative entirely. Future work could employ computational and behavioral methods to better characterize the content of learning, and how that might

contribute to subsequent suboptimal behavior, as rats learn to perform individual components of the optimal task behavior (e.g. perseverate at nosepoke port at zero blink change vs. respond at the left port after blink rate increases).

Furthermore, we limited the information available to rats in the task to the current price of each stock (blink rate) and the outcome of a trade/hold option (tone and reward volume). It should be noted that while the differences in reward volume between payouts were at times small (albeit in line with other tasks, see e.g. Constantinople et al., 2019a and b), the linear trend depicted in Figure 3b demonstrates that rats were capable of discriminating between outcome volumes of reward. While the task design still afforded rats the ability to develop a personal 'trading' history of reinforcement, such a restrictive model did not allow rats to directly associate changing blink rates with other rats' actions or to ascertain an overview of current portfolio holdings. Thus, rats were forced to trade incrementally and to measure outcomes based on very small changes in reward. It is therefore notable that these behavioral biases can arise in a task in which the potential for both theory of mind (Whiten, 1991) and portfolio optimization (Markowitz, 1959) have been virtually precluded. The prospect that rats exhibit behavioral biases in this context implies not only that the behaviors are underpinned by neural processes that both humans and rats share, but also that they are not artifacts of culture, education, or even notions of how currency works.

In the current version of the task, price was determined at any given moment based on the cumulative 'demand' from the cohort of four rats. This allows for serial

874 autocorrelation in prices. There exists a long-standing debate about whether or not the
875 average stock's movement over time carries momentum, meaning it is more likely to
876 continue on its current trajectory than to reverse directions, or whether its movement
877 more closely approximates a random walk. While many argue that asset fluctuations are
878 entirely random (notably, Malkiel, 1973 and Fama, 1965), others provide evidence for
879 serial autocorrelation (Bondt & Thaler, 1985; Campbell, Lo, & MacKinlay, 1997; Shiller,
880 2003), suggesting that psychological factors such as herding behaviour lead to serial
881 correlation in prices over time. Future versions of the task could help resolve this debate
882 by contrasting rats' behavior in a condition with demand-based pricing (as employed
883 here) with controlled pseudo-random prices simulated *a priori*.

884

885 The rat model presented here represents a novel means of interrogating potential
886 links between market mechanisms, investor behavior, and investor cognition. One area
887 rife for future study remains the exploration of putative interactions between model-based
888 and model-free strategies in rodents that are free of confounding 'human' cognitive
889 factors, such as preconceived notions of how a stock market works or individual
890 differences in culture and education. The development of a rat stock market task opens up
891 the opportunity to establish a computational and neurobiological account of decision-
892 making that has the potential to lend predictive power to current economic and financial
893 models. The finding that the behavioral biases observed here likely arise from a
894 combination of learning strategies provides key insight into the mechanisms that may be
895 governing investor biases in the brain. Areas commonly implicated in reinforcement
896 learning, such as the dopaminergic midbrain (Schultz et al., 1997) and in serial pattern

recognition, such as NMDR receptor mediated plasticity within the hippocampus (Fountain & Rowan, 2007), represent natural targets for future investigations into stock market traders' "inner rat."

References

Tables

Table 1

Summary of multinomial logistic regression coefficients of stock choice

Parameter	*Coefficient*	*SE*	*t*	Exp(*B*)	95% *CI*
Avg. Previous Reward	2.981***	0.417	7.146	19.711	
Stock | S1					
Intercept	-0.080	0.054	-1.478	0.923	
Hole Location					
3	0.130	0.075	1.736	1.139	
4	1.183***	0.196	6.030	3.265	
Perseveration	0.127*	0.058	2.182	1.135	
Blink Rate (Hz)	-0.539***	0.095	-5.690	0.583	
Stock | S2					
Intercept	-0.115*	0.054	-2.115	0.891	
Hole Location					
3	0.294***	0.074	3.989	1.342	
4	0.357	0.198	1.802	1.429	
Perseveration	0.125*	0.058	2.160	1.133	
Blink Rate (Hz)	-0.111	0.093	-1.191	0.895	

Note. Stock reference value: Stock 3; Nosepoke hole reference value: Hole 2; Perseveration reverence value: Switch. R2 = 0.01, Log likelihood = -8343.00, LR test = 116.78***. **p<.01, ***p<.001

Table 2

Summary of multinomial logistic regression coefficients of option choice

Parameter	*b*	*SE*	*t*	Exp(*B*)	95% *CI*
Avg. Previous Reward	-2.338***	0.676	-3.460	0.097	-3.663 / -1.014
Option | Buy					
Intercept	-1.582***	0.041	-39.004	0.206	-1.661 / -1.502
Previous Loss	0.184***	0.056	3.298	1.202	0.075 / 0.294
Δ Blink Rate (Hz)					
<= -.15	3.670***	0.528	6.947	39.253	2.635 / 4.705
-.15 to -.01	0.985***	0.055	18.036	2.679	0.878 / 1.093
.01 to .15	1.194***	0.051	23.434	3.301	1.094 / 1.294
>= .15	3.385***	0.407	8.324	29.510	2.588 / 4.182
Option | Sell					
Intercept	-1.682***	0.044	-38.555	0.186	-1.767 / -1.596
Previous Loss	0.153**	0.057	2.676	1.165	0.041 / 0.265
Δ Blink Rate (Hz)					
<= -.15	3.615***	0.533	6.788	37.139	2.571 / 4.658
-.15 to -.01	1.255***	0.053	23.469	3.508	1.150 / 1.360
.01 to .15	1.069***	0.054	19.867	2.913	0.964 / 1.175
>= .15	3.652***	0.406	8.986	38.554	2.855 / 4.449

Note. Option reference value: Hold; Previous Loss reference value: Gain (0); Δ Blink Rate reference value: 0. $R^2 = 0.054$, Log likelihood = -13,488.030, LR test = 1540.443*** (df = 13). ***p*<.01, ***p<.001.

Table 3

Summary of binomial logistic regression coefficients of optimal response

Parameter	*Coefficient*	*SE*	*t*	Exp(*B*)	95% *CI*
Optimal					
Intercept	-0.653***	0.045	-14.521	0.521	-0.741/ -0.565
Previous Outcome					
Loss	-0.337***	0.086	-3.898	0.714	-0.506 / -0.168
Current Option					
Trade	-1.249***	0.041	-30.658	0.287	-1.329 / -1.169
Session Quartile					
2	0.140**	0.045	3.124	1.150	0.052 / 0.228
3	0.134**	0.049	2.744	1.143	0.038 / 0.230
4	0.154**	0.052	2.954	1.166	0.052 / 0.256
Absolute Blink Change					
0	1.736***	0.043	40.491	5.676	1.652 / 1.820
2	0.909***	0.071	12.780	2.483	0.770 / 1.049
3	1.113***	0.115	9.660	3.043	0.887 / 1.339
4	1.555***	0.201	7.719	4.734	1.160 / 1.950
5	1.959***	0.294	6.669	7.093	1.383 / 2.535
6	2.445***	0.437	5.594	11.528	1.588 / 3.301
Previous Outcome × Current Option					
Previous Loss × Trade	0.834***	0.093	8.934	2.302	0.651 / 1.017
Previous Outcome × Absolute Blink Change					
Previous Loss × 0	-0.145	0.101	-1.431	0.865	-0.344 / 0.054
Previous Loss × 2	-0.596***	0.146	-4.082	0.551	-0.883 / -0.310
Previous Loss × 3	-0.195	0.232	-0.842	0.823	-0.649 / 0.259
Previous Loss × 4	-0.385	0.487	-0.790	0.680	-1.340 / 0.570
Previous Loss × 5	0.045	0.918	0.049	1.046	-1.754 / 1.844
Previous Loss × 6	-0.666	0.856	-0.778	0.514	-2.344 / 1.011

Note. Optimal reference value: Suboptimal; Previous Outcome reference value: Gain (0); Session Quartile reference value: 1; Absolute Value of Blink Change reference value: 1. R2 = 0.165, Log likelihood = -9683.952, LR test = 3824.850*** (df = 19). ***p*<.01, ****p*<.001

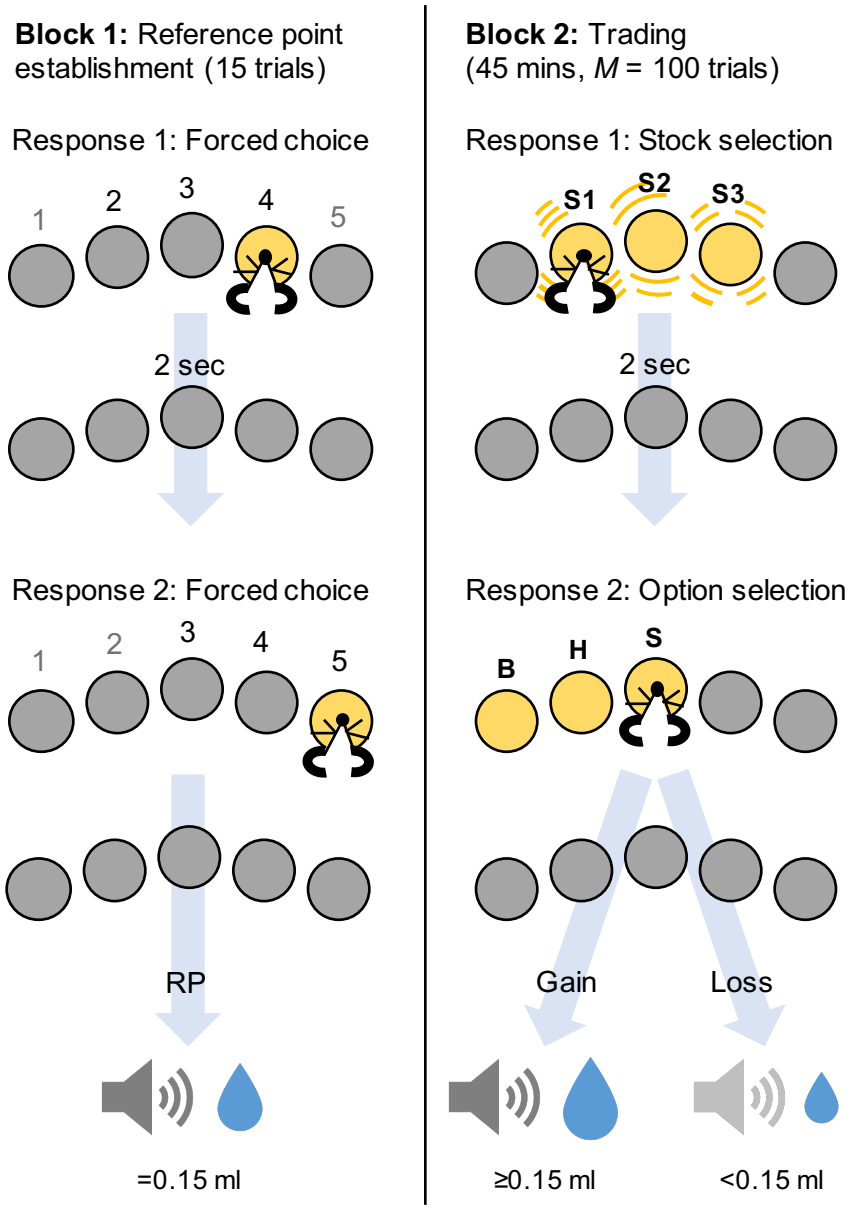
Table 4

Summary of Hazard Analysis of Return on Sale

Parameter	*Coefficient*	*SE*	*t*	*HR*	95% *CI*
Return					
< -22%	-1.155*	0.477	-2.420	0.004	-2.420 / -0.220
-22% − -18%	1.650	1.011	1.632	0.067	-0.331 / 3.630
-18% − -14%	2.128*	0.878	2.424	0.108	0.408 / 3.848
-14% − -10%	3.164***	0.903	3.504	0.304	1.394 / 4.933
-10% − -6%	3.312***	0.896	3.695	0.353	1.555 / 5.069
-6% − -2%	3.939***	0.895	4.401	0.660	2.185 / 5.693
-2% − 2%	4.354***	0.893	4.874	1.000	2.603 / 6.105
2% − 6%	5.229***	0.897	5.828	2.399	3.471 / 6.988
6% − 10%	5.692***	0.911	6.250	3.811	3.907 / 7.477
10% − 14%	5.803***	0.948	6.123	4.257	3.945 / 7.660
14% − 18%	6.093***	1.016	5.999	5.692	4.103 / 8.084
≥18%	5.316***	1.225	4.340	2.617	2.916 / 7.717
Sales Count	-0.147***	0.005	-30.044	0.863	-0.157 / -0.138
Counterfactual Reward	-6.193***	1.271	-22.856	0.002	-6.725 / -5.662

Note. Reference value: -2% – 2%. Based on random subject slopes. Likelihood ratio test(14) = 3352, p <.001. *p<.05, ***p<.001

957 **Figure Legends**



958

Blink Rate		Payoff		
trial t-1	trial t	Buy	Sell	Hold
2 Hz	2 Hz	0.15 ml	0.15 ml	0.17 ml
2 Hz	3 Hz	0.05 ml	0.25 ml	0.17 ml
3 Hz	2 Hz	0.25 ml	0.05 ml	0.17 ml

Fig 1: Task Schematic. 6 cohorts of 4 rats (N=24) completed two blocks of trials in which they made 2 nosepokes into illuminated nosepoke holes within a 5-hole array. At task onset, in block 1 (left panel) each rat was given fifteen trials to habituate to the reference point (RP) of 0.15 ml reward, where rats made a first nosepoke response in one of three illuminated holes for 'stock' selection (here, hole 4). All lights immediately extinguished for 2 seconds once the nosepoke response was initiated. After the delay, one of three holes (either the hole selected in the 'stock selection' response or one of the adjacent holes to either side of it; here, either hole 3, 4, or 5), was illuminated to indicate the second forced-choice response hole was available. Rats nosepoked into the illuminated hole (here, hole 5) to mimic 'option' selection, at which point all hole lights were immediately extinguished and the conditioned tone-light stimuli signalled the availability of sweet sodium saccharine reward at the reward spigot. A volume of 0.15 ml of reward was delivered for every trial in block 1, irrespective of the rats' choices. Once all four rats had completed all fifteen trials in block 1, the cohort advanced to block 2 (right panel) as a group. In free choice trials (as depicted here) the three center holes blinked simultaneously at variable rates to indicate the 'prices' of stocks 1, 2, and 3. Rats nosepoked into a blinking hole to select a stock (here, 'S1'), then experienced an immediate 2 second delay with all lights extinguished. Then, three holes were illuminated and rats nosepoked into either the 'buy' (B), 'hold' (H), or 'sell' (S) hole to make their option selection. All hole lights were again extinguished and either a gain tone or a loss tone, paired with the light at the reward magazine, indicated whether the reward volume would be greater than or less than the reference point (0.15 ml). Reward was available immediately upon presentation of the tone-light stimuli, and movement time (MT) was measured from tone onset to lick onset at the reward spigot. The volume of reward delivered was greater than the reference point (RP) for trades resulting in a gain and less than the RP for trades resulting in a loss.

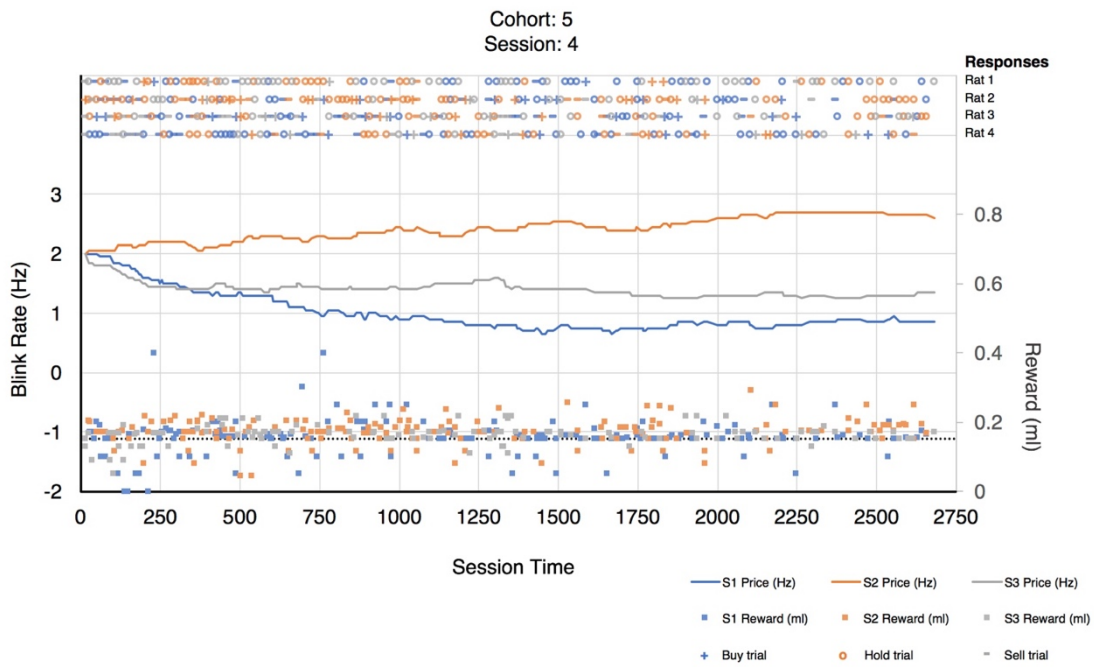


Fig 2: Example Behavior. Blink rate (Hz) fluctuations during a single testing session from one cohort of four rats. Colored lines represent price evolution (in arbitrary units) operationalized by blink rate along the left axis for each of the three stocks during the 45-min session (block 2). Stock prices moved both up and down over time, depending on the cumulative choices made by the four rats. The top panel represents individual choices made on each trial by each of the four rats over the course of the session. 'Buy' responses ('+') increased the price of that color stock, 'Sell' responses ('-') decreased the price of that color stock, and 'Hold' responses ('o') had no effect on stock price. The bottom panel represents reward values (ml) from a response in that colored stock and is measured along the right axis. The dotted line represents the reference point of reward (RP) at 0.15 ml. Readers are referred to the SI for further such graphs of other cohorts as well as the same cohort in other sessions.

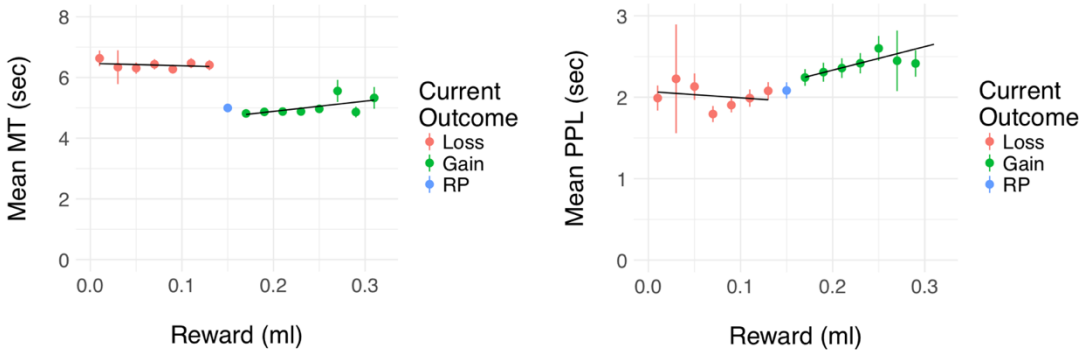


Fig 3: Effects of a current loss on behavior pre- and post-reward consumption. (A) Rats' movement times (MT) to collect reward were measured as an indicator of

motivation for reward. Upon hearing a gain tone, rats' mean movement time to collect a gain reward was 4.87 secs, which was significantly more quickly than a loss tone, (M_{LOSS} = 6.43 sec, p<.001). Despite hearing the same gain tone, rats approached the reference point (RP, blue) 0.14 secs more slowly than a gain (p<.01). There was no linear relationship between MT and reward at either a loss (t=-1.05, p>.05) or at a gain (t=1.70, p>.05). Error bars represent SEs. **(B)** The time spent post-pump licking (PPL) was measured as an indicator of 'savoring' reward post-consumption. In contrast to MT behavior, rats' average PPL at the RP (M= 2.11, SE= 0.10 sec) was significantly different than the average gain (M_{DIFF}=-0.18, SEM=0.02 secs, p<.001) but not the average loss (M_{DIFF}=0.13, SE=0.05 secs, p>.05). We found a significant interaction between reward volume and outcome (t=4.64, p<.001), whereby PPL increased linearly with the volume of the gain reward (t=13.18, p<.001), but not a loss reward (t=1.34, p>.05). PPL at the RP is 0.18 secs shorter than at a gain (t=7.31, p<.001) and 0.12 secs longer than a loss (t=-2.48, p <.05). Error bars represent SEs**. (C)** There was a significant interaction between previous outcome (gain or loss) and trial type (congruent: win-stay/lose-shift = optimal; incongruent: win-stay/lose-shift ≠ optimal) on rats' choice of the optimal response, ($F_{(1,23)}$=9.15, p<.01). Win-stay/lose-shift was evaluated based on location of the nosepoke hole. On average, rats chose the optimal response ~5% more when the optimal option was a win-stay, although post-hoc tests showed that this fell short of significance (t(23) = 1.89, p>.05). Intriguingly, rats chose the optimal option nearly 10% less after a loss on congruent trials (t(23)=2.66, p<.05). Error bars represent 95% CI's.

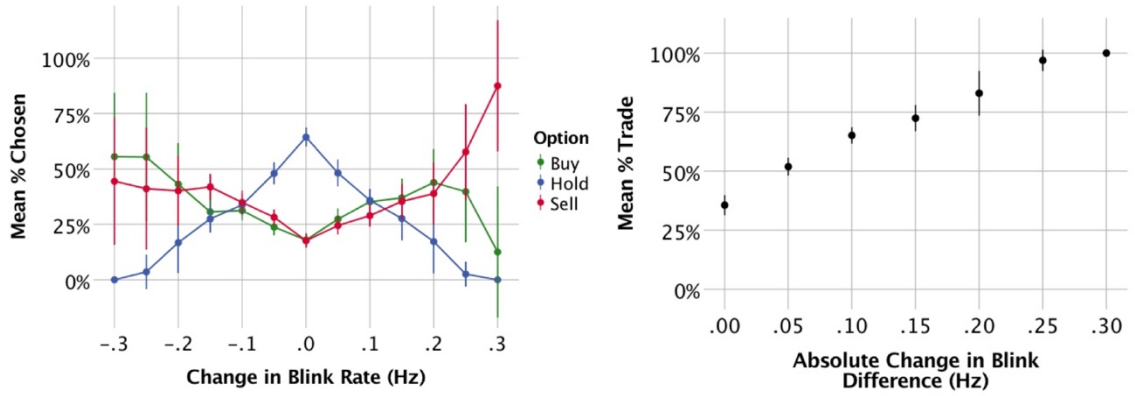


Fig 4: Option Choice. (A) The optimal option choice on any given trial varied depending on the change in blink rate. Each 0.05 Hz change in blink rate represented a price difference of one share (i.e. one buy/sell decision). At zero blink rate change, the optimal option was to 'hold' (blue). Rats chose the optimal hold option at zero blink change on nearly 70% of trials, which was significantly more likely than either the 'buy' (t=39.00, p<.001) or 'sell' options (t=38.56, p<.001). Negative changes in blink rate represented decreasing share prices, which indicated that the buy option (green) was the optimal choice. Alternatively, positive changes in blink rate represented increasing prices, which indicated that the sell (red) option was the optimal choice. Despite being suboptimal, rats continued to choose the safer hold option most frequently at smaller changes in blink rate until the change became larger than .15 Hz, at which point rats were more likely to buy at negative changes (t=8.99, p<.001) and sell at positive changes (t=6.95, p<.001). We did not find a significant difference in choice of the buy option

relative to the sell option at negative blink rates across all stocks (see SI for choices within individual stocks), *p*=NS. However, rats did demonstrate a significantly greater choice of the optimal sell option relative to the buy option at larger positive blink rates (t=2.11, p<.05). **(B)** Rats choice of either of the trade options compared to the hold option increased as the absolute change in blink rate increased. This indicates that although rats may not have learned to fully maximize the optimal 'buy' and 'sell' strategies on all trials as illustrated in Fig 4a, they were able to discriminate between blink rate changes and use this information to guide choice away from the perseverative center 'hold' option and toward one of the left or right nosepoke holes.

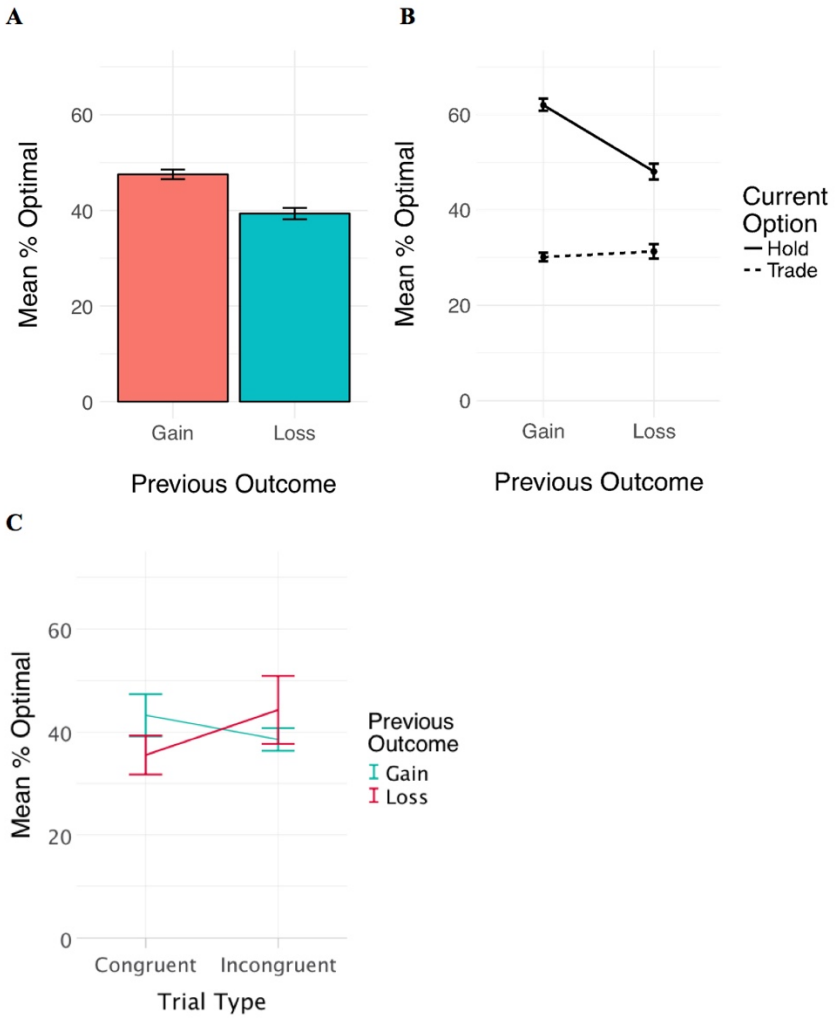


Fig 5: Effects of a previous loss on optimal choice. (A) Rats were 8.2% less likely to choose the optimal outcome on trials that were immediately preceded by a loss compared to a gain (paired-sample t-test, $t(23)$=-5.07, p<.001). **(B)** We found an interaction between previous outcome and current choice of the hold or trade options (t=8.93, p<.001). On trials that were immediately preceded by a loss, optimal performance was 13.1% lower when choosing the hold option ($t(23)$ = -7.56, p < .001). Conversely, a

previous loss led to a marginal non-significant increase (1.2%) in optimal choice of the trade options, $t(23)= 0.79$, $p > .05$. Error bars represent SEs.

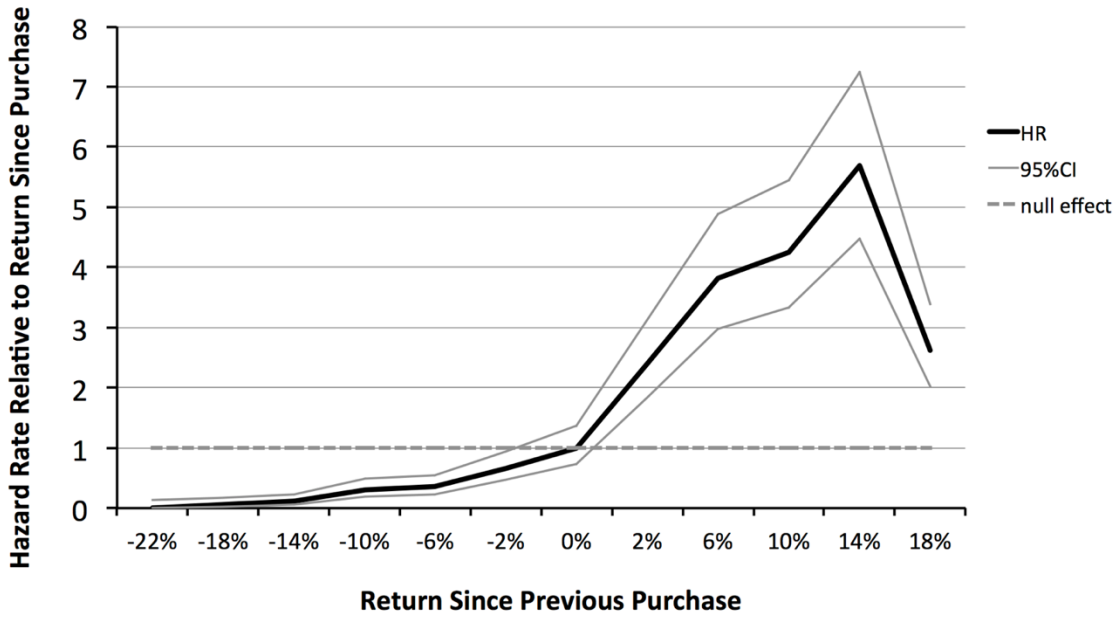


Fig 6: Disposition Effect. In this analysis, the hazard rate for each return bin, spanning 4% returns each, was calculated relative to zero return. A bin with a hazard ratio of 1 (dotted line) corresponds to a null effect on rats' choice of the sell option. Hazard ratios lying significantly above 1 indicate that the likelihood of selling was higher in that bin, while hazard ratios below 1 denote a reduced likelihood of selling in the given bin. Peaking at around 14% return, rats were up to 500% more likely to sell a stock when returns were positive. The opposite was true in the case of negative returns, whereby subjects demonstrated a decreasing disposition to realize losses as returns become more negative. Error bars (solid gray lines) represent 95%CI's.

Amsel, A. (1958). The role of frustrative nonreward in noncontinuous reward situations. *Psychological Bulletin*, *55*, 102-119.

Barber, B. M., & Odean, T. (2011). The behavior of individual investors. *SSRN*, 187221, Article 1872211. http://ssrn.com/abstract=187221

Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. In G. M. Constantinides, M. Harris, & R. M. Stultz (Eds.), *Handbook of Economics and Finance* (1st ed., Vol. 1, pp. 1053-1128). Elsevier.

Benartzi, S., & Thaler, R. (1995). Myopic loss aversion and the equity premium puzzle. *The Quarterly Journal of Economics*, *110*(1), 73-92.

Bhatti, M., Jang, H., Kralik, J. D., & Jeong, J. (2014). Rats exhibit reference-dependent choice behavior. *Behavioural brain research*, *267*, 26-32. https://doi.org/http://dx.doi.org/10.1016/j.bbr.2014.03.012

Bodenhausen, G. V., Sheppard, L. A., & Kramer, G. P. (1994). Negative affect and social judgment: The differential impact of anger and sadness. *European Journal of Social Psychology*, *24*(1), 45-62. https://doi.org/10.1002/ejsp.2420240104

Bondt, W. F. M. D., & Thaler, R. (1985). Does the stock market overreact? *The Journal of Finance*, *40*(3), 793-805. https://doi.org/10.2307/2327804

Bower, G. H. (1981). Mood and memory. *American Psychologist*, *36*(2), 129-148.

Boysen, S. T., & Berntson, G. G. (1995). Responses to quantity: perceptual versus cognitive mechanisms in chimpanzees (Pan troglodytes). *Journal of experimental psychology. Animal behavior processes*, *21*(1), 82-86. https://doi.org/10.1037//0097-7403.21.1.82

Capaldi, E. J., Verry, D. R., & Davidson, T. (1980). Memory, serial anticipation pattern learning, and transfer in rats. *Animal Learning & Behavior*, *8*, 575-585.

Carlson, S. M., Davis, A. C., & Leach, J. G. (2005). Less is more: Executive function and symbolic representation in preschool children. *Psychological Science*, *16*(8), 609-616. https://doi.org/10.1111/j.1467-9280.2005.01583.x

Charness, G., & Levin, D. (2005). When optimal choices feel wrong: A laboratory study of bayesian updating, complexity, and affect. *The American economic review*, *95*(4), 1300-1309. https://doi.org/10.2307/4132717

Chen, G., Kim, K., Nofsinger, J. R., & Rui, O. M. (2007). Trading performance, disposition effect, overconfidence, representativeness bias, and experience of emerging market investors. *Journal of Behavioral Decision Making*, *20*(4), 425-451. SSRN: http://ssrn.com/abstract=957504

Chen, M. K., Lakshminarayanan, V., & Santos, L. R. (2006). How basic are behavioral biases? Evidence from capuchin monkey trading behavior. *Journal of Political Economy*, *114*(3), 517-537.

Choi, J. J., Laibson, D., Madrian, B. C., & Metrick, A. (2009). Reinforcement learning and savings behavior. *The Journal of Finance*, *64*(6), 2515-2534. https://doi.org/10.1111/j.1540-6261.2009.01509.x

Christ, S. E., White, D. A., Mandernach, T., & Keys, B. A. (2001). Inhibitory Control Across the Life Span. *Developmental Neuropsychology*, *20*(3), 653-669. https://doi.org/10.1207/S15326942DN2003_7

Constantinople, C. M., Piet, A. T., Bibawi, P., Akrami, A., Kopec, C., & Brody, C. D. (2019). Lateral orbitofrontal cortex promotes trial-by-trial learning of risky, but not spatial, biases. *elife*, *8*, e49744. https://doi.org/10.7554/eLife.49744

Constantinople, C. M., Piet, A. T., & Brody, C. D. (2019). An Analysis of Decision under Risk in Rats. *Current Biology*, *29*(12), 2066-2074.e2065. https://doi.org/https://doi.org/10.1016/j.cub.2019.05.013

Cox, D. R., & Oakes, D. (1984). *Analysis of Survival Data*. Chapman and Hall.

Croissant, Y. (2013). *mlogit: multinomial logit model*. In (Version R package version 0.2-4) https://CRAN.R-project.org/package=mlogit

Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor Psychology and Security Market Under- and Overreactions. *The Journal of Finance*, *53*(6), 1839-1885. https://doi.org/10.1111/0022-1082.00077

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control [10.1038/nn1560]. *Nat Neurosci*, *8*(12), 1704-1711. https://doi.org/http://www.nature.com/neuro/journal/v8/n12/suppinfo/nn1560_S1.html

Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, *88*, 848-881.

Fama, E. F. (1965). Random walks in stock market prices. *Financial Analysts Journal*, *21*(5), 55-59.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25*(2), 383-417. https://doi.org/10.1111/j.1540-6261.1970.tb00518.x

Feng, L., & Seasholes, M. (2005). Do investor sophistication and trading experience eliminate behavioral biases in financial markets? *Review of Finance*, *9*, 305-351.

Ferland, J.-M. N., Adams, W. K., Murch, S., Wei, L., Clark, L., & Winstanley, C. A. (2018). Investigating the influence of 'losses disguised as wins' on decision making and motivation in rats. *Behavioural pharmacology*, *29*(8). https://journals.lww.com/behaviouralpharm/Fulltext/2018/12000/Investigating_the_influence_of__losses_disguised.10.aspx

Forgas, J. P. (1989). Mood effects on decision making strategies. *Australian Journal of Psychology*, *41*(2), 197-214. https://doi.org/10.1080/00049538908260083

Fountain, S. B., & Rowan, J. D. (2007). Differential impairments of rat serial-pattern learning and retention induced by MK-801, an NMDA receptor antagonist. *Psychobiology*, *28*, 32-44.

Frydman, C., Barberis, N., Camerer, C., Bossaerts, P., & Rangel, A. (2014). Using neural data to test a theory of investor behavior: An application to realization utility. *The Journal of Finance*, *69*(2), 907-946.

Frydman, C., & Camerer, C. F. (2016). The Psychology and Neuroscience of Financial Decision Making. *Trends in cognitive sciences*, *20*(9), 661-675. https://doi.org/https://doi.org/10.1016/j.tics.2016.07.003

Fuster, A., Laibson, D., & Mendel, B. (2010). Natural expectations and macroeconomic fluctuations. *Journal of Economic Perspectives*, *24*(4), 67-84. https://doi.org/10.1257/jep.24.4.67

Garlick, D., Fountain, S. B., & Blaisdell, A. P. (2017). Serial pattern learning in pigens: Rule-based or associative. *Journal of Experimental Psychology: Animal Learning and Cognition*, *43*(1), 30-47.

Grinblatt, M., & Keloharju, M. (2001). What makes investors trade? *The Journal of Finance*, *56*(2), 589-616. https://doi.org/10.2307/222575

Grinblatt, M., Keloharju, M., & Linnainmaa, J. T. (2012). IQ, trading behavior, and performance. *Journal of Financial Economics*, *104*(2), 339-362. https://doi.org/http://dx.doi.org/10.1016/j.jfineco.2011.05.016

Gutiérrez-Roig, M., Segura, C., Duch, J., & Perelló, J. (2016). Market imitation and win-stay lose-shift strategies emerge as unintended patterns in market direction guesses. *Plos One*, *11*(8), e0159078. https://doi.org/10.1371/journal.pone.0159078

Hershberger, W. (1986). An approach through the looking-glass. *Animal Learning & Behavior*, *14*(4), 443-451. https://doi.org/10.3758/BF03200092

Hirshleifer, D. (2001). Investor Psychology and Asset Pricing. *The Journal of Finance*, *56*(4), 1533-1597. https://doi.org/10.1111/0022-1082.00379

Hulse, S. H. (Ed.). (1973). *The psychology of learning and motivation: Advances in research and theory* (Vol. 7). Academic Press.

Johnson, E. J., & Tversky, A. (1983). Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology*, *45*(1), 20-31.

Kagel, J. H., Battalio, R. C., Rachlin, H., & Green, L. (1981). Demand curves for animal consumers. *The Quarterly Journal of Economics*, *96*(1), 1-15. https://doi.org/10.2307/2936137

Kagel, J. H., Battalio, R. C., Rachlin, H., Green, L., Basmann, R. L., & Klemm, W. R. (1975). Experimental studies of consumer demand behavior using laboratory animals. *Economic Inquiry*, *13*(1), 22-38. https://doi.org/10.1111/j.1465-7295.1975.tb01101.x

Kagel, J. H., MacDonald, D. M., Battalio, R. C., White, S., & Green, L. (1986). Risk aversion in rats (Rattus norvegicus) under varying levels of resource availability. *Journal of Comparative Psychology*, *100*(2), 95.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263-291.

Kalenscher, T., & van Wingerden, M. (2011). Why We Should Use Animals to Study Economic Decision Making – A Perspective. *Frontiers in Neuroscience*, *5*, 82. https://doi.org/10.3389/fnins.2011.00082

Kaustia, M. (2010). Prospect Theory and the disposition effect. *Journal of Financial and Quantitative Analysis*, *45*(3), 791-812. https://doi.org/doi:10.1017/S0022109010000244

Kaustia, M., & Knüpfer, S. (2008). Do investors overweight personal experience? Evidence from IPO subscriptions. *Jounral of Finance*, *63*, 2679-2702.

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. https://doi.org/10.3819/ccbr.2015.100003

Koot, S., Adriani, W., Saso, L., van den Bos, R., & Laviola, G. (2009). Home cage testing of delay discounting in rats. *Behavior research methods*, *41*(4), 1169-1176. https://doi.org/10.3758/BRM.41.4.1169

Kuhnen, C. M., & Knutson, B. (2005). The neural basis of financial risk taking. *Neuron*, *47*(5), 763-770.

Lakshminarayanan, V., Chen, M. K., & Santos, L. R. (2011). The evolution of decision-making under risk:

Framing effects in monkey risk preferences. *Journal of Experimental Social Psychology*, *47*, 689-693.

Lakshminaryanan, V., Chen, M. K., & Santos, L. R. (2008). Endowment effect in capuchin monkeys. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *363*(1511), 3837-3844. https://doi.org/10.1098/rstb.2008.0149

Lawson, R., & Marx, M. H. (1958). A comparison of some presumed frustrating and secondary-reinforcing operations. *Journal of Comparative and Physiological Psychology*, *51*(6), 742-746.

Li, Y. J., Kenrick, D. T., Griskevicius, V., & Neuberg, S. L. (2012). Economic decision biases and fundamental motivations: How mating and self-protection alter loss aversion. *Journal of Personality and Social Psychology*, *102*(3), 550-561. https://doi.org/10.1037/a0025844

Malkiel, B. G. (1973). *A Random Walk Down Wall Street*. Norton. https://books.google.co.uk/books?id=oZWNMQEACAAJ

Markowitz, H. M. (1959). *Portfolio Selection: Efficient Diversification of Investments*. Jon Wiley & Sons.

Marshall, A. T., & Kirkpatrick, K. (2015). Relative Gains, Losses, and Reference Points in Probabilistic Choice in Rats. *Plos One*, *10*(2), e0117697. https://doi.org/10.1371/journal.pone.0117697

Marshall, A. T., & Kirkpatrick, K. (2017). Reinforcement learning models of risky choice and the promotion of risk-taking by losses disguised as wins in rats. *Journal of Experimental Psychology: Animal Learning and Cognition*, *43*(3), 262-279. https://doi.org/10.1037/xan0000141

Mendl, M., Burman, O. H. P., Parker, R. M. A., & Paul, E. S. (2009). Cognitive bias as an indicator of animal emotion and welfare: Emerging evidence and underlying mechanisms. *Applied Animal Behaviour Science*, *118*(3), 161-181.

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

Muller, M. D., & Fountain, S. B. (2010). Concurrent cognitive processes in rat serial pattern learning: Item memory, serial position, and pattern structure. *Learn. Motiv.*, *41*(4), 252-272.

Muller, M. D., & Fountain, S. B. (2016). Concurrent cognitive processes in rat serial pattern learning: II. Discrimination learning, rule learning, chunk length, and multiple-item memories. *Journal of Experimental Analysis of Behavior*, *105*(1), 155-175.

Odean, T. (1998). Are investors reluctant to realize their losses? *Journal of Finance*, *53*, 1775-1798.

Pais-Vieira, M., Lima, D., & Galhardo, V. (2007). Orbitofrontal cortex lesions disrupt risk assessment in a novel serial decision-making task for rats. *Neuroscience*, *145*(1), 225-231.

Payzan-LeNestour, E., & Bossaerts, P. (2015). Learning about unstable, publicly unobservable payoffs. *The Review of Financial Studies*, *28*(7), 1874-1913.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & Team, R. C. (2017). *nlme: Linear and Nonlinear Mixed Effects Models*. In (Version R package version 3.1-131) https://CRAN.R-project.org/package=nlme

Poterba, J. M., & Summers, L. H. (1988). Mean reversion in stock prices: Evidence and Implications. *Journal of Financial Economics*, *22*(1), 27-59. https://doi.org/http://dx.doi.org/10.1016/0304-405X(88)90021-9

Rivalan, M., Ahmed, S. H., & Dellu-Hagedorn, F. (2009). Risk-Prone Individuals Prefer the Wrong Options on a Rat Version of the Iowa Gambling Task. *Biological Psychiatry*, *66*(8), 743-749. https://doi.org/10.1016/j.biopsych.2009.04.008

Rivalan, M., Valton, V., Seriès, P., Marchand, A. R., & Dellu-Hagedorn, F. (2013). Elucidating Poor Decision-Making in a Rat Gambling Task. *Plos One*, *8*(12), e82052. https://doi.org/10.1371/journal.pone.0082052

Samuelson, P., A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, *6*(2).

Santos, L. R., & Platt, M. L. (2014). Evolutionary anthropological insights into neuroeconomics: What non-human primates can tell us about human decision-making strategies. In E. a. P. W. G. Fehr (Ed.), *Neuroeconomics (Second Edition)* (pp. 109-122). Academic Press. https://doi.org/http://dx.doi.org/10.1016/B978-0-12-416008-8.00007-3

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*,

275(5306), 1593-1599. https://doi.org/10.1126/science.275.5306.1593

Shapira, Z., & Venezia, I. (2001). Patterns of behavior of professionally managed and independent investors. *Journal of Banking & Finance*, *25*(8), 1573 - 1587.

Shefrin, H., & Statman, M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of Finance*, *40*(3), 777-790. https://doi.org/10.1111/j.1540-6261.1985.tb05002.x

Shiller, R. J. (2000). Measuring Bubble Expectations and Investor Confidence. *Journal of Psychology and Financial Markets*, *1*(1), 49-60. https://doi.org/10.1207/S15327760JPFM0101_05

Shiller, R. J. (2003). From efficient markets theory to behavioral finance. *The Journal of Economic Perspectives*, *17*(1), 83-104. https://doi.org/10.2307/3216841

Shleifer, A., & Summers, L. H. (1990). The Noise Trader Approach to Finance. *The Journal of Economic Perspectives*, *4*(2), 19-33. http://www.jstor.org/stable/1942888

Shoda, Y., Mischel, W., & Peake, P. K. (1990). Predicting adolescent cognitive and self-regulatory competencies from preschool delay of gratification: identifying diagnostic conditions. *Dev Psychol*, *26*(6), 978-986.

Shumway, T., & Wu, G. (2005). *Does the disposition effect drive momentum?* AFA 2006 Boston Meeting,

Simon, N. W., Gilbert, R. J., Mayse, J. D., Bizon, J. L., & Setlow, B. (2009). Balancing risk and reward: a rat model of risky decision making. *Neuropsychopharmacology: Official Publication Of The American College Of Neuropsychopharmacology*, *34*(10), 2208-2217. http://search.ebscohost.com/login.aspx?direct=true&db=cmedm&AN=19440192&site=eds-live

St Onge, J. R., & Floresco, S. B. (2009). Dopaminergic modulation of risk-based decision making. *Neuropsychopharmacology*, *34*(3), 681-697. https://doi.org/10.1038/npp.2008.121

Steiner, A. P., & Redish, A. D. (2014). Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task. *Nat Neurosci*, *17*, 995-1002.

Stephens, D. W., & Krebs, J. R. (1986). *Foraging Theory*. Princeton University Press.

Strahilevitz, M. A., Odean, T., & Barber, B. M. (2011). Once burned, twice shy: How naive learning, counterfactuals, and regret affect the repurchase of stocks previously sold. *Journal of Marketing Research*, *48*(SPL), S102-S120. https://doi.org/doi:10.1509/jmkr.48.SPL.S102

Sutton, R., & Barto, A. (1998). *Reinforcement Learning*. MIT Press.

Thaler, R. H., Tversky, A., Kahneman, D., & Schwartz, A. (1997). The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test. *The Quarterly Journal of Economics*, *112*(2), 647-661. http://www.jstor.org/stable/2951249

Therneau, T. M. (2014). *Survival: A package for survival analysis in R. R package version 2.37-7*. In http://CRAN.R-project.org/package=survival

Thorndike, R. (1911). *Animal intelligence: Experiental studies*. McMillan.

Tiedens, L. Z., & Linton, S. (2001). Judgment under emotional certainty and uncertainty: the effects of specific emotions on information processing. *Journal of Personality and Social Psychology*, *81*(6), 973-988.

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*(4), 1184-1195. https://doi.org/10.3758/s13415-014-0271-1

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*(211), 453-458.

van Wingerden, M., Marx, C., & Kalenscher, T. (2015). Budget constraints affect male rats' choices between differently priced commodities. *Plos One*, *10*(6), e0129581.

Whiten, A. (Ed.). (1991). *Natural theories of mind: Evolution, development and simulation of everyday mindreading*. Basil Blackwell.

Wilson, D. I., Laidlaw, A., Butler, E., Hofmann, D., & Bowman, E. M. (2006). Development of a behavioral task measuring reward "wanting" and "liking" in rats. *Physiology & Behavior*, *87*(1), 154-161. https://doi.org/10.1016/j.physbeh.2005.09.016

Xiong, W. (2013). Bubbles, Crises, and Heterogeneous Beliefs. In J.-P. Fouque & J. A. Langsam (Eds.), *Handbook on Systemic Risk* (pp. 663-713). Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9781139151184.033

Zalocusky, K. A., Ramakrishnan, C., Lerner, T. N., Davidson, T. J., Knutson, B., & Deisseroth, K. (2016).

Nucleus accumbens D2R cells signal prior outcomes and control risky decision-making. *Nature*, *531*(7596), 642-646. https://doi.org/10.1038/nature17400

Zeeb, F. D., Robbins, T. W., & Winstanley, C. A. (2009). Serotonergic and dopaminergic modulation of gambling behavior as assessed using a novel rat gambling task. *Neuropsychopharmacology (New York, N.Y.)*, *34*(10), 2329-2343.