Development and employment of discounting tasks in 16p11.2 hemideletion mice

A Dissertation

SUBMITTED TO THE FACULTY OF THE

UNIVERSITY OF MINNESOTA

BY

Gerardo Raul Rojas

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Nicola M. Grissom

December 2023

Acknowledgements

I want to thank my mentor Nicola Grissom for providing me with excellent mentorship and for providing an environment where I was comfortable in a personal and professional capacity. Nicola taught me to think about the diversity of behavior and to think critically about how to interpret those behaviors analytically. Nicola has been patient and empathetic during critical moments in my career and helped push me along toward the finish line. I am truly happy and grateful to have had an advisor, who I can state with the utmost confidence, was invested in all aspects graduate education and in my personal success.

I want to thank my committee members Iris Donga Vilares, Vanessa Lee and Jocelyn Richard who all came together and helped me complete my degree. Additionally, I want to thank Ben Hayden and Sarah Heilbronner for providing important feedback during my preliminary exam. I appreciate everyone's effort and investment in my success.

I want to thank everyone in the Grissom lab including Cathy, Evan, Dana, Madison, Erin, Angelica, Nic, and Solo for sating my curiosity with excellent discussions and for the constant exploration of diverse research. Thank you everyone for being my colleagues and friends.

Last but not least, I want to thank my family for being supportive in all aspects of my life and celebrating important graduation milestones with me. Thank you for giving me the confidence to push forward through tough times and for rooting for me in many ways, including driving to conferences to spend time with me. I appreciate you all.

Abstract

Humans must wait or make risky decisions to obtain different commodities of varying value. These decisions rely on learning the value of those commodities and adapting choices to maximize rewards while minimizing costs. As costs increase, humans discount the value of large costly rewards when a small, less costly option presents itself. The assessment of these cost-benefit analyses shifts according to individual factors such as sex and neurodivergence. It is then imperative to understand how different costs such as delay or probabilistic risk affect those choices. This dissertation has the goals of explaining how factors such as learning shape cost-associated decisions and how these decisions change according to individual factors. The first study focuses on tracking the development of discounting behavior as a function of delay and probability. Comparing the development of these behaviors across cost types reveal specific adaptations to different cost types in both sexes. Females were more sensitive to order effects than males. Overall learning of these tasks was well tracked by choice variability, even when discounting value estimations were unstable. The second study reveals how a copy number variant associated with neurodevelopmental disorders affects sensitivity to delay and probability in a sex-biased manner. Delay induced greater large reward preference in male carriers compared to noncarriers, but probabilistic risk instead induced small reward preference in male carriers. Male carriers in particular use environmental cues more than noncarriers to control behavior when rewards are delayed, but not when rewards are risky. These results highlight how this copy number variant affects choice according to uncertainty. Taken together, these studies reveal how sex and sources of neurodiversity contribute to decision-making.

Table of Contents

List of Figures

1. Introduction

Humans are constantly presented with choices accompanied by different types of costs. It may be excruciatingly tempting to pull up to the nearest McDonald's after a long day of work and immediately eat, however we often make the trade-off of waiting in order to obtain a better, more nutritious meal at home. Additionally, said person may have forgotten to do grocery shopping, so there is a risk that they do not have sufficient ingredients to make a satisfying meal. If the cost is too high on either occasion, that person will choose the immediate and guaranteed McDonald's meal. The propensity to weigh temporal and probabilistic costs and prefer guaranteed small rewards as costs increase is known as discounting. Humans and animals alike exhibit choices that favor costly large rewards until the price exceeds an individuals' willingness to endure costs (Green & Myerson, 2004). The source of discounting is hotly contested, with some suggesting discounting arises due to a change in value of the reward as a function of the cost (Green et al., 2014; Mazur, 1987; Odum, 2011a; Richards et al., 1999), others suggesting a role for reward maximization in incentivizing small reward preference (Blanchard et al., 2013; Daw & Touretzky, 2000; Stephens et al., 2004), and others emphasizing how recent reward history shifts choice preference to options that increase reward rate in short time spans (Namboodiri et al., 2014).

These decisions are studied in the lab as binary choice tasks that ask participants to pick between smaller sooner and larger later rewards in delay discounting or between smaller safe and larger risky rewards in probability discounting. Humans are willing to wait for rewards, but waiting too long incurs an intolerance to the cost of time. You may be willing to forgo $10 now for $20 in a week, but if the choice changed to $10 now and $20 in two weeks, you would be less willing to wait. Similarly, you might be willing to take a 75% gamble for the $20, but less willing to

take a 25% gamble. When a person consistently faces these choices, they form an idea of their delay and risk tolerance. If instead, a person is asked to choose between $10 in ten weeks and $20 in twelve weeks a person would be more willing to wait than the other previously mentioned two week difference question. By reframing the time difference, the large reward option became more salient of a choice. In another case, what if a person learns that in a time limited session, they can earn more money by constantly picking the small immediate reward? In that case, they learn to maximize the amount of reward they earn by quickly and repeatedly selecting the low cost option. Depending on the perspective, someone could interpret that behavior as "optimal" — they acted in a way to earn the most rewards. Another person might interpret that behavior as "impulsive" — they were unable to resist the temptation of immediacy to earn larger quantities of reward.

Early discounting research adopted this notion of self-control in order to make links to different neuropsychiatric disorders. For instance, a drug addict makes the decision to gain the immediate relief of withdrawal symptoms by taking drugs now and forgoes long-term health benefits of not taking drugs. People with attention-deficit hyperactivity disorder are distracted by a text message in class now and forgo the knowledge they could have received from ignoring the text. It is these instances of self-control that researchers attempt to model using discounting tasks (Green & Myerson, 2013; Strickland & Johnson, 2021). This may be the case historically, but as mentioned above behaviors that look one way can be motivated by an entirely different factor. Discounting behavior is linked to different influences, and parsing between them has become important in research.

Animal models have been critical in exploring the different ways in which discounting occurs. The discounting phenomenon occurs across species in the same task, implying some common decision-making process is being captured (Heilbronner,

2017; Vanderveldt et al., 2016). Animal studies have been essential in elucidating different factors that shift discounting preferences in the presence of costs. However, a mouse version of delay and probability discounting has yet to be thoroughly explored. Mice are especially valuable subjects because of the availability of genetic tools, but two barriers have muddied the assessment of discounting behavior: 1) common measures of discounting limit the interpretation of how discounting occurs and 2) alternate explanations for discounting, such as differences in adaptations to tasks or attentional adjustments, are commonly ignored. These are significant sources of variability to consider when assessing how a genotype modeling a disorder affects acquisition and mastery of discounting behavior or rates.

This report aims to tackle these problems by providing parallels in the discounting literature between humans and animals. I will introduce common frameworks to understand decision-making across species. Throughout this first section I will highlight: 1) theoretical explanations for discounting behavior in humans and animals, 2) mathematical approaches developed to describe discounting and 3) how discounting behavior is related to decision-making in neuropsychiatric disorders. Next, I will provide evidence for the importance of a mouse version of discounting tasks and how that can be utilized to maximize the advantage of the wide array of genetic tools available for mice. In order to demonstrate this, I will explain: 1) current discounting tasks employed across species, 2) how certain disorder-related genotypes may shift decision-making and 3) why we want to study the effect of the copy number variant 16p11.2 hemideletion on decision-making. Lastly, I will present my experiments explaining the development of a mouse version of discounting and how 16p11.2 hemideletion mice adapt to different types of costs.

1.1 Delay and Probability Discounting

Humans and animals have evolutionarily learned to forage for food. In these situations, animals weigh the cost of time investment and likelihood of food being in a known patch against the potential time investment and risk of food being available in a new, unexplored patch (Hayden, 2018; Pyke & Stephens, 2019). Quicker food acquisition is preferred in order to minimize environmental uncertainty (e.g., predator risk). There may be some shared uncertainty in the availability of food and in the delay to receive food, implying probabilistic and delayed rewards share some common sources of processing (Hayden & Platt, 2007; Heilbronner et al., 2010; Story et al., 2023). Research that directly compares the tasks by examining fMRI activity confirms there is overlapping activity in general reward processing areas like ventral striatum, but differential activity in other parts of the brain (Peters & Büchel, 2009). There are benefits to looking at both types of discounting, especially when trying to deconstruct the different contributing factors and strategies that contribute toward different choices (Green & Myerson, 2013; Strickland & Johnson, 2021). Today, humans still rely on assessments of known versus unknown costs in order to guide decision-making.

Discounting tasks as a psychological concept originate from economic observations of human behavior. Utility functions were produced in order to determine how temporal costs change the subjective value of goods (Grüne-Yanoff, 2015). A couple of simple observations were made: 1) humans prefer immediate rewards compared to delayed rewards and 2) people will prefer large delayed rewards over small delayed rewards when both are sufficiently delayed, but will switch to choose the small reward more often when the delays become shorter. An exponential model was formed which accounted for preference of immediate gratification (Grüne-Yanoff, 2015). The model makes the assumption that the rate of discounting is constant; preferences do not switch over time (Green & Myerson, 2004). The exponential model fails to cover point two above: that humans will prefer the small reward when the large

reward is sufficiently delayed (i.e., preference reversals). A landmark paper incorporating animal and human research demonstrated that both observations mentioned above were better captured by a hyperbolic model of temporal discounting (Ainslie, 1975; Hayden, 2016). This model accounted for the seemingly irrational behavior that people switch to preferring a smaller sooner reward compared to a larger later reward as the delay decreases. Ainslie noted that reversals caused by the temptation of immediacy are related to a lack of self-control and surmised that this can lead to impulsive choices. Researchers then were interested in applying exponential and hyperbolic models to other types of discounting in order to determine if the different discounting rates followed similar predictions. Probability discounting was found to be better described by a hyperbolic model as well (Rachlin et al., 1991). Although, some have argued for alternative formulations on the basis that delay and probability discounting represent different types of costs with their own sensitivities (Green & Myerson, 2004; Killeen, 2023b). Other discounting tasks have also been attempted to be modeled in either similar models or in a unified model (Białaszek et al., 2019).

Other types of discounting tasks are important as well, such as effort discounting and social discounting. Effort discounting is the case where the amount of work (i.e., cognitive or physical) needed to obtain rewards is manipulated (Botvinick et al., 2009; Westbrook et al., 2013). Social discounting tests the willingness of a person to pass up on a large reward and give it to someone else, costs increase by manipulating social distance to the giver (Rachlin & Jones, 2008). Some efforts have been made to study social costs in animals, but it is not developed enough, so I will focus on effort discounting here. Effort discounting has some parallels to delay discounting, take for instance the patch foraging paradigm mentioned above. Part of the delay to get to the next patch requires some amount of effort expended on the individual to traverse the

landscape (Białaszek et al., 2019). Effort-based tasks need time based yoked controls to dissociate the effects from delay to reward (Mitchell, 2017). Effort discounting, while useful, didn't suit our purposes because we wanted to study parallel tasks whose parameters are more readily comparable. Additional concerns need to be addressed in order to properly control for time-related influences on effort-discounting. Further, we wanted to study discounting tasks that probe fundamental questions about learning and decision-making in neurodevelopmental disorders in order to challenge mouse models. Incorporation of uncertainty into the processing of rewards is thought to be altered in neurodevelopmental disorders (Sinha et al., 2014). This makes delay and probability discounting tasks comparable based on their possibly shared, but dissociable uncertainty (Peters & Büchel, 2009; Prévost et al., 2010). A common currency theory of cost-benefit decision-making predicts effort discounting accesses a similar domain-general network in the brain to assess rewards as a function of cost in a manner similar to delay and probability discounting (Levy & Glimcher, 2012). The nucleus accumbens is among those reward processing areas that have been implicated in committing to effortful decisions in humans (Botvinick et al., 2009; Schouppe et al., 2014) and animals (Ghods-Sharifi & Floresco, 2010). Although, other studies find no involvement of the ventral striatum generally (Prévost et al., 2010). The lack of striatal involvement may be due to subregion differences in the encoding of cost and reward, resulting in the change being masked (Suzuki et al., 2021). Still, it is evident that more studies need to be conducted on effort discounting especially because of the broad heterogeneity of effortful tasks (e.g., timing difficulty, action requirements, and experience effects). There have however been recent efforts to translate this task in mouse touchscreen chambers (Lopez-Cruz et al., 2023). Next, I will focus on delay and probability discounting and explain how they are studied in the laboratory.

1.2 Human and animal models of discounting

Human Tasks Used to Measure Discounting

Human discounting tasks largely employ surveys to study choice behavior (e.g., "Would you rather have $1 now or $1000 after a month?"). Researchers gain a wide distribution of responses by increasing or decreasing the delay to rewards and observing when a conflict in preference occurs (i.e., the indifference point). Researchers systematically often specifically change the delay of the large reward in order to gain a curve of multiple indifference points, which are thought to reflect the discounted value of the reward (Odum, 2011a; Rachlin et al., 1991). An alternative method is to instead adjust the amount of reward to create choice conflict (Mazur, 1987). Extensive analysis suggests indifference points obtained through different methods are similar to each other (Holt et al., 2012). The slope of said curve is *k* or the discounting rate. A similar process is used to measure probability discounting by replacing delay with probabilistic risk (Rachlin et al., 1991). Recent advances in human discounting tasks have cleverly come up with other ways to generate discounting rates without the use of indifference points. The Kirby Money Questionnaire is a well-accepted alternative method to derive discounting curves by using different ratios of small and large reward amounts at different delays to find the *k* value where a switch in preference occurs (Kirby, 2009; Kirby et al., 1999). A similar questionnaire has also been developed to look at *h* decay rates in probability discounting (Gray et al., 2016; Madden et al., 2009).

Questionnaires are useful, but there are some important methodological implications to consider. Questionnaires: 1) are usually hypothetical (i.e., the participant does not experience the delay with their choice, nor do they experience the receipt of reward) and 2) usually use money as potential rewards. One early study compared temporal discounting rates in select hypothetical and real-reward discounting experiments and presented evidence that discounting rates were steeper in

real-reward tasks (Kirby, 1997). This prompted others to similarly explore the experience of reward within discounting tasks, where some found similar discounting rates between hypothetical and real-reward delay discounting rates (Dixon et al., 2013; Madden et al., 2003). Madden et al. (2003) had posited this could have been due to learning effects caused by the within-subject nature of their design (i.e., doing both the hypothetical and real-reward versions in the same subject). A later study confirmed order effects for within-subject hypothetical and real-reward tasks do influence discounting steepness (Hinvest & Anderson, 2010). Further, Madden et al. (2003) used potential rewards rather than guaranteed rewards, which added a layer of uncertainty and possibly serving as another source of confounds. Researchers were additionally motivated to examine if the experience of costs is important to the discounting effect, or if hypothetical costs were sufficient. One study manipulated post-reward delays in order to examine the effect of trial length on experiential delay discounting compared to hypothetical delay (Dixon et al., 2013). The authors noticed increased steepness only when they didn't control trial length, possibly indicative of promoting a rate-maximizing strategy or alternatively boredom (Smits et al., 2013). Another group systematically looked at the role of experiential delay in preference formulation and found that as long as delays are experienced, discounting will occur whether a reward is real or hypothetical (Steele et al., 2019). These results mean that in order to adequately gain an idea of the cost-opportunity tradeoff, it is fundamental to experience the delay. When both cost and reward are hypothetical, it is feasible that a more trait-like, long-term decision-making process (Green et al., 1994; Koffarnus et al., 2013; Odum, 2011b). Experiential tasks have increasingly become sought after because they are thought to access state-like decision-making that better reflects choices in response to environmental challenges. Further, experiential tasks might be better at probing certain sensitivities in neuropsychiatric disorders (Horan et al., 2017). The experience of the reward and its properties are also important to understand.

An additional common concern regarding these tasks is the nature of money as a reinforcer. Monetary rewards pose an interesting problem in that they allow for the experience of the receipt of rewards, but not the consummatory behaviors related to something like food rewards (Berridge & Robinson, 2016). The finding that humans discount food rewards more steeply than monetary rewards is robust, at least for delay discounting (Estle et al., 2007; Odum et al., 2006, 2020). Discounting tasks are used to detect maladaptive shifts in choice, but sometimes those effects are outcome specific, such as with drugs (Odum et al., 2020). Money is a secondary reinforcer that has ambiguous value and payoff, which makes it more flexible as a reinforcer and through associativity acts on the same parts of the brain as primary reinforcers (McClure et al., 2007). Studies that have looked at other types of non-food commodities find similar behavioral results, although correlations do exist between food and non-food rewards (Charlton & Fantino, 2008). These results do not diminish money as a motivating outcome but do point out the nuance needed when interpreting those results without other outcome types. Using food rewards have the added benefit of enabling more direct cross-species comparisons with animals.

Animal Tasks Used to Measure Discounting

All animal discounting tasks are experiential. Animals will always experience costs associated with primary rewards like sugar or drugs. Many labs use operant chambers and present choices that can be made through nosepokes or lever-presses (Mar & Robbins, 2007). These tasks resemble human experiential tasks: one option leads to a small amount of reward that is immediate and guaranteed, and the other option results in a large amount of reward that is delayed or risky. Standard procedures suggest training animals until “stable” discounting is achieved. Stability is defined in different ways, but most often refers to a lack of influence of training on choice

behavior following weeks of training. Discounting rates are then attained by systematically incrementing reward sizes or costs.

Delay and probability discounting tasks in animals resemble human discounting, where the amount of reward or delay is manipulated in order to generate indifference points. Similar to humans, indifference points generated from adjusted-delay and adjusted-reward methods are comparable to each other in animals (Green et al., 2007). Within the rodent literature on discounting, fixed procedures are popular. In this paradigm, animals are presented with a small guaranteed and a costly large reward that cycles through blocks of trials (Evenden & Ryan, 1996). Instead of calculating indifference points by adjusting costs or rewards, discounted value is estimated by measuring the proportion of large choices made throughout discounting blocks. Methods were developed so that the costly option either began with no cost and increased throughout a session or began with a large cost and decreased throughout a session. Researchers compared adjusting-delay and increasing-delay methods and found that these methods produced similar indifference points (Craig et al., 2014). Another benefit to these fixed cost methods is that they provide a cost-free period for the large delayed option. This provides a built-in comparison of the magnitude difference of reinforcers, a comparison of how different manipulations (e.g., chemogenetics or drugs) affect rewards free of costs, and a comparison to other blocks to determine where preference starts to deviate from a cost-free period. Within-subject analyses are also empowered by such designs and can reveal learning effects in models of neurodevelopmental disorders (E. Sjoberg et al., 2023). Cost order by dopaminergic drug interaction effects are consistent for different types of costs (Krebs & Anderson, 2012; Maguire et al., 2014; St Onge & Floresco, 2009; Tanno et al., 2014). Dopaminergic drugs seem to increase the stickiness or perseveration in choice, depending on the initial delay or cost experienced (St Onge et al., 2010). The effect is

consistent in male and female rodents in probability discounting, although under baseline conditions female rats are more sensitive to risk than males (Islas-Preciado et al., 2020; see Orsini & Setlow, 2017). The opposite pattern was observed in delay discounting, with sex differences emerging after the application of drugs (Eubig et al., 2014). While cost presentation is important, outcome type can influence discounting as well.

Similar to humans, reinforcer types matter. Common rewards in rodent paradigms include pellets (e.g., grain or sucrose) and liquid (e.g., sucrose or milk) reinforcers. One study looked at commonly studied strains of mice and demonstrated greater maintenance of high levels of responding for milk than pellets in a simple fixed-ratio design (Hutsell & Newland, 2013). Indeed, there is an existing literature to support differences in satiety between liquid and solid food that could influence motivation to perform tasks (Almiron-Roig et al., 2003; Stribiţcaia et al., 2020). Reinforcer types are worth considering when conducting discounting tasks, even in animals, although it seems to have less of an effect on discounting rates than in humans (Calvert et al., 2010). Another aspect to consider is if pursuit of these reinforcers match naturalistic settings. Different environmental pressures such as reward type and availability can influence choice sampling and thus warrant exploration in settings like foraging tasks.

Foraging tasks have been explored in parallel to discounting tasks in order to measure temporal and risky costs in a more naturalistic paradigm. In sequential designs, animals have to balance quite a few costs. Exploiting one patch diminishes the possible prospects of accumulating rewards. Animals can explore another patch, but they incur the cost of travel time to the next patch and uncertainty whether the next patch will be more rewarding or not than the current patch (Hall-McMaster et al., 2021; Kilpatrick et al., 2021). Animals will stay at the same, small reward patches in order to maximize their rate of reward (Hayden, 2016). In a foraging context, animals

will prioritize long-term rate maximization, even if that means exploiting small reward patches (Stephens et al., 2004). It is this reward-maximization behavior that can contribute to hyperbolic choice patterns, where if animals have a hard time identifying a post-reward delay it can result in small reward exploitation (Blanchard et al., 2013). Some studies have however demonstrated that rodents specifically devalue rewards to delays and not in consequence of post-reward waiting periods (E. A. Sjoberg et al., 2021). Taken together, these results and others suggest reward rate maximization may play an additional role in standard delay discounting tasks. These are considerations to take into account when analyzing and modeling choice behavior in discounting tasks.

1.3 Measures of Delay and Probability Discounting

Exponential and Hyperbolic Discounting

Researchers borrowed from economic theories of decision-making and found that an exponential decay rate describes a large amount of discounting behavior (Green & Myerson, 2004). As mentioned above, the exponential equation was developed to describe tradeoffs made in the face of increasing costs. In other words, how costs associated with large rewards and the availability of alternative low cost rewards "discount" the value of large rewards. The exponential equations are:

$$V = Ae^{-kD}$$

$$V = Ae^{-h(\frac{1}{P}-1)}$$

In delay discounting, V is the discounted value of A amount reward after D amount of time. In probability discounting, V and A are the same and P is the probability of the reward. Both equations have a free parameter *k*/*h* that describe the steepness of discounting, in other words the discounting rate. While these models were able to explain a large amount of variance in discounting data, some assumptions of the model made it harder to detect deviations from its predictions. These models assumed

a constant rate of discounting, which made it difficult to describe discounting behavior with changes in discounting rate over short and long delays due to preference reversals (Green & Myerson, 2004). This satisfied the observation that small reward alternatives shift preference but did not account for the psychological account of preference reversals (Green et al., 2014; Grüne-Yanoff, 2015).

In order to accommodate the extra variability in the data, researchers developed the hyperbolic forms of discounting (Mazur, 1987):

$$V = \frac{A}{1 + kD}$$

$$V = \frac{A}{1 + h(\frac{1}{P} - 1)}$$

The variables from this hyperbolic equation are the same as the exponential form, except for the predictions. Now, this model accounted for cases where the discounting rate changes according to how great the cost is (e.g., steeper for small costs, shallower for great costs).

Value estimates are derived from indifference points. Indifference points are obtained depending on the nature of the task. Indifference points are used to estimate the discounted value of the reward by comparing the points where the cost-free small reward and costly large reward come into conflict. In adjusting-delay types of tasks, either the short delay or the long delay option is adjusted according to choice behavior until subjects demonstrate equal preference for the two options (Mazur, 2000). Adjusting-amount designs are similar, but the amount of reward is changed instead (Frye et al., 2016; Richards et al., 1997). These indifference points are then used to estimate discounting rates. Instead of fitting curves to indifference points, fixed-delay and fixed-probability procedures fit discounting curves to the proportion of large choice (Evenden & Ryan, 1996). Fixed procedures are desirable for the reasons in the

previous section, mostly because the cost-free period adds a control period for magnitude effects free from cost (i.e., small versus large reward) and for comparisons of deviations from that cost-free period. Fixed procedures are especially prevalent in drug and animal studies for these reasons (Craig et al., 2014). Previous research findings indicate that fixed schedules are susceptible to cost presentation order effects (Robles et al., 2009), but some studies demonstrate they are comparable (Craig et al., 2014; Rodzon et al., 2011). There are also important ideas that can be gleaned from order effects, such as adjustments to switches (e.g., anchoring effects) and learning to differentiate the orders (Fox et al., 2008). Area under the curve (AUC) estimations are another method to obtain a measure of discounting derived from indifference points or percent of large choice throughout cost blocks.

Area Under the Curve

Discounting behavior can be explained by hypothesis driven models that have taken multiple forms across the literature. AUC analyses have been proposed as an alternative measure of discounting behavior because it does not make predictions about the shape of the data and is thus theory neutral (Myerson et al., 2001). Making no assumptions about the shape of the curve the data forms is valuable because not all subjects conform to the same fit, often subjects are excluded for not following a hyperbolic form. AUC is frequently used as a supplementary measure of discounting in order to obtain a comparable metric across research and to avoid problems with fit according to model assumptions (Yoon et al., 2017). AUC typically correlates well with discounting rate across different outcomes because discounting regularly takes a hyperbolic form (Odum et al., 2020). A theory informed AUC has been recently developed to derive AUC from the discounting rate (Killeen, 2023a). Discounting rate and AUC are typically reported together to have complementary measures that give insight into discounting sensitivity. These approaches are important in assessing

discounting behavior in clinical populations and are used to analyze discounting behavior in animal models as well.

1.4 Altered Decision-Making in Neurodevelopmental Disorders

Altered Decision-Making in Neurodevelopmental Disorders

Neurodevelopmental disorder diagnosis criteria are partly reflected by an alteration in reward motivated behaviors. These symptoms are typically grouped under a general collection of externalizing disorders, where a large proportion of the symptoms stem from impaired behavioral control (Karlsson Linnér et al., 2021). Decision-making can be influenced by increased amounts of repetitive behaviors and decreased ability to inhibit inappropriate behaviors (Balleine & Dezfouli, 2019), but to a limited extent (Vandaele & Ahmed, 2020). Decisions rely on learning the consequences of actions and adjusting behavior to explore new options or exploit highly rewarding options (Chen et al., 2021). An impaired ability to judge the value of an action and the appropriate timing of an action can be detrimental to building an optimal strategy congruent with the task at hand (Aparicio et al., 2019; Bergh et al., 2006).

These observations imply decisions are partially governed by learning, motivation, and cognition. Abnormally steep discounting is often attributed to increased impulsivity and has been attempted to explain a general increase in impulsive motor behaviors (Aparicio et al., 2019; Moon et al., 2006; Silverman et al., 2010). Impulsivity is thought to be a unifying construct for explaining impaired decision-making across tasks different costs. Previous research with clinical cases of attention-deficit hyperactivity disorder have found greater temporal discounting in children and adults (Castro Paiva et al., 2019; Jackson & MacKillop, 2016), but the findings are mixed for autism-spectrum disorder (Carlisi et al., 2017; Demurie et al.,

2012; Warnell et al., 2019). Other studies on reinforcement learning in attention-deficit hyperactivity disorder however give the alternate explanation that this behavior could result as a consequence of learning style (e.g., decreased use of model-based/model-free strategies) rather than an inherent tendency to choose small rewards (Nissan et al., 2023). Disruption in the balance between goal-directed and habitual control has been suggested to possibly be related to impulsivity (Hogarth et al., 2012). As discussed in previous sections, discounting can also arise through rate-maximization and thus probably represents a combination of factors rather than a sole impulsivity factor. It is probably this combination of factors that explains how discounting tasks lack predictive validity, even when accounting for the heterogeneity or severity of a disorder (Bailey et al., 2021). As some research points out, delay discounting has some predictive value, but does not completely capture impulsive decision-making or even specific criteria of neuropsychiatric disorders. Discounting effects are replicable (Stein et al., 2022) but even at their best cannot serve as diagnostic criteria for the presence of a disorder (Bailey et al., 2022). However, the way in which decision-making tasks like discounting are analyzed should be informed by the defining behaviors observed in clinical populations in order to better understand how they are engaging with these tasks.

One hypothesis of neurodevelopmental disorders posits that the flexibility of decisions can shift according to predictive relationships in different contexts (Sinha et al., 2014). Contingent associations presumably become harder to acquire, in some cases requiring additional training. Once acquired, however, that information becomes the best possible way to reduce uncertainty and people stick to it. People with neurodevelopmental disorders are hypothesized to struggle with incorporating violations of expectation to alter existing contingencies (D'Cruz et al., 2013; Miller et al., 2015). Behavioral flexibility is decremented because of problems with updating

previous learning (Weiss et al., 2020). These predictive properties have neural origins in reward learning structures like the striatum (Pavăl, 2017). Neurodevelopmental disorders arise partly due to genetic influences, which broadly alter brain functions and result in the previously mentioned behaviors.

Genetic Variants Associated With Neurodevelopmental Disorders

Genetic variation can be a significant source of behaviors associated with neurodevelopmental disorders. Human studies recognize that genetic variance contributes to the development of a disorder (Kreek et al., 2005; Niemi et al., 2018; Savatt & Myers, 2021). Deletion or duplication of genes important to developmental functions of the brain, which can alter development of neural circuits. Neurodevelopmental disorders involve reward and movement circuits, which are often compromised as a consequence of missing or duplicated genes (Fuccillo, 2016; Rein & Yan, 2020).

Alterations in neural circuits of reward and movement have behavioral consequences and explain some sources for core criteria of neurodevelopmental disorders. Fragile X syndrome, for instance, is caused by deletion of the FMR1 gene and is characterized by a series of developmental delays (Varghese et al., 2017). Deletion of SHANK genes are associated with behaviors, like repetitive actions, characteristic of neurodevelopmental disorders (Kalueff et al., 2016; Varghese et al., 2017). Both of these gene deletions are highly associated with characteristics of autism-spectrum disorder. 16p11.2 hemideletion — the deletion of one copy of the 16th chromosome — is a copy number variant that is also associated with autism-spectrum disorder (Niarchou et al., 2019).

Carriers of 16p11.2 hemideletion typically display language deficits and repetitive behaviors before any sort of formal diagnosis is made (Hanson et al., 2015). Further,

research has shown several aspects of social behavior are impaired in 16p11.2 hemideletion carriers (Benedetti et al., 2022). Previous research has discovered how patients of neurodevelopmental disorders, specifically autism-spectrum disorder, show deficits in the processing of social rewards (Scott-Van Zeeland et al., 2010), further reinforcing how 16p11.2 hemideletion is linked to important diagnostic phenotypes. In addition, there has been growing evidence supporting the validation of animal models for 16p11.2 hemideletion and their translational value to understanding neurodevelopmental disorders in humans (Bertero et al., 2018).

1.5 Neurological and Behavioral Outcomes in a Mouse Model of 16p11.2 Hemideletion

Behavior of a Mouse Model of 16p11.2 hemideletion

In an effort to capture the consequences of genetic deletions on a reproducible large scale, a mouse model of 16p11.2 hemideletion was developed. A group of researchers developed and provided evidence for a mouse model of 16p11.2 hemideletion (Horev et al., 2011). Horev and colleagues targeted a chromosomal region with conserved genes to that of human 16p11.2 hemideletion. The researchers reported some preliminary behavioral analyses, such as the development of stereotypic cage climbing behavior reminiscent of repetitive behaviors induced by nigrostriatal lesions (an observation we have noticed in unpublished findings) and hyperactivity. These results are promising because 16p11.2 hemideletion is highly associated with common behaviors in neurodevelopmental disorders, as mentioned in the human findings above. Due to the strong human parallels, 16p11.2 hemideletion mice are an important model for exploring the neurological and behavioral consequences of those neurodevelopmental conditions. This set up future research to explore changes in the brain caused by 16p11.2 hemideletion that cause these behavioral changes to arise.

Another research group used this model in order to explore system-wide structural changes in the brain and to further characterize the behavioral profile of 16p11.2 hemideletion mice (Portmann et al., 2014). Portmann and colleagues found increases in basal ganglia size compared to wildtype mice when controlling for relative volume, however an overall decrease in size in absolute volume. The researchers investigated possible changes in medium spiny neuron expression as a consequence of structural differences and found equal amounts of dopamine D1 receptor gene expression in male 16p11.2 hemideletion and wildtype mice, but increased dopamine D2 receptor gene expression in 16p11.2 hemideletion mice only. Their behavioral observations suggested hyperactivity occurs in familiar contexts such as home cages, but reduced activity occurs in novel environments. Portmann et al. (2014) administered a dopamine D2 receptor antagonist in mice and found 16p11.2 hemideletion movement to be less inhibited than wildtype mice. Collectively, these results imply 16p11.2 hemideletion mice behavioral alterations can partly be explained by alterations in basal ganglia signaling. This study only employed male mice, which is problematic because the consequences of neurodevelopmental disorders in females are underexplored.

Our lab further characterized cellular and behavioral changes in dopamine function in 16p11.2 hemideletion mice and demonstrated how these changes are sex dependent (Grissom et al., 2018). Male 16p11.2 hemideletion mice exhibited slowed acquisition of fixed-ratio 1 learning compared to wildtype mice. Female mice exhibited similar levels of reinforcement learning. One possible reason for this difference is that male 16p11.2 hemideletion mice were slower than male wildtype mice to learn their action of a center nosepoke led to reward. Grissom and colleagues critically determined this effect was not due to a difference in preference for the reinforcer because mice equally consumed it when it was offered ad libitum. Male 16p11.2 hemideletion mice committed less nonreinforced nosepokes which could be inferred to

be as a result of altered learning or willingness to engage in other actions. Further, when trained on a 5-choice serial reaction time task, male 16p11.2 hemideletion mice had a decreased number of correct responses due to responding on a wrong option during the stimulus period. Grissom et al. (2018) extended Portmann and colleagues' (2014) findings by showing male but not female 16p11.2 hemideletion mice had increased dopamine D2 receptor expression. These data reinforced previous findings of Portmann et al. (2014) by showing evidence for altered reinforcement learning, which is mediated by dopaminergic circuits. Female 16p11.2 hemideletion mice seemed to be unaffected in these contexts, but other research points to possible sensitivities they may exhibit.

While the previous studies found changes specific to male mice, female 16p11.2 hemideletion mice have recently been shown to have increased sensitivity to stressors and an anxiety-like phenotype (Giovanniello et al., 2021). Female 16p11.2 hemideletion spent less time in an open arm during an elevated plus maze task compared to female wildtype mice following a fear-inducing event. Giovanniello et al. (2021) followed up their behavioral results by recording from central amygdala neurons and found a greater magnitude of miniature excitatory postsynaptic potentials in female 16p11.2 hemideletion mice compared to wildtype mice. They demonstrated increased excitability was present in central amygdala neurons projecting to globus pallidus externa neurons, a circuit implicated in fear generalization. These results demonstrated the importance of female 16p11.2 hemideletion specific vulnerabilities that were not probed by previously mentioned tasks.

These results collectively suggest 16p11.2 hemideletion affects motivated behaviors and learning in a sex-biased manner. Goal-directed learning is important to everyday decision-making and helps with flexibility to multiple challenges. However, altered decision-making processes have yet to be explored in carriers of 16p11.2

hemideletion. Future studies should construct tasks that access differing types of costs in order to detect possible sensitivities to those costs.

2. Discounting rate in delay and probability discounting is sculpted by decision noise in mice

Sequential delay and probability discounting tasks in mice reveal anchoring effects partially attributable to decision noise

Gerardo R. Rojas, Lisa S. Curry-Pochy, Cathy S. Chen, Abigail T. Heller, Nicola M. Grissom

Department of Psychology, University of Minnesota, Minneapolis, MN 55455, USA

Abstract

Delay discounting and probability discounting decision making tasks in rodent models have high translational potential. However, it is unclear whether the discounted value of the large reward option is the main contributor to variability in animals' choices in either task, which may limit translation to humans. Male and female mice underwent sessions of delay and probability discounting in sequence to assess how choice behavior adapts over experience with each task. To control for "anchoring" (persistent choices based on the initial delay or probability), mice experienced "Worsening" schedules where the large reward was offered under initially favorable conditions that became less favorable during testing, followed by "Improving" schedules where the large reward was offered under initially unfavorable conditions

that improved over a session. During delay discounting, both male and female mice showed elimination of anchoring effects over training. In probability discounting, both sexes of mice continued to show some anchoring even after months of training. One possibility is that "noisy", exploratory choices could contribute to these persistent anchoring effects, rather than constant fluctuations in value discounting. We fit choice behavior in individual animals using models that included both a value-based discounting parameter and a decision noise parameter that captured variability in choices deviating from value maximization. Changes in anchoring behavior over time were tracked by changes in both the value and decision noise parameters in delay discounting, but by the decision noise parameter in probability discounting. Exploratory decision making was also reflected in choice response times that tracked the degree of conflict caused by both uncertainty and temporal cost, but was not linked with differences in locomotor activity reflecting chamber exploration. Thus, variable discounting behavior in mice can result from changes in exploration of the decision options rather than changes in reward valuation.

Keywords: delay discounting, probability discounting, modeling, touchscreen, mice

2.1 Introduction

Delay discounting tasks measure value assessments against a temporal cost, while probability discounting tasks measure value assessments against risky reward (Green et al., 2014; Odum, 2011a). These tasks have been important tools in assessing dysregulated reward processing in neuropsychiatric disorders such as addiction or neurodevelopmental disorders (Andrade & Petry, 2012) (Dalley et al., 2011; Richards et al., 1999; Rung et al., 2019). Because of this, translational animal versions of these tasks are of high interest (Mitchell, 2014; St Onge & Floresco, 2009). However, it is

unclear if animals use similar discounting strategies to those used by humans (Vanderveldt et al., 2016), for two reasons.

One issue arises from the fact that behavior in each of these tasks alone are thought to reflect choice impulsivity in animals(Acheson et al., 2006), even though these tasks contribute in different ways to assessing an impulsive profile (Strickland & Johnson, 2021). It has recently been shown in humans that multiple distinct discounting tasks are needed to better capture common traits (Białaszek et al., 2019); methods testing both delay and probability in the same animals are therefore of strong interest, but not widely available or used, especially in mice.

A second issue is that discounting tasks are typically modeled in both humans and rodents using economic value functions (e.g., *k*-values and *h*-values) which assume the main relevant factor in choices is the current discounted value of the reward (Odum, 2011a). However, recent evidence from the literature on reinforcement learning and decision making strongly implicates choice history and exploration as important variables in how both humans and animals perform value-based decision tasks (Chen et al., 2020; Cinotti et al., 2019; Daw et al., 2011; Speekenbrink & Konstantinidis, 2015). Importantly, animals often engage in non-reward seeking behaviors that are typically described as exploration (Findling et al., 2019; Gershman, 2019). This decision "noise" is rarely considered as a contributor to choices in discounting tasks despite exploratory events being necessary for animals to learn new task rules (Ebitz et al., 2018, 2019). In humans, exploration in other decision making tasks correlates with the degree of delay discounting shown (Sadeghiyeh et al., 2020). We have recently identified exploration as a key driver of sex differences in other decision making tasks (Chen et al., 2020, 2021). Because discounting tasks are often used to compare groups of animals modeling neuropsychiatric risk factors or other individual differences such as sex differences (Bos et al., 2014; Grissom & Reyes, 2019; Orsini et al., 2016; Orsini

& Setlow, 2017; Weafer & Wit, 2014), it is imperative to identify methods that allow us to distinguish whether differences in behavior are due to value judgements putatively reflecting impulsivity, or if exploration is a latent contributor to behavior.

One way to address these issues is to develop a method allowing within-subjects comparison of delay and probability discounting functions. This approach would allow for comparing overall performance within and between groups and permit computational modeling across tasks incorporating a decision noise parameter in addition to a value parameter. In the present study, we describe a sequence of delay and probability discounting tasks in mice achieving these goals. Mice are increasingly used for cognitive task batteries because of their high genetic tractability. Recent advancements in technology for mouse operant testing available through touchscreens have substantially improved the ease of training mice (Horner et al., 2013), enabling us to develop matched versions of probability and delay discounting for mice.

Here, we describe a novel battery of sequential delay and probability discounting schedules in touchscreens tested in male and female wildtype mice. One key factor previously shown to affect choices in these tasks is the order of presentation of delays or probabilities on the large reward (Koffarnus et al., 2011; Maguire et al., 2014; St Onge et al., 2010). We alternated mice between Worsening and Improving schedules within each discounting task, and demonstrated that these order effects are substantial in both male and female mice. These order effects are fully eliminated in both sexes with extensive training in delay discounting. However, order effects remain in probability discounting, especially in female mice, consistent with sex differences in risk processing but not "impulsivity" *per se* (Grissom & Reyes, 2019). We analyzed reward strategy via both win-stay/lose-shift analyses in probability discounting, and for both tasks with computational models incorporating both value and decision noise parameters in order to better understand how mice adapt choice strategies between

delay and probability discounting. Win-stay/lose-shift analyses suggest schedule differences in probability discounting emerged because females learned to adjust win-stay behavior consistently over experience with probability discounting. Delay discounting is captured by the combination of a value and decision noise parameter, but probability discounting is better explained by the choice parameter. Analysis of choice response times and locomotor behavior suggest that reductions in exploratory choices with increased experience are linked with increased deliberation between choices, and are likely not due to simple changes in task engagement. These results demonstrate that value functions may capture one aspect of impulsivity (i.e. overall reward preference), but that exploratory or "noisy" decisions are significantly contributors to mouse behavior in these tasks, especially as choice behavior changes between tasks or across multiple experiences of the same task.

2.2 Results

Age-matched male and female wildtype mice (n=15, 8 males, 7 females, strain B6129SF1/J) were trained to perform sequential discounting tasks using touchscreen operant chambers (**Figure 2.0A**). This permitted us to test the extent to which choice preferences and discounting were "anchored" by the initial delay/probability of the large reward. Trials were paced to require 30 seconds minus the length of the delay before the next trial could be initiated to remove the ability to complete all trials more quickly (Pearson et al., 2010) that may contribute to prior demonstrations of greater action impulsivity. Repeated sessions of delay and probability discounting allowed us to study anchoring effects and choice strategy, as well as ask questions about sex differences and individual differences in preferences for delay and probability simultaneously.

2.2.1 Delay discounting

Anchoring effects to delayed rewards are reduced with experience. Rodent models of delay discounting have previously used latin square design of delays (Mitchell, 2014), but structured delay schedules have not widely been used in mice (e.g. ascending or descending schedules). Because of the novelty of these schedules, we initially put mice through delay discounting in order to test whether sex is an important factor in anchoring effects induced by shifting schedules (i.e. "Worsening" and "Improving"). Research suggests uncertainty in discounting tasks can produce sex-specific effects (Grissom & Reyes, 2019), thus we wanted to test whether sex was important in anchoring responses to Worsening and Improving schedules. Here, we present the data from each round of Worsening and Improving schedules grouped together.

The first time mice experienced both the Worsening and Improving delay discounting schedules (Delay Discounting I), their preferences for the large reward across the entire sessions were heavily anchored by the initial delay experienced (**Figure 2.1A**, main effect of schedule, $F(1, 15.01) = 55.26$, $p < 0.001$). Mice learned to shift their choice from the delayed side to the immediate side as delay increased (**Figure 2.1A**, main effect of delay, $F(4, 60.17) = 52.26$, $p < 0.001$). Evidence suggested mice exhibited delay specific sensitivity according to schedule (**Figure 2.1A**, schedule x delay interaction, $F(4, 1081.77) = 5.84$, $p < 0.001$). Male mice had a large choice preference at all delays on the Worsening schedule compared to the Improving schedule (0s: $p = 0.0316$; 4s: $p < 0.001$; 12s: $p < 0.001$; 20s: $p = 0.0129$; 28s: $p = 0.0170$). Female mice had a large choice preference on the Worsening schedule compared to the Improving schedule except for at the 28s delay (0s: $p < 0.001$; 4s: $p < 0.001$; 12s: $p = 0.0138$; 20s: $p = 0.0406$). Mice exhibited this anchoring effect both at the beginning of training (**Figure 2.1A**, days 1-3: main effect of schedule, $p < 0.001$) and end of training (**Figure 2.1A**, days 6-8: main effect of

schedule, $p = 0.0136$). Although increased large choice preference across delays of the Worsening schedule was specific to the beginning of training (**Figure 2.1A**, days 1-3: schedule x delay interaction, $p < 0.001$). These results indicate temporal uncertainty induced by a schedule shift immediately caused anchoring effects. Preference for the large reward at all delays was greater when it was "anchored" by an initial 0 second delay than an initial 28 second delay.

Despite a second round of testing (**Figure 2.1B**, Delay Discounting II), anchoring effects persisted (**Figure 2.1B,** main effect of schedule, $F(1, 15) = 6.81$, $p = 0.0197$). Animals continued to show strong discounting to each transition of delay (**Figure 2.1B**, main effect of delay, $F(4, 60.21) = 116.42$, $p < 0.001$). Mice adjusted choice across delays according to schedule (**Figure 2.1B**, schedule x delay interaction, $F(4, 1251.92) = 2.98$, $p = 0.018$). Anchoring was especially apparent in female mice (**Figure 2.1B**, sex x schedule x delay interaction, $F(4, 1251.92) = 5.21$, $p < 0.001$). Anchoring in female mice was specific to the smallest delays (0s: $p = 0.0465$; 4s: $p < 0.001$; 12s: $p = 0.0057$). Mice retained anchoring early into training (**Figure 2.1B**, days 1-3: main effect of schedule, $p = 0.0123$), but the effect was reduced with additional training (**Figure 2.1B**, days 6-8: main effect of schedule, $p = 0.207$). However evidence suggests female mice retained sensitivity to differences in schedule order with experience (**Figure 2.1B**, days 6-8: sex x schedule interaction, $p = 0.0238$; days 6-8: sex x schedule x delay interaction ns, $p = 0.0577$). These data suggest that female mice have an increased sensitivity to anchoring effects and/or increased sensitivity to unexpected changes in the task rules.

By the time animals were tested on Delay Discounting III (**Figure 2.1C**), there were no longer any anchoring effects or differences apparent in choice (main effect of schedule ns, $p > 0.05$). Animals showed robust discounting to each delay (**Figure 2.1D**, main effect of delay, $F(4, 60) = 175.55$, $p < 0.001$) that had some suggestion of delay

specific effects between schedules (**Figure 2.1D**, schedule x delay interaction, $F(4, 2355) = 5.79$, $p < 0.001$) but did not reveal any post hoc effects. Mice expressed anchoring effects early into training still (**Figure 2.1C**, days 1-3: main effect of schedule, $p = 0.0282$) of which was more apparent in females (**Figure 2.1C**, days 1-3: sex x schedule x delay interaction, $p = 0.0140$). Increased training mitigated anchoring effects (**Figure 2.1C**, days 6-8: main effect of schedule ns, $p > 0.05$; days 6-8: schedule x delay interaction, $p = 0.0064$). Taken together, our results suggest mice have the ability to form similar choice preferences across two different schedule orientations. Anchoring effects are eliminated by the end of discounting for both male and female mice. However, females exhibited persistent anchoring until the penultimate round of discounting.

2.2.2 Probability discounting

Anchoring effects are persistent when discounting risky rewards. We put mice through probability discounting with "Worsening" and "Improving" schedules in order to challenge anchoring in response to uncertain large rewards. We tested mice on these schedules to see if risky rewards produced anchoring effects in a similar manner. If males and females are differently affected by uncertainty, it would stand to reason that those differences may be most reflected in the anchoring effects. We present the data from each round of Worsening and Improving schedules grouped together.

Unlike delay discounting, the first time mice experienced both the Worsening and Improving probability discounting schedules (Probability Discounting I), their preferences for the large reward across the entire sessions were not heavily anchored overall by the initial probability experienced (**Figure 2.2A**, no main effect of schedule, $F(1, 15.01) = 0.42$, $p = 0.527$). All mice showed significant discounting on both schedules, as measured by changes in their preferences for the large reward (**Figure**

2.2A, main effect of probability, $F(4, 60.88) = 123.43$, $p < 0.001$). However, mice showed schedule specific preferences depending on the probability (**Figure 2.2A**, schedule x probability interaction, $F(4, 1122.52) = 21.15$, $p < 0.001$) and differences in the magnitude of preference between schedules (**Figure 2.2A**, sex x schedule interaction, $F(1, 15.01) = 4.83$, $p = 0.0441$). Males chose the large reward more often on the Worsening schedule compared to the Improving schedule when the chance of winning risky rewards was at 25% ($p = 0.0016$) and 50% ($p = 0.0102$). Females increased large choice preference on the Worsening schedule compared to the Improving schedule when large rewards were guaranteed (100% chance, $p = 0.0121$). Anchoring did not appear throughout training (**Figure 2.2A**, days 1-3 and days 6-8, main effect of schedule ns, $p > 0.05$). However, differences in preference at different probabilities between schedules were present at the start of training (**Figure 2.2A**, days 1-3, schedule x probability interaction, $p < 0.001$) and the end of training (**Figure 2.2A**, days 6-8, schedule x probability interaction, $p = 0.0144$). Male specific differences in preference seemed to appear early on (**Figure 2.2A**, days 1-3, sex x schedule x probability interaction ns, $p = 0.0542$). These data indicate mice were not as affected by schedule effects compared to delay discounting, especially with greater experience.

As mice continued to gain more experience with the task design during the second round of testing (**Figure 2.2B**, Probability Discounting II), anchoring effects became more apparent (**Figure 2.2B**, main effect of schedule, $F(1, 15) = 37.74$, $p < 0.001$). All mice continued to show strong discounting to each transition of probability (**Figure 2.2B**, main effect of probability, $F(4, 60) = 128.56$, $p < 0.001$). These anchoring effects were specific to different probabilities of risky rewards (**Figure 2.2B**, schedule x probability interaction, $F(4, 1140) = 9.08$, $p < 0.001$). Female mice significantly reduced their large choice preference on the Improving schedule compared to the

Worsening schedule when rewards were risky (12.5 % chance, $p = 0.0149$; 25% chance, $p < 0.001$; 50% chance, $p = 0.0025$; 75% chance, $p = 0.0082$). Male mice started to show reduced large choice preference on the Improving schedule compared to the Worsening schedule at 25% chance of reward ($p < 0.001$) and 50% chance of reward ($p < 0.001$). While the overall data suggests no sex effects, male and female mice did have a difference in large reward preference early into training (**Figure 2.2B**, days 1-3: main effect of sex, $p = 0.0163$). Anchoring effects were present early into training (**Figure 2.2B**, days 1-3: main effect of schedule, $p < 0.001$) and at the end of training (**Figure 2.2B**, days 6-8: main effect of schedule, $p = 0.0013$). Early training emphasized a difference in large preference at different probabilities between schedules (**Figure 2.2B**, days 1-3, schedule x probability interaction, $p < 0.001$) and differences in overall preference between schedules (**Figure 3B**, days 1-3, sex x schedule interaction, $p < 0.001$). Mice shifted to differences in large preference based on specific probabilities rather than generally between the schedules (**Figure 2.2B**, days 6-8, sex x probability interaction, $p = 0.0188$). These data indicate anchoring effects reappeared after mice gained more experience where reward receipt was probabilistic.

By the time animals were tested on Probability Discounting III (**Figure 2.2C**), anchoring effects were still pervasive (**Figure 2.2D**, main effect of schedule, $F(1, 14.22) = 7.80$, $p = 0.0142$). All mice showed significant discounting (**Figure 2.2D**, main effect of probability, $F(4, 60.57) = 315.46$, $p < 0.001$). Mice again showed schedule specific preferences depending on the probability (**Figure 2.2D**, schedule x probability interaction, $F(4, 2198.20) = 45.12$, $p < 0.001$) and within sex (**Figure 2.2D**, sex x schedule interaction, $F(1, 14.22) = 4.98$, $p = 0.0422$). Female mice reduced risky choices on the Improving schedule compared to the Worsening schedule when uncertainty was high (12.5% chance, $p = 0.0106$; 25% chance, $p < 0.001$) and

increased risky choices on the Improving schedule compared to the Worsening schedule when uncertainty was low (75% chance, $p = 0.0102$; 100% chance, $p < 0.001$). Male mice also made risk-based adjustments but only to maximize rewards on the Improving schedule compared to the Worsening schedule when uncertainty was at or above chance level (50% chance, $p = 0.0350$; 75% chance, $p < 0.001$; 100% chance, $p < 0.001$). Mice showed a reduction of anchoring early into training, but anchoring resurfaced with additional exposure to probabilistic rewards (**Figure 2.2C**, days 6-8: main effect of schedule, $p < 0.001$). Large reward preference remained sensitive to the schedule type and specific probability experienced (**Figure 2.2C**, days 1-3: schedule x probability interaction, $p < 0.001$; days 6-8: schedule x probability interaction, $p < 0.001$). Schedule directed large reward preference was present early into training (**Figure 2.2C**, days 1-3: sex x schedule interaction, $p < 0.001$) but not late into training. These results demonstrate that females and males made schedule specific adjustments in avoiding losses around an immediately risky schedule (i.e. Improving probability discounting) and that mice continued to remain sensitive to schedule effects with extended training.

2.2.3 Win-stay/lose-shift Adaptations to Risk Order

Probability discounting schedules that decrease trial block uncertainty promote win-stay choices. Win-stay/lose-shift behaviors are important indicators of strategy specific adaptations to wins and losses. Win-stay ratios were calculated by dividing how often mice stayed on the same risky side after being rewarded divided by all rewarded risky responses. Lose-shift ratios were calculated by dividing how often mice switched to the small guaranteed side after not receiving a large risky reward divided by all losses on the risky side. We wanted to study if females and males made specific adaptations to wins and losses in response to different risk orientations, which can be a source of choice differences (Stopper et al., 2014).

At the beginning of probability discounting, mice increased win-stay behavior on the Improving schedule (**Figure 2.3A**, main effect of schedule, $F(1, 255) = 34.38$, $p < 0.001$), but a sex x schedule interaction (**Figure 2.3A**, $F(1, 255) = 21.77$, $p < 0.001$) indicates only females increased win-stay behavior on the Improving schedule ($p < 0.001$). Mice also increased lose-shift behavior on the Improving schedule (**Figure 2.3B**, main effect of schedule, $F(1, 255) = 76.87$, $p < 0.001$). Both male ($p < 0.001$) and female ($p < 0.001$) mice exhibited increased lose-shift behavior on the Improving condition. These results suggest both male and female mice adjust their behavior to increased initial uncertainty in a similar way, but females specifically showed a schedule specific distinction in maximizing rewards.

Mice continued to make win-stay choice adaptations throughout the second round of probability discounting (**Figure 2.3A**, main effect schedule, $F(1, 270) = 10.50$ $p = 0.0013$) which was again specific to female mice (Worsening II > Improving II, $p = 0.0116$) and continued to make loss specific adaptations (**Figure 2.3B**, main effect of schedule, $F(1, 255) = 32.50$, $p < 0.001$). Increased lose-shift behavior on the Improving schedule was found again in both males ($p < 0.001$) and females ($p = 0.0016$). Female mice again differentiated themselves from male mice because females shape their choice behavior around adjustments in win-stay behavior.

As mice finished probability discounting, male ($p < 0.001$) and female ($p < 0.001$) mice continued to increase win-stay behavior on the Improving schedule (**Figure 2.3A**, main effect of schedule, $F(1, 460.28) = 55.94$, $p < 0.001$). A main effect of schedule was found for lose-shift behavior (**Figure 2.3B**, $F(1, 461.30) = 26.65$, $p < 0.001$), but only male mice exhibited decreased aversion to losses (**Figure 2.3B**, $F(1, 461.30) = 5.16$, $p = 0.0235$) on the Improving schedule ($p = 0.0024$). Females made schedule specific adaptations to wins even with minimal experience whereas males required extended training to show win-stay schedule effects. Males and females were equally

loss avoidant for the most of training, extended training reduced schedule specific effects in females whereas males remained sensitive to losses.

2.2.4 Computational models of choice variability in delay and probability discounting

Decision noise is a major contributor to delay and probability discounting. Win-stay/lose-shift analyses revealed some learning specific effects, but did not explain changes in preference over renditions of the task. We posited that snapshot win-stay/lose-shift analyses might not capture broader trends driving choice behavior. There is a growing amount of evidence suggesting exploration is a key latent variable driving choice behavior (Ebitz et al., 2018, 2019). Therefore, we pursued two discounting models for delay and probability, each incorporating a value parameter (k for delay, h for probability; (Green et al., 2014)) and an inverse temperature parameter capturing variability in choice around these value preferences (β). This allowed us to track adaptations in value and choice within and between tasks. We excluded sex as a factor to increase power and better detect schedule specific adaptations. Data were fit to an exponential model rather than a hyperbolic model for delay and probability discounting because those provided better fits according to AIC (Delay Discounting Hyperbolic AIC (Improving = 31055.57 & Worsening = 29689.31) > Delay Discounting Exponential AIC (Improving = 29910.06 & Worsening = 28364.77); Probability Discounting Hyperbolic AIC (Improving = 52026.82 & Worsening = 54768.72) > Probability Discounting Exponential AIC (Improving = 47774.92 & Worsening = 52824.56)).

Analysis of the delay discounting rate parameter (k) revealed mice had smaller discounting rates for the Worsening condition (**Figure 2.4A**, main effect of schedule, $F(5, 75) = 8.53$, $p < 0.001$). Mice had a steeper discounting rate initially on the Improving condition compared to the Worsening condition (**Figure 2.4A**, $p < 0.001$),

possibly due to uncertainty induced by a switch in orientation. Mice were steeper still for the second round of Improving and Worsening schedules (**Figure 2.4A**, $p = 0.0371$). Mice became more willing to endure delays for large rewards between the first and third round of Improving schedules (**Figure 2.4A**, $p = 0.0171$). Probability discounting rates (h) were also schedule dependent (**Figure 2.4A**, $F(5, 75) = 8.12$, $p < 0.001$), especially for the second session of discounting (**Figure 2.4A**, Improving II > Worsening II, $p < 0.001$). Mice became steeper between for the third round of Worsening discounting compared to the second (**Figure 2.4A**, $p = 0.0104$). Mice became more cost enduring with training for the large reward in delay discounting, but were generally cost avoidant for the large reward in probability discounting.

Between-task analyses of the inverse temperature in our discounting model revealed delay and probability discounting noise were similar in magnitude (**Figure 2.4B**, main effect of task ns, $p > 0.05$). However there were specific task x schedule interactions which were present (**Figure 2.4B**, $F(5, 75) = 8.37$, $p < 0.001$). Mice engaged in increased repetitive choice on Worsening III of delay discounting compared to probability discounting (**Figure 2.4B**, $p < 0.001$).

Within-task comparisons of the choice parameter revealed mice adapt to initial temporal uncertainty caused by an orientation switch through increased sampling of reward options (**Figure 2.4B**, Worsening I > Improving I, $p = 0.0021$). Mice learned to decrease decision noise throughout their experience with Improving delay discounting (**Figure 2.4B**, Improving I < Improving II, $p = 0.0430$; Improving I < Improving III, $p < 0.001$). Mice similarly learned to decrease decision noise with Improving probability discounting (**Figure 2.4B**, Improving I < Improving III, $p = 0.0065$; Improving II < Improving III, $p < 0.001$). By the end of training, decision noise was greater for Worsening probability discounting compared to Improving probability discounting (**Figure 2.4B**, Improving III > Worsening III, $p < 0.001$).

Mice choice preference in delay discounting demonstrates that differences in behavior can at least in part be explained by changes in *k*, the parameter describing the discounted value of the large reward. That is, when we model the behavior of these animals, the best fit model adjusts the *k* value over the progression of Delay Discounting (I, II, III), meaning that differences in animals' choices across the progression of delay discounting tasks reflect in part true changes in the discounted value of the large reward. When animals' choice curves shift, this reflects changes in how much they value the reward, not only changes in how variable or noisy they are in responding. In contrast, in the probability discounting task, the comparable *h* value, the discounted value of the large reward, is stable - meaning that changes in behavior in this task are not described well by models adjusting the value of the reward. Rather, adjustments in the inverse temperature parameter, which measures how noisy animals are in adhering to their values, provide the best fit - suggesting that exploratory noise is a major driver in decision making in the probability task.

2.2.5 Choice response times during delay and probability tasks

Examination of response times and locomotion as indexes of cognitive uncertainty and changing task engagement, respectively. Exploration has a number of different meanings in different fields. Our computational model contains an inverse temperature parameter that reflects the "noisiness" of adhering to a value based decision rule, and this noisiness could emerge from one of several sources. First, it could reflect cognitive uncertainty and exploration processes; these can be reflected in increasing choice response times as decisions become more costly or more uncertain (Fontanesi et al., 2019; McDougle & Collins, 2021; Steverson et al., 2019). Alternatively, noisiness in behavior could reflect changing engagement with the tasks over time, which can be reflected in the locomotor behavior or unobserved behaviors of animals as they complete the tasks (Musall et al., 2019; Thompson et al., 2018). As

the exploration in both of our tasks goes down with experience, we should see changes in *choice response times* with task experience if this exploration reflects aspects of cognitive uncertainty, versus changes in *locomotor behavior* if exploration reflects task disengagement. We addressed each of these possibilities by examining choice response times and locomotor behavior incidentally captured in the chambers during collection of the above task performance data.

2.2.5.1 Choice response times track temporal uncertainty only in animals experienced with delay discounting.

Choice response times are subject to change according to schedule orientation in discounting tasks (Robles & Vargas, 2007). Research suggests discounting choice response times increase when the subjective value of the large delayed option equals the delay-free option (Basile & Toplak, 2015; Robles & Vargas, 2007). These researchers suggested reaction time reflects conflict of equally valued options, something that should reveal itself with increasing or decreasing costs. Response times then could reflect the difficulty present in our within-subject tasks where choice preference is anchored to what is learned in the preceding schedule. If differences in exploration over training on the delayed discounting task do reflect increasing conflict as animals better understand the temporal cost of a choice, then we should see the emergence of delay-dependent increases in choice response times over our training schedules, particularly as anchoring effects go away. Indeed, we found that choice response times tracked the degree of temporal uncertainty of each delay in the most highly trained sections of the task (Delay Discounting III) compared to naive performance of the task (Delay Discounting I).

In order to demonstrate this, we took the time from choice presentation to decision made as an index of choice response time. These data are depicted in Figure 6. The first round of delay discounting showed immediate evidence of the effect of delay

on response time. Anchoring has an effect on choice response times where expectation of an initial delay-free period (i.e. from the Worsening schedule) causes mice to speed up on the Improving schedule. It is likely that small choice response times in this case reflect the learning effect of experiencing delay for the first time. Both the small immediate option (**Figure 2.5A**, main effect of delay, $F(4, 57.42) = 4.29$, $p = 0.0042$) and large delayed option (**Figure 2.5A**, main effect of delay, $F(4, 40.32) = 33.41$, $p < 0.001$) showed evidence of increased response time with delay. Small choice response times showed evidence of anchoring (**Figure 2.5A**, main effect of schedule, $F(1, 14.93) = 45.06$, $p < 0.001$). Small choice schedule differences in delay (**Figure 2.5A**, schedule x delay interaction, $F(4, 5460.37) = 4.82$ $p < 0.001$) were apparent as soon as rewards were delayed in which mice were faster to respond for small rewards on the Improving schedule compared to the Worsening schedule (4s: $p < 0.001$; 12s: $p < 0.001$; 20s: $p < 0.001$; 28s: $p < 0.001$). Large choice response times also showed evidence of anchoring (**Figure 2.5A**, main effect of schedule, $F(1, 14.85) = 4.91$, $p = 0.0427$). Large choice schedule effects per delay (**Figure 2.5A**, schedule x delay interaction, $F(4, 5603.65) = 26.99$, $p < 0.001$) were found across most delays. Mice generally were slower at smaller delays when the schedule was decreasing in delay (Improving > Worsening, 0s: $p < 0.001$; 4s: $p = 0.0393$), but sped up at larger delays (Worsening > Improving, 20s: $p < 0.001$; 28s: $p < 0.001$). Delay differently affected males and females (**Figure 2.5A**, sex x schedule x delay interaction, $F(4, 5603.65) = 3.85$, $p = 0.00395$) where females specifically slowed down on the Improving schedule when there was no delay (female Improving > Worsening, 0s: $p < 0.001$), but sped up at larger delays (female Worsening > Improving, 20s: $p < 0.001$; 28s: $p < 0.001$). Females were faster than males on the Improving schedule at the 28s delay (male Improving > female Improving, $p = 0.0067$). Males only sped up on the Improving schedule at the 20s delay (male Worsening > male Improving, $p = 0.0091$).

The second round of delay discounting was still subject to delay effects on both the small option (**Figure 2.5B**, main effect of delay, $F(4, 50.39) = 9.14$, $p < 0.001$) and the large option (**Figure 2.5B**, main effect of delay, $F(4, 53.01) = 12.51$, $p < 0.001$). Further experience with delay orientations showed mice mitigated choice response time differences between schedules as they learned to distinguish between the two schedules. Small choice response times had some suggestion of anchoring (**Figure 2.5B**, main effect of schedule, $F(1, 18.60) = 10.29$, $p = 0.0047$), but at no specific delays. Large choice response times suggested no anchoring, but delay specific effects (**Figure 2.5B**, schedule x delay interaction, $F(4, 7552.42) = 5.25$, $p < 0.001$). Mice slowed down on the Improving schedule during the 20s delay (Worsening > Improving, $p = 0.0189$). Females were slower than males when collapsed across schedules (**Figure 2.5B**, sex x delay interaction, $F(4, 53.01) = 4.63$, $p = 0.0027$) at the 20s delay ($p = 0.0347$) and 28s delay ($p = 0.0094$). Mice need to slow down choice time on the Improving schedule to make a large choice when the delay was initially high. Choice time was adjusted to delays in general, but also to the presentation of delays. This indicated increased choice response times are deliberate and are adjusted to not only the cost of the task, but also to the order of the cost.

Schedule effects became more pronounced with extended training. Compared to previous rounds, increased schedule effects were feasible because choice response time tracks cost deliberation according to the order of delays. Mice adjusted response times according to delay on both the small option (**Figure 2.5D**, main effect of delay, $F(4, 51.79) = 7.36$, $p < 0.001$) and the large option (**Figure 2.5D**, main effect of delay, $F(4, 45.03) = 12.51$, $p < 0.001$). Small choice response times were sensitive to delay according to the schedule (**Figure 2.5D**, schedule x delay interaction, $F(4, 11368.97) = 3.65$, $p = 0.0057$) and sex (**Figure 2.5D**, sex x schedule x delay interaction, $F(4, 11368.97) = 7.18$, $p < 0.001$). Mice were slower on the Worsening schedule at the 0s

delay (Worsening > Improving, $p = 0.0034$), female mice specifically were slower on the Worsening schedule at the 12s delay (female Worsening > female Improving, $p = 0.0187$). Large choice response times also had specific delay effects (**Figure 2.5D**, schedule x delay interaction, $F(4, 13465.12) = 20.00$, $p < 0.001$) with slower response times on the Improving schedule at the 4s (Improving > Worsening, $p = 0.004$) and the 12s (Improving > Worsening, $p < 0.001$) delays, but also slower response times on the Worsening schedule at the 28s (Worsening > Improving, $p < 0.001$) delay. These effects were however sex driven (**Figure 2.5D**, sex x schedule x delay interaction, $F(4, 13465.12) = 6.62$, $p < 0.001$) with females being slower on the Improving schedule at the 4s (female Improving > female Worsening, $p < 0.001$), 12s (female Improving > female Worsening, $p < 0.001$) and 20s (female Improving > female Worsening, $p = 0.0383$) delays, and males being slower on the Worsening schedule at the 28s (male Worsening > male Improving, $p < 0.001$) delay. Males were significantly faster than females when comparing Improving schedule large choice response times at the 4s delay (female Improving > male Improving, $p < 0.001$). Extended experience on delay discounting increased schedule differences in large choice response times. Females in particular slowed down on a schedule that is immediately costly for large choices, which is in line with previous rounds of choice behavior demonstrating anchoring effects are especially prominent in females.

2.2.5.2 Choice response times reflect increasing reward delivery uncertainty throughout probability discounting.

Reward uncertainty can modulate the speed in which choices are made. Assessment of uncertainty can lead to differences in response speed, leading to generally slower choices in situations of greater conflict. A previous report of Worsening probability discounting has demonstrated Long-Evans female rats slowed down choices in general compared to male rats (Braunscheidel et al., 2019; Islas-

Preciado et al., 2020). We have previously observed that differences in choice response time can be attributed to individual differences in explorative/exploitative strategy engagement (Chen et al., 2020, 2021). Decision noise contributes to the type of strategy used, something we have demonstrated to change throughout experience with probability discounting. If differences in exploration over training on the probability discounting task reflect increasing conflict as animals better understand the risky cost of a choice, then we should see the emergence of probability-dependent increases in choice response times over our training schedules. Indeed, we found that choice response times tracked the degree of reward uncertainty of each probability block in the most highly trained sections of the task (Probability Discounting III) compared to naive performance of the task (Probability Discounting I).

The introduction of a risky uncertain large reward option caused mice to reassess schedule differences. Mice sped up choice response times as the large reward became more probable regardless of whether they were selecting the small reward side (**Figure 2.6A**, main effect of probability, $F(4, 61.21) = 4.23$, $p = 0.0044$) or large reward side (**Figure 2.6A**, main effect of probability, $F(4, 61.36) = 72.78$, $p < 0.001$). Anchoring was apparent both in small choice (**Figure 2.6A**, main effect of schedule, $F(1, 18.69) = 31.53$, $p < 0.001$) and large choice (**Figure 2.5A**, main effect of schedule, $F(1. 13.97) = 57.35$, $p < 0.001$). Schedule effects in small choice response times were apparent at different probabilities (**Figure 2.6A**, schedule x probability interaction, $F(4, 7631.93) = 3.42$, $p = 0.0084$) for both males and females (**Figure 2.6A**, sex x schedule x probability interaction, $F(4, 7631.93) = 3.57$, $p = 0.0065$). General schedule differences for small choice were found at all probabilities except for when the trial block had 100% delivery rate on the large risky side (Worsening > Improving, 12.5%: $p < 0.001$; 25%: $p < 0.001$; 50%: $p < 0.001$; 75%: $p < 0.001$). Females were slower to make a small choice on the Worsening schedule when the probability of large rewards was at

12.5% (female Worsening > female Improving, $p < 0.001$) and 25% (female Worsening > female Improving, $p = 0.0022$). Males were slower to make a small choice on the Worsening schedule at all probabilities (male Worsening > male Improving, 12.5%: $p = 0.0031$; 25%: $p < 0.001$; 50%: $p < 0.001$; 75%: $p = 0.0022$; 100%: $p = 0.0093$). Males were slower to make a small choice on the Worsening schedule than females when large rewards were always delivered (male Worsening > female Worsening, 100%: $p = 0.0055$). General schedule effects for large risky response times were found (**Figure 2.6A**, schedule x probability interaction, $F(4, 61.38) = 72.78$, $p < 0.001$). Mice had slower large choice response times on the Worsening schedule at most probabilities except when chance of delivery was 50% (Worsening > Improving, 12.5%: $p < 0.001$; 25%: $p < 0.001$; 75%: $p = 0.0068$; 100%: $p = 0.0062$). Mice sped up when safety increased within a schedule (i.e. Improving) compared to when risk increased within a schedule (i.e. Worsening) for both small and large options. Males seemed to be most affected by risk orientation while females were sensitive to only the riskiest trial blocks (i.e. 12.5% and 25% probabilities).

Mice were quick to adjust to risky large rewards after Probability Discounting II. As noted by the win-stay/lose-shift data, male mice adapted primarily by shifting to the small choice in a schedule that promotes large initial uncertainty. This is how male mice combat uncertainty and how they made that choice easier (i.e. reduced response time difference). Females dealt with uncertainty by adjusting both win-stay and lose-shift behavior, which created increased difficulty of choice especially when uncertainty was large. Mice responding still slowed down the more uncertain a reward became for large choice (**Figure 2.6B**, main effect of probability, $F(4, 58.53) = 62.70$, $p < 0.001$) but not for small choice. Small choice response times were still adjusted according to schedule (**Figure 2.6B**, main effect of schedule, $F(1, 15.46) = 13.69$, $p = 0.0020$) but not large choice response times. Mice showed schedule effects for small choice

(**Figure 2.6B**, schedule x probability interaction, $F(4, 9006.65) = 2.48$, $p = 0.0422$) as soon as the large option became risky (Worsening > Improving, 12.5%: $p < 0.001$; 25%: $p < 0.001$; 50%: $p = 0.0071$; 75%: $p = 0.0172$). Female small choice response times were especially sensitive to schedule effects (**Figure 2.6B**, sex x schedule x probability interaction, $F(4, 9006.65) = 10.14$, $p < 0.001$) at 12.5% (female Worsening > female Improving, $p < 0.001$) and 25% (female Worsening > female Improving, $p < 0.001$) probabilities. Mice large choice response times were sensitive at 12.5% (Worsening > Improving, $p < 0.001$) and 25% (Worsening > Improving, $p = 0.0408$). Females slowed down on the Worsening schedule when large reward receipt was unlikely (female Worsening > female Improving, 12.5%: $p < 0.001$; 25%: $p = 0.0151$), as well as males (male Worsening > male Improving, 12.5%: $p = 0.0161$). Mice, especially females, showed difficulty in choice when the objective reward value of small choice equaled or was larger than large choice (i.e. 1:1 ratio of reward volume for small choice or large choice only; 25% reward probability).

After extended training on probability discounting, choice response times were no longer sensitive to anchoring. Anchoring in choice was still prominent because probability discounting is a noisy task that mice constantly adjust to, but that aspect becomes less apparent in choice response time. Small choice response times became sensitive to probability again (**Figure 2.6D**, main effect of probability, $F(4, 59.40) = 3.22$, $p = 0.0186$) and large choice response times remained sensitive (**Figure 2.6D**, main effect of probability, $F(4, 65.57) = 96.56$, $p < 0.001$). Mice were still generally affected by uncertainty for small choice (**Figure 2.6D**, schedule x probability interaction, $F(4, 17455.37) = 15.83$, $p < 0.001$) and large choice (**Figure 2.6D**, schedule x probability interaction, $F(4, 33459.35) = 41.12$, $p < 0.001$). Mice sped up small rewards on the Worsening schedule when large rewards were guaranteed (Improving > Worsening, $p = 0.0038$), but slowed down as risk increased (Worsening >

Improving, 12.5%: $p = 0.0015$; 25%: $p < 0.001$; 50%: $p = 0.0062$). Mice sped up for large rewards on the Worsening schedule when chance of reward delivery was 50% (Improving > Worsening, $p = 0.0101$), but slowed down when probability of reward was at its riskiest (Worsening > Improving, 12.5%: $p < 0.001$). Our results suggest extended experience with probability discounting helped mitigate sex-specific differences in choice response time speed.

2.2.6 Locomotor activity in delay and probability discounting

Patterns of locomotor activity are task dependent, but unrelated to changes in choice over experience with the tasks. It is possible that the shifts in preference for the large reward seen over multiple sessions in these schedules reflect simpler contributions to behavior than changes in decision making processes such as deliberation and cognitive exploration. For example, although animals are familiar with the chamber, perhaps continued locomotor exploration is also changing during this time and influencing the choices animals make. If this were true, then we should see locomotor activity within the chamber changing in synchrony with changes in preference. To test this hypothesis, we measured locomotor activity across all days of testing using infrared beams located at the front (close to the touchscreen) and back (close to the magazine) of the chamber. We did not find support for changes in locomotor behavior that could explain changes in decision exploration over experience.

Over delay discounting, we found evidence of different locomotor strategies between Improving and Worsening versions of the task; however, this did not resolve over time, despite stabilization of choice behavior over that same period. Large temporal gaps between choice and reward delivery at the start of the schedule shifted locomotor behavior from the touchscreen to the magazine. In the first round of delay

discounting, mice demonstrated an increase in locomotor activity (**Figure 2.7A**, main effect of break location, $F(1, 430) = 3.82$, $p < 0.001$) that was schedule dependent (**Figure 2.7A**, main effect of schedule, $F(1, 430) = 21.46$, $p < 0.001$; schedule x break location interaction, $F(1, 430) = 34.01$, $p < 0.001$). Mice focused locomotor behavior toward the touchscreen on a Worsening schedule (Worsening front > Worsening back, $p < 0.001$), but switched focus towards the magazine on an Improving schedule (Improving back > Improving front, $p < 0.001$). Mice increased activity near the magazine on the Improving schedule compared to the Worsening schedule (Improving back > Worsening back, $p < 0.001$) and decreased locomotor activity near the touchscreen (Worsening front > Improving front, $p < 0.001$).

Similar patterns of locomotor behavior persisted for the second and third round of delay discounting. Mice changed their pattern of locomotor activity in a schedule dependent manner (**Figure 2.7B**, main effect of schedule, $F(1, 502) = 1119.82$, $p < 0.001$; schedule x break location interaction, $F(1, 503) = 1430.63$, $p < 0.001$). Delays again induced large schedule differences which were not reflective of their choice behavior. Locomotor behavior was focused towards the touchscreen on a Worsening schedule (Worsening front > Worsening back, $p < 0.001$), but shifted towards the magazine on an Improving schedule (Improving back > Improving front, $p < 0.001$). Mice again checked the back of the chamber more often for the reward on the Improving schedule compared to the Worsening schedule (Improving back > Worsening back, $p < 0.001$) which came at the cost of decreased activity near the touchscreen (Worsening front > Improving front, $p < 0.001$). Increased waiting time for the reward caused mice to again narrow locomotor activity near the magazine. Reward delivery delay changes locomotor behavior, but not choice behavior (i.e. anchoring effects are eliminated at this point). Male and female mice started to diverge in locomotor allocation by Delay Discounting III (**Figure 2.7D**, main effect of sex,

$F(1, 13) = 6.68$, $p = 0.0226$). Mice still altered locomotor activity (**Figure 2.7D**, main effect of break location, $F(1, 916) = 13.97$, $p < 0.001$) in a schedule specific way (**Figure 2.7D**, schedule x break location interaction, $F(1, 916) = 946.03$, $p < 0.001$). Male and female mice by this point settled on different locomotor patterns (**Figure 2.7D**, sex x schedule x break location interaction, $F(1, 916) = 65.07$, $p < 0.001$). Males and females still showed increases in locomotion for the back of the chamber on the Improving schedule (male and female Improving back > male and female Worsening back, $p < 0.001$), but only males increased their locomotor activity to the front of the chamber on the Worsening schedule (male Worsening front > male Improving front, $p < 0.001$). Males and females continued to suppress locomotor activity to the front of the chamber on the Improving schedule (male and female Improving back > male and male and female Improving front), but only males demonstrated more interest for the front of the chamber on the Worsening schedule compared to the Improving schedule (male Worsening front > male Improving front, $p < 0.001$). Males made more beam breaks than females on the Worsening schedule in the front of the chamber (male Worsening front > female Worsening front, $p = 0.002$). Throughout discounting, locomotor activity patterns remained similar despite choice behavior mice constantly adjusting choice behavior throughout renditions of the task. This suggested locomotor activity was not indicative of learning in delay discounting.

In probability discounting, despite constant changes in choice preference there were stable locomotor behaviors across the tasks over the entire testing period, arguing against locomotion as the driver of behavioral change. This was apparent in the first round (**Figure 2.8A**, main effect of break location, $F(1, 446.09) = 543.83$, $p < 0.001$). There was some suggestion that mice were sensitive to schedule type again (**Figure 2.8A**, main effect of schedule, $F(1, 446.09) = 11.54$, $p < 0.001$), but instead mice experienced general increases in locomotor behavior towards the touchscreen (**Figure**

2.8A, sex x break location interaction, $F(1, 446.09) = 11.54$, $p < 0.001$; male and female front > male and female back, $p < 0.001$). There was no longer a distinction between schedules depending on where the breaks were made.

Mice continued to commit a similar pattern of locomotor changes during Probability Discounting II (**Figure 2.8B**, main effect of break location, $F(1, 450.02) = 714.08$, $p < 0.001$). This general increase in locomotor activity again does not give insight into choice, which was influenced by schedule effects (i.e. anchoring). Mice increased locomotor behavior towards the touchscreen (**Figure 2.8B**, sex x break location interaction, $F(1, 450.02) = 16.58$, $p < 0.001$; male and female front > male and female back, $p < 0.001$). Mice generally chose to explore more of the front of the chamber compared to the back of the chamber. Probability Discounting III was characterized by a similar increase in locomotion toward the touchscreen (**Figure 2.8D**, main effect of break location, $F(1, 917.99) = 2718.16$, $p < 0.001$). Anchoring at this point was still present. Schedule specific difficulties in choice are not represented in locomotor data. Mice directed behavior towards the touchscreen (**Figure 2.8D**, sex x break location interaction, $F(1, 917.99) = 244.96$, $p < 0.001$; sex x schedule x break location interaction, $F(1, 917.99) = 7.00$, $p = 0.0083$; male and female front > male and female back, $p < 0.001$). Male mice shifted their front of the chamber locomotor preference to the Worsening schedule (male Worsening front > male Improving front, $p < 0.001$). This increase in locomotion for the males on the Worsening schedule was now greater than locomotion for the females on the Worsening schedule (male Worsening front > female Worsening front, $p = 0.0073$). The volatility or probability discounting reflected in our choice data and computational modeling are not captured by our locomotor activity. Locomotor activity again suggests a consistent pattern of movement in the chamber that is not reflective of the choice data.

2.3 Discussion

We trained male and female mice in a novel battery of delay and probability discounting schedules, in order to assess 1) if mice exhibit stable choice behavior across these tasks, and 2) if sex differences affected stabilization of choice behavior. These tasks are high-priority goals for cross-species translation, and there is some controversy over whether these two tasks test similar or distinct constructs. Overall, we found mice showed substantial discounting, indicating sensitivity to the structure of both tasks. Mice formed stable choice behavior in delay discounting over time, but continued to show anchoring effects throughout probability discounting. To understand why there might be differences in the persistence of anchoring effects between these tasks, we examined win-stay/lose-shift strategies in probability discounting, and across both tasks, fit discounting models to animal data that included both value parameters (k/h) and an inverse temperature parameter to capture decision noise (β). Win-stay/lose-shift analysis hinted at the presence of schedule-dependent shifts in choice, similar to the anchoring effects prominent throughout probability discounting. Discounting models revealed that mice learned throughout training on both tasks to reduce decision noise. However, the volatility of probability discounting led mice to never really solidify one strategy for the task. Choice response times tracked decision conflict and learning throughout delay and probability discounting, corroborating our findings that changes in choices reflected noisy or exploratory decision making. Our results indicate exploratory decision noise may be underappreciated contributors to behavior in animal models in reward-guided decision making tasks.

Human discounting analyses use value-based models (i.e. k and h values) to determine the extent of discounting behavior according to the value of the reward. Discounting research has demonstrated how hyperbolic models of value best explain discounting behavior of both humans and animals (Green et al., 2014; Vanderveldt et al., 2016). Discounting steepness is believed to follow a value rule that is liable to

change in response to individual factors such as risk tolerance and choice impulsivity (Odum, 2011a; Simon et al., 2009). Our results are in line with previous research where modeling around a value parameter describes stable discounting behavior (Green et al., 2014; Odum, 2011b). However, recent research into animal discounting shows that animals do not always optimize for the discounted value of the reward and implies discounting can arise from multiple individual sources (Blanchard et al., 2013; Hayden & Niv, 2020). In our data, adding an inverse temperature (β) parameter substantially improved our model fit. There were distinct differences for the role of value in delay and probability discounting. Adjustments in observed behavior in delay discounting were in part captured by experience-linked changes in the value parameter *k*, suggesting that as animals shifted their choices for the large reward over training, they were in fact shifting in their *preference* for that large reward. The same could not be said in probability discounting. Instead, in probability discounting, the value parameter *h* was relatively stable, suggesting that preferences for the large uncertain reward were stable, and that variation in behavior was instead captured by changes in the inverse temperature parameter, suggesting that animals were changing *how much they adhered to their value estimates* when making choices over experience with the task. Our modeling results are in line with our observations of persistent anchoring effects in probability discounting even after extended training, which we interpret as being due to inconsistent adherence to value estimates. These findings are reflective of the volatility with trial by trial reward uncertainty. Our choice to add an inverse temperature parameter to capture decision noise allowed us to better capture adaptations to schedules across tasks, and provide an explanation for differences in choice preference for both tasks. This suggests that decision noise is an underappreciated contributor to value-based decisions in animal models.

Sex effects have not been consistently observed in reward-guided delay or probability discounting tasks (Grissom & Reyes, 2019; Orsini & Setlow, 2017). Sex differences appear in discounting tasks depending on the type of uncertainty or consequences (Orsini et al., 2016; Orsini & Setlow, 2017), but generally do not appear when the risk is the loss of a reward (Grissom & Reyes, 2019). Studies have more recently found sex differences in a Worsening schedule of delay discounting (Hernandez et al., 2020) and in a Worsening schedule of probability discounting (Islas-Preciado et al., 2020) in rats. Previous studies however have noted the possible nuances when relating discounting results across species (Vanderveldt et al., 2016) and strains of rats (Islas-Preciado et al., 2020). Supporting the idea of the importance of genetic background, Bagley et al. (2022) have recently demonstrated sex effects in reward guided behaviors are not consistent across mouse strains. Instead, it is genotype background that dictates the strength and direction of a sex effect (sex by genotype interaction). In line with this thought, strain differences in delay discounting have previously been found (Helms et al., 2006; Isles et al., 2004). In the current set of data, animals were able to learn a consistent pattern of behavior to delay regardless of the order of presentation, but probability discounting did not lead to a consistent pattern of choices across sexes. Despite differences in the degree of anchoring across probabilities seen in males versus females in our choice data (Figure 2.2), we were unable to capture those effects in our computational model (Figure 2.4). Given these findings, we consider three possibilities for why we do not observe sex differences. First, species and/or strain appear to affect the strength of the sex difference, and may play a role here. Second, it is possible that we were underpowered to detect sex differences in decision making given that our lab has observed it in the past (Chen et al., 2021). Third, sex differences might be better captured with an additional latent variable not defined within our model.

While delay and probability discounting are often both thought to measure an aspect of choice impulsivity (Dalley et al., 2011; Green & Myerson, 2004; MacKillop, 2016; Strickland & Johnson, 2021), they are found to be weakly correlated even in humans (Green et al., 2014; Strickland & Johnson, 2021). As such, straightforward value models may miss hidden traits specific to delay or risk. In attempting to model these tasks in rodents, one source of variability in reward preference could arise from differences in choice patterns outside of the optimal choice. We included a decision noise parameter in order to capture decision noise hidden in the value parameters (Nussenbaum & Hartley, 2019). Our results suggest mice not only make value-based decisions, but they also adapt decision noise around the discounted value of a reward. Differences in choices across schedules are better explained by changes in decision noise as opposed to the value parameter. Decision noise is therefore a significant contributor to discounting behavior in our mice, and may be an underrecognized contributor to rodent choice behavior in other contexts.

While we designed our tasks to capture adaptations in value and decision noise, some caveats come with the sequential design of our tasks. Delay and probability discounting rates are significantly impacted by the order of presentation of delayed or probabilistic uncertainties (Craig et al., 2014; Fox et al., 2008; Slezak & Anderson, 2009) and could also be affected by whether animals experience delay or probability discounting first (Rung et al., 2019; St Onge & Floresco, 2010; Tanno et al., 2014). Fox et al. (2008) exposed rats to delays in a Worsening-Improving order similar to what the current set of experiments and found similar anchoring results between schedules. Rats were prone to prefer the large reward when the delays got longer within-session, producing similar anchoring effects as we describe here. Slezak & Anderson (2009) however used a random-chance delay exposure paradigm which constantly exposed rats to either order depending on chance. Under this mixed exposure design, schedule

differences were mitigated. It is worth noting that our mice did stabilize their performance in delay discounting, despite experiencing two different orders of delay. This could be due to differences in the length of training compared to previous studies like @Fox2008-oe. The order of exposure to delay and probability discounting could account for some of the persistent anchoring we observe with probability discounting. Our delay and probability discounting tasks are structurally similar and thus delay discounting strategies probably influenced choice behavior during at least early probability discounting (Neville et al., 2020). We did, however, use a novel house light cue introduced during probability discounting to help mice distinguish between both tasks. Further, each task promotes different types of uncertainty (Garr et al., 2020) which provides an opportunity for new learning. Our choice data and computational results support these ideas as mice showed differences in adaptation across both tasks. Still, we cannot fully rule out order effects, but this seems less likely given that our behavior and modeling parameter fits indicate that animals continued to adjust their behavior across the duration of the probability discounting task.

Our results support a role for computational modeling in identifying latent variables that contribute to decisions in rodent tasks. As noted above, while k and h parameters can be used to reflect value in discounting tasks, they are not able to capture the contributions of other variables that might influence choices. Probabilistic tasks, including discounting, are amenable to analyses of choice patterns following wins and losses (e.g. win-stay and lose-shift) but lack a parallel in delay discounting. A global parameter to compare decision noise across tasks is important when assessing choice behavior and testing for factors contributing toward impulsive behavior (Dalley et al., 2011; MacKillop et al., 2016; Strickland & Johnson, 2021). The inverse temperature parameter we included in our computational model helps bridge the gap between the two tasks by allowing for task comparisons and examining influences of

order presentation and group differences in valuation of rewards. Our results demonstrate how computational models that account for decision noise are better at detecting different sources of behavioral variability, such as anchoring effects, demonstrated in sequential versions of decision making tasks.

2.4 Materials and methods

2.4.1 Subjects.

8 male and 7 female BL6129SF1/J mice (from Jackson Laboratories) took part in both delay discounting and probability discounting. Mice were housed in groups of 3-4 (2 groups of males, 2 groups of females). Mice began experiments at approximately 70 days of age. Mice had free access to water and were food restricted with their home chow at 85-90% of their baseline weight. Mice were pre-exposed to the operant reinforcer, vanilla flavored Ensure, in their home cage for one day prior to training. Ensure was freely available to be licked from a bottle. Each group of mice were verified to have consumed a full bottle of Ensure (148 ml). Behavioral testing took place Monday to Friday, and on Fridays, mice had free access to home chow. Animals were housed on a reverse light dark cycle (9am-11pm) and were tested during the dark period. Animals were cared for in accordance with National Institute of Health guidelines and were approved by the University of Minnesota Institutional Animal Care and Use Committee.

2.4.2 Apparatus.

16 identical triangular touchscreen operant chambers (Lafayette Instrument Co., Lafayette, IN) were used for training and testing. The touchscreen was housed in the front while the food delivery magazine in the back. Information on individual touches on touchscreens throughout sessions were recorded via ABET-II software. Touchscreens were limited by masks with holes which allowed responding in 5 square

holes. Liquid reinforcer (50% diluted Ensure) was pumped via a peristaltic pump (1000 ms or 250 ms duration, corresponding to volumes of approximately 25 μl and 6.25 μl). ABET-II software (Lafayette Instrument Co., Lafayette, IN) was used to program operant schedules and analyze all data from training and testing.

2.4.3 Behavioral procedures.

Magazine training. Mice received free 7 μl of Ensure every 30 seconds for 30 minutes in operant chambers for 5 days. Mice learned to approach the magazine to obtain Ensure.

Center Hole Fixed-Ratio 1. Mice were initially trained to nosepoke the center hole of a 5-hole mask on the touchscreen chamber on a Fixed-Ratio 1 schedule for 10 days, 30 minutes each day. 7 μl of Ensure was delivered immediately following a nosepoke. Only the center hole was illuminated during these sessions. The magazine holding the Ensure was illuminated until mice interrupted an infrared beam when their head entered the reward port, and this allowed them to move to the next trial. All mice were moved on to the next phase once they could reach the max number of trials needed to complete delay discounting schedules (60 trials).

Chaining Center to Left and Right. For this phase of training, hole 3 (center) illuminated and a nosepoke resulted in 7 μl of Ensure. After a center nosepoke and its reward, on the next trial, holes 2 (left) and 4 (right) illuminated, and mice learned that nosepoking either side resulted in a large amount (28 μl) of Ensure. Training lasted 29 days with 30 minute sessions. Responses on holes 1 and 5 were counted as non-reinforced touches. Mice were moved on to the next schedule when they could consistently complete over 60 trials.

Responding on Sides Only. Mice learned to nosepoke only the left and right holes for Ensure at the same volumes as above for 13 days. The left and right holes were the only holes to light up during these sessions. Center hole nosepokes no longer delivered the reinforcer. Sessions ended after 30 minutes had elapsed. Mice were moved on to the next schedule when they could complete 60 trials.

Responding on Chained Sides Followed by ITI. Mice then learned to chain a center hole nosepoke to a left or right hole nosepoke for 13 days. The action sequence of center-to-left or center-to-right led to a large reward (25 μl) or small reward (6.25 μl). This phase of the testing was counterbalanced; half of the mice experienced the large reward on the left and the other half on the right. An inter-trial interval (ITI) of 30 seconds followed in order to suppress the reward rate. Animals were limited to no more than two trials in a row selecting one side before being forced to try the other side, to ensure they experienced the small reward side as well as the large reward. Sessions ended after 30 minutes had passed. Mice were moved on to the next schedule when they could complete 60 trials.

Improving and Worsening Delay Discounting. To test the influence of anchoring effects, we tested mice on a Worsening schedule that was at the start of the session was initially favorable and became unfavorable as the session proceeded, then reversed the order of delays for the Improving schedule, then reversed again 4 additional times. We used the criteria set by (Mar & Robbins, 2007) to judge if mice were ready to move on from the initial discounting round. We ran mice for extended periods of time (≥ 9 sessions) in order to ensure behavior had stabilized between rounds. Mice underwent 10 days of Worsening Delay Discounting I followed by 9 days of Improving Delay Discounting I, 9 days of Worsening Delay Discounting II and 11 days of Improving Delay Discounting II, and 18 days of Worsening Delay Discounting III and 18 days of Improving Delay Discounting III. One side delivered a large but

delayed reward (25 μl) or a small and immediate reward (6.25 μl). Each delay block of trials began with 2 forced trials where mice had to choose the large delayed side in order to be reminded of the delay. Mice then had 10 free choice trials. The side with the large reward was matched to the chained side training. If mice responded for the small reward, an ITI occurred based on the delay for that session. If mice chose the large reward, the center hole blinked for the duration of the delay. On the Worsening schedule, mice experienced 12 trials of increasing delays – 0s, 4s, 12s, 20s, 28s delays – within one session. The Improving schedule was similar but in the reverse orientation – 28s, 20s, 12s, 4s, 0s delays – within one session. Mice could only move on to the next delay if they responded on all 12 trials for that delay and collected the reward. In order to track progression of delay discounting stability, we have included subplots with fewer days according to the start and end of training (day matched with the last round of training and probability discounting).

Worsening and Improving Probability Discounting. Animals transitioned directly from the last delay discounting schedule to the first probability discounting schedule because the structure of the task (location of large and small reward, response order) did not change, only the rule governing payout of the large reward. Mice underwent 10 days of Worsening Probability Discounting I and followed by 8 days of Improving Probability Discounting I, 9 days of Worsening Probability Discounting II and 9 days of Improving Probability Discounting II, and 16 days of Worsening Probability Discounting III and 18 days of Improving Probability Discounting III. As in delay discounting, the same side resulted in a large reward or small reward. Both rewards were immediate, but the large reward was delivered probabilistically. All probability blocks consisted of free choice trials. On the Worsening schedule, mice experienced 20 trials of decreasing probability of reward delivery – 100%, 75%, 50%, 25%, 12.5% chance of reward – within one session. The Improving schedule was

similar but in the reverse orientation – 12.5%, 25%, 50%, 75%, 100% probabilities – within one day. If the trial was not rewarded, a feedback house light would blink indicating a non-rewarded trial. We used the house light instead of the screen in order to help mice distinguish between discounting tasks (i.e. delay discounting used screen to signal delay, probability discounting used house light to signal reward loss). Mice could only move on to the next probability if they responded to all 20 trials for that probability and collected the reward. In order to track progression of probability discounting stability, we have included subplots with fewer days according to the start and end of training (day matched with the last round of training and delay discounting).

2.4.4 Computational modeling.

To quantitatively examine how the value of rewards vary as a function of delay, we fit an exponential discounting model (Odum, 2011a), shown as in the equation below:

$$V = Ae^{-kD},$$

where *V* is the subjective or discounted value of the delayed reward, *A* is the amount or magnitude of the delayed reward, and *D* is the length of delay. *k* is a free parameter that reflects the discounting rate: the larger *k* is, the steeper the discounting of reward value; the smaller *k* is, the slower the discounting of reward value. *k* is determined by the fit of the model to the actual data.

Then, we fit a similar exponential model (Richards et al., 1999) to examine how uncertainty of reward affects the value of reward.

$$V = Ae^{-h(\frac{1}{P}-1)},$$

In this model, *p* is the probability of reward and *A* is the magnitude of the reward. The free parameter *h* dictates how steep the change in the value of reward is as a function of the reward probability. Thus, *h* in the probability discounting model and *k* in the above delay discounting model both describe how rapidly the value of a reward is discounted, either by uncertainty or delay of the reward.

For both models, the action selection was performed based on a Softmax probability distribution:

$$P(A) = \frac{e^{V_A * \beta}}{\sum_j e^{V_j * \beta}}$$

Where V_A corresponds to the subjective reward value of action A, and a second free parameter inverse temperature β determines the level of decision noise. When inverse temperature is high, the decision noise is low, which means more exploitation of the action with high subjective value; when inverse temperature is low, the decision noise is large, which means more random exploration regardless of value. The optimized parameters were obtained through minimization of the negative log likelihood of the models.

2.4.5 Statistical analysis.

Delay discounting and probability discounting data were analyzed using R Studio using the lme4 package. Linear mixed models were fit to preference proportion data for both delay discounting and probability discounting with fixed effects of sex (males and females), schedule (Improving and Worsening), delay/probability and random effect for subjects (**Figures 2.1 & 2.2**). Win-stay scores were calculated by taking the number of times animals stayed on the large risky side after receiving Ensure on the previous trial divided by the total number of times animals received Ensure on the large side. Lose-shift scores were calculated by taking the number of times animals did not

receive Ensure on the large side and switched to the small certain side divided by the total number of times animals did not receive rewards on the large side (regardless if they shifted or not). These scores were then analyzed using linear mixed models with fixed factors of sex and schedule and random effect for subjects (**Figure 2.3**). Computational modeling results were fit to discounting and inverse temperature parameters (**Figure 2.4**) and were analyzed with linear-mixed models with task and schedule as the fixed factors and random effects for subjects.

Response time for small and large choice data were fit with the same factors (**Figure 2.5 & 2.6**). Response times greater than 3 standard deviations away from the mean were eliminated from analysis. Beam break data were fit with fixed factors of sex, schedule, and break location (front and back) and random effect for subjects (**Figure 2.7 & 2.8**). 2 mice (1 male, 1 female) were excluded from beam break analysis because of a malfunction in infrared beams. Beam breaks greater than 3 standard deviations aways from the mean were eliminated. In cases where sex was not significant based on a log-likelihood ratio test, sex was removed from the linear mixed models. Data were fit for all days in the main plots and for days 1-3 and days 6-8 in the subplots in order to track the progression of discounting stability. For all statistics, an alpha of 0.05 was used.

2. 5 Figures

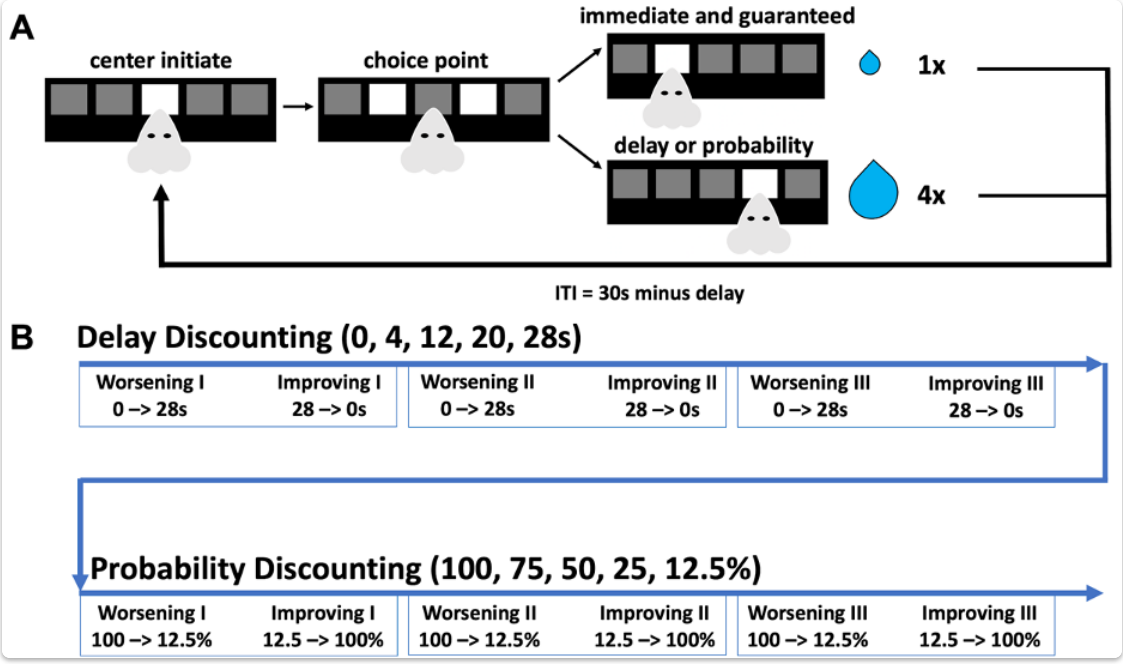

Figure 2.0 Schematic of delay discounting and probability discounting.

A. Mice responded for large and small rewards after initiating the trial by poking the center hole. Mice then chose one side for a small immediate reward (6.25 μl) or a large delayed/probabilistic reward (25 μl). Small choices on the delay schedule were followed by an ITI to time match the delay of the large reward. **B.** Mice began discounting tasks on the Worsening schedule (i.e. increasing delay) for 10 days and then were switched to the Improving schedule (i.e. decreasing delay) for 9 days. Schedules continued to switch two more times. Mice then experienced probability discounting on the Worsening schedule (i.e. increasing uncertainty) for 10 days and then were switched to the Improving schedule (i.e. decreasing uncertainty) for 8 days. Schedules again switched two more times.

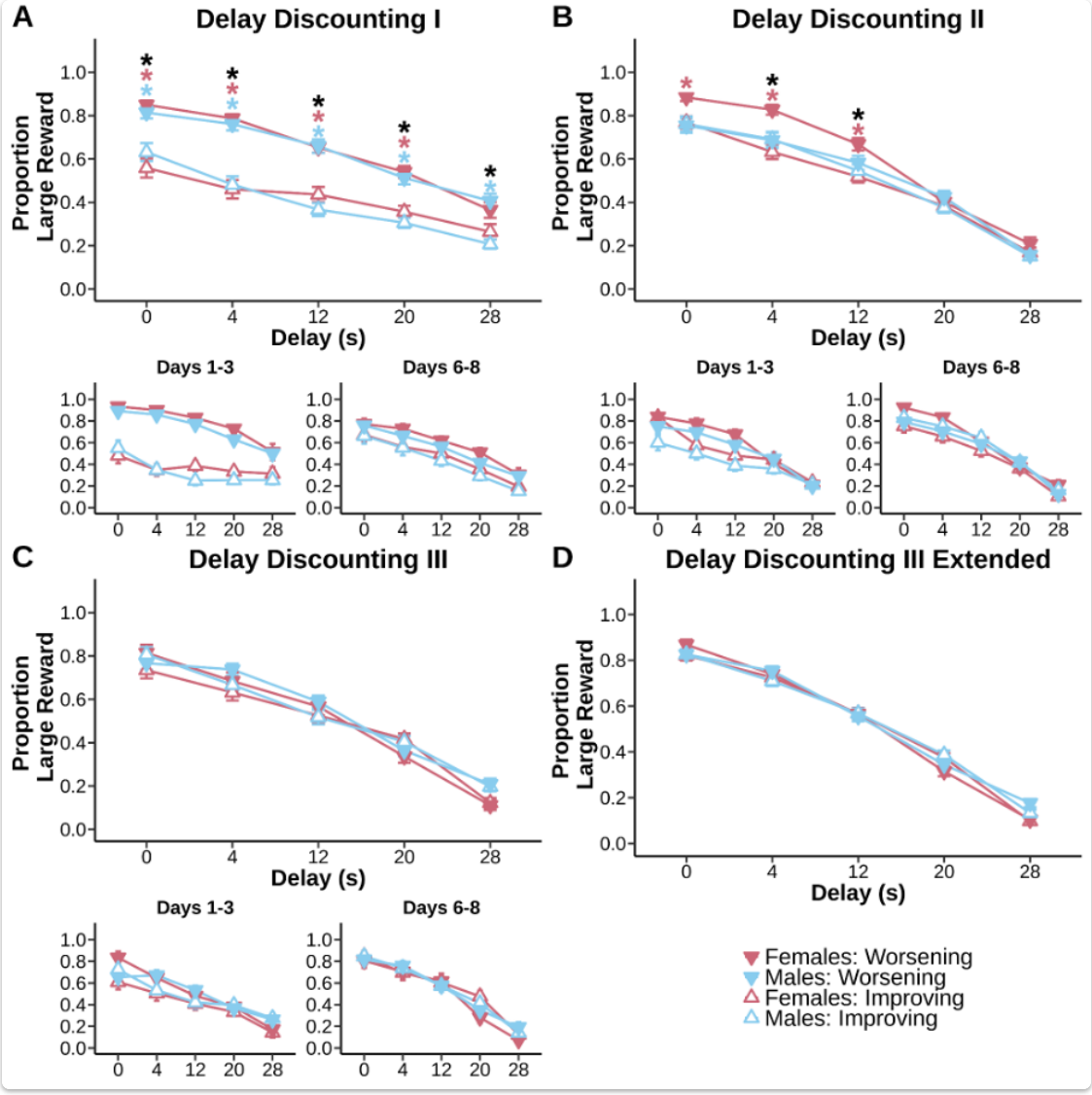

Figure 2.1 Anchoring effects in delay discounting are eliminated with extended experience.

Main plots represent a summary of all days of training. Subplots represent day matched periods of training. **A:** Mice of both sexes responded less for the large reward at the longest delays compared to the shortest, even with minimal experience on the tasks. However, there was a significant anchoring effect, such that mice in the Improving (started at 28s delay) condition had a persistent reduction in choosing the large reward compared to the Worsening (started at 0s delay) condition. Anchoring was persistent from the beginning to the last days of training as indicated by schedule x

delay effects (Worsening > Improving: 0s, $p < 0.001$; 4s, $p < 0.001$; 12s, $p < 0.001$; 20s, $p < 0.001$; 28s, $p = 0.0020$). **B:** Anchoring effects lingered into the second session of delay discounting (Worsening > Improving: 4s, $p = 0.0021$; 12s, $p = 0.0025$), but were most pronounced in females. Anchoring became more apparent during the last days of training in females. **C&D:** Data in figure C represent the first 8 days, data in figure D represent all days of training. With continued experience, females and males reached similar discounting rates and no longer anchored their preference according to whether sessions started with a long or short delay. Anchoring was present early into training but disappeared with continued experience. Figures depict mean ± SEM, black asterisks indicate significant anchoring effects (schedule x delay interactions) while colored asterisks indicate significant sex effects (planned post hoc comparisons of Worsening and Improving schedules within a sex at each timepoint) of $p < 0.05$.

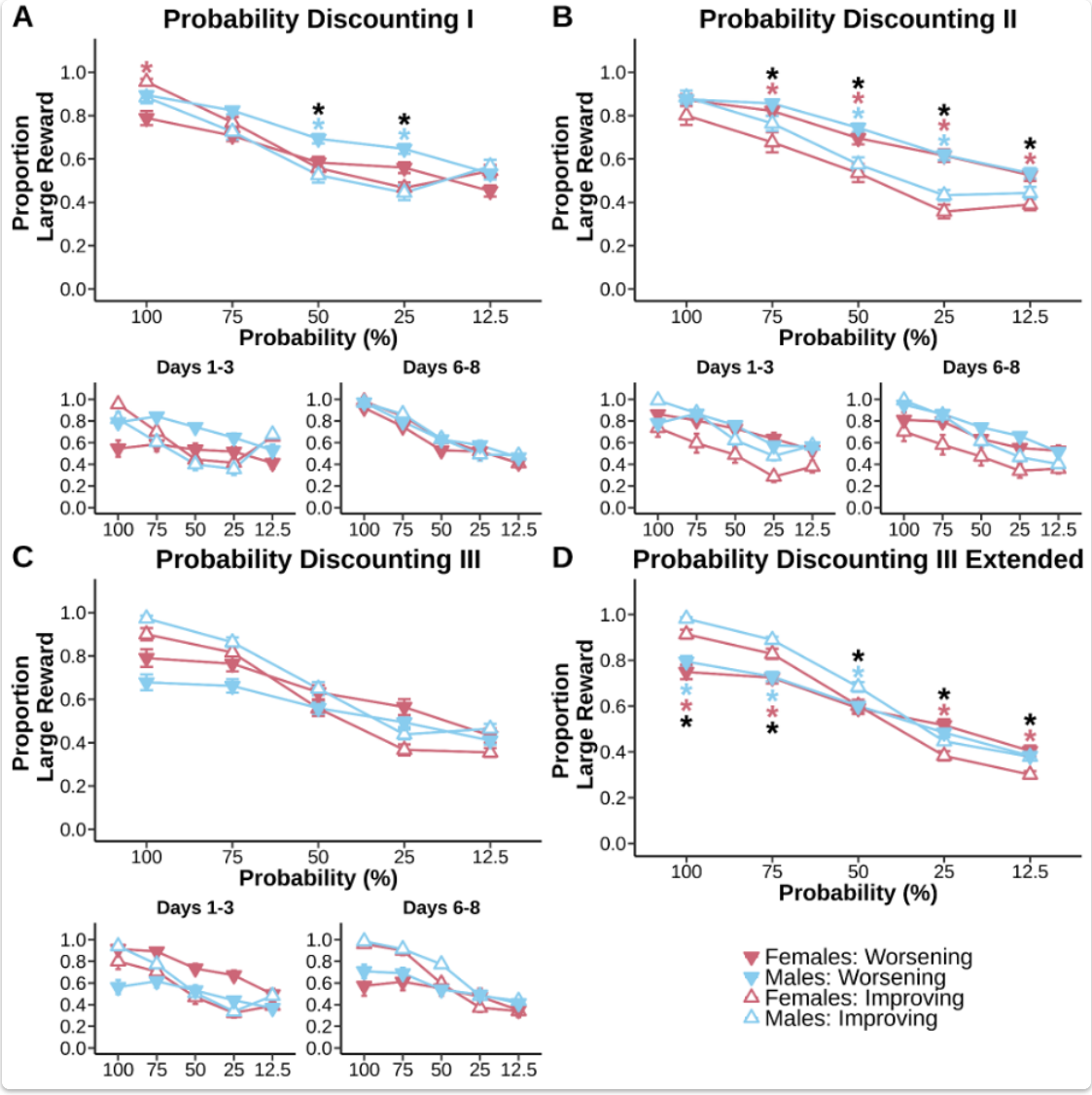

Figure 2.2 Anchoring effects continue to influence probability discounting behavior after extended experience.

Main plots represent a summary of all days of training. Subplots represent day matched periods of training. **A:** Regardless of uncertainty orientation, males and females choose the large reward less often with increasing risk. Interaction effects of schedule x probability (Worsening > Improving) were specific to the 25% ($p = 0.001$) and 50% ($p = 0.0210$) probabilities. **B:** For the second round of probability discounting, mice showed significant discounting of rewards and evidence of anchoring. Interaction effects of schedule x probability (Worsening > Improving) were

specific to the 12.5% ($p < 0.001$), 25% ($p < 0.001$), 50% ($p < 0.001$), and 75% ($p < 0.001$) probabilities. Females were sensitive to uncertainty throughout the session (Worsening > Improving) whereas males were only sensitive at the 50% and 25% probabilities (Worsening > Improving). **C&D:** Data in figure C represent the first 8 days, data in figure D represent all days of training. With extended training, mice were sensitive to schedule effects at all probabilities where large choice preference was greater on the Worsening schedule than the Improving schedule at probabilities below 50% chance (12.5%: $p = 0.0230$, 25%: $p < 0.001$), but switched to a greater large choice preference on the Improving schedule compared to the Worsening schedule when probabilities were at or above 50% chance (50%: $p = 0.0410$, 75%: $p < 0.001$, and 100%: $p < 0.001$). Females showed decreased risky choice preference on the Improving schedule at risky probabilities, but increased risky choice preference on the Improving schedule at safe probabilities. Male mice made schedule specific adjustments only at safer probabilities (50% probability and higher). Figures depict mean ± SEM, black asterisks indicate significant anchoring effects (schedule x probability interactions) while colored asterisks indicate significant differences within a sex in responding to specific probabilities (planned post hoc comparisons) of $p < 0.05$.

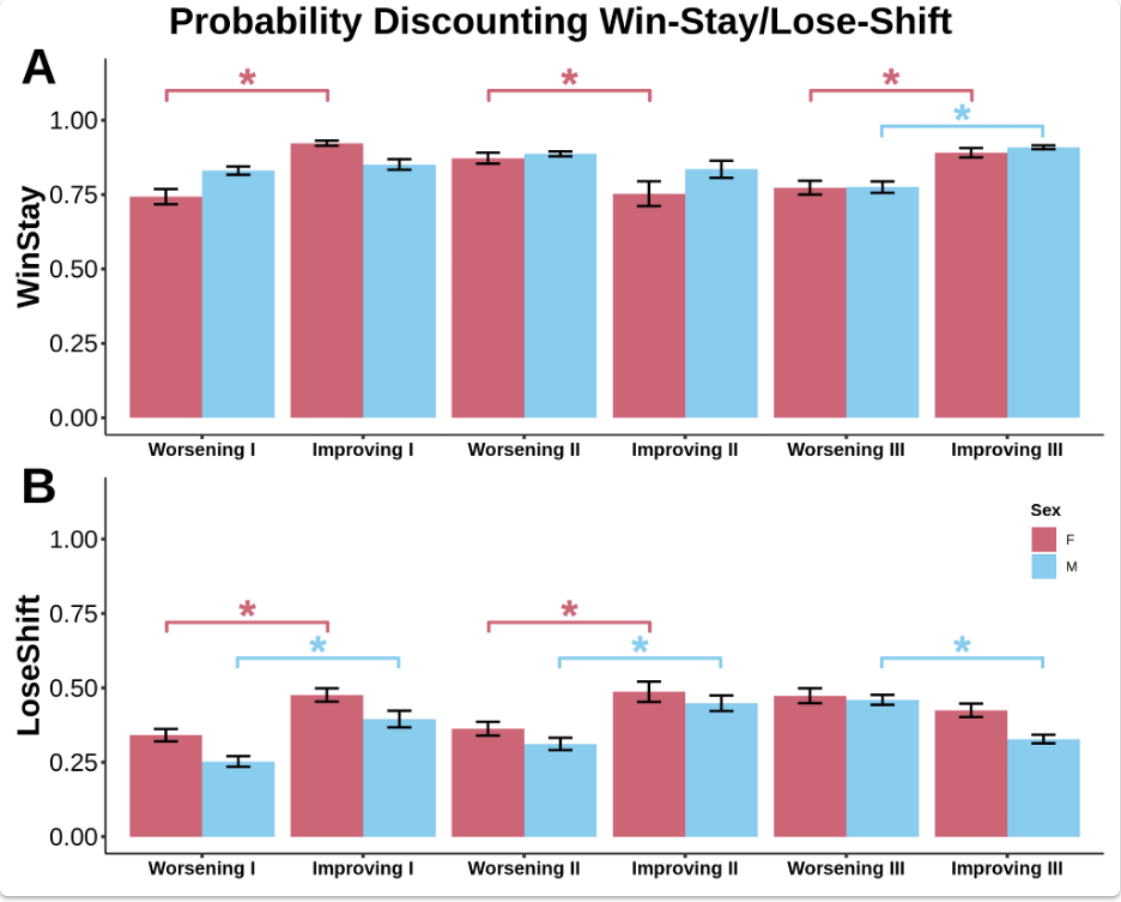

Figure 2.3 Adaptations to wins and losses with experience.

We examined whether tendencies to win-stay/lose-shift explained sex differences in probabilistic responding. Win-stay ratios were calculated by dividing how often mice stayed on the same risky side after being rewarded divided by all rewarded risky responses. Lose-shift ratios were calculated by dividing how often mice switched to the small guaranteed side after not receiving a large risky reward divided by all losses on the risky side. **A:** Female mice specifically learned to increase win-stay behavior after their first experience with risky rewards. Females continuously make adjustments to win-stay behavior throughout their exposure to both Worsening and Improving schedules. By the end of probability discounting, males and females learned to make more win-stay choices on the Improving schedule. **B:** Male and female mice adapted to risky rewards by increasing their lose-shift behavior on an initially risky schedule (i.e.

Improving). By the end of probability discounting, only males showed lose-shift specific adaptations. Figures depict mean ± SEM, colored asterisks indicate significant within sex schedule effects (planned post hoc comparisons) of $p < 0.05$.

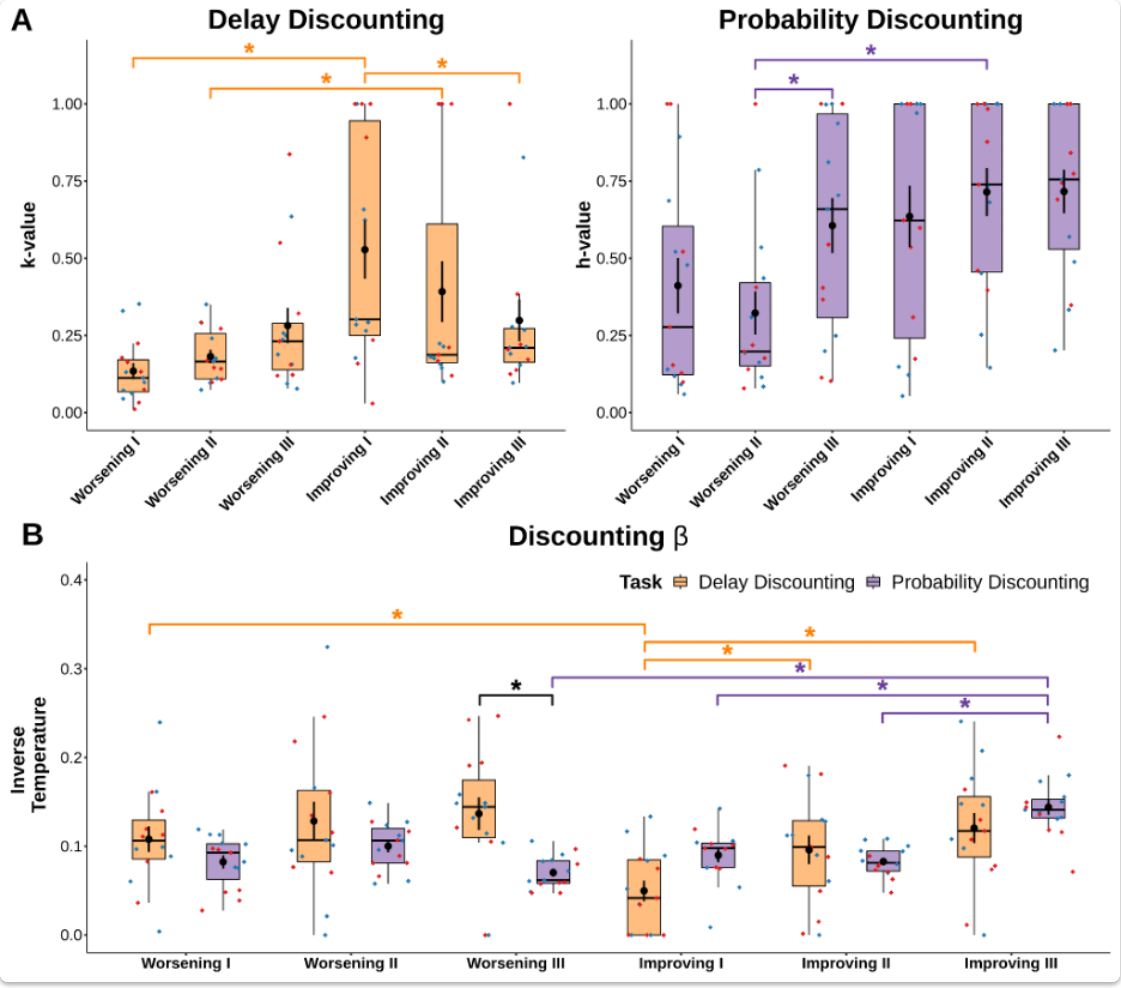

Figure 2.4 Discounting as explained by value and decision noise.

We modeled discounting rates (*k* and *h*) and decision noise (inverse temperature, β) for each mouse across all days of testing. **A:** Mice alter *k* as they continuously adapt to differences in delay order. Mice quickly converge on a similar *h* value despite changes in probability order. **B:** Inverse temperature (β, reflecting decision noise) parameters for delay and probability discounting reveal large changes in the noisiness of choices as schedules change. Inverse temperature increased over time with extended

experience with both tasks (especially on Improving schedules), indicating decreased decision noise in choices with experience. Figures depict individual points for males (blue) and females (red) and box-and-whisker plots with mean ± SEM overlaid; asterisks indicate significance of $p < 0.05$.

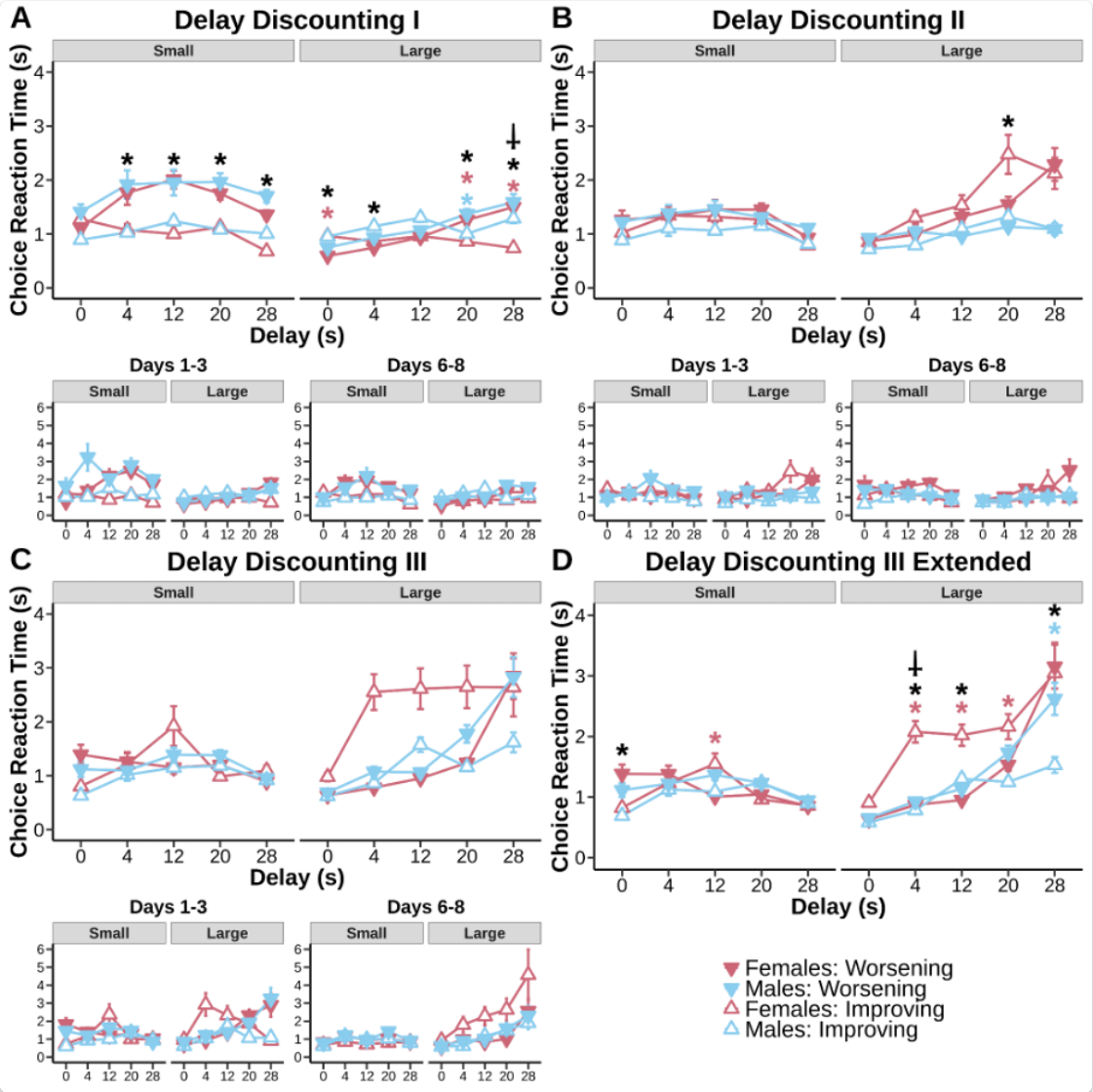

Figure 2.5 Response times become more influenced by delay cost on a choice as animals gain experience with delay discounting.

Main plots represent a summary of all days of training. Subplots represent day matched periods of training. **A:** Mice made slightly slower choice responses for both

large and small rewards when the delays were Worsening relative to when they were Improving. **B:** Mice converged choice response times for small and large choice response times for Delay Discounting II. The clearest delay effect was noticeable at the 20s delay. **C&D:** Data in figure C represent the first 8 days, data in figure D represent all days of training. Extended experience with delay discounting induced schedule effects in small and large choice response times. All mice were slower on the Improving schedule 0s delay, but only female mice were slower on the Improving schedule at the 12s delay. Females slowed down for large choices when the schedule was Improving for the 4s, 12s, and 20s delays. Only males slowed down at the largest delay on the Worsening schedule. Figures depict mean ± SEM. Black asterisks indicate significant schedule effects (schedule x delay interactions), black daggers pointing up indicate between sex effects on the Improving schedule (sex x schedule x delay interactions) of $p < 0.05$. Colored asterisks indicate significant schedule effects within a sex (sex x schedule x delay interactions) of $p < 0.05$.

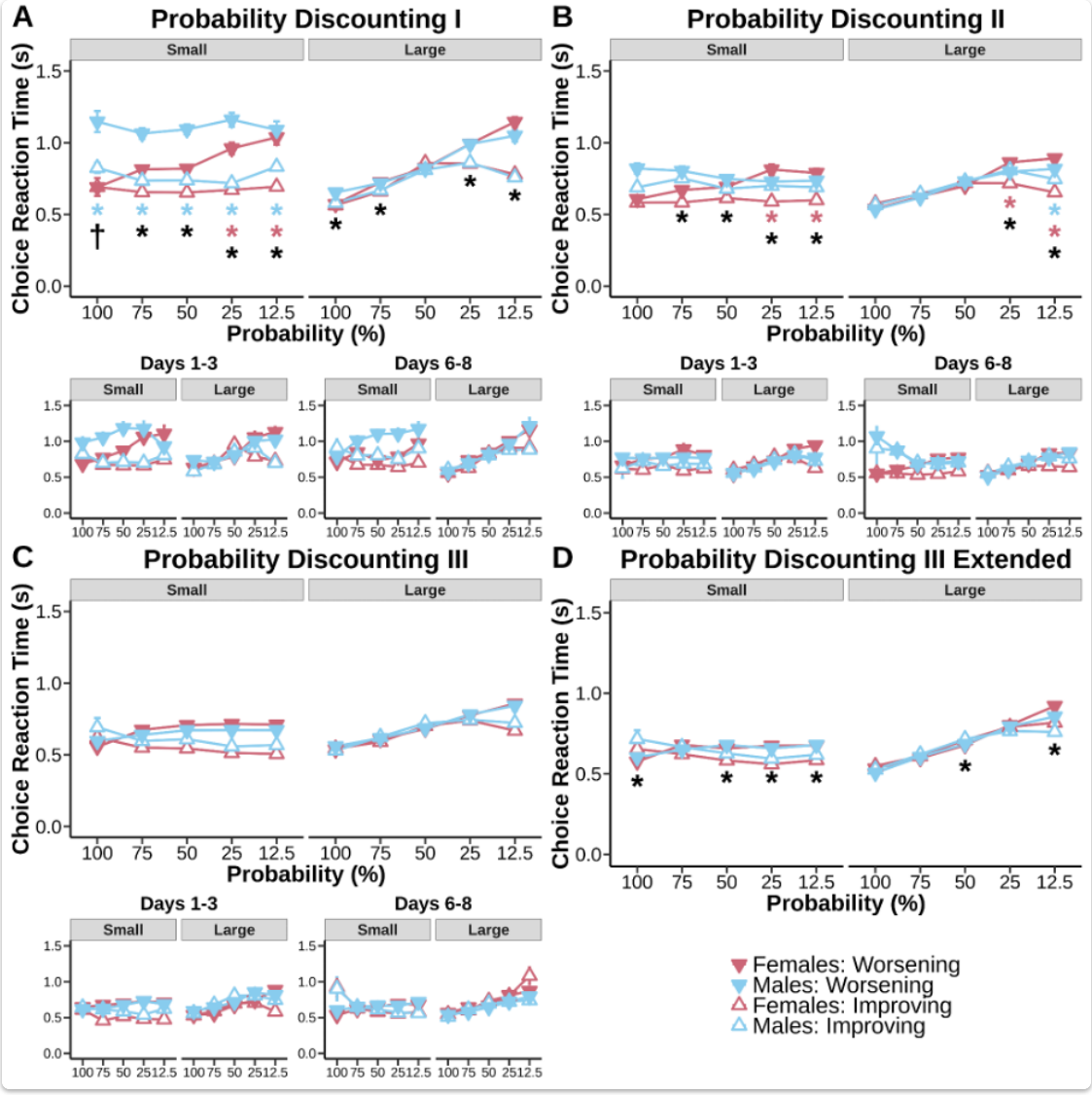

Figure 2.6 Response times somewhat stabilize but remain sensitive to high uncertainty over probability discounting.

Main plots represent a summary of all days of training. Subplots represent day matched periods of training. **A:** As mice first experienced probability discounting, females were slower to respond for small rewards on the Worsening schedule at small probabilities and males were also slower but at all probabilities. Schedule effects appeared for large choice response times across all probabilities except for 50% probability chance of large reward. **B:** Small choice response time schedule effects carried into Probability Discounting II. Mice were consistently slower for Worsening

small choices, but females specifically were slower when the probability for the trial blocks were 12.5% and 25% chance for large reward. Both males and females were slower for Worsening large choices at 12.5% probability, but only females were also sensitive at 25% probability. **C&D:** Data in figure C represent the first 8 days, data in figure D represent all days of training. Sex specific effects at delays diminished with the last round of discounting. General schedule effects at delays were still present. Small choice schedule effects were pervasive for all probabilities of large reward except 75%, large choice differences arose only at 12.5% and 50% probabilities. Figures depict mean ± SEM. Black asterisks indicate significant schedule effects (schedule x delay interactions), black daggers pointing down indicate between sex effects on the Worsening schedule (sex x schedule x delay interactions) of $p < 0.05$. Colored asterisks indicate significant schedule effects within a sex (sex x schedule x delay interactions) of $p < 0.05$.

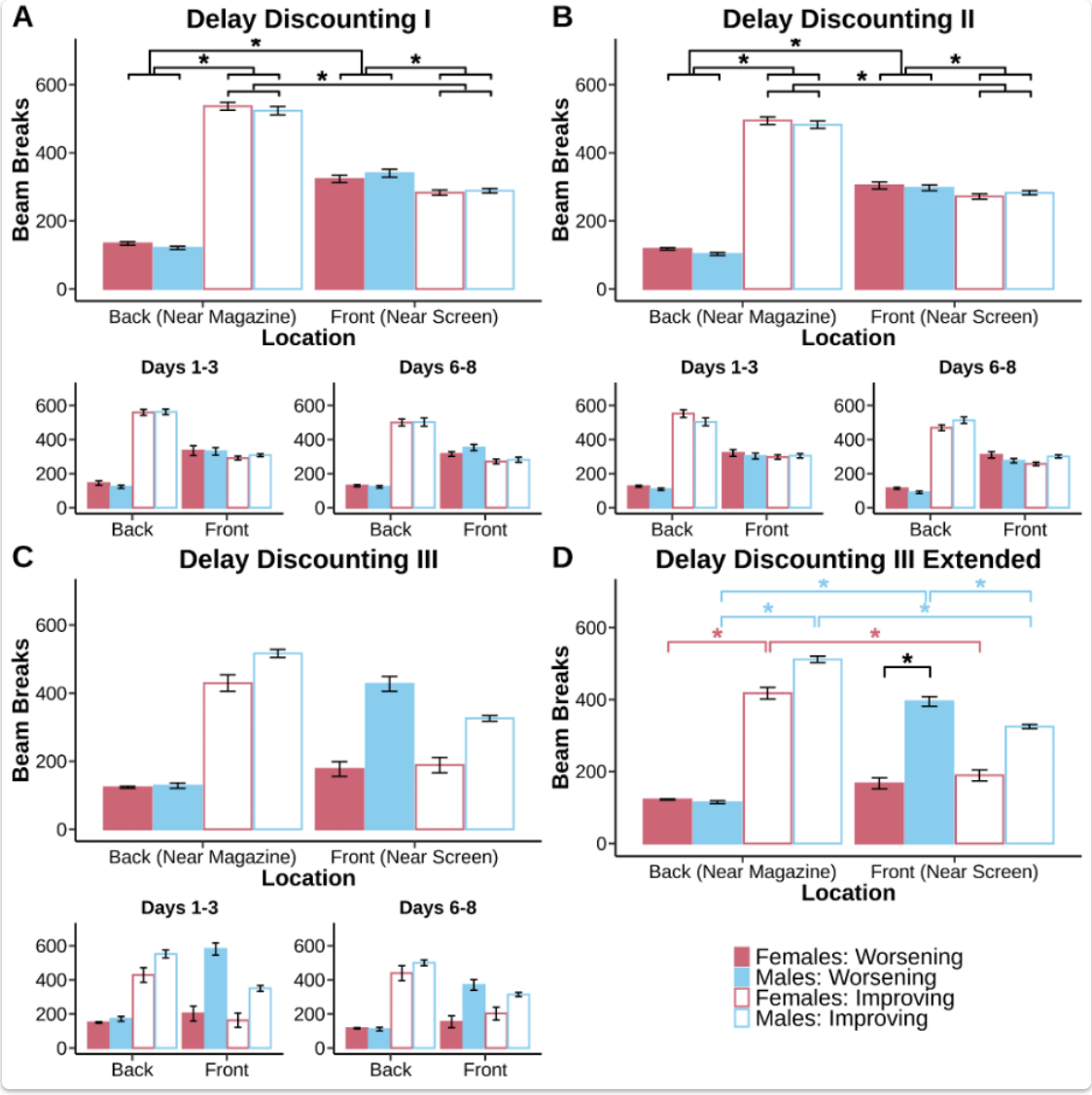

Figure 2.7 Mice adjust locomotor behavior near the magazine and screen according to the delay discounting schedule.

Main plots represent a summary of all days of training. Subplots represent day matched periods of training. Mice broke an infrared beam whenever they crossed to the back of the chamber (near the magazine) or the front of the chamber (near the screen). Shifts in delay discounting schedule orientation changed where mice chose to spend most of their time. **A:** When mice experienced a schedule with increasing delays, exploration of the chamber shifted toward the front of the chamber. However, when the schedule switched orientation to decreasing delays, mice shifted screen preference to

magazine preference. **B:** Mice showed a similar pattern of exploration where mice preferred the front of the chamber when delays were Worsening, but shifted that preference to the back of the chamber when delays were Improving. **C&D:** Extended experience with delay discounting caused females to reduce front of chamber exploration regardless of schedule type. Females maintained increased interest in the back of the chamber on the Improving schedule. Males continued to shift their behavior based on schedule type. Figures depict mean ± SEM. Black asterisks indicate significant schedule x break location interactions for Figure A and Figure B, but a significant sex effect for Figure D of $p < 0.05$. Colored asterisks indicate significant differences within a sex effects (sex x schedule x break location interactions) of $p < 0.05$.

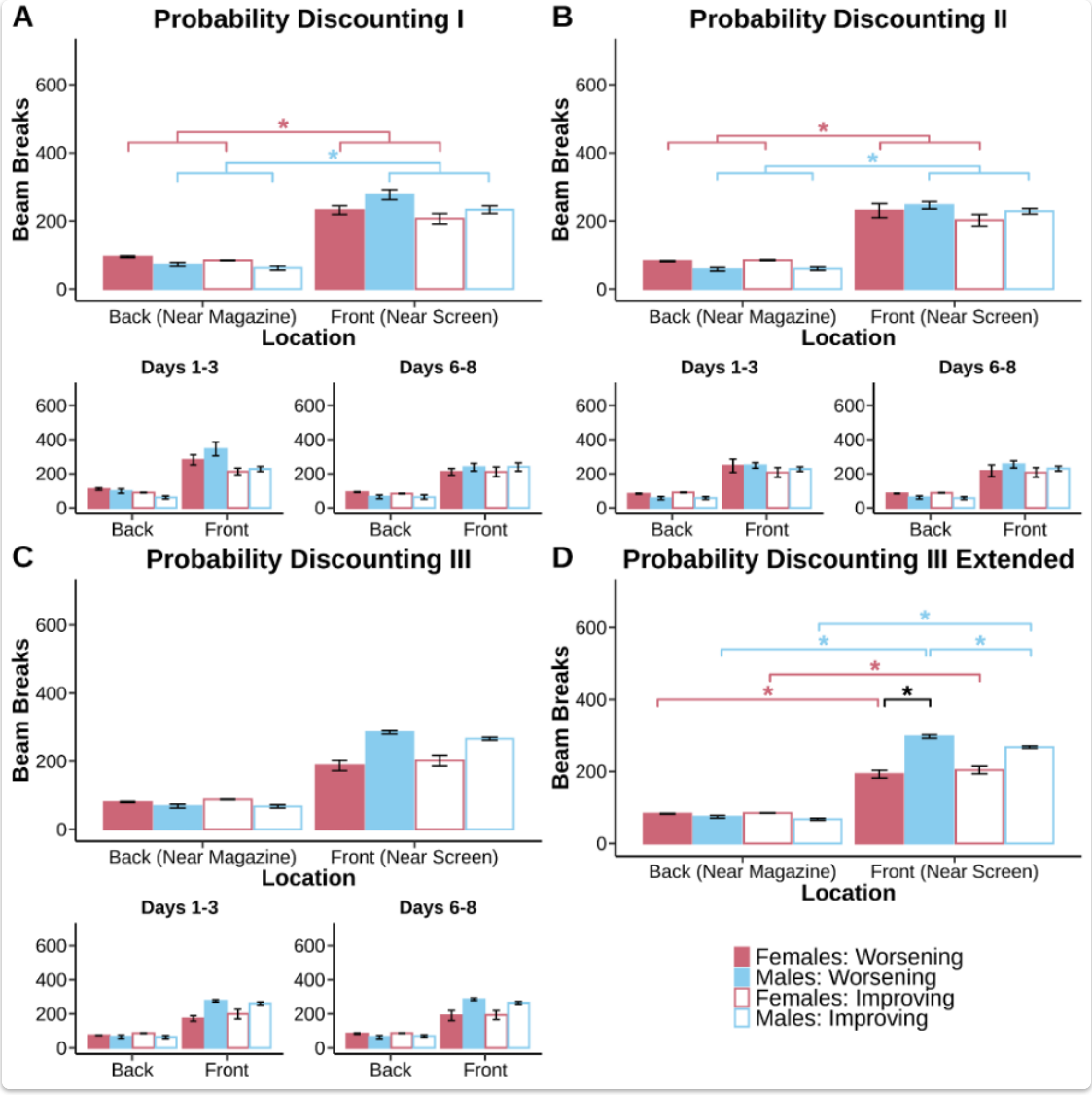

Figure 2.8 Mice prefer to stay near the screen during probability discounting.

Main plots represent a summary of all days of training. Subplots represent day matched periods of training. Movement throughout the chamber is entirely schedule dependent throughout the first round of probability discounting (**A**), the second round of probability discounting (**B**), and the last round of discounting (**C&D**). Male mice however made more beam breaks near the screen on the Worsening schedule compared to females on the Worsening schedule and compared to themselves on the Improving schedule (**C&D**). Figures depict mean ± SEM. Black asterisks indicate significant sex

effects of $p < 0.05$. Colored asterisks indicate significant differences within a sex effects (sex x schedule x break location interactions) of $p < 0.05$.

3. Delay and probability in 16p11.2 hemideletion mice divergently modulates preference

Differential processing of delay versus uncertainty in male but not female 16p11.2 hemideletion mice.

Gerardo R. Rojas, Abigail T. Heller, Nicola M. Grissom

Department of Psychology, University of Minnesota, Minneapolis, MN 55455, USA

Abstract

Neurodevelopmental disorders are associated with differences in learning and motivation that can influence executive function, including behavioral flexibility and decision making. 16p11.2 hemideletion is a chromosomal copy number variant that is linked to neurodevelopmental disorders. 16p11.2 hemideletion in mice has been previously found to produce male-biased changes in reward learning, but the link between this and altered flexible decision making is poorly understood. We challenged 16p11.2 hemideletion mice with two reward-guided decision making tasks assessing flexible decision making under cost, delay and probability discounting. Both tasks elicited long-term changes in flexible decision making that separated 16p11.2 hemideletion males from wildtype males. In delay discounting,16p11.2 hemideletion males had a stronger, less flexible preference for the large reward at long delays, and

this effect was reduced as wildtype males adjusted their preference to match that of the hemideletion males. In probability discounting, 16p11.2 hemideletion males initially had a similar preference for seeking improbable large rewards as did wildtype males, but over time began to prefer certainty to a greater extent than did wildtype males. Female mice discounted similarly for delayed or risky rewards regardless of the presence of the copy number variant. We have previously seen that male 16p11.2 hemideletion mice commit fewer nonreinforced responses than male wildtype mice in an operant setting, which we replicate here in delay discounting, while the introduction of risky rewards eliminates genotype differences in nonreinforced responses. Overall, these data suggest that 16p11.2 hemideletion in males leads to differential processing of costs of delay versus inconsistency, with greater aversion to uncertainty than delays, and greater behavioral control by cues that consistently predict an outcome.

Keywords: Mice, Mouse Model, Neurodevelopmental Disorders, Delay Discounting, Probability Discounting

3.1 Introduction

Understanding how neurodevelopmental disorder-linked genes impact flexible decision making may shed light on the connections between fundamental neurobiology and diversity in cognition. One area of interest has been how neurodevelopmental disorders influence decision-making processes. It has been repeatedly observed that autism spectrum disorder (ASD) symptoms may be partly explained by differential processing of environmental uncertainty (Lawson et al., 2017; Minassian et al., 2007; Sinha et al., 2014). Neurodevelopmental disorders are known to impact fundamental processes that support decision-making such as learning, motivation and attention (Dichter et al., 2012). Studies that have focused on the processing of rewards in ASD find reduced social motivation (Chevallier et al., 2012; Clements et al., 2018; Kohls et

al., 2012), deficits in reward processing (Scott-Van Zeeland et al., 2010) and deficits in specific reward epochs like reward anticipation (Baumeister et al., 2023; Clements et al., 2018). In more complex situations which require probabilistic learning (e.g. Iowa Gambling Task), those with ASD show a broad trend of choices that resembles risk avoidance (South et al., 2014), but when closely examined could represent a difference in strategy (Zeif et al., 2023). This raises the question of whether applying different decision making tasks in the same individuals can reveal greater specificity in whether reward processing, costs, or uncertainty are most central in driving differences in flexible decision making.

Deletion of one copy of chromosomal region 16p11.2 in humans has been linked to diagnosis of autism and attention-deficit hyperactivity disorder (ADHD), and these individuals can exhibit language delays, social communication issues, and motor patterns typical of neurodevelopmental disorders regardless of whether or not specific diagnostic criteria are reached (Hanson et al., 2015; Rein & Yan, 2020; Walsh & Bracken, 2011). In mice, the 16p11.2 region is highly conserved (Horev et al., 2011), and mice modeling this hemideletion show impacts in basal ganglia function (Portmann et al., 2014) and decreases NMDA receptor activity in the prefrontal cortex (Wang et al., 2018), both of which contribute to aspects of decision-making. Neurodevelopmental disorders such as ASD and ADHD are diagnosed at a higher rate in males than females (Loomes et al., 2017; Posserud et al., 2021). While some impacts of 16p11.2 hemideletion, including hyperactivity, are seen across sexes (Angelakos et al., 2017), sex-biased increases in several neurodevelopmental-disorder relevant domains, including sleep disturbances, anxiety-like behaviors, and reward learning alterations have been seen (Angelakos et al., 2017; Giovanniello et al., 2021; Grissom et al., 2018). Collectively, these data suggest that the impact of this copy

number variant may be male-biased, but the extent to which this is true in flexible decision making is unknown.

Recently, we demonstrated mice are able to learn both delay and probability discounting (Rojas et al., 2022). Delay and probability discounting are valuable because they are complementary tasks that challenge animals with temporal or risky costs (Green & Myerson, 2004; McKerchar & Renda, 2012). In order to determine if 16p11.2 hemideletion mice are more sensitive to one type of cost, we tested mice on delay and probability discounting tasks. We had mice first undergo "Worsening" and then "Improving" versions of delay and probability discounting because the acclimation to an Improving schedule (i.e. going from no delay training to large delay testing) can obscure sensitivity to delay in models of neurodevelopmental disorders (Sjoberg et al., 2023). Each schedule promotes a different choice pattern, which gives us the ability to compare whether delay or risk orientation alters choice in 16p11.2 hemideletion mice. We found that each task induced differences between 16p11.2 and wildtype males, but not females, but the differences depended on whether the task engaged the cost of delay or uncertainty. In delay discounting,16p11.2 hemideletion males had a stronger, less flexible preference for the large reward at long delays, and this effect was reduced as wildtype males adjusted their preference to match that of the hemideletion males. In contrast in probability discounting, 16p11.2 hemideletion males initially had a similar preference for seeking improbable large rewards as did wildtype males, but over time began to prefer certainty to a greater extent than did wildtype males. We have previously seen that male 16p11.2 hemideletion mice commit fewer nonreinforced responses than male wildtype mice in an operant setting, which we replicate here in delay discounting, while the introduction of risky rewards eliminates genotype differences in nonreinforced responses. Overall these data suggest that 16p11.2 hemideletion in males leads to differential processing of costs of delay versus

inconsistency, with greater aversion to uncertainty than delays, and greater behavioral control by cues that consistently predict an outcome.

3.2 Results

3.2.1 Delay Discounting

Male 16p11.2 hemideletion mice resist devaluation effects of large delay.

16p11.2 hemideletion mice in the past have been demonstrated to show hyperactivity and a delayed rate of reward learning (Angelakos et al., 2017; Grissom et al., 2018), both of which can correlate with a difference in reward valuation. 16p11.2 hemideletion mice have also been shown to stick to a choice rule once formed (Yang et al., 2015). In order to examine if this difference in learning extends to decision-making, we had 16p11.2 hemideletion mice experience both forms in order to identify possible vulnerabilities to delayed rewards. We present data from delay discounting progression to identify possible acquisition and mastery effects.

Across all of delay discounting, delay significantly reduced large reward preference (Figure 3.1A & 3.1B: main effect of Delay, $p<0.001$). Transitions from Worsening to Improving resulted in anchoring effect after initial training (Figure 3.1A & 3.1B (left): main effect of Schedule, $F(1, 36) = 14.894$, $p<0.001$) and remained after additional training (Figure 3.1A & 3.1B (right): main effect of Schedule, $F(1, 36) = 11.773$, $p=0.002$). Anchoring effects were prevalent at specific delays (Figure 2A & 2B: Schedule x Delay interaction, $p<0.001$).

Within the first round of Worsening and Improving delay discounting, 16p11.2 hemideletion mice showed evidence of resisting devaluation effects caused by delays to reward (Figure 3.1A & 3.1B (left): Genotype x Delay interaction, $F(4, 288) = 3.207$, $p=0.0134$). Upon comparing genotype differences in discounting within sexes, we found male 16p11.2 hemideletion mice repeatedly chose large delayed rewards more

often than male wildtype mice when the delay was largest (Figure 3.1A (left): male 20s delay, p=0.0721 NS; male 28s delay, p=0.0237). Although evidence suggests there might be sex differences in delay discounting (Figure 3.1A & 3.1B (left): Sex x Delay x Schedule interaction, $F(4, 288) = 2.589$, p=0.0370), post-hoc comparisons do not support within schedule differences.

In order to determine the strength of this male-specific difference in delay discounting, we compared preference rates across multiple days. Male 16p11.2 hemideletion mice chose the large reward more often than male wildtype mice specifically early into training, but adjusted their discounting to become as sensitive to delays as wildtype mice (Figure 3.1C: 28s delay: day 2, $t(13.604) = -2.576$, p=0.022; day 3, $t(15.585) = -2.565$, p=0.021; day 4, $t(15.197) = -2.898$, p=0.011). These data indicate male-specific differences in discounting are minimized as mice gain more experience on the task. This is in stark contrast to female mice who consistently chose the large reward at a similar opportunity cost ratio (Figure 3.1C: $p>0.05$). In order to determine if the difference in discounting at the 28s delay within the first round of Worsening discounting was due to a change in the male wildtypes or male 16p11.2 hemideletion mice, we compared their training history and grouped them into a Minimal phase (days 2-4) and an Extended phase (days > 4). We discovered that it was male 16p11.2 hemideletion mice that significantly shifted their preference over the course of the first round of Worsening delay discounting (Figure 3.1D: Genotype x Training interaction, $F(1, 18) = 5.027$, p=0.0378). Post-hoc comparisons confirm this was only significant within male 16p11.2 hemideletion mice (Figure 3.1D: p=0.0215)

Male 16p11.2 hemideletion mice commit fewer trial initiation errors, indicative of action inhibition through a nonreinforcement period.

Male 16p11.2 hemideletion mice have previously been shown to use different response strategies to attain rewards in that they tend to make significantly less

nonreinforced responses in a simple continuous reinforcement task (Grissom et al., 2018). As such, we measured those nonreinforced touches during different epochs of the task in order to determine if nonreinforced touches were differently influenced between genotypes as a function of task progression. We present nonreinforced responses during the choice period throughout the delay discounting session as a measure of initiation perseveration.

Consistent with previous research, 16p11.2 hemideletion mice made significantly fewer nonreinforced nosepokes in the center (initiation) hole during the choice period for both the first half of discounting (Figure 3.2 (left half of males and females): main effect of Genotype, $F(1, 32) = 5.711$, $p=0.0227$) and second half of discounting (Figure 3.2 (right half of males and females): main effect of Genotype, $F(1, 34.742) = 8.010$, $p=0.0077$). Nonreinfored nosepokes were initially increased on the Improving Schedule compared to the Worsening Schedule (Figure 3.2 (left half of males and females): main effect of Schedule, $F(1, 33) = 4.263$, $p=0.0469$) but that difference diminished with additional experience. Post-hoc tests revealed this difference was mainly due to more nonreinforced touches made by female 16p11.2 hemideletion mice on the Improving Schedule (Figure 3.2 (left half of females), $p=0.0178$). Male 16p11.2 hemideletion mice made less nonreinforced nosepokes for both the first round of discounting (Figure 3.2 (left half of males): Worsening, $p=0.0514$ NS; Improving, $p=0.0441$) and the second round of discounting (Figure 3.2 (right half of males): Worsening, $p=0.0015$; Improving, $p=0.0219$). Females exhibited no such genotype differences ($p>0.05$).

3.2.2 Probability Discounting

Male genotype differences in risk preference arise with experience of large risk.

Reward uncertainty is another form of cost that can modulate the value of a reward. A prominent theory within the study of neurodevelopmental disorders implicates the importance of stable cues in learning environments (Sinha et al., 2014). We challenged 16p11.2 hemideletion mice with probability discounting in order to determine if they can adjust to the uncertainty of reward delivery coming off of a deterministic task like delay discounting. We again present discounting data to emphasize the progression of learning.

For all probability discounting sessions, we found an expected significant main effect of probability indicating a tendency to reduce large reward preference as risk increased (Figure 3.3A & 3.3B: main effect of Probability, $p<0.001$). Anchoring effects were clearly present within the first round of probability discounting (Figure 3.3A & 3.3B (left): main effect of Schedule, $F(1, 37) = 6.663$, $p=0.0139$), but the influence of anchoring diminished over additional training (Figure 3.3A & 3.3B (right): $p>0.05$). However, for both rounds there was clear evidence of probability specific differences in discounting between schedules (Figure 3.3A & 3.3B: Schedule x Probability interaction, $p<0.001$). Male mice generally endured risk more often for large rewards than female mice for the first round of probability discounting (Figure 3.3A & 3.3B (left): main effect of Sex, $F(1, 37) = 4.571$, $p=0.0392$), a trend which was present but not significant with additional training (Figure 3.3A & 3.3B (right): main effect of Sex, $F(1, 37) = 3.942$, $p=0.0546$). These increased male risky decisions were prominent at multiple probabilities (Figure 3.3A & 3.3B: Sex x Probability interaction, $p<0.001$).

Male mice displayed differences in risk processing in the second half of probability discounting. Male 16p11.2 hemideletion mice were more risk averse when uncertainty of reward delivery was high on the Worsening Schedule (Figure 3.3A (right): male 12.5% Probability, $p=0.0369$) and the Improving Schedule (Figure 3.3A (right): male 25% Probability, $p=0.0274$).

We compared discounting preferences across days in order to assess the consistency of these effects. Differences do arise early and late into training of probability discounting when risk is greatest (Figure 3.3C: 12.5%: day 6, $t(8.799) = 2.649$, $p=0.027$; day 25, $t(12.359) = 2.657$, $p=0.020$). Interestingly, differences were more apparent when the expected value of the large risky option was in conflict with the value of the small option (Figure 3.3C: 25%: day 8, $t(8.131) = 2.407$, $p=0.042$; day 10, $t(8.814) = 2.524$, $p=0.033$; day 25, $t(9.843) = 5.435$, $p<0.001$; day 28, $t(10.751) = 2.649$, $p=0.023$).

Uncertainty of reward delivery erases male genotype differences in nonreinforced responding.

Delay discounting adds a delay to reinforcement but is ultimately still a deterministic task. Animals tend to use different strategies when delivery of reinforcement is uncertain. Therefore we measured nonreinforced nosepokes through all epochs of probability discounting to determine if the previously measured genotype difference is still present. We present nonreinforced responses during the choice period throughout the probability discounting session as a measure of initiation perseveration and to compare to delay discounting.

Male genotype differences in nonreinforced nospokes committed in the center hole during the choice period were not significant, but were still present early into probability discounting (Figure 3.4 (left half of males): Worsening, $p=0.0997$ NS; Improving, $p=0.0818$ NS) and became less apparent as mice gained additional experience. While females were still similar in their pattern of nonreinforced nosepokes, female wildtypes adjusted to the uncertainty of the task by the end of training by reducing their nonreinforced nospokes (Figure 3.4 (right half of females): $p=0.0284$). Comparing these findings to the delay discounting nonreinforced

nosepokes suggests task uncertainty disrupts or influences male 16p11.2 hemideletion specific action patterns.

3.3 Discussion

At the beginning of our study, we set out to challenge decision-making in 16p11.2 hemideletion mice in order to determine the impact of a copy number variant related to neurodevelopmental disorders. We chose delay and probability discounting tasks because they assess different aspects of reward-guided decision-making, of which is altered in neurodevelopmental disorders (Clements et al., 2022; Damiano et al., 2012; Mosner et al., 2017; Nissan et al., 2023). Using our battery of decision-making tasks, we provide evidence that 16p11.2 hemideletion impacts choices in two types of discounting tasks. Previous research found male 16p11.2 hemideletion mice are slower to form action-outcome relationships, resulting in delayed instrumental learning in a simple reinforcement task (Grissom et al., 2018). Our results corroborate those findings and expand upon them in a choice paradigm where task demands progressively shift within a session. We found that male 16p11.2 hemideletion mice exhibited resistance to temporal costs and delayed sensitivity to probabilistic costs compared to male wildtype mice. Male genotype differences in response to temporal delays occur initially with minimal experience, while the effects of reward uncertainty emerge with extensive experience. Patterns of nonreinforced responses during the choice period shift between tasks where male 16p11.2 hemideletion mice noticeably commit less trial initiation errors. Our data suggests male 16p11.2 hemideletion mice engage with their environment in different ways than do male wildtype mice and display temporarily enduring patterns of choice.

Our results are in agreement with previous findings on 16p11.2 hemideletion mice emphasizing slow learning and response inflexibility. Previous studies report delayed

operant learning of which we find a parallel to here with male 16p11.2 hemideletion mice taking longer than male wildtype mice to shift their choice in response to large delays (Grissom et al., 2018). We found evidence of initially rigid choice similar to previous studies emphasizing inflexibility and perseverative choice in 16p11.2 hemideletion mice (Yang et al., 2015). These papers suggest a general challenge with adapting to reward-guided tasks, but do not elucidate why. Slower or different adaptation to costs is a phenotype observed in these mice and in neurodevelopmental disorders generally (Mussey et al., 2015; Yechiam et al., 2010; Zeif et al., 2023). Some research suggests increased sensitivity to losses can lead to a difference in choice in probabilistic tasks (Gosling & Moutier, 2018). Uncertainty avoidance can form when predictions are violated, resulting in what appears to be increased sampling or a tendency to shift towards certainty (Bervoets et al., 2021; Sinha et al., 2014; Zeif et al., 2023). Similarly, when reward certainty decreases in probability discounting, male 16p11.2 hemideletion mice significantly shift their choices toward the small certain choice. These results suggest exploratory behavior and decisions in 16p11.2 hemideletion mice could be modulated by uncertainty or risk of reward, which might also make it difficult to acquire new responses. Taken together, our results suggest male 16p11.2 hemideletion mice assess reward value in a manner that is different from wildtype mice and that is experience and uncertainty dependent. Future research could examine decision-making in 16p11.2 hemideletion mice in a dynamic setting such as in bandit tasks (Chen et al., 2020, 2021). This will enable researchers to determine the stability of 16p11.2 hemideletion-induced strategies and periods of possible perseveration as reward contingencies shift.

Next, we decided to measure nonreinforced touches throughout discounting in order to examine if exploratory behavior or approaches to learning about choice contingencies differed between 16p11.2 hemideletion and wildtype mice. One feature

of operant behavior in 16p11.2 hemideletion that we observed here is a differential rate of nonreinforced responding, where hemideletion males make fewer nonreinforced responses than do males. Grissom et al. (2018) previously saw a similar pattern in a different task (5-choice serial reaction time task) and different operant chamber format (9-hole nosepokes versus touchscreen). This suggests that one source of male-biased impacts of 16p11.2 hemideletion is a reduction of unnecessary actions that wildtype males exhibit. Two possible explanations for nonreinforced responding in general that may differ across genotypes are differences in motivation/hyperactivity, or differences in attentional enhancement. To address the first hypothesis, anticipatory responses leading up to choice can be thought to reflect a type of general "impulsivity" (Dalley et al., 2011; Hogarth et al., 2012), while reduced motivation and vigor of actions are elicited when rewards are less frequent or more delayed (Ko & Wanat, 2016; Mohebi et al., 2019; Nicola, 2010). Through this lens, male 16p11.2 hemideletion mice may exhibit less effort towards unnecessary responses when rewards are temporally distal (in delay discounting) compared to when they are more immediate and uncertain (in probability discounting). Prior work has shown reduced responding in 16p11.2 hemideletion male mice in a progressive ratio task, consistent with this hypothesis (Grissom et al., 2018). The second hypothesis is that 16p11.2 hemideletion males may have greater attentional control and/or ability to inhibit prepotent responses. Openshaw et al. (2023) recently used a continuous performance task and found that male 16p11.2 hemideletion mice performed with greater accuracy as measured by hit rate (i.e. responding during a correct stimulus) compared to male wildtype mice. Because the nonreinforced responding we measured occurred after the center hole was no longer illuminated, it may be that wildtype males have their responding under less stringent control of the cue than do 16p11.2 hemideletion males. One strategy for rational agents in discounting tasks is to decrease attention during nonrewarding periods and increase attention when the reinforcer becomes available again (Mikhael et al., 2021). For male

wildtypes, extra responses may be one way in which to combat the ambiguity of the ITI period. Male 16p11.2 hemideletion mice rely more on the presence of response cues that signal for trial phases. However, when the reliability of reinforcement decreases such as in probability discounting, male 16p11.2 hemideletion mice start to adopt the same behavioral pattern as male wildtype mice. Future task designs may wish to explicitly test whether changes in motivation or changes in attentional or cue-regulated control form the greater contribution of behavioral differences in 16p11.2 hemideletion males.

Our work supports growing evidence that 16p11.2 hemideletion impacts males more than females especially in reward-related scenarios. Males show genotype mediated differences in discounting while females are generally unaffected in their choice behavior across both tasks similar to other types of reinforcement learning settings (Grissom et al., 2018). Contrary to findings of increased hyperactivity in 16p11.2 hemideletion mice, we found evidence of increased behavioral control (Angelakos et al., 2017). However as the authors point out themselves, 16p11.2 hemideletion-induced hyperactivity is context dependent. General locomotion in small operant chambers tend to be similar regardless of genotype (Grissom et al., 2018). Instead, 16p11.2 hemideletion mice committed fewer nonreinforced nosepokes similar to previous findings (Grissom et al., 2018). One group found evidence of anxiety-like behavioral and neurological responses after a fear-inducing event in female 16p11.2 hemideletion mice (Giovanniello et al., 2021). Probability discounting is one task in which risk attitudes affect discounting rates (Shead & Hodgins, 2009), so we expected female genotype differences to potentially be exposed there. However, contrary to those expectations female mice discounted at similar rates indicating the need for more overt punishments for exposing possible genotype differences. These results highlight how this copy number variant is affected by background, genotype and sex (Grissom et

al., 2018; Horev et al., 2011; Portmann et al., 2014). Given that 16p11.2 hemideletion in humans has different impacts based on other genotypes and across genders (Chawner et al., 2019; Duyzend & Eichler, 2015; Hanson et al., 2015), this highlights the importance of understanding cognitive and neural impact of neuropsychiatric-disorder linked copy number variants as a function of individual differences.

3.4 Methods

3.4.1 Subjects

Male 16p11.2 hemideletion mice (stock #013128) and female wildtype mice (stock #101043) were obtained from Jackson Laboratories and bred in order to generate mice of both genotypes for the experiments. Mice were housed in groups of 2-5 of mixed genotypes. Mice began experiments at approximately >120 days of age. Mice had free access to water and were food restricted with their home chow at 85–90% of their baseline weight. Mice were pre-exposed to the operant reinforcer, vanilla flavored Ensure, in their home cage for one day prior to training. Ensure was freely available to be licked from a bottle. Each group of mice were verified to have consumed a full bottle (148 ml) of Ensure. Behavioral testing took place Monday to Friday, and on Fridays, mice had free access to home chow. Animals were housed on a reverse light-dark cycle (9 am-11 pm) and were tested during the dark period. Animals were cared for in accordance with National Institute of Health guidelines and were approved by the University of Minnesota Institutional Animal Care and Use Committee.

3.4.2 Apparatus

We used 16 identical triangular touchscreen chambers, 8 per sex separated in different rooms for training and testing (Lafayette Instrument Co., Lafayette, IN). The touchscreen apparatus was located at the front of the chamber while liquid rewards

were delivered at the back of the chamber. Schedules were administered and interactions with the screen were recorded via ABET-II software. All data were exported using ABET-II software. Touchscreens were limited by masks with five evenly spaced square holes. Ensure was diluted to 50% and delivered via peristaltic pump.

3.4.3 Touchscreen Training

Training was similar to our previous experiment (Rojas et al., 2022). Mice began with magazine training for 3 days and moved on to training schedules. For all training schedules, mice were only able to advance in trials if they performed the required action(s) designated by illumination by a 5-choice mask and collected the reward in the magazine at the back of the chamber. The touchscreen illuminated in the holes closest to the center where mice were required to nosepoke in order to advance the trial. Any responses on the outer two holes always resulted in no reinforcement at all stages. Reward delivery was accompanied by a light cue and the sound of the pump.

Center Only Fixed-Ratio 1. Mice experienced FR1 training for 40 days until a majority of them readily acquired a basic understanding to nosepoke for 7 µl of Ensure. Responses to the illuminated center hole resulted in reward delivery, all other responses were recorded but did not lead to reward delivery. Sessions ended after 30 minutes had elapsed.

Progressive Ratio. Mice intermittently experienced two sessions of progressive ratio. Mice were still required to selectively respond in the center hole for reinforcement. The first session occurred after 14 days of FR1 training. The second session occurred 37 days after FR1 training. Sessions advanced in an arithmetic sequence (1, 2, 4, 7, 11, 16, 21, …n). Each ratio had to be completed for three trials before advancing to the next in the sequence. Sessions terminated after animals failed

to nosepoke for 5 minutes or after 60 minutes had elapsed. The last ratio completed became the animal's breakpoint.

Chaining Center (Reinforced) to Left and Right. Animals began training to chain their response in order to build up to choices in discounting tasks. Sessions began with an illuminated center hole. Nosepoking the center hole resulted in 7 μl of Ensure, which upon collection illuminated the holes immediately to the left and right half of the center hole. Mice were then required to respond on either side for the same amount of Ensure on each side (28 μl). Training lasted for 24 days and sessions were 30 minutes in length.

Chaining Center (Unreinforced) to Left and Right. The next phase aimed to teach mice to respond in the center purely as an initiation response, resulting in no Ensure. The mice chose between left and right options which were still reinforced at the same volume. Mice experienced this schedule for 40 days and sessions were 40 minutes in length.

Chaining Center to Small and Large Magnitude Reward Options with ITI. The last stage of training before discounting was conducted before delay discounting in order to teach mice one option immediately left or right to the center hole resulted in a greater amount of Ensure (20 μl) than the other (5 μl). Additionally, the now large option resulted in a longer intertrial interval (ITI) of 30s. However, animals could choose no more than 2 of the same option in a row before they were forced to sample the other option. Mice learned these new concepts for 20 days and sessions ended after 60 minutes had elapsed or after mice cleared 60 trials.

3.4.4 Delay Discounting Phase

MIce underwent sequences of Worsening and Improving schedules as previously described (Rojas et al., 2022). One deviation from the previous study is that mice did not receive forced-choice trials in order to prioritize completion of delay blocks. Mice learned the Worsening schedule initially in which sessions began with a 0s delay for delivery Ensure after the large option, which temporally increased in cost as trials progressed (Figure 3.0). Sessions consisted of 40 free-choice trials divided into 5 blocks per delay (0, 4, 12, 20, and 28s delays). There was no delay to reinforcement for the small option, but mice did need to wait additional time to initiate the next trial (matched to the length of the delay trial). Mice proceeded to learn the Improving schedule in which sessions began with a 28s delay cost for Ensure after the large option, which temporally decreased in cost as trials progressed (28, 20, 12, 4, 0s delays). Small reward choices were again balanced to large option trial length by increasing the ITI length. Mice underwent 13 days of the Worsening schedule, 24 days of the Improving schedule, 10 days of a return to the Worsening schedule, and 10 days of a return to the Improving schedule.

3.4.5 Probability Discounting Phase

Following delay discounting, mice then were exposed to the probability discounting task. Probability discounting training started with the Worsening schedule where 100% of trials resulted in delivery of the large reward, but uncertainty increased as trials continued (Figure 3.0). Sessions initially consisted of 40 trials for the first round of Worsening and Improving schedules, but switched to 80 trials in order to promote increased discounting of the reward. Sessions were divided into five probability blocks (100, 75, 50, 25, and 12.5% probability of large reward). Mice went on to learn the Improving schedule where odds of winning a large reward began at 12.5%, but probabilistically increased as trials progressed (12.5, 25, 50, 75, 100% probability of large reward). Mice underwent 10 days of the Worsening schedule, 10

days of the Improving schedule, 8 days of an initial return to the Worsening schedule, 15 days of an initial return to the Improving schedule, 10 days of a secondary return to the Worsening schedule, and 10 days of a secondary return to the Improving schedule. Data from the first round of probability discounting are excluded because of the decreased trial counts compared to the other rounds (40 vs 80) and because animals failed to show significant devaluation effects with increased risk.

3.4.6 Statistical Analysis

We constructed linear mixed models in order to analyze choice data and compare nonreinforced nosepokes. For analysis of large reward preference in delay discounting, we included fixed factors of Sex, Genotype, Schedule, Delay and their interactions (Sex x Genotype, Sex x Schedule, Sex x Delay, Genotype x Schedule, Genotype x Delay, Delay x Schedule, Sex x Genotype x Delay, Sex x Genotype x Schedule, Sex x Delay x Schedule, Genotype x Delay x Schedule, Sex x Genotype x Delay x Schedule). Random factors included a random slope of Schedule and random intercept of Subject. For analysis of large reward preference in probability discounting, we included fixed factors of Sex, Genotype, Schedule, Probability and their interactions (Sex x Genotype, Sex x Schedule, Sex x Probability, Genotype x Schedule, Genotype x Probability, Delay x Schedule, Sex x Genotype x Probability, Sex x Genotype x Schedule, Sex x Probability x Schedule, Genotype x Probability x Schedule, Sex x Genotype x Probability x Schedule). Random factors included a random slope of Schedule and random intercept of Subject. Comparisons of delay discounting and probability discounting preference within days were assessed using Welch's t-tests (uncorrected for multiple comparisons). If mice were unable to reach at least 60% completion of trials on average across all days, those mice were excluded from analysis. Additional analysis of Worsening delay discounting within 28s delay of male mice was conducted with a linear mixed model consisting of main effects of Genotype

and Training, and a Genotype x Training interaction term. There was also a random factor of Subject. Final group sizes for delay discounting analysis were as follows: 9 male wildtype mice, 9 male hemideletion mice, 10 female wildtype mice and 8 female hemideletion mice. Final group sizes for probability discounting analysis were as follows: 9 male wildtype mice, 9 male hemideletion mice, 10 female wildtype mice and 9 female hemideletion mice.

Analysis of nonreinforced touches in delay discounting and probability discounting included fixed factors of Sex, Genotype, Schedule and their interactions (Sex x Genotype, Sex x Schedule, Genotype x Schedule, Sex x Genotype x Schedule). There was a random intercept for Subject. Nonreinforced touches two standard deviations away from the mean were excluded from the analysis.

The R package *lmerTest* was used to fit linear mixed models. The Satterthwaite method was used to estimate degrees of freedom for omnibus F tests (main effects and interactions). Post-hoc comparisons were made using the *emmeans* R package with Tukey adjustment and Satterthwaite method for estimation of degrees of freedom. All graphs were produced using the *ggplot2* R package.

3.5 Figures

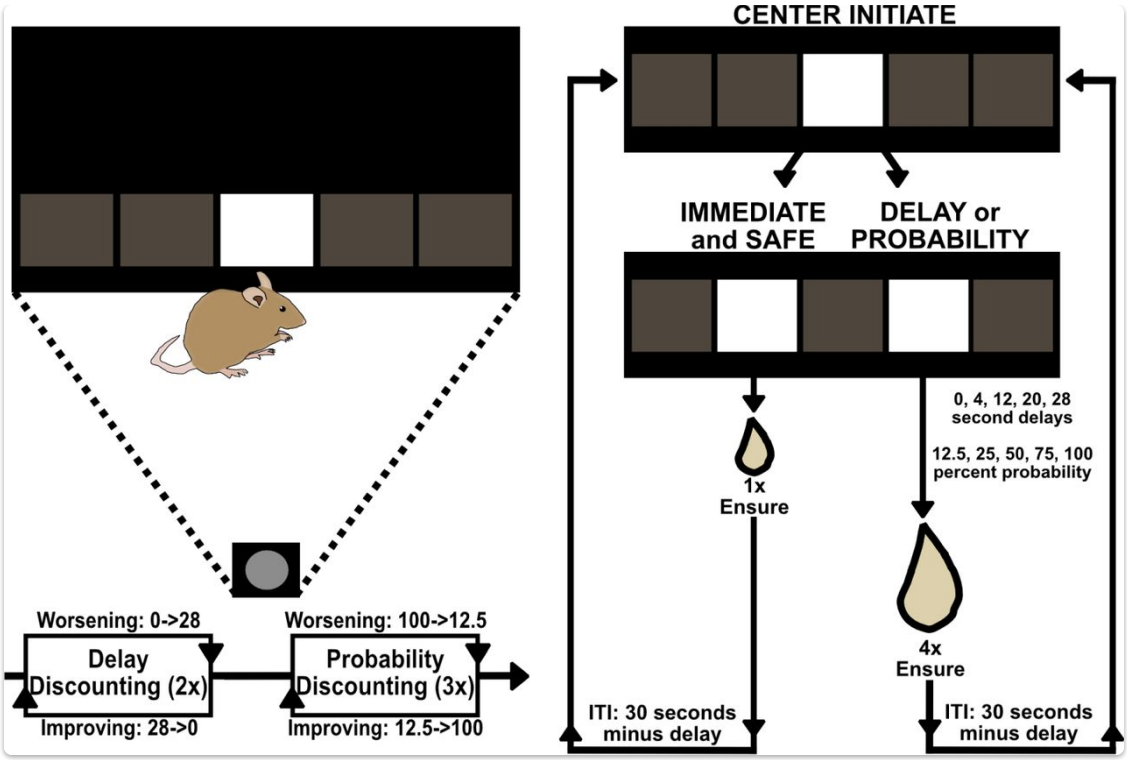

Figure 3.0 Sequencing of delay discounting and probability discounting.

(Left-top) Mice were required to respond in illuminated areas of the touchscreen, limited to the 5 hole mask. The front of the chamber contained the touchscreen, the back of the chamber contained the magazine where Ensure was delivered. (Left-bottom) Mice experienced the Worsening schedule of delay discounting for 13 days before experiencing the Improving schedule for 24 days. They then repeated each schedule one more time before moving to probability discounting. Mice experienced the Worsening schedule for 10 days before experiencing the Improving schedule for 10 days. Mice then repeated each schedule two times before completing discounting. (Right) Sessions of discounting began with a center initiation nosepoke on a 5 hole mask. Mice were then presented with a small choice in which Ensure delivery was always safe and immediate (5 µl) and a large choice in which Ensure was 4x the magnitude (20 µl) but delivered in a delayed or probabilistic manner. ITIs were

matched for trial length in delay and probability discounting. In delay discounting, ITIs were 30s minus the delay. In probability discounting, ITIs were 3s.

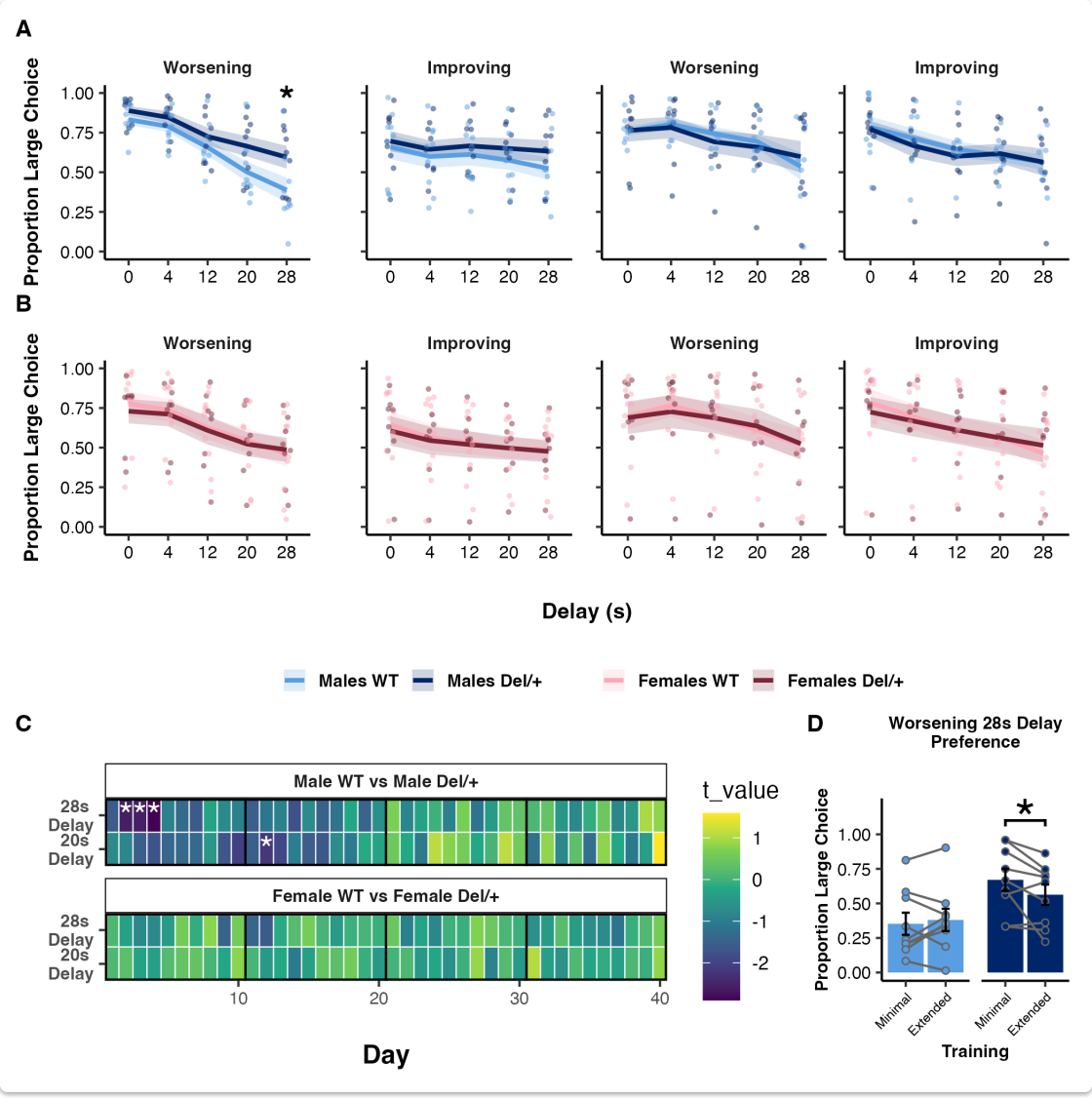

Figure 3.1 Male 16p11.2 hemideletion mice learned to shift reward preference with large delays.

Mice transitioned from Worsening schedules (0s -> 28s delay) to Improving schedules (28s -> 0s delay) in order to assess schedule specific learning. (A) Male 16p11.2 hemideletion mice were relatively insensitive to longer delays compared to male wildtype mice (28s delay: p=0.0246). However, as male 16p11.2 hemideletion

mice gained more experience with the task, they no longer exhibited delay specific differences compared to male wildtype mice. (B) Female mice exhibited similar discounting rates throughout all of delay discounting. (C) Discounting data is presented throughout days where the black vertical line indicates a change in schedule (Worsening to Improving, Improving to Worsening). When comparing learning history of those delays throughout training, it was evident that male 16p11.2 hemideletion mice preferred the large delayed option compared to male wildtype mice early but not late into training. In agreement with the overall trend with extended training, male genotype differences disappeared with additional training. D) Training data was split into a minimal timeframe (days 2-4) and an extended time frame (days > 4) to determine if wildtype mice became more like 16p11.2 hemideletion mice or vice versa within the first round of Worsening delay discounting. Our data confirms male 16p11.2 hemideletion mice but not wildtypes significantly alter their preference point ($p=0.0215$). Shaded areas in A and B and error bars represent standard error from the mean. Asterisks represent statistical significance of $p<0.05$.

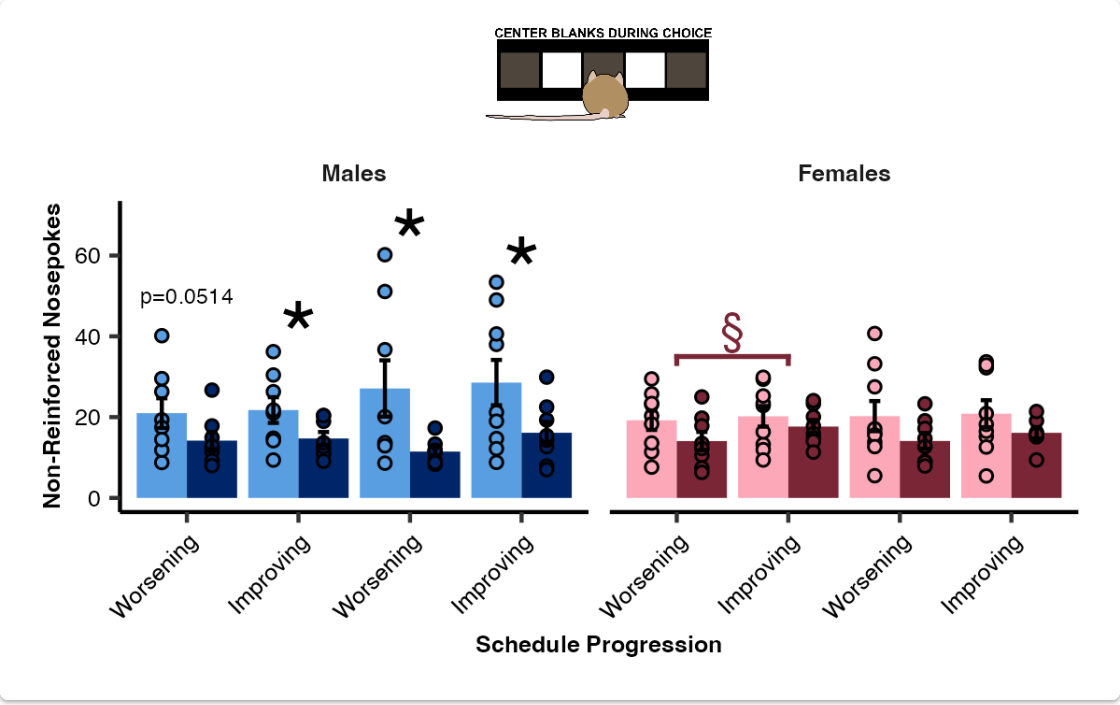

Figure 3.2 Male 16p11.2 hemideletion mice withheld center nosepokes during the choice epoch of delay discounting.

For all of delay discounting, male 16p11.2 hemideletion mice make less initiation errors than male wildtype mice. Female 16p11.2 hemideletion mice make more nonreinforced nosepokes when the schedule starts with a large delay and progressively becomes shorter (Improving). Error bars represent standard error from the mean. An asterisk represents a genotype effect of $p<0.05$, a double s represents a schedule effect within a genotype (indicated by color) of $p<0.05$.

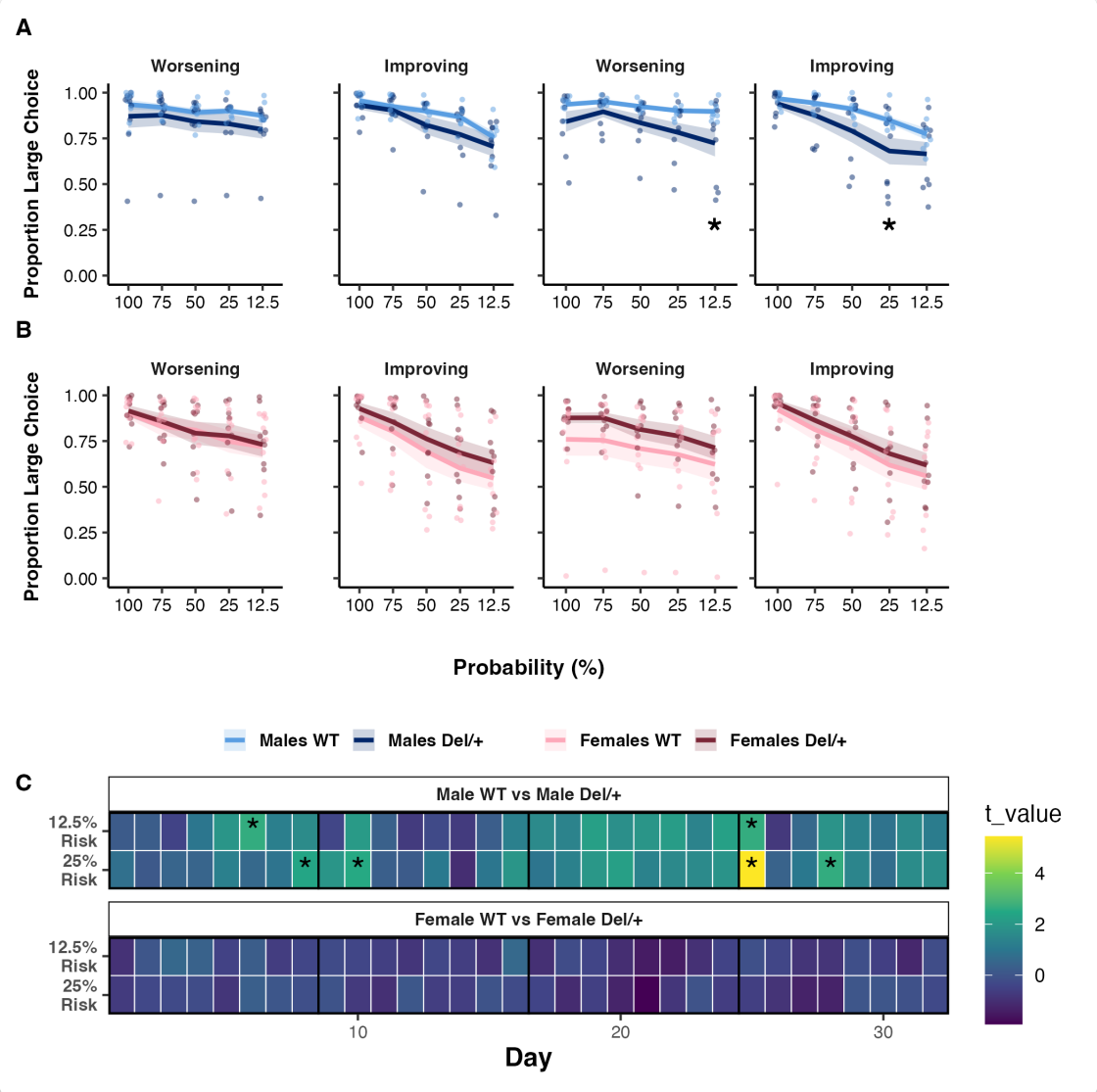

Figure 3.3 Male 16p11.2 hemideletion mice became slightly more risk avoidant with probability discounting training.

Mice transitioned from Worsening schedules (100% -> 12.5% probability) to Improving schedules (12.5% probability -> 100% probability) in order to assess schedule specific learning. (A) Mice show greater cost endurance for large uncertain rewards compared to delayed large delayed rewards. 16p11.2 hemideletion and wildtype mice discount similarly initially on both Worsening and Improving schedules. (B) However, male 16p11.2 hemideletion mice become steeper than male wildtype mice as risk increases. Specifically, male 16p11.2 hemideletion mice prefer smaller rewards more often than male wildtype mice when risk of reward delivery is highest (12.5%: $p=0.0354$). (C) Discounting data is presented throughout days where the black vertical line indicates a change in schedule (Worsening to Improving, Improving to Worsening). Training history analysis does hint at some transient increased risk sensitivity that becomes more apparent with additional training. Shaded areas in A and B and error bars represent standard error from the mean. Asterisks represent statistical significance of $p<0.05$.

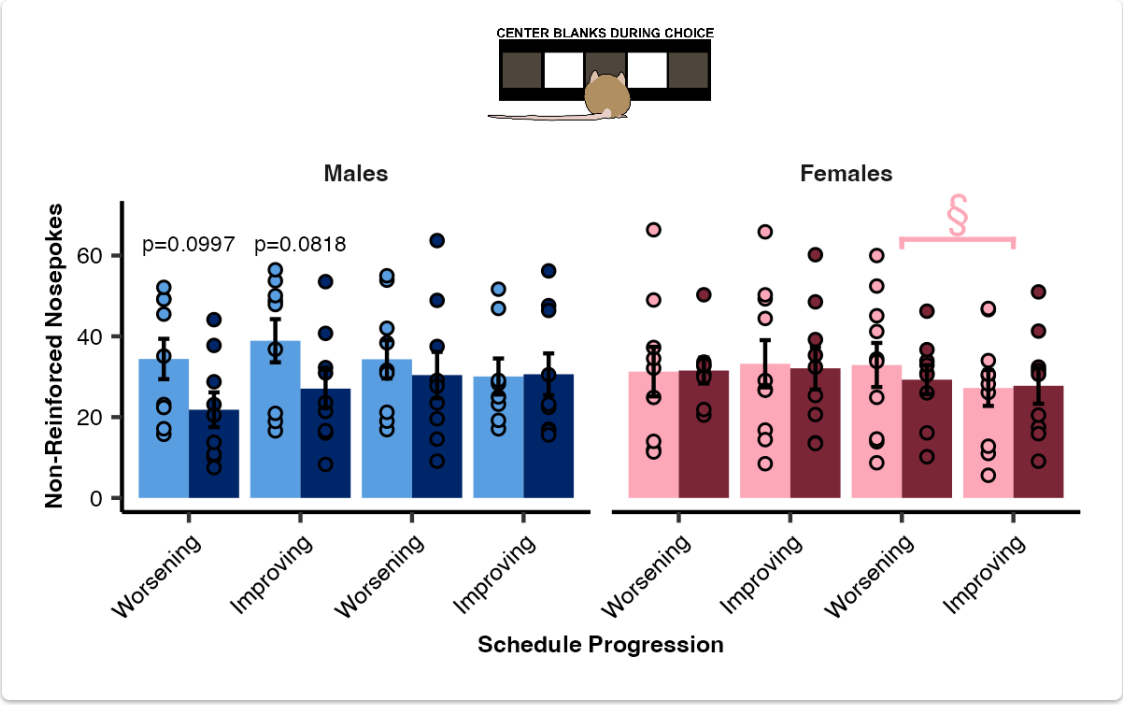

Figure 3.4 Male 16p11.2 hemideletion initiation mistakes increased upon training on probability discounting.

Upon switching from delay discounting to probability discounting, male 16p11.2 hemideletion mice committed just as many initiation errors as male wildtype mice. Female wildtype mice adapt to the uncertainty of the task at the end of probability discounting by reducing the amount of initiation errors they commit. Error bars represent standard error from the mean. Double s represents a schedule effect within a genotype (indicated by color) of $p<0.05$.

4. Discussion

The current set of studies sought to: 1) develop discounting tasks suitable for mice and provide analytical methods to study sources of discounting, and 2) explore the effects of 16p11.2 hemideletion on the acquisition and mastery of a mouse version of delay and probability discounting tasks. We designed delay and probability discounting tasks in order to determine how mice learn to shift their choices in the face of different costs, and to determine if 16p11.2 hemideletion mice are more likely to endure one type of cost over another. 16p11.2 hemideletion has relevance to behaviors observed in neurodevelopmental disorders, such as causing delayed language development in 71% of individuals and significant presence of autism-related behaviors in one clinical population (Hanson et al., 2015). Study 1 established that mice are able to learn to discount rewards, mice adapt to changes in schedule orientation, and mice are able to discern between delayed and probabilistic costs. We applied discounting models that estimated the value of the large reward through a discounting parameter, and improved on that analysis adding an inverse temperature parameter in order to observe how choice history influences those value estimations. Implementing this model elucidated the role of decision noise in discounting. Study 2 built upon the foundation of the previous experiment by using that same discounting task battery and challenging 16p11.2 hemideletion mice with delay and probability. Through this training regime, we discovered male 16p11.2 hemideletion mice engaged with the structure of the tasks and adapted their choices to reward cost types differently than male wildtype mice. This experiment built upon previous findings that male 16p11.2 hemideletion mice exhibit instrumental learning delays and struggle to switch from an established correct choice when learning reversals (Grissom et al., 2018; Yang et al., 2015). Similar to those studies, male 16p11.2 hemideletion mice acquired a greater choice preference for

the large reward than wildtype mice, and temporarily refused to adjust their preference over delay discounting. Probability discounting revealed male 16p11.2 hemideletion mice choice patterns and task structure attention is disrupted by probabilistic uncertainty. Our studies suggest mice make separate value assessments for delay and probability and show that 16p11.2 hemideletion mice deviate from those assessments based on experience and uncertainty. I will now explain how these two experiments provide important implications for how discounting behavior forms and the additional epochs in discounting tasks that should be probed for differences in learning.

4.1 Summary of Study 1 (Modeling Discounting Choices in Mice)

In order to probe decision-making in mice, I began by producing "Worsening" and "Improving" variants of delay and probability discounting. Mice were presented with the choice of either a small reward with no consequences or a large reward with programmed delays (0s, 4s, 12s, 20s, or 28s delay to reward) or probabilistic risk (100%, 75%, 50%, 25%, or 12.5% odds of reward). Mice needed to learn periods in which the large reward had no cost, and to shift their preference once the cost was too hefty. Schedules either began with no cost to the large reward that progressively got worse within a session (i.e., "Worsening") or started with the largest cost to the large reward that progressively improved within a session (i.e., "Improving"). Mice learned to do this through multiple repetitions in order to solidify choice preferences.

This study had the following goals: 1) prove mice can successfully discount rewards according to the cost of the reward, 2) test a novel computational approach to discounting behavior, and 3) track fluctuations in discounting behavior and computational parameters throughout iterations of the tasks. Mouse versions of discounting tasks are valuable because discounting takes a similar form between humans and rodents (Vanderveldt et al., 2016) and discounting tasks enables

researchers to study different forms of decision-making as a consequence of mice modeling genes related to neurological disorders (Navabpour et al., 2020; Silverman et al., 2022; Tan & Zoghbi, 2019). By studying the development of different types of value-based decisions, we can build a profile of "trait-like" or stable longitudinal decision-making and make comparisons across cost types (Alabi et al., 2019).

We gained clear evidence of the first goal — mice were able to show within-session sensitivity to delay and probability. Our long training history revealed mice initially demonstrated significant "anchoring" of choice due to schedule switches (i.e., previous initial delay or probability experience affected current choices), but learned to adapt to schedule changes with extended training. Additionally, mice showed changes in choice response times according to cost, indicating choices became harder as costs increased and further supporting that mice were actively incorporating costs into their value assessment (Busemeyer et al., 2019). Our preliminary analysis of choices through win-stay and lose-shift comparisons revealed mice adapted their strategies according to schedule switches, prompting us to further analyze their behavior through computational models.

Such tasks are amenable to different computational approaches, such as discounting and reinforcement learning models. One analysis we had in mind to tackle goals two and three was to incorporate a noise parameter into the standard discounting models. Reinforcement learning often uses two parameter models to capture learning rate and decision deviations (Nussenbaum & Hartley, 2019). Inverse temperature captures deviations in choices, which is important to assess how animals learn about choices in different environments (Chen et al., 2021). Standard models already have a value parameter; thus we determined an inverse temperature would tell us more about how animals arrive at that value parameter. Incorporating an inverse temperature parameter (Ebitz et al., 2019) into a standard value-based model (Green & Myerson,

2004) revealed mice systematically build a value assessment based on their exploration of cost-reward tradeoffs. A significant insight revealed, no matter how unstable the discount parameter was, we were able to track changes in preference over iterations of the tasks based on the inverse temperature parameter. One interesting observation from delay discounting was that mice defaulted to a greater large choice strategy but became noisier and shifted when schedules initially changed from Worsening to Improving. Decision noise allows for opportunities for learning by encouraging exploration, which was probably needed in order to adapt to schedule changes (Ebitz et al., 2019). These results provided critical evidence for the importance of tracking the development of discounting and a computational approach to explaining shifts in discounting.

4.2 Summary of Study 2 (Adaptations to Costs in 16p11.2 Hemideletion Mice)

The second experiment utilized the now validated mouse discounting tasks to test the development of discounting in a mouse model of 16p11.2 hemideletion. We used the same Worsening and Improving schedules in order to determine if genotype differences were present on certain variants of the tasks. In addition, we examined other epochs of the tasks (e.g., the intertrial interval period) besides the choice period in order to ascertain if 16p11.2 hemideletion induces changes in task engagement and environmental exploration.

Study 2 had the following goals: 1) determine if sex-biased 16p11.2 hemideletion reinforcement effects extend to decision-making paradigm (Grissom et al., 2018), 2) apply our computational factor in order to determine gross differences in choice variability in 16p11.2 hemideletion mice, and 3) look for periods of perseveration induced by task structures (Yang et al., 2015). 16p11.2 hemideletion mice have exhibited behavioral abnormalities in several studies pertaining to motor behavior, motivation, and learning. Recent studies consistently find increases in hyperactivity

induced by 16p11.2 hemideletion (Angelakos et al., 2017; Horev et al., 2011). 16p11.2 hemideletion male mice in operant settings have exhibited deficits in motivation as measured by progressive ratio performance and simple continuous reinforcement tasks (Grissom et al., 2018). Additionally, 16p11.2 male hemideletion mice exhibit decreased nonreinforced responses (Grissom et al., 2018). In a discounting paradigm, this could affect their willingness to obtain rewards in the face of costs, their understanding of rewarding periods (e.g., contingency learning), and their exploration of the chamber. Many discounting studies focus solely on the choice behavior, but previous studies indicate there may be more subtle but important expressions of changes in learning and motivation (Grissom et al., 2018). We began our analysis with choice behavior in order to determine if previous sex-specific changes in learning extended to a decision-making setting.

We measured choice behavior throughout experience of delay and probability discounting and confirmed that differences in discounting behavior appeared only in male mice. Male 16p11.2 hemideletion mice specifically temporarily showed a stronger preference for large rewards compared to male wildtype mice, especially at larger delays. We know male 16p11.2 hemideletion initially stuck with that choice preference and shifted after a couple of days towards the delay-free small reward, indicating male 16p11.2 hemideletion mice were slower to adjust to increasing costs than wildtype mice were. However, male 16p11.2 hemideletion mice were more receptive to probabilistic rewards, as they exhibited a decreased tendency to endure large reward risk compared to wildtype mice. Female mice did not exhibit such differences in discounting despite some studies indicating female 16p11.2 hemideletion mice are more sensitive to aversive situations, which could have extended to the type of uncertainty related to reward delivery risk (Giovanniello et al., 2021). Still, our results confirmed previous male-specific findings. Next, we decided to look at our

established computational model to determine how choice variability contributed to preference. However, we had limited success with the models for a few reasons. The models were able to capture overall decision noise, but they were unable to produce the differences we observed in preferences over blocks. One immediate problem is that the task is already limited on trials and proportionally the transient effect we observed in discounting made up less of those trials. The model is less robust with fewer trials. Additionally, the effects we observed were at specific delays which could be hidden by the overall curve produced by the models. To further add evidence, even a simple area-under-the-curve analysis to all delays was unable to capture the difference in discounting we observed. So instead, we focused on the last goal to look through different periods of the discounting tasks in order to see if we observe differences in the way 16p11.2 hemideletion mice engaged with the tasks.

The third goal was achieved and reproduced previous findings that male 16p11.2 hemideletion mice produce less nonreinforced responses than male wildtype mice. Specifically, male wildtype mice produced significantly more nonreinforced responses in the center initiation hole when choices were presented compared to male 16p11.2 hemideletion mice. This pattern of responding was significant throughout delay discounting, but less prevalent in probability discounting. We have two competing theories for why this might be. Reduced nonreinforced responding could be due to a motivational difference in responding, as has been previously found in male 16p11.2 hemideletion mice (Grissom et al., 2018). It could be that the nonreinforced responding reflects response vigor, which would explain a response deficit in male 16p11.2 hemideletion mice. A second competing theory we had was that nonreinforced responding during this period reflects an attentional difference. Male 16p11.2 hemideletion mice have previously been shown to have an increased capacity for identifying periods in which to respond (Openshaw et al., 2023). The intertrial interval

period is hypothesized to be a period of reduced attention in animals, which immediately precedes the response initiation period (Blanchard et al., 2013). If male 16p11.2 hemideletion mice display increased attentional capabilities in other tasks, it would follow that they would also show an enhanced capability to identify response requirements as determined by the task structure. Importantly, this difference in nonreinforced responding diminished in probability discounting, and we believe this is due to responses being less indicative of guaranteed rewards. Taken together, these results indicate predictability of the reward modulate how male 16p11.2 hemideletion mice learn and adapt to costs. Male 16p11.2 hemideletion mice potentially pay more attention to stimuli that signal for task progression and display fixed responses to them if they are highly predictive of reward.

4.3 Significance Statement

My research has provided a within-subject approach to studying discounting in delay and probability discounting in mice. This approach is especially important to genetic mouse models because some choice or response patterns emerge over different periods of delay and probability discounting. My significant contributions are: 1) development of discounting tasks in mice, 2) leveraging a computational approach to provide a pipeline for understanding alternative sources of value in discounting tasks, and 3) demonstrating reinforcement related phenotypes observed in male 16p11.2 hemideletion mice appear in choice tasks and manifest based on the predictability of reinforcement.

Rodent versions of discounting tasks enable researchers to explore genetic models in two common types of decisions that have shown to be impacted by the presence of individual factors like sex and neurodevelopmental disorders in humans (Demurie et al., 2012; Story et al., 2015; Weafer & Wit, 2014). It is essential to identify cases in

which these factors interact to evaluate the impact on decision-making processes (Grissom & Reyes, 2019). In humans, there is still contention as to the experimental parameters and designs best suited for describing influences on decision-making (Madden & Johnson, 2010; Steele et al., 2019). Significant efforts however have been made to identify sources of discounting and provide evidence for a general discounting rate across task structures (Craig et al., 2014; Patt et al., 2021; Steele et al., 2019). Rodent tasks have been utilized to explore the experiential version of discounting, which is argued to reflect decision-making in real-time (Odum, 2011). Cross-species analyses have demonstrated animal discounting in general (including rodents) reflects human discounting (Vanderveldt et al., 2016). This translatability empowers computational approaches to understand drivers of discounting other than reward value (Chen et al., 2021). Such an approach may be valuable when assessing discounting behavior (Schoemann & Scherbaum, 2019; Story et al., 2015). My research has provided additional support for this approach by demonstrating how choice behavior shapes the discounting function in decision-making tasks. Such an approach should help ameliorate some cross-task implications of choice behavior and provide an alternative explanation for how discounting rates emerge and stabilize (Green & Myerson, 2013; Strickland & Johnson, 2021).

Rodent models have been essential to exploring neurodevelopmental influences on behavior, such as 16p11.2 hemideletion models. 16p11.2 hemideletion is frequently associated with developmental issues and is conserved genetically across species in rodents (Hanson et al., 2015; Horev et al., 2011; Rein & Yan, 2020). Although there have been comprehensive efforts to understand changes in learning and alterations in the underlying neurobiology of 16p11.2 hemideletion mice, these studies have yet to explore whether observed changes in learning and task engagement appear in more complex situations such as opportunity-cost scenarios (Angelakos et al., 2017; Grissom

et al., 2018; Portmann et al., 2014). My research adds to this growing literature by demonstrating that male 16p11.2 hemideletion mice are more willing to endure temporal costs than probabilistic costs compared to male wildtype mice. Surprisingly, the nonreinforcement phenotype found by Grissom et al. (2018) in a simple reinforcement scenario was replicated in my study where animals have to track different phases of trial progression. These findings implicate enduring learning and motivational processes that are replicated consistently across different rewarding scenarios.

4.4 Future Directions

The future of behavioral and neurological analysis of 16p11.2 hemideletion mice is exciting and retains strong implications for neurodevelopmental disorders. In this section, I will briefly explore some alternative approaches to understanding the source of some behaviors explained in this dissertation.

One immediate direction comes from the parallels of discounting and foraging. Some possible explanations from foraging might aid in understanding the stickiness in choice during delay discounting in 16p11.2 hemideletion mice. For instance, organisms in a foraging context "overharvest" or stay too long on a depleting patch. However, evidence suggests that as task structure is learned and experience is gained, subjects will learn to "overharvest" less and subsequently shift their choice more (Harhen & Bornstein, 2023). In delay discounting tasks, there can exist a deficit in task structure comprehension that can inadvertently produce steep discounting (Blanchard et al., 2013). Explicit cueing of an intertrial period can produce shallower discounting (Pearson et al., 2010). One interesting implication is that, because male 16p11.2 hemideletion mice 1) persevere for large rewards associated with large delays and 2) produce less nonreinforced responses during choice than wildtype mice, they may have

a better understanding of the task structure than wildtype mice through their utilization of discounting epoch cues. Increased task understanding may also aid in their choice behavior (e.g., exploiting a non-depleting large reward). In a capped deterministic scenario, this can result in a greater accumulation of rewards. When that sense of environmental volatility increases, such as in a probabilistic setting, male 16p11.2 hemideletion mice instead stick to guaranteed options. Additionally, male 16p11.2 hemidetion mice potentially pay less attention to trial cues and compensate by adopting a male wildtype pattern of trial initiation. One method to isolate the effects of intertrial interval versus delay would be to have mice experience them separately. One group found discounting to only occur when delays were incurred rather and that trial length did not have devaluation effects on the large reward (Sjoberg et al., 2021). Another thing they do is explicitly cue the intertrial interval, which could aid in attention and perhaps reduce the need for wildtype mice to commit extra nonreinforced responding. In order to explore the effect of reward probability on choice, future research can explore stay-leave decisions in a bandit style task as well, in cases where reward probability shifts throughout sessions.

The decision-making and nonreinforced responding of male 16p11.2 hemideletion mice were shown to be modulated by increased uncertainty of reward delivery. One intriguing avenue for future studies would be to challenge 16p11.2 hemideletion mice with a bandit task. In a restless bandit, probabilities for rewards are constantly shifting as animals progress through a schedule (Chen et al., 2021; Speekenbrink & Konstantinidis, 2015). This means that there are periods in which reward certainty for one option is high and periods where it is low. Such environmental changes to volatility tests how male 16p11.2 hemideletion mice would adapt to phases of uncertainty and helps determine global strategies used to make choices. Additionally, bandit tasks have been modified to incorporate reward epochs like an initiation hold

response used to assess attention and motivation (Faust et al., 2023). This alternative path would allow for additional exploration into those phenotypes of male 16p11.2 hemideletion mice. Our lab has thoroughly developed this method, and this would be a logical next step.

5. Bibliography

Acheson, A., Farrar, A. M., Patak, M., Hausknecht, K. A., Kieres, A. K., Choi, S., Wit, H. de, & Richards, J. B. (2006). Nucleus accumbens lesions decrease sensitivity to rapid changes in the delay to reinforcement. *Behavioural Brain Research*, *173*(2), 217–228. https://doi.org/10.1016/j.bbr.2006.06.024

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

Andrade, L. F., & Petry, N. M. (2012). Delay and probability discounting in pathological gamblers with and without a history of substance use problems. *Psychopharmacology*, *219*(2), 491–499. https://doi.org/10.1007/s00213-011-2508-9

Angelakos, C. C., Watson, A. J., O'Brien, W. T., Krainock, K. S., Nickl-Jockschat, T., & Abel, T. (2017). Hyperactivity and male-specific sleep deficits in the 16p11.2 deletion mouse model of autism. *Autism Research: Official Journal of the International Society for Autism Research*, *10*(4), 572–584. https://doi.org/10.1002/aur.1707

Almiron-Roig, E., Chen, Y., & Drewnowski, A. (2003). Liquid calories and the failure of satiety: How good is the evidence? *Obesity Reviews: An Official Journal of the International Association for the Study of Obesity*, *4*(4), 201–212. https://doi.org/10.1046/j.1467-789x.2003.00112.x

Aparicio, C. F., Hennigan, P. J., Mulligan, L. J., & Alonso-Alvarez, B. (2019). Spontaneously hypertensive (SHR) rats choose more impulsively than Wistar-Kyoto (WKY) rats on a delay discounting task. *Behavioural Brain Research*, *364*, 480–493. https://doi.org/10.1016/j.bbr.2017.09.040

Bailey, A. J., Romeu, R. J., & Finn, P. R. (2021). The problems with delay discounting: A critical review of current practices and clinical applications. *Psychological Medicine*, *51*(11), 1799–1806. https://doi.org/10.1017/S0033291721002282

Bailey, A. J., Romeu, R. J., & Finn, P. R. (2022). The fundamental questions left unanswered: Response to commentary on the 'problems with delay discounting'. *Psychological Medicine*, 1–2. https://doi.org/10.1017/S0033291721005572

Bagley, J. R., Khan, A. H., Smith, D. J., & Jentsch, J. D. (2022). Extreme phenotypic diversity in operant response to intravenous cocaine or saline infusion in the hybrid mouse diversity panel. *Addiction Biology*, *27*(3). https://doi.org/10.1111/adb.13162

Balleine, B. W., & Dezfouli, A. (2019). Hierarchical action control: Adaptive collaboration between actions and habits. *Frontiers in Psychology*, *10*, 2735. https://doi.org/10.3389/fpsyg.2019.02735

Basile, A. G., & Toplak, M. E. (2015). Four converging measures of temporal discounting and their relationships with intelligence, executive functions, thinking dispositions, and behavioral outcomes. *Frontiers in Psychology*, *6*, 728. https://doi.org/10.3389/fpsyg.2015.00728

Baumeister, S., Moessnang, C., Bast, N., Hohmann, S., Aggensteiner, P., Kaiser, A., Tillmann, J., Goyard, D., Charman, T., Ambrosino, S., Baron-Cohen, S., Beckmann, C., Bölte, S., Bourgeron, T., Rausch, A., Crawley, D., Dell'Acqua, F.,

Dumas, G., Durston, S., … EU-AIMS LEAP Group. (2023). Processing of social and monetary rewards in autism spectrum disorders. *The British Journal of Psychiatry: The Journal of Mental Science*, *222*(3), 100–111. https://doi.org/10.1192/bjp.2022.157

Benedetti, A., Molent, C., Barcik, W., & Papaleo, F. (2022). Social behavior in 16p11.2 and 22q11.2 copy number variations: Insights from mice and humans. *Genes, Brain, and Behavior*, *21*(5), e12787. https://doi.org/10.1111/gbb.12787

Bergh, F. S. van den, Bloemarts, E., Chan, J. S. W., Groenink, L., Olivier, B., & Oosting, R. S. (2006). Spontaneously hypertensive rats do not predict symptoms of attention-deficit hyperactivity disorder. *Pharmacology, Biochemistry, and Behavior*, *83*(3), 380–390. https://doi.org/10.1016/j.pbb.2006.02.018

Berridge, K. C., & Robinson, T. E. (2016). Liking, wanting, and the incentive-sensitization theory of addiction. *The American Psychologist*, *71*(8), 670–679. https://doi.org/10.1037/amp0000059

Bertero, A., Liska, A., Pagani, M., Parolisi, R., Masferrer, M. E., Gritti, M., Pedrazzoli, M., Galbusera, A., Sarica, A., Cerasa, A., Buffelli, M., Tonini, R., Buffo, A., Gross, C., Pasqualetti, M., & Gozzi, A. (2018). Autism-associated 16p11.2 microdeletion impairs prefrontal functional connectivity in mouse and human. *Brain: A Journal of Neurology*, *141*(7), 2055–2065. https://doi.org/10.1093/brain/awy111

Bervoets, J., Milton, D., & Van de Cruys, S. (2021). Autism and intolerance of uncertainty: An ill-fitting pair. *Trends in Cognitive Sciences*, *25*(12), 1009–1010. https://doi.org/10.1016/j.tics.2021.08.006

Białaszek, W., Ostaszewski, P., Green, L., & Myerson, J. (2019). On four types of devaluation of outcomes due to their costs: Delay, probability, effort, and social discounting. *The Psychological Record*, *69*(3), 415–424. https://doi.org/10.1007/s40732-019-00340-x

Blanchard, T. C., Pearson, J. M., & Hayden, B. Y. (2013). Postreward delays and systematic biases in measures of animal temporal discounting. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(38), 15491–15496. https://doi.org/10.1073/pnas.1310446110

Bos, R. van den, Koot, S., & Visser, L. de. (2014). A rodent version of the iowa gambling task: 7 years of progress. *Frontiers in Psychology*, *5*, 203. https://doi.org/10.3389/fpsyg.2014.00203

Botvinick, M. M., Huffstetler, S., & McGuire, J. T. (2009). Effort discounting in human nucleus accumbens. *Cognitive, Affective & Behavioral Neuroscience*, *9*(1), 16–27. https://doi.org/10.3758/CABN.9.1.16

Braunscheidel, K. M., Okas, M. P., Hoffman, M., Mulholland, P. J., Floresco, S. B., & Woodward, J. J. (2019). The abused inhalant toluene impairs medial prefrontal cortex activity and Risk/Reward Decision-Making during a probabilistic discounting task. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *39*(46), 9207–9220. https://doi.org/10.1523/JNEUROSCI.1674-19.2019

Calvert, A. L., Green, L., & Myerson, J. (2010). Delay discounting of qualitatively different reinforcers in rats. *Journal of the Experimental Analysis of Behavior*, *93*(2), 171–184. https://doi.org/10.1901/jeab.2010.93-171

Carlisi, C. O., Norman, L., Murphy, C. M., Christakou, A., Chantiluke, K., Giampietro, V., Simmons, A., Brammer, M., Murphy, D. G., MRC AIMS consortium, Mataix-Cols, D., & Rubia, K. (2017). Comparison of neural substrates of temporal discounting between youth with autism spectrum disorder and with obsessive-compulsive disorder. *Psychological Medicine*, *47*(14), 2513–2527. https://doi.org/10.1017/S0033291717001088

Castro Paiva, G. C. de, Souza Costa, D. de, Malloy-Diniz, L. F., Marques de Miranda, D., & Jardim de Paula, J. (2019). Temporal reward discounting in children with attention Deficit/Hyperactivity disorder (ADHD), and children with autism spectrum disorder (ASD): A systematic review. *Developmental Neuropsychology*, *44*(6), 468–480. https://doi.org/10.1080/87565641.2019.1667996

Charlton, S. R., & Fantino, E. (2008). Commodity specific rates of temporal discounting: Does metabolic function underlie differences in rates of discounting? *Behavioural Processes*, *77*(3), 334–342. https://doi.org/10.1016/j.beproc.2007.08.002

Chawner, S. J. R. A., Owen, M. J., Holmans, P., Raymond, F. L., Skuse, D., Hall, J., & Bree, M. B. M. van den. (2019). Genotype-phenotype associations in children with copy number variants associated with high neuropsychiatric risk in the UK (IMAGINE-ID): A case-control cohort study. *The Lancet. Psychiatry*, *6*(6), 493–505. https://doi.org/10.1016/S2215-0366(19)30123-3

Chen, C. S., Ebitz, R. B., Bindas, S. R., Redish, A. D., Hayden, B. Y., & Grissom, N. M. (2020). Divergent strategies for learning in males and females. *Current Biology: CB*. https://doi.org/10.1016/j.cub.2020.09.075

Chen, C. S., Knep, E., Han, A., Ebitz, R. B., & Grissom, N. M. (2021). Sex differences in learning from exploration. *eLife*, *10*. https://doi.org/10.7554/eLife.69748

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, *16*(4), 231–239. https://doi.org/10.1016/j.tics.2012.02.007

Cinotti, F., Fresno, V., Aklil, N., Coutureau, E., Girard, B., Marchand, A. R., & Khamassi, M. (2019). Dopamine blockade impairs the exploration-exploitation trade-off in rats. *Scientific Reports*, *9*(1), 6770. https://doi.org/10.1038/s41598-019-43245-z

Clements, C. C., Ascunce, K., & Nelson, C. A. (2022). In context: A developmental model of reward processing, with implications for autism and sensitive periods. *Journal of the American Academy of Child and Adolescent Psychiatry*. https://doi.org/10.1016/j.jaac.2022.07.861

Clements, C. C., Zoltowski, A. R., Yankowitz, L. D., Yerys, B. E., Schultz, R. T., & Herrington, J. D. (2018). Evaluation of the social motivation hypothesis of autism: A systematic review and meta-analysis. *JAMA Psychiatry*, *75*(8), 797–808. https://doi.org/10.1001/jamapsychiatry.2018.1100

Craig, A. R., Maxfield, A. D., Stein, J. S., Renda, C. R., & Madden, G. J. (2014). Do the adjusting-delay and increasing-delay tasks measure the same construct: Delay discounting? *Behavioural Pharmacology*, *25*(4), 306–315. https://doi.org/10.1097/FBP.0000000000000055

Dalley, J. W., Everitt, B. J., & Robbins, T. W. (2011). Impulsivity, compulsivity, and top-down cognitive control. *Neuron*, *69*(4), 680–694. https://doi.org/10.1016/j.neuron.2011.01.020

Damiano, C. R., Aloi, J., Treadway, M., Bodfish, J. W., & Dichter, G. S. (2012). Adults with autism spectrum disorders exhibit decreased sensitivity to reward parameters when making effort-based decisions. *Journal of Neurodevelopmental Disorders*, *4*(1), 13. https://doi.org/10.1186/1866-1955-4-13

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Daw, N. D., & Touretzky, D. S. (2000). Behavioral considerations suggest an average reward TD model of the dopamine system. *Neurocomputing*, *32–33*, 679–684. https://doi.org/10.1016/S0925-2312(00)00232-0

D'Cruz, A.-M., Ragozzino, M. E., Mosconi, M. W., Shrestha, S., Cook, E. H., & Sweeney, J. A. (2013). Reduced behavioral flexibility in autism spectrum disorders. *Neuropsychology*, *27*(2), 152–160. https://doi.org/10.1037/a0031721

Demurie, E., Roeyers, H., Baeyens, D., & Sonuga-Barke, E. (2012). Temporal discounting of monetary rewards in children and adolescents with ADHD and autism spectrum disorders. *Developmental Science*, *15*(6), 791–800. https://doi.org/10.1111/j.1467-7687.2012.01178.x

Dichter, G. S., Damiano, C. A., & Allen, J. A. (2012). Reward circuitry dysfunction in psychiatric and neurodevelopmental disorders and genetic syndromes: Animal models and clinical findings. *Journal of Neurodevelopmental Disorders*, *4*(1), 19. https://doi.org/10.1186/1866-1955-4-19

Dixon, M. R., Lik, N. M. K., Green, L., & Myerson, J. (2013). Delay discounting of hypothetical and real money: The effect of holding reinforcement rate constant.

Journal of Applied Behavior Analysis, *46*(2), 512–517. https://doi.org/10.1002/jaba.42

Duyzend, M. H., & Eichler, E. E. (2015). Genotype-first analysis of the 16p11.2 deletion defines a new type of "autism." *Biological Psychiatry*, *77*(9), 769–771. https://doi.org/10.1016/j.biopsych.2015.02.032

Ebitz, R. B., Albarran, E., & Moore, T. (2018). Exploration disrupts Choice-Predictive signals and alters dynamics in prefrontal cortex. *Neuron*, *97*(2), 450-461.e9. https://doi.org/10.1016/j.neuron.2017.12.007

Ebitz, R. B., Sleezer, B. J., Jedema, H. P., Bradberry, C. W., & Hayden, B. Y. (2019). Tonic exploration governs both flexibility and lapses. *PLoS Computational Biology*, *15*(11), e1007475. https://doi.org/10.1371/journal.pcbi.1007475

Estle, S. J., Green, L., Myerson, J., & Holt, D. D. (2007). Discounting of monetary and directly consumable rewards. *Psychological Science*, *18*(1), 58–63. https://doi.org/10.1111/j.1467-9280.2007.01849.x

Eubig, P. A., Noe, T. E., Floresco, S. B., Sable, J. J., & Schantz, S. L. (2014). Sex differences in response to amphetamine in adult Long-Evans rats performing a delay-discounting task. *Pharmacology, Biochemistry, and Behavior*, *118*, 1–9. https://doi.org/10.1016/j.pbb.2013.12.021

Evenden, J. L., & Ryan, C. N. (1996). The pharmacology of impulsive behaviour in rats: The effects of drugs on response choice with varying delays of reinforcement. *Psychopharmacology*, *128*(2), 161–170. https://doi.org/10.1007/s002130050121

Findling, C., Skvortsova, V., Dromnelle, R., Palminteri, S., & Wyart, V. (2019). Computational noise in reward-guided learning drives behavioral variability in volatile environments. *Nature Neuroscience*, *22*(12), 2066–2077. https://doi.org/10.1038/s41593-019-0518-9

Fontanesi, L., Gluth, S., Spektor, M. S., & Rieskamp, J. (2019). A reinforcement learning diffusion decision model for value-based decisions. *Psychonomic Bulletin & Review*, *26*(4), 1099–1121. https://doi.org/10.3758/s13423-018-1554-2

Fox, A. T., Hand, D. J., & Reilly, M. P. (2008). Impulsive choice in a rodent model of attention-deficit/hyperactivity disorder. *Behavioural Brain Research*, *187*(1), 146–152. https://doi.org/10.1016/j.bbr.2007.09.008

Frye, C. C. J., Galizio, A., Friedel, J. E., DeHart, W. B., & Odum, A. L. (2016). Measuring delay discounting in humans using an adjusting amount task. *Journal of Visualized Experiments: JoVE*, *107*. https://doi.org/10.3791/53584

Fuccillo, M. V. (2016). Striatal circuits as a common node for autism pathophysiology. *Frontiers in Neuroscience*, *10*, 27. https://doi.org/10.3389/fnins.2016.00027

Garr, E., Bushra, B., Tu, N., & Delamater, A. R. (2020). Goal-directed control on interval schedules does not depend on the action-outcome correlation. *Journal of Experimental Psychology. Animal Learning and Cognition*, *46*(1), 47–64. https://doi.org/10.1037/xan0000229

Gershman, S. J. (2019). Uncertainty and exploration. *Decisions*, *6*(3), 277–286. https://doi.org/10.1037/dec0000101

Ghods-Sharifi, S., & Floresco, S. B. (2010). Differential effects on effort discounting induced by inactivations of the nucleus accumbens core or shell. *Behavioral*

Neuroscience, *124*(2), 179–191. https://doi.org/10.1037/a0018932

Giovanniello, J., Ahrens, S., Yu, K., & Li, B. (2021). Sex-Specific Stress-Related behavioral phenotypes and central amygdala dysfunction in a mouse model of 16p11.2 microdeletion. *Biological Psychiatry Global Open Science*, *1*(1), 59–69. https://doi.org/10.1016/j.bpsgos.2021.01.001

Gosling, C. J., & Moutier, S. (2018). Brief report: Risk-Aversion and rationality in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *48*(10), 3623–3628. https://doi.org/10.1007/s10803-018-3616-8

Gray, J. C., Amlung, M. T., Palmer, A. A., & MacKillop, J. (2016). Syntax for calculation of discounting indices from the monetary choice questionnaire and probability discounting questionnaire. *Journal of the Experimental Analysis of Behavior*, *106*(2), 156–163. https://doi.org/10.1002/jeab.221

Green, L., Fry, A. F., & Myerson, J. (1994). Discounting of delayed rewards: A Life-Span comparison. *Psychological Science*, *5*(1), 33–36. https://doi.org/10.1111/j.1467-9280.1994.tb00610.x

Green, L., & Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin*, *130*(5), 769–792. https://doi.org/10.1037/0033-2909.130.5.769

Green, L., & Myerson, J. (2013). How many impulsivities? A discounting perspective. *Journal of the Experimental Analysis of Behavior*, *99*(1), 3–13. https://doi.org/10.1002/jeab.1

Green, L., Myerson, J., Shah, A. K., Estle, S. J., & Holt, D. D. (2007). Do adjusting-amount and adjusting-delay procedures produce equivalent estimates of subjective

value in pigeons? *Journal of the Experimental Analysis of Behavior*, *87*(3), 337–347. https://doi.org/10.1901/jeab.2007.37-06

Green, L., Myerson, J., & Vanderveldt, A. (2014). Delay and probability discounting. In F. K. McSweeney (Ed.), *The wiley blackwell handbook of operant and classical conditioning , (pp* (Vol. 738, pp. 307–337). https://doi.org/10.1002/9781118468135.ch13

Grissom, N. M., McKee, S. E., Schoch, H., Bowman, N., Havekes, R., O'Brien, W. T., Mahrt, E., Siegel, S., Commons, K., Portfors, C., Nickl-Jockschat, T., Reyes, T. M., & Abel, T. (2018). Male-specific deficits in natural reward learning in a mouse model of neurodevelopmental disorders. *Molecular Psychiatry*, *23*(3), 544–555. https://doi.org/10.1038/mp.2017.184

Grissom, N. M., & Reyes, T. M. (2019). Let's call the whole thing off: Evaluating gender and sex differences in executive function. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *44*(1), 86–96. https://doi.org/10.1038/s41386-018-0179-5

Grüne-Yanoff, T. (2015). Models of temporal discounting 1937-2000: An interdisciplinary exchange between economics and psychology. *Science in Context*, *28*(4), 675–713. https://doi.org/10.1017/S0269889715000307

Hall-McMaster, S., Dayan, P., & Schuck, N. W. (2021). Control over patch encounters changes foraging behavior. *iScience*, *24*(9), 103005. https://doi.org/10.1016/j.isci.2021.103005

Hanson, E., Bernier, R., Porche, K., Jackson, F. I., Goin-Kochel, R. P., Snyder, L. G., Snow, A. V., Wallace, A. S., Campe, K. L., Zhang, Y., Chen, Q., D'Angelo, D., Moreno-De-Luca, A., Orr, P. T., Boomer, K. B., Evans, D. W., Kanne, S., Berry,

L., Miller, F. K., … Simons Variation in Individuals Project Consortium. (2015). The cognitive and behavioral phenotype of the 16p11.2 deletion in a clinically ascertained population. *Biological Psychiatry*, *77*(9), 785–793. https://doi.org/10.1016/j.biopsych.2014.04.021

Hayden, B. Y. (2016). Time discounting and time preference in animals: A critical review. *Psychonomic Bulletin & Review*, *23*(1), 39–53. https://doi.org/10.3758/s13423-015-0879-3

Hayden, B. Y. (2018). Economic choice: The foraging perspective. *Current Opinion in Behavioral Sciences*, *24*, 1–6. https://doi.org/10.1016/j.cobeha.2017.12.002

Hayden, B., & Niv, Y. (2020). *The case against economic values in the orbitofrontal cortex (or anywhere else in the brain)*. https://doi.org/10.31234/osf.io/7hgup

Hayden, B. Y., & Platt, M. L. (2007). Temporal discounting predicts risk sensitivity in rhesus macaques. *Current Biology: CB*, *17*(1), 49–53. https://doi.org/10.1016/j.cub.2006.10.055

Heilbronner, S. R. (2017). Modeling risky decision-making in nonhuman animals: Shared core features. *Current Opinion in Behavioral Sciences*, *16*, 23–29. https://doi.org/10.1016/j.cobeha.2017.03.001

Heilbronner, S. R., Hayden, B. Y., & Platt, M. L. (2010). Neuroeconomics of risk-sensitive decision making. In G. J. Madden (Ed.), *Impulsivity: The behavioral and neurological science of discounting , (pp* (Vol. 453, pp. 159–187). American Psychological Association, xvi. https://doi.org/10.1037/12069-006

Helms, C. M., Reeves, J. M., & Mitchell, S. H. (2006). Impact of strain and d-amphetamine on impulsivity (delay discounting) in inbred mice.

Psychopharmacology, *188*(2), 144–151. https://doi.org/10.1007/s00213-006-0478-0

Hernandez, C. M., Orsini, C., Wheeler, A.-R., Ten Eyck, T. W., Betzhold, S. M., Labiste, C. C., Wright, N. G., Setlow, B., & Bizon, J. L. (2020). Testicular hormones mediate robust sex differences in impulsive choice in rats. *eLife*, *9*. https://doi.org/10.7554/eLife.58604

Hinvest, N. S., & Anderson, I. M. (2010). The effects of real versus hypothetical reward on delay and probability discounting. *Quarterly Journal of Experimental Psychology*, *63*(6), 1072–1084. https://doi.org/10.1080/17470210903276350

Hogarth, L., Chase, H. W., & Baess, K. (2012). Impaired goal-directed behavioural control in human impulsivity. *Quarterly Journal of Experimental Psychology*, *65*(2), 305–316. https://doi.org/10.1080/17470218.2010.518242

Holt, D. D., Green, L., & Myerson, J. (2012). Estimating the subjective value of future rewards: Comparison of adjusting-amount and adjusting-delay procedures. *Behavioural Processes*, *90*(3), 302–310. https://doi.org/10.1016/j.beproc.2012.03.003

Horan, W. P., Johnson, M. W., & Green, M. F. (2017). Altered experiential, but not hypothetical, delay discounting in schizophrenia. *Journal of Abnormal Psychology*, *126*(3), 301–311. https://doi.org/10.1037/abn0000249

Horev, G., Ellegood, J., Lerch, J. P., Son, Y.-E. E., Muthuswamy, L., Vogel, H., Krieger, A. M., Buja, A., Henkelman, R. M., Wigler, M., & Mills, A. A. (2011). Dosage-dependent phenotypes in models of 16p11.2 lesions found in autism. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(41), 17076–17081. https://doi.org/10.1073/pnas.1114042108

Horner, A. E., Heath, C. J., Hvoslef-Eide, M., Kent, B. A., Kim, C. H., Nilsson, S. R. O., Alsiö, J., Oomen, C. A., Holmes, A., Saksida, L. M., & Bussey, T. J. (2013). The touchscreen operant platform for testing learning and memory in rats and mice. *Nature Protocols*, *8*(10), 1961–1984. https://doi.org/10.1038/nprot.2013.122

Hutsell, B. A., & Newland, M. C. (2013). A quantitative analysis of the effects of qualitatively different reinforcers on fixed ratio responding in inbred strains of mice. *Neurobiology of Learning and Memory*, *101*, 85–93. https://doi.org/10.1016/j.nlm.2013.01.005

Islas-Preciado, D., Wainwright, S. R., Sniegocki, J., Lieblich, S. E., Yagi, S., Floresco, S. B., & Galea, L. A. M. (2020). Risk-based decision making in rats: Modulation by sex and amphetamine. *Hormones and Behavior*, *125*, 104815. https://doi.org/10.1016/j.yhbeh.2020.104815

Isles, A. R., Humby, T., Walters, E., & Wilkinson, L. S. (2004). Common genetic effects on variation in impulsivity and activity in mice. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *24*(30), 6733–6740. https://doi.org/10.1523/JNEUROSCI.1650-04.2004

Jackson, J. N. S., & MacKillop, J. (2016). Attention-Deficit/Hyperactivity disorder and monetary delay discounting: A Meta-Analysis of Case-Control studies. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, *1*(4), 316–325. https://doi.org/10.1016/j.bpsc.2016.01.007

Kalueff, A. V., Stewart, A. M., Song, C., Berridge, K. C., Graybiel, A. M., & Fentress, J. C. (2016). Neurobiology of rodent self-grooming and its value for translational neuroscience. *Nature Reviews. Neuroscience*, *17*(1), 45–59. https://doi.org/10.1038/nrn.2015.8

Karlsson Linnér, R., Mallard, T. T., Barr, P. B., Sanchez-Roige, S., Madole, J. W., Driver, M. N., Poore, H. E., Vlaming, R. de, Grotzinger, A. D., Tielbeek, J. J., Johnson, E. C., Liu, M., Rosenthal, S. B., Ideker, T., Zhou, H., Kember, R. L., Pasman, J. A., Verweij, K. J. H., Liu, D. J., … Dick, D. M. (2021). Multivariate analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction. *Nature Neuroscience*, *24*(10), 1367–1376. https://doi.org/10.1038/s41593-021-00908-3

Killeen, P. R. (2023a). From data through discount rates to the area under the curve. *Journal of the Experimental Analysis of Behavior*. https://doi.org/10.1002/jeab.888

Killeen, P. R. (2023b). Variations on a theme by rachlin: Probability discounting. *Journal of the Experimental Analysis of Behavior*, *119*(1), 140–155. https://doi.org/10.1002/jeab.817

Kilpatrick, Z. P., Davidson, J. D., & El Hady, A. (2021). Uncertainty drives deviations in normative foraging decision strategies. *Journal of the Royal Society, Interface / the Royal Society*, *18*(180), 20210337. https://doi.org/10.1098/rsif.2021.0337

Kirby, K. N. (1997). Bidding on the future: Evidence against normative discounting of delayed rewards. *Journal of Experimental Psychology. General*, *126*(1), 54–70. https://doi.org/10.1037/0096-3445.126.1.54

Kirby, K. N. (2009). One-year temporal stability of delay-discount rates. *Psychonomic Bulletin & Review*, *16*(3), 457–462. https://doi.org/10.3758/PBR.16.3.457

Kirby, K. N., Petry, N. M., & Bickel, W. K. (1999). Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *Journal of*

Experimental Psychology. General, *128*(1), 78–87. https://doi.org/10.1037//0096-3445.128.1.78

Ko, D., & Wanat, M. J. (2016). Phasic dopamine transmission reflects initiation vigor and exerted effort in an action- and Region-Specific manner. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(7), 2202–2211. https://doi.org/10.1523/JNEUROSCI.1279-15.2016

Koffarnus, M. N., Jarmolowicz, D. P., Mueller, E. T., & Bickel, W. K. (2013). Changing delay discounting in the light of the competing neurobehavioral decision systems theory: A review. *Journal of the Experimental Analysis of Behavior*, *99*(1), 32–57. https://doi.org/10.1002/jeab.2

Koffarnus, M. N., Newman, A. H., Grundt, P., Rice, K. C., & Woods, J. H. (2011). Effects of selective dopaminergic compounds on a delay-discounting task. *Behavioural Pharmacology*, *22*(4), 300–311. https://doi.org/10.1097/FBP.0b013e3283473bcb

Kohls, G., Chevallier, C., Troiani, V., & Schultz, R. T. (2012). Social 'wanting' dysfunction in autism: Neurobiological underpinnings and treatment implications. *Journal of Neurodevelopmental Disorders*, *4*(1), 10. https://doi.org/10.1186/1866-1955-4-10

Krebs, C. A., & Anderson, K. G. (2012). Preference reversals and effects of d-amphetamine on delay discounting in rats. *Behavioural Pharmacology*, *23*(3), 228–240. https://doi.org/10.1097/FBP.0b013e32835342ed

Kreek, M. J., Nielsen, D. A., Butelman, E. R., & LaForge, K. S. (2005). Genetic influences on impulsivity, risk taking, stress responsivity and vulnerability to drug

abuse and addiction. *Nature Neuroscience*, *8*(11), 1450–1457. https://doi.org/10.1038/nn1583

Lawson, R. P., Mathys, C., & Rees, G. (2017). Adults with autism overestimate the volatility of the sensory environment. *Nature Neuroscience*, *20*(9), 1293–1299. https://doi.org/10.1038/nn.4615

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: A neural common currency for choice. *Current Opinion in Neurobiology*, *22*(6), 1027–1038. https://doi.org/10.1016/j.conb.2012.06.001

Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What is the Male-to-Female ratio in autism spectrum disorder? A systematic review and Meta-Analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*, *56*(6), 466–474. https://doi.org/10.1016/j.jaac.2017.03.013

Lopez-Cruz, L., Phillips, B. U., Hailwood, J. M., Saksida, L. M., Heath, C. J., & Bussey, T. J. (2023). Refining the study of decision-making in animals: Differential effects of d-amphetamine and haloperidol in a novel touchscreen-automated Rearing-Effort discounting (RED) task and the Fixed-Ratio effort discounting (FRED) task. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*. https://doi.org/10.1038/s41386-023-01707-z

MacKillop, J. (2016). The behavioral economics and neuroeconomics of alcohol use disorders. *Alcoholism, Clinical and Experimental Research*, *40*(4), 672–685. https://doi.org/10.1111/acer.13004

MacKillop, J., Weafer, J., C Gray, J., Oshri, A., Palmer, A., & Wit, H. de. (2016). The latent structure of impulsivity: Impulsive choice, impulsive action, and impulsive

personality traits. *Psychopharmacology*, *233*(18), 3361–3370. https://doi.org/10.1007/s00213-016-4372-0

Madden, G. J., Begotka, A. M., Raiff, B. R., & Kastern, L. L. (2003). Delay discounting of real and hypothetical rewards. *Experimental and Clinical Psychopharmacology*, *11*(2), 139–145. https://doi.org/10.1037/1064-1297.11.2.139

Madden, G. J., Petry, N. M., & Johnson, P. S. (2009). Probability discounting questionnaire. *Experimental and Clinical PsychopharmacologyJournal of Abnormal Psychology*. https://doi.org/10.1037/t10481-000

Maguire, D. R., Henson, C., & France, C. P. (2014). Effects of amphetamine on delay discounting in rats depend upon the manner in which delay is varied. *Neuropharmacology*, *87*, 173–179. https://doi.org/10.1016/j.neuropharm.2014.04.012

Mar, A. C., & Robbins, T. W. (2007). Delay discounting and impulsive choice in the rat. *Current Protocols in Neuroscience / Editorial Board, Jacqueline N. Crawley ... [et Al.]*, *Chapter 8*, Unit 8.22. https://doi.org/10.1002/0471142301.ns0822s39

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. *The Effect of Delay and of Intervening Events on Reinforcement Value.*, *344*, 55–73. https://psycnet.apa.org/fulltext/1986-98701-003.pdf

Mazur, J. E. (2000). Tradeoffs among delay, rate, and amount of reinforcement. *Behavioural Processes*, *49*(1), 1–10. https://doi.org/10.1016/s0376-6357(00)00070-x

McClure, S. M., Ericson, K. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2007). Time discounting for primary rewards. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(21), 5796–5804. https://doi.org/10.1523/JNEUROSCI.4246-06.2007

McDougle, S. D., & Collins, A. G. E. (2021). Modeling the influence of working memory, reinforcement, and action uncertainty on reaction time and choice during instrumental learning. *Psychonomic Bulletin & Review*, *28*(1), 20–39. https://doi.org/10.3758/s13423-020-01774-z

McKerchar, T. L., & Renda, C. R. (2012). Delay and probability discounting in humans: An overview. *The Psychological Record*, *62*(4), 817–834. https://doi.org/10.1007/BF03395837

Mikhael, J. G., Lai, L., & Gershman, S. J. (2021). Rational inattention and tonic dopamine. *PLoS Computational Biology*, *17*(3), e1008659. https://doi.org/10.1371/journal.pcbi.1008659

Miller, H. L., Ragozzino, M. E., Cook, E. H., Sweeney, J. A., & Mosconi, M. W. (2015). Cognitive set shifting deficits and their relationship to repetitive behaviors in autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(3), 805–815. https://doi.org/10.1007/s10803-014-2244-1

Minassian, A., Paulus, M., Lincoln, A., & Perry, W. (2007). Adults with autism show increased sensitivity to outcomes at low error rates during decision-making. *Journal of Autism and Developmental Disorders*, *37*(7), 1279–1288. https://doi.org/10.1007/s10803-006-0278-8

Mitchell, S. H. (2014). Assessing delay discounting in mice. *Current Protocols in Neuroscience / Editorial Board, Jacqueline N. Crawley ... [et Al.]*, *66*, Unit 8.30.

https://doi.org/10.1002/0471142301.ns0830s66

Mitchell, S. H. (2017). Devaluation of outcomes due to their cost: Extending discounting models beyond delay. *Nebraska Symposium on Motivation*. *Nebraska Symposium on Motivation*, *64*, 145–161. https://doi.org/10.1007/978-3-319-51721-6_5

Mohebi, A., Pettibone, J. R., Hamid, A. A., Wong, J.-M. T., Vinson, L. T., Patriarchi, T., Tian, L., Kennedy, R. T., & Berke, J. D. (2019). Dissociable dopamine dynamics for learning and motivation. *Nature*, *570*(7759), 65–70. https://doi.org/10.1038/s41586-019-1235-y

Moon, J., Beaudin, A. E., Verosky, S., Driscoll, L. L., Weiskopf, M., Levitsky, D. A., Crnic, L. S., & Strupp, B. J. (2006). Attentional dysfunction, impulsivity, and resistance to change in a mouse model of fragile X syndrome. *Behavioral Neuroscience*, *120*(6), 1367–1379. https://doi.org/10.1037/0735-7044.120.6.1367

Mosner, M. G., Kinard, J. L., McWeeny, S., Shah, J. S., Markiewitz, N. D., Damiano-Goodwin, C. R., Burchinal, M. R., Rutherford, H. J. V., Greene, R. K., Treadway, M. T., & Dichter, G. S. (2017). Vicarious Effort-Based Decision-Making in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *47*(10), 2992–3006. https://doi.org/10.1007/s10803-017-3220-3

Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S., & Churchland, A. K. (2019). Single-trial neural dynamics are dominated by richly varied movements. *Nature Neuroscience*, *22*(10), 1677–1686. https://doi.org/10.1038/s41593-019-0502-4

Mussey, J. L., Travers, B. G., Klinger, L. G., & Klinger, M. R. (2015). Decision-making skills in ASD: Performance on the iowa gambling task. *Autism Research:*

Official Journal of the International Society for Autism Research, *8*(1), 105–114. https://doi.org/10.1002/aur.1429

Myerson, J., Green, L., & Warusawitharana, M. (2001). Area under the curve as a measure of discounting. *Journal of the Experimental Analysis of Behavior*, *76*(2), 235–243. https://doi.org/10.1901/jeab.2001.76-235

Namboodiri, V. M. K., Mihalas, S., Marton, T. M., & Hussain Shuler, M. G. (2014). A general theory of intertemporal decision-making and the perception of time. *Frontiers in Behavioral Neuroscience*, *8*, 61. https://doi.org/10.3389/fnbeh.2014.00061

Neville, V., King, J., Gilchrist, I. D., Dayan, P., Paul, E. S., & Mendl, M. (2020). Reward and punisher experience alter rodent decision-making in a judgement bias task. *Scientific Reports*, *10*(1), 11839. https://doi.org/10.1038/s41598-020-68737-1

Niarchou, M., Chawner, S. J. R. A., Doherty, J. L., Maillard, A. M., Jacquemont, S., Chung, W. K., Green-Snyder, L., Bernier, R. A., Goin-Kochel, R. P., Hanson, E., Linden, D. E. J., Linden, S. C., Raymond, F. L., Skuse, D., Hall, J., Owen, M. J., & Bree, M. B. M. van den. (2019). Psychiatric disorders in children with 16p11.2 deletion and duplication. *Translational Psychiatry*, *9*(1), 8. https://doi.org/10.1038/s41398-018-0339-8

Nicola, S. M. (2010). The flexible approach hypothesis: Unification of effort and cue-responding hypotheses for the role of nucleus accumbens dopamine in the activation of reward-seeking behavior. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*(49), 16585–16600. https://doi.org/10.1523/JNEUROSCI.3958-10.2010

Niemi, M. E. K., Martin, H. C., Rice, D. L., Gallone, G., Gordon, S., Kelemen, M., McAloney, K., McRae, J., Radford, E. J., Yu, S., Gecz, J., Martin, N. G., Wright, C. F., Fitzpatrick, D. R., Firth, H. V., Hurles, M. E., & Barrett, J. C. (2018). Common genetic variants contribute to risk of rare severe neurodevelopmental disorders. *Nature*, *562*(7726), 268–271. https://doi.org/10.1038/s41586-018-0566-4

Nissan, N., Hertz, U., Shahar, N., & Gabay, Y. (2023). Distinct reinforcement learning profiles distinguish between language and attentional neurodevelopmental disorders. *Behavioral and Brain Functions: BBF*, *19*(1), 6. https://doi.org/10.1186/s12993-023-00207-w

Nussenbaum, K., & Hartley, C. A. (2019). Reinforcement learning across development: What insights can we draw from a decade of research? *Developmental Cognitive Neuroscience*, *40*, 100733. https://doi.org/10.1016/j.dcn.2019.100733

Odum, A. L. (2011a). Delay discounting: I'm a k, you're a k. *Journal of the Experimental Analysis of Behavior*, *96*(3), 427–439. https://doi.org/10.1901/jeab.2011.96-423

Odum, A. L. (2011b). Delay discounting: Trait variable? *Behavioural Processes*, *87*(1), 1–9. https://doi.org/10.1016/j.beproc.2011.02.007

Odum, A. L., Baumann, A. A. L., & Rimington, D. D. (2006). Discounting of delayed hypothetical money and food: Effects of amount. *Behavioural Processes*, *73*(3), 278–284. https://doi.org/10.1016/j.beproc.2006.06.008

Odum, A. L., Becker, R. J., Haynes, J. M., Galizio, A., Frye, C. C. J., Downey, H., Friedel, J. E., & Perez, D. M. (2020). Delay discounting of different outcomes:

Review and theory. *Journal of the Experimental Analysis of Behavior*, *113*(3), 657–679. https://doi.org/10.1002/jeab.589

Openshaw, R. L., Thomson, D. M., Bristow, G. C., Mitchell, E. J., Pratt, J. A., Morris, B. J., & Dawson, N. (2023). 16p11.2 deletion mice exhibit compromised fronto-temporal connectivity, GABAergic dysfunction, and enhanced attentional ability. *Communications Biology*, *6*(1), 557. https://doi.org/10.1038/s42003-023-04891-2

Orsini, C. A., & Setlow, B. (2017). Sex differences in animal models of decision making. *Journal of Neuroscience Research*, *95*(1–2), 260–269. https://doi.org/10.1002/jnr.23810

Orsini, C. A., Willis, M. L., Gilbert, R. J., Bizon, J. L., & Setlow, B. (2016). Sex differences in a rat model of risky decision making. *Behavioral Neuroscience*, *130*(1), 50–61. https://doi.org/10.1037/bne0000111

Pavăl, D. (2017). A dopamine hypothesis of autism spectrum disorder. *Developmental Neuroscience*, *39*(5), 355–360. https://doi.org/10.1159/000478725

Peters, J., & Büchel, C. (2009). Overlapping and distinct neural systems code for subjective value during intertemporal and risky decision making. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(50), 15727–15734. https://doi.org/10.1523/JNEUROSCI.3489-09.2009

Pearson, J. M., Hayden, B. Y., & Platt, M. L. (2010). Explicit information reduces discounting behavior in monkeys. *Frontiers in Psychology*, *1*, 237. https://doi.org/10.3389/fpsyg.2010.00237

Portmann, T., Yang, M., Mao, R., Panagiotakos, G., Ellegood, J., Dolen, G., Bader, P. L., Grueter, B. A., Goold, C., Fisher, E., Clifford, K., Rengarajan, P., Kalikhman,

D., Loureiro, D., Saw, N. L., Zhengqui, Z., Miller, M. A., Lerch, J. P., Henkelman, M., … Dolmetsch, R. E. (2014). Behavioral abnormalities and circuit defects in the basal ganglia of a mouse model of 16p11.2 deletion syndrome. *Cell Reports*, *7*(4), 1077–1092. https://doi.org/10.1016/j.celrep.2014.03.036

Posserud, M.-B., Skretting Solberg, B., Engeland, A., Haavik, J., & Klungsøyr, K. (2021). Male to female ratios in autism spectrum disorders by age, intellectual disability and attention-deficit/hyperactivity disorder. *Acta Psychiatrica Scandinavica*, *144*(6), 635–646. https://doi.org/10.1111/acps.13368

Prévost, C., Pessiglione, M., Météreau, E., Cléry-Melin, M.-L., & Dreher, J.-C. (2010). Separate valuation subsystems for delay and effort decision costs. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*(42), 14080–14090. https://doi.org/10.1523/JNEUROSCI.2752-10.2010

Pyke, G. H., & Stephens, D. W. (2019). Optimal foraging theory: Application and inspiration in human endeavors outside biology. In J. C. Choe (Ed.), *Encyclopedia of animal behavior (second edition)* (pp. 217–222). Academic Press. https://doi.org/10.1016/B978-0-12-809633-8.90161-4

Rachlin, H., & Jones, B. A. (2008). Social discounting and delay discounting. *Journal of Behavioral Decision Making*, *21*(1), 29–43. https://doi.org/10.1002/bdm.567

Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective probability and delay. *Journal of the Experimental Analysis of Behavior*, *55*(2), 233–244. https://doi.org/10.1901/jeab.1991.55-233

Rein, B., & Yan, Z. (2020). 16p11.2 copy number variations and neurodevelopmental disorders. *Trends in Neurosciences*, *43*(11), 886–901. https://doi.org/10.1016/j.tins.2020.09.001

Richards, J. B., Mitchell, S. H., Wit, H. de, & Seiden, L. S. (1997). Determination of discount functions in rats with an adjusting-amount procedure. *Journal of the Experimental Analysis of Behavior*, *67*(3), 353–366. https://doi.org/10.1901/jeab.1997.67-353

Richards, J. B., Zhang, L., Mitchell, S. H., & Wit, H. de. (1999). Delay or probability discounting in a model of impulsive behavior: Effect of alcohol. *Journal of the Experimental Analysis of Behavior*, *71*(2), 121–143. https://doi.org/10.1901/jeab.1999.71-121

Robles, E., & Vargas, P. A. (2007). Functional parameters of delay discounting assessment tasks: Order of presentation. *Behavioural Processes*, *75*(2), 237–241. https://doi.org/10.1016/j.beproc.2007.02.014

Robles, E., Vargas, P. A., & Bejarano, R. (2009). Within-subject differences in degree of delay discounting as a function of order of presentation of hypothetical cash rewards. *Behavioural Processes*, *81*(2), 260–263. https://doi.org/10.1016/j.beproc.2009.02.018

Rodzon, K., Berry, M. S., & Odum, A. L. (2011). Within-subject comparison of degree of delay discounting using titrating and fixed sequence procedures. *Behavioural Processes*, *86*(1), 164–167. https://doi.org/10.1016/j.beproc.2010.09.007

Rojas, G. R., Curry-Pochy, L. S., Chen, C. S., Heller, A. T., & Grissom, N. M. (2022). Sequential delay and probability discounting tasks in mice reveal anchoring effects partially attributable to decision noise. *Behavioural Brain Research*, *431*, 113951. https://doi.org/10.1016/j.bbr.2022.113951

Rung, J. M., Frye, C. C. J., DeHart, W. B., & Odum, A. L. (2019). Evaluating the effect of delay spacing on delay discounting: Carry-over effects on steepness and the

form of the discounting function. *Journal of the Experimental Analysis of Behavior*, *112*(3), 254–272. https://doi.org/10.1002/jeab.556

Sadeghiyeh, H., Wang, S., Alberhasky, M. R., Kyllo, H. M., Shenhav, A., & Wilson, R. C. (2020). Temporal discounting correlates with directed exploration but not with random exploration. *Scientific Reports*, *10*(1), 4020. https://doi.org/10.1038/s41598-020-60576-4

Savatt, J. M., & Myers, S. M. (2021). Genetic testing in neurodevelopmental disorders. *Frontiers in Pediatrics*, *9*, 526779. https://doi.org/10.3389/fped.2021.526779

Schouppe, N., Demanet, J., Boehler, C. N., Ridderinkhof, K. R., & Notebaert, W. (2014). The role of the striatum in effort-based decision-making in the absence of reward. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(6), 2148–2154. https://doi.org/10.1523/JNEUROSCI.1214-13.2014

Scott-Van Zeeland, A. A., Dapretto, M., Ghahremani, D. G., Poldrack, R. A., & Bookheimer, S. Y. (2010). Reward processing in autism. *Autism Research: Official Journal of the International Society for Autism Research*, *3*(2), 53–67. https://doi.org/10.1002/aur.122

Shead, N. W., & Hodgins, D. C. (2009). Probability discounting of gains and losses: Implications for risk attitudes and impulsivity. *Journal of the Experimental Analysis of Behavior*, *92*(1), 1–16. https://doi.org/10.1901/jeab.2009.92-1

Silverman, J. L., Yang, M., Lord, C., & Crawley, J. N. (2010). Behavioural phenotyping assays for mouse models of autism. *Nature Reviews. Neuroscience*, *11*(7), 490–502. https://doi.org/10.1038/nrn2851

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. https://doi.org/10.1038/npp.2009.48

Sinha, P., Kjelgaard, M. M., Gandhi, T. K., Tsourides, K., Cardinaux, A. L., Pantazis, D., Diamond, S. P., & Held, R. M. (2014). Autism as a disorder of prediction. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(42), 15220–15225. https://doi.org/10.1073/pnas.1416797111

Sjoberg, E. A., Ramos, S., López-Tolsa, G. E., Johansen, E. B., & Pellón, R. (2021). The irrelevancy of the inter-trial interval in delay-discounting experiments on an animal model of ADHD. *Behavioural Brain Research*, *408*, 113236. https://doi.org/10.1016/j.bbr.2021.113236

Sjoberg, E., Ottåsen, H. M., Wilner, R. G., & Johansen, E. B. (2023). Previous experience with delays affects delay discounting in animal model of ADHD. *Behavioral and Brain Functions: BBF*, *19*(1), 4. https://doi.org/10.1186/s12993-022-00199-z

Slezak, J. M., & Anderson, K. G. (2009). Effects of variable training, signaled and unsignaled delays, and d-amphetamine on delay-discounting functions. *Behavioural Pharmacology*, *20*(5–6), 424–436. https://doi.org/10.1097/FBP.0b013e3283305ef9

Smits, R. R., Stein, J. S., Johnson, P. S., Odum, A. L., & Madden, G. J. (2013). Test-retest reliability and construct validity of the experiential discounting task.

Experimental and Clinical Psychopharmacology, *21*(2), 155–163. https://doi.org/10.1037/a0031725

South, M., Chamberlain, P. D., Wigham, S., Newton, T., Le Couteur, A., McConachie, H., Gray, L., Freeston, M., Parr, J., Kirwan, C. B., & Rodgers, J. (2014). Enhanced decision making and risk avoidance in high-functioning autism spectrum disorder. *Neuropsychology*, *28*(2), 222–228. https://doi.org/10.1037/neu0000016

Speekenbrink, M., & Konstantinidis, E. (2015). Uncertainty and exploration in a restless bandit problem. *Topics in Cognitive Science*, *7*(2), 351–367. https://doi.org/10.1111/tops.12145

St Onge, J. R., Chiu, Y. C., & Floresco, S. B. (2010). Differential effects of dopaminergic manipulations on risky choice. *Psychopharmacology*, *211*(2), 209–221. https://doi.org/10.1007/s00213-010-1883-y

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

St Onge, J. R., & Floresco, S. B. (2010). Prefrontal cortical contribution to risk-based decision making. *Cerebral Cortex*, *20*(8), 1816–1828. https://doi.org/10.1093/cercor/bhp250

Steele, C. C., Gwinner, M., Smith, T., Young, M. E., & Kirkpatrick, K. (2019). Experience matters: The effects of hypothetical versus experiential delays and magnitudes on impulsive choice in delay discounting tasks. *Brain Sciences*, *9*(12). https://doi.org/10.3390/brainsci9120379

Stein, J. S., MacKillop, J., McClure, S. M., & Bickel, W. K. (2022). Unsparing self-critique strengthens the field, but bailey et al. Overstate the 'problems with delay discounting'. *Psychological Medicine*, 1–2. https://doi.org/10.1017/S0033291721005286

Stephens, D. W., Kerr, B., & Fernández-Juricic, E. (2004). Impulsiveness without discounting: The ecological rationality hypothesis. *Proceedings. Biological Sciences / The Royal Society*, *271*(1556), 2459–2465. https://doi.org/10.1098/rspb.2004.2871

Steverson, K., Chung, H.-K., Zimmermann, J., Louie, K., & Glimcher, P. (2019). Sensitivity of reaction time to the magnitude of rewards reveals the cost-structure of time. *Scientific Reports*, *9*(1), 20053. https://doi.org/10.1038/s41598-019-56392-0

Stopper, C. M., Green, E. B., & Floresco, S. B. (2014). Selective involvement by the medial orbitofrontal cortex in biasing risky, but not impulsive, choice. *Cerebral Cortex*, *24*(1), 154–162. https://doi.org/10.1093/cercor/bhs297

Story, G. W., Kurth-Nelson, Z., Moutoussis, M., Iigaya, K., Will, G.-J., Hauser, T. U., Blain, B., Vlaev, I., & Dolan, R. J. (2023). Discounting future reward in an uncertain world. In J. Decision (Ed.). https://doi.org/10.1037/dec0000219

Stribiţcaia, E., Evans, C. E. L., Gibbons, C., Blundell, J., & Sarkar, A. (2020). Food texture influences on satiety: Systematic review and meta-analysis. *Scientific Reports*, *10*(1), 12929. https://doi.org/10.1038/s41598-020-69504-y

Strickland, J. C., & Johnson, M. W. (2021). Rejecting impulsivity as a psychological construct: A theoretical, empirical, and sociocultural argument. *Psychological Review*, *128*(2), 336–361. https://doi.org/10.1037/rev0000263

Suzuki, S., Lawlor, V. M., Cooper, J. A., Arulpragasam, A. R., & Treadway, M. T. (2021). Distinct regions of the striatum underlying effort, movement initiation and effort discounting. *Nature Human Behaviour*, *5*(3), 378–388. https://doi.org/10.1038/s41562-020-00972-y

Tanno, T., Maguire, D. R., Henson, C., & France, C. P. (2014). Effects of amphetamine and methylphenidate on delay discounting in rats: Interactions with order of delay presentation. *Psychopharmacology*, *231*(1), 85–95. https://doi.org/10.1007/s00213-013-3209-3

Thompson, S. M., Berkowitz, L. E., & Clark, B. J. (2018). Behavioral and neural subsystems of rodent exploration. *Learning and Motivation*, *61*, 3–15. https://doi.org/10.1016/j.lmot.2017.03.009

Vandaele, Y., & Ahmed, S. H. (2020). Habit, choice, and addiction. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*. https://doi.org/10.1038/s41386-020-00899-y

Vanderveldt, A., Oliveira, L., & Green, L. (2016). Delay discounting: Pigeon, rat, human–does it matter? *Journal of Experimental Psychology. Animal Learning and Cognition*, *42*(2), 141–162. https://doi.org/10.1037/xan0000097

Varghese, M., Keshav, N., Jacot-Descombes, S., Warda, T., Wicinski, B., Dickstein, D. L., Harony-Nicolas, H., De Rubeis, S., Drapeau, E., Buxbaum, J. D., & Hof, P. R. (2017). Autism spectrum disorder: Neuropathology and animal models. *Acta Neuropathologica*, *134*(4), 537–566. https://doi.org/10.1007/s00401-017-1736-4

Walsh, K. M., & Bracken, M. B. (2011). Copy number variation in the dosage-sensitive 16p11.2 interval accounts for only a small proportion of autism incidence: A systematic review and meta-analysis. *Genetics in Medicine: Official*

Journal of the American College of Medical Genetics, *13*(5), 377–384. https://doi.org/10.1097/GIM.0b013e3182076c0c

Wang, W., Rein, B., Zhang, F., Tan, T., Zhong, P., Qin, L., & Yan, Z. (2018). Chemogenetic activation of prefrontal cortex rescues synaptic and behavioral deficits in a mouse model of 16p11.2 deletion syndrome. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *38*(26), 5939–5948. https://doi.org/10.1523/JNEUROSCI.0149-18.2018

Warnell, K. R., Maniscalco, S., Baker, S., Yi, R., & Redcay, E. (2019). Social and delay discounting in autism spectrum disorder. *Autism Research: Official Journal of the International Society for Autism Research*. https://doi.org/10.1002/aur.2085

Weafer, J., & Wit, H. de. (2014). Sex differences in impulsive action and impulsive choice. *Addictive Behaviors*, *39*(11), 1573–1579. https://doi.org/10.1016/j.addbeh.2013.10.033

Weiss, E. O., Kruppa, J. A., Fink, G. R., Herpertz-Dahlmann, B., Konrad, K., & Schulte-Rüther, M. (2020). Developmental differences in probabilistic reversal learning: A computational modeling approach. *Frontiers in Neuroscience*, *14*, 536596. https://doi.org/10.3389/fnins.2020.536596

Westbrook, A., Kester, D., & Braver, T. S. (2013). What is the subjective cost of cognitive effort? Load, trait, and aging effects revealed by economic preference. *PloS One*, *8*(7), e68210. https://doi.org/10.1371/journal.pone.0068210

Yang, M., Lewis, F. C., Sarvi, M. S., Foley, G. M., & Crawley, J. N. (2015). 16p11.2 deletion mice display cognitive deficits in touchscreen learning and novelty recognition tasks. *Learning & Memory*, *22*(12), 622–632. https://doi.org/10.1101/lm.039602.115

Yechiam, E., Arshavsky, O., Shamay-Tsoory, S. G., Yaniv, S., & Aharon, J. (2010). Adapted to explore: Reinforcement learning in autistic spectrum conditions. *Brain and Cognition*, *72*(2), 317–324. https://doi.org/10.1016/j.bandc.2009.10.005

Yoon, J. H., De La Garza, R., 2nd, Newton, T. F., Suchting, R., Weaver, M. T., Brown, G. S., Omar, Y., & Haliwa, I. (2017). A COMPARISON OF MAZUR'S k AND AREA UNDER THE CURVE FOR DESCRIBING STEEP DISCOUNTERS. *The Psychological Record*, *67*(3), 355–363. https://doi.org/10.1007/s40732-017-0220-9

Zeif, D., Yakobi, O., & Yechiam, E. (2023). Choice behavior in autistic adults: What drives the extreme switching phenomenon? *PloS One*, *18*(3), e0282296. https://doi.org/10.1371/journal.pone.0282296