ON THE SUBJECTIVE VALUATION OF EFFORT: COMPUTATIONAL INSIGHTS INTO MOTOR LEARNING AND CONTROL

by

CHADWICK MICHAEL HEALY

B.S., United States Naval Academy, 2009

M.S., University of Colorado, 2010

M.S., University of Colorado, 2016

M.S., University of Colorado, 2021

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy

Biomedical Engineering Program

2024

Committee Members:

Alaa A. Ahmed

Torin Clark

Roger Enoka

Andrew Tan

Cara Welker

Healy, Chadwick Michael (Ph.D., Biomedical Engineering)

ON THE SUBJECTIVE VALUATION OF EFFORT: COMPUTATIONAL INSIGHTS INTO MOTOR LEARNING AND CONTROL

Thesis directed by Alaa A. Ahmed, Ph.D.

ABSTRACT

Every movement is a reflection of a person's subjective values, manifested as a willingness to expend energy to achieve a desired outcome. Economically speaking, a person should maximize the utility of their movements by acquiring as much reward as early as possible with the least amount of effort. However, effort alone paints an incomplete picture of the costs associated with movement. There are other factors at play, such as error, time, and risk, that can modulate how much and when effort is expended. Here, I explore how the subjective valuation of effort relative to other costs can influence movement decisions.

The first study investigates the effect of age on preference for error over effort during motor learning. Older adults performed an arm-reaching adaptation task with greater error than younger adults, which is traditionally interpreted as a deficit in learning. However, I show that a difference in subjective cost of effort can manifest these same larger errors while still learning as much as younger adults.

The second study investigates the interaction of time and effort on preference. The traditional view of effort is that we should be agnostic to when periods of high effort are performed—only total energy matters so long as the goal is accomplished. Using an isometric arm-pushing task, I show that subjects preferred earlier high physical effort to later and that the strength of their preference is also reflected in how they indicated their preference.

The third study pilot tests how a virtual reality system could be used to investigate the effect of reward on walking speed. The fourth and final study builds and extends this pilot study to investigate how immediate reward, a history of reward, and baseline effort costs influence walking speed. I find that subjects walked faster not only for higher immediate reward but also when they experienced a history of higher reward. In addition, subjects walked slower when effort costs were higher.

Collectively, these studies deepen our understanding of how subjective effort cost interacts with error, time, and reward, and how these interactions manifest in our movement decisions.

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the support and encouragement from so many of you. I would like to express my sincerest gratitude to you all. Thank you:

To my wife, Corinne—your unwavering love and support always bring a smile to my face. Thank you for embracing our move to Boulder and for standing by me every step of the way throughout this journey.

To my parents, Greg and Kristi, who instilled a sense of curiosity in me, and encouraged me to do challenging things. Thank you for believing in me.

To my sister, Shanay, who always took the time to listen and give the biggest hugs. Your kindness and wisdom continue to help me navigate the toughest times.

To all my friends in Boulder and afar, who provided a listening ear, an escape, and a good laugh through the ups and downs of graduate school.

To my labmates, past and present, in the Neuromechanics Lab. Megan O'Brien, Erik Summerside, Alison Pienciak-Siewert, Gary Bruening, Shruthi Sukumar, Robbie Courter, Sam Rubaker, Colin Korbisch, Rachel Marbaker, and Caelan Thom. Thank you for enduring my long-winded weekly updates, for your critiques, and for your camaraderie.

To my many mentors during my winding path through academia—Admiral Craig Steidle, Dr. Dick Fahey, Dr. David Klaus, Astronaut Joe Tanner, Astronaut Jim Voss, Dr. Rodger Kram, Dr. Alena Grabowski, Rachel Forman, Mike Westenhaver, and many others. Thank you for exposing me to the fields of science and engineering and shaping me into a better researcher.

To my committee members, Dr. Torin Clark, Dr. Roger Enoka, Dr. Andrew Tan, and Dr. Cara Welker. Thank you for sharing your expertise and critical feedback on my thesis.

Finally, and most importantly, thank you, Dr. Alaa Ahmed. You are not only an incredible scientist but also a thoughtful and compassionate leader. Thank you for taking a chance on me, letting me wander, and welcoming me back for a second shot at finishing my PhD. I have learned so much from you.

CONTENTS

TABLES

FIGURES

CHAPTER I

THESIS INTRODUCTION

1.1 Introduction

Think of what actions you have taken today that brought you to the point of reading this thesis. Perhaps it included walking to your computer, sitting in a chair, typing in your memorized username and password, delicately manipulating a cursor on your computer screen, then clicking on this document. You had to breathe, walk, sit, reach, and gaze—all of which you have learned to perform accurately and efficiently. These coordinated movements were performed with relative ease despite their incredibly complexity. The ability to move is a vital skill for most animals, yet the mechanisms for how we decide to move and learn to move remain unclear. The field of neuromechanics is attempting to synthesize biomechanics with neural control (Burdet et al., 2013), and Moravec's paradox elegantly summarizes the difficulty of investigating this field: the oldest human skills such as motor control—which appear to be effortless or subconscious—are actually the most difficult to characterize and recreate. "It is comparatively easy to make computers exhibit adult level performance on intelligence tests or playing checkers, and difficult or impossible to give them the skills of a one-year-old when it comes to perception and mobility." (Moravec, 1988). It appears the skills that we have evolved to perform over thousands of years are significantly harder to reverse engineer. More than 35 years after Moravec's proclamation, significant discoveries have been made in the field, but we still have

not designed robots that can move as elegantly as a human, signaling that we do not yet understand how the brain decides to move the body.

If you are anything like the author, a strong cup of coffee is a necessary companion when reading an academic article. Let us reflect on your hypothetical reach for that warm cup of joe. When reaching for your mug, how do you decide to move? Your goal could have been achieved by moving in an infinite number of ways, yet you chose one that is not so different from the rest of the human population. The observation that individuals tend to move the same led to the hypothesis that humans are utilizing some common function for selecting and controlling movements. Though these functions are not directly measurable, by quantifying observable behaviors, we are able to back out an objective function and control methodologies for human movements. By determining the underlying functions that we use to control our movements, adapt our movements, and decide how to move, we may better understand the neural mechanisms underlying those processes, leading to a theoretical foundation for treating movement disorders, improving rehabilitation, and designing more human-like robots.

1.2 Introduction to sensorimotor control

The motor system must solve the problem of what commands to send (i.e., neural activation of muscles) to achieve a desired outcome (e.g., moving the body to a new position). However, this problem is overdetermined, unconstrained, and nonlinear. There are many more degrees of freedom and muscles than are necessary to achieve most movements. For instance, only considering the end state of grasping

the coffee mug's handle, there are an infinite number of joint angles that achieve this position. Further, assuming a single limb position is desired, the trajectory selected to reach that state has infinitely many solutions in both space and time (i.e., the path and how long you take to reach the mug). In addition, command inputs do not directly control limb position. Changes in position are accomplished indirectly through neural activations of muscles that generate forces that result in torques across joints that ultimately result in acceleration of the limbs. The same command inputs are state-dependent, where the same input can yield different accelerations depending on the state of the limbs and muscles (i.e., muscle force output is dependent on its length, velocity, and other factors such as fatigue) (Hill, 1953). Thus, there is not a unique set of controls to achieve that change in position; the same net acceleration can be achieved through varying levels of co-contraction of muscles. To make the problem more complex, muscle commands are highly noisy, and sensory information is delayed (Faisal et al., 2008; Jones et al., 2002). Given the complexity of the problem that the motor system must solve, one might expect humans to choose very different solutions, yet humans still appear to move similarly across individuals. In arm movements, humans reach similarly across different directions, distances, and orientations, following roughly straight paths with bell-shaped velocity curves (Morasso, 1981). This suggests that humans are all optimizing for some common variables, and we see patterns in movement that reflect this.

Modeling movement trajectories begins with developing a dynamical model. Most simply, in cases such as arm reaches, the end effector can be modeled as a point

mass. Increasing biological realism, arms and legs can be modeled as multi-link rigid bodies with torques as activators at each joint. And even more realistically, models can incorporate muscles across each joint that generate forces and exhibit viscoelastic properties of the muscles and tendons themselves. The appropriate level of detail is a matter of judgment and is based on movement modality and the specific research question. Trade-offs between biological realism, model interpretability, and generalizability are typically considered when choosing a dynamical model.

Types of control in the field of motor control are generally split into two categories: feedforward and feedback. In a feedforward control policy, coordination is achieved through prediction of motor commands to state outcome, where action is taken before sensory information can be integrated. It is open loop, and errors are not monitored. Importantly, the existence of a forward model and feedforward control is an essential compensatory mechanism for the existence of sensory delays. Open loop feedforward control is used in very fast movements such as saccades. Conversely, feedback control incorporates sensory feedback and makes corrections or updates based on the difference between the predicted state and the sensed state.

Sensory feedback is essential for feedback control. When a movement begins, humans are reliant on visual and proprioceptive sensory information to sense the state of their bodies and compare it with the predicted state. However, sensory feedback is notoriously noisy and delayed. Investigations into how we process sensory feedback suggest that we do so optimally using Bayesian integration, where sensors are combined based on a prior model of the sensor's variability (Körding & Wolpert,

2004). In feedback control, the estimated state derived from sensory integration is compared to a predicted state based on a forward model.

Revisiting the hypothesis that because humans decide to move the same, there is likely a common set of factors being optimized, this idea can be applied to motor control. Early models investigated an optimal control policy that maximized the smoothness of our movements. Flash and Hogan proposed a control policy that minimized jerk that accurately produced the bell-shaped velocity profile observed in arm reaches and eye movements (Flash & Hogan, 1985). A more dynamical version of this model minimizes the square sum of the derivatives of torques about a joint (Uno et al., 1989), which produces a similar bell-shaped velocity profile. However, due to the nonlinearities in joint angle mappings, the two models produce slightly different results. While both were the first models to broadly describe human movement using optimization, they both fell short in that they did not incorporate any kinematics. Flanagan and Rao designed a point-to-point reaching task to investigate reliance on kinematic feedback (Flanagan & Rao, 1995). In one case, cursor position was determined by shoulder and elbow angle. Subjects adopted a more costly curved hand movement to produce straighter trajectories in the shoulder-elbow joint defined space, showing that the nervous system does not minimize dynamic criterion alone and considers visual, kinematic criteria. More recently, Harris and Wolpert proposed a sensorimotor control system that minimizes endpoint variance. If obtaining the goal requires an accurate endpoint (e.g., grabbing your mug of coffee), then a natural cost that is minimized should be endpoint error. By incorporating

signal-dependent noise and the penalty on endpoint variance, saccades and arm reach trajectories exhibit the characteristic skew of the bell-shaped velocity profiles observed in the data. Motor control and motor learning appear to be optimizing for a combination of error and effort minimization, leading to more recent formulations of control policies that penalize motor commands.

Thus far, the control policies discussed have only been feedforward. For slower movement types (i.e., not saccades), sensory feedback can be incorporated for online correction. The natural evolution of these optimizing control policies is to develop an optimal feedback control policy (Todorov & Jordan, 2002). In this framework, the optimal trajectory and motor commands are recalculated at each step based on the current sensed state. The control policy calculates the next step of control inputs based on the newly planned desired trajectory to reach the end goal at a minimized cost. By incorporating feedback, control inputs try to cancel out the effects of noise and external perturbations in order to accomplish the end goal with minimal effort. This framework elegantly describes several movement modalities and solves the motor redundancy by utilizing a minimum intervention policy—control inputs are only made for task-relevant goals. Optimal control theory is a powerful framework in that can explain a wide range of behaviors with a single structure, and importantly, these models are interpretable, where relative costs on specific states can be backed out into an objective function for that task. In Chapter III of this thesis, I employ an optimal control model for this very reason.

Optimal control theory has garnered some criticism, as it is difficult to disprove. If an optimal control model fails to match data, does this disprove optimal control as a model, or is it that improper dynamical or cost function models were used? Conversely, with enough complexity, optimal control theory could describe almost any behavior; thus, researchers must be careful in tuning cost functions to reflect biological realism.

Most optimal feedback control models utilize a finite horizon where the movement duration is a predetermined value. This formulation cannot describe continuous motor behaviors like tracking or drawing, nor can it exhibit the non-smooth trajectories of slow reaches observed in infants (von Hofsten, 1991) and adults forced to reach slowly (>1 second). These velocity profiles suggest movements are comprised of a series of sub-movements, where the overall velocity profile appears to be a summation of bell-shaped velocity profiles updated at roughly 200 ms intervals (Milner, 1992).

Recent modeling approaches attempt to address these issues by developing a model with an infinitely receding time horizon, where the task goal is periodically updated (Guigon, 2021). This model can describe continuous movements like drawing and can describe the non-smooth velocity profiles of slow reaches. This model is one step closer to a more unified description of motor control; however, there are still unanswered questions that require further investigation: are there internal representations of control policy in the brain, and how does the brain represent movement and control inputs?

1.3 Introduction to sensorimotor learning

While motor control can be viewed as an optimization problem as summarized above, so can sensorimotor learning. When learning a new task, subjects initially adopt different strategies for solving the task, then converge to almost identical types of behavior (Nagengast et al., 2009), suggesting some common feedback mechanisms that drive motor behavior towards an optimal solution. Similar to optimal feedback control, motor learning requires prediction and sensory feedback of the movement. By investigating how we learn and what drives learning, we gain insight into what factors are at play in our movements' objective function.

Learning is not only required to obtain new skills or behaviors, but also to adapt to changing environments and body dynamics, to improve on previously developed behaviors (Tseng et al., 2007). Sensorimotor learning can generally be divided into two categories: motor adaptation and motor skill acquisition. The former is a process in which the motor system adapts to perturbations, taking place over short periods of time, and does not involve learning a new control policy. The latter involves learning a new control policy over a longer time period (Caramiaux et al., 2020).

Numerous studies in motor learning focus on the adaptation paradigm, where a well-learned task, such as walking or reaching, is perturbed or altered, causing errors. One such method of investigation adaptation is through visuomotor rotation, where visual feedback is rotated or altered. Initially, this was done with prism goggles (Helmholtz, 1962) but more recently, it has been investigated in arm reaches with

hand position displayed virtually and digitally manipulated (Krakauer et al., 2000; Pellizzer & Georgopoulos, 1993). Another common approach to investigate motor adaptation is through altering the dynamics in which a task is performed. In arm reaches, this is commonly investigated using a viscous "curl field," where subjects' hands are perturbed by an outside force perpendicular and proportional to their hand's velocity (Shadmehr & Mussa-Ivaldi, 1994) or reaching in a Coriolis field by rotating subjects (Lackner & Dizio, 1994). Altering dynamics has also been investigated in the walking domain via a split-belt treadmill (Dietz et al., 1994). In all these adaptation paradigms, subjects initially experience large errors (typically measured by kinematic deviations from baseline performance), then systematically and exponentially reduce these errors over several repetitions. When returned to the baseline, normal state, subjects exhibit the opposite error, then gradually de-adapt back to baseline performance. It is clear from these paradigms that error is a driving signal for learning and that the learning process is optimizing for error reduction. What constitutes an error? What is being learned, and how do you measure how much has been learned?

The amount in which a person has learned is not a directly measurable state—it can only be inferred from other observable behaviors. An internal model can be thought of as a hidden representation or mapping in the brain. An internal model that predicts sensory consequences of motor commands is known as an internal forward model. This type of model incorporates a dynamical model of the limbs, muscles, and environment (e.g., acceleration/gravity). An internal inverse model does

the opposite: it maps a desired state to desired motor commands (Kawato, 1999). The existence of an inverse internal model has been superseded by the concept of a control policy described earlier. Nevertheless, both the control policy and forward model are latent states that are not directly measurable. Motor learning can be modeled as a process of updating the internal model using sensory feedback (Shadmehr, Smith, et al., 2010).

Collectively, sensitivity to error and formulation of an internal forward model give rise to a descriptive model of the learning process: the state space model (Donchin et al., 2003; Thoroughman & Shadmehr, 2000). This model is a discrete-time, linear dynamical system that predicts a subject will update their motor commands on the next trial proportional to the error experienced on the current trial. This formulation can explain the observed exponential decay of errors, and by fitting a parameter for sensitivity to error, it allows for the comparison of learning rates across test conditions and populations.

The state space model is also easily extendable to describe observed phenomena in motor learning. First, we look at retention of adaptation, where there is a tendency to return to baseline behavior in the absence of error feedback. By including a "forgetting factor," Galea and colleagues were able to model motor commands that gradually decayed back to baseline (Galea et al., 2011). Secondly, Shadmehr and colleagues explored the concept of spontaneous recovery, first observed in saccades (Kojima et al., 2004), where the rate of returning to baseline was accelerated compared to adapting to a new environment (Smith et al., 2006). By

extending the state space model to include both a fast and a slow rate of error sensitivity, spontaneous recovery behavior can be accurately modeled. Ongoing research has yet to confirm that these two separate processes exist in the brain, but researchers hypothesize that the fast process could be sensitive to kinematic error reduction, while the slow process updates predictive feedforward commands. While these models have been criticized as just a descriptive model of behavior, importantly, they show that prediction error drives learning.

Another natural extension of the state space model is through the incorporation of uncertainty and prior beliefs. By doing so, the state space model becomes a Kalman filter, where learning is a process of updating a prior belief using prediction error (Korenberg & Ghahramani, 2002). This model predicts that learning rate is a consequence of the variability of the task and environment, not just an arbitrarily fit parameter. Interestingly, this model predicts that as the uncertainty of feedback increases, the learning rate should become slower, and as uncertainty in the error or perturbation increases, the learning rate should increase. This model can explain the process of error attribution, where learning only occurs for task-relevant errors. Investigations have successfully shown this in crediting transient versus persistent error (Wei & Körding, 2009), body versus world error (Berniker & Kording, 2008), and motor execution versus sensory miscalibration error (Haith et al., 2008.).

State space and Bayesian models provide powerful descriptions of how error drives adaptation; however, there are other factors that impact learning. In addition to error, studies have shown that effort is reduced through the learning process

(Huang et al., 2012; Pienciak-Siewert et al., 2016; Selinger et al., 2015), suggesting a simultaneous optimization of both these factors (Emken et al., 2007). Studies have also shown learning rate is increased by increasing rewards (Nikooyan & Ahmed, 2015) and increasing the consequence of errors (Trent & Ahmed, 2013). These studies, in conglomeration, show that there is more at play in motor learning than just error reduction and suggest that future development of models that incorporate these factors is necessary to accurately capture the motor learning process.

In Chapter III, I investigate how subjective costs on error and effort could influence performance in a motor learning task.

1.4 Introduction to sensorimotor decision-making

Drawing from the field of neuroeconomics, movements can be framed as a decision that weighs factors such as reward, effort, time, and risk. These factors combine to form the overall *utility* of a movement, where a person pays the cost of effort in order to achieve some reward. We find another optimization problem: do humans and animals maximize the utility of their movements? If so, how? To investigate these questions, it is natural to split the concept of movement decisions into two parts: 1) what is the preferred goal of the movement? and 2) how should that movement be performed? Critically, these decisions are not independent of one another: the same factors that determine the goal also impact the vigor in which the movement is performed.

Let us first define the concept of reward and consider this variable independently. A reward is an incentive. In the wild, a reward can be any number of

things: food, reproduction, survival. In the laboratory, for animal studies, a reward is typically food, water, or juice. For humans in the laboratory, a reward is typically monetary but could be as simple as fictitious points or a rewarding sound. Well-designed experiments have rewards that are discernable and quantifiable. Logically, when given a choice between a large reward and a small reward, humans choose the larger reward (Levy & Glimcher, 2011). Yet, humans do not objectively value rewards. In the field of economics, the marginal utility diminishes as reward increases. For example, the difference between receiving $10 and $15 is subjectively larger than the difference between receiving $110 and $115, even though objectively, they are equal. This is typically modeled as a power function where the objective reward is raised to some power value between 0 and 1. O'Brien and Ahmed confirmed the subjective value of reward in movement decisions is similar to values commonly reported in economic tasks (O'Brien & Ahmed, 2014).

Now, let us consider the concept of cost. This is the anti-reward. First, let's explore the costs of time and uncertainty of reward. Both animal and human studies have shown that earlier rewards are preferred to later rewards, suggesting there is a cost on time, formalized as a temporal discounting factor on reward. In both humans and animals, there are varying degrees of impulsivity. Some studies have seen subjects willing to accept smaller, earlier rewards over larger, later rewards (Green et al., 2004). Temporal discounting is typically modeled as either exponential (Hull, 1943) or hyperbolic (Mazur et al., 1987). The latter has become the more widely accepted model, as it can describe the reversal of preferences regarding time, whereas

the former cannot. Similar to the diminishing marginal utility, we see diminishing effects of time as it increases. Consider the choice of receiving $5 in a week or $6 in two weeks, which would you choose? Let's assume you would take the $5 in one week. Now consider the choice of receiving $5 in a year or $6 in a year and one week. The choice is the same net difference in time, but a person may be more inclined to wait the extra week for the added dollar in the second scenario.

Next, let's consider effort. Like time, effort is a cost that discounts reward. Let us constrain the concept of effort to purely physical for the purposes of movement decisions, though cognitive effort is a cost worth considering. Though the neural correlates for effort are still undetermined, numerous studies attest that effort can be quantified by metabolic cost (Shadmehr et al., 2016). Metabolic cost is a measure of the body's expenditure of energy through the conversion of chemical energy into mechanical work. Typically, this is measured using expired gas analysis, which compares inhaled oxygen to exhaled carbon dioxide. Alternative measures such as the doubly labeled water technique can measure average energy expenditure over longer time periods (Speakman, 1998), or control inputs such as muscle forces act as correlates to effort. Despite the uncertainty of how the brain represents physical effort, there is clear evidence that we are sensitive to and aware of the effort of our movements. In locomotion studies, it appears both humans and animals choose to walk at a speed that minimizes energy consumption (Hoyt & Taylor, 1981; Ralston, 1958). There exists a metabolically optimal reaching speed, which humans roughly

choose to reach (Shadmehr et al., 2016). These findings show we are optimizing for and minimizing our movement effort.

Reward, time, and effort all combine into an overall optimization problem to determine the movement speed that maximizes the rate of utility. The concept of optimizing a rate of reward, or utility, in behavior, is not new. With roots in ecology, optimal foraging theory explores how animals move between patches of food in a way that maximizes the rate of energetic intake minus energetic expenditure (Charnov, 1976). Recent research extends this theory to human movements, where similarly, a movement should aim to maximize the net utility over time, represented as the sum of the rewards minus energetic costs, hyperbolically discounted by time. If reward diminishes in subjective value as time passes, logically, that reward should be captured as early as possible. However, in order to reduce the cost on time, additional effort cost must be expended. The choice of how quickly to move can then be conceptualized as optimizing for a balance between the reward value, the cost of time, and the effort required to acquire that reward. The optimization manifests in a decision of movement speed, duration, peak velocity, peak force, and/or reaction time (Shadmehr & Ahmed, 2020).

This formulation makes some interesting predictions that have been observed in discrete human and animal movements. First, we should move faster to higher rewards. Indeed, movements toward more rewarding targets are performed with more vigor. Both humans and other animals saccade faster to targets when that target is more rewarding (Takikawa et al., 2002; Xu-Wilson et al., 2009). The vigor of

arm reaches also increases with increased reward in the form of points and audiovisual stimuli (Summerside, Shadmehr, et al., 2018). However, this has not been shown in human walking (a clue for what is ahead in Chapters V and VI).

Next, with a constant reward, we should choose the movement that requires less effort and move more quickly (Shadmehr et al., 2019). Indeed, when people are given the choice of two reach directions they choose the direction that has lower effective mass (Cos et al., 2011) and also perform that reach with higher velocity (Gordon et al., 1994). Conversely, higher effort costs would predict decreases in vigor. Investigations have shown that added effort in longer arm reaches results in slower reaction times (Rosenbaum, 1980) and longer movement durations (Bruening et al., 2024; Shadmehr et al., 2016).

Interestingly, movement vigor also depends on the background state of the environment in terms of average expected utility. Niv and colleagues were perhaps the first to explore ties to decision-making and movement vigor. In their model, they incorporate an opportunity cost on time, where time spent harvesting diminishing rewards is now at odds with time spent acquiring potential future rewards (Niv et al., 2007). A history of higher rewards should lead to faster movements because the opportunity cost of time is greater. Conversely, a history of low rewards should lead to slower movements. A history of high effort, lowering the average utility and opportunity cost, predicts a slower speed. Conversely, a history of lower effort would predict a faster speed. Recently, this sensitivity to history of reward and history of

effort has been observed in saccades (Yoon et al., 2018) and an arm-reaching foraging task (Sukumar et al., 2024).

With effort discounted hyperbolically in the proposed utility equation, this also predicts some temporal sensitivity, where, as opposed to wanting reward as early as possible, one would procrastinate effort as late as possible. Some investigations support this hypothesis, using a model of utility that subtracts physical effort costs from reward, with the net value discounted by time (Bautista et al., 2001; Bruening et al., 2024; Rigoux & Guigon, 2012; Shadmehr et al., 2016; Sukumar et al., 2024; Summerside, Kram, et al., 2018). However, other recent investigations contradict this. In Rosenbaum et al. 2014, subjects preferred to complete earlier subgoals at the expense of extra physical effort (Rosenbaum et al., 2014). This result could be formulated as a freeing up cognitive resources required by the subgoal, in the case of Rosenbaum et al., picking up a bucket. Another possible explanation for choosing to pre-crastinate physical effort is to view it similarly to pain, where a better ending is preferred (Kahneman et al., 1993). Taken together, how time and effort interact is not well understood. Chapter IV explores the idea of temporal sensitivity of physical effort and explicitly gauges preference for when periods of high effort occur.

1.5 Summary

In summary, the processes of motor learning, motor control, and movement decision-making can all be thought of as optimization problems. Traditional views of movement costs are largely assumed to be objective and purely energetic, but findings continue to show that there is more to our movement decisions than Joules alone. The

brain weighs movement costs such as effort, error, and time, but it is unclear how these relate to one another to determine an optimal movement path. This thesis takes strides in exploring how the cost of effort is subjectively valued relative to other movement costs like time and error and how the subjective valuation of effort manifests in differences in motor learning and control.

CHAPTER II

THESIS OBJECTIVES

2.1 Motivation

Understanding how and why we move the way we do is essential to improving interventions for aging adults and patient populations. Energy expenditure plays a significant role in determining how we move, but it is only one of many costs associated with a movement. To date, how other costs interplay with effort and if their relationship changes with age or disease is not well characterized. In the following studies, we leverage mathematical frameworks to describe movement costs. Using this approach, we can better understand the relationship between movement costs and how they influence movement preferences, motor control, and motor learning.

2.2 Specific Aims

- Aim 1: Determine the effect of age on preference for reducing movement error over reducing effort expenditure during learning of a novel motor task.
- Aim 2: Determine the interaction of time and effort on movement preference.
- Aim 3: Develop a virtual reality system to probe the influence of reward on preferred walking speed.
- Aim 4: Determine the influence of immediate reward, history of reward, and baseline effort on preferred walking speed.

2.3 Outline

In the following chapters of this thesis, I present and summarize evidence for each of the specific aims above.

- Chapter III employs a unique computational modeling approach to show that caring more about effort relative to error can describe kinematic differences normally associated with a lack of adaptation. Using this technique, I show that older adults could have learned as well as younger adults in a classic arm-reaching adaptation task. This study has been published (Healy et al., 2023).
- Chapter IV shows that subjects care about when periods of high effort are performed within a movement. I use a unique, isometric arm-pushing task to tease apart the temporal sensitivity of effort costs and task-related costs and show that subjects tend to prefer to perform high effort earlier rather than later. This study has been published (Healy & Ahmed, 2024).
- Chapter V lays the foundation for a novel experimental setup to probe the effect of reward on preferred walking speed. This chapter details how I integrated a virtual reality system with a self-paced treadmill, presents analyzed pilot data, and summarizes necessary changes and lessons learned ahead of formal data collection.
- Chapter VI leverages the aforementioned virtual reality and treadmill experimental setup to investigate the effect of immediate reward, history of reward, and baseline effort on preferred walking speed in healthy young

adults. These results further expand recent findings of movement motivation and invigoration into the realm of human walking.

Following these chapters, I summarize my findings and their significance to the field in Chapter VII.

2.4 Significance

The following investigations build on our understanding of the subjective costs of movements, specifically how subjective effort ties into movement decisions, how it relates to other movement costs, and how it might vary across populations. These series of studies show that energetic cost is not the only factor that determines movement speed. With a more complete understanding of the cost function for movement control and how it might translate across different movement modalities, this work can aid in the understanding of how motor functions in older adults and patient populations change and help develop better diagnostic approaches and interventions to assist in the rehabilitation and maintenance of motor function.

CHAPTER III

LEARNING VERSUS MINDING: HOW SUBJECTIVE COSTS CAN MASK MOTOR LEARNING[1]

3.1 Abstract

When learning new movements some people make larger kinematic errors than others, interpreted as a reduction in motor-learning ability. Consider a learning task where error-cancelling strategies incur higher effort costs, specifically where subjects reach to targets in a force field. Concluding that those with greater error have learned less has a critical assumption: everyone uses the same error-canceling strategy. Alternatively, it could be that those with greater error may be choosing to sacrifice error reduction in favor of a lower effort movement. Here, we test this hypothesis in a dataset that includes both younger and older adults, where older adults exhibited greater kinematic errors. Utilizing the framework of optimal control theory, we infer subjective costs (i.e., strategies) and internal model accuracy (i.e., proportion of the novel dynamics learned) by fitting a model to each population's trajectory data. Our results demonstrate trajectories are defined by a combination of the amount learned and strategic differences represented by relative cost weights. Based on the model fits, younger adults could have learned between 65-90% of the novel dynamics. Critically, older adults could have learned between 60-85%. Each model fit produces trajectories that match the experimentally observed data, where a lower proportion learned in the model is compensated for by increasing costs on

[1] *The work from this chapter has been published in PLOS One (Healy et al., 2023).*

kinematic errors relative to effort. This suggests older and younger adults could be learning to the same extent, but older adults have a higher relative cost on effort compared to younger adults. These results call into question the proposition that older adults learn less than younger adults and provide a potential explanation for the equivocal findings in the literature. Importantly, our findings suggest that the metrics commonly used to probe motor learning paint an incomplete picture, and that to accurately quantify the learning process the subjective costs of movements should be considered.

3.2 Introduction

When people are introduced to a novel environment, they initially experience large errors relative to their intended performance. Through repeated practice, they systematically reduce these errors, learning how best to perform the task (Ahmed & Wolpert, 2009; Finley et al., 2013; Flanagan & Wing, 1997; Huang et al., 2012; Krakauer et al., 1999; Morton & Bastian, 2006; Nikooyan & Ahmed, 2015; Pienciak-Siewert et al., 2016; Reisman et al., 2005; Shadmehr & Mussa-Ivaldi, 1994). Because the amount a person has learned is not directly assessable, we are reliant on indirect measures such as error that reflect the latent state of how much a person has learned. However, error-reduction in and of itself does not necessarily equate to learning a model of the task. First, numerous studies have shown that a model-free process, such as impedance control, can cancel error without the necessity of learning an accurate model of the task (Crevecoeur et al., 2019; A. M. Haith & Krakauer, 2013). Second, even when assuming similar model-free approaches, when comparing across

subjects or populations, variation in strategies must be considered. In this study, we focus on the latter point. We ask whether greater error may not be a result of learning less, but rather, a strategic trade-off between movement errors and other similarly important factors.

Movement decisions and learning are not solely driven by error reduction. People optimize for other factors such as reward, effort, time, and risk. When people move faster, they move with greater error (Fitts, 1954) and tend to have trajectories which minimize the variance of the end or target position (Harris & Wolpert, 1998). Yet, people still choose to move faster in the presence of greater reward in both arm reaches (Summerside, Shadmehr, et al., 2018) and saccades (Xu-Wilson et al., 2009). Reward is discounted by time, thus people and animals are willing to expend more effort to achieve an earlier reward (Manohar et al., 2015; Rigoux & Guigon, 2012; Shadmehr et al., 2016). In addition to error, studies have shown that effort is reduced through the learning process (Finley et al., 2013; Huang et al., 2012), suggesting a simultaneous optimization of both these factors (Emken et al., 2007). Studies have also shown learning rate is increased by increasing rewards (Nikooyan & Ahmed, 2015), increasing consequence of an error (Pienciak-Siewert et al., 2020), or by manipulating the level of uncertainty of state or the environment (Gonzalez Castro et al., 2014; Wei & Koerding, 2010). Traditionally interpreted as learning to a lower extent, increased residual error in an adaptation task is influenced by the consequences of an error (Trent & Ahmed, 2013) or the level of effort (Emken et al., 2007). These studies, in conglomeration, show that there is more at play in motor

learning than just error reduction. When evaluating the extent of learning in motor learning studies, residual errors do not necessarily mean a reduced amount of learning has occurred, but rather, subjects could be optimizing for factors in addition to error.

Optimal feedback control theory offers a formalization of these trade-offs and has been used to describe human movements (Franklin & Wolpert, 2011; Izawa et al., 2008; Nagengast et al., 2009; Todorov, 2005). Using this framework, subjective costs are quantified and used to develop a control policy, offering insight beyond kinematics into the subjective strategies and decisions of each individual. Izawa and colleagues showed that in an adaptation task, higher error is acceptable, and indeed optimal, if the total movement cost includes not only error but effort as well (Izawa et al., 2008). Others have used an optimal control model to formalize trade-offs between reward and effort that could predict not only which movement would be made, but also how that movement would be performed (Rigoux & Guigon, 2012). Optimal control models offer insight that is deeper than a comparison of *how* the kinematics of two movements differ; by investigating the subjective costs used, these models offer an explanation for *why* the movements differ.

Many motor learning studies implicitly assume populations use identical strategies when learning new motor tasks, thus kinematic differences between populations are evidence for a difference in the ability to learn. Using error as a correlate for the amount learned, prior literature shows mixed results as to whether motor learning capability declines with age. A number of studies have concluded that

our ability to learn new motor tasks declines with age (Buch et al., 2003; Etnier & Landers, 1998; Huang & Ahmed, 2014; McNay & Willingham, 1998; Seidler-Dobrin & Stelmach, 1998; Vandevoorde & Orban de Xivry, 2019), attributing this observation to underlying causes such as motor variability (Bruijn et al., 2012; Cooke et al., 1989), sensory deficits (Christou, 2011; Lord, 2006) and attention (Goble et al., 2009). However, the causes are inconsistent and are highly dependent on the task structure, complexity, and familiarity (Voelcker-Rehage, 2008). In some cases, older adults learn to the same extent as younger adults (Cesqui et al., 2008; Kitchen & Miall, 2021; Reuter et al., 2018; Trewartha et al., 2014; Wishart et al., 2002). In other studies, we see older adults exhibit different movement strategies than younger adults (Cesqui et al., 2008; Morgan et al., 1994). One study has shown that older adults put higher subjective value on effort than their younger counterparts (Sidney & Shephard, 1977). The observed kinematic differences between older and younger adults (Cesqui et al., 2008), combined with their difference in strategies and subjective cost values, suggest that it is important to isolate and investigate these variables separately.

Specific to reach adaptations to a curl field, a few recent studies have found no differences in learning between older and younger adults (Kitchen & Miall, 2021; Reuter et al., 2018; Trewartha et al., 2014), however each use slightly different protocols. Comparing protocols to the study conducted by Huang and Ahmed (Huang & Ahmed, 2014), which shows significant differences, protocols differed in curl field gains, reach distances, reach directions, and/or inter-trial intervals. These differences

in protocol may be the reason for why differences are observed in Huang and Ahmed where others have not: the task demands more effort which may have accentuated the effect of age-related changes in sensitivity to effort. Huang and Ahmed's protocol uses a higher value for curl gain (-20 N-s/m), longer reaches (20 cm), and only brief pauses (no robot-guided reset to the starting point, reaches starting every two seconds).

Here we ask whether observed age-related errors during motor learning in a velocity-dependent force field can be explained by a difference in subjective costs rather than a difference in learning. We focus specifically on a dynamics learning task, wherein subjects must reach in a velocity-dependent force field. In this task, learning involves a trade-off between effort and error reduction. Some recent studies have shown that older and younger adults perform similarly in adapting to a velocity dependent force field. However, Huang and Ahmed observe a difference between age groups in their 2014 study (Huang & Ahmed, 2014). We use data from this previously published study and fit an optimal feedback control model to the younger and older adult reach trajectories. Using a simple model, we first demonstrate how differences in strategies, which we quantify through subjective costs, can give rise to changes in learned behavior usually interpreted as reduced learning. Next, we determine the range of strategies that can explain the observed behavior. Together, the results demonstrate a large overlap across the younger and older adults in the proportion learned if their subjective costs are different. In particular, the two groups appear to

learn similar amounts, but older adults place a higher subjective cost on effort required to reduce kinematic errors.

3.3 Materials and methods

3.3.1 Modeling subjective cost trade-offs

Trade-offs between subjective costs and the amount of learning can produce equivalent control laws, and thus trajectories. This concept is best illustrated with a simple model. Let us consider a one-dimensional linear dynamical system:

$$\dot{x} = ax + bu \tag{1}$$

where x represents the state, u, the control, and a and b parameterize the dynamics. We can model learning as the process of estimating an internal model of the state dynamics where $\hat{a}$ is the estimate of the true value a, scaled by ε, the "proportion learned".

$$\hat{a} = \varepsilon a \tag{2}$$

Given this internal model, we define the conventional linear quadratic cost, J, which penalizes state deviations from zero (kinematic error) and control (effort) with weights q and r, respectively.

$$J = J_q + J_r = \int_0^\infty qx^2 dt + \int_0^\infty ru^2 dt + = \int_0^\infty (qx^2 + ru^2)\, dt \tag{3}$$

The solution to this cost is the well-known LQR solution, and the control, u, is expressed below:

$$u = -Kx = \frac{-1}{b}\left(\hat{a} + \sqrt{\hat{a} + \frac{b^2 q}{r}}\right)x = \frac{-1}{b}\left(\varepsilon a + \sqrt{\varepsilon a + \frac{b^2 q}{r}}\right)x \tag{4}$$

From this equation we see that similar control laws and resulting trajectories can be created through different combinations of the parameters, q, r, and ε. To illustrate, the controller obtained for one particular value of q, r, and ε can be identical to another controller with a reduced ε by increasing the penalty on control input, r, or decreasing the penalty on state, q, or a combination of the two. Accordingly, two different trajectories do not necessarily indicate a different internal model of the dynamics; this could be a result of having differing costs. In other words, there is no unique mapping from the control law and trajectory to the proportion of the dynamics learned. These trade-offs are visualized in Figure 1. In the case of this simple model, q and r represent the subjective costs of a person or population and ε represents the proportion of the dynamics that person or population has learned. This relationship clearly demonstrates how changes in subjective costs across individuals can mask differences in how much they have learned.

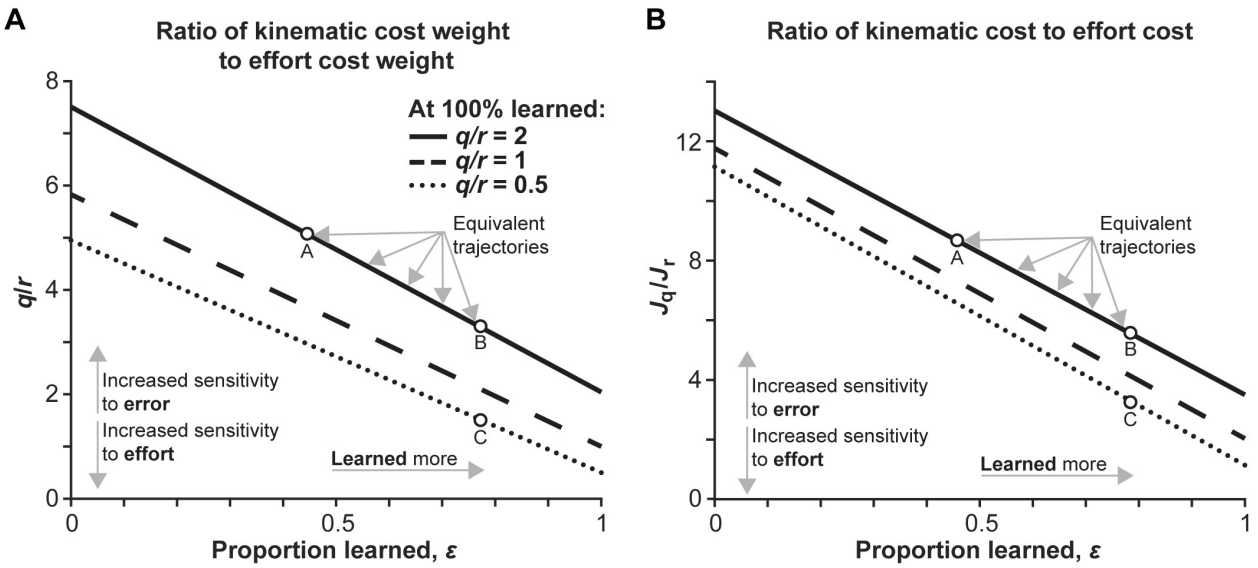

Figure 1: Trade-offs between subjective costs and proportion learned can produce equivalent trajectories. Each line represents a unique solution produced by ratios of kinematic cost weight to effort cost weight of 2, 1 and 0.5 with perfect knowledge of the dynamics. As the extent of learning is reduced, moving left along the x-axis, we see that in order to produce an equivalent trajectory, represented by each line, the ratio of kinematic error costs to effort costs would need to increase. *A:* shows the ratio of specific cost weights, q and r. *B:* shows the ratio of the total kinematic cost (J_q) combining q and x to the total effort cost combining r and u (J_r). Both *A* and *B* show that a trajectory with a lower proportion learned, but with a ratio of kinematic costs to effort costs that is greater could be identical to a trajectory with a higher proportion learned but with a lower ratio of kinematics costs to effort costs. Conversely, if two groups used the same strategy (ratio of cost weights) and produced different trajectories, then they must have learned a different amount. For example, Point A and B are both equivalent trajectories because they fall along the same line, however, Point A has learned less than Point B. Thus, to produce the same trajectory as Point B, the ratio of kinematic costs to effort costs is higher for Point A. Next, Point B and Point C have learned to the same extent, but because Point B has a higher relative kinematic cost to effort cost than Point C, they produce different trajectories (i.e., they are not on the same line). These two comparisons show that kinematic observations alone do not correlate to a specific proportion learned.

3.3.2 Experimental setup

Using the example illustrated above, we extend the model to apply it to an experimental dataset of subjects performing arm reaches in a motor learning task (Huang & Ahmed, 2014). In that study, younger and older adults made planar arm reaches while holding onto the handle of a robotic manipulandum (shown in Figure 2). After a baseline period of 200 reaching movements between a start position and

target circle, the robot applied a velocity-dependent force to the hand, acting perpendicular to its velocity by the following equation:

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = b \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix} \quad (5)$$

In Equation 5, b is the curl field gain, which was held constant at -20 N-s/m during the learning phase. As is typical for this paradigm, the force perturbation led to large deviations (i.e., errors) from the baseline movements that were subsequently reduced over many reaches. Abrupt removal of the force field led to large deviations in the opposite direction. To measure the horizontal force exerted, subjects were exposed to a force channel trial every five trials throughout the experiment, where reaches appeared to travel in a straight-line path towards the target. While the trends in error onset and reduction were similar in both younger and older adults, there were distinct differences in how they chose to reach the target.

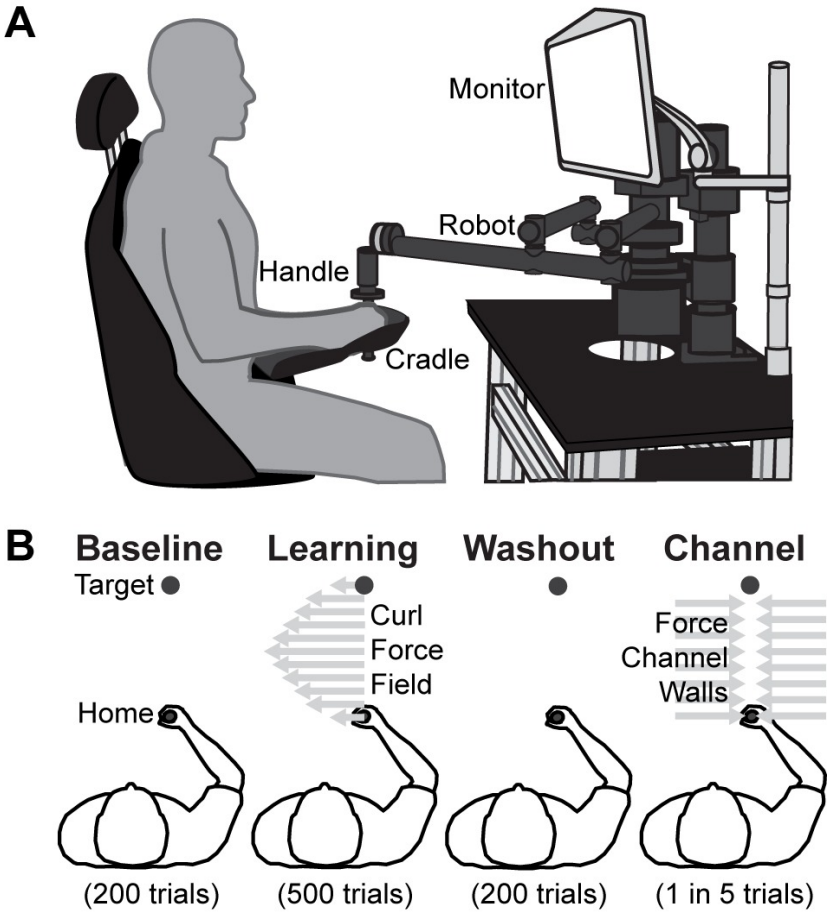

Figure 2: Experimental setup and conditions. *A*: subjects controlled a cursor on a computer screen by making reaches while grasping a robotic manipulandum, reaching anteriorly from one circle to a single target circle. *B*: subjects experienced 900 trials of varying dynamics. Baseline and washout phases have no perturbations, while learning imparted a curl field force specified in Equation 5 and represented by the gray arrows pushing them perpendicular to the direction of their reach. Channel trials were interspersed throughout the experiment, where subjects moved along a straight-line path to the target within a force channel. These trials are used to estimate the extent to which subjects were compensating for the dynamics by measuring the force that subjects pushed into the force channel wall, represented in gray. Details of how the channel trial is implemented are summarized in the Methods.

Throughout our analyses, we quantify performance using three common metrics of learning: maximum perpendicular error, maximum perpendicular force, and a trajectory-derived adaptation index. Maximum perpendicular error is measured in non-channel trials and is calculated as horizontal deviation from a straight-line trajectory from the start position to the target. Maximum perpendicular force is measured in channel trials and is a coarse reflection of learning as it is a

measure of the anticipatory force the subject is generating to counter the force field. The adaptation index normalizes the time profiles of anticipatory force by the velocity of the movement, purportedly correlating to a subject's estimate of the curl field gain. However, this metric is a reflection of desired error cancellation: if force is perfectly correlated to velocity in a curl field, then reaches will have no horizontal deviation. As shown by Izawa and colleagues (Izawa et al., 2008), some horizontal deviation is optimal even with perfect adaptation, thus adaptation index is influenced by subjective strategies. We focus on performance at four phases of the experiment: the last ten trials of the baseline period (late baseline), the first five and last ten trials in the learning period (early learning and late learning), and the first five trials after the perturbation was removed (early washout). To increase confidence by including more data, ten trials were included during the late phases, as these were less variable trial to trial. Because adaptation and de-adaptation occur so quickly, only the first five trials were considered during early learning and early washout.

This experiment used data from Huang and Ahmed 2014, which investigated differences in learning between older and younger adults (Huang & Ahmed, 2014). We will briefly review the experiment here and refer the reader to the original publication for greater detail. Huang and Ahmed's study (Huang & Ahmed, 2014) utilized data from eleven older adults (mean ± s.d., age 73.8 ± 5.6 years) and 15 younger adults (23.8 ± 4.7 years). The analyses in this study uses data from 15 older adults (age 72.9 ± 5.44, 7 female, 8 male) and 13 younger adults (age 24.7 ± 4.6, 13 female, 2 male). In the original submission, four older adults were not included due

to poor quality of metabolic data. However, because this study analyzed reaching trajectories only, these subjects' data could be included. Additionally, two younger adult subjects' data were not used due to corrupted or inaccessible reach trajectory data. Despite these different subject numbers, conclusions from the original manuscript are the same.

Each subject made targeted reaching movements while grasping the handle of a robotic arm. The experimental setup is visualized in Figure 2A subjects were seated with their right forearm cradled and the computer screen set at eye level. Reaches were made in the anterior and posterior directions. Successful reaches needed to reach the within a movement time between 300 and 600 ms. Figure 2B outlines the various dynamics subjects experienced throughout the course of the experiment. Each subject made 900 reaches with the robotic arm, 450 anteriorly, 450 posteriorly. For the middle 500 reaches, subjects were exposed to a velocity-dependent force (curl) field. The forces imparted by the robotic arm per Equation 5, where the forces, f_x and f_y, imparted on the hand were proportional and perpendicular to the velocity of the hand, v_x and v_y, scaled by the curl field gain, b. In this experiment, b is -20 N-s/m. The progression of trials was broken into three blocks: 200 trials with no forces (baseline), 500 trials with the curl field on (learning), and a final 200 trials with no forces (washout).

Throughout the 900 trials, subjects were exposed to one force channel trial every five trials, pseudo-randomly dispersed. The channel trial forced subjects to reach in a straight line, while simultaneously measuring the forces exerted on the

robotic arm. The robot arm enforced the channel trial using a horizontal force relative to the horizontal position and velocity, summarized in the equation below:

$$f_x = -2000p_x - 50v_x \quad (6)$$

As subjects adapted their reaches to the curl field, they began to anticipate the perturbing forces. On channel trials, subjects' anticipatory compensation resulted in exerting force against the channel, opposite the direction of the curl field perturbation. The measured force trace, often analyzed in conjunction with the velocity trace, provides insight into how well subjects have learned the novel dynamics. The exact process of estimating this value is detailed in the Learning Metrics section below.

3.3.3 Data preparation

3.3.3.1 Trajectory data

To analyze the performance of our model, we considered only the outward reaches from the final ten trials in the late baseline phase and late learning phase, and the outward reaches from the first five trials of the early learning and early washout phase. All position, velocity and force data were collected at 200 Hz. To be included in the analysis, the subject must have reached the target between 250 ms (50 samples) and 1500 ms (300 samples) after the target appeared. Movement onset was defined as when the anterior velocity was ≥ 0.03 m/s towards the target, and movement termination was defined as when the cursor was within the target area and anterior velocity was ≤ 0.03 m/s. If the trial never reached the target area, and/or did not sufficiently slow down, it was still included, but subjected to the reach time

criteria mentioned above. Collectively across experimental phases and subject groups, 96% of trials met these criteria. Younger adults had 5.8% of trials that did not meet the time criteria and older adults, 2.9%. Of the failed trials, 74% were in the Early Learning phase, and 21% in the Early Washout phase. Finally, if data were corrupted or the robotic arm setup experienced any issues, these trials were thrown out (<0.1% of trials). The trials that met these criteria were averaged across each subject with each resampled to the mean trial length of that subject's trajectories. This set of trajectories was used for fitting individual subject data. To obtain group averages for the younger and older subject groups, the mean trial length across subjects was calculated, each subject's mean trajectory was normalized to that trial length, then averaged across subjects. Channel trials were considered separately from the non-channel trials but used the same criteria and processing method. Due to the pseudo-randomly dispersed trials in the original experiment and the criteria set for an acceptable reach, there was typically one or zero channel trials for a subject in the set of trials considered for each phase. However, it was essential to only consider the first five reaches in early adaptation and early washout, because adaptation and de-adaptation occur very quickly. We chose these stricter inclusion criteria to represent model assumptions more accurately, thus performance metric values presented here are slightly different than the numbers in the original manuscript; however, they do not alter the conclusions made in the previous manuscript.

3.3.3.2 Learning Metrics

From the time-normalized average trajectories, we calculated commonly used learning metrics: maximum perpendicular error, maximum perpendicular force, and adaptation index. Maximum perpendicular error is calculated as the largest absolute perpendicular deviation from a theoretical straight line that connects the start position to the target position. Often, the net perpendicular deviation is considered, where the average perpendicular deviation in the training phase, or baseline phase, is subtracted from the perpendicular error in the novel environment. Using the net deviation accounts for any natural curvature in reaches due to the biomechanical constraints of the arm and allows for a better within-subject analyses. In this experiment, however, we were more concerned about the comparison between subject groups, so total perpendicular error was a more appropriate metric.

Because this experiment was a force-adaptation task, it was also useful to consider how the anticipatory force changed over the course of the experiment. Using data collected from channel trials, we calculated the maximum force that each subject pushed against the channel, perpendicular to the direction of movement. A positive value indicates a force in the right-hand direction against the perturbing force, while a negative value indicates a force in the left-hand direction with the perturbing force. Accordingly, this value can be compared to an expected maximum perpendicular force, calculated from the maximum velocity and curl field gain.

A more comprehensive metric for adaptation, assuming the strategy is to reach as straight as possible, is to compare the entire force trace to the entire velocity trace

in a channel trial. In a curl field trial, the horizontal force is applied to the hand proportional to the vertical velocity through the scalar value, b, as per Equation 5. If a trajectory accurately anticipates and compensates for these forces, the force trace measured in a channel will be approximately equal to b times the velocity profile. If the dynamics are underestimated, the force profile will be equal to the velocity profile scaled by some value less than b. Thus, using the measured horizontal force trace and dividing it by the vertical velocity trace, the scalar value that best approximates this linear relationship is an estimate for the amount of adaptation that has occurred, which we call the adaptation index.

3.3.4 Arm Reach Model

We employed a discrete-time, two-dimensional, finite-horizon, linear quadratic optimal control model using a symmetrical point mass to describe the arm reaching trajectories. The model included an internal estimate of the state dynamics to calculate the control law (referred to as the proportion learned, above), where all states were deterministically observable and there was no system variability or uncertainty. The model also includes higher derivative terms analogous to muscle activation filters (Nagengast et al., 2009). The system dynamics are governed by the following equation:

$$\boldsymbol{x}_{t+1} = \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{B}\boldsymbol{u}_t \tag{7}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are the dynamics of the system, and $\boldsymbol{x}$ is the state vector defined as follows:

$$\boldsymbol{x} = \left[p_x\ p_y\ v_x\ v_y\ f_x\ f_y\ \dot{f}_x\ \dot{f}_y\ T_x\ T_y\right]^T \tag{8}$$

The variables p_x and p_y represent hand position, v_x and v_y, hand velocities, f_x and f_y are hand forces, $\dot{f}_x$ and $\dot{f}_y$ are the rate of change of hand forces, and T_x and T_y are target position. In this paradigm, target position is constant. The motor commands, u_x and u_y are the second derivative of force. The matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ encapsulate the dynamics of the curl field and channel trials, as outlined in Equations 5 and 6. Additionally, a separate, but similar matrix, $\hat{\boldsymbol{A}}$, represents the internal model of the dynamics, which includes a subject's estimate of the curl field gain instead of the true value.

The cost function used to calculate the control law is defined as:

$$J = \sum_{t=1}^{T-1} (\boldsymbol{x}_t^T \boldsymbol{Q} \boldsymbol{x}_t + \boldsymbol{u}_t^T \boldsymbol{R} \boldsymbol{u}_t) + \boldsymbol{x}_T^T \boldsymbol{\Phi} \boldsymbol{x}_T \tag{9}$$

where the matrices $\boldsymbol{Q}$, $\boldsymbol{R}$, and $\boldsymbol{\Phi}$ are symmetric matrices, which penalize state tracking, control input and terminal state, respectively. The matrices represent these penalties as cost weights along the diagonal of each matrix. Within $\boldsymbol{Q}$, there are separate costs for horizontal deviation from target/centerline (($\boldsymbol{Q}$(1,1) = $\boldsymbol{Q}$(9,9) = -$\boldsymbol{Q}$(1,9) = -$\boldsymbol{Q}$(9,1)), vertical distance from target ($\boldsymbol{Q}$(2,2) = $\boldsymbol{Q}$(10,10) = -$\boldsymbol{Q}$(2,10) = -$\boldsymbol{Q}$(10,2)), horizontal velocity ($\boldsymbol{Q}$(4,4)), force ($\boldsymbol{Q}$(5,5) = $\boldsymbol{Q}$(6,6)), and rate of force ($\boldsymbol{Q}$(7,7) = $\boldsymbol{Q}$(8,8)). Within $\boldsymbol{R}$, there is a cost on the second derivative of force ($\boldsymbol{R}$(1,1) = $\boldsymbol{R}$(2,2)). And within $\boldsymbol{\Phi}$, there are costs on the end position relative to target position ($\boldsymbol{\Phi}$(1,1) = $\boldsymbol{\Phi}$(2,2) = $\boldsymbol{\Phi}$(9,9) = $\boldsymbol{\Phi}$(10,10) = -$\boldsymbol{\Phi}$(1,9) = -$\boldsymbol{\Phi}$(9,1) = -$\boldsymbol{\Phi}$(2,10) = -$\boldsymbol{\Phi}$(10,2)), the end velocity relative to zero ($\boldsymbol{\Phi}$(3,3) = $\boldsymbol{\Phi}$(4,4)), end force relative to zero ($\boldsymbol{\Phi}$(5,5) = $\boldsymbol{\Phi}$(6,6)), and end rate of force relative to zero ($\boldsymbol{\Phi}$(7,7) = $\boldsymbol{\Phi}$(8,8)). All other elements in these

matrices were zero. Collectively, these matrices represent the subjective costs for a movement and are and are used to formulate a movement plan where the control sequence, determined by the control law, is calculated with these cost matrices and the estimated dynamics, $\hat{\boldsymbol{A}}$ and $\boldsymbol{B}$. Trajectories were simulated by running the control sequence forward in time through the true dynamics (either the with or without a curl field or within a force channel).

3.3.5 Trajectory Matching

When fitting the model to the experimental results, we approached the trajectory matching problem using optimization techniques. We sought to minimize an objective function that described the data through varying the cost weights in $\boldsymbol{Q}$, $\boldsymbol{R}$, and $\boldsymbol{\Phi}$ used to calculate the optimal control model and trajectory. In some cases, the internal model of the dynamics (i.e., proportion learned), was allowed to vary as well. We used MATLAB's constrained minimization function, *fmincon*, designed for nonlinear optimization problems. The minimizing solution was obtained by comparing results from multiple restarts with randomized initial parameter values. Of the multiple restarts, the solution that resulted in the smallest value of the objective function was chosen.

3.3.5.1 Objective Function

Our initial analysis of group data considered four phases of the experiment: late baseline, early learning, late learning and early washout, and the fit sensitivity and cost ratio analyses considered only late learning and early washout. Within each phase, both channel trials and non-channel trials were considered in the objective

function. Within the objective function, the weighted sum of the z-scores of the data's end point, maximum perpendicular position (error), adaptation index, and maximum perpendicular force are penalized. When analyzing trajectories without confidence intervals, as is the case with the model validation and individual data, the objective function penalized these same metrics but did not z-score them.

To find the trajectory that minimizes the objective function, the proportion learned varied in some cases, and the weights on each cost parameter were varied. Cost parameters that penalized the horizontal hand position, vertical hand position, hand force, rate of change of hand force, second derivative of hand force (control input), terminal position, terminal velocity, terminal force, and terminal rate of change of force were allowed to vary. Each of these parameters were along the diagonal of cost matrices $\boldsymbol{Q}$, $\boldsymbol{R}$, and $\boldsymbol{\Phi}$. Constraining these matrices to be diagonal limited the quality of fits thus our solutions represent a lower-bound on the quality of fits achievable. In the initial fit across all four phases, the value of the internal model's proportion learned for early learning, late learning and early washout were allowed to vary in finding the best fit, while late baseline was fixed at a value of zero (i.e., no learning). In the subsequent analyses of group data, we find how sensitive the best fit trajectory was, the internal model of the dynamics was fixed at values from 40 to 120% of the curl field gain in the late learning phase, while the other parameters were still allowed to vary. Ultimately, this produced a range of learning values and subjective costs that accurately describe the data, where the learning metrics must fall within the set maximum confidence interval (initially 95%). The

intersection of proportion learned between subject groups was investigated further using the same method described above, but with more stringent success criteria. When analyzing individual subject data, proportion learned and cost weights could vary, and the objective function did not z-score the learning metrics.

3.3.5.2 Search Function Performance

We took an incremental approach to verifying search function's performance, where the search function could vary cost weights, proportion learned, or both, to fit a set model-generated trajectories. We tested the search function's ability to fit to combinations of two proportions learned (60% and 100%) and the high and low cost ratios (-0.0175 and -3.36, respectively). Initially, we fit only a single trajectory: the "learning" phase reach in the curl field, then we tested the search function with two trajectories: the "learning" phase in the curl field and "washout" phase in the null field. The latter, with two trajectories, is what we use throughout the paper except for the initial fit which considered all four major phases. Within each category we performed the search in three ways: holding costs constant and searching for the proportion learned, holding proportion learned constant and searching for costs, and finally letting both costs and proportion learned vary. For each search, we found ten best fit solutions, then calculated the mean cost ratio and proportion learned across these solutions.

3.3.5.3 Model Sensitivity Analysis

Simulated reaches whose learning metrics fell within the specified confidence intervals (initially 95%) of the experimentally obtained metrics were considered

statistically indistinguishable from the data. Only solutions that had all learning metrics that fell within this range for both late learning and early washout were considered acceptable solutions. After these solutions were found, they were checked by comparing their spatial plots to that of the data.

First, we compared the resulting ranges of acceptable values of proportion learned with 95% c.i. If there was no overlap, then we could conclude the two populations had learned a different amount. If there was overlap, then investigating how the specific cost parameters differ between subject groups could explain the differences in reaching behavior. Next, we tested the robustness of this result by enforcing more stringent success criteria. Following the same procedure, a solution was only deemed acceptable if the learning metrics fell within an incrementally decreasing confidence bound. We chose to investigate 90% c.i., 80% c.i., 70% c.i, and 60% c.i. These bounds corresponded to incrementally decreasing the acceptable range of learning metric values. As with the previous analyses, only a solution that had learning metrics fall within these bounds was deemed acceptable.

To more easily interpret the motor strategy differences, individual costs were categorized and combined into two types: kinematic or effort costs. Kinematic costs included costs on position (perpendicular error, distance to target, end position). Effort costs included force and the derivatives of force states (f_x, f_y, $\dot{f}_x$, $\dot{f}_y$, $\boldsymbol{u}$). Costs for each state were normalized by the sum of the squared states for to account for the difference in number of samples between subject groups. Each category of costs was summed together, then to account for a potential uniform scaling of cost weights,

normalized kinematic cost is divided by the normalized effort cost to create a metric that encapsulated each group's subjective value of kinematic versus effort costs, referred to as the cost ratio. For each subject group, the log transform of the set of all cost weight ratios for all proportions learned were sampled with replacement 10,000 times to provide the estimated 95% confidence intervals.

3.3.5.4 Fitting to Individual Subjects

To analyze trajectories for individual subjects, we used the same inclusion criteria as the averaged group data. The same windows of trials were used: the last ten reaches of the late learning phase and first five reaches of early washout. Anterior reach trajectories were averaged across these windows to provide a single trajectory for late learning and a single trajectory for early washout per subject. In taking this approach, data is more sparse, where some individuals have only one or no successful catch trials for a given phase, which are essential for calculating maximum perpendicular force and adaptation index. Rather than not consider a number of subjects due to lack of data, and to be consistent across how we fit all subjects' trajectories, we use a more constrained objective function that does not z-score the learning metrics and only penalized deviation from the mean trajectories' maximum perpendicular error and end point accuracy. The search function was allowed to vary cost weights and a single value proportion learned that was used for both late learning and early washout to find the best fit solution. Once a solution was found that met the maximum allowable objective function, we combined the cost ratios and

proportions learned across subjects, then compared distributions between younger and older adults.

3.4 Results

3.4.1 Differences in reaching behavior between older and younger adults

Younger adults exhibited smaller perpendicular position errors than older adults in late learning (-0.553 ± 0.169 cm versus -1.20 ± 0.185 cm, $t(26) = 2.53$, $P = 0.0176$). Younger adults also exhibited greater maximum perpendicular force at late learning (11.4 ± 1.45 N versus 8.06. ± 1.15 N, $t(26) = 1.82$, $P = 0.0797$), and a higher adaptation index at late learning (0.851 ± 0.0398 versus 0.693 ± 0.0510, $t(26) = 2.38$, $P = 0.0248$). Collectively, these metrics have been interpreted as older adults learning less than younger adults (Huang & Ahmed, 2014). However, this conclusion does not consider the potential strategic differences between older and younger adults that may also cause these observed differences.

3.4.2 Validating the arm reach model

To investigate model behavior and predict possible findings between older and younger adults, we manually manipulated costs and proportions of learning and inspected the resulting trajectories run through the forward model (Figure 3A). As the cost ratio of error to effort increases, meaning the subject would care more about reducing error than expending effort, we see reduced horizontal deviations at low levels of learning in the curl field. At high amounts of learning, with low error-to-effort costs, we see trajectories with horizontal deviations into the curl field as predicted by Izawa and colleagues (Izawa et al., 2008). With a high cost ratio and

high amounts of learning, we see straight trajectories that do not deviate horizontally. During washout, we see the same trend: at high amounts of learning, lower cost ratios deviate more than higher cost ratios. These trends present in the trajectory matches expectations, and shows the model is producing the expected results from our manual manipulations.

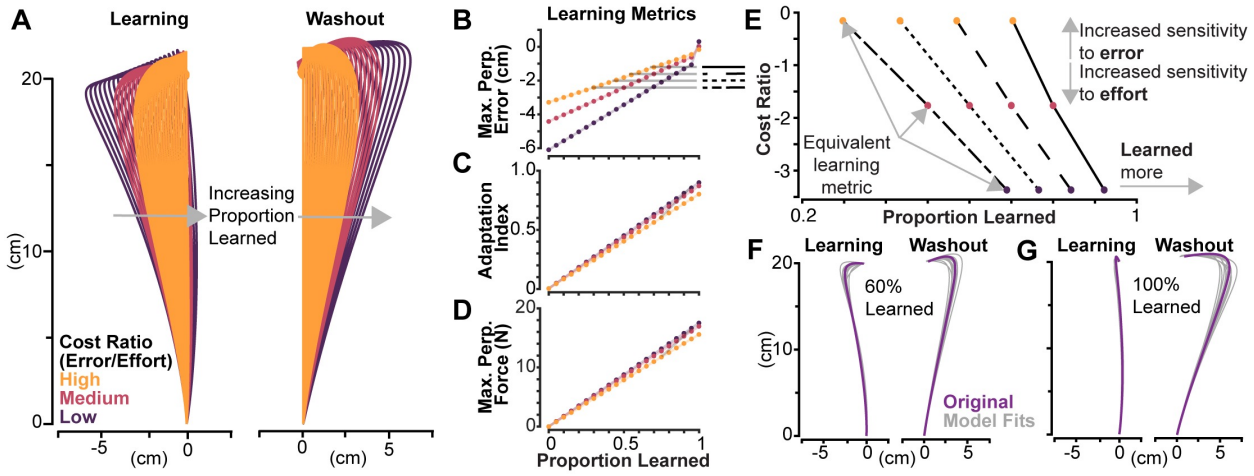

Figure 3: Model-based predictions and validation. *A*: trajectories in the curl field, i.e., "Learning" and null field "Washout" with various proportion learned are shown for three types of cost ratios: "high", "medium" and "low" in yellow, red, and purple, respectively. *B, C,* and *D*: the corresponding learning metrics maximum perpendicular error, adaptation index, and maximum perpendicular for combinations of cost ratios and proportion learned. *E*: lines connecting trajectories with equivalent maximum perpendicular are shown in. These lines correspond to the gray horizontal lines in *B*, where the solid black line corresponds to a proportion learned of 80% for the medium cost ratio, and the dashed, dotted, and dash-dot to 70%, 60%, and 50%, respectively. The same learning metric can be produced by increasing sensitivity to error and a lower proportion learned, or vice versa, showing the same trend predicted by the simple model. *F* and *G*: trajectories show model fits in gray relative to the original trajectory in purple for 60% and 100% proportion learned in the low-cost ratio condition. These results were used to validate the search function's ability to back out cost ratios and proportion learned.

Figure 3B-D visualizes the trends in learning metrics for trajectories in the learning phase across cost ratios. As the proportion learned increases, we see maximum perpendicular error steadily decrease but go beyond zero in lower cost ratios. Additionally, we see adaptation index and maximum perpendicular force increase as learning increases. For this particular set of cost ratios, lower ratios have

a slightly greater slope than higher cost ratios in Figure 3C and Figure 3D, though this trend is not consistent across all combinations of costs. In some cases, the opposite can be true, where higher cost ratios have a greater slope. Clearly shown in Figure 3B-D, evidenced by their differing slopes, a given learning metric value is not uniquely specified by a proportion learned. Instead, this value is a function of both cost ratio and proportion learned.

Using the generated sets of trajectories shown in Figure 3A, we observe the same predictions made in the simple model and Figure 1: equivalent trajectories can be produced by decreasing cost ratio and increasing learning, or vice versa. In Figure 3E, values of 50%, 60%, 70%, and 80% proportion learned for the medium cost ratio were taken as baselines, then we find proportion learned values that produced equivalent maximum perpendicular error values in the high and low cost ratios. Values for the cost ratio were calculated by taking the log transform of the sum of the kinematic or "error" costs divided by the effort costs. The less negative the value (e.g., a "high" cost ratio), the more sensitive to error. The solutions show consistent negative slopes, where the high cost ratio solution produces an equivalent maximum perpendicular error with a lower proportion learned, and the lower cost ratio could produce the same maximum perpendicular error with a higher proportion learned. These predictions match those of the simple model and show that, though more complex and not directly solvable, the arm reaching model exhibits the same behavior.

To extract cost ratios and proportion learned from trajectory data, we utilize a search function that minimizes the difference between the experimental trajectories and model-produced trajectories of which the cost ratios and proportion learned are known. The search function's ability to accurately extract these values from existing trajectories is essential, thus we sought to verify its performance using an incremental approach. To verify performance, we tested the ability to fit model-generated trajectories instead of experimental data, where the true proportion learned and cost ratios are known. We chose to test example trajectories near the extremes shown in Figure 3A. We tested the search function's ability to fit to combinations of two proportions learned (60% and 100%) and the high and low cost ratios (-0.0175 and -3.36, respectively).

The tests verified the search function was able to accurately recover cost ratios and proportioned learned. In the most complex case, where learning and washout trajectories were fit and both costs and learning were allowed to vary, the search function accurately captured these values across each condition, summarized in Table 1 below. The mean found proportion learned across all sets of trajectories roughly varied by only 1%, while cost ratios were slightly more variable, but within the correct order of magnitude. This also gives confidence that finding only 10 model solutions is sufficient to accurately capture the data on average. An example set of solutions for 60% and 100% proportion learned for the high cost ratio are shown in Figure 3F and Figure 3G.

Table 1: Search function verification results.

True Cost Ratio	True Proportion Learned	Found Cost Ratio (mean ± s.d.)	Found Proportion Learned (mean ± s.d.)
Low (-3.36)	60%	-3.37 ± 0.511	60.7 ± 1.08%
Low (-3.36)	100%	-3.36 ± 0.723	100.7 ± 0.860%
High (-0.0175)	60%	-0.124 ± 0.917	60.5 ± 1.14%
High (-0.0175)	100%	-0.0437 ± 0.656	98.9 ± 0.590%

3.4.3 Model fits to reach trajectories

With confidence that our search function can accurately find solutions, we then fit model-produced trajectories to the experimentally observed trajectories, where a single controller was used to describe all four major phases of the experiment (late baseline, early learning, late learning, and early washout), and only the model's value for proportion learned was allowed to vary between phases. Critically, we assume that within a population, the strategy remains the same throughout the course of the experiment where each phase uses the same cost parameters.

Using this method, we found model-generated trajectories that accurately described the experimental data for both younger and older adults. The results from the best fit model for each age group are summarized in Figure 4. The model-produced trajectories are compared to the experimental data for each major phase of the experiment (late baseline, early learning, late learning, and early washout). Where the spatial trajectories are shown in Figure 4A and each of the corresponding learning metrics for each trajectory and phase are shown in Figure 4B. For both younger and older subjects, the model-generated trajectories in late learning and early washout had all learning metrics fall within the 95% confidence intervals of the experimentally observed trajectories.

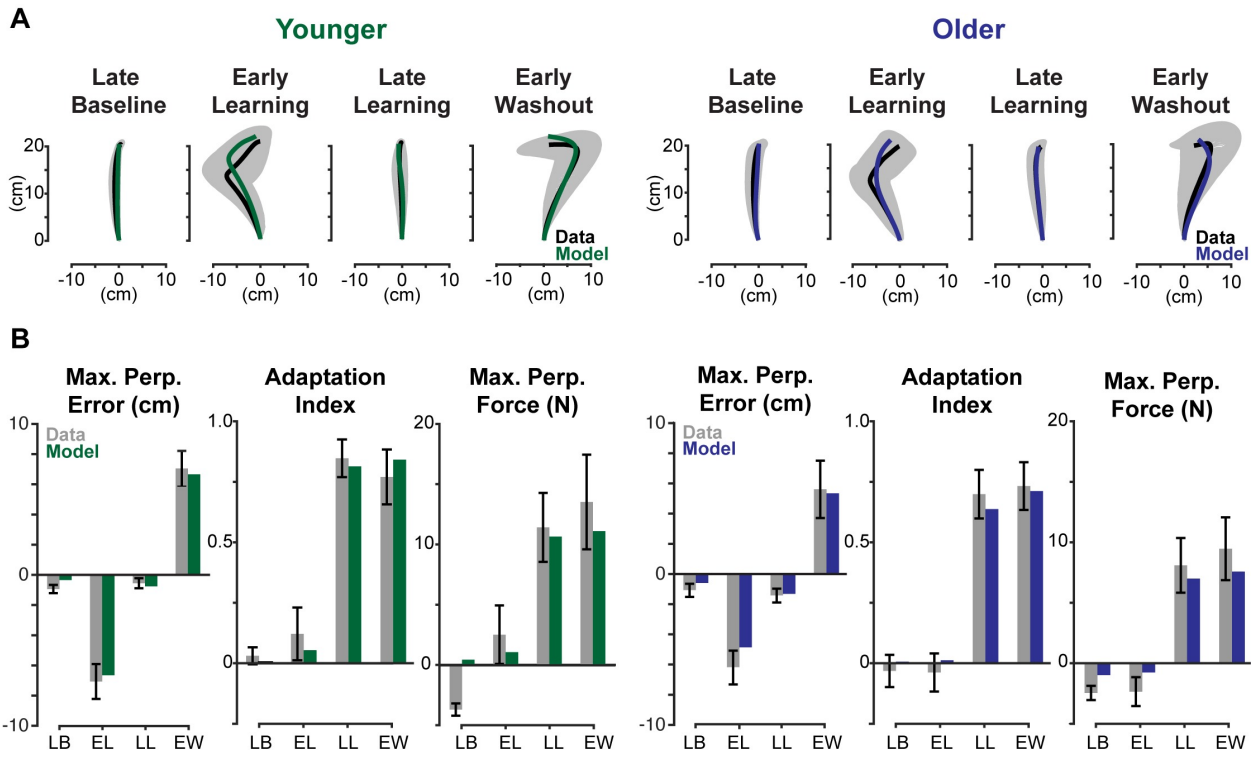

Figure 4: Model fits for younger and older adults. A: model-produced trajectories for late baseline, early learning, late learning, and early washout are compared to data from older and younger adults. Data from the experiment are shown in black with the 95% confidence interval shaded in gray. Model fits are shown in green for younger adults and blue for older adults. B: maximum perpendicular error, adaptation index, and maximum perpendicular force are compared for each phase. Experimental data represented in gray bars with error bars representing 95% confidence intervals, and model fits are represented in green and blue for younger and older, respectively.

Some learning metrics did not fall within the 95% confidence interval for the late baseline and early learning phases. The maximum perpendicular error and maximum perpendicular force for both older and younger adults fell outside this range. These phases, however, are exemplary in capturing the limitations of our model. Some of the natural curvature seen in the trajectories in late baseline, a symptom of biomechanical constraints and the dynamics of the robotic manipulandum, is difficult to capture with a simple point mass. These differences are small, as shown by the resulting spatial plots of the trajectories and are acceptable because our focus is on later phases of the experiment (late learning and early

washout), where the model captures subject behavior more reliably. Taken together, these results demonstrate that a model that assumes subjective costs do not change over the course of the experiment can reliably capture subject behavior.

The model-based proportion learned, as previously defined, provides a latent metric for the internal model of the dynamics. When fitting all four phases, we find solutions that qualitatively match the learning process. The late baseline phase was fixed at a proportion learned of zero for both younger and older adults. We found younger adults had proportion learned values 5.3%, 79.3% and 82.3% for early learning, late learning, early washout respectively, while older adults learned 0.0565%, 76.6% and 84.7%. These findings match expectations, where, for each subject group, early learning should be close to zero and late learning and early washout should be roughly equivalent. The proportions learned found for late learning suggest that younger adults and older adults learn close to the same amount. Investigating the cost ratios which sum the kinematic costs (e.g., horizontal deviation) and divide them by the sum of effort costs (e.g., force and rate of force), we found younger adults had a log transform of the cost ratio of error to effort of -4.495, and older adults, -5.369. A more negative number, as seen in older adults, indicates a higher cost on effort relative to error. Interestingly, we see a bigger separation between cost ratios than with proportion learned in late learning. This suggests that the observed kinematic differences are due to younger adults caring more about error relative to effort than older adults, rather than learning slightly more.

3.4.4 Model-derived range of learning

The previous solution presents a single best fit, but we sought to determine how robust this solution was. If we assumed the proportion learned was a particular value other than those found above, could we find a set of cost weights that could adequately describe the data? If predictions from the simple model in Figure 1 are accurate, then there may be several combinations of learning and costs that could describe the experimental data. To determine the robustness of the results above, we manually derived a confidence interval for the proportion learned. We asked whether different amounts of learning could predict similar learning metrics as previously analyzed (maximum perpendicular error, maximum perpendicular force, and adaptation index). For each subject group, we re-ran the search over late learning and early washout trajectories, where we held the proportion learned constant and searched only for a set of the subjective cost weights that could describe the data. We repeated this analysis for a range of proportion learned values, from 0.4 to 1.2 in 0.05 increments.

We set a stringent set of criteria for if a model fit adequately described the data. A model fit was only deemed acceptable if the model-generated trajectories for both phases had learning metrics that fell within 95% confidence intervals (c.i.) of the real trajectory data. Statistically, those might be deemed similar, or rather, not significantly different. Using these criteria, we found several solutions that did not adequately describe the data, leaving a small window of acceptable proportion learned values for each subject group. For younger adults, adequate solutions were

found with a proportion learned ranging from 65% to 90%, and for older adults, 60% to 85% of the dynamics: an overlap in proportion learned from 65% to 85%. These model fits, visualized in Figure 5 suggest that the two populations could had learned the same amount.

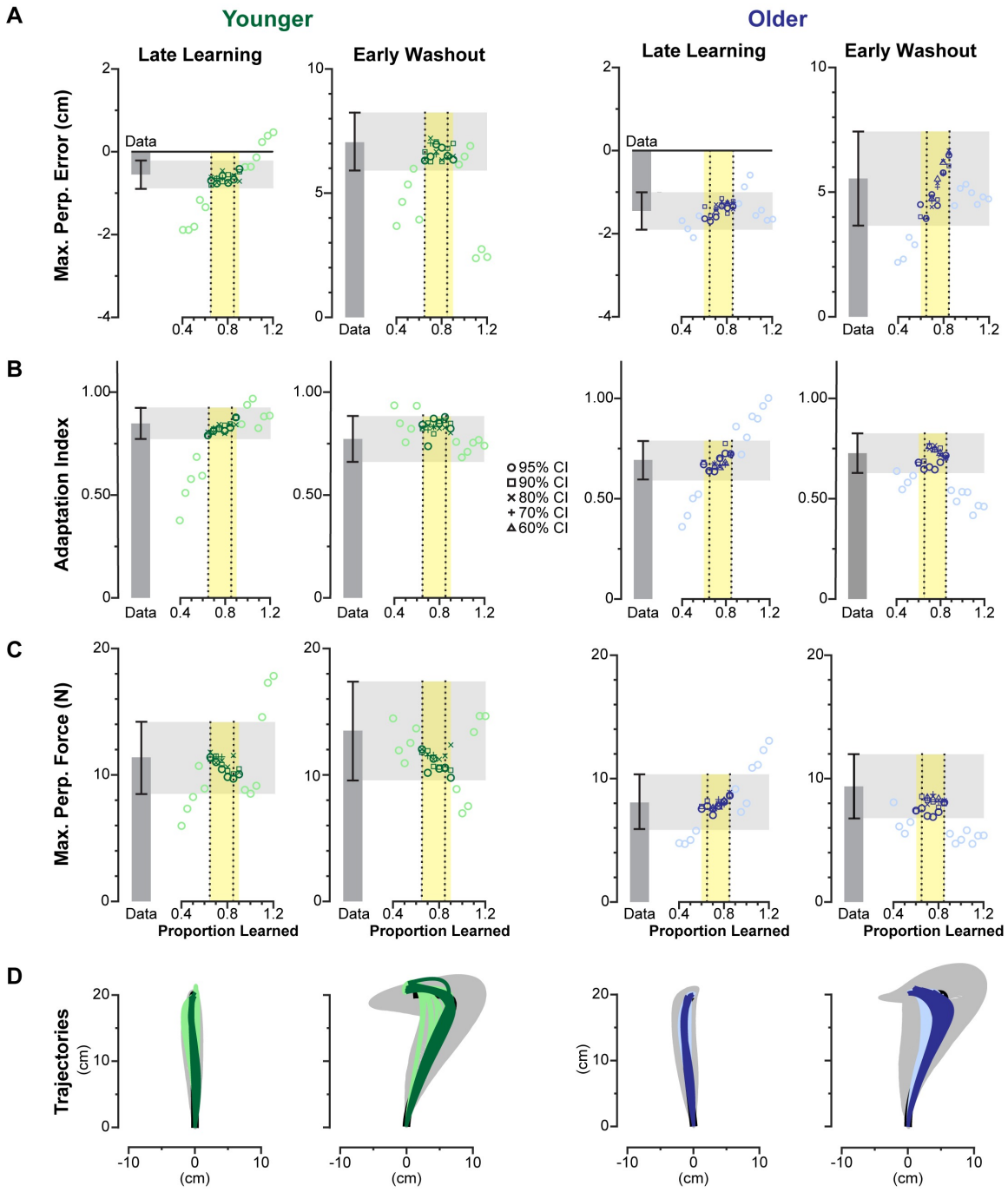

Figure 5: Adaptation metrics for model fits across different amounts of learning. Learning metrics from the data are shown as dark gray bars with black 95% c.i. The light gray region highlights the 95% c.i. bounds, deemed for the model to be statistically similar to the data for that phase and metric. The yellow area highlights the region that has statistically similar model fits across all three learning metrics, for both phases within a subject group (younger: green, older: blue). Each model fit is represented as a point, where acceptable model fits are highlighted in bold and unacceptable model fits are a lighter shade. Fits for each level of success criteria (95% c.i. down to 60% c.i.) are represented as different shapes per the legend. Fits for all three learning metrics shown in *A:* maximum perpendicular error, *B:* adaptation index, and *C:* maximum perpendicular error. *D:* each fit's resulting trajectories are shown with the actual trajectories shown in black with 95% confidence ellipses in gray.

It is important to acknowledge that 95% c.i. gives the model a significant amount of margin to work within. We also sought to test how little margin we could give and still arrive at the same result, so we repeated the process above, but with increasingly stringent acceptability criteria. Instead of using 95% c.i., we attempted to fit trajectories across same range of proportion learned, but with learning metrics falling within the 90% c.i. (1.645 x s.e.), 80% c.i. (1.281 x s.e.), 70% c.i. (1.036 x s.e.), and 60% c.i. (0.842 x s.e.). Critically, this approach assumes that we have accurately captured the mean and are simply increasing the confidence in the data by reducing the c.i. widths. As we reduced bounds of the success criteria, the range of learning that could accurately describe the data also decreased. Shown in Figure 6, 90% c.i. had the same range of acceptable fits as 95% c.i. for both subject groups. Older adults' most stringent successful fit was at 60% c.i., where acceptable solutions were found for 70%, 75%, and 80% proportion learned. For younger adults, the most stringent criteria where solutions were found was 70% c.i. At this level, we found solutions for 70% and 75% proportion learned. An overlap between younger and older still existed at 70% c.i., suggesting our results are robust to an increased confidence in trajectory data. At the 70% c.i. level, we see both age groups solutions converge towards 70-75% proportion learned, rather than diverge into different bins of proportion learned, further suggesting that older and younger adults may have learned the same amount. This result also corresponds well to the best-fit solution, which predicted younger adults learned 79.3% and older adults, 76.6%.

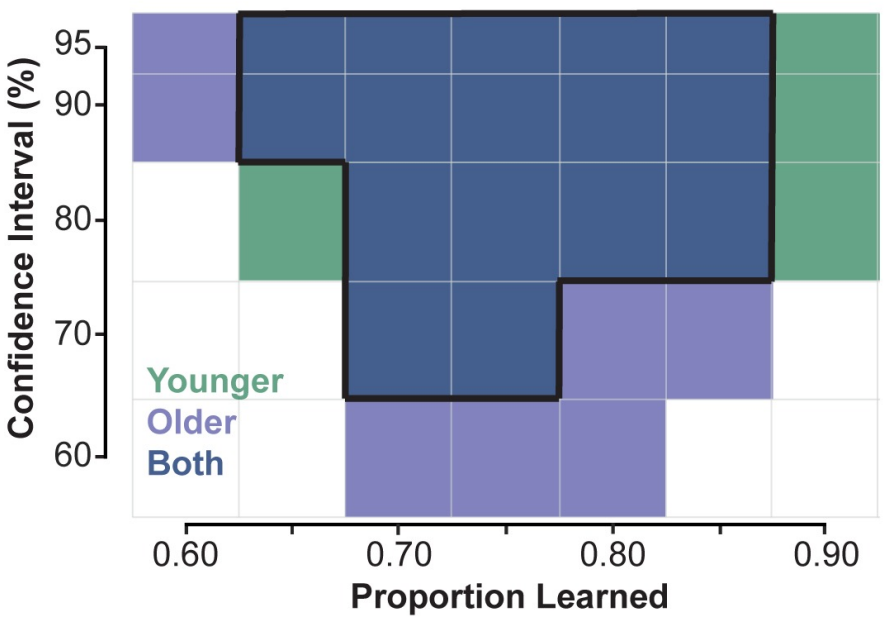

Figure 6: Successful model fits across different amounts of learning with increasingly stringent success criteria. Model fits with all adaptation metrics falling within the corresponding confidence interval highlighted in green (Younger), light blue (Older) and dark blue (both Younger and Older). Overlap between both subject groups is outlined in black.

This analysis has shown there are several solutions for both younger and older adults with equal proportions learned. If our analysis had shown there was no overlap between the proportion learned of older and younger adults, then we could posit that the two populations had learned different amounts. However, because the resulting ranges overlapped, and did so with more stringent success criteria that mimics a reduced standard error, the differences in reaching behavior between younger and older adults could be due to a difference in subjective costs, rather than a difference in ability to learn.

3.4.5 The differences in subjective costs

Model fits suggest that older and younger adults may have learned the same amount, and differences in their subjective costs resulted in different reaching

behaviors. We speculated whether the interaction of these subjective costs involved a consistent trade-off between effort and error, that could help define a general strategy for each population. To investigate this interaction, we used the same method to produce numerous additional model fits for both younger and older adults across their overlapping range of proportion learned (65% to 85%). We then analyzed the ratio of their kinematic costs (position and velocity terms) relative to their effort costs (force, rate of force, and control input) to investigate how these costs and proportion learned interact. If the predictions matched those presented in Figure 1 and Figure 3E, we would expect that greater relative costs on kinematic error could mask a deficit in learning. Similarly, we would observe that older adults, who exhibit higher error, would have consistently higher costs on effort relative to kinematic errors than younger adults.

First, we see that for equivalent trajectories within each subject group, as the proportion learned increases, the ratio of kinematic costs to effort costs decreases (Figure 7). Using a simple linear regression model to test if proportion learned significantly predicted cost ratio. The overall regression was significant for both subject groups (younger: $R^2 = 0.0954$, $F(1,48) = 5.06$, $P = 0.0291$; older: $R^2 = 0.328$, $F(1,48) = 23.4$, $p < 0.001$), where both younger and older adults have significantly negative slopes (younger: slope = -7.65, $P = 0.0291$; older: slope = -16.9, $p < 0.001$). This result matches the predictions laid out by the simple model described in Figure 1 and the arm model in Figure 3E, validating that the more complex model fit to real data exhibits this same predicted behavior. Additionally, the ratio of kinematic to

effort costs across this range of learning is significantly greater for younger adults than older adults, with bootstrapped 95% confidence intervals for younger: [-2.36,-1.39] and older: [-5.04,-3.87], which do not intersect. This further supports our explanation of the observed differences between younger and older adults: older adults could have learned to the same extent as younger adults; however, older adults placed a greater premium on reducing effort compared to reducing kinematic errors than younger adults.

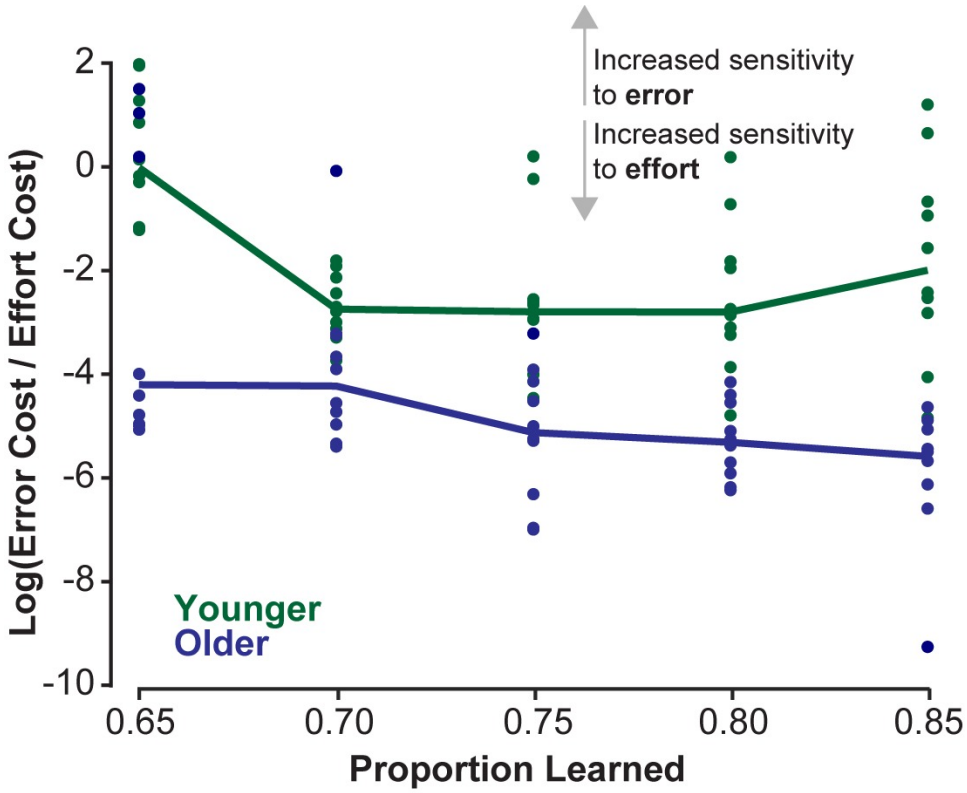

Figure 7: Older adults have a lower cost on kinematics relative to effort than younger adults. The ratio of kinematic costs to effort costs are shown for younger (green) and older (blue) adults. Each dot represents a model fit that has learning metrics which fall within 95% c.i. of the experimental data for that subject group for the prescribed proportion learned. Solid lines connect the mean cost ratio for each proportion learned per subject group. A higher ratio of kinematic costs to effort costs (larger value in the y-axis) can be interpreted as having an increased sensitivity to error, where a lower cost ratio could be interpreted as an increased sensitivity to effort.

3.4.6 Analyzing individual performance

Another approach to investigating strategic differences is to extract subjective costs and proportions learned for each individual, then collect and compare across the subject groups. For each subject, the search function varied cost weights and a single proportion learned value that was the same for each phase. Two example subjects, one younger and one older, are shown in Figure 8A, where we see the model can accurately describe these subjects' trajectories. One younger adult did not have an acceptable trajectory for late learning, thus was left out of the analysis. One younger adult and two older adults did not have acceptable model fits and were left out of the statistical comparisons between groups. Compiling the findings from each individual fit, we find younger adults learned slightly more on average than older adults (mean ± s.e., younger: 81.9 ± 4.87% older: 72.9 ± 5.52% $t(22) = 1.20$, $P = 0.242$, Figure 8B), though the groups are not significantly different. Additionally, we find that younger adults cost ratio is slightly higher (-4.29 ± 0.286 versus -5.05 ± 0.238, $t(22) = 2.07$, $P = 0.0503$, Figure 8C) though it is also not significantly different between younger and older adults.

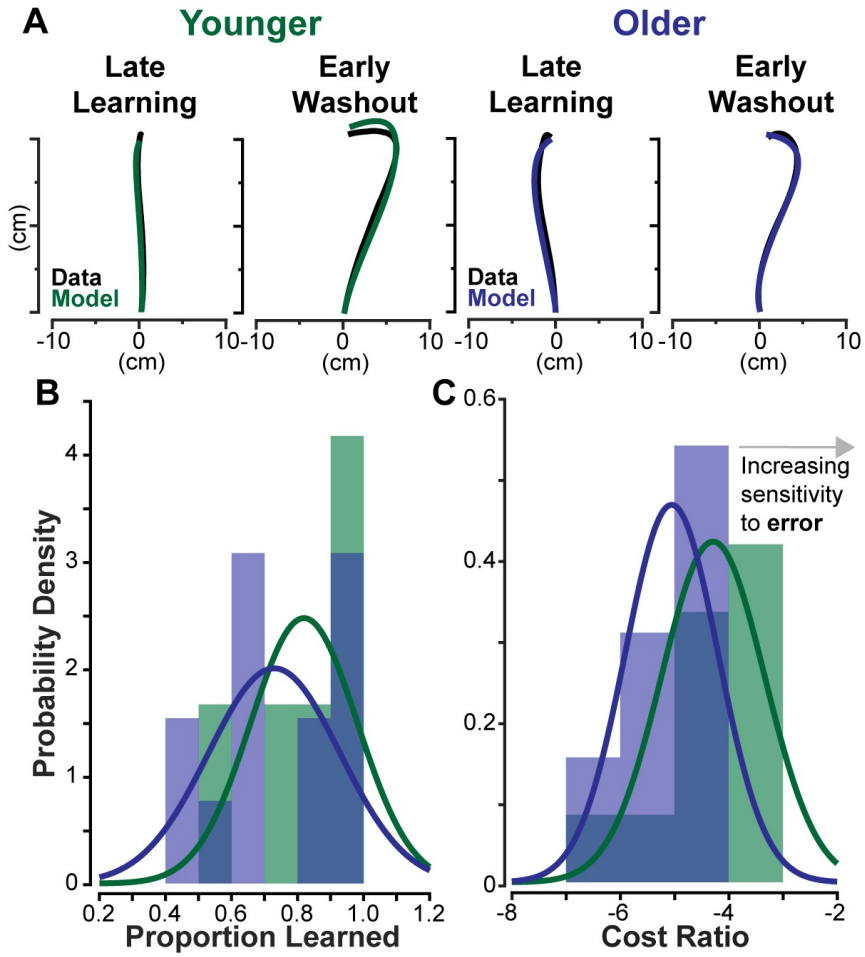

Figure 8: Individual subject trajectory fits. For each subject, trajectories were fit to late learning and early washout phases. *A*: an example fit for a younger subject is shown on the left and for an older subject on the right. The original data is shown in black, while the model produced trajectories are shown in green and blue. *B*: the distribution of proportion learned for each subject is plotted, where data for younger adults are shown in green, and older adults in blue. A normal distribution was fit to each dataset and is overlayed in bold. *C*: the same labeling and color scheme as *B* shows the distribution of cost ratios for each subject group, where less negative values demonstrate an increased weight on error costs relative to effort.

Due to the small sample size and limited number of trials, the individual analysis is limited in its utility for this particular dataset. While it does not find any statistical differences between groups, it weakly suggests that both learning and cost ratios may differ, with younger adults learning more and having a higher cost on kinematic error than their older counterparts. Moreover, the average proportion learned and average cost ratio for each group roughly matches the values we obtained

through the group analysis. At a minimum, analyzing individuals offers an alternative framework to determine subjective costs and learning and enable within subject comparisons. However, for this particular dataset, the group-averaged approach seems more appropriate in order to avoid issues with losing data.

Collectively, our analyses suggest that younger adults could have different subjective costs than older adults, and this may be the cause for observed kinematic differences. When analyzing averaged trajectories for each group, we found similar proportioned learned between younger and older adults (79.3% and 76.6%), but different cost ratios (-4.495 and -5.369). This was further supported by investigating the range of learning that could accurately describe the learning and washout phases for each group, resulting in significant overlap between subject groups (65% to 85%). Further, when the bounds for an acceptable fit became more stringent, both groups converged to the same amount of learning (70-75%). Investigating the set solutions for each subject group, across all values of proportion learned, younger adults had a significantly higher cost ratio of error relative to effort. Finally, when analyzing individual subjects, we see again that proportion learned did not significantly differ between younger and older adults, though this may be due to the small sample size and lack of power. However, the values found for both proportioned learned and cost ratios mirror analyses performed on the group average data quite well. This suggests that the kinematic differences observed between younger and older adults may be a result of different sensitivities to effort expenditure and error, rather than a difference in learning.

3.5 Discussion

Our analysis suggests that larger kinematic errors do not necessarily imply less learning, and that we must consider subjective strategies when assessing learning. Using our model, we find that older and younger adults reaching trajectories can be explained with similar amounts of learning, despite their large kinematic differences. These differences in motor behavior may be attributed to older adults caring more about effort relative to kinematic error than younger adults. Our model assumes that objective effort is similar in younger and older adults, which is consistent with measured metabolic power in Huang and Ahmed 2014 (Huang & Ahmed, 2014); thus, our results suggest that it is the subjective weighting of effort that differs (Summerside & Ahmed, 2021). An alternative explanation is that the objective effort costs are higher in older adults compared to younger adults. Locomotion studies have shown older adults incur greater metabolic cost, an objective measure of effort, compared to younger adults (Larish et al., 1988; Martin et al., 1992). While metabolic cost has been measured in both younger and older adults performing motor learning tasks, we await a direct comparison of the metabolic cost between groups. Our findings, in their current form, cannot distinguish between a higher subjective effort cost versus a higher objective effort cost in older adults.

Additionally, our results question the validity of adaptation index used in Huang and Ahmed 2014 (Huang & Ahmed, 2014) as a measure of learning in curl field experiments. We show that the same trajectory can be produced across different proportions learned, and because the adaptation index is calculated using trajectory

data, the adaptation index will not correlate to the model-based, latent state of proportion learned. This suggests that adaptation index may be a poor indicator of learning between groups, as it inherently assumes a single strategy where kinematic error-canceling is more highly weighted than effort.

Of note, our finding may be unique to the force-field adaptation paradigm. Using visuomotor rotations may eliminate any effort-centric strategic differences. However, visuomotor rotation experiments probe different mechanisms in the motor learning domain. Force-field adaptation tasks use proprioceptive feedback to make online and trial-to-trial corrections, while visuomotor rotations use visual information as a feedback signal. Because force-field adaptation tasks probe both effort and error simultaneously, the paradigm may better emulate motor tasks encountered in the real world.

Additionally, the significant difference in learning metrics between older and younger adults may be unique to the protocol used for this force-field adaptation task. The Huang and Ahmed 2014 study employed a protocol that may have put an added premium on the effort of the task (Huang & Ahmed, 2014). Due to the desire to collect metabolic data, subjects reached nearly continuously in both the anterior and posterior directions during each trial block. Other studies that did not find a difference between age groups used different reach distances, reach time constraints, curl field strengths, reach directions, and/or inter-trial intervals (Kitchen & Miall, 2021; Reuter et al., 2018; Trewartha et al., 2014). All these nuanced differences in

protocols could have increased the effect observed by Huang and Ahmed (Huang & Ahmed, 2014).

The data from Huang and Ahmed (Huang & Ahmed, 2014), while a unique finding in geriatric research, does have some drawbacks due to its limited sample size. Though the sample size was sufficient to find statistically significant difference in learning metrics, the confidence bounds combined with the degrees of freedom of the model may provide enough flexibility so that an overlap in learning would necessarily exist. An increased sample size would be helpful in drawing stronger conclusions. This shortcoming is especially present in the individual analysis, where our objective function was necessarily different due to only some of the subjects experiencing a channel trial in the early washout phase. With a different protocol or an increased sample size, results in comparing between older and younger adults may be more conclusive than what is offered in this analysis. Nevertheless, this example dataset shows how these methods can be applied to extract latent variables and offer an alternative explanation to assumed differences between age groups.

It is worth acknowledging that our model was purposefully simple, so differences between subject groups are more easily interpretable. However, as with all modeling studies, there is a trade-off between biological realism and model complexity. It could be the case that a different, or more realistic model would result in different findings. A higher fidelity model, reducing the influence of underlying assumptions, could be accomplished through a few different means.

The first improvement lies in the dynamical model. For instance, the use of a non-linear model of the arm could offer improvements over a point mass. As already discussed, this could improve the fits in the late baseline phase. This type of model could exhibit some natural deviation from the centerline and reduce costs penalizing non-straight reaches. Additionally, a more complex model that allows for co-contraction could also offer further insight into effort costs and error-canceling strategies that are independent of the perturbation. Our model only considers motor commands as a basis of effort, but subjects could be exerting more effort through co-contraction that is unable to be captured by our model. Another improvement could be to include off-diagonal terms in our cost functions to capture interaction terms between state variables. This would improve the quality of our fits, but in turn, make the model prone to over-fitting and make the subjective costs less interpretable. Finally, developing a model that incorporates motor noise, sensory uncertainty or delay may offer an alternative explanation to the differences in behavior (Sherback et al., 2010). This could be a driving force in their movement strategy that could be reflected in learning rate or trajectory differences, and potentially mask differences in learning or subjective costs. Overall, these improvements to the model could help tease out the specific differences between subject groups, and ultimately determine whether populations are learning less or compensating in a different manner.

While it is important to extract certain behaviors from observed trajectories such as arm reaches, using metrics such as maximal values cannot account for the temporal, stochastic, and highly dynamic factors surrounding human motor control.

Strong conclusions about how two populations learn differently should be extremely thorough and consider multiple metrics. The powerful framework of optimal control enables us to compare temporal data to temporal data, extract valuable information from these models, and estimate the hidden value of how much a person has learned. Deducing whether a difference in behavior is due to a difference in learning, a difference in subjective costs, or a combination of the two is still unanswered; however, we offer a framework which can probe these differences.

Our results show that subjective movement strategies can mask the latent variable of how much a person has learned. Additionally, we have shown that both older and younger adults adapt their reaches to a curl field, but whether they definitively learn to different extents remains unclear. If younger and older adults learn to the same extent, our model offers a plausible explanation that behavior differences between older and younger adults are caused by older adults caring more about effort relative to kinematic errors. We show that using learning metrics alone gives insufficient insight into the adaptation process. In future studies investigating how much a person or population has learned, it is imperative to consider their implicit strategies.

CHAPTER IV

PHYSICAL EFFORT PRECRASTINATION DETERMINES PREFERENCE IN AN ISOMETRIC TASK[2]

4.1 Abstract

How the brain decides when to invest effort is a central question in neuroscience. When asked to walk a mile to a destination, would you choose a path with a hill at the beginning or the end? The traditional view of effort suggests we should be indifferent—all joules are equal so long as it does not interfere with accomplishing the goal. Yet, when total joules are equal across movement decisions, the brain's sensitivity to the temporal profile of effort investment remains poorly understood. Here, we sought to parse the interaction of time and physical effort by comparing subjective preferences in an isometric arm-pushing task that varied the duration and timing of high and low effort. Subjects were presented with a series of two-alternative forced choices, where they chose the force profile they would rather complete. Subjects preferred earlier physical effort (i.e., to precrastinate) but were idiosyncratic about preference for task timing. A model of subjective utility that includes physical effort costs, task costs, and independent temporal sensitivity factors described subject preferences best. Interestingly, deliberation time and response vigor covary with the same subjective utility model for preference, suggesting a utility that underlies both decision making and motor control. These results suggest

[2] *The work from this chapter has been published in the Journal of Neurophysiology (Healy & Ahmed, 2024).*

physical effort costs are temporally sensitive, with earlier investment of effort preferred to later investment, and that the representation of effort is based on not only the total energy required but also when it is required to invest that energy.

4.2 Introduction

When asked to walk a mile to a destination, would you choose a path with a hill at the beginning or a path with a hill at the end? Would you ever choose a route with a bigger hill so that you could walk up that hill at a specific time? Traditional models of effort-based decision-making suggest we minimize the total amount of physical effort (Alexander, 1997; Hull, 1943; Ralston, 1958), but these models are indifferent to when periods of high physical effort occur so long as it completes the task. Yet, when posed this simple question, most people have a preference as to when to climb a hill: something the traditional model would not predict. Only recently have investigations shown that humans and animals have a preference for when physical effort costs or task-specific cognitive costs occur (see Rosenbaum and Sauerberger 2022 for a thorough review). Here, we infer the underlying relationship between physical effort and time through subjects' preferences in an isometric arm-pushing task.

Understanding how and when we decide to move (i.e., invest physical effort) is a fundamental question in movement neuroscience. Characterizing how the healthy brain represents movement costs such as physical effort and time is a necessary step in ultimately understanding the changes that occur with age (Healy et al., 2023; Lamb et al., 2016; Summerside et al., 2024) and clinical populations such as patients

with Parkinson's disease (Mazzoni et al., 2007) or multiple sclerosis (Courter et al., 2023; Goldman et al., 2008) patients. To date, many questions remain about how the brain subjectively values effort and how effort interacts with other movement costs to determine movement decisions.

In contrast to movement costs, the interaction of reward and time is better defined. In reward-based decision-making, reward is temporally discounted, where its value decays hyperbolically with time (A. M. Haith et al., 2012; Mazur et al., 1987; Myerson & Green, 1995; Shadmehr, Orban de Xivry, et al., 2010). Earlier rewards are preferred to later rewards of the same magnitude. In addition, a reward that requires less physical effort to obtain is preferred to the same reward that requires more physical effort (Hartmann et al., 2013; Klein-Flügge et al., 2015). Effort discounting of reward has been traditionally modeled as multiplicative (Klein-Flügge et al., 2015; Prévost et al., 2010; Sugiwaka & Okouchi, 2004), but recently has been shown that subtracting effort from reward can better reconcile findings in motor control with optimal foraging theory (Shadmehr et al., 2016). Logically, if reward is discounted by time, and effort discounts reward, then effort costs should be minimized and performed as late as possible. In other words, effort should be procrastinated. Some investigations support this hypothesis, using a model of utility that includes physical effort costs as negative reward, with the net value discounted by time (Bautista et al., 2001; Bruening et al., 2024; Rigoux & Guigon, 2012; Shadmehr et al., 2016; Sukumar et al., 2024; Summerside, Kram, et al., 2018). Indeed, it has been shown that cognitive effort discounting of reward decreases the

further into the future the effort is required (Johnson & Most, 2023). In sum, many of the current models of decision-making optimization would predict effort procrastination: later effort is preferred to earlier effort.

However, other recent observations contradict this: humans and animals prefer to incur some costs earlier. In a study by Rosenbaum et al., participants preferred to complete earlier subgoals at the expense of extra physical effort (Rosenbaum et al., 2014), coining the term *precrastination* to describe the subjects' alacritous approach towards task completion. In these experiments, subjects were asked to walk down an alley, pick up one of two buckets, and carry it to the end. The positions of the buckets varied relative to one another. Surprisingly, most subjects chose to pick up the bucket closer to them, needing to carry the bucket further and expend more energy than the alternative. Following this initial study, numerous experiments have reaffirmed that subjects tend to precrastinate certain tasks (Blinch & DeWinne, 2019; Fournier, Coder, et al., 2019; Fournier, Stubblefield, et al., 2019; McBride et al., 2023; Patterson & Kahan, 2020; Raghunath et al., 2021; Rosenbaum & Dettling, 2023; Schwob et al., 2022; VonderHaar et al., 2019; Wasserman & Brzykcy, 2015).

Many studies that observe precrastination are task-centric, if not purely cognitive, and have credited freeing up cognitive resources as the underlying factor for precrastination (VonderHaar et al., 2019). However, it is possible that it is the physical effort cost itself that is precrastinated. Indeed, other costs exist, such as pain, in which subjects precrastinate. Research has shown that memory of pain

follows the peak-end rule, where, in retrospect, people tend to neglect duration and instead remember the most intense moment of an experience and how that experience ends. In cold water baths (Kahneman et al., 1993), a significant majority of participants preferred the longer, objectively more painful trial that ended at a slightly warmer temperature. Patients' pain evaluation of a colonoscopy procedure correlated to how it ended (Redelmeier & Kahneman, 1996). In addition, in object manipulation tasks, people adopt awkward, hard-to-control positions at first in order to arrive at comfortable, easy-to-control positions later, essentially precrastinating discomfort (Cohen & Rosenbaum, 2004). Subjects might be precrastinating high physical effort so they end on a better note.

In many investigations of precrastination, expending higher physical effort is tied to subtask completion. If subjects chose to empty heavier buckets first (Fournier, Stubblefield, et al., 2019) or carry more dodgeballs earlier (Rosenbaum & Dettling, 2023), is this about precrastinating physical effort or about subtask completion? Here, we design a set of effortful tasks that decouple when high physical effort and subtask completion occur to investigate how the timing of physical effort costs influences preference.

Collectively, what people prefer regarding the timing of physical effort costs is not well characterized. We sought to clarify this by comparing subjective preferences across isometric force generation tasks. Drawing upon traditional models of physical effort, we predicted that subjects would be sensitive to the duration of high physical effort, preferring shorter durations, and that subjects would be agnostic to the timing

of high physical effort. We tested two effort profile types–one with a low effort baseline and brief periods of high effort, and the other, with high baseline effort and brief periods of low effort. We manipulated the duration and onset time of these brief periods to decouple the timing of physical effort and task costs and tease apart which costs may be time-sensitive. We tested this in a novel motor task yet to be explored in studies of subjective valuation of effort. We controlled for performance accuracy and explicit information about the task required. Across two sessions, subjects chose between effortful isometric arm-pushing tasks that vary in timing and duration of high effort. From each subject's preferences, we inferred a model of subjective utility used as a basis for these decisions. In a group-wise analysis, we found a generalizable model of subjective utility and determined the time-sensitivity of physical effort costs.

4.3 Materials and methods

4.3.1 Participants

Twenty-five right-handed, healthy subjects (mean ± s.d., age 24.0 ± 5.5 years, 14 females, 11 males) participated in this study. Subjects were required to be between 18 and 40 years of age, able to understand English, and right-hand dominant. Exclusion criteria included any major musculoskeletal or ophthalmological impairments, low or sedentary activity level (self-reported), and any neurological or vestibular diseases. All subjects provided written informed consent, and the experimental protocol was approved by the Institutional Review Board of the University of Colorado Boulder.

4.3.2 Experimental setup

Subjects sat in front of a computer monitor (1,280x1,024 px) and used their right arm to grasp the handle of a robotic manipulandum (Interactive Motion Technologies Shoulder-Elbow Robot, Cambridge, MA, Figure 9A). The handle was secured so it would not move, allowing subjects to isometrically push on the handle with their arm. A force sensor (ATI Industrial Automation, Apex, NC) housed inline with the robot arm was used to measure forces exerted by the subject onto the handle. A monitor mounted in front of the subject displayed interfaces that varied depending on the phase of the experiment. Subjects interacted with the displayed interfaces by exerting a force onto the handle which moved a cursor or changed the length of an arrow on the screen.

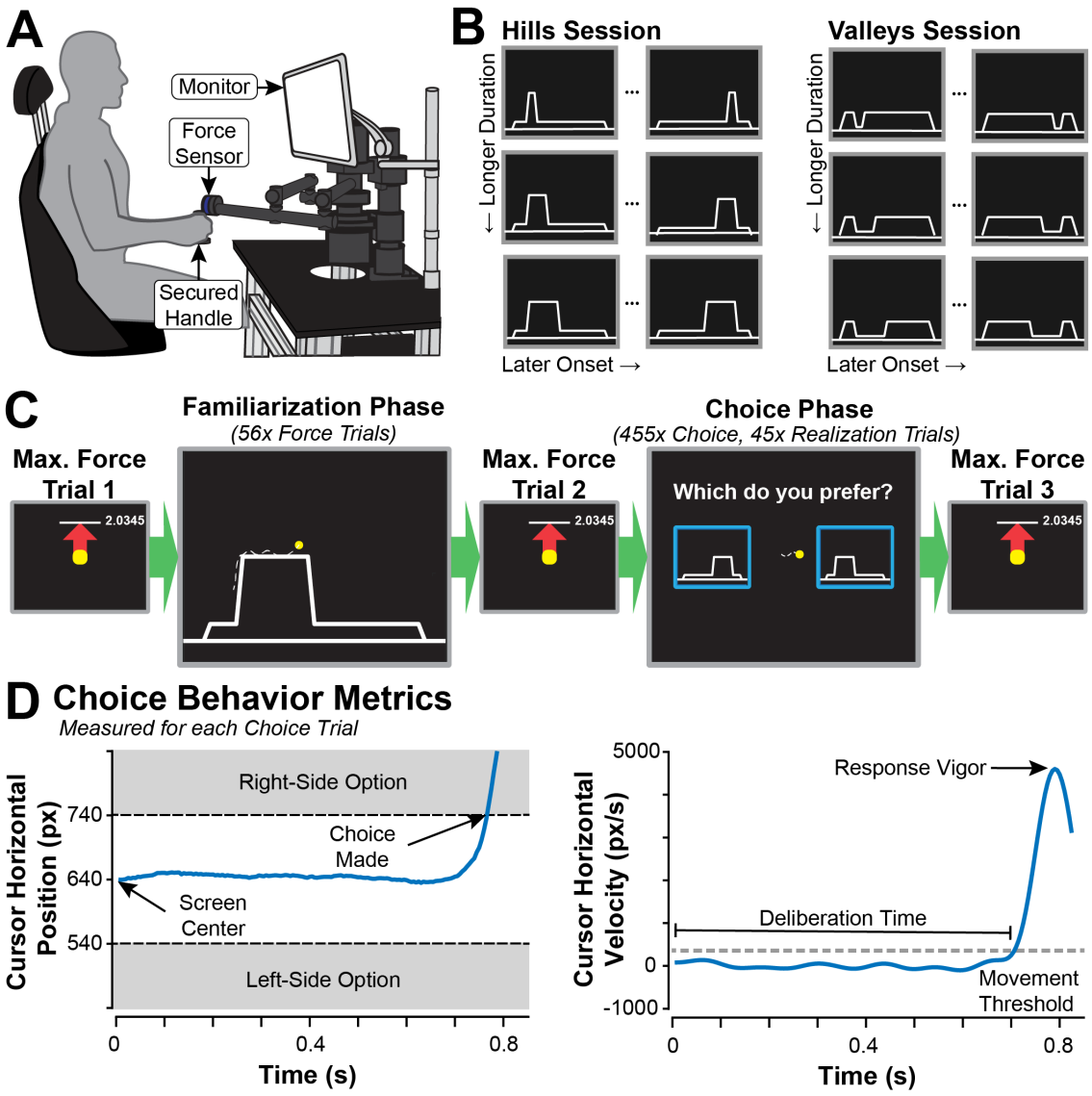

Figure 9. Experimental setup, protocol, and measures. *A:* subjects sat in front of a computer monitor and controlled the cursor's position by exerting force on a handle to a static, secured robotic arm. *B:* subjects completed two sessions with identical procedures but different types of force profiles. The Hills Session tested force profiles with brief periods of high effort, while the Valleys Session tested force profiles with brief periods of low effort. Both sessions had 14 unique force profiles determined from combinations of hill/valley duration and onset time. *C:* both sessions followed the same five-phase procedure, a Maximum Force Trial was completed three times during a session: at the beginning of the session, between the experiment's Familiarization Phase and Choice Phase, and at the end of the session. The Familiarization Phase consisted of 56 force trials, and the Choice Phase consisted of 455 choice trials with 45 force realization trials randomly dispersed. *D:* data from an example choice trial that shows how deliberation time and response vigor are calculated. *Left*: the cursor position starts in the middle of the screen (640 pixels) and then deviates to the right or left, depending on which force profile is preferred. In this example, the subject chooses the right option, moving into the righthand box with a border 100 pixels from the center (740 pixels). *Right*: deliberation time ends when the cursor velocity exceeds a 200 px/sec threshold. Response vigor is calculated as the peak cursor velocity as it moves toward the chosen force profile.

4.3.3 Experimental procedure

4.3.3.1 Effort task design

Each subject completed two sessions (Hills and Valleys) that followed an identical procedure (Figure 9C). Each session occurred on different days roughly one week apart (mean ± s.d.: 8.32 ± 5.91 days; min: 1 day; max: 22 days). The order of sessions was randomized and counterbalanced across subjects.

Each session tested a unique set of force profiles (Figure 9B) which subjects matched using isometric contractions of the shoulder and elbow. The Hills Session tested profiles with brief periods of high effort (i.e., hills) with a low baseline effort required. The Valleys Session tested profiles with a high baseline physical effort and brief periods of low physical effort (i.e., valleys). Hills Sessions' base level of effort was 10% of the subject's maximum force, with brief hills at 50% of the subject's maximum force. Valleys Sessions' base level of physical effort was 25% of the subject's maximum force, with brief valleys at 5% of the subject's maximum force.

For both sessions, we varied the duration of the hill or valley, as well as the time of onset. All force profiles were the same 17 seconds in duration. The hill or valley was one, three, or five seconds in duration and began between 2.5 and 12.5 seconds after the start of the trial. For one-second hills and valleys, we tested six different onset times: 2.5, 4.5, 6.5, 8.5, 10.5 and 12.5 seconds; for three-second hills and valleys, four different onset times: 2.5, 4.5, 8.5, 10.5; and for five-second hills and valleys, four different onset times: 2.5, 4.5, 6.5, 8.5. Not all combinations were possible (e.g., starting a five-second hill at 12.5 seconds exceeds the 17-second

duration), resulting in 14 possible combinations of hill/valley onset times and durations per session. A few example profiles for the hill and valley sessions are shown in Figure 9B which visualizes the extremes of duration and onset time.

With this approach, we independently manipulated the timing of two types of costs: *physical effort costs* and *task costs*. Because this is an isometric task, no true mechanical work is performed, so for a given trial, we represented *physical effort costs* as the force-time integral that has been shown to roughly linearly correlate with energy expenditure (Kushmerick & Paul, 1976; Russ et al., 2002). Hills and valleys were similar in their magnitude of our model of physical effort: trials with three-second hills and three-second valleys had equal force-time integrals. A profile with a one-second hill had a slightly smaller force-time integral than a five-second valley, and a five-second hill had a slightly larger force-time integral than a one-second valley. We defined *task costs* as a change in force level, where potential additional resources are required to accurately trace the line during these changes in force level (increasing force to climb a hill or out of a valley, or reducing force to descend from a hill or into a valley). Specific to our paradigm, this is modeled as the squared rate of force. In the framework of Rosenbaum et al.'s bucket task, a hill would look like picking up a bucket, walking, then setting it back down, where the task costs are incurred during both picking up and setting down the bucket (Rosenbaum et al., 2014). In our paradigm, valleys incurred lower magnitude task costs than hills, because the rate of change of force is lower, though both profiles require changes over similar time windows (0.5 seconds to ramp up or down). Inspecting the force profiles

(Figure 9B), a majority of each profile was a flat, constant force level (i.e., no task cost); however, at the beginning and end of a hill or valley, there was a sloped line and high task cost. With hills, periods of high physical effort and high task cost roughly coincided, making it difficult to tease out which costs the subject may value. However, in valley trials, periods of high task cost roughly coincided with low physical effort. If we only tested one type of session, we could not separate these costs. For instance, preferring early hills could be a preference for early high task costs or early high physical effort costs. However, preferring early hills and late valleys shows a clear preference for early high physical effort, because the preference for task costs is early for hills and late for valleys. With this structure, we could dissociate when the two types of costs occur and better tease apart the temporal sensitivity of each type of cost.

4.3.3.2 Maximum force trial

In each session, subjects were asked to complete three *Maximum Force Trials* throughout the experiment: once at the beginning, once between the *Familiarization Phase* and *Choice Phase,* and once at the conclusion of the experiment. For this task, subjects were asked to make three attempts to push as hard as possible anteriorly against the handle. On the monitor, subjects saw a red arrow grow in height as they pushed harder into the handle, and a white line marked the arrow's maximum height for the trial (i.e., the maximum force exerted) across all attempts. The maximum force exerted was used to normalize the force experienced across subjects and to assess objective fatigue throughout the experiment. Following the subject's maximum

exertion, they were asked to rate their subjective fatigue on a scale from one to five, where one is "no fatigue" and five is "unable to continue." Measuring fatigue was essential to ensure that subjects did not fatigue to exhaustion, which could skew decision-making ability.

4.3.3.3 Familiarization phase

Following the first *Maximum Force Trial*, each session began with the *Familiarization Phase*. Subjects experienced each of the 14 force profiles in a session four times (a total of 56 *Force Trials*). Each force profile type was presented in random order, where it was performed three times in a row. After each force profile type was completed three times, subjects again saw each of the 14 profiles once more in randomized order. Before the final set of 14 trials, subjects were reminded to consider which profile types they preferred. Subjects were given a 10-second rest period every 10 trials. Each trial displayed the force profile as a white line, which subjects were instructed to trace with the cursor. Subjects were given a countdown from three down to zero, displayed numerically, and reinforced with audio tones. During the countdown, the yellow cursor remained locked in place at the lefthand side of the force profile. Following the countdown, the yellow cursor started to move horizontally from left to right at a fixed rate. Subjects controlled the cursor's vertical position by exerting force isometrically against the fixed handle, where a higher anterior force resulted in a higher vertical cursor position. The force trial concluded when the subject's cursor reached the righthand side of the profile, when the new profile would appear, and the countdown began again. After each trial, the subject's performance

was evaluated by taking the average difference between vertical cursor position and the expected cursor position in units of percent maximum force. If subjects completed a profile with greater than 7% average error, they were required to repeat it immediately following their prior attempt, however, no subjects exceeded 7% error on any trial. Following the *Familiarization Phase*, subjects completed their second *Maximum Force Trial.*

4.3.3.4 Choice phase

After the second *Maximum Force Trial*, subjects completed the *Choice Phase*, consisting primarily of *Choice Trials* where subjects performed a two-alternative forced choice task. Unique pairs of the 14 force profiles were presented in a right-left orientation, and subjects were asked to choose which profile they preferred. Each possible combination of profiles for a session (91 unique combinations) was presented five times in random order, totaling 455 choices per session. For a given combination, the order of the profiles (either right or left sides) was randomized. For 10% of choices made (45 per session, randomly dispersed), subjects were required to trace the force profile they selected (*Realization Trials*), reinforcing that their choices had consequences and should be taken seriously. Subjects were given 10-second rest periods every 40 choices.

Each *Choice Trial* began with the subject's cursor in the middle of the screen with the two options displayed simultaneously with the text prompt "Which do you prefer? Move the cursor to choose." Deliberation time was unconstrained; subjects were neither forced nor encouraged to choose quickly. Subjects chose which profile

type they preferred by exerting a right or left force on the handle moving the cursor to their preferred profile. The horizontal cursor position had the same dynamics as tracing the force profiles. Cursor position away from the middle of the screen was directly proportional to the force they exerted on the handle, normalized to the subject's maximum force exerted anteriorly. Once the cursor crossed a border of the profile they preferred (±100 px from the center, 12.1% maximum force exerted), an audio cue was given, the profiles disappeared, and a circular target in the center of the screen reappeared with the text "Move back to the center." After they chose, the cursor was required to be returned to the center of the screen (i.e., no force exerted on the handle) before the next pair of profiles to choose between would appear.

Following a choice trial, 10% of the time, the choice the subject made would need to be performed (*Realization Trial). Realization Trials* were identical to a *Force Trial* in the *Familiarization Phase*, with a countdown, tracing task, and accuracy enforcement. If subjects performed the task with more than 7% average error, they were required to repeat it. Again, no subjects exceeded 7% error on any *Force Realization* trial. Following the *Choice Phase,* participants completed a third and final *Maximum Force Trial*.

4.3.4 Data collection and processing

A six-axis force sensor embedded in the robot handle measured forces at 200 Hz for all phases of the experiment. Axes of the force sensor were aligned to the robot arm. The angle and orientation of the robot arm were measured and used to rotate forces exerted on the handle to match the orientation of the subject: forces generated

in the anterior-posterior direction of the subject moved the cursor vertically, and forces generated to the right or left of the subject moved the cursor right or left on the screen. The handle was secured in a fixed position, so forces were exerted isometrically.

Measured forces were smoothed using a three-point average and used to determine cursor position. Cursor position was recalculated at approximately 10 Hz, and the display refreshed at approximately 60 Hz. Cursor position was linearly mapped to the smoothed forces, scaled to screen size, and normalized by the maximum anterior force exerted for each subject measured during each *Maximum Force Trial,* where 100% of the subject's maximum force equated to 840 px of deviation from the starting/zero-force point.

All data collected were analyzed and processed using custom scripts and built-in functions in MATLAB.

4.3.4.1 Maximum force trials

Only the peak force exerted anteriorly from the subject was used during the *Maximum Force Trial*. Maximum force exertions were hand-recorded from the voltage value displayed on the screen. Subjective fatigue ratings were solicited verbally and hand-recorded by the experimenter. These measures were collected and analyzed by experiment phase to investigate if subjects fatigued.

4.3.4.2 Force trials and realization trials

The main measure for *Force Trials* and *Realization Trials* was the accuracy with which the profiles were performed. When subjects traced profiles, forces exerted

anteriorly were used to control the cursor's vertical position, so subjects could visually compare cursor position to the displayed force profile (white lines) in real-time. The difference between the force profile and force exerted was used to calculate the average absolute error, scaled to the proportion of maximum force (82 pixels was approximately 10% of each subject's maximum force). Error calculations used cursor position data calculated and written at approximately 10 Hz. When subjects completed tracing a profile, the calculated error was used to determine if the subjects were sufficiently accurate. Any profile completed with more than 7% average error was required to be repeated. The accuracy of each subject's performance during the *Familiarization Task* was averaged across each unique profile type (four trials each). This metric was used to investigate whether accuracy was a predictor for profile preference, as measured in the *Choice Phase*. In addition, the accuracy of each subject's *Realization Trials* was analyzed to investigate any fatigue throughout the *Choice Phase* as well as look at performance differences within a profile.

4.3.4.3 Choice trials

During each *Choice Trial,* we measured the choice made (preferred profile), deliberation time, and choice vigor. These were calculated from the recorded cursor position, forces exerted, and time elapsed for each trial. The cursor position, resulting from forces exerted, was used to determine preference. Once the cursor crossed the border of the preferred profile (±100 px from the center, 12.1% maximum force exerted), the corresponding profile (right or left) was deemed their preferred option. In some infrequent cases, the robot computer would skip to the subsequent trial

without letting the subject choose (0.25% of trials). These choices were identified, discarded, and not included in any analyses. Choice behavior measures were determined from the rate of force, which corresponds to vertical cursor velocity (Figure 9D). This was calculated post hoc by regenerating the raw cursor position from measured forces, filtering using a fourth-order Butterworth filter, and then differentiating by time. Deliberation time was calculated as the time of choice appearance until the cursor velocity first exceeded 200 px/sec. Choice vigor was calculated as the maximum cursor speed, which was equivalent to the maximum rate of force normalized to a subject's maximum force exerted. For the choice behavior analysis, outliers that exceeded the 2.5^{th} and 97.5^{th} percentiles for a given subject were removed. For deliberation time, 3.36% of total choices were not considered in the analysis, and for choice vigor, 2.93% of total choices.

4.3.5 Statistical analysis

Our primary methods of analysis were through linear and non-linear models to describe observed behaviors. Below, we detail the structure and measures for each model. Unless specified, results for all models are reported with 95% confidence intervals for parameter estimates.

4.3.5.1 A linear model of profile preference

Our primary variable of interest was the subject's choice on each trial, as reported by the onscreen cursor's movement into the preferred option. We first investigated how subjects' preferences were determined by the variables we intentionally manipulated: duration, onset time, and session type. We coded choices

as a binomial dataset, *ChoseLeft*, where 1 indicates the left option was chosen, and 0, the right. We employed a logit-linked binomial generalized linear mixed-effects model to investigate the fixed covariates and interactions of session type (*SessionType*, categorical with two levels: hill or valley), the difference in hill or valley duration (*ΔDuration*, continuous, seconds), the difference in hill or valley onset time (*ΔOnset,* continuous, seconds), and *Subject* (categorical) as a random intercept term. The model is summarized in Equation 10 below, where $ChoseLeft_{ij}$ is the *j*th choice of subject *i*, and $Subject_i$ is normally distributed with mean 0 and variance σ^2.

$$\begin{aligned} \text{ChoseLeft}_{ij} &\sim \text{Binomial}(\theta_{ij}) \\ \text{E}[\text{ChoseLeft}_{ij}] &= \theta_{ij} \\ \text{logit}(\theta_{ij}) = \text{SessionType}_{ij} &+ \Delta\text{Duration}_{ij} + \Delta\text{Onset}_{ij} \\ &+ \text{SessionType}_{ij} \times \Delta\text{Duration}_{ij} \\ &+ \text{SessionType}_{ij} \times \Delta\text{Onset}_{ij} \\ &+ \Delta\text{Duration}_{ij} \times \Delta\text{Onset}_{ij} \\ &+ \text{SessionType}_{ij} \times \Delta\text{Duration}_{ij} \times \Delta\text{Onset}_{ij} \\ &+ \text{Subject}_i \\ \text{Subject}_i &= N(0, \sigma^2) \end{aligned} \tag{10}$$

4.3.5.2 Profile preferences using models of subjective utility

We used a more complex but more generalizable model for subject preference. Here, choices were modeled as a function of the difference in subjective utility between profile types, where subjective utility was determined from the actual force profile versus time, instead of protocol-specific factors like duration or onset time of a single hill or valley. We compared four different models of subjective utility that incorporated different specific costs (physical effort costs, task costs, and time sensitivity) and their interactions. With this approach, we could infer more specifically what costs were driving subjects' preferences.

For all models of subjective utility, we quantified the decision-making process as a function of the difference between the subjective utility of each option. The probability of choosing the left profile was structured as a sigmoidal function of the difference between the subjective utility for each choice, where subjective utility was represented by U for option *left* or *right,* and a subject-specific temperature parameter, τ:

$$p(choose\ left|\ U_{left}, U_{right}) = \left(1 + e^{-\tau(U_{left} - U_{right})}\right)^{-1} \quad (11)$$

We calculated utility as reward, R, minus costs, E. In this specific experiment, there was no explicit reward thus no difference in reward across profiles. Our model assumes there was no intrinsic reward for performing the effortful task. Mathematically, we modeled reward to be zero resulting in a utility that was simply the negative of the subjective effort, E:

$$U_{subjective} = R - E = -E \quad (12)$$

We distinguished between two types of costs: *physical effort* and *task*. Physical effort costs for a given profile were represented as force normalized to a subject's maximum and the force-time integral of a one-second hill profile, $K_{physical}$ (Equation 13). Task costs were represented by the rate of force squared normalized to a subject's maximum, and the force-time integral of a one-second hill profile, K_{task} (Equation 14). This normalization was important so both types of effort were the same order of magnitude. Examples of physical effort costs and task costs in different profile types are shown in Figure 10.

$$E_{physical}(t) = \frac{F(t)}{K_{physical}}, \tag{13}$$

$$E_{task}(t) = \frac{\dot{F}^2(t)}{K_{task}} \tag{14}$$

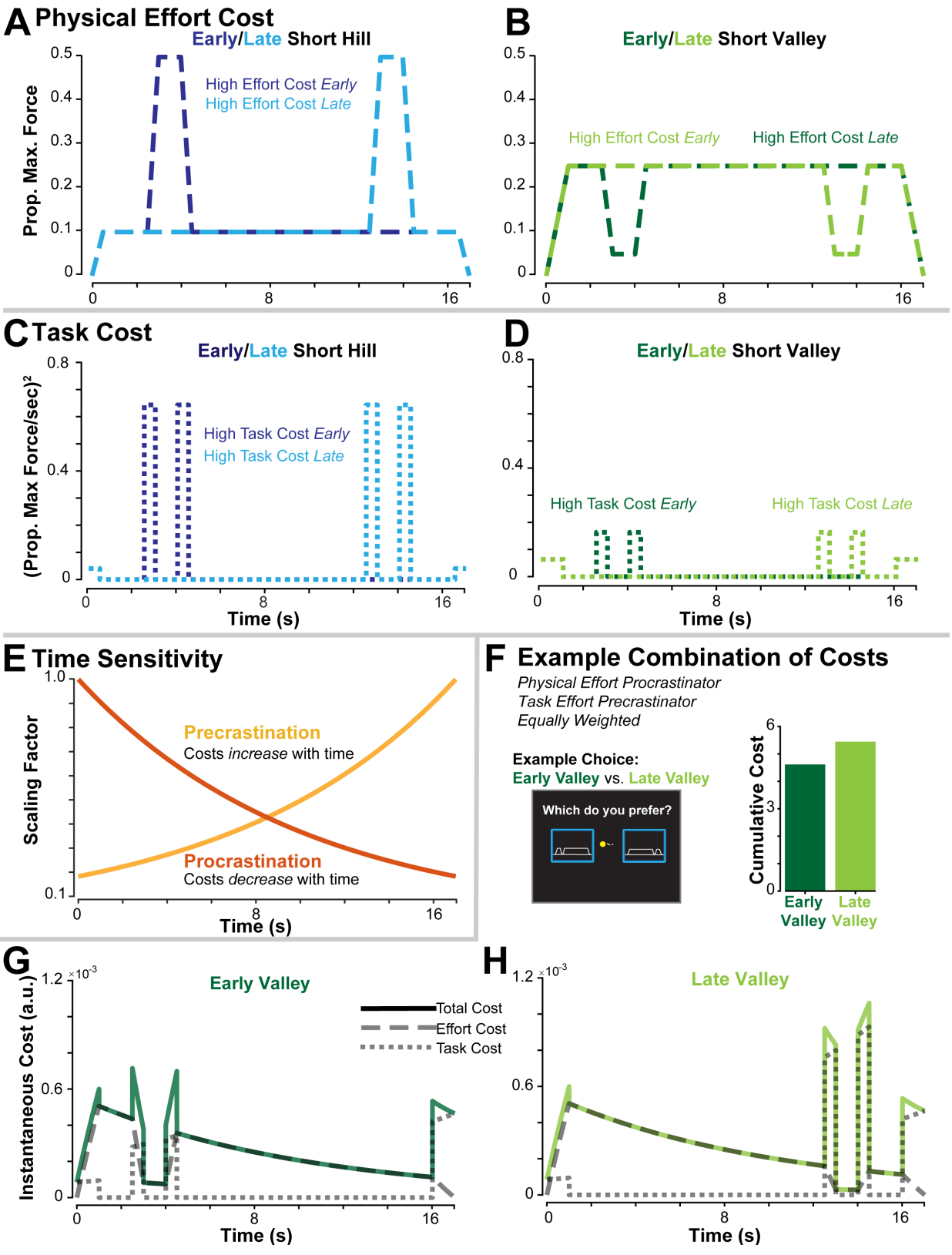
A Physical Effort Cost
Early/Late Short Hill
High Effort Cost Early
High Effort Cost Late
Prop. Max. Force
B
Early/Late Short Valley
High Effort Cost Early
High Effort Cost Late
C Task Cost
Early/Late Short Hill
High Task Cost Early
High Task Cost Late
(Prop. Max Force/sec)²
Time (s)
D
Early/Late Short Valley
High Task Cost Early
High Task Cost Late
Time (s)
E Time Sensitivity
Scaling Factor
Precrastination
Costs increase with time
Procrastination
Costs decrease with time
Time (s)
F Example Combination of Costs
Physical Effort Procrastinator
Task Effort Precrastinator
Equally Weighted
Example Choice:
Early Valley vs. Late Valley
Which do you prefer?
Cumulative Cost
Early Valley
Late Valley
G
Early Valley
Total Cost
Effort Cost
Task Cost
Instantaneous Cost (a.u.)
Time (s)
H
Late Valley
Time (s)

Figure 10. Models of subjective costs. How subjective effort costs were calculated from the force profiles is shown, including an example choice. Physical effort cost (dashed lines) was represented as the force-time integral. *A:* an early, short hill (dark blue) had high physical effort cost early and low physical effort cost late, and a late, short hill (light blue) had high physical effort cost late and low physical effort cost early. *B:* an early, short valley (dark green) had low physical effort cost early and high physical effort cost late, and a late, short valley (light green) had high physical effort cost early and low physical effort cost late. Task costs (dotted lines), represented as rate of change of force squared, corresponded to when the hill or valley occur. *C:* for both early short hills (dark blue) and *D:* early short valleys (dark green), high task costs occurred early, while late short hills (light blue) and late short valleys both had high tasks costs late. *E: t*ime sensitivity was represented with two possible categories: precrastination (yellow) and procrastination (orange). This cost function multiplied with physical effort costs or task costs to combine into a time sensitivity for when those costs occurred. For precrastination, costs increased with time, and for procrastinators, costs decreased with time. Using this structure to represent effort costs, they combined per the various utility functions (Equations 16 through 21). Using the sixth model of utility (Equation 21), an example of how the costs combine is illustrated. In this case, we show how someone who was an effort cost procrastinator, task cost precrastinator, with equal weights on types of costs. *G:* how that subject might have experienced an early valley. *H:* how that subject might have experienced a late valley. Total costs are shown in bold color with the effort costs shown as a dashed line, and task costs, a dotted line. Each profile's cost was integrated, and whichever profile has the lower cumulative cost was the preferred option. *F:* in this case, we see the early valley (dark green) has lower cumulative cost and was the preferred option.

We modeled the cost of time as a function that could be multiplied with effort costs to exhibit temporal sensitivity for those costs (Figure 10E). Previous studies model temporal sensitivity as an exponential (Rigoux & Guigon, 2012) or hyperbolic discounting function (A. M. Haith et al., 2012). For this study, we used a piecewise exponential formulation (Equation 15), where γ is the sensitivity factor, t is instantaneous time, and T is the end time (in our case, always 17 seconds). A positive γ value is a cost procrastinator, where costs further in the future incur a smaller penalty than costs closer to the present. Alternatively, a negative γ value is a cost precrastinator, where costs in the future are a larger consequence than costs expended in the near term. Using the piecewise function structure allows for the effect of positive and negative γ-values to be equally distributed from zero.

$$s(t,\gamma) = \begin{cases} e^{-\gamma t}, \gamma < 0 \\ e^{\gamma(T-t)}, \gamma \geq 0 \end{cases} \quad (15)$$

We explored six different subjective utility models, where each was a different combination of costs from Equations 13, 14, and 15. First, we considered a traditional model of effort: not time-sensitive and simply the force-time integral of the profile (Equation 16). This model of subjective utility, U_1, could predict a preference for shorter hills and longer valleys but indifference as to when those hills and valleys occur.

$$U_1 = -\int_0^T E_{physical}(t)dt \tag{16}$$

The second model of subjective utility incorporated time sensitivity and physical effort (Equation 17). This model, U_2, could predict subjects who prefer when periods of physical effort occur. For example, this model could describe subjects who preferred early hills and late valleys and would classify them as physical effort precrastinators.

$$U_2 = -\int_0^T E_{physical}(t) \cdot s\left(t, \gamma_{physical}\right)dt \tag{17}$$

We also considered an alternative model that included only time-sensitive task costs (Equation 18). This model was agnostic to the force-time integral, and instead, cared about when a hill or valley starts and ends. This model could describe a preference for shorter early valleys over a long valley because the task is completed earlier.

$$U_3 = -\int_0^T E_{task}(t) \cdot s(t, \gamma_{task})dt \tag{18}$$

Building from the previous model, an alternative, U_4, included a sensitivity to physical effort that is time-neutral in addition to time-sensitive task costs (Equation

19). A weighting parameter, α, varied between 0 and 1, that shifted the importance of physical effort versus task costs. This model could predict a preference for shorter hills to longer hills and temporal sensitivity driven by task onset and completion time.

$$U_4 = -\int_0^T \left(\alpha E_{physical}(t) + (1-\alpha)E_{task}(t) \cdot s(t, \gamma_{task})\right) dt \quad (19)$$

Next, we considered a fifth model, U_5, that incorporated task costs, physical effort costs, and time sensitivity, $\gamma_{combined}$, (Equation 20). We constrained this model so that both were temporally sensitive in the same way. This model could describe subjects who preferred early hills and early valleys, classifying them as precrastinators of both physical effort and task costs with a higher weighting on task costs.

$$U_5 = -\int_0^T \left(\alpha E_{physical}(t) + (1-\alpha)E_{task}(t)\right) s(t, \gamma_{combined})\, dt \quad (20)$$

Finally, we considered a model of utility, U_6, that included physical effort and task costs with independent temporal-sensitivity factors, $\gamma_{physical}$ and γ_{task} (Equation 21). This model offered an alternative explanation to U_5 for a subject who prefers early hills and early valleys: they could be task cost precrastinators and physical effort procrastinators, but care more about task costs.

$$U_6 = -\int_0^T \left(\alpha E_{physical} \cdot s\left(t, \gamma_{physical}\right) + (1-\alpha)E_{task}(t) \cdot s(t, \gamma_{task})\right) dt \quad (21)$$

4.3.5.3 Predicting subject preferences with utility functions

The models summarized above are nonlinear, thus we employed MATLAB's constrained minimization function, *fmincon*, that is designed for nonlinear optimization problems. For each subject, we minimized an objective function

(Equation 22), which is the negative log-likelihood of the probabilistic model's prediction for each subject's choices (Equation 11). For the total number of valid trials, N, this function compares if the subject chooses the left option (binary outcome variable, y) to the model's predicted choice, p, for a given profile combination. For a given utility model, each subject was fit independently. Parameters were constrained but allowed to vary within set bounds. Temporal sensitivity factors, $\gamma_{physical}$, γ_{task}, and $\gamma_{combined}$, were constrained to [-1.5,1.5]; the weighting parameter, a, [0,1], and the temperature parameter, τ, [0,10^6]. Temporal sensitivity boundary limits were determined through exploratory fits, where the bounds allowed for the smallest search space that did not include a solution that was at the boundary limits. The minimizing solution was obtained for each fit by comparing results from multiple restarts with randomized initial parameter values. The solution with the smallest objective function (i.e., the greatest negative log likelihood) was chosen. Parameters for each subject and model were then collated and compared, and the aggregate model performance across subjects was analyzed to see if a single model best described subject choice data.

$$\text{Min } -\sum_{i=1}^{N} y_i \log\left(p(choose\ left \mid U_{left,i}, U_{right,i})\right) + (1-y_i)\log\left(1 - p(choose\ left \mid U_{left,i}, U_{right,i})\right) \quad (22)$$

4.3.5.4 Utility model selection and estimating population parameters

The most likely model was determined using Bayesian model selection (BMS) (Rigoux et al., 2014; Stephan et al., 2009). Using the calculated log-likelihood for each subject-model pair, the Akaike Information Criteria (AIC) was calculated. Model

posterior means and protected exceedance probabilities (pxp) were computed using AICs with the Statistical Parametric Mapping software (SPM, version 12). Protected exceedance probability is the probability that a given model is more likely, tested against the null hypothesis that all models are equally likely. The model with the highest pxp was selected as the winning model.

For each model, we estimated population distribution for each parameter by bootstrapping each set of fit parameters 10,000 times with replacement. From these confidence intervals, we determined how the population may be weighting the various costs of physical effort costs, task costs, and time.

4.3.5.5 Choice behavior and preference

Both deliberation time and choice vigor were fit using linear mixed-effects models. Using the winning model of utility, the absolute difference in subjective utility was min-max normalized for each subject using Equation 23. The normalized difference in subjective utility was used as the independent variable to explain the dependent variables, deliberation time and choice vigor.

$$|\Delta U|_{normed} = \frac{|\Delta U| - \min\ |\Delta U|}{\max\ |\Delta U| - \min\ |\Delta U|} \tag{23}$$

Deliberation time was modeled using a simple linear mixed-effects model with a Gaussian response variable (Equation 24). We investigated the fixed effect of the normalized difference in subjective utility ($|\Delta U|_{normed}$, continuous from 0 to 1), a fixed intercept term, and a random effect of subject (*Subject*, categorical) on the intercept. In the model, $DelibTime_{ij}$ is the deliberation time of the jth choice of subject i, and $Subject_i$ is normally distributed with mean 0 and variance σ^2.

$$\begin{gathered}\text{DelibTime}_{ij} \sim N(\mu_{ij}, \sigma) \\ E(\text{DelibTime}_{ij}) = \mu_{ij} \\ \mu_{ij} = |\Delta U|_{normed,ij} + \text{Subject}_i \\ \text{Subject}_i \sim N(0, \sigma^2)\end{gathered} \tag{24}$$

We investigated choice vigor using an identically structured linear mixed-effects model with a Gaussian response variable (Equation 25). Again, we included the fixed effect of the normalized difference in subjective utility ($|\Delta U|_{normed}$, continuous from 0 to 1), a fixed intercept term, and a random effect of subject (*Subject*, categorical) on the intercept. In the model, $ChoiceVigor_{ij}$ is the vigor of the jth choice of subject i, and $Subject_i$ is normally distributed with mean 0 and variance σ^2.

$$\begin{gathered}\text{ChoiceVigor}_{ij} \sim N(\mu_{ij}, \sigma) \\ E(\text{ChoiceVigor}_{ij}) = \mu_{ij} \\ \mu_{ij} = |\Delta U|_{normed,ij} + \text{Subject}_{ij} \\ \text{Subject}_{ij} \sim N(0, \sigma^2)\end{gathered} \tag{25}$$

Finally, we tested if deliberation time could be predicted by choice vigor using a linear mixed-effects model (Equation 26). We modeled deliberation time (*DelibTime*, continuous) as a Gaussian response variable, a fixed effect of choice vigor (*ChoiceVigor,* continuous), a fixed intercept term, and a random effect of subject (*Subject,* continuous) on the intercept.

$$\begin{gathered}\text{DelibTime}_{ij} \sim N(\mu_{ij}, \sigma) \\ E(\text{DelibTime}_{ij}) = \mu_{ij} \\ \mu_{ij} = \text{ChoiceVigor}_{ij} + \text{Subject}_{ij} \\ \text{Subject}_{ij} \sim N(0, \sigma^2)\end{gathered} \tag{26}$$

4.3.5.6 Force trial performance and a predictor of preference

We investigated if subject accuracy when performing a profile predicted their preferred profile. Because not all profiles were performed during the *Realization*

Trials in the *Choice Phase,* we used the average error for each profile type from the *Force Trials* completed in the *Familiarization Phase*. Using the binomial choice data, *ChoseLeft*, we used the same logit-linked binomial generalized linear mixed-effects model as Equation 10 but added a fixed effect of difference in error between profile types (*ΔError*, continuous).

In addition, we investigated if there was a correlation between error and the force profile characteristics (onset, duration, session type, and their interactions) using a linear mixed-effects model with a random intercept term for subject (Equation 27). This model used the set of all possible choice combinations from the *Familiarization Phase.* We investigated the fixed covariates and interactions of session type (*SessionType*, categorical with two levels: hill or valley), hill or valley duration (*Duration*, continuous, secs), hill or valley onset time (*Onset,* continuous, secs), and *Subject* (categorical) as a random intercept term. The model is summarized in Equation 27 below, where $Error_{ij}$ is the jth choice combination of profile types for subject i, and $Subject_i$ is normally distributed with mean 0 and variance σ^2.

$$\begin{gathered}
\mathrm{Error}_{ij} \sim N(\mu_{ij,}\sigma) \\
E[\mathrm{Error}_{ij}] = \mu_{ij} \\
\mu_{ij} = \mathrm{SessionType}_{ij} + \mathrm{Duration}_{ij} + \mathrm{Onset}_{ij} + \mathrm{SessionType}_{ij} \times \mathrm{Duration}_{ij} \\
+ \mathrm{SessionType}_{ij} \times \mathrm{Onset}_{ij} + \mathrm{Duration}_{ij} \times \mathrm{Onset}_{ij} \\
+ \mathrm{SessionType}_{ij} \times \mathrm{Duration}_{ij} \times \mathrm{Onset}_{ij} + \mathrm{Subject}_i \\
\mathrm{Subject}_i = N(0, \sigma^2)
\end{gathered} \tag{27}$$

Next, we investigated how subjects completed *Realization Trials* during the *Choice Phase*. We analyzed the average force before, during, and after a hill or valley. Given the differences in demanded force levels of hills and valleys (Figure 9A and Figure 9B), there may have been acute fatigue within a trial that could have led to a

temporal preference of high effort. We ran a linear mixed effects model, summarized in Equation 28, below. For each force measure, we investigated the fixed covariates and interactions of session type (*SessionType*, categorical with two levels: hill or valley), hill or valley duration (*Duration*, continuous, secs), hill or valley onset time (*Onset,* continuous, secs), trial number (*TrialNumber),* and *Subject* (categorical) as a random intercept term.

$$
\begin{aligned}
\text{AvgForce}_{ij} &\sim N(\mu_{ij}, \sigma) \\
\text{E}[\text{AvgForce}_{ij}] &= \mu_{ij} \\
\mu_{ij} = {} & \text{SessionType}_{ij} + \text{Duration}_{ij} + \text{Onset}_{ij} + \text{TrialNumber}_{ij} \\
& + \text{SessionType}_{ij} \times \text{Duration}_{ij} + \text{SessionType}_{ij} \times \text{Onset}_{ij} \\
& + \text{SessionType}_{ij} \times \text{TrialNumber}_{ij} + \text{Duration}_{ij} \times \text{Onset}_{ij} \\
& + \text{Duration}_{ij} \times \text{TrialNumber}_{ij} + \text{Duration}_{ij} \times \text{TrialNumber}_{ij} \\
& + \text{SessionType}_{ij} \times \text{Duration}_{ij} \times \text{Onset}_{ij} \\
& + \text{SessionType}_{ij} \times \text{Duration}_{ij} \times \text{TrialNumber}_{ij} \\
& + \text{SessionType}_{ij} \times \text{Onset}_{ij} \times \text{TrialNumber}_{ij} \\
& + \text{Duration}_{ij} \times \text{Onset}_{ij} \times \text{TrialNumber}_{ij} \\
& + \text{SessionType}_{ij} \times \text{Duration}_{ij} \times \text{Onset}_{ij} \times \text{TrialNumber}_{ij} \\
& + \text{Subject}_{i} \\
\text{Subject}_{i} &= N(0, \sigma^2)
\end{aligned}
\tag{28}
$$

4.3.5.7 Subject fatigue by experiment phase

We used the measured maximum force exerted and subjective fatigue rating collected during the *Maximum Force Trial* to determine if subjects physically fatigued throughout the experiment. We investigated how the two measures of fatigue change across the three measurements (*ExpPhase,* ordinal, three values) and session type (categorical, hill or valley). Each measure used an identical linear mixed-effects model shown in Equation 29, where $Fatigue_{ij}$ is jth observation either the maximum force exerted or the subjective fatigue rating for subject i.

$$\begin{gathered}\text{Fatigue}_{ij} \sim N(\mu_{ij}, \sigma) \\ E[\text{Fatigue}_{ij}] = \mu_{ij} \\ \mu_{ij} = \text{SessionType}_{ij} + \text{ExpPhase}_{ij} + \text{Subject}_i \\ \text{Subject}_i = N(0, \sigma^2)\end{gathered} \tag{29}$$

Next, we looked at secondary measures that may indicate subjects were fatiguing, physically or cognitively, through the *Choice Phase*. We investigated a possible effect of trial number of the choice behavior (choice vigor and deliberation time), the average error of *Realization Trials* performed during the *Choice Phase*, and if preferences changed between the first half of the *Choice Phase* versus the second half. Models for vigor, deliberation time, and average error used a simple linear mixed-effects model with a Gaussian response variable with a fixed effect of trial number (*TrialNumber*), a fixed intercept term, and a random effect of subject (*Subject*, categorical) on the intercept. Equation 30 shows an example of the model used, specifically for deliberation time. $DelibTime_{ij}$ is the deliberation time of the jth choice of subject i, and $Subject_i$ is normally distributed with mean 0 and variance σ^2. This same variable type and model structure was used for investigating choice vigor ($ChoiceVigor_{ij}$) and realization trial error ($Error_{ij}$).

$$\begin{gathered}\text{DelibTime}_{ij} \sim N(\mu_{ij}, \sigma) \\ E(\text{DelibTime}_{ij}) = \mu_{ij} \\ \mu_{ij} = \text{TrialNumber}_{ij} + \text{Subject}_{ij} \\ \text{Subject}_{ij} \sim N(0, \sigma^2)\end{gathered} \tag{30}$$

We built upon the linear model for preference shown in Equation 10 to investigate if there was a change of preferences over the *Choice Phase*. Shown in Equation 31, we added a binary variable, *HalfSession,* to describe whether the trial occurred in the first half of a session or second half of a session (trials 1-277, and 278-

455, respectively) and included an interaction term between hill or valley sessions and trial number, in case one session type was more fatiguing than the other.

$$\begin{aligned}
&\text{ChoseLeft}_{ij} \sim \text{Binomial}(\theta_{ij}) \\
&\quad \text{E}[\text{ChoseLeft}_{ij}] = \theta_{ij} \\
\text{logit}(\theta_{ij}) = &\ \text{SessionType}_{ij} + \Delta\text{Duration}_{ij} + \Delta\text{Onset}_{ij} \\
&+ \text{SessionType}_{ij} \times \Delta\text{Duration}_{ij} \\
&+ \text{SessionType}_{ij} \times \Delta\text{Onset}_{ij} \\
&+ \Delta\text{Duration}_{ij} \times \Delta\text{Onset}_{ij} \\
&+ \text{SessionType}_{ij} \times \Delta\text{Duration}_{ij} \times \Delta\text{Onset}_{ij} \\
&+ \text{HalfSession}_{ij} + \text{HalfSession}_{ij} \times \text{SessionType}_{ij} \\
&+ \text{Subject}_i \\
&\quad \text{Subject}_i = N(0, \sigma^2)
\end{aligned} \tag{31}$$

4.4 Results

4.4.1 Profile preference

We first investigated underlying factors for subject preferences (Table 2) using a generalized linear mixed-effects model. Notably, we observed a significant effect of hill or valley duration (mean and 95% confidence interval: β = 0.905 [0.872, 0.938], P < 0.0001), where the positive coefficient indicates a preference for longer valleys (Figure 11, top right). There was also a significant interaction between duration and session type (β = -1.761 [-1.81, -1.72], P < 0.0001), with a reversal of sign demonstrating that preference was shifted towards shorter hills (Figure 11, top left). Interestingly, there was a significant preference for later valleys (β = 0.0448 [0.0338, 0.0558], P < 0.0001) and a significant interaction between onset time and session type (β = -0.0922 [-0.107, -0.0768], P <0.0001) that indicates a preference for earlier hills. However, the effect size of onset time is much smaller than that of duration. There was also a significant effect of the interaction between duration, onset time, and

session type (β = -0.0122 [-0.0217, -2.63x10^{-3}], P = 0.0124), which could suggest there was a stronger preference for later valleys than earlier hills (Figure 11, middle).

Table 2: Generalized linear mixed-effects model of subject preference.

Fixed Effect	**Estimate (mean, 95% CI)**	**P-value**
Intercept	-0.156 [-0.307, -0.00513]	0.0427
SessionType=Hill	0.152 [0.0783, 0.225]	<0.0001
ΔDuration	0.905 [0.872, 0.939]	<0.0001
ΔOnset	0.0448 [0.0338, 0.0558]	<0.0001
ΔDuration x SessionType=Hill	-1.76 [-1.81, -1.72]	<0.0001
ΔOnset x SessionType=Hill	-0.0922 [-0.107, -0.0768]	<0.0001
ΔDuration x ΔOnset	4.95x10^{-3} [-1.94x10^{-3}, 1.18x10^{-2}]	0.159
ΔDuration x ΔOnset x SessionType=Hill	-0.0122 [-0.0217, -2.63x10^{-3}]	0.0124

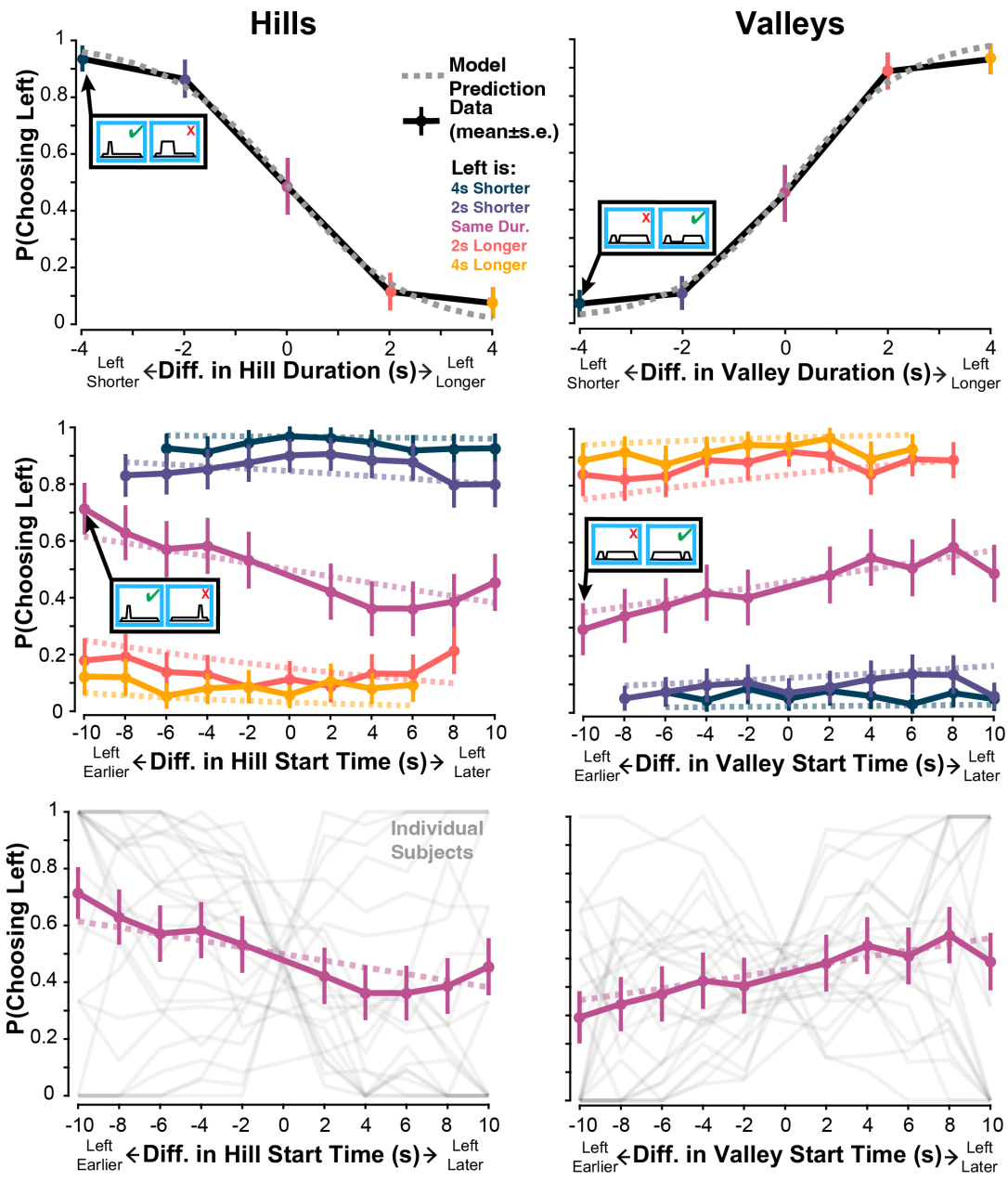

Figure 11: Subject choice behavior and linear model predictions. *Top*: The probability of choosing a hill (*left*) or valley (*right*), based on the difference in duration between options (*left-right*). Mean and s.e. choice data are shown with colored dots and bars connected by black lines. The dashed line is the predicted probability from the linear model. Across subjects, there was a strong preference for shorter hills and longer valleys. *Middle*: The probability of choosing a hill (*left*) or valley (*right*) depends on the difference in onset time. Choices are stratified by difference in duration, where blue is four seconds shorter, purple is two seconds shorter, pink is equal duration, orange is two seconds longer, and yellow is four seconds longer. Mean data are shown with dots connected by bold, colored lines, and model predictions are shown as dashed lines. The dominant effect on preference was duration, but when hill or valley duration was equal, there was a higher probability of choosing an earlier hill or a later valley. Bottom: The probability of choosing a hill or valley of equal duration, with individual subjects' preference curves plotted in gray.

These results make sense; subjects were sensitive to total duration of physical effort and preferred less effort to more effort. Interestingly, these results also show that subjects were sensitive to when effort was invested, not just for how long. For effort of similar duration, subjects tended to precrastinate and prefer high physical effort earlier than later (earlier hills and later valleys).

4.4.2 Preferences suggest time sensitivity of physical effort and task costs

Subjects' choices were fit using the various models of subjective utility (Equations 16-21). The four-parameter model of utility (Equation 21, U_6), which includes physical effort costs, task costs, and independent time sensitivity factors, best explained subject choices (Table 3). This model had the highest pxp of 0.995, as determined from the group-level BMS. Comparing AICs for each individual and model pair, this model (U_6) best described data from 12 of the 25 subjects compared with four subjects for U_5, four subjects for U_2, two for U_3, two for U_4, and one subjects for U_1.

Table 3: Posterior means, protected exceedance probabilities (pxp and bootstrapped parameter values for each model of subjective utility.

Model	# Param	Posterior Means	pxp	τ	$\gamma_{physical}$	γ_{task}	α
Physical Effort Only (U_1)	1	0.0376	$9.58\text{x}10^{-5}$	[9.86, 28.8]	N/A	N/A	N/A
Time-Sensitive Physical Effort (U_2)	2	0.130	$1.10\text{x}10^{-3}$	[8.23, 20.5]	[-0.0763, $-5.72\text{x}10^{-3}$]	N/A	N/A
Time-Sensitive Task Costs (U_3)	2	0.0580	$1.22\text{x}10^{-4}$	[$5.44\text{x}10^{3}$, $3.80\text{x}10^{4}$]	N/A	[-0.408, 0.0868]	N/A
Time-Neutral Physical Effort and Time-Sensitive Task Costs (U_4)	3	0.0416	$9.58\text{x}10^{-5}$	[995, $3.44\text{x}10^{3}$]	N/A	[-0.253, 0.0201]	[0.297, 0.648]
Time-Sensitive, Combined Physical Effort and Tasks Cost (U_5)	3	0.171	$3.93\text{x}10^{-3}$	[$7.54\text{x}10^{3}$, $1.43\text{x}10^{5}$]	[-0.0574, $-4.65\text{x}10^{-3}$]	Equal to $\gamma_{physical}$	[0.349, 0.660]
Independent Time-Sensitivity, Physical Effort and Tasks Cost (U_6)	4	0.562	0.995	[$1.70\text{x}10^{3}$, $1.15\ \text{x}10^{4}$]	[-0.0878, $-4.47\text{x}10^{-3}$]	[-0.172, 0.149]	[0.236, 0.580]

We investigated the results of the winning model Figure 12, where we saw average choice behavior matched model predicted choice behavior (Figure 12A). Because individual subjects were fit independently, we inspected individual fit performance. Here, we show a particular subject who was classified as a physical effort precrastinator and task cost precrastinator, with an emphasis on physical effort over task costs ($\gamma_{physical}$ = -0.0202; γ_{task} = -0.524; α = 0.999). In Figure 12B, we show the model accurately predicts the subject's choice behavior as well as a function of the difference in subjective utility. In Figure 12C, we sorted each profile by the model-predicted subjective utility to intuit what the subject cares about. For this particular subject, a physical effort precrastinator, we see that early, short hills were most preferable, and late, long hills were least preferable. Comparing valleys, we see that

long late valleys were preferable and short early valleys were least preferable. This ranking by utility matches the description of a physical effort precrastinator.

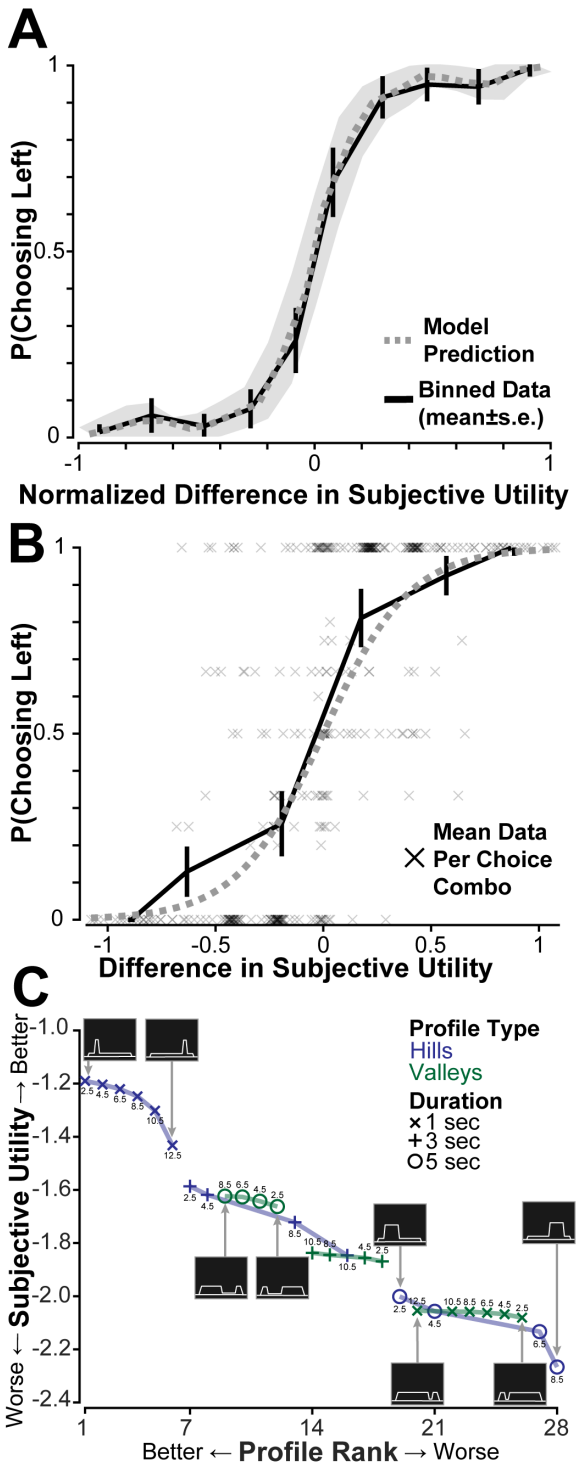

Figure 12: Utility model predictions. *A:* model predicted probability of choosing the left option, given the normalized difference in subjective utility (left-right) for the winning model averaged across subjects. Model predictions with 95% confidence intervals are shown in gray. Choice data were binned and averaged, with mean and standard error bars shown in black. Data and model predictions from a representative subject are shown in *B* and *C***.** This subject was characterized as a physical effort precrastinator ($\gamma_{physical}$ = -0.0202) and task cost precrastinator (γ_{task} = -0.524), with a strong emphasis on physical effort over task cost (α = 0.999) *B:* Model predicted probability of choosing the left option for the example subject is shown in gray, using subjective utility from the winning model. Mean probability of choosing left for a given combination (two unique profile types and which profile is on the left) is shown as a black x**.** Binned choice data was averaged, with the mean and standard deviation bars shown in black. *C:* Using the subjective utility, each profile was ranked against one another. Hill profiles are shown in blue, and Valley profiles are shown in green. The different durations of hills and valleys are shown using different markers per the legend, with the onset time indicated numerically. For this subject, the most preferable options were short, early hills, and the least preferable options were short, early valleys and late, long hills. This subject's pattern of preference was consistent, where, for a given duration of hill or valley, preference decreased as high physical effort occurred later.

4.4.3 Subjects precrastinate physical effort costs

To investigate parameter estimates for each utility model, we bootstrapped each set of parameter values with replacement 10,000 times to estimate the population distribution (Table 3). For time-sensitivity factors, a negative value indicated *precrastination,* whereas a positive value indicated *procrastination.* For the weighting term, a value of α = 0.5 equally valued *physical effort* and *task costs*, α = 1 valued only *physical effort costs*, and α = 0 valued only *task costs*.

The winning model, U_6, suggests that subjects precrastinate physical effort costs (CI: [-0.0878, -4.47×10^{-3}]), where high physical effort is preferred earlier rather than later. Interestingly, this trend was consistent across all models that included a time sensitivity of physical effort parameter. U_2 and U_5 also predicts significant precrastination of physical effort. Time sensitivity of task costs was nearly centered at zero (CI: [-0.172, 0.149]), showing no consistent trend across subjects. On an individual basis, subjects were time sensitive to task costs, hence the fact that U_6 was the preferred model and not U_2. Weighting between physical effort and task costs also

supported this (CI: [0.236, 0.580]), where subjects tended to value task costs equally, if not slightly more than physical effort alone. With time sensitivity of task costs being a significant factor, the lack of a trend in the confidence interval, and the weighting factor slightly favoring tasks cost, these results suggest that subjects valued task cost timing but were much more idiosyncratic about preferences across the group. In contrast, underlying these idiosyncratic preferences for task cost timing is a consistent tendency to precrastinate physical effort.

4.4.4 Choice behavior reflects subjective utility

One critique may be that the choice behavior is simply the result of a forced choice task, and subjects choose an arbitrary heuristic that is adhered to consistently. To further investigate this, we also characterized *how* subjects made and reported their choices. Specifically, we looked at the deliberation time and the response vigor for each choice and how these measures correlated to the difference in subjective utility. If subjects chose arbitrarily, choice behavior may not correlate to the subjective utility used to model subject preferences.

Deliberation times varied across subjects (mean ± s.e.: 0.629 ± 0.127), yet consistently decreased as the difference in utility increased. We observed a significant effect of the difference in normalized subjective utility, where a negative slope indicated a decreasing (or quicker) deliberation time as the difference in subjective utility increased (β = -0.206 [-0.231, -0.181], P <0.0001, Figure 13A). Our data reflected the same trends as previously observed: easier choices are made more quickly (Festinger, 1943).

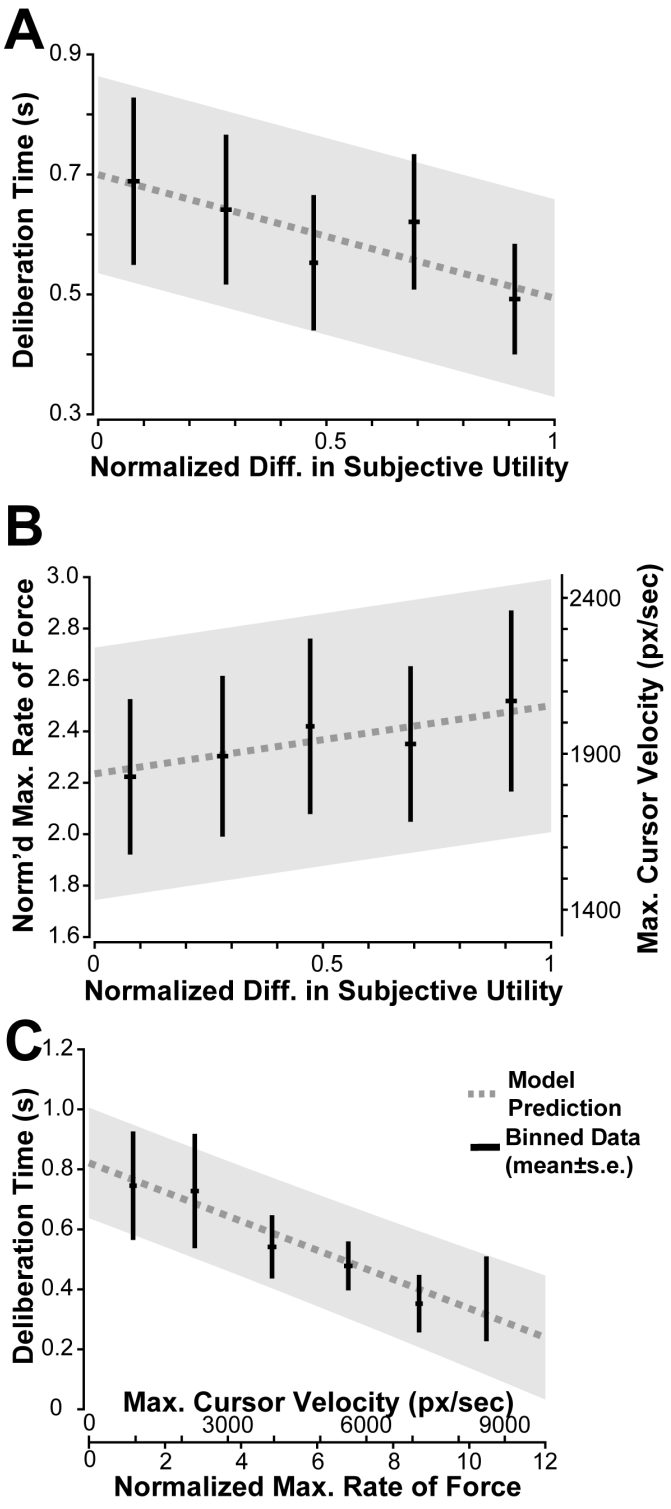
A
Deliberation Time (s)
0.9
0.7
0.5
0.3
0
0.5
1
Normalized Diff. in Subjective Utility
B
Norm'd Max. Rate of Force
3.0
2.8
2.6
2.4
2.2
2.0
1.8
1.6
2400
1900
1400
Max. Cursor Velocity (px/sec)
0
0.5
1
Normalized Diff. in Subjective Utility
C
Deliberation Time (s)
1.2
1.0
0.8
0.6
0.4
0.2
0
Model Prediction
Binned Data (mean±s.e.)
Max. Cursor Velocity (px/sec)
0
3000
6000
9000
0
2
4
6
8
10
12
Normalized Max. Rate of Force

Figure 13: Choice behavior. *A:* for the model of subjective utility that best predicts subject preference (U_6), the mean and standard error for binned deliberation time (black) is plotted against the normalized absolute difference in subjective utility between force profiles. *B:* Deliberation time decreased as the difference in normalized subjective utility increased. The mean and standard error for binned choice vigor is plotted against the normalized difference in subjective utility between force profiles. Choice vigor is the maximum rate of force, manifested as the maximum cursor velocity for a choice. Choice vigor increased as the difference in normalized subjective utility increased. *C:* as choice vigor increased, deliberation time decreased. Mean and standard error for deliberation time binned by choice vigor is shown in black. In all panels, the linear model prediction with 95% confidence intervals is shown in gray.

Interestingly, choices with a greater difference in utility were also made more vigorously (Figure 13B). Choice vigor was fit using the winning model of utility with a linear mixed-effects model that included a random effect of subject. We found choice vigor had a significantly positive correlation to the difference in utility (β = 0.266 [0.214, 0.318], $P < 0.0001$).

Independent of subjective utility models, we found that choices that were made faster (i.e., shorter deliberation time) were also made more vigorously (β = -0.0489 [-0.0588, -0.0390], $P < 0.0001$), shown in Figure 13C.

In summary, variations in deliberation time and response vigor were explained by the same model of subjective utility that best explained subject preferences, supporting the idea that choices were not just made with an arbitrary, consistently executed heuristic. Collectively, these results reinforce our findings that a subjective model of utility includes independently time-sensitive penalties on physical effort costs and task costs and further suggest there is a link between an individual's subjective valuation of utility, their decisions, and their movement control.

4.4.5 Subjects do not fatigue

We examined if subjects fatigued during the experiment. Using a linear model, we found that there were no significant effects of experiment phase (start and end of

the *Familiarization Phase*: $\beta = 2.83 \times 10^{-3}$ [-0.122, 0.127], $P = 0.964$; start and end of the Session: $\beta = 0.119$ [-0.00641, 0.244], $P = 0.0627$, start and end of the *Choice Phase*: β=-0.116 [-0.00923, 0.241], $P = 0.0692$), nor between hill or valley session types (β=-4.27×10^{-3} [-0.106, 0.0978], $P = 0.934$) on subject's maximum force exerted (Figure 14A). The lack of difference suggests that subjects were not decreasing in their ability to perform the task.

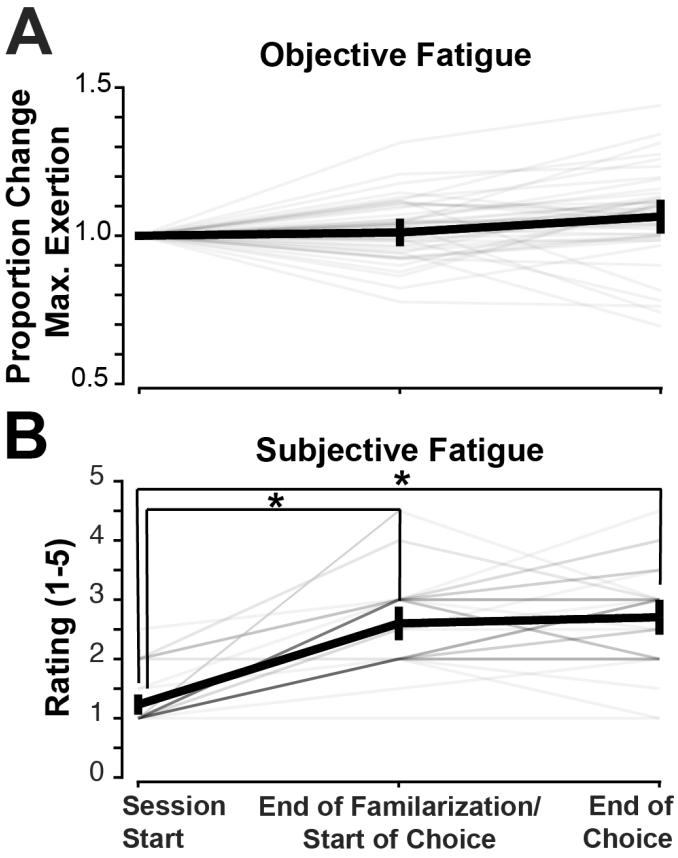

Figure 14: Measures of fatigue. *A*: objective physical fatigue did not change throughout the experiment. The proportion of change of the maximum force exerted is plotted by phase. Subject means and standard error bars are shown in black with individual subjects plotted in gray. *B:* Subjective fatigue ratings (1-5) are plotted by phase. Subject means and standard error bars are shown in black with individual subjects in gray. *significant effect of phase, seen between the start of the session and the end of familiarization and start of the session and end of the full session.

Physical fatigue was also assessed via self-report using a Likert scale rating from one to five (Figure 14B). Subjects generally began with little to no fatigue (most

rated their fatigue as a 1), with a significant increase in rating between the start and end of the *Familiarization Task* (β = 1.37 [1.16, 1.58], $P < 0.0001$) and start and end of the Session (β = 1.48 [1.28, 1.69], $P < 0.0001$), but no difference between start and end of the *Choice Task* (β = 0.113, [-0.0945, 0.320], $P = 0.284$) nor of session type (β = -0.158 [-0.327, 0.0108], $P = 0.0664$). These results suggest that subjects' maximum performance did not degrade objectively nor did they subjectively fatigue throughout the choice period.

We analyzed potential secondary indications of physical or cognitive fatigue throughout the *Choice Phase*, which could result in a decrease in performance. If subjects were fatigued, we could expect to see longer deliberation times, lower choice vigor, and/or higher error in *Realization Trials* later in the session. Interestingly, we found that choice vigor was significantly positively correlated to trial number (β = $4.46\text{x}10^{-4}$ [$3.43\text{x}10^{-4}$, $5.48\text{x}10^{-4}$], $P < 0.0001$), where subjects had a higher peak force at later trials, opposite of what we would expect if subjects were fatiguing. Similarly, we found that choice deliberation time was significantly negatively correlated to trial number ($\beta = -2.82\text{x}10^{-4}$ [$-3.31\text{x}10^{-4}$, $-2.33\text{x}10^{-4}$], $P < 0.0001$), where subjects had shorter deliberation times at later trials, again, opposite of what we would expect if subjects were fatiguing. Finally, we found that error during *Realization Trials* was significantly positively correlated to trial number (β = $9.21\text{x}10^{-6}$ [$5.73\text{x}10^{-6}$, $1.26\text{x}10^{-5}$], $P < 0.0001$), where subjects had slightly more error in later trials (<0.5% average increase in error between trials 1 and 455). However, all subjects still performed within the required accuracy for all realization trials.

If subjects were fatiguing, we may have also seen a change in preference from the beginning to the end of the *Choice Phase*. Using the linear model in Equation 31, we found that there was not a significant difference between preferences in the first half versus the second half of the session (HalfSession: β = -0.0736 [-0.108, 0.0933], P = 0.886; HalfSession x SessionType = Hill: β = -0.0691 [-0.210, 0.0719], P = 0.377). Taken as a whole, these measures support that subjects did not fatigue in a way that significantly impacted performance or preference throughout the experiment.

4.4.6 Force trial performance as a predictor for preference

All subjects performed the isometric arm-pushing tasks with less than the maximum 7% error. No subject failed a *Force Trial* in the *Familiarization Phase* nor the *Realization Trials* during the *Choice Phase*. During the *Familiarization Phase*, subjects performed trials with an average error of 1.88 ± 0.180% (mean ± s.e.), and in the *Choice Phase*, with 2.35 ± 0.285% (mean ± s.e.). In both phases, subjects met the accuracy requirements for all trials.

We investigated if there were any differences in the types of profiles and how they were performed. We found that the duration of a hill or valley had a significant correlation to error, where longer duration hills and shorter duration valleys were performed with less accuracy (SessionType = Valley, $\beta = 5.04\text{x}10^{-3}$ [$2.39\text{x}10^{-3}$, $7.68\text{x}10^{-3}$], P = 0.000191; Duration, $\beta = 1.04\text{x}10^{-3}$ [$4.39\text{x}10^{-4}$, $1.64\text{x}10^{-3}$], P = 0.000709; Duration x SessionType = Valley, $\beta = -1.12\text{x}10^{-3}$ [$-1.97\text{x}10^{-3}$, $-2.68\text{x}10^{-4}$], P = 0.00998), however, the difference in onset time did not correlate with error ($\beta = 1.58\text{x}10^{-4}$ [$-8.08\text{x}10^{-5}$, $3.96\text{x}10^{-3}$], P = 0.195). When the difference in error between force profiles was added

as a factor to the linear model for preference in Equation 10, the difference in error was significantly related to the probability of choosing a profile (β = -8.02, [-14.5, -1.58], P = 0.0146), where the profile that was more accurately performed was preferred. However, all other factors remained significant from the original model without accuracy as a factor, including duration. Though error and duration covary, they both predicted preference. Because no subject failed a trial due to error, we believe the difference in hill or valley duration was the driver behind preference but happens to covary with error. However, we cannot verify the direction of causality statistically.

Finally, we looked at average force before, during, and after a hill or valley. If this varied between profiles with different durations or onset times or with trial number across the choice phase, it could be an indication that subjects may be acutely fatiguing within a force trial, and perhaps an explanation for why subjects might be precrastinating. However, we found no significant correlation with duration, onset time, trial number, or their interactions, suggesting that acute performance differences could not predict preference.

4.5 Discussion

Here, we found that subjects are sensitive to the timing of physical effort costs and prefer to precrastinate physical effort. This finding is consistent across models of utility tested. We also found that the best model of utility includes a task cost that is also time sensitive, though subjects were idiosyncratic in preference. Subjective

utility also correlates well with choice deliberation time and vigor, suggesting a common utility underlying motor control and decision making.

This study builds on a number of others investigating the subjective valuation of physical effort. Although much work in biomechanics has demonstrated the importance of metabolic cost as a measure of physical effort, recent work has suggested the cost of effort is more than energetic cost alone. Subjective effort costs appear to grow nonlinearly with energy or force (Hartmann et al., 2013; Klein-Flügge et al., 2015; Morel et al., 2017; O'Brien & Ahmed, 2019) and the subjective valuation seems to vary widely across individuals (Summerside & Ahmed, 2021). Here, we demonstrate that the time of effort investments also influences decision making over and above physical effort costs alone.

Unsurprisingly, subjects were sensitive to the total amount of physical effort, in line with other studies that show that subjects prefer shorter durations of high force and the lower total force-time integral overall (Körding et al., 2004). However, the time sensitivity results are novel. An energetic model would predict no time sensitivity, and an economical model might predict procrastination. However, we found that physical effort costs were precrastinated and subjects were inconsistent about time preference of task costs. Why do subjects precrastinate physical effort?

Because physical effort investment is not tied to task completion in our paradigm, Rosenbaum's hypothesis of reducing cognitive effort cannot explain our results (Rosenbaum et al., 2014). Precrastination of physical effort could align with pain studies that observe a form of precrastination, preferring to end on a better note

(Kahneman et al., 1993; Redelmeier & Kahneman, 1996). However, in these studies, pain is experienced without predictability and preference is evaluated in retrospect, relying on a memory of an expended cost. Our study has complete insight and decisions are made prospectively, so it is not likely due to a recency bias due to retrospection but could be due to discomfort associated with physical effort exertion.

Perhaps discomfort or fatigue increases within a force profile. The same force level might feel easier at the beginning of a profile compared with the end, causing a preference for high effort early. Indeed, the subjective value of effort increases when subjects are fatigued (Hogan et al., 2020; Iodice et al., 2017). If a state of fatigue can impact subjective valuation of effort, it is plausible that precrastinated effort within a task is preferable to avoid the possibility of fatigue. Our analyses of average force for a given segment of a profile shows that objectively, there was no difference in how subjects performed due to duration or onset time. Thus, a difference in force output behavior cannot predict preference, and this possible explanation for physical effort precrastination is not verifiable based on the data we collected. Furthermore, we did not observe fatigue throughout a session, nor a change in behavior or preference indicative of fatigue. Although discomfort or acute fatigue was not specifically observable, this remains a possible explanation for subject precrastination.

Exerting effort itself might be intrinsically rewarding, transforming effort into reward and driving it toward precrastination. However, if this were the case, we would have observed subjects choosing longer hills or shorter valleys, consistently

seeking to maximize effort expenditure. This was not observed in our experiment, thus is an unlikely explanation for precrastination.

In this experiment, subjects must perform the effortful task and they knew that they could perform the profile. There was no delay and necessity of performing the effortful task was certain. Force profiles were accurately visually represented, thus there was no uncertainty about what amount of effort the task required. Could this certainty about the necessity to perform physical effort and the ability to do so, cause subjects to precrastinate? Recent evidence suggests precrastination is the "automatic" response when costs are low, but as costs scale up, subjects care more about efficiency and begin to procrastinate. However, in many of these studies, procrastination is also conflated with the minimum energy solution. With bucket carrying (Rosenbaum et al., 2014) and water cup carrying (Raghunath et al., 2021) performing the task later is also less effortful because the distance carried is less. As demand for either physical effort costs or task costs increases, where performance uncertainty becomes a factor, costs seem to be procrastinated and temporal preferences are overruled. It could be an interesting follow-up experiment to scale up the physical effort, in either intensity or duration, or manipulate uncertainty, to see if there is shift in preference towards procrastination.

4.5.1 Choice behaviors reflect utility

In this study, we show that the degree of preference (i.e., the difference in subjective utility of the options) is reflected in choice behavior. In perceptual discrimination tasks, the more similar two stimuli are, the longer it takes to

distinguish between them; thus, more difficult decisions are made more slowly (Festinger, 1943). In our study, the closer in value the subjective utility is between two options, the longer it took subjects to make that choice. A more novel finding is that a change in the vigor of the response readout is also found to vary with decision difficulty, in line with some recent studies (E. Smith & Peters, 2022). Previous studies have shown saccades are made faster to more rewarding targets (Takikawa et al., 2002; Xu-Wilson et al., 2009), and saccade vigor reflects the subjective utility of a decision when the effort is varied (Shadmehr et al., 2019). Recently, Korbisch et al. show that a difference in utility between options is reflected in the vigor of saccades made during the decision-making process, where the greater the difference between the utility of choices, the greater the difference in peak saccade velocity between each option (Korbisch et al., 2022). Our study supports this finding, where the vigor with which an option was chosen increased as the difference in subjective utility between the two choices increased.

4.5.2 Limitations

One critique of our experiment design is that we sampled too coarsely. We intentionally tested discrete values of hill and valley duration and onset time for test efficiency and the sake of between-subject comparison, but perhaps these were too coarsely sampled with a minimum of two-second differences in onset time and/or duration between profiles for a given choice trial. We did not find a precise indifference point between time preference and duration, though we could predict where this might occur using the subjective utility model. Taking an adaptive

sampling approach, we may have been able to more precisely define how much extra physical effort is worth performing at a given time.

Our formulation of task costs may be a combination of physical effort and cognitive effort. To date, a precise measure of objective cognitive effort remains unclear. Using the rate of force squared as a representation of task costs is in line with a control perspective, where a change of state requires control input. This formulation may be conflated with physical effort cost because force generation (i.e., a positive rate of force) incurs a metabolic cost over and above force maintenance (Russ et al., 2002). This formulation could also be aligned with the attentional demands in motor control, where tracing complexity requires more effort from the subject. How we model task costs is task-centric but may not be a clean break between cognitive and physical demands. Nevertheless, as evinced by the winning model of utility, this formulation of costs is better off included in the overall utility.

Our results are similar to many other precrastination studies, where a majority of subjects are precrastinators, but it is not unanimous (Blinch & DeWinne, 2019; Fournier, Coder, et al., 2019; Kahneman et al., 1993; Patterson & Kahan, 2020; Rosenbaum et al., 2014). We also see about two-thirds of our subjects precrastinate (16 of 25 subjects precrastinate physical effort). Currently, there are no conclusive measures that can stratify and predict who in a sampled population will decide to precrastinate or procrastinate. With further investigation, we may better understand the underlying mechanisms of precrastination beyond task-specific descriptors.

4.5.3 Conclusions

In this study, we designed an isometric arm-pushing task to tease apart subjects' preferences for physical effort, tasks costs, and time sensitivity. We found that subjects exhibit a preference for when both physical effort costs and task costs occur, adding to the body of evidence that costs in motor control are indeed time-sensitive. We found that, on average, subjects choose to precrastinate physical effort but are idiosyncratic about the timing of task costs. We propose a generalizable descriptive model of subjective utility, which predicts subject preferences as well as choice deliberation time and choice vigor.

CHAPTER V

SELF-PACED WALKING IN VIRTUAL REALITY

5.1 Foreword

The following chapter is a summary of the combined efforts from myself and Mr. J. Sam Rubaker. Mr. Rubaker built the original experimental interface consisting of a virtual reality system and instrumented treadmill, and he collected pilot data. This work is summarized in his Master's thesis (Rubaker, 2023). I built and designed the self-paced treadmill interface, integral to the experimental setup. Following pilot data collection, I further analyzed data and prepared figures. Interpreting findings from the pilot data, I revised the experimental interface and experiment design. I authored this chapter, which summarizes the mechanics of the virtual reality experimental setup, the protocol for the pilot study, new analysis and results, and the lessons learned from the pilot study. Importantly, these findings lay the foundation for the formal investigations reported in Chapter VI.

5.2 Abstract

Recent findings have shown that the speed at which we move reflects the value of what we hope to acquire. We reach faster and move our eyes faster to objects we value more; however, there is currently little understanding about how value influences self-selected walking speed. To probe the effect of reward on walking speed, we integrated a motion capture system, a self-paced treadmill, and a virtual reality (VR) headset to immerse participants in a realistic environment. We conducted a pilot study, where preliminary results suggested that higher rewards are captured with

greater walking speed. Interestingly, we also saw an influence of a history of reward, but not in the way we would expect. Instead of a history of high rewards increasing walking speed, we saw participants move faster to a high reward after a history of low rewards, suggesting an invigorating effect of reward prediction error. From this pilot study, our results suggested that walking vigor is modulated by reward value and perhaps a history of reward, though more formal investigation is needed. The experimental interface used for the pilot study demonstrated promise for future investigation but necessitated some improvements, namely several revised visual features in the virtual environment as well as a new control algorithm for the self-paced treadmill to enable hands-free, unassisted walking while wearing the VR headset.

5.3 Introduction

Expectation of reward has been shown to influence both decision making and motor control. We not only prefer a more valuable option, but we also move faster towards it (Korbisch et al., 2022; Sackaloo et al., 2015; Summerside, Shadmehr, et al., 2018; Takikawa et al., 2002; Xu-Wilson et al., 2009; Yoon et al., 2020). Though the pace at which someone walks has garnered much attention of researchers and clinicians, this sensitivity to value has not been shown in human locomotion. Here, we explore how to build a controlled experimental interface to answer the question of whether value can influence walking speed.

Below, we provide a brief overview of the literature behind the influence of reward on movement speed; however, Chapter VI will cover this in greater detail.

Preferred walking speed has been thought to be determined by energetic costs, where humans and other animals tend to walk at speeds that minimize the metabolic cost of movement per unit distance, or cost of transport (Browning et al., 2006; Donelan et al., 2001; Hoyt & Taylor, 1981; Ralston, 1958; Selinger et al., 2015). In contrast, optimal foraging theory posits that animals use strategies to maximize their energy intake while minimizing the associated costs, such as energy expenditure and time spent searching (Charnov, 1976). Recent research extends this theory to human movements, where similarly, a movement should aim to maximize the net utility over time, represented as the sum of the rewards minus energetic costs, all divided by time. This theory predicts we should move faster to a higher value reward. Additionally, movements should be influenced by a history of rewards, where this history forms an expectation for upcoming reward. A history of higher rewards should lead to faster movements because the opportunity for future rewards is greater, or an "opportunity cost" of time. In contrast, a history of reward might create a prediction for an upcoming reward, and an error signal based on the difference between the actual reward value and what was expected, or a "reward prediction error" (RPE). This could lead to an opposing effect to opportunity cost, where in a highly rewarding environment, there is a higher expectation of reward, thus if a lower reward value is offered, it leads to disappointment and a negative RPE.

How do we create a controlled environment to manipulate immediate reward and a history of reward, but still allow participants to walk freely at their own pace? Value-driven effects on eye and reaching movements are studied in seated

participants in video game like environments (Korbisch et al., 2022; Summerside, Shadmehr, et al., 2018; Takikawa et al., 2002). To create a similar game-based environment when a person is walking, we must use virtual or augmented reality. Virtual reality (VR) provides a way to immerse participants in naturalistic environments and allow them to perform goal-directed movements in an environment where experimenters can manipulate what is experienced.

Relatively few studies have investigated walking speed as an outcome measure with a VR headset on a treadmill. Greater gait variability and instability were observed while wearing a VR headset (Hollman et al., 2007). One study showed participants walked more slowly with a VR headset (Wodarski et al., 2020). Walking with VR has become a popular rehabilitation tool, where a gamified VR environment has been shown to increase a (manually adjusted) walking speed in healthy controls and patients with multiple sclerosis or stroke (Winter et al., 2021).

Measuring preferred walking speed on a treadmill has typically been accomplished using one of two methods: manual adjustments of a fixed-paced treadmill, or automatically adjusted by measuring the participant's position and/or speed relative to the treadmill. The technology to accomplish this has evolved from the use of photocells (Hughes & Goldman, 1970) to utilizing instrumented treadmills and high-speed motion capture systems (Pimentel et al., 2022; Sloot et al., 2014; Song et al., 2020). The latter approach of automatic adjustment is essential for us to observe walking speed changes that vary subconsciously and potentially on timescales not observable if a manual-adjustment technique is utilized.

Below, we present pilot data and detail the development of a virtual reality walking system that can probe the effect of reward on walking speed. Participants walked at their own pace through a virtual environment collecting monetary rewards symbolized as apples, where we manipulated the value of rewards and frequency that different values occur. We hypothesized that both immediate reward and the history of reward will influence walking speed. We predicted that the presence of immediate reward would increase walking speed proportional to the reward value. We also predicted that a history of higher reward would increase walking speed. Following pilot data analyses, we organized lessons learned and implemented required changes ahead of formal data collection.

5.4 Pilot study methods

5.4.1 Pilot study participants

Eight healthy young adults (n = 8, 2 Females, Age: 27.5 ± 6.5 yrs) participated in the pilot testing. To qualify, participants were required to be between 18 and 35 years of age, able to understand English and with vision normal or corrected to 20/20. Exclusion criteria included any major musculoskeletal or ophthalmological impairments, low or sedentary activity level (self-reported), any previously experienced side effects from VR including motion sickness (self-reported), any neurological or vestibular diseases, and any medications that could impair the ability to complete the walking task. Participants were informed that they would not be compensated for pilot testing the experiment, but that they would be able to earn candy throughout the experiment.

5.4.2 Pilot study experimental setup

For this study, we built and integrated a self-paced treadmill with a virtual reality headset. This system integrated numerous components into a single, immersive, self-paced walking experience. We used an Oculus Quest VR Headset, a Vicon motion capture system, a Bertec force-instrumented treadmill, and a virtual environment built in Unity, all controlled with custom MATLAB scripts.

Participants donned a position marker belt fitted to their waist and positioned just above the sacrum (Figure 15A). Reflective markers were placed on the position marker belt and the VR headset, which were read into the system via the Vicon motion capture cameras and Vicon Tracker software to track head position and participant position. The participant's waist position was used to determine the forward speed at which the participant was moving through the environment. This position was read by the self-pacing algorithm to update the speed of the treadmill, which was, in turn, read by Unity to translate the environment forward for a given timestep and speed.

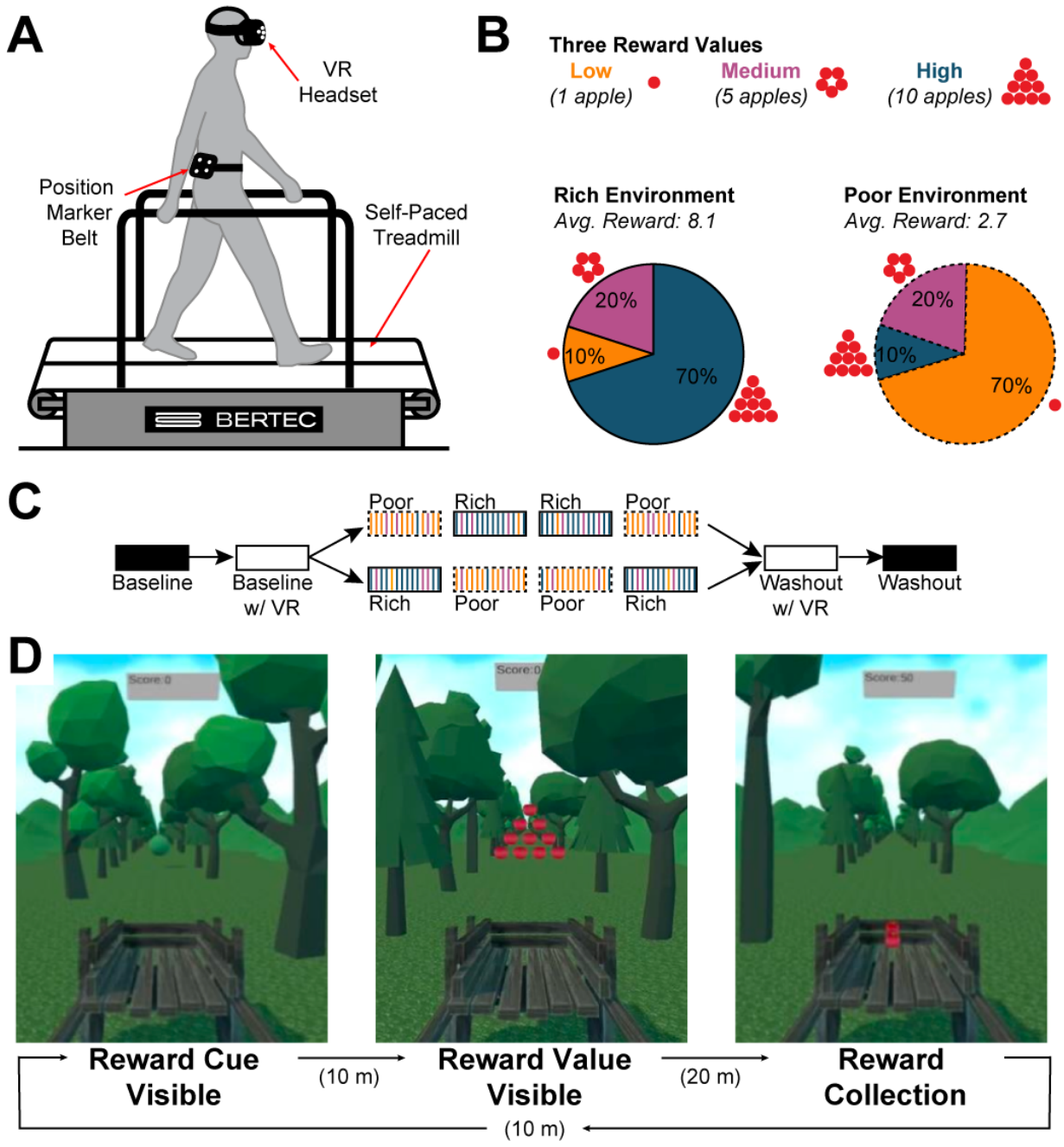

Figure 15: Experimental setup and protocol. *A:* participants walked on an instrumented treadmill while wearing a position marker belt and VR headset. *B*: In a rewarding trial, participants could see one of three reward values: low/1-apple (yellow), medium/5-apple (purple), and high/10-apple (blue). A rich environment had an average reward value of 8.1 apples, and consisted of 70% high rewards, 20% medium rewards, and 10% low rewards. A poor environment had an average reward of 2.9 apples, and consisted of 10% high rewards, 20% medium rewards, and 70% low rewards. *C*: Within a session, participants completed eight walking trials, each six minutes in length. Participants were placed in one of two counter-balanced groups, where one group saw reward environments ordered as poor-rich-rich-poor, and the other group, rich-poor-poor-rich. *D:* each rewarding trial followed the same repeating flow. After 10 meters of walking a green reward cue would appear, then 10 meters of walking later, the value of the reward was revealed, then after 20 meters of walking, the reward was collected.

5.4.2.1 Self-paced treadmill

The self-paced treadmill ran using a custom control algorithm based loosely on previous work from Pimentel et al. 2022 and Song et al. 2020. (Pimentel et al., 2022; Song et al., 2020). Generally, as a participant sped up, moving forward on the treadmill, the treadmill would speed up bringing the participant back to the center; as a participant slowed down, moving backwards on the treadmill, the treadmill would slow down until the participant moved forward towards to the center. The algorithm is a bit more nuanced and considers both the position and velocity of the participant ($p_{participant}$ and $v_{participant}$, respectively). If position was between 2.5 cm in front or behind the center, the controller did not update the speed. This 5 cm deadzone allowed for some natural variability in walking speed that might not be intentional. If the participant was outside the deadzone, the algorithm considered their velocity. If a participant was in front of the center of the treadmill ($p_{treadmillcenter}$, constant), but moving backwards, the treadmill did not update its speed. This was an attempt to prevent overshoot in the negative direction, potentially sending a participant off the back of the treadmill. If the participant was either behind the center and outside the deadzone or ahead of center and with a non-negative velocity, the treadmill speed ($V_{treadmill}$) was updated as follows, where k_p, the position gain, is equal to 1:

$$V_{treadmill,t+1} = V_{treadmill,t} + k_p\left(p_{participant,t} - p_{treadmillcenter}\right)^3 \quad (32)$$

The treadmill velocity update logic was as follows:

```
if position outside the deadzone then
        if (position is forward of center and velocity is negative) then
                Do not update speed.
        else
                Update speed, per Equation 32.
        end
else
        Do not update speed.
end
```

We also implemented some restrictions to make the self-paced treadmill usable with the VR headset. Participants were instructed to use the handrails while walking, and speed could not exceed 2 m/s nor go backwards.

5.4.2.2 Virtual environment

The virtual environment was built to emulate an orchard for participants to forage for rewards, symbolized as apples. Participants walked along a tree-lined path in the virtual orchard. In front of the participant, a wheelbarrow was displayed, where the handles of the wheelbarrow were aligned with the handrails of the treadmill. Additionally, a scoreboard was displayed overhead that counted the total number of points acquired during a trial. With 10 seconds left in a trial, the scoreboard also displayed a countdown timer.

A variety of sounds were played during a trial. In the background, ambient forest noises played to enhance immersion and obscure some of the treadmill and fan noise. Pleasant tones were played after each reward was collected, and the tones varied based on the value of the reward. In addition, an audio alert played when the 10-second trial countdown timer began.

As the participant walked, they encountered different reward values, (i.e., apple quantities). A reward value could be one of three values: High/10-apple, Medium/5-apple, and Low/1-apple (Figure 15B). In addition, we varied the quality of the environment between "rich" and "poor." In a rich environment, rewards were randomly distributed as 70% High/10-apple, 20% Medium/5-apple, and 10% Low/1-apple; in a poor environment, 10% High/10-apple, 20% Medium/5-apple, and 70% Low/1-apple. These distributions led to an average reward value of 8.1 apples in the rich environment and 2.7 apples in the poor.

Each reward trial consisted of repeating cycles of reward events spaced evenly every 40 meters (Figure 15D). After participants walked 10 meters, a green orb "cue" would appear, signaling the location of the reward 30 meters ahead, without revealing its value. After participants walked 10 meters closer, the cue disappeared and revealed how much reward they could collect (High, Medium, or Low). After walking the final 20 meters, participants collected the reward by simply walking through the apples. Upon collection, a pleasant sound was played that varied based on the value of the reward. The collected apples appeared in the wheelbarrow, and the score was updated overhead. This pattern repeated until time expired for the trial.

5.4.3 Pilot study protocol

At the beginning of the session, participants were familiarized with the equipment, including the VR headset, position marker belt, self-paced treadmill and virtual environment. Participants donned the position marker belt and were given

instructions on how to interact with the self-paced treadmill. They were instructed to always hold on to the handrails, but once they felt comfortable, they were allowed to loosely grasp the handrails so they could move effectively forward and backward. They were told that as they walked down the path, they would encounter a green cue that as they got closer would reveal a certain number of apples. Each individual apple was worth five points, and for every 100 points, they would earn a piece of candy. The experimenter instructed participants that their goal was to earn as many points as possible.

The test session consisted of eight six-minute trials (Figure 15C). The first two trials gradually introduced participants to walking in a VR environment. In the first trial, "baseline", participants walked with no VR headset, and next, "baseline with VR", they walked in VR but without rewards. In each of the baseline trials, participants were instructed to explore the feeling of accelerating and decelerating on the treadmill for the first 60 seconds, then settle into a natural walking pace. The middle four trials were the focus of the pilot study, during which participants walked while collecting rewards in either a rich or poor environment. The order was counter-balanced across participants; either rich-poor-poor-rich or poor-rich-rich-poor. The final two blocks were the same as the first two trials, but in opposite order: no reward in VR ("washout with VR") then walking without VR ("washout").

5.4.4 Data processing and statistical analyses

Our primary outcome measure was walking speed, which was inferred from the treadmill speed, recorded at roughly 10 Hz in MATLAB. The state of the virtual

environment was recorded at 10 Hz from Unity. These data included participant position and velocity within the virtual environment coordinate frame, as well as when cues and rewards were visible. The time stamps of the specific events in the virtual environment were used to parse walking speed data recorded in MATLAB. All data collected were analyzed and processed using custom scripts and built-in functions in MATLAB.

We used linear mixed-effects models to determine the effects on walking speed. We investigated the fixed covariates and interactions of immediate reward value (*RewardValue*; ordinal, Low, Medium, High), environment (*Environment*, categorical, Rich or Poor), and *Participant* (categorical) as a random intercept term. The model is summarized in Equation 33 below, where $Speed_{ij}$ is the speed at a given moment of interest during the *j*th reward instance for participant *i*, and $Participant_i$ is normally distributed with mean 0 and variance σ^2.

$$\begin{gathered} \text{Speed}_{ij} \sim N\left(\mu_{ij,}\sigma\right) \\ E\left[\text{Speed}_{ij}\right] = \mu_{ij} \\ \mu_{ij} = \text{RewardValue}_{ij} + \text{Environment}_{ij} \\ +\text{RewardValue}_{ij} \times \text{Effort}_{ij} + \text{Participant}_i \\ \text{Participant}_i = N(0, \sigma^2) \end{gathered} \tag{33}$$

We also used a linear mixed-effects model to test a correlation of walking speed to RPE. For this analysis, we calculated RPE as the difference between the reward value, *R*, and the average reward for the environment, $E_{environment}$.

$$\begin{gathered} RPE = R - E_{environment} \\ \text{where: } E_{rich} = 8.1 \text{ and } E_{poor} = 2.7 \end{gathered} \tag{34}$$

For example, a high reward in a rich environment would have an RPE of +1.9, whereas a high reward in a poor environment would have an RPE of +7.3. Conversely,

a low reward in a rich environment would have an RPE of -7.1, whereas a low reward in a poor environment would have an RPE of -1.7.

5.5 Pilot study results

5.5.1 Walking speed

First, we inspected how walking speeds vary within a given trial and across participants (Figure 16). Within a trial, we saw considerable variation, that could be varying systematically based on event (cue, reveal, collection). Figure 16A shows an example trial in a poor environment. Interestingly, we saw this participant accelerate considerably after seeing a rare high reward.

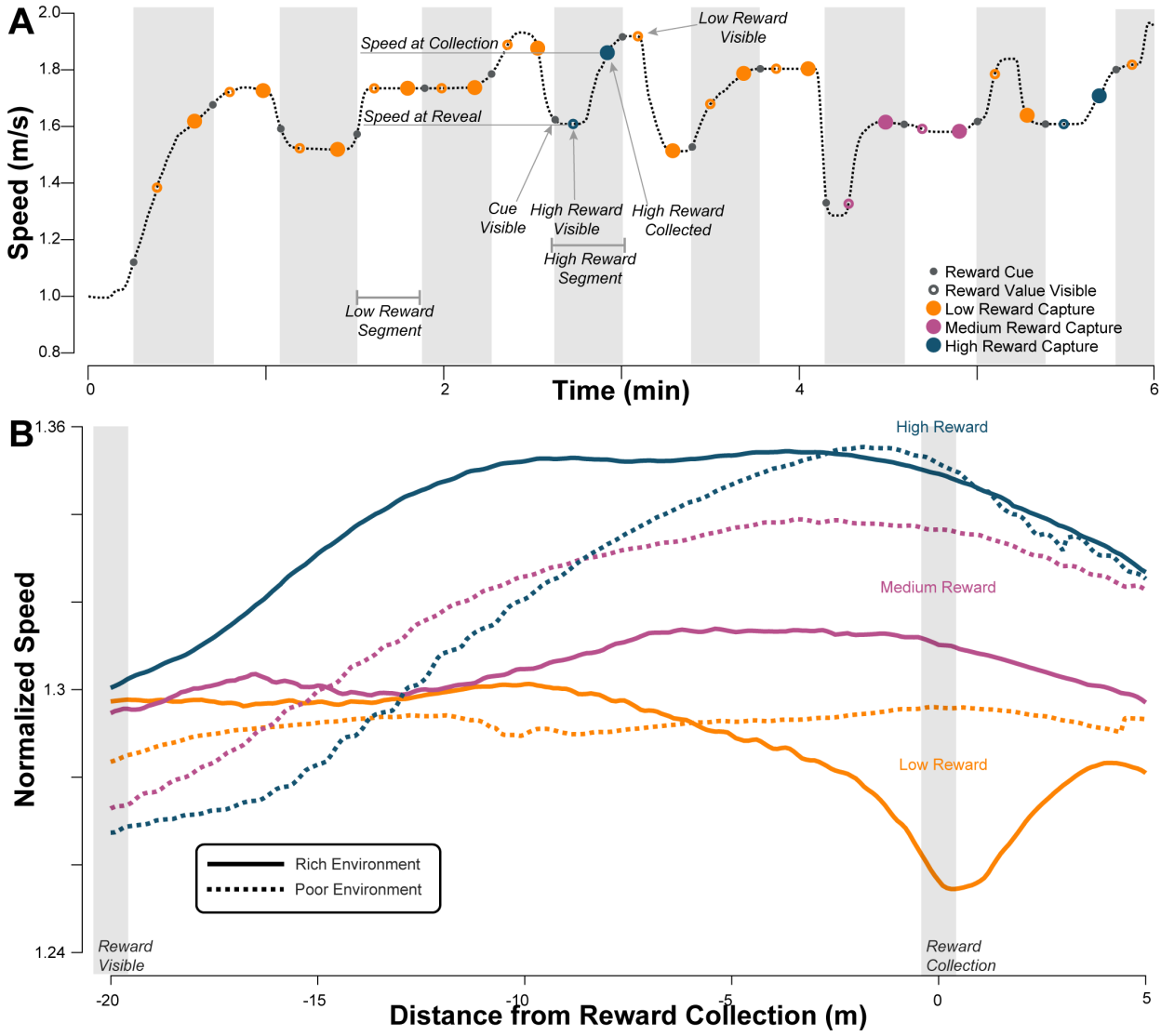

Figure 16: Example trial and average walking speed. *A*: an example trial in a poor environment. Each reward segment is highlighted by vertical white or gray segments. Reward value is indicated by color, where high is blue, medium is purple, and low is yellow. Throughout a trial, speed is not constant. *B*: average walking speeds, normalized to each participants' baseline speed is plotted relative to distance from reward collection. Each line represents a combination of reward value and environment, where reward values follow the same color scheme, and a rich environment is shown in solid lines, and a poor environment, in dashed lines. Inspecting these averages, trends start to emerge.

We segmented each instance of reward collection and categorized them by the value of reward and the reward environment. Averaging within a participant, then across participants, we plotted each of these curves in Figure 16B. From these curves we start to see some trends emerge. First, before reward value is known, we see a

grouping of speeds by environment, where the rich environment appears to be a slightly faster walking speed than the poor environment. Following reward value reveal, we see acceleration that matches the reward value, where high rewards are collected at faster speeds than medium rewards, and medium rewards, faster than low. Interestingly though, we saw that speed at reward collection in the poor environment appeared to be faster than the rich. This, of course, needed to be verified statistically.

First, we looked at the effect of immediate reward on walking speed. We compared the instantaneous speed at reward collection across environments and rewards. Preliminary results supported our first hypothesis: higher rewards were collected at faster walking speeds (Medium: $\beta = 0.0784$ [0.00388, 0.153], $P = 0.0393$, High: $\beta = 0.122$ [0.0591, 0.185], $P = 1.61 \times 10^{-4}$, Figure 17A). We also found that there was a significant effect of the environment on the walking speed (Poor: $\beta = 0.0659$, [0.00275, 0.129], $P = 0.0409$); however, it was opposite of our hypothesis: walking speed for the same reward was slower in the rich environment than the poor environment.

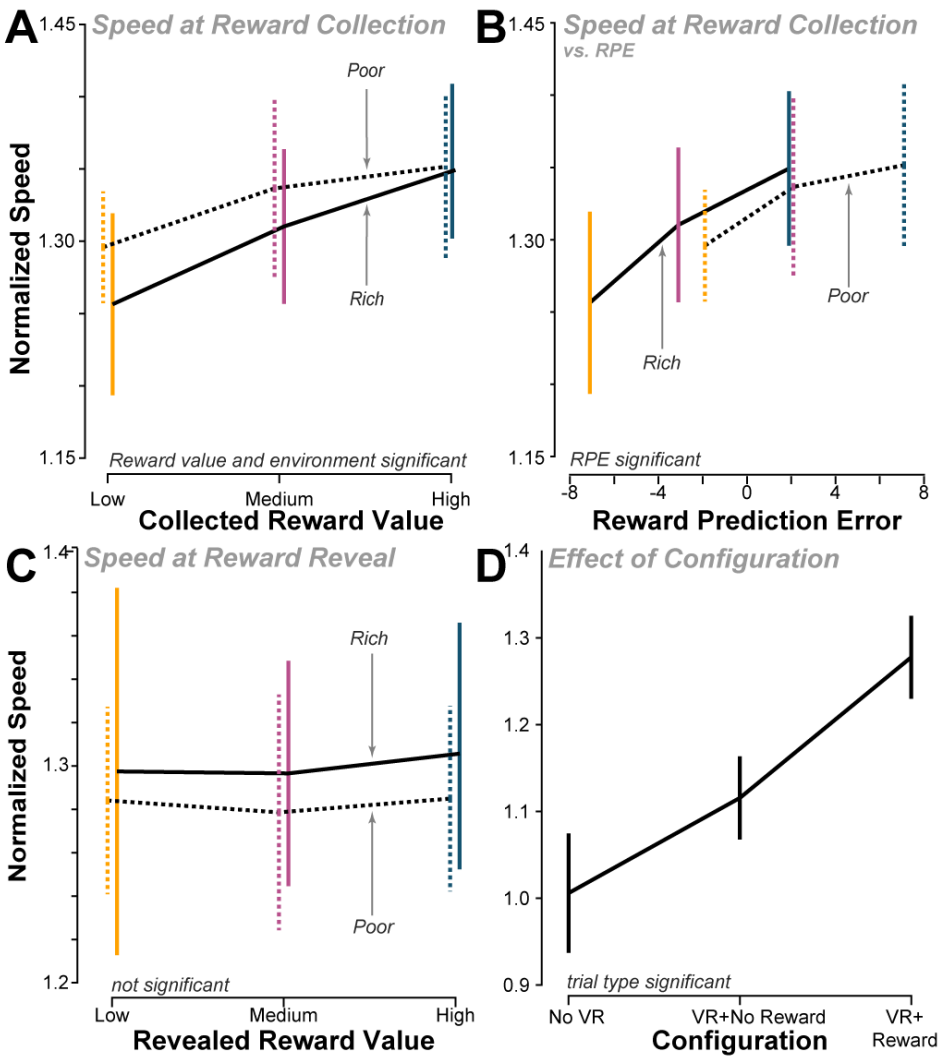

Figure 17: Walking speed data by event and configuration. *A:* normalized walking speed at reward collection, averaged across participants and separated by rewarding environment. *B:* normalized walking speed at reward collection but plotted as a function of RPE instead of reward value. *C:* normalized walking speed at reward value reveal, separated by reward environment. *D:* average speeds by configuration.

Because of the directionality of this effect, instead of an opportunity cost of time, there might be an effect of RPE, where the same high reward has a higher RPE in a poor environment than a rich environment. Indeed, we found that walking speed at reward collection is strongly positively correlated to RPE (β = 0.0103 [0.00571, 0.0148], P < 0.0001, Figure 17B), suggesting that a history of reward may be

influencing walking speed through this mechanism rather than an opportunity cost of time.

Next, we looked at the instantaneous speed at reward reveal (Figure 17C). At this moment of time, the participant did not yet know the value of the reward. As expected, we did not see an influence of the upcoming reward value. Though there appears to be some separation between rich and poor environments, this effect is not significant ($P = 0.978$). We also investigated if we might see an influence on *average* walking speed prior to reveal. Again, we found no significant effects ($P = 0.360$). It appears that, in our task, history of reward is not being considered in terms of an invigorating opportunity cost on time as hypothesized, but instead, in a formulation of expectation of reward.

Finally, we looked at gross averages across trial configuration types (Figure 17D). We compared the average walking speeds based on configuration. We grouped baseline and washout trials into the categories "No VR" and "VR+NoReward", and both rich and poor environments into "VR+Reward". We found that the VR with reward trials were significantly faster than other configurations (No VR: $\beta = -0.350$ [-0.443, -0.257], $P < 0.0001$, VR+No Reward: $\beta = -0.211$ [-0.304, -0.118], $P < 0.0001$). In addition, the VR+No Reward was significantly faster than the No VR condition ($\beta = 0.139$ [0.0312, 0.246], $P = 0.0123$).

Collectively, we saw variations in walking speed based on immediate reward value and a history of reward on reward collection speed, which could be an effect of

RPE. We also saw changes in speed due to configuration, with the fastest speeds occurring on reward trials.

5.5.2 Other Findings

Preliminary walking speeds results were promising, but we also identified several improvements that needed to be made ahead of formal data collection. We found that, through data analysis, some of our data points were too sparsely sampled. Between the trial duration and proportions of rewards in each environment, we were not able to gain confidence in our measures for the high reward/poor environment combination and the low reward/rich environment combination. In some trials, only a single instance of that reward combination was sampled. There were numerous ways to solve this, many of which we employed. First, we extended the trial time of the rewarding trials from 6 minutes to 9 minutes. Next, we adjusted the proportional breakdown of rewards to 20% low reward / 30% medium reward / 50% high reward in a rich environment, and 50% low reward / 30% medium rewards / 20% high reward in a poor environment. These two adjustments would triple our number of data points in the extreme reward-environment combinations, while not sacrificing too many samples in the other combinations.

Other minor feedback was on the user experience side. Some participants reported that the green cue was too difficult to distinguish, which we updated to a high contrast purple cue. The scoreboard was also small and sometimes obscured by the virtual trees, so the scoreboard was repositioned and updated to display in a larger font size and with greater contrast (black on white). Additionally, the apples

in the wheelbarrow bounced around, and were reportedly distracting. For this, we made the apples static in the wheelbarrow, and neatly ordered in rows and columns. Updates to the virtual environment are shown in Figure 18.

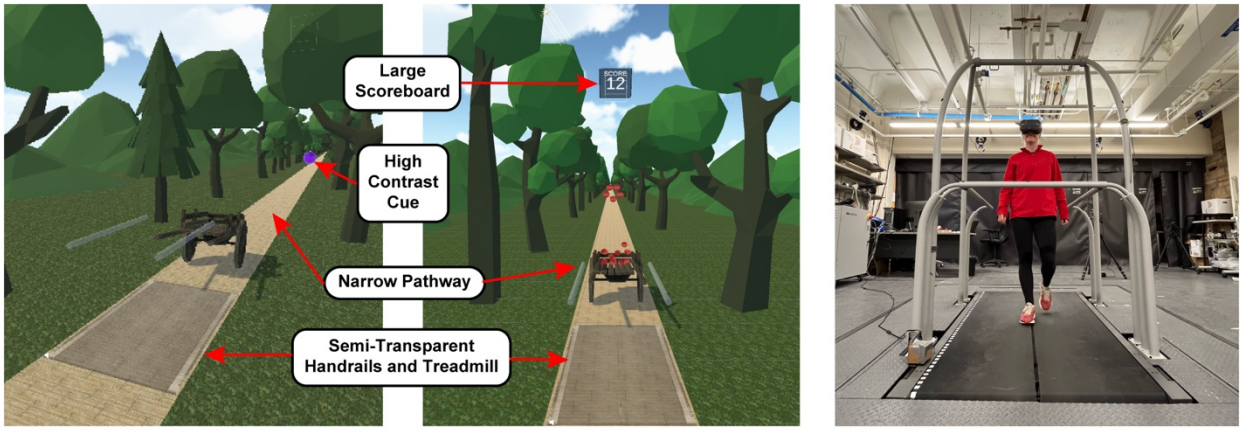

Figure 18: Improvements to the virtual environment. New visual features were added to the virtual environment to improve task awareness and enable hands-free walking. *Left* and *Middle:* new visual features include a larger scoreboard, a higher-contrast cue, the addition of a visible pathway, and a semi-transparent representation of the treadmill. The orderly apple stacking is not pictured. *Right:* a test participant demonstrates walking hands-free while wearing the VR headset.

During follow up pilot testing, we observed participants relying on the handrails as an assistive device rather than a safety backup for stability. One participant held on tightly to the handrails and walked at the maximum speed of the treadmill for all the rewarding trials. Overall, these observations raised concerns that handrails could make natural speed changes too salient and be used in a manner that would prevent us from acquiring the natural variation in walking speed we desired. Thus, we opted to create a system where walking with the VR headset without handrails felt safe.

Initial pilot tests with participants not holding on to the handrails failed. Participants would inevitably feel the need to grasp the handrails as the treadmill controller hunted (and often overshot) bringing the participant to the center of the

treadmill. Thus, we quickly recognized that the control algorithm needed to be updated.

5.6 Enabling unassisted walking with a virtual reality headset

To enable unassisted walking with a VR headset, most of this effort was spent tuning the treadmill controller to be responsive but also stable enough to walk safely on with a VR headset. When walking with a VR headset, a person's field of view is much narrower (Sauer et al., 2022) and walking kinematics increase variability (Hollman et al., 2007). However, we also made small changes to the visual interface that assisted with walking comfortably in the virtual environment. These findings are summarized below.

The first new visual interface was a safety feature. We included an accurate spatial representation of the treadmill platform and handrails (Figure 18). If participants needed to grasp the handrails for stability, they would be able to quickly find them without needing to remove the headset or go into passthrough mode. In addition, the treadmill platform position was there for peace of mind and awareness of the edges of the treadmill in case participants were anxious about stepping off the front or falling off the back.

The next important visual features helped keep participants roughly centered on the treadmill. Through A/B testing and subjective feedback, we found that a pathway approximately the width of the treadmill was helpful to keep participants aligned. The other important visual cue was the wheelbarrow. We initially removed it when starting to develop a hands-free walking experience, because its original

purpose was to improve immersion by making participants feel like their hands were on a wheelbarrow instead of a treadmill. However, this near-field object turned out to be a significant aid in maintaining lateral position while walking in VR, acting as a feedback mechanism for lateral deviation. The wheelbarrow was brought back, but was resized and positioned, so that it no longer matched the handrail position.

5.6.1 Algorithm tuning experiment

We sought to refine the self-pacing control algorithm for our system so that participants could walk confidently without using the handrails. This experiment was essential to test with humans in the loop. Because the participant is adapting and controlling their movement alongside the treadmill which is actively adapting to the change in movement from the participant, traditional controls analysis was inadequate. Additionally, beyond objective biomechanical performance, subjective measures such as trust in the system was paramount. Thus, my approach was to develop an experimental interface that could adjust controller gains in real time with a person actively walking on the treadmill.

Specific control values could be changed in real time while the treadmill was running (Figure 19). We limited the analysis to the following structure of control, where the velocity at the next time step, V_{t+1} was updated as a function of a position error, $\Delta p_{participant}$, and velocity error, $\Delta v_{participant}$:

$$V_{t+1} = V_t + k_p \Delta p_{participant,t}{}^{N_p} + k_d \Delta v_{participant,t}{}^{N_d} \tag{35}$$

The position error, $\Delta p_{participant}$, was the signed distance from the edge of the deadzone, and $\Delta v_{participant}$ was the difference in velocity from bounds of the velocity deadzone.

The tunable parameters were the gain values (k_p and k_d), exponents (N_p and N_d), and deadzones for position and velocity. The gain values could range from 0 to infinity. The exponents could be odd values from 1 or greater, and the deadzones could range from 0 to infinity.

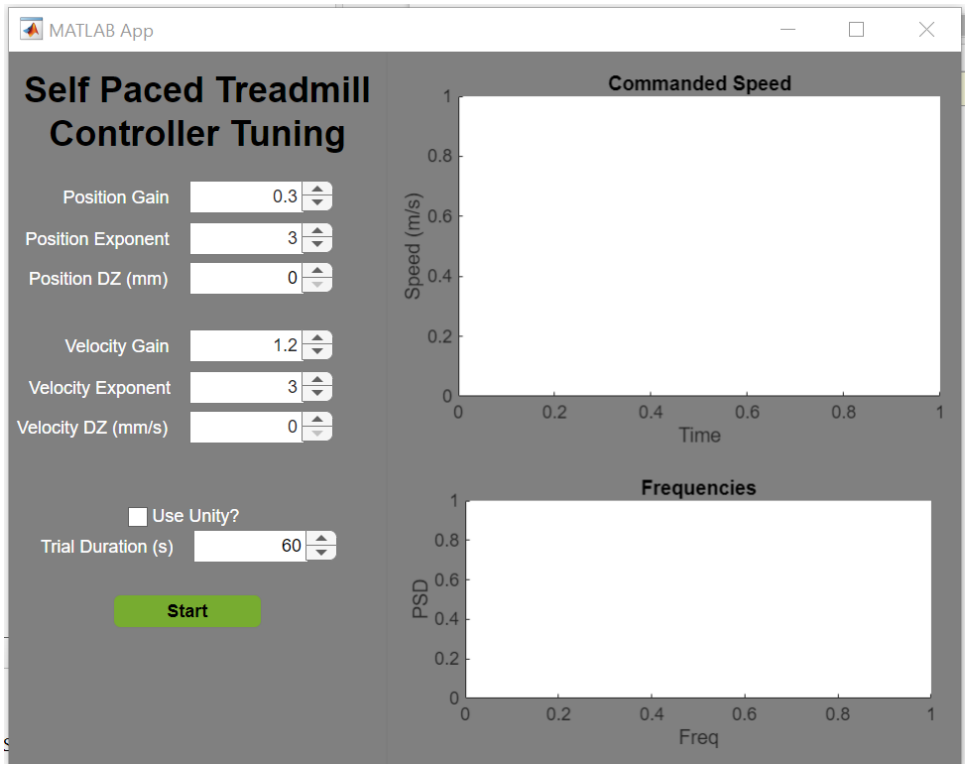

Figure 19: Algorithm tuning interface. Control parameters could be updated in real time by updating the values on the left. The treadmill commanded speed history is plotted in real time in the upper right, and every 10 seconds a frequency distribution plot is updated in the bottom right.

If position gain was too high relative to velocity gain, it would lead to large overshoots and oscillations, which we saw with the previous iteration of control algorithm that necessitated the velocity flag and use of the handrails. Conversely, a controller that over penalized velocity might be overdamped, could be too jerky and not adequately position the participant on the treadmill.

A defined position and velocity deadzone could help filter out some of the noise of natural variations in movement and from the motion capture sensing. Alternatively, raising the error terms to a power greater than one can create a naturally smooth deadzone, where small deviations are comparatively penalized much less than large deviations.

We tested with three participants (n = 3, 1 Female). The first two participants walked on the treadmill as the test conductor varied combinations of parameters and collected feedback. Their data was used to converge on a set of parameters. The third participant had no experience with self-paced treadmills, and we tested how a naïve participant might respond to the new algorithm.

We performed all tests with the VR headset on, walking in the virtual orchard with no rewards. The protocol started with participants walking at their preferred walking speed, attempting to hold a steady pace for 1 minute. Following the first minute, participants were asked to accelerate to their fastest comfortable speed and maintain it for 30 seconds. Following the high speed segment, participants were then asked to decelerate to a slow pace and maintain that speed for 30 seconds. Then participants were asked to accelerate back to their natural preferred walking speed. Following this 2-3 minute cycle, adjustments were made to the controller parameters at the discretion of the experimenter. If the set of parameters failed to stabilize the participant, then the protocol was paused, adjustments were made to the parameters, and the cycle started over.

We considered numerous measures when testing the controller: maximum excursions from the center of the treadmill, subjective ratings of stability, trust, and how natural the walking experience felt, absence of resonant frequencies, and smoothness of data. Ultimately, we converged on a set of parameters that were positively rated by the participants, and did not lead to any instances of grasping the handrails nor stepping off the edges of the treadmill.

5.6.2 Algorithm tuning results

We found that the following controller structure and gains allowed for a stable experience that minimized drastic overshoots and resulted in adequately smooth velocity traces. Utilizing the structure of Equation 35, the velocity and position errors were penalized cubically ($N_p = 3$ and $N_d = 3$), neither a position nor velocity deadzone was required, and gains were tuned to $k_p = 0.3$ and $k_d = 1.2$, resulting in the following:

$$V_{t+1} = V_t + k_p\left(p_{participant,t} - p_{treadmill\ center}\right)^3 + k_d {v_{participant,t}}^3 \qquad (36)$$

Interestingly, we found that a deadzone for position and velocity were not necessary, as the cubic relationship made for a gradually increasing penalty on position and velocity error. Additionally, the deadzone led to a more jagged velocity profile of treadmill speed, when the deadzone was exceeded, the velocity update would jump up. Though not always perceptible to the participant, it was not conducive to the necessary analyses of speed data.

Below, we show an example participant walking with the new algorithm, with the original algorithm using handrails, and original algorithm not using handrails (Figure 20). The new algorithm outperformed the original algorithm, with very little

overshoot and no instances of crossing the boundary of the treadmill. This is seen by the flatter lines and near rectangular patterns corresponding to speed changes by the participant. The old algorithm, when used without handrails, performed the worst. The controller chased the speed of the participant, causing them to step off the front and be sent off the back. Even when the participant was able to stay on the treadmill, it oscillated in a manner that was reportedly uncomfortable. The old algorithm with use of handrails performed with greater stability but required assistance. At times, the participant used the handrails to partially lift their weight and hold a constant position so the algorithm would settle on a pace. Collectively, we saw that performance with the new algorithm was a vast improvement.

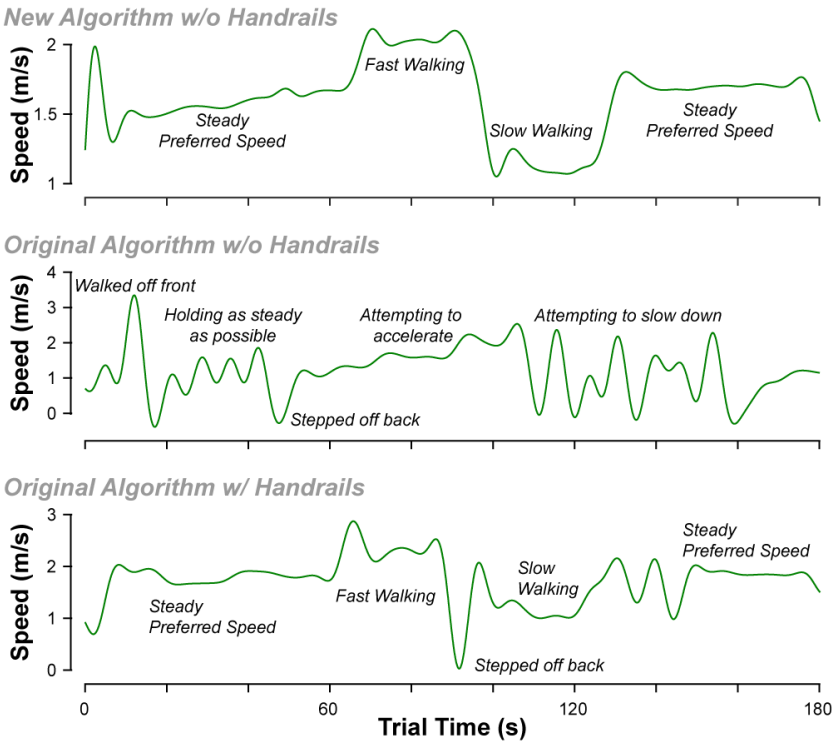

Figure 20: Comparison of new and old algorithms. Walking speed data for three minute trials, comparing different algorithms. Each trial consisted of a period of walking at their preferred pace, fast walking, slow walking, then their preferred pace again. *Top:* the new algorithm shows steady stable changes and can adapt to rapid speed changes from the subject. *Middle:* using the original algorithm and no handrails performed the worst, with the subject walking off the front and falling off the back. The algorithm oscillated and chased position, where walking steadily was difficult. *Bottom:* using the old algorithm but with handrails shows a much steadier experience, while it overshot some accelerations, utilization of the handrails enabled more stabilized walking than without.

5.7 Discussion

In this chapter, we summarized the pilot testing and development efforts to build an effective experimental setup that can probe at the effect of reward on walking speed. Results from our pilot data showed promise for our approach. Even with a small sample size, we see some effects of reward on walking speed. Here, we also detailed lessons learned from pilot testing – an aspect that often goes unpublished but could be useful for researchers exploring the same types of experimental approaches. Below, we summarize and discuss our major findings.

5.7.1 Immediate reward increases walking speed

In this study, we found a clear influence of immediate reward on walking speed. In locomotion, where energetics is thought to primarily dictate walking speed, we saw large changes in speed influenced by reward taking place on the order of seconds. The finding is quite novel, as no prior research has shown that reward could influence walking speed. This finding is in line with other movement modalities that show invigoration due to reward. Eyes saccade faster to more rewarding targets (Takikawa et al., 2002; Xu-Wilson et al., 2009) and reaches are performed more quickly to higher value targets (Summerside, Shadmehr, et al., 2018). It is intriguing that this same phenomenon generalizes to walking.

5.7.2 Effects of history of reward

We saw an influence of history of reward, but in the opposite direction we hypothesized. A history of high reward is hypothesized to further invigorate movement (Niv et al., 2007; Sukumar et al., 2024; Yoon et al., 2018). However, a positive RPE has also been shown to invigorate movement (Sedaghat-Nejad et al., 2019). In our dataset, we predominately observe the latter effect of RPE. Though this is only pilot data with a low number of subjects, it leaves us with questions about the magnitude of these effects, and when the invigorating effects occur in a movement.

First, the distribution of the environment might provide for a large effect of RPE that could overshadow an opportunity cost of time. The rarity of a high reward in a poor environment, and low reward in rich environment might influence actions more so than if these rewards were more frequent. Speculatively, opportunity cost

may be diminished because trial duration was too short, or perhaps because the only choice subjects are making is the speed they move, in comparison to foraging tasks where a stay-switch decision is also made. Regardless, this observation is interesting and warrants further investigation before an effect of RPE on walking speed vigor can truly be formalized.

5.7.3 Virtual reality increases walking speed

We found that walking with VR and reward was significantly faster than VR without reward, suggesting a motivational effect of reward on walking speed. Interestingly we also found that participants walked slightly faster in VR even without reward. Speculatively, this could be due to the reduced peripheral vision when wearing a VR headset impacting perception of optical flow (Banton et al., 2005), and the influence of optical flow on walking speed (Warren et al., 2001).

5.7.4 Self-paced, hands-free walking with a VR headset is possible

Despite the added gait variability and decreased visual sensory information that comes with walking with a VR headset, we found a control algorithm can keep persons stable while they walk at their own variable pace on the treadmill. Though many self-paced treadmill algorithms exist, none are marketed or verified as functional with a VR headset. Here, we openly share that data with the hope that future researchers can build on our findings.

5.7.5 Future Directions

Importantly, the ability to walk hands-free on the treadmill in VR enables numerous exciting experiments that would be less easily accomplished if hands were

constrained to the treadmill handrails. One avenue of experiments is testing participants carrying added mass, where they would not be able to use handrails for assistance. Perhaps more exciting is enabling actual reach and grasp interactions with the environment. Participants could make choices between rewarding patches by use of their hands, or this could even be used to create a truly realistic foraging task, where participants walk between patches, but gather rewards with their hands while stopped.

5.7.6 Conclusion

We demonstrated a feasible approach to probing at the influence of reward on walking speed. Leveraging the power of virtual reality coupled with a self-paced instrumented treadmill, we were able to create a naturalistic and immersive walking experience in the controlled laboratory setting.

CHAPTER VI

WALKING TOWARDS RICHES: REWARD PAYS THE COST OF EFFORT

6.1 Abstract

Slowing of movements is a symptom of numerous motor and psychiatric disorders, yet our understanding of what determines our movement speed is not well characterized. The preferred speed of walking is thought to be determined primarily by energetic cost, but can reward influence walking speed? Recent findings have shown that the speed at which we move reflects the value of what we hope to acquire in eye movements and arm reaches, suggesting a link between the neural processes that control movements and those that assign value. People reach and saccade faster to objects that promise greater reward and when there is greater opportunity cost of time. However, there is currently little understanding of how upcoming reward and history influence walking speed, a measure often thought to be prescribed by energetic cost alone. Here, we integrate a self-paced treadmill with a virtual reality system to immerse subjects in a realistic environment that probes the effects of reward, reward history, and effort on walking speed. Subjects completed two sessions with different baseline effort conditions. In the high effort session, subjects donned a weight vest with approximately 15% of their body mass. During each session, subjects completed a series of walking trials in virtual reality that involved walking along a path while collecting varied rewards visualized as apples. We manipulated the value of each reward, as well as the history of reward by changing the average payout of an

environment. We found that walking speeds increased as immediate reward value increased and were faster in a rich environment compared to a poor environment, suggesting there is an opportunity cost of time. In addition, we found that subjects walked slower in the high effort condition versus the low effort, confirming that movement speed considers the utility of a movement instead of reward alone. Our results build on the recent findings that both reward and baseline effort can modulate movement speed. These results show significant promise that walking speed can provide a non-invasive marker of the implicit value the brain assigns to the world around us.

6.2 Introduction

The pace at which we walk is thought to be determined primarily by energetic costs, but why might we run to greet a loved one and only walk to greet a stranger? If walking speed was determined only by effort expenditure, a person should not change their pace depending on who they are walking towards. However, recent findings have shown that the speed at which we move reflects the value we hope to acquire (Shadmehr et al., 2016, 2019), suggesting a neural link between the parts of the brain controlling movement and making valuation judgments. We reach faster (Sackaloo et al., 2015; Summerside, Shadmehr, et al., 2018) and move our eyes faster (Korbisch et al., 2022; Takikawa et al., 2002; Xu-Wilson et al., 2009; Yoon et al., 2020) to objects we value more, and even move faster to the same value in more rewarding environments (Sukumar et al., 2024; Yoon et al., 2018). Selection of how quickly to make a reach or a saccade appears to be a function of not only the effort the movement

will cost but the reward to be obtained. However, it is an open question whether preferred walking speed, a measure canonically linked to effort minimization alone, will be susceptible to reward-dependent modulation.

Humans and other animals tend to walk at speeds that minimize the metabolic cost of movement per unit distance, or cost of transport (Browning et al., 2006; Donelan et al., 2001; Hoyt & Taylor, 1981; Ralston, 1958; Selinger et al., 2015). When the metabolic cost of locomotion increases due to steeper inclines, the speed that minimizes cost of transport is slower. Correspondingly, walking speed slows overall (Minetti et al., 2002; Wickler et al., 2000) and people still choose a pace that minimizes cost of transport (Gidley & Lankford, 2021). The effect of carrying extra mass on preferred gait speed in humans is much less studied. Though carrying a load increases the overall energetic cost, curiously, it does not shift the walking speed that minimizes the energetic cost per unit distance (Bastien et al., 2005). Despite this, humans tend to walk slower than optimal at extreme loads (Hughes & Goldman, 1970). Though energetic cost makes some powerful predictions about preferred walking speed, recent investigations suggest there could be other influences.

Optimal foraging theory, rooted in ecology, traditionally explains how animals optimize their search for resources in a patchy environment by maximizing the rate of energy intake while minimizing energetic costs (Charnov, 1976). Recent research extends this theory to human movements, where similarly, a movement should aim to maximize the net utility over time, represented as the sum of the rewards minus energetic costs, all divided by time. To maximize the rate of reward, it follows logically

that reward should be captured as quickly as possible. However, to reduce the cost on time, movements must be faster leading to greater expenditure of energetic cost. The choice of how quickly to move can then be conceptualized as optimizing for a balance between the reward value, the cost of time, and the effort required to acquire the reward within the given time.

This formulation makes some interesting predictions, which have been observed in discrete human and animal movements such as reaches and saccades. First, we should move faster to higher immediately available rewards. (We use the term "immediate" here to contrast with "history," discussed in the next paragraphs). Indeed, movements toward more rewarding targets are performed with more vigor. Both humans and other animals saccade faster to targets when that target is more rewarding (Takikawa et al., 2002; Xu-Wilson et al., 2009). The vigor of arm reaches also increases with increased reward in the form of points and audiovisual stimuli (Summerside, Shadmehr, et al., 2018).

Opposite reward, varying levels of effort should also change movement vigor. For the same value of reward, one should choose the movement that requires less effort and move more quickly (Shadmehr et al., 2019). Indeed, when people are given the choice of two reach directions, they choose the direction that has a lower effective mass (Cos et al., 2011) and also perform that reach with higher velocity (Bruening et al., 2024; Gordon et al., 1994). Conversely, higher effort costs would predict decreases in vigor. Indeed, added effort in arm reaches results in slower reaction times (Rosenbaum, 1980) and longer movement durations (Bruening et al., 2024; Shadmehr

et al., 2016). If a person is carrying a heavy backpack, would they still run to greet their loved one, or perhaps would they choose to walk because the effort required is greater? Studying walking speed provides a unique opportunity to manipulate baseline effort naturally and continuously.

Movement vigor also depends on the background state of the environment in terms of average expected utility, i.e., a history of reward and a history of effort. The past experience of reward leads to an opportunity cost of time. Niv and colleagues were perhaps the first to explore ties to decision-making and movement vigor. In their model, they incorporated an opportunity cost on time, where time spent between reward opportunities was at odds with the effort spent acquiring potential future rewards (Niv et al., 2007). A history of higher rewards should lead to faster movements because the opportunity cost of time is greater. Conversely, a history of low rewards should lead to slower movements. A history of high effort, lowering the average utility and opportunity cost, predicts a slower pace. Conversely, a history of lower effort predicts a faster pace. Recently, this sensitivity to history of reward and history of effort has been observed in saccades (Yoon et al., 2018) and an arm-reaching foraging task (Sukumar et al., 2024). For the same upcoming reward, reaches and saccades were slower following a poor history of experiences (low reward or high effort), compared to a richer one (high reward or low effort).

A classic study measured preferred walking speed by pedestrians and found that those in large cities were twice as fast as those in smaller towns (Bornstein & Bornstein, 1976), and a follow-up study further dissected this phenomenon and found

that average income was the greatest predictor (Levine & Norenzayan, 1999). Anecdotally, based on these findings, preferred walking speed might change based on the opportunity cost of time.

In addition to forming an opportunity cost of time, a history of reward may also form an expectation for the amount or probability of reward received in the future. Studies in rats have shown that dopamine neurons encode this expectation of reward value before a reward is given, and following the reward (or no reward), dopamine neuron firing rates correlate to the difference between the reward received and the reward expected (Mohebi et al., 2019; Schultz et al., 1997). In essence, these dopamine neurons are reflecting a reward prediction error (RPE). This has also been seen in monkey studies, where if a collected reward is much larger than expected, many neurons fire, but if what is received matches expectations, the neurons do not fire (Bayer & Glimcher, 2005; Schultz et al., 1997). Only recently has RPE been shown to influence movement vigor. In a human study, saccades and reaches were made faster when there was a greater positive RPE (a higher value than expected) and more slowly with a negative RPE (a lower value than expected) (Korbisch & Ahmed, 2022; Sedaghat-Nejad et al., 2019). No study has shown RPE influencing walking speed.

Here, we developed a virtual reality walking system that probes the effect of reward and effort on walking speed. Subjects walked at their own pace through a virtual environment collecting monetary rewards symbolized as apples, where we manipulated the value of rewards and frequency that different values occur, as well as the baseline effort of walking. Our hypothesis is that the desire to maximize net

rate of reward influences the choice of preferred walking speed. Thus, three economic variables would affect movement vigor: immediate reward, the history of reward, and baseline effort. We predicted that the presence of immediate reward would increase walking speed proportional to the reward value. Second, we predicted that a history of higher reward would also increase walking speed, even when immediate reward was not present. Finally, we expected that higher baseline effort would reduce walking speeds.

6.3 Methods and materials

6.3.1 Participants

Sixteen healthy young adults (n = 16; 7 Females; Age: 22.4 ± 4.95 yrs) participated in this study. Subjects were required to be between 18 and 35 years of age, able to understand English, and with vision normal or corrected to 20/20. Exclusion criteria included any major musculoskeletal or ophthalmological impairments, low or sedentary activity level (self-reported), any previously experienced side effects from VR including motion sickness (self-reported), any neurological or vestibular diseases, and any medications that could impair the ability to complete the walking task. All subjects provided written informed consent, and the experimental protocol was approved by the Institutional Review Board of the University of Colorado Boulder.

6.3.2 Experimental setup

Subjects walked on a self-paced treadmill through a virtual environment. In addition to compensating them for their participation, they received $0.02/reward

collected to incentivize them to collect as many rewards as possible. The virtual reality walking system integrated numerous components to enable an immersive, safe, self-paced walking experience. We used an Oculus Quest VR Headset, a Vicon motion capture system, a Bertec force-instrumented treadmill, and a virtual environment built in Unity, all controlled with custom MATLAB scripts.

Throughout the experiment, subjects wore a position marker belt fitted to their waist and positioned just superior to the sacrum (Figure 21A). Subjects also donned a Polar Verity optical heart rate monitor, which was worn snuggly on their upper right arm. During trials with high effort, subjects wore a weight vest (Fitness Gear, Dick's Sporting Goods) with approximately 15% of their body mass, rounded down to the nearest 2.5-lb increment (actual average: 14.4 ± 0.500%, mean ± s.d.).

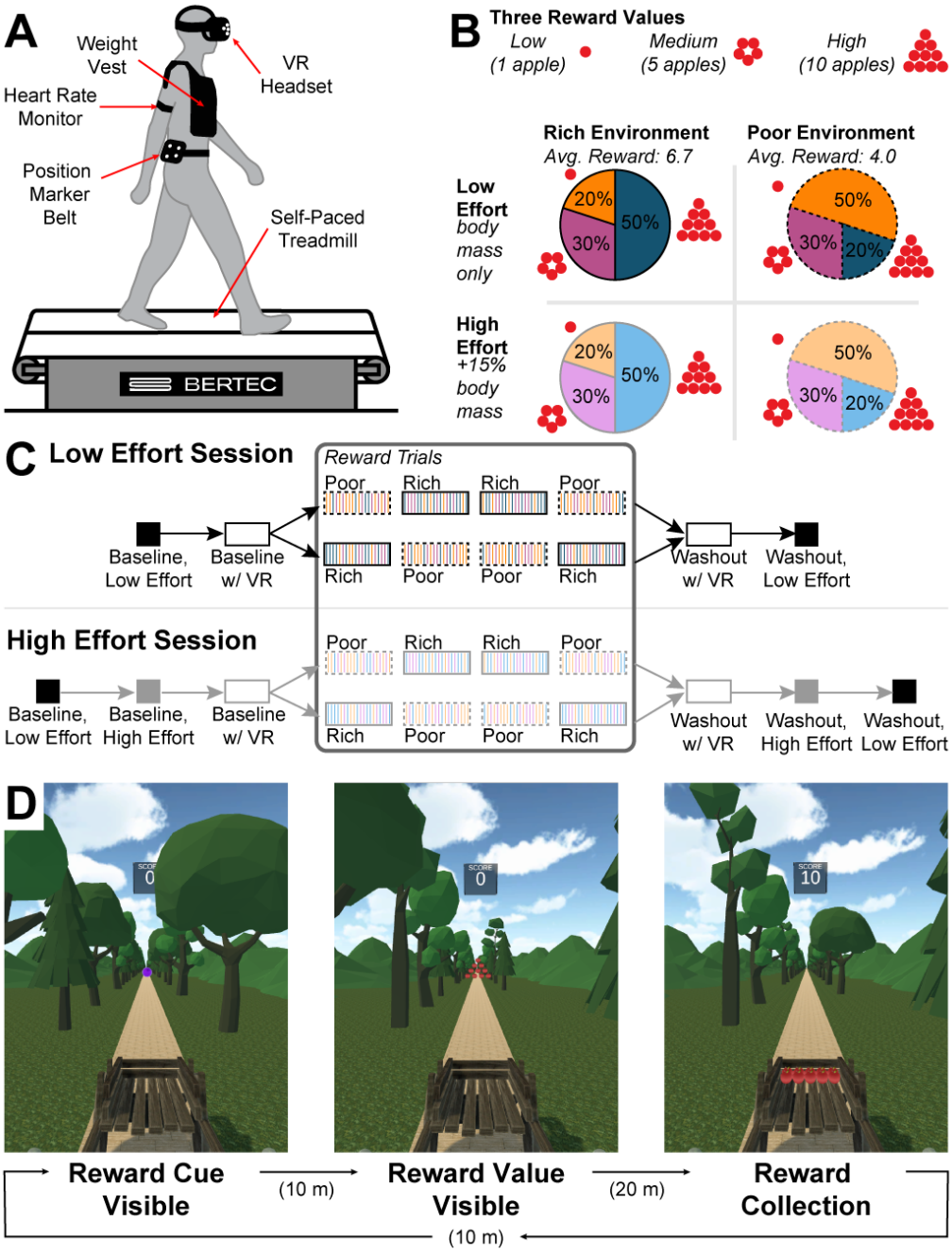

Figure 21. Experimental setup and protocol. *A*: subjects walked on a self-paced treadmill while wearing a VR headset and heart rate monitor. In high effort trials, subjects donned a weight vest weighing approximately 15% of their body mass. The subject's position was tracked using a position marker belt. *B*: within a reward trial, the subject saw three types of reward values: high (10 apples), medium (5 apples), and low (1 apple). Subjects experienced different combinations of reward environments and baseline effort conditions. Low effort conditions are shown in bold/dark colors, and high effort conditions are shown in light colors. Rich environments are shown with solid lines, and poor, with dashed lines. *C*: the low effort session consisted of eight walking trials, and the high effort session consisted of ten walking trials, with the middle four being reward trials. *D*: Each reward trial followed the same repeating flow of cue appearance, reward reveal, and reward collection.

6.3.2.1 Position tracking and self-pacing

Subject waist position and headset position were tracked throughout a trial using Vicon cameras and Vicon Tracker. Motion capture markers installed on the VR headset tracked head position in the global coordinate frame and were read into Unity to translate and rotate the virtual environment on the VR headset displays. The subject's position was used to determine the forward speed at which the subject was moving through the environment. This position was read by the self-pacing algorithm to update the speed of the treadmill, which was, in turn, read by Unity to translate the environment forward for a given timestep and speed.

The self-paced treadmill utilized a custom control algorithm that enabled subjects to vary their speed naturally throughout a trial. The algorithm and setup are based loosely on previous work from Pimentel et al. 2022 and Song et al. 2020. (Pimentel et al., 2022; Song et al., 2020). The position tracker, consisting of motion capture markers mounted to a belt and positioned just superior to the sacrum, tracked the subject's position in global coordinates. The position of the subject relative to the center of the treadmill was read into the controller and also differentiated to find the subject's velocity relative to the treadmill. We used a proportional-derivative controller that cubically penalized both the subject distance from the center of the treadmill, $p_{subject}$, and speed, $v_{subject}$, (Equation 37). A positive difference from treadmill center was further forward on the treadmill, and a positive velocity value was forward motion relative to the treadmill. The treadmill speed, $v_{treadmill}$, was updated each timestep, which ran at approximately 10 Hz.

$$v_{treadmill,t} = v_{treadmill,t+1} + k_p\left(p_{subject} - p_{treadmill\ center}\right)^3 + k_d {v_{subject}}^3 \qquad (37)$$

Through pilot testing we tuned control constants and exponents and found that k_p = 0.3 and k_d = 1.2 led to a stable and comfortable experience for participants.

Qualitatively, if a subject sped up and moved forward on the treadmill, the treadmill belts would increase speed and ultimately bring them back to the center of the treadmill. If a subject slowed down, traveling backward on the treadmill, the treadmill belts would slow down, again letting them move forward to the center of the treadmill.

The development of this controller is noteworthy, as it enabled subjects to walk without using handrails for assistance and was also responsive enough to prevent subjects from walking off either end of the treadmill, all while wearing the VR headset with a limited field of view and blind to the real world. Hands-free walking was important to the experiment design for two reasons. First, using the handrails could act as a feedback mechanism for a subject changing speed, though they might be unaware of it otherwise. Second, the handrails could be used to assist in suspending some of the added weight and diminishing the effect of the high effort condition.

6.3.2.2 Virtual Environment

The virtual environment was built in Unity 3D and mimicked a virtual orchard. We constructed a simulated apple "foraging" task, where subjects collected varying quantities of apples, corresponding to rewards, as they walked along the virtual orchard. The virtual orchard was an infinitely long tree-lined pathway. In

front of the subject, a virtual wheelbarrow traveled along the path with the subject and contained the virtual apples as the subject accumulated them. In addition, the virtual environment included semi-transparent, accurate representations of the treadmill belts and handrails. If subjects needed to grab the handrails, they could quickly find them while keeping the VR headset on. The virtual environment also displayed a scoreboard overhead that counted the total number of apples collected within that trial. With 10 seconds left in a trial, the scoreboard displayed a countdown timer. The virtual world was three-dimensional, and the subject's position was accurately represented within the world, including small lateral movements on the treadmill. However, we designed the task to work with the unidirectional nature of the treadmill.

A variety of sounds were played during a trial. In the background, ambient forest noises played to enhance immersion and obscure some of the treadmill and fan noise. Pleasant tones were played after each reward was collected, and the tones varied based on the value of the reward. In addition, an audio alert played when the 10-second trial countdown timer began.

Each reward trial consisted of nine minutes of walking. As the subject walked, they encountered different reward opportunities, (i.e., apple quantities). A reward opportunity could be one of three values: High/10-apples, Medium/5-apples, and Low/1-apple (Figure 21B). In addition, we varied the quality of the rewarding trial environment between "rich" and "poor." In a rich environment, reward opportunities were randomly distributed as 50% High/10-apples, 30% Medium/5-apples, and 20%

Low/1-apple; in a poor environment, reward opportunities were randomly distributed as 20% High/10-apples, 30% Medium/5-apples, and 50% Low/1-apple. These distributions led to an average reward opportunity value of 6.7 apples in the rich environment and 4.0 apples in the poor. The actual distribution of the rewards was 6.71 ± 0.225 (mean ± s.d.) in the rich environment and 4.00 ± 0.205 (mean ± s.d.) in the poor environment.

Each reward trial consisted of repeating cycles of reward events spaced evenly every 40 meters (Figure 21D). Each cycle began with the subject walking along the path. After subjects walked 10 meters, a purple orb "cue" would appear, signaling the location of an upcoming reward opportunity 30 meters ahead, without revealing its value. After subjects walked 10 meters closer, the cue disappeared and revealed the value of the reward they could collect (High, Medium, or Low). After walking the final 20 meters, subjects collected the reward by simply walking through the apples. Reward collection was not performance dependent; as long as the subject continued walking, the reward was collected. Upon collection, a pleasant sound was played that varied based on the value of the reward. The collected apples appeared in the wheelbarrow, and the score (the total number of apples) was updated overhead. After collection, the next cycle immediately began. The subject was not required to stop or slow down. This pattern repeated until time expired for the trial.

6.3.3 Protocol

Subjects completed two sessions: one tested walking with low effort, and the other, with high effort. Each session was conducted on separate days, spaced no less

than 24 hours apart (mean ± s.d., 3.66 ± 2.79 days). Both sessions followed a similar flow, visualized in (Figure 21C). The low effort session consisted of eight walking trials, and the high effort, ten trials. Trials in each session varied between three, six, and nine minutes per trial, and varied in configuration (VR headset, rewards, and/or weight vest). In the first session, subjects were familiarized with the equipment, including the VR headset, position marker belt, heart rate monitor, weight vest, and virtual environment. They were instructed to "collect as many apples as possible," and were not told about the different reward environments nor the randomization and frequency of rewards.

At the beginning of each session, subjects donned the heart rate monitor and motion tracking belt. Next, they completed one quiet, 90-second stand to measure resting heart rate. Following the quiet stand, subjects completed one 3-minute baseline trial without wearing the VR headset or the weight vest (Baseline, Low Effort). Subjects were instructed to explore how the treadmill changed speeds as they changed their own walking speed for up to the first 60 seconds, then settle into a comfortable walking pace. In the high effort session, subjects donned the weight vest and completed an additional 3-minute baseline trial (Baseline, High Effort).

Next, subjects completed a 6-minute familiarization trial wearing the VR headset (Baseline w/VR). Subjects walked in the virtual environment, but did not receive any cues or rewards. The next four trials incorporated the reward opportunities (reward trials) and were the main interest of the study. Each reward trial was 9 minutes long and varied what rewarding environment they were in. The

order of the environments was either rich-poor-poor-rich or poor-rich-rich-poor to balance any potential order effects.

Each session ended with washout trials that were identical to the baseline trials but opposite in order. Subjects completed a 6-minute trial, walking in the virtual environment with no cues or rewards (Washout, w/VR). In the high effort session, subjects completed an additional 3-minute washout trial without the VR headset with the weight vest on (Washout, High Effort). In both sessions, the final trial was a 3-minute washout trial without the VR headset and no weight vest on (Washout, Low Effort).

Following completion of all walking trials, subjects doffed the remaining equipment (the heart rate monitor and motion tracking belt).

Subjects were pseudo-randomly placed into one of eight groups to counterbalance any order effects in the study. Each group was a unique combination of which order they experienced the high and low effort sessions, and which order they experienced the rewarding environments in each of the sessions.

Each walking trial followed a specific series of events. Subjects started and ended each walking trial with hands on the handrails and were given a countdown until the treadmill started moving and again as the treadmill was about to stop. Each walking trial started at a speed of 1 m/s but quickly adjusted to the subject's preferred speed. After each walking trial, subjects reported a rating of perceived exertion (Borg RPE) and took a short break of approximately three minutes.

6.3.4 Data collection and processing

For each subject, we collected demographic and anthropometric information, including age, sex, height, weight, leg length, and previous experience with virtual reality headsets.

Our primary outcome measure was walking speed, which was inferred from the treadmill speed, recorded at 10 Hz. For each walking trial, we threw out the first and last 10 seconds, as this was when the subject was likely accelerating or decelerating and was free to use the handrails. Walking acceleration was calculated by differentiating walking speed data and filtering using a 4^{th}-order Butterworth filter.

Heart rate data was measured at 1 Hz using the Polar Verity, and ground reaction forces were acquired from the treadmill via Vicon Tracker at 1000 Hz. The state of the virtual environment was recorded at 10 Hz. These data included subject position and speed in the environment as well as specific events such as when a cue or reward was visible. Importantly, the time stamps of these events were used to parse the walking speed data to investigate specific windows of time.

The Borg Rating of Perceived Exertion (Borg RPE), a numerical evaluation of subjective effort with values ranging from 6 to 20, was solicited from subjects immediately following the completion of each walking trial (Borg, 1982).

The total time for each session and the total number of apples collected for each rewarding trial were recorded and used for subject compensation.

All data collected were analyzed and processed using custom scripts and built-in functions in MATLAB.

6.3.5 Statistical analyses

6.3.5.1 Effects of immediate reward, history of reward, and baseline effort

We employed linear mixed-effects models to investigate the effects of reward and effort on walking speed, time, and acceleration across various phases. For each model, we represented reward value as the specific number of apples (1, 5, or 10) and the environment value as the average number of apples for a given environment (Rich: 6.7; Poor: 4.0). We investigated the fixed covariates and interactions of immediate reward value (*RewardValue*; continuous), environment (*Environment*, continuous), effort condition (*Effort*, categorical, "high" or "low"), session number (*Session,* continuous, 1 or 2), trial time (*Time*, continuous) and *Subject* (categorical) as a random intercept term. The model is summarized in Equation 38 below, where $Speed_{ij}$ is the jth reward instance for subject i, and $Subject_i$ is normally distributed with mean 0 and variance σ^2.

$$\begin{gathered}
\mathrm{Speed}_{ij} \sim N(\mu_{ij}, \sigma) \\
E[\mathrm{Speed}_{ij}] = \mu_{ij} \\
\begin{aligned}
\mu_{ij} = {} & \mathrm{RewardValue}_{ij} + \mathrm{Environment}_{ij} + \mathrm{Effort}_{ij} \\
& + \mathrm{Session}_{ij} + \mathrm{Time}_{ij} + \mathrm{RewardValue}_{ij} \times \mathrm{Effort}_{ij} \\
& + \mathrm{Environment}_{ij} \times \mathrm{Effort}_{ij} + \mathrm{Subject}_i
\end{aligned} \\
\mathrm{Subject}_i = N(0, \sigma^2)
\end{gathered} \tag{38}$$

This structure and all factors were consistent across the models we tested investigating the effects of immediate reward, history of reward, and effort. We generically represented the dependent variable as *Speed* in the equation above; however, this also could be replaced by *Acceleration* or *Time,* depending on the

variable of interest. We looked at average values within specific epochs of interest. For instance, we looked at average speed across a small window (1 meter before and 1 meter after) surrounding each event shown in Figure 21D (cue appearance, reward value reveal, and reward collection). We also looked at average speed and acceleration across each of three major segments. First, we were interested in the segment between cue appearance and reward value reveal, where we predicted to see an effect of the environment and the effort condition, but not an effect of the reward value which had not yet been revealed. Second, we analyzed the segment between reward value reveal and reward collection, where we predicted to see an effect of immediate reward value as well as environment and effort condition. Finally, we looked at the segment between reward collection and the next cue appearance, where we expected to see residual effects of the previously collected reward as well as the environment and effort condition.

In addition, we were interested in an alternative formulation of history of reward. Instead of using environment as a predictor, could the actual history of rewards influence speed? Building on the linear model in Equation 38, we replaced the single value for *Environment* with new factors representing explicitly received rewards. In the model below, *PrevNRewardValue* represents the value of the *N*th previous reward. In our model, we tested the four previous rewards, resulting in the following linear model:

$$
\begin{gathered}
\text{Speed}_{ij} \sim N(\mu_{ij,}\sigma) \\
E[\text{Speed}_{ij}] = \mu_{ij} \\
\mu_{ij} = \text{RewardValue}_{ij} + \text{Prev1RewardValue}_{ij} + \text{Prev2RewardValue}_{ij} \\
+ \text{Prev3RewardValue}_{ij} + \text{Prev4RewardValue}_{ij} + \text{Effort}_{ij} \\
+ \text{Session}_{ij} + \text{Time}_{ij} + \text{RewardValue}_{ij} \times \text{Effort}_{ij} \\
+ \text{Environment}_{ij} \times \text{Effort}_{ij} + \text{Subject}_{i} \\
\text{Subject}_{i} = N(0, \sigma^2)
\end{gathered}
\tag{39}
$$

6.3.5.2 Acceleration and reward prediction error

For this analysis, we took a closer look at acceleration following reward value reveal rather than the gross average over the segment between reward value reveal and collection. When does acceleration correlate to reward prediction error? We calculated averages of acceleration over 2-meter bins, from 0-2 meters following reveal, 2-4 meters, etc., until 18-20 meters at which the reward was collected. These averages were then input as predictors into a linear mixture model to see if the acceleration value could predict reward prediction error. We modeled reward prediction error, "RPE", as the difference between the revealed reward value and the average reward for an environment. We used the term "RPE" here to represent reward prediction error and distinguish it from the Borg rating of perceived exertion, which we refer to as "Borg RPE". For a high reward of 10 apples, the RPE in a rich environment would be +3.3, but for a poor environment, the same high reward would have an RPE of +6.0. In essence, the high reward in a poor environment was more surprising, thus the higher RPE. The linear model we tested was as follows:

$$
\begin{gathered}
RPE_{ij} \sim N(\mu_{ij,}\sigma) \\
E[\mathrm{RPE}_{ij}] = \mu_{ij} \\
\mu_{ij} = \mathrm{Accel0to2}_{ij} + \mathrm{Accel2to4}_{ij} + \mathrm{Accel4to6}_{ij} + \mathrm{Accel6to8}_{ij} \\
+\mathrm{Accel8to10}_{ij} + \mathrm{Accel10to12}_{ij} + \mathrm{Accel12to14}_{ij} + \mathrm{Accel14to16}_{ij} \\
+\mathrm{Accel16to18}_{ij} + \mathrm{Accel18to20}_{ij} + \mathrm{Subject}_i \\
\mathrm{Subject}_i = N(0, \sigma^2)
\end{gathered}
\tag{40}
$$

6.3.5.3 Comparison of measures of effort

To investigate how different measures of effort varied over the course of the experiment and between each other, we compared walking speed, heart rate, and Borg RPE. A subject could control their level of effort expenditure by choosing a walking speed. The subject's heart rate was an objective measure of effort expenditure, and Borg RPE was a subjective measure of how much effort the subject felt they were expending. We looked for differences in the averages of these measures over each trial type in the experiment (*TrialType,* categorical) and effort condition (*Effort,* categorical). We tested this using a linear mixed model below:

$$
\begin{gathered}
\mathrm{EffortMeasure}_{ij} \sim N(\mu_{ij,}\sigma) \\
E[\mathrm{EffortMeasure}_{ij}] = \mu_{ij} \\
\mu_{ij} = \mathrm{TrialType}_{ij} + \mathrm{Effort}_{ij} + \mathrm{Subject}_i \\
\mathrm{Subject}_i = N(0, \sigma^2)
\end{gathered}
\tag{41}
$$

6.3.5.4 Effects of virtual reality on walking speed

We tested for effects of high-level experimental configurations and phases on gross averages of walking speed rather than specific event windows. Again, we tested for these effects using a linear mixed effects model. We grouped each trial type into configuration-specific categories: No VR, VR without reward, and VR with reward (categorical, *Configuration*). Baseline and washout trials without VR are grouped together into the first category, baseline and washout trials with VR are grouped into

the second category, and all rewarding VR trials, regardless of reward environment, are grouped into the third and final category. We tested the effect of configuration using the following linear mixed model:

$$\begin{aligned} \text{Speed}_{ij} &\sim N(\mu_{ij,}\sigma) \\ E[\text{Speed}_{ij}] &= \mu_{ij} \\ \mu_{ij} &= \text{Configuration}_{ij} + \text{Effort}_{ij} \\ &\quad + \text{Configuration}_{ij} \times \text{Effort}_{ij} + \text{Subject}_i \\ \text{Subject}_i &= N(0, \sigma^2) \end{aligned} \tag{42}$$

6.3.5.5 Effects of session number

The effect of session number is a factor in the linear models we tested using Equation 38. In addition to these tests, we took a more direct approach to measure the effects of session number. We compared the average walking speed for the very first and very last trial of each session: the baseline trial and washout trial conducted without the VR headset and no weight vest. These were performed in both Session 1 and Session 2, regardless of whether it was the high effort or low effort session. The baseline trials were the essential comparison for observing an effect of session (i.e., if subjects were significantly faster on the very first trial in Session 2 versus the first trial of Session 1). We also included washout trials to better characterize speed changes: were subjects getting faster within a Session (baseline versus washout), was there a difference in speed between the end of Session 2 and beginning of Session 1 (washout Session 1 vs baseline Session 2), and did subjects change speed differently in Session 2 (baseline versus washout, Session 1 versus Session 2)? We tested the effect of phases (*Phase,* categorical: baseline versus washout) and employed the following linear model to test the effect of session:

$$\begin{gathered} \text{Speed}_{ij} \sim N(\mu_{ij,}\sigma) \\ E[\text{Speed}_{ij}] = \mu_{ij} \\ \mu_{ij} = \text{Phase}_{ij} + \text{Session}_{ij} + \text{Subject}_i \\ \text{Subject}_i = N(0, \sigma^2) \end{gathered} \tag{43}$$

6.3.5.6 Effects of time within a trial

We examined the effect of time within a trial in the numerous linear models tested using Equation 38. We supplemented these findings by directly investigating differences in speed that were agnostic to the event-focused approach from before. We analyzed the effect of time in the reward trials (the middle four trials of each session) by averaging the speed in the first and last minute of each reward trial. We represented the first minute and last minute as a categorical variable, *StartOrEnd,* and tested these effects on average walking speed using the following linear model:

$$\begin{gathered} \text{Speed}_{ij} \sim N(\mu_{ij,}\sigma) \\ E[\text{Speed}_{ij}] = \mu_{ij} \\ \mu_{ij} = \text{StartOrEnd}_{ij} + \text{Environment}_{ij} + \text{Effort}_{ij} + \text{Subject}_i \\ \text{Subject}_i = N(0, \sigma^2) \end{gathered} \tag{44}$$

Additionally, we sought to understand how prolific the influence of time was and what might correlate to the magnitude of speed change. For instance, was speed change specific to reward trials, or perhaps an effect of virtual reality? To perform this analysis, again, we took the average walking speed from the first minute and the last minute of every trial to find the difference (*ΔSpeed*). We fit a linear mixed model to trial duration (*TrialDuration*: 3, 6 or 9 min), effort condition, and session number, per the equation below. Due to experiment design, trial duration happened to covary with configuration (all baseline and washout trials without VR were three minutes, baseline and washout with VR were six minutes, and reward trials were nine

minutes), so it was impossible to test these effects separately. However, we did run the same statistical test but replaced trial duration with trial type as a categorical variable (categorical: baseline w/VR, baseline w/o VR, rich reward, poor reward, washout w/VR, and washout w/o VR). This approach further separates data (i.e., six trial types instead of three trial duration values) to observe if a specific trial type observed a change in speed.

$$\begin{gathered}\Delta\text{Speed}_{ij} \sim N\left(\mu_{ij,}\sigma\right) \\ E\left[\Delta\text{Speed}_{ij}\right] = \mu_{ij} \\ \mu_{ij} = \text{TrialDuration}_{ij} + \text{Effort}_{ij} \\ +\text{TrialDuration}_{ij} \times \text{Effort}_{ij} \\ +\text{SessionNum}_{ij} + \text{Subject}_{i} \\ \text{Subject}_{i} = N(0, \sigma^2)\end{gathered} \tag{45}$$

6.4 Results

6.4.1 Model predictions

First, we built a dynamical model to predict how walking speed might change based on immediate reward value, a history of reward, and baseline effort. Foundational studies have found that time discounts reward hyperbolically (A. M. Haith et al., 2012; Jimura et al., 2009; Kobayashi & Schultz, 2008; Mazur et al., 1987; Myerson & Green, 1995; Shadmehr, Orban de Xivry, et al., 2010) where a reward of value R, is discounted by a temporal sensitivity factor, γ, and the delay time, T:

$$R_{discounted} = \frac{R}{1 + \gamma T} \tag{46}$$

The utility of a decision or action is represented as the reward at stake subtracted by the effort required to achieve the reward (Shadmehr et al., 2016). We combine this into a single utility collectively discounted by time:

$$U = \frac{R - E}{1 + \gamma T} \tag{47}$$

Effort, in our case, is best represented by the metabolic cost of walking. Ralston 1958, Bobbert 1960 and numerous others have shown that the metabolic rate varies quadratically as a function of average walking speed, v (Equation 48, Bobbert, 1960; Ralston, 1958).

$$\dot{E} = a + bv^2 \tag{48}$$

We propose that the speed with which a person moves should optimize the utility equation above, where, for a given distance, the optimal time can be determined and, thus, the average movement speed. We convert the model of rate of energy expenditure into total energy and represent movement speed in terms of distance and time, resulting in the following equation:

$$E = amT + \frac{bmD^2}{T} \tag{49}$$

We combine Equations 48 and 49 and replace walking speed with distance, D, and time, T. The utility equation becomes:

$$U(d) = \frac{kR - (amT + \frac{bmD^2}{T})}{1 + \gamma T} \tag{50}$$

To find the optimal movement speed for a given reward value, R, and distance to reward, D, we find the optimal time, T^*, that maximizes the utility. The optimal average movement speed is simply the distance to reward divided by that time.

Finally, we add an opportunity cost that represents a cost on time that varies with the average rate of reward. We model this cost by solving for Equation 50 where

R is the average reward for a given environment (rich or poor). In our case, we have four distinct opportunity costs, combining the two reward environments, rich and poor, and the two effort conditions (represented by additional mass, m, in the case of high effort). We set the distance equal to the average distance to the reward (a constant 40 m in our protocol). We find the movement time that maximizes utility, then we take the maximum utility and divide it by the optimal time to arrive at an average utility rate. We add the opportunity cost term to the utility term in Equation 50 to arrive at our final formulation of utility:

$$U(d,T) = \frac{k_1 R(d) - (amT + \frac{bm(D_{rew} - d)^2}{T})}{1 + \gamma T} - k_2 \dot{U}_{avg} T \tag{51}$$

In the equation above, we add a new constant and delineate between k_1 and k_2, where the former still scales reward value relative to effort, and the latter scales opportunity cost relative to the baseline utility. Reward value, $R(d)$, is modeled as a function of distance throughout the trial. The reward function is zero unless the reward value has been revealed. We do not associate a reward value when only the cue is visible.

To generate predicted walking speeds throughout a reward cycle, we propagate our model forward in time, starting at a distance of zero (d = 0m) and ending once the total trial time is nine minutes. For each timestep, the optimal time and corresponding speed are recalculated based on whether a reward was visible, the value of that reward, the distance to that reward, and the reward environment (i.e., history of reward). From this simple model, we make predictions about how a subject

might respond differently to different reward values, effort conditions, and history of rewards (Figure 22).

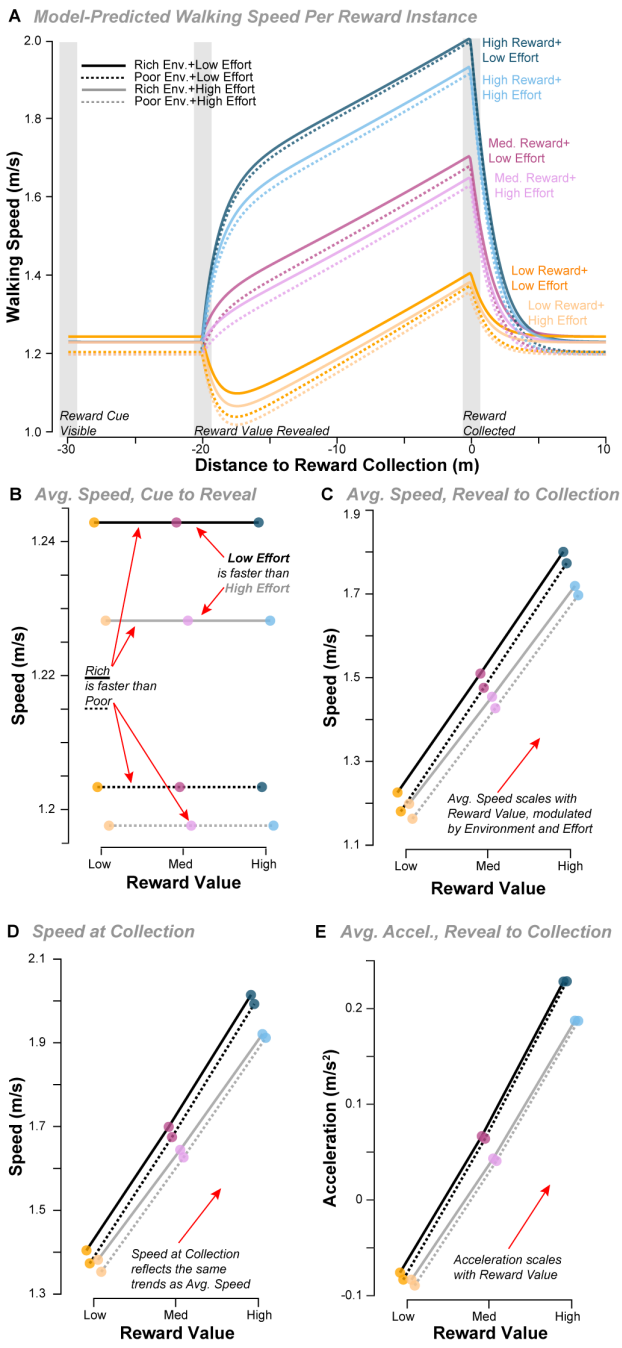

Figure 22: Model predictions. *A*: predicted walking speed for one reward cycle (reward cue through the next cue) as a function of distance to reward (mean with shaded standard deviation). Immediate reward values are shown in blue for high, red for medium, and yellow for low. A rich environment is a solid line, and a poor environment is dashed. The high effort condition is shaded light, and the low effort condition is shaded dark. Key events are highlighted with vertical gray bars. *B*: the model predicts speed differences due to effort conditions and environment on the average speed between cue and reward value reveal *C*: following reward reveal, speed is modulated by the immediate reward value, the environment, and the effort condition. *D*: inspects the instantaneous speed at reward collection, which reflects the same trends as *C*. *E*: acceleration following reward reveal is affected by the reward value and the effort condition, while the reward environment has a reduced effect.

To test our model, we tuned the constants (k_1 = 1000, k_2 = 0.2, and γ = 0.02) and assumed a subject mass of 72 kg in the low effort condition. Importantly, qualitative model predictions held across a wide range of parameter values. We then propagated the model forward in time with rewards randomized and evenly spaced as in the actual protocol. We generated ten trials, each nine minutes long. Within a trial, rewards were randomly ordered based on the distribution of the environment. Each of these reward instances was then parsed into spans of 40 m before collection (the instant of the previous reward being collected) to 10 m after the current reward was collected. The predicted walking speeds for each combination of reward value, effort condition, and environment were interpolated to fixed distance intervals of 0.1 meters and then averaged. The mean walking speed profile for each combination is plotted in Figure 22A.

The model makes interesting predictions. Prior to reward reveal, walking speed reflects both the quality of the environment and the effort condition. Average walking speed between cue and reveal (Figure 22B) is faster in a rich environment than in a poor environment and slower in a high effort condition than a low effort condition. Perhaps it is obvious, but before the immediate reward value is revealed, the reward value does not affect walking speed.

Once the reward value is revealed, walking speed and acceleration are modulated by the value of the reward. Average speed between reveal and collection scales with reward value (Figure 22C), but appears to also be influenced by effort condition and environment, where average speeds are faster in the rich environment

and faster in a low effort condition. Ultimately, speed is higher at reward collection for higher reward values and exhibits the same effects as average speed regarding environment and effort condition (Figure 22C and D). Importantly, we also see persistent effects of the environment and effort condition across all measured speeds (Figure 22B-D), where walking speeds to collection (Figure 22C and D) are faster in a rich environment and slower in a poor environment as well as faster in a low effort condition than a high effort condition. Thus, for the same immediate reward value, walking speed is nonetheless influenced by the history of reward (rich or poor) and effort condition (high or low). Interestingly, walking acceleration is greater for higher rewards (Figure 22E) but appears to be modulated by effort condition, where the high effort condition accelerates less. Though difficult to tell from model predictions, there may be a small influence of environment, where acceleration is higher in the rich environment than in the poor. Nevertheless, the average acceleration is greater for higher immediate reward, as anticipated by the predicted higher average speeds (Figure 22C).

This simple model allows us to make predictions about what we might see in the data, as well as provide a possible explanation if these predictions were confirmed. First, walking speed should increase as immediate reward value increases. Second, walking speed should increase overall due to a history of high rewards, or rather, in a rich environment, as represented by an opportunity cost of time. An alternative model without opportunity cost would not predict a difference

between environments at cue, reveal, or collection. Finally, our model predicts that walking speed should decrease with added effort.

6.4.2 The effects of reward and effort on walking speed and acceleration

Within a given trial, subjects were free to vary their speed. An example of walking speed variance is shown in the three-minute snippet in Figure 23A. We separated each rewarding instance, anchoring them in time or distance from one of the events shown in Figure 21D (reward cue, reward value reveal, and reward collection). These speeds were separated by effort condition, reward environment, and immediate reward value (Figure 23B and C). We normalized each subject's walking speed to their average speed during the baseline trial without the weight vest and no VR headset at the beginning of the session. Plotted by distance from reward collection, some patterns begin to emerge. On average, across subjects, we saw that walking speed in the high effort condition was slower than in the low effort condition. Before the reward value was revealed, rich environments were faster than poor environments, and following reward value reveal, we saw a change in speed proportional to the reward value where high was faster than medium and low. Investigating normalized subject averages versus distance showed the continuity and variability in the speed data and further validated statistical findings summarized in the subsequent sections.

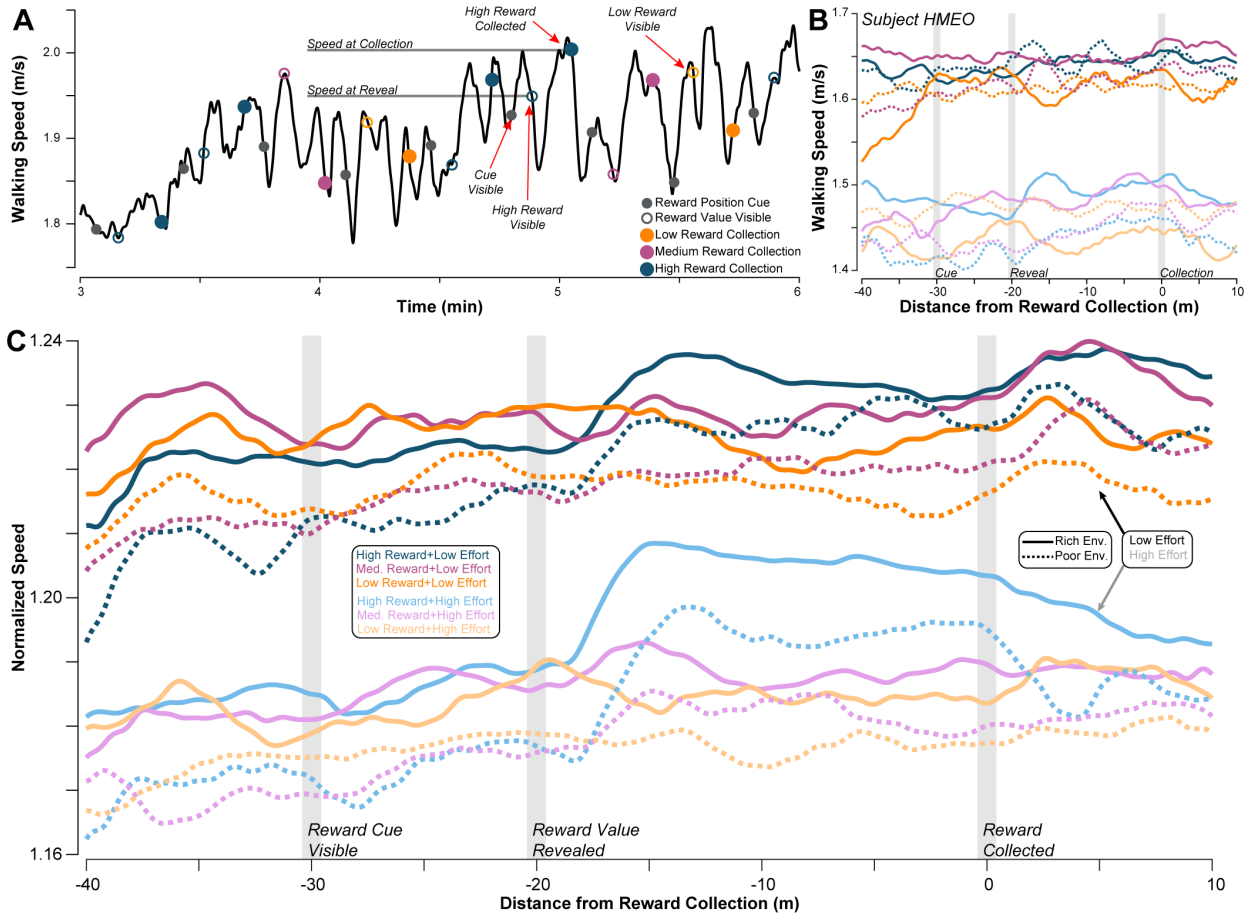

Figure 23: Example walking speed data and average behavior. *A*: an example of walking speed varying over the course of a trial. A three-minute window is shown, with specific events highlighted, including cue visibility (solid gray circle), reward value reveal (colored open circle), and reward collection (colored filled circle). *B*: an example subject, with walking speed behavior averaged across each reward instance, categorized by the reward value (high reward is blue, medium is red, low is yellow), the reward environment (rich environment is solid lines, poor is dashed), and the effort condition (low effort conditions are dark colors, high effort are light colors). For this specific subject, we see a clear separation between effort conditions and some organization between rich and poor environments. *C*: walking speed is averaged for each reward instance within a subject. We normalized each subject's walking speed and then averaged across subjects. High effort appears slower than low effort. Before the reward value is revealed, rich environments are faster than poor environments, and following reward value reveal, we see a change in speed proportional to the reward value where high is faster than medium and low.

6.4.2.1 Reward cue to reward value reveal

We begin our analysis by focusing on the period between when the cue appears and the reward value is revealed. We focus here on the effect of immediate reward, reward environment, and effort on the average speed, speed surrounding reveal, and average acceleration (Figure 24A-C, respectively). We investigated event-specific

(near-instantaneous) measures of walking speed in addition to averages over longer windows. First, we looked at the average speed between the cue appearance and reward value reveal (Figure 24A). As predicted by our model, we expected walking speed to reflect the quality of the environment and effort condition but not reward value, as it had not been revealed yet.

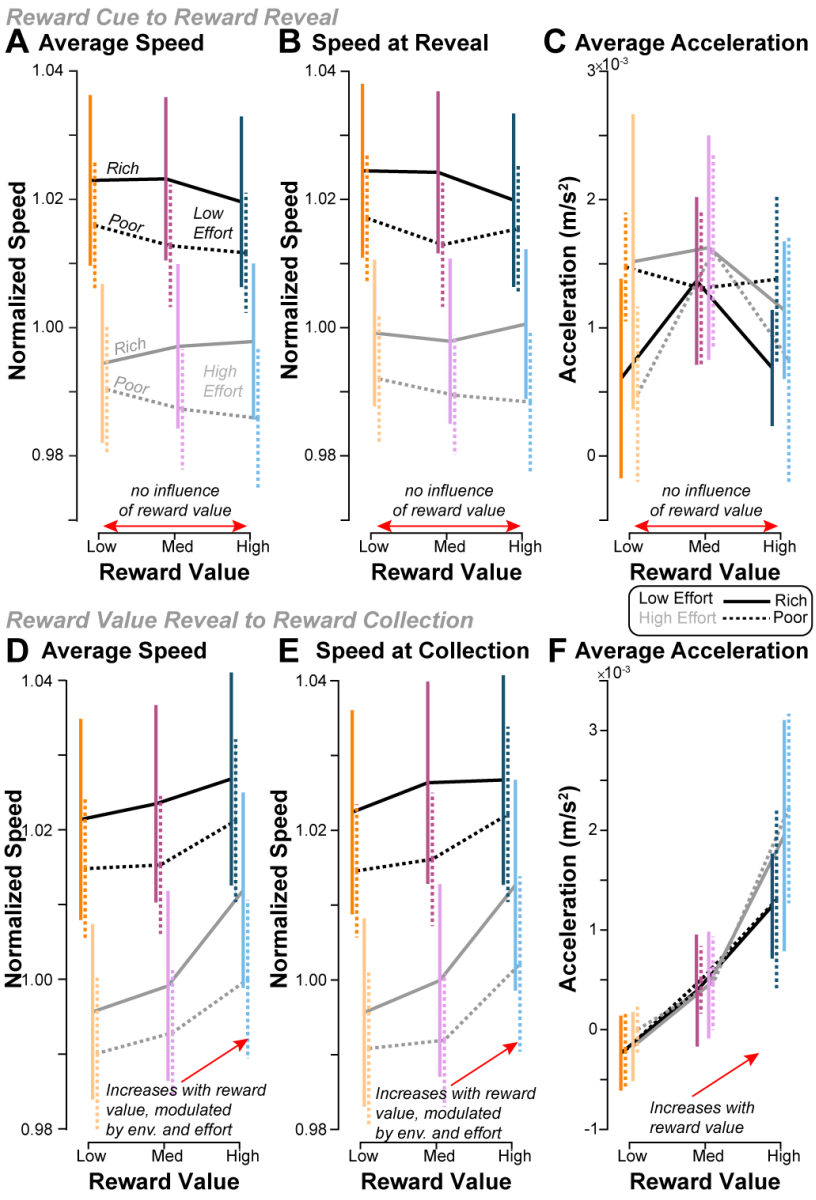

Figure 24: Average walking speed and acceleration. Averaged walking speeds and accelerations are separated by reward environment and effort condition and plotted by reward value. Speeds are normalized by the average speed for all rewarding trials for a given subject and then averaged across subjects. Data points represent means across subjects with error bars indicating standard error (mean ± s.e.). Reward values are shown in blue, red, and yellow for high, medium, and low, respectively. Rich environment is shown in solid lines, and poor is shown in dashed. Low effort is shown in dark/bold colors, and high effort is shown in light colors. *Top Row* focuses on the segment between when the reward cue becomes visible and when the reward value is revealed. *A*: Average speed between cue visible and reward value reveal shows an effect of effort condition and reward environment. *B*: the average speed between one meter before and one meter after reward reveal shows the same effect. *C*: average acceleration shows no significant trend. *Bottom Row* focuses on the segment following reward value reveal until reward capture. *D*: shows an effect of reward value on average walking speed, in addition to effort condition and rewarding environment. *E*: the average speed spanning one meter before and one meter after reward collection shows the same effect. *F*: average acceleration shows an effect of reward value but not of environment or effort condition.

Indeed, our predictions were confirmed: average speed was influenced by effort condition, where low effort was faster (Low Effort: β = 0.0336 [0.00368, 0.0636], P = 0.0278) and influenced by a history of higher average reward (i.e., in the rich reward environment, β = 0.0053 [0.00117, 0.00945], P = 0.0119), and not influenced by reward value.

Next, we looked at walking speed surrounding the moment of reveal (Figure 24B), averaging speed between 1 meter before and 1 meter after. Testing the same factors as above, we found that the same predictors as average speed over the entire window significantly predicted this speed. Low effort was faster (Low Effort: β = 0.0349 [0.00524, 0.0646], P = 0.0211), rich environment was faster (β = 0.00567 [0.00157, 0.00977], P = 0.00674). Both the average speed between cue appearance and reward reveal and the speed at reveal reflected the same variables and were in alignment with our model predictions.

Next, we looked at the average acceleration between cue and reward reveal (Figure 24C). Because the cue is associated with future reward, we thought subjects might begin accelerating following the appearance of the cue. However, acceleration was not influenced by reward value, environment, or effort. Additionally, there was an insignificant intercept term (β = -6.49×10^{-4} [-00277, 0.00147], P = 0.548), so it appeared that subjects were not significantly accelerating during this window. Separately, we ran a comparison between the speed when the cue appears and the speed when the reward value is revealed. We found that the speeds were not significantly different (P = 0.385). Acceleration was not affected by reward history or

effort, and velocity did not change between cue and reveal, suggesting that the certainty of impending reward did not change their speed. Speculatively, subjects may have learned the structure of the task such that the cue could be ignored, and they expected a reward even before the cue had appeared.

In summary, average walking speed after the reward cue appeared but before the reward value was revealed is modulated by effort and environment condition. Speeds were faster when effort was low and the environment was rich and did not change between the cue and reward reveal. As expected, neither speed nor acceleration were influenced by upcoming reward value.

6.4.2.2 Reward value reveal to reward collection

Shifting focus to the segment following reward value reveal until reward collection, we looked at the effects of immediate reward, reward environment and effort on average speed, speed surrounding collection, and average acceleration (Figure 24D, E, and F, respectively). From model predictions, we expected to see a response to the immediate reward value.

Indeed, the average speed was faster to a higher immediate reward value ($\beta =$ 0.00236 [0.00101, 0.00371], $P < 0.0001$). Average speed was also higher for a rich environment than a poor ($\beta =$ 0.00495 [0.00115, 0.00874], $P = 0.0107$) and faster in the low effort condition (Low Effort: $\beta =$ 0.0351 [0.00766, 0.0626], $P =$ 0. 0122).

These effects were also maintained at reward collection. Higher reward was collected at higher speeds ($\beta =$ 0.00236 [0.00101, 0.00371], $P <$ 0.0001). Walking speed at reward collection was also faster in the low effort condition (Low Effort: $\beta =$ 0.0332

[0.00273, 0.0616], P = 0.0323) and in a rich environment (β = 0.00477 [6.88 x 10^{-4}, 0.00885], P =0.0220).

Average acceleration between reveal and collection was not influenced by effort nor reward environment; however, it was influenced by immediate reward value (β = 2.68 x 10^{-4} [1.43 x 10^{-4}, 3.94 x 10^{-4}], P < 0. 0001). Greater immediate reward was accompanied by greater average acceleration. It appears as if the acceleration was closely tied to reward value but not influenced by either reward environment or effort. Our model predicted some effect of reward environment and a larger effect of effort condition on acceleration (Figure 22G) that we did not see in the data (Figure 24F). We will return to this point in the Discussion.

Collectively, we saw that immediate reward had an effect on walking speed after the reward value was revealed. Subjects walked faster to greater immediate reward. History of reward, modeled as the environment, also had a persistent influence on walking speed, where speeds were faster in a richer environment. We also saw that a higher baseline effort slowed walking speed. Finally, the effects of the environment and effort condition seemed to persist across all speeds, whether the immediate reward value was visible or not.

6.4.2.3 Reward collection to the next cue

Next, we investigated what happened following reward collection and prior to the appearance of the subsequent cue. Were there residual effects from the reward just collected? We focused on average walking speed and average acceleration between reward collection and the next cue (Figure 25A and B, respectively). Average

speed was significantly predicted by the reward environment (β = 0.00470 [8.51 x 10^{-4}, 0.00856], P = 0.0167) and effort condition (Low Effort: β = 0.0293 [0.00146, 0.0571], P = 0.0392), but not the preceding reward value (β = 0.00110 [-2.70 x 10^{-4}, 0.00247], P = 0.115).

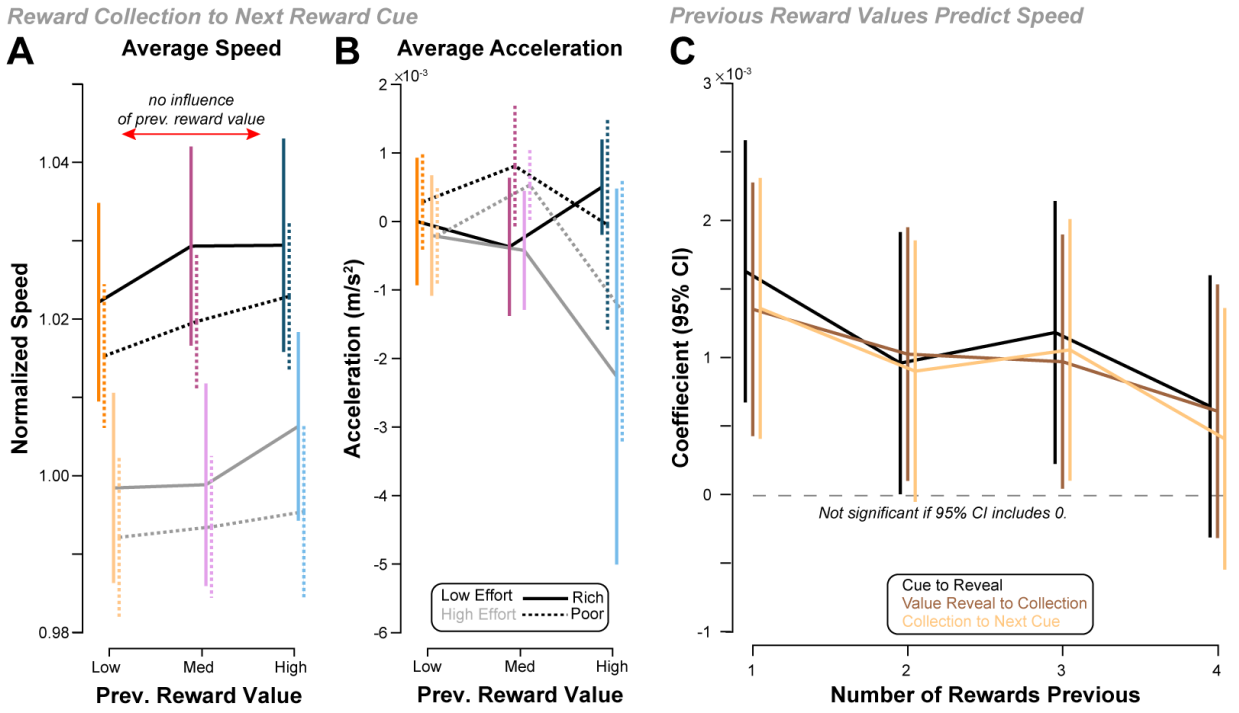

Figure 25: Response following reward collection. *A*: average speeds following collection are sensitive to the reward environment and effort condition but not the preceding reward value. *B*: average acceleration shows a significant correlation to reward value, environment, or effort condition. Data points in *A* and *B* represent means across subjects with standard error bars (mean ± s.e.). *C*: investigates an alternative formulation for a history of reward, looking at the effects of explicit values of previously received reward as a predictor for average walking speed over specific windows. Cue to reward value reveal is shown in black, reward value reveal to reward collection is shown in brown, and reward collection to the next cue is shown in tan. Across all three windows, the first and the third reward predicted speeds, and the fourth reward value did not.

There was not a significant effect of reward value on average acceleration at the significance level tested (β = -1.51 x 10^{-4} [-3.17 x 10^{-4}, 1.62 x 10^{-5}], P = 0.0768, Figure 25B). Speculatively, the lack of correlation to the previous reward could have been due to a quick period of deceleration compared to the relatively long segment we were averaging over.

Collectively, these findings suggest that there may not be a measurable influence of previously received rewards on the ensuing movement speed. However, this could be due to the presence of reward environment in our models, which accurately represented the average reward history.

6.4.3 Explicit history of reward as a predictor

In the prior analyses, we tested the effect of history of reward by using the average expected reward value for a given environment. This representation is in line with the literature on how opportunity cost influences invigoration of movement (Niv et al., 2007), but it assumes that subjects have a perfectly accurate model of the environment payout and awareness regarding which environment they are in.

Here, we explored an alternative representation of the history of reward. Instead of an average for the entire environment, we modeled history of reward as what was explicitly collected. We looked at the effect from previously collected reward values (e.g., whether walking speed is influenced differently if you receive three high rewards in a row versus a medium, then two high rewards). If so, how many previous rewards influence current walking speed (in a sense, gauging the near-term memory of history of rewards)? Over the nine-minute trial, subjects collected a median of 20 rewards per trial (mean ± s.d., 20.5 ± 2.77). For this analysis, we considered up to four prior collected rewards and their values that could influence walking speed. We looked for influences due to previous rewards on the average walking speed for a given segment (Figure 24A, D, and Figure 25A). Using a linear model, we find if any of the previous four rewards significantly predicted average walking speed (Figure

25C). Average speed between cue and reward value reveal significantly predicted the first through third prior reward, but not the fourth (First Prior Reward: β = 0.00163 [6.71 x 10^{-4}, 0.00258], P = 0.000859; Second Prior Reward: β = 9.58 x 10^{-4} [1.66 x 10^{-6}, 0.00192], P = 0.0496; Third Prior Reward: β = 0.00118 [2.23 x 10^{-4}, 0.00214], P = 0.0157). Average speed between reward value reveal and reward collection was similar: the first through third prior reward significantly predicted walking speed and the fourth did not (First Prior Reward: β = 0.00135 [4.26 x 10^{-4}, 0.00228], P = 0.00421; Second Prior Reward: β = 0.00102 [9.89 x 10^{-5}, 0.00195], P = 0.0301; Third Prior Reward: β = 9.69 x 10^{-4} [4.18 x 10^{-5}, 0.00190], P = 0.0405). The average walking speed between collection and next cue broke this pattern, with the first and third prior rewards still significant, but not the second or fourth (First Prior Reward: β = 0.00136 [4.06 x 10^{-4}, 0.00231], P = 0.00516; Second Prior Reward: β = 9.00 x 10^{-4} [-5.41 x 10^{-5}, 0.00185], P = 0.0645; Third Prior Reward: β = 0.00106 [1.00 x 10^{-4}, 0.00201], P = 0.0303). Together, this suggests that for our task, a history of reward could be a short-term moving average rather than inferring the structure and average payout of an environment.

These results suggest a viable alternative to how history of reward might influence walking speed and provides some evidence for how a model of the reward environment might develop and influence walking speed.

6.4.4 Time to reveal and collection

Traditional observations of foraging behavior focus on time spent within a rewarding patch and time spent transiting between patches. In our paradigm, with

rewards set at fixed distances, average speed is analogous to time. To be thorough, we tested if immediate reward value, reward environment, and effort impacted the time spent on a segment. Measuring the time between cue appearance and reward reveal (Figure 26A), we saw an effect of reward environment (β = -0.0216 [-0.0401, -0.00310], P = 0.0222) and effort condition (β = -0.181 [-0.315,-0.0468], P = 0.00820), wherein higher effort conditions and/or in the poor environment, subjects take more time to traverse this segment. As expected, we did not see an effect of reward value.

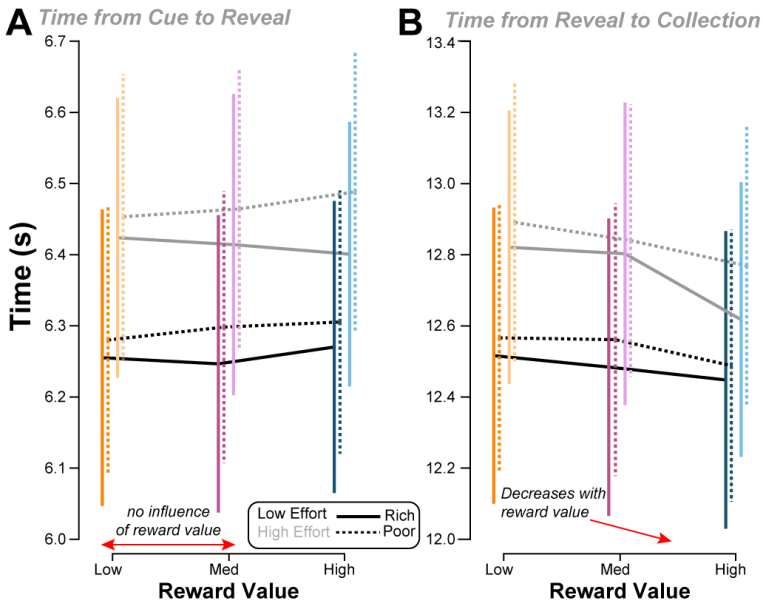

Figure 26: Alternative measures and predictors. *A*: the average time between cue appearance and reward value reveal was longer for high effort environments and longer for poor environments than rich, reflecting the same trends as average speed. *B*: Time from reveal to collection shows an effect of reward value: times were longer for low rewards than medium and high rewards. In addition, high effort was longer than low effort, and poor was longer than a rich environment. Data points represent means across subjects with error bars indicating standard error (mean ± s.e.).

Next, we looked at the segment following reward reveal until reward collection (Figure 26B). Here, we saw that higher immediate reward values were collected in less time (β = -0.0189 [-0.0305, -0.00725], P = 0.00147). Additionally, reward was collected in less time in the low effort condition (Low Effort: β = -0.326 [-0.562, -

0.0991], $P = 0.00677$), and reward in a rich environment was obtained in less time as well ($\beta =$ -0.0331 [-0.0657, -3.98 x 10^{-4}], $P = 0.0473$). These findings reinforce those regarding average speed.

6.4.5 Reward prediction error and acceleration following reward reveal

In the prior analyses of acceleration, we found that average acceleration was correlated to the immediate reward value after being revealed (Figure 24F). Here, we looked specifically at the acceleration response following reward reveal at finer increments to investigate if there were temporal or magnitude differences.

Comparing average acceleration curves, we found that acceleration increases with increasing reward immediately following reveal, then settles back to near zero after about 8 meters on average (Figure 27A). Interestingly, we observed a difference between rich and poor environments, where peak acceleration exhibited a pattern that reflects reward prediction error (RPE), calculated as the difference between reward value and the average reward value for the environment. Peak acceleration differences from Figure 27A are plotted versus RPE in Figure 27B, revealing a linear trend. Peak acceleration for a high reward in a poor environment was greater than peak acceleration for the same high reward in a rich environment. Conversely, we saw little to no acceleration toward a low reward in a poor environment; however, we saw deceleration toward a low reward in a rich environment.

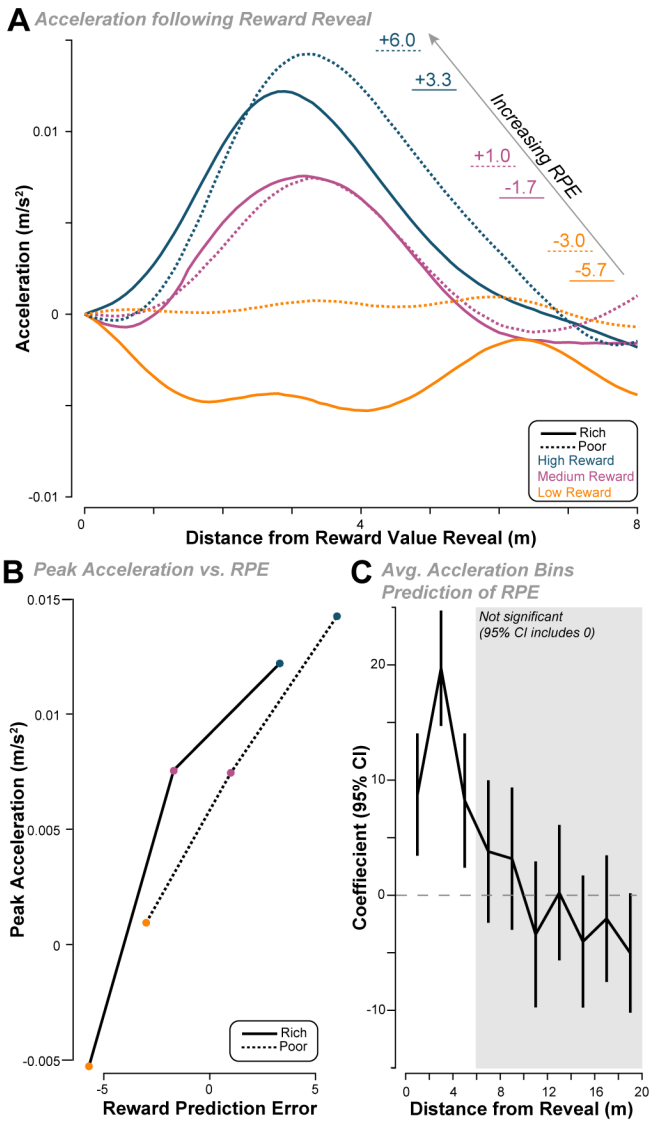

Figure 27: Acceleration following reward reveal. *A*: shows the difference in acceleration following reward value reveal. Acceleration is averaged across both effort conditions for visualization purposes. These curves show not only a trend with reward value but also a prediction error, where a response is greater for the high reward in a poor environment than a rich environment, and conversely, the response is more negative for the low reward in a rich environment than a poor environment. *B*: peak acceleration following reward reveal averaged across subjects, from *A*, is plotted versus RPE, revealing a linear trend. *C*: investigates at what distance following reward value reveal does acceleration predict RPE. Coefficients for average acceleration over two-meter windows were regressed with RPE values. Acceleration windows between 0 to 2, 2 to 4, and 4 to 6 meters significantly predicted RPE, but nothing after that window.

Statistically, we regressed RPE with the peak difference in acceleration and found they are correlated (β = 8.22 x 10^{-4} [3.00 x 10^{-4}, 0.00134], P = 0.00203), where higher RPE has a higher peak acceleration. This hints at an influence of RPE on the

difference in acceleration following reveal. However, it is difficult to disentangle an influence of reward alone from RPE because reward value is embedded in how RPE is calculated. We tested the standard model from Equation 38 and found that reward is also correlated with the difference in peak acceleration (β = 9.24 x 10^{-4} [1.29 x 10^{-4}, 0.00195], P = 0.0227). Comparing these two models, we used Akaike Information Criteria (AIC) and found a small difference (ΔAIC = 4.16) that moderately favors the model that regresses RPE. Collectively, it appears that an influence of RPE might be present in walking acceleration.

Next, we investigated what duration following reveal is acceleration correlated to RPE. We used a linear mixed model and found that accelerations between 0 to 2 meters, 2 to 4 meters, and 4 to 6 meters can predict RPE (reward prediction error) (0 to 2: β = 8.74 [3.42, 14.0], P = 0.00128; 2 to 4: β = 19.7 [14.7, 24.7], P < 0.0001; 4 to 6: β = 8.22 [2.39, 14.0], P = 0.00570). Beta coefficient confidence intervals for all acceleration intervals are plotted in Figure 27C, showing that only the coefficients before 6 meters are significant. So, it appears this response happens shortly after reveal (~3-5 seconds), then fades.

6.4.6 Comparing measures of effort expenditure

We sought to manipulate the baseline effort of our task by adding mass to subjects but left walking speed, which dictates effort expenditure as necessarily free to vary. Added mass is task-agnostic: it makes accomplishing the goal more difficult. Increasing walking speed is also more effortful but is task-relevant. As previously discussed, subjects are willing to pay the cost of extra effort to achieve the goal earlier.

Though effort cannot be directly measured, we infer effort expenditure by measuring average heart rate and by collecting Borg RPE scores for each trial. In Figure 28, we compare gross averages of walking speed, heart rate, and Borg RPE across experiment phases.

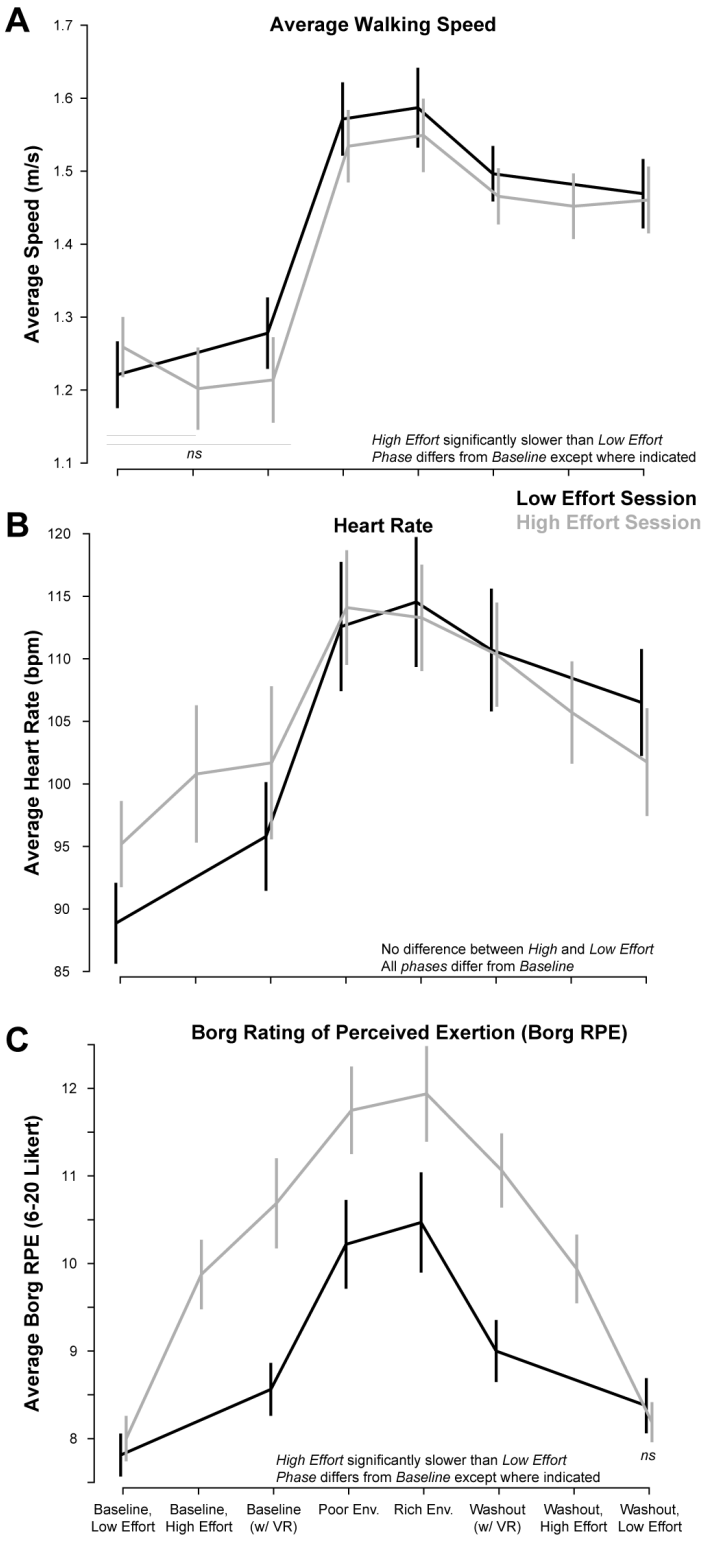

Figure 28: Comparison of effort measures by experiment phase. *A*: average walking speed (mean ± s.e.) across subjects for each phase. Data from the low effort session is shown in black, and the high effort session in gray. *B*: average heart rate data in beats per minute (bpm) is shown across experiment phases. *C*: average Borg RPE scores across subjects for each phase.

Using linear mixed effects models, we tested the effects of experiment phase and effort condition for the three measures of effort expenditure. We found a significant effect of high effort on average walking speed (High Effort: β = -0.0370 [-0.0677, -6.23x10^{-3}], P = 0.0186, Figure 28A) and a significant increase in speed from baseline for all phases except Baseline with VR (Poor environment: β = 0.339 [0.290, 0.388], P < 0.0001; Rich: β = 0.354 [0.306, 0.404], P < 0.0001; Washout with VR: β = 0.266 [0.207, 0.324], P < 0.0001; Washout without VR: β = 0.239 [0.286, 0.291], P < 0.0001). Intriguingly, average speed increased in trials with reward and stayed high in subsequent washout phases.

Interestingly, there was no effect of high effort on average heart rate (High Effort: β = 2.29 [-1.06, 5.63], P = 0.180, Figure 28B). Similar to walking speed, there was an increase across all phases compared to Baseline except Baseline w/ VR (Poor environment: β = 18.01 [12.7, 23.7], P < 0.0001; Rich environment: β = 18.6 [13.2, 23.9], P < 0.0001; Washout with VR: β = 15.2 [8.81, 21.6], P < 0.0001; Washout without VR: β = 9.70 [4.02, 15.4], P = 8.93x10^{-4}). This suggests that subjects may have expended a similar energetic output with the added mass by walking more slowly. It also suggests that large changes in walking speed, as seen in the different phases, track well with average heart rate.

Next, we investigated subjective ratings of effort. There was an effect of high effort where higher (harder) ratings were given for the high effort condition (High Effort: β = 1.73 [1.44, 2.02], P < 0.0001, Figure 28C). In addition, there was a significant difference between experiment phases for every phase except Washout

without VR (Baseline with VR: $\beta = 0.774$ [0.229, 1.32], $P = 0.00550$; Poor environment: $\beta = 2.13$ [1.68, 2.59], $P < 0.0001$; Rich environment: $\beta = 2.35$ [1.90, 2.81], $P < 0.0001$; Washout with VR: $\beta = 1.18$ [0.636, 1.72], $P < 0.0001$; Washout without VR: $\beta = 0.271$ [-0.214, 0.756], $P = 0.273$). The Borg RPE matches increases in both walking speed and heart rate well, but we saw a large divergence in scores when comparing high effort and low effort, compared to the small differences in walking speed and no difference in heart rate.

As a whole, we found objective and subjective changes in effort expenditure depending on effort condition as well as experiment phase.

6.4.7 Temporal and virtual reality effects on walking speed

In the previous analyses, in addition to the effects of reward and effort, we also saw consistent effects of time within a trial and session number. We report these findings in Table 4, below. Subjects were faster in Session 2 compared to Session 1 and faster near the end of a reward trial compared to the beginning. In this section, we better characterize these effects. Is there an order effect, and how does it manifest in walking speed? Are subjects only speeding up in rewarding trials, or is it configuration dependent? Finally, is walking speed influenced by configuration (no VR, VR without reward, VR with reward) or reward alone?

Table 4: Effects of session number and time within a trial (95% CI).

Measure and Trial Segment	Session Number	Time within a Trial
Average Speed, Reward Cue to Reward Reveal (Figure 24A)	[0.104, 0.119]	[1.67 x 10^{-4}, 2.17 x 10^{-4}]
Average Speed, Reward Reveal to Reward Collection (Figure 24D)	[0.108, 0.122]	[1.31 x 10^{-4}, 1.76 x 10^{-4}]
Average Speed, Reward Collection to Next Cue (Figure 25A)	[0.102, 0.115]	[1.03 x 10^{-4}, 1.49 x 10^{-4}]
Average Speed at Reward Reveal (Figure 24B)	[0.105, 0.119]	[1.55 x 10^{-4}, 2.04 x 10^{-4}]
Average Speed at Reward Collection (Figure 24E)	[0.108, 0.123],	[1.11 x 10^{-4}, 1.60 x 10^{-4}]
Average Acceleration, Reward Cue to Reward Reveal (Figure 24C)	[-2.00 x 10^{-4}, 0.00129], $P = 0.152$	[-6.30 x 10^{-6}, -1.33 x 10^{-6}]
Average Acceleration, Reward Reveal to Reward Collection (Figure 24F)	[-9.30 x 10^{-5}, 0.00116], $P = 0.0937$	[-4.00 x 10^{-6}, 1.62 x 10^{-6}], $P = 0.0883$
Average Acceleration, Reward Collection to Next Cue (Figure 25B)	[-0.00253, -8.77 x 10^{-4}]	[-3.89 x 10^{-6}, 2.72 x 10^{-7}], $P = 0.406$
Time from Cue to Reveal, (Figure 26A)	[-0.483, -0.417]	[-9.63 x 10^{-4}, -7.44 x 10^{-4}]
Time from Reveal to Collection, (Figure 26B)	[-0.971, -0.855]	[-0.00151, -0.00112]

First, we looked at time within a trial. We saw a consistently significant effect of time within a trial on walking speeds, suggesting that subjects were walking faster later in rewarding trials. The coefficients and significance for each model fit are summarized in detail in Table 4. We verified this finding by taking a different approach to investigating the effect of time within a trial. We compared the average speed for the first minute and the average speed for the last minute of each trial. Interestingly, we found there is a significant increase in speed between the first minute and last minute of every trial type, regardless of configuration. The magnitude of the difference was significantly influenced by the trial duration, where longer trials had larger differences in speed (β = 0.00686 [0.00101, 0.0127], P = 0.0216), and session number, where the second session, subjects sped up less (β = -0.0279 [-0.0514, -0.00439], P = 0.0202), but there was no influence due to weight

condition. To visualize this effect, we plot and test the averages of the first and last minute for trials in VR with reward (Figure 29A). Unsurprisingly, for this subset of trials, there is still a difference between the start and finish ($\beta = 0.171$ [0.142, 0.200], $P < 0.0001$), but there was no difference between rich and poor environments. Collectively, there appeared to be a subtle increase in speed agnostic to configuration even when reward was not present. Importantly, this trend is accounted for in our statistical analyses.

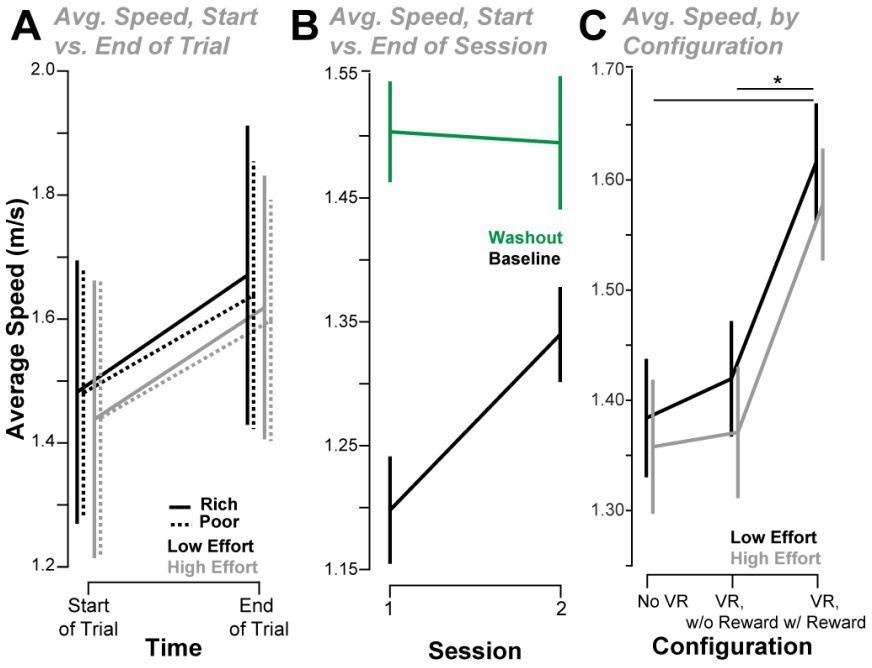

Figure 29: Effects of configuration and test session on walking speed. *A*: average walking speed in the first minute of a rewarding trial with VR compared to the last minute. Walking speeds increase in both rewarding environments and effort conditions, demonstrating a significant effect of time in a trial. *B*: average walking speeds for trials without the VR headset, separated by session. Washout trials are faster than baseline trials but do not vary across sessions. Baseline speeds for the second session were faster than baseline speeds in the first session. *C*: average walking speed was higher in the virtual environment when reward was present compared to baseline walking speed without the VR headset and higher than with the VR headset when no reward was present. Trials in the high effort condition are shown in gray, and the low effort condition in black. All plots are across subject means with bars representing standard error (mean ± s.e.).

In the previous sections investigating reward and effort influences on walking speed, we also saw a consistent effect of session number (Table 4), where the second

session walking speeds were significantly faster than the first session across all measures of walking speed. While the previous statistical tests were focused on the trials with VR and rewards, we expanded this finding, looking at the first and last trials of each session. We saw an effect of session when comparing average speeds for the baseline and washout trials completed without the weight vest and VR headset (Figure 29B). Baseline walking speeds in the second session were significantly faster than the first session ($\beta = 0.142$ [0.0818, 0.202], $P < 0.0001$). Washout trials from both sessions were faster than the baseline ($\beta = 0.456$ [0.322, 0.590], $P < 0.0001$), but there was not an effect of session. Thus, within a given session, subjects were walking faster over time, and by the second session, they began the session at a faster pace.

We compared configuration-specific categories of trials to assess if VR by itself influences walking speed. We grouped baseline and washout trials without the VR headset and grouped the baseline and washout trials with the VR headset. We also grouped all rewarding trials conducted with VR in both rich and poor environments (Figure 29C). Using a linear mixed effects model with a random effect of subject, we compared walking speeds for these gross categories and effort conditions. We did not see a difference between the No VR condition and VR with no reward condition (VR without Reward: $\beta = 0.0262$ [-0.0248, 0.0772], $P = 0.313$), suggesting that walking in VR on its own did not slow or quicken walking speeds. Additionally, we found there was no effect of effort condition at the significance level we tested (High Effort: $\beta = -0.0370$ [-0.0746, 6.83×10^{-3}], $P = 0.0543$). This may be due to the coarse grouping and

gross averages used for this analysis, whereas we saw clear effects of effort in the previously presented analyses.

Interestingly we found a strong effect of the presence of reward on walking speeds. Trials in VR with reward were significantly faster than both the VR without reward and no VR conditions (VR with Reward: $\beta = 0.228$ [0.185, 0.271], $P < 0.0001$). Compared to baseline walking conditions, rewarding trials were 14.4% faster on average. These findings suggest that while the VR system and virtual environment do not influence walking speed, the presence of reward does. Simply adding rewards to the environment led to nearly a 15% increase in average walking speed.

6.5 Discussion

How fast should one walk to a reward? If we optimized for energetic costs alone, we should only go at the speed that minimizes effort. Extensions of marginal value theory predict that our movement speeds should maximize a global utility that considers the reward, effort, and time (Yoon et al., 2018). Our findings show that these predictions hold even in the selection of preferred walking speeds, a measure thought to be primarily governed by energetic costs.

Here, walking speeds were sensitive to the immediate value of reward, as well as the history of reward. In addition, walking speeds were sensitive to baseline effort expenditure. Collectively, these findings demonstrate that locomotor decisions were dictated by optimizing a utility function that included reward and effort, discounted by time. This study is the first to our knowledge to formally show these effects in human walking.

6.5.1 Immediate reward invigorates walking speed

We see that higher reward values result in greater acceleration and faster walking speeds. This finding is in line with numerous other studies that show invigoration due to reward. Eyes saccade faster to more rewarding targets (Takikawa et al., 2002; Xu-Wilson et al., 2009) and reaches are performed more quickly to higher value targets (Summerside, Shadmehr, et al., 2018).

Interestingly, we did not see invigoration following the reward cue. We speculate this could be due to the deterministic nature of the task, where reward distances are predictable, and the presence of a reward is certain. Subjects may have learned that the cue is not important and does not change the outcome. So, while other task paradigms in animal studies show a response at the cue, where movement is invigorated at a rate equivalent to the average reward rate for the environment, our subjects disregarded this cue and only responded following reward reveal.

6.5.2 A history of high rewards invigorates walking speed

Our results show that the effect of the environment quality (rich or poor) is present across walking speeds, where subjects walked faster in a richer environment than in a poor environment. Importantly, even for the same immediate reward value, walking speed was slower following a poor history of reward compared to a rich one. Similar to Niv et al., we modeled history of reward as the average payout for an environment and represented it as an opportunity cost of time (Niv et al., 2007). This formulation seems to hold up well. However, we were also able to describe results using explicitly experienced reward collection events. In addition, there may be a

more structured learning approach where one of the two environments is detected rather than continuously integrated. Though it is clear there is an influence of history of reward from our data, changing the experimental approach so there are not two discrete environments may be able to better tease out what is being encoded as a history of reward and how it influences movement vigor. Collectively, our results build on findings that a history of reward influences movement vigor. What has been seen in saccades (Yoon et al., 2018) and arm reaches (Sukumar et al., 2024), is also present in human locomotion.

6.5.3 Increased walking effort decreases walking speed

In our experiment, subjects slowed their walking speed in the high effort condition. This effect was seen across gross averages of walking speed as well as in short segments between reward-specific events (cue, reveal, collection). This supports the theory that movement vigor considers the utility of a movement, not just energetic cost, nor is it solely based on reward value alone. Baseline effort could be viewed similarly to a history of effort, shown to affect saccades (Yoon et al., 2018) and arm reaches (Sukumar et al., 2024). In our case, the added effort cost decreased the opportunity cost of time, thus predicting slower movement speeds.

We manipulated baseline effort, increasing the energetic cost to cover the distance. Based on findings from Bastien et al., walking speed should be the same (Bastien et al., 2005) if subjects were solely optimizing for cost of transport. The lower average walking speed in combination with higher Borg RPE scores, suggests subjective effort incorporates more than energetics alone.

In our analyses of walking speed data, we tested for interaction effects with effort condition, but we did not observe any. We were curious if, for the same reward value and environment, in a high effort condition, we might see smaller changes in speed because of the decreased utility. This interaction may not exist, or we may not be able to observe them if the effects are small in comparison to the variability of walking speed. Further investigation is required to truly assess whether these interactions exist.

6.5.4 Reward prediction error accelerates walking speed

Interestingly, during a short window following reward reveal, we saw peak walking accelerations RPE. In a poor environment, a high reward is rarer and thus elicits a higher peak acceleration than a high reward in a rich environment. In the opposite scenario, we saw that subjects have a negative peak acceleration when a low reward is revealed in a rich environment. This finding was surprising given the variability in our data. Large accelerations for walking may only occur during specific gait phases, and the timing of the gait phases is uncontrolled relative to the time of reward reveal. The timescale of this effect is also noteworthy. Acceleration only reflects RPE in the first 6 meters following reveal (approximately 3-5 seconds depending on walking speed), then the effect diminishes. Average acceleration, on the other hand, is determined by reward value only. Thus, the effect of RPE is smaller or on a shorter time scale than the effects of immediate reward value and a history of reward. Though dopaminergic responses to RPE have been shown across numerous studies, the influence of RPE on vigor has only recently been shown in saccade

reaction times (Sedaghat-Nejad et al., 2019) and speed of arm reaches (Korbisch & Ahmed, 2022). In conjunction with our results, perhaps the effect of RPE is more globally present in movement invigoration.

6.5.5 Limitations

One limitation is that there is a potential for bleed-over effects from one reward segment to the next. Many of the previous findings in movement vigor study arm reaches or eye movements that start and end with zero velocity. However, in this experiment, subjects walked continuously and were not required to stop at or between rewards. Thus, the starting walking speed for a given reward segment may have differed based on the previously experienced reward. This created some difficulty in determining if the effect of the environment we saw before a reward cue and reward reveal is truly an effect of the expectation of future rewards. Alternatively, this could be more of a leaky integrator effect, where an effect of a rich environment has a higher frequency of high rewards, higher rewards are collected at higher speeds, and the time to decelerate is slow enough that walking speed does not return to baseline between reward bouts. However, in our analyses, we found that there was no effect of the previous reward on the walking speed at the next cue when environment was included as a predictor.

In addition, our task was more deterministic than other studies. Subjects were given a reward (at least one apple), with certainty every 40 meters. This repeating structure could have led to predictability and a diminished effect of the reward cue, of which, we do not see a change of speed towards. Additionally, we only have one

type of cue and the different reward environments are separated by distinct trials. It would be an interesting extension to introduce more stochastic outcomes, such as multiple cue types with variable outcomes in reward value (including no reward) and immediate effort.

One minor limitation is that we did not directly measure a subject's walking speed; we inferred it from treadmill speed. There might be some lag between true changes in walking speed and what is detected and commanded by the treadmill. Overall, this effect is negligible, as we clearly found observable differences in walking speed.

Finally, our model is necessarily simple. Though it makes many accurate predictions, it failed to make some of the predictions we observed in our data. We did not see a difference in acceleration based on effort condition following reward reveal in our data. Perhaps this is because the model treats all accelerations as equal, though accelerations from higher speeds are more energetically costly. Walking speeds in the higher effort condition are lower, thus perhaps the cost to accelerate the same amount is energetically lower than at a higher speed, as seen in the low effort condition. A more energetically accurate model of walking speed might improve predictions made by the current model.

6.5.6 Conclusions

Here, we found that reward and effort influence the speed that humans choose to walk. We developed a novel self-paced VR system to probe at the effects of immediate reward, history of reward, and baseline effort. We found that the incentive

of reward can increase walking speed. Because walking speed is such an accessible metric, our results demonstrate promise for applications in the field and clinic to better characterize neurological differences inferred through locomotion, as well as develop rehabilitation techniques to restore locomotor function.

CHAPTER VII

THESIS CONCLUSIONS

7.1 Summary of Findings

1. In my first study, I investigated if a difference in subjective effort costs relative to error can mask motor learning in an arm reaching adaptation task. I found that:
    a. Kinematic behaviors, such as reaching errors, are not uniquely specified by an amount learned.
    b. Older adults could have learned as well as younger adults when adapting their reaches in a curl field.
    c. If younger and older adults were to have learned the same amount, older adults care more about effort relative to error than younger adults.
    d. Future investigations should consider the subjective costs, or implicit strategies, when comparing movement behaviors between groups or populations.
2. In my second study, I developed a unique isometric arm pushing task to tease apart the cost of physical effort, task costs, and the time sensitivity of these costs. By measuring subject preferences, I found that:
    a. Subjects exhibit a preference for both when physical effort and task costs occur, adding to the body of evidence that costs in motor control are time sensitive.

b. On average, subjects preferred to perform high physical effort earlier rather than later but were idiosyncratic about when task costs should be performed.

c. How subjects indicated their preferences, including movement vigor and deliberation time, reflected the degree of their preference, where larger differences in utility were made more vigorously and with shorter deliberation times.

3. In the third study, I integrate a VR system with a self-paced treadmill to study the effects of reward on preferred walking speed. In this study, I found:

   a. Subjects exhibited sensitivity to reward, where higher rewards were collected at faster speeds.

   b. There may be an effect of reward prediction error, where higher rewards in a poorer environment are more invigorating.

   c. Modifications to the experimental setup were necessary to enable hands-free walking, including the incorporation of visual cues like a pathway and semi-transparent treadmill, as well as tuning the control algorithm to ensure a stable and responsive experience when walking with the VR headset.

4. In the fourth and final study, I employ our VR system and self-paced treadmill to investigate the effects of immediate reward, history of reward, and baseline effort on preferred walking speed. I found that:

a. Walking speed is influenced by both reward and effort, not just energetic cost alone, and follows many of the predictions made by optimal foraging theory.
b. Immediate reward invigorates walking speed, proportional to the value of the reward.
c. Walking speed increases when there is a history of higher reward.
d. Acceleration in response to reward reveal is proportional to the reward value minus the average reward payout for the environment, hinting that reward prediction error plays a role in walking invigoration.
e. Walking speed is slower with increasing baseline effort but does not necessarily interact with the influence of reward.
f. I propose a normative model that predicts the observations seen in the walking speed data, further justifying that movement utility is a combination of reward discounted by effort and time and includes an opportunity cost of time.

Collectively, the presented studies help further our understanding of the subjective valuation of effort and how it interplays with reward, time, and error. Importantly, these studies offer conclusive evidence of human behavior that motor control and motor learning are not simply minimizing energetic costs, but rather are complex interplays of costs and benefits, manifested into the decision of how to move.

7.2 Significance

Our brains, whether consciously or subconsciously, are constantly making decisions about how to move our bodies. The evidence presented in this thesis contributes to an overall understanding of the underlying determinants of human movement. Integral to the decision of how to move is how effort expenditure is weighed against other factors. Collectively, this work demonstrates that effort expenditure is not just about energetic cost alone—how and when that energy is expended matters.

Behaviorally, this collection of work suggests that the brain does not make movement decisions solely based on minimizing energy expenditure. Instead, other factors such as error and time can influence how much effort a person is willing to expend. How a person values effort relative to error can influence how they appear to learn a new task. People may prefer to exert more effort earlier in a task rather than later, even if total energy is the same, which revises the assumption that costs such as effort should be put off as long as possible. Additionally, walking speed, often considered to be primarily determined by energetic costs, can be influenced by factors like reward and a history of reward. This suggests that our motivation and the value we place on the world around us can override our natural tendency to conserve energy.

Methodologically, our results advance the computational techniques used for researchers. This research highlights the importance of considering subjective costs, such as effort, when studying motor learning and control. Traditional models often

assume that people learn and control their movements by minimizing error or effort independently, but here, I show that this may not always be the case. People may be willing to tolerate higher error if it means reducing effort, or expend higher effort to achieve a reward in less time. This has implications for how researchers design experiments and interpret their data. This thesis highlights the descriptive power of computational models that incorporate subjective costs. Leveraging approaches such as optimal control models, researchers can better understand the factors that influence movement decisions.

Clinically, understanding the role of subjective effort cost can lead to better diagnostic approaches and interventions for motor rehabilitation. For example, the finding that older adults may not have a learning deficit but rather prioritize minimizing effort has implications for how clinicians design rehabilitation programs for this population. Instead of focusing solely on improving performance, clinicians may need to consider strategies that reduce the perceived effort of movements to encourage participation and adherence to therapy. The discovery that walking speed can be a marker of implicit value assigned by the brain holds promise for the diagnosis and treatment of psychiatric disorders. Monitoring changes in walking speed could provide insights into a patient's motivation, reward sensitivity, and overall mental state, potentially offering a non-invasive and easily measurable tool for assessing and tracking the effectiveness of interventions.

It is my sincere hope that this work accelerates advances in understanding how people make movement decisions, that it informs researchers on how to build from

these findings and adopt these methods, and that it inspires clinicians and engineers to apply these findings and improve the lives of others.

REFERENCES

Ahmed, A. A., & Wolpert, D. M. (2009). Transfer of Dynamic Learning Across Postures. *Journal of Neurophysiology*, *102*(5), 2816–2824. https://doi.org/10.1152/jn.00532.2009

Alexander, R. McN. (1997). A minimum energy cost hypothesis for human arm trajectories. *Biological Cybernetics*, *76*(2), 97–105. https://doi.org/10.1007/s004220050324

Banton, T., Stefanucci, J., Durgin, F., Fass, A., & Proffitt, D. (2005). The Perception of Walking Speed in a Virtual Environment. *Presence*, *14*(4), 394–406. Presence. https://doi.org/10.1162/105474605774785262

Bastien, G. J., Willems, P. A., Schepens, B., & Heglund, N. C. (2005). Effect of load and speed on the energetic cost of human walking. *European Journal of Applied Physiology*, *94*(1), 76–83. https://doi.org/10.1007/s00421-004-1286-z

Bautista, L. M., Tinbergen, J., & Kacelnik, A. (2001). To walk or to fly? How birds choose among foraging modes. *Proceedings of the National Academy of Sciences*, *98*(3), 1089–1094. https://doi.org/10.1073/pnas.98.3.1089

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain Dopamine Neurons Encode a Quantitative Reward Prediction Error Signal. *Neuron*, *47*(1), 129–141. https://doi.org/10.1016/j.neuron.2005.05.020

Berniker, M., & Kording, K. (2008). Estimating the sources of motor errors for adaptation and generalization. *Nature Neuroscience*, *11*(12), 1454–1461. https://doi.org/10.1038/nn.2229

Blinch, J., & DeWinne, C. R. (2019). Pre-crastination and procrastination effects occur in a reach-to-grasp task. *Experimental Brain Research*, *237*(5), 1129–1139. https://doi.org/10.1007/s00221-019-05493-3

Bobbert, A. C. (1960). Energy expenditure in level and grade walking. *Journal of Applied Physiology*, *15*(6), 1015–1021. https://doi.org/10.1152/jappl.1960.15.6.1015

Borg, G. A. (1982). Psychophysical bases of perceived exertion. *Medicine and Science in Sports and Exercise*, *14*(5), 377–381.

Bornstein, M. H., & Bornstein, H. G. (1976). The pace of life. *Nature*, *259*(5544), 557–559. https://doi.org/10.1038/259557a0

Browning, R. C., Baker, E. A., Herron, J. A., & Kram, R. (2006). Effects of obesity and sex on the energetic cost and preferred speed of walking. *Journal of Applied Physiology*, *100*(2), 390–398. https://doi.org/10.1152/japplphysiol.00767.2005

Bruening, G. W., Courter, R. J., Sukumar, S., O'Brien, M. K., & Ahmed, A. A. (2024). Disentangling the effects of metabolic cost and accuracy on movement speed. *PLOS Computational Biology*, *20*(5), e1012169. https://doi.org/10.1371/journal.pcbi.1012169

Bruijn, S. M., Van Impe, A., Duysens, J., & Swinnen, S. P. (2012). Split-belt walking: Adaptation differences between young and older adults. *Journal of Neurophysiology*, *108*(4), 1149–1157. https://doi.org/10.1152/jn.00018.2012

Buch, E. R., Young, S., & Contreras-Vidal, J. L. (2003). Visuomotor Adaptation in Normal Aging. *Learning & Memory*, *10*(1), 55–63. https://doi.org/10.1101/lm.50303

Burdet, E., Franklin, D. W., & Milner, T. E. (2013). *Human robotics: Neuromechanics and motor control*. The MIT Press.

Caramiaux, B., Françoise, J., Liu, W., Sanchez, T., & Bevilacqua, F. (2020). Machine Learning Approaches for Motor Learning: A Short Review. *Frontiers in Computer Science*, *2*, 16. https://doi.org/10.3389/fcomp.2020.00016

Cesqui, B., Macrì, G., Dario, P., & Micera, S. (2008). Characterization of age-related modifications of upper limb motor control strategies in a new dynamic environment. *Journal of NeuroEngineering and Rehabilitation*, *5*(1), 31. https://doi.org/10.1186/1743-0003-5-31

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, *9*(2), 129–136. https://doi.org/10.1016/0040-5809(76)90040-X

Christou, E. A. (2011). Aging and Variability of Voluntary Contractions. *Exercise and Sport Sciences Reviews*, *39*(2), 77–84. https://doi.org/10.1097/JES.0b013e31820b85ab

Cohen, R. G., & Rosenbaum, D. A. (2004). Where grasps are made reveals how grasps are planned: Generation and recall of motor plans. *Experimental Brain Research*, *157*(4), 486–495. https://doi.org/10.1007/s00221-004-1862-9

Cooke, J. D., Brown, S. H., & Cunningham, D. A. (1989). Kinematics of arm movements in elderly humans. *Neurobiology of Aging*, *10*(2), 159–165. https://doi.org/10.1016/0197-4580(89)90025-0

Cos, I., Bélanger, N., & Cisek, P. (2011). The influence of predicted arm biomechanics on decision making. *Journal of Neurophysiology*, *105*(6), 3022–3033. https://doi.org/10.1152/jn.00975.2010

Courter, R. J., Alvarez, E., Enoka, R. M., & Ahmed, A. A. (2023). Metabolic costs of walking and arm reaching in persons with mild multiple sclerosis. *Journal of Neurophysiology*, *129*(4), 819–832. https://doi.org/10.1152/jn.00373.2022

Crevecoeur, F., Scott, S. H., & Cluff, T. (2019). Robust Control in Human Reaching Movements: A Model-Free Strategy to Compensate for Unpredictable Disturbances. *Journal of Neuroscience*, *39*(41), 8135–8148. https://doi.org/10.1523/JNEUROSCI.0770-19.2019

Dietz, V., Zijlstra, W., & Duysens, J. (1994). Human neuronal interlimb coordination during split-belt locomotion. *Experimental Brain Research*, *101*(3). https://doi.org/10.1007/BF00227344

Donchin, O., Francis, J. T., & Shadmehr, R. (2003). Quantifying Generalization from Trial-by-Trial Behavior of Adaptive Systems that Learn with Basis Functions: Theory and Experiments in Human Motor Control. *The Journal of Neuroscience*, *23*(27), 9032–9045. https://doi.org/10.1523/JNEUROSCI.23-27-09032.2003

Donelan, J. M., Kram, R., & Arthur D., K. (2001). Mechanical and metabolic determinants of the preferred step width in human walking. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *268*(1480), 1985–1992. https://doi.org/10.1098/rspb.2001.1761

Emken, J. L., Benitez, R., Sideris, A., Bobrow, J. E., & Reinkensmeyer, D. J. (2007). Motor Adaptation as a Greedy Optimization of Error and Effort. *Journal of Neurophysiology*, *97*(6), 3997–4006. https://doi.org/10.1152/jn.01095.2006

Etnier, J. L., & Landers, D. M. (1998). Motor Performance and Motor Learning as a Function of Age and Fitness. *Research Quarterly for Exercise and Sport*, *69*(2), 136–146. https://doi.org/10.1080/02701367.1998.10607679

Faisal, A. A., Selen, L. P. J., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience*, *9*(4), 292–303. https://doi.org/10.1038/nrn2258

Festinger, L. (1943). Studies in decision: I. Decision-time, relative frequency of judgment and subjective confidence as related to physical stimulus difference. *Journal of Experimental Psychology*, *32*(4), 291–306. https://doi.org/10.1037/h0056685

Finley, J. M., Bastian, A. J., & Gottschall, J. S. (2013). Learning to be economical: The energy cost of walking tracks motor adaptation. *The Journal of Physiology*, *591*(4), 1081–1095. https://doi.org/10.1113/jphysiol.2012.245506

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*(6), 381–391. https://doi.org/10.1037/h0055392

Flanagan, J. R., & Rao, A. K. (1995). Trajectory adaptation to a nonlinear visuomotor transformation: Evidence of motion planning in visually perceived space. *Journal of Neurophysiology*, *74*(5), 2174–2178. https://doi.org/10.1152/jn.1995.74.5.2174

Flanagan, J. R., & Wing, A. M. (1997). The Role of Internal Models in Motion Planning and Control: Evidence from Grip Force Adjustments during Movements of Hand-Held Loads. *Journal of Neuroscience*, *17*(4), 1519–1528. https://doi.org/10.1523/JNEUROSCI.17-04-01519.1997

Flash, T., & Hogan, N. (1985). The coordination of arm movements: An experimentally confirmed mathematical model. *Journal of Neuroscience*, *5*(7), 1688–1703. https://doi.org/10.1523/JNEUROSCI.05-07-01688.1985

Fournier, L. R., Coder, E., Kogan, C., Raghunath, N., Taddese, E., & Rosenbaum, D. A. (2019). Which task will we choose first? Precrastination and cognitive load in task ordering. *Attention, Perception, & Psychophysics*, *81*(2), 489–503. https://doi.org/10.3758/s13414-018-1633-5

Fournier, L. R., Stubblefield, A. M., Dyre, B. P., & Rosenbaum, D. A. (2019). Starting or finishing sooner? Sequencing preferences in object transfer tasks. *Psychological Research*, *83*(8), 1674–1684. https://doi.org/10.1007/s00426-018-1022-7

Franklin, D. W., & Wolpert, D. M. (2011). Computational Mechanisms of Sensorimotor Control. *Neuron*, *72*(3), 425–442. https://doi.org/10.1016/j.neuron.2011.10.006

Galea, J. M., Vazquez, A., Pasricha, N., Orban de Xivry, J.-J., & Celnik, P. (2011). Dissociating the Roles of the Cerebellum and Motor Cortex during Adaptive Learning: The Motor Cortex Retains What the Cerebellum Learns. *Cerebral Cortex*, *21*(8), 1761–1770. https://doi.org/10.1093/cercor/bhq246

Gidley, A. D., & Lankford, D. E. (2021). Cost of transport at preferred walking speeds are minimized while walking on moderately steep incline surfaces. *Human Movement Science*, *79*, 102849. https://doi.org/10.1016/j.humov.2021.102849

Goble, D. J., Coxon, J. P., Wenderoth, N., Van Impe, A., & Swinnen, S. P. (2009). Proprioceptive sensibility in the elderly: Degeneration, functional consequences and plastic-adaptive processes. *Neuroscience & Biobehavioral Reviews*, *33*(3), 271–278. https://doi.org/10.1016/j.neubiorev.2008.08.012

Goldman, M. D., Marrie, R. A., & Cohen, J. A. (2008). Evaluation of the six-minute walk in multiple sclerosis subjects and healthy controls. *Multiple Sclerosis Journal*, *14*(3), 383–390. https://doi.org/10.1177/1352458507082607

Gonzalez Castro, L. N., Hadjiosif, A. M., Hemphill, M. A., & Smith, M. A. (2014). Environmental Consistency Determines the Rate of Motor Adaptation. *Current Biology*, *24*(10), 1050–1061. https://doi.org/10.1016/j.cub.2014.03.049

Gordon, J., Ghilardi, M. F., & Ghez, C. (1994). Accuracy of planar reaching movements. *Experimental Brain Research*, *99*(1), 97–111. https://doi.org/10.1007/BF00241415

Green, L., Myerson, J., Holt, D. D., Slevin, J. R., & Estle, S. J. (2004). Discounting of Delayed Food Rewards in Pigeons and Rats: Is There a Magnitude Effect? *Journal of the Experimental Analysis of Behavior*, *81*(1), 39–50. https://doi.org/10.1901/jeab.2004.81-39

Guigon, E. (2021). A computational theory for the production of limb movements. *Psychological Review*, No Pagination Specified-No Pagination Specified. https://doi.org/10.1037/rev0000323

Haith, A. M., & Krakauer, J. W. (2013). Model-Based and Model-Free Mechanisms of Human Motor Learning. In M. J. Richardson, M. A. Riley, & K. Shockley (Eds.), *Progress in Motor Control* (pp. 1–21). Springer. https://doi.org/10.1007/978-1-4614-5465-6_1

Haith, A. M., Reppert, T. R., & Shadmehr, R. (2012). Evidence for Hyperbolic Temporal Discounting of Reward in Control of Movements. *Journal of Neuroscience*, *32*(34), 11727–11736. https://doi.org/10.1523/JNEUROSCI.0424-12.2012

Haith, A., Miall, C., Jackson, C., & Vijayakumar, S. (n.d.). *Unifying the Sensory and Motor Components of Sensorimotor Adaptation*. 8.

Harris, C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature*, *394*(6695), Article 6695. https://doi.org/10.1038/29528

Hartmann, M. N., Hager, O. M., Tobler, P. N., & Kaiser, S. (2013). Parabolic discounting of monetary rewards by physical effort. *Behavioural Processes*, *100*, 192–196. https://doi.org/10.1016/j.beproc.2013.09.014

Healy, C. M., & Ahmed, A. A. (2024). Physical effort precrastination determines preference in an isometric task. *Journal of Neurophysiology*, *132*(5), 1395–1411. https://doi.org/10.1152/jn.00040.2024

Healy, C. M., Berniker, M., & Ahmed, A. A. (2023). Learning vs. minding: How subjective costs can mask motor learning. *PLOS ONE*, *18*(3), e0282693. https://doi.org/10.1371/journal.pone.0282693

Helmholtz, H. von. (1962). *Treatise on Physiological Optics*. Courier Corporation.

Hill, A. V. (1953). The mechanics of active muscle. *Proceedings of the Royal Society of London. Series B - Biological Sciences*, *141*(902), 104–117. https://doi.org/10.1098/rspb.1953.0027

Hogan, P. S., Chen, S. X., Teh, W. W., & Chib, V. S. (2020). Neural mechanisms underlying the effects of physical fatigue on effort-based choice. *Nature Communications*, *11*(1), 4026. https://doi.org/10.1038/s41467-020-17855-5

Hollman, J. H., Brey, R. H., Bang, T. J., & Kaufman, K. R. (2007). Does walking in a virtual environment induce unstable gait?: An examination of vertical ground reaction forces. *Gait & Posture*, *26*(2), 289–294. https://doi.org/10.1016/j.gaitpost.2006.09.075

Hoyt, D. F., & Taylor, C. R. (1981). Gait and the energetics of locomotion in horses. *Nature*, *292*(5820), 239–240. https://doi.org/10.1038/292239a0

Huang, H. J., & Ahmed, A. A. (2014). Older adults learn less, but still reduce metabolic cost, during motor adaptation. *Journal of Neurophysiology*, *111*(1), 135–144. https://doi.org/10.1152/jn.00401.2013

Huang, H. J., Kram, R., & Ahmed, A. A. (2012). Reduction of Metabolic Cost during Motor Learning of Arm Reaching Dynamics. *Journal of Neuroscience*, *32*(6), 2182–2190. https://doi.org/10.1523/JNEUROSCI.4003-11.2012

Hughes, A. L., & Goldman, R. F. (1970). Energy cost of "hard work." *Journal of Applied Physiology*, *29*(5), 570–572. https://doi.org/10.1152/jappl.1970.29.5.570

Hull, C. L. (1943). *Principles of behavior: An introduction to behavior theory* (pp. x, 422). Appleton-Century.

Iodice, P., Calluso, C., Barca, L., Bertollo, M., Ripari, P., & Pezzulo, G. (2017). Fatigue increases the perception of future effort during decision making. *Psychology of Sport and Exercise*, *33*, 150–160. https://doi.org/10.1016/j.psychsport.2017.08.013

Izawa, J., Rane, T., Donchin, O., & Shadmehr, R. (2008). Motor Adaptation as a Process of Reoptimization. *Journal of Neuroscience*, *28*(11), 2883–2891. https://doi.org/10.1523/JNEUROSCI.5359-07.2008

Jimura, K., Myerson, J., Hilgard, J., Braver, T. S., & Green, L. (2009). Are people really more patient than other animals? Evidence from human discounting of real liquid rewards. *Psychonomic Bulletin & Review*, *16*(6), 1071–1075. https://doi.org/10.3758/PBR.16.6.1071

Johnson, S. T., & Most, S. B. (2023). Taking the path of least resistance now, but not later: Pushing cognitive effort into the future reduces effort discounting. *Psychonomic Bulletin & Review*, *30*(3), 1115–1124. https://doi.org/10.3758/s13423-022-02198-7

Jones, K. E., Hamilton, A. F. de C., & Wolpert, D. M. (2002). Sources of Signal-Dependent Noise During Isometric Force Production. *Journal of Neurophysiology*, *88*(3), 1533–1544. https://doi.org/10.1152/jn.2002.88.3.1533

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When More Pain Is Preferred to Less: Adding a Better End. *Psychological Science*, *4*(6), 401–405. https://doi.org/10.1111/j.1467-9280.1993.tb00589.x

Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, *9*(6), 718–727. https://doi.org/10.1016/S0959-4388(99)00028-8

Kitchen, N. M., & Miall, R. C. (2021). Adaptation of reach action to a novel force-field is not predicted by acuity of dynamic proprioception in either older or younger adults. *Experimental Brain Research*, *239*(2), 557–574. https://doi.org/10.1007/s00221-020-05997-3

Klein-Flügge, M. C., Kennerley, S. W., Saraiva, A. C., Penny, W. D., & Bestmann, S. (2015). Behavioral Modeling of Human Choices Reveals Dissociable Effects of Physical Effort and Temporal Delay on Reward Devaluation. *PLOS Computational Biology*, *11*(3), e1004116. https://doi.org/10.1371/journal.pcbi.1004116

Kobayashi, S., & Schultz, W. (2008). Influence of Reward Delays on Responses of Dopamine Neurons. *Journal of Neuroscience*, *28*(31), 7837–7846. https://doi.org/10.1523/JNEUROSCI.1600-08.2008

Kojima, Y., Iwamoto, Y., & Yoshida, K. (2004). Memory of Learning Facilitates Saccadic Adaptation in the Monkey. *Journal of Neuroscience*, *24*(34), 7531–7539. https://doi.org/10.1523/JNEUROSCI.1741-04.2004

Korbisch, C. C., & Ahmed, A. A. (2022, November 11). *Vigor of movement to probabilistic reward tracks reward prediction error*. Advances in Motor Learning and Motor Control, San Diego, CA.

Korbisch, C. C., Apuan, D. R., Shadmehr, R., & Ahmed, A. A. (2022). Saccade vigor reflects the rise of decision variables during deliberation. *Current Biology*, *32*(24), 5374-5381.e4. https://doi.org/10.1016/j.cub.2022.10.053

Körding, K. P., Fukunaga, I., Howard, I. S., Ingram, J. N., & Wolpert, D. M. (2004). A Neuroeconomics Approach to Inferring Utility Functions in Sensorimotor Control. *PLOS Biology*, *2*(10), e330. https://doi.org/10.1371/journal.pbio.0020330

Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, *427*(6971), 244–247. https://doi.org/10.1038/nature02169

Korenberg, A. T., & Ghahramani, Z. (n.d.). *A Bayesian view of motor adaptation*. 28.

Krakauer, J. W., Ghilardi, M.-F., & Ghez, C. (1999). Independent learning of internal models for kinematic and dynamic control of reaching. *Nature Neuroscience*, *2*(11), 1026–1031. https://doi.org/10.1038/14826

Krakauer, J. W., Pine, Z. M., Ghilardi, M.-F., & Ghez, C. (2000). Learning of Visuomotor Transformations for Vectorial Planning of Reaching Trajectories. *The Journal of Neuroscience*, *20*(23), 8916–8924. https://doi.org/10.1523/JNEUROSCI.20-23-08916.2000

Kushmerick, M. J., & Paul, R. J. (1976). Aerobic recovery metabolism following a single isometric tetanus in frog sartorius muscle at 0 degrees C. *The Journal of Physiology*, *254*(3), 693–709. https://doi.org/10.1113/jphysiol.1976.sp011253

Lackner, J. R., & Dizio, P. (1994). Rapid adaptation to Coriolis force perturbations of arm trajectory. *Journal of Neurophysiology*, *72*(1), 299–313. https://doi.org/10.1152/jn.1994.72.1.299

Lamb, D. G., Correa, L. N., Seider, T. R., Mosquera, D. M., Rodriguez, J. A., Salazar, L., Schwartz, Z. J., Cohen, R. A., Falchook, A. D., & Heilman, K. M. (2016). The aging brain: Movement speed and spatial control. *Brain and Cognition*, *109*, 105–111. https://doi.org/10.1016/j.bandc.2016.07.009

Larish, D. D., Martin, P. E., & Mungiole, M. (1988). Characteristic patterns of gait in the healthy old. *Annals of the New York Academy of Sciences*, *515*, 18–32. https://doi.org/10.1111/j.1749-6632.1988.tb32960.x

Levine, R. V., & Norenzayan, A. (1999). The Pace of Life in 31 Countries. *Journal of Cross-Cultural Psychology*, *30*(2), 178–205. https://doi.org/10.1177/0022022199030002003

Levy, D. J., & Glimcher, P. W. (2011). Comparing Apples and Oranges: Using Reward-Specific and Reward-General Subjective Value Representation in the Brain. *Journal of Neuroscience*, *31*(41), 14693–14707. https://doi.org/10.1523/JNEUROSCI.2218-11.2011

Lord, S. R. (2006). Visual risk factors for falls in older people. *Age and Ageing*, *35*(suppl_2), ii42–ii45. https://doi.org/10.1093/ageing/afl085

Manohar, S. G., Chong, T. T.-J., Apps, M. A. J., Batla, A., Stamelou, M., Jarman, P. R., Bhatia, K. P., & Husain, M. (2015). Reward Pays the Cost of Noise Reduction in Motor and Cognitive Control. *Current Biology*, *25*(13), 1707–1716. https://doi.org/10.1016/j.cub.2015.05.038

Martin, P. E., Rothstein, D. E., & Larish, D. D. (1992). Effects of age and physical activity status on the speed-aerobic demand relationship of walking. *Journal of Applied Physiology*, *73*(1), 200–206. https://doi.org/10.1152/jappl.1992.73.1.200

Mazur, J. E., Commons, M. L., Nevin, J. A., & Rachlin, H. (1987). *The Effect of Delay and of Intervening Events on Reinforcement Value: Quantitative Analyses of Behavior, Volume V*. Psychology Press.

Mazzoni, P., Hristova, A., & Krakauer, J. W. (2007). Why Don't We Move Faster? Parkinson's Disease, Movement Vigor, and Implicit Motivation. *Journal of Neuroscience*, *27*(27), 7105–7116. https://doi.org/10.1523/JNEUROSCI.0264-07.2007

McBride, D. M., Villarreal, S. R., & Salrin, R. L. (2023). Precrastination in cognitive tasks. *Current Psychology*, *42*(17), 14984–15002. https://doi.org/10.1007/s12144-022-02750-7

McNay, E. C., & Willingham, D. B. (1998). Deficit in learning of a motor skill requiring strategy, but not of perceptuomotor recalibration, with aging. *Learning & Memory*, *4*(5), 411–420. https://doi.org/10.1101/lm.4.5.411

Milner, T. E. (1992). A model for the generation of movements requiring endpoint precision. *Neuroscience*, *49*(2), 487–496. https://doi.org/10.1016/0306-4522(92)90113-G

Minetti, A. E., Moia, C., Roi, G. S., Susta, D., & Ferretti, G. (2002). Energy cost of walking and running at extreme uphill and downhill slopes. *Journal of*

Applied Physiology, *93*(3), 1039–1046. https://doi.org/10.1152/japplphysiol.01177.2001

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

Morasso, P. (1981). Spatial control of arm movements. *Experimental Brain Research*, *42*(2). https://doi.org/10.1007/BF00236911

Moravec, H. P. (1988). *Mind children: The future of robot and human intelligence*. Harvard University Press.

Morel, P., Ulbrich, P., & Gail, A. (2017). What makes a reach movement effortful? Physical effort discounting supports common minimization principles in decision making and motor control. *PLOS Biology*, *15*(6), e2001323. https://doi.org/10.1371/journal.pbio.2001323

Morgan, M., Phillips, J. G., Bradshaw, J. L., Mattingley, J. B., Iansek, R., & Bradshaw, J. A. (1994). Age-Related Motor Slowness: Simply Strategic? *Journal of Gerontology*, *49*(3), M133–M139. https://doi.org/10.1093/geronj/49.3.M133

Morton, S. M., & Bastian, A. J. (2006). Cerebellar Contributions to Locomotor Adaptations during Splitbelt Treadmill Walking. *Journal of Neuroscience*, *26*(36), 9107–9116. https://doi.org/10.1523/JNEUROSCI.2622-06.2006

Myerson, J., & Green, L. (1995). Discounting of delayed rewards: Models of individual choice. *Journal of the Experimental Analysis of Behavior*, *64*(3), 263–276. https://doi.org/10.1901/jeab.1995.64-263

Nagengast, A. J., Braun, D. A., & Wolpert, D. M. (2009). Optimal Control Predicts Human Performance on Objects with Internal Degrees of Freedom. *PLoS Computational Biology*, *5*(6), e1000419. https://doi.org/10.1371/journal.pcbi.1000419

Nikooyan, A. A., & Ahmed, A. A. (2015). Reward feedback accelerates motor learning. *Journal of Neurophysiology*, *113*(2), 633–646. https://doi.org/10.1152/jn.00032.2014

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology*, *191*(3), 507–520. https://doi.org/10.1007/s00213-006-0502-4

O'Brien, M. K., & Ahmed, A. A. (2014). Take a stand on your decisions, or take a sit: Posture does not affect risk preferences in an economic task. *PeerJ*, *2*, e475. https://doi.org/10.7717/peerj.475

O'Brien, M. K., & Ahmed, A. A. (2019). Asymmetric valuation of gains and losses in effort-based decision making. *PLOS ONE*, *14*(10), e0223268. https://doi.org/10.1371/journal.pone.0223268

Patterson, E. E., & Kahan, T. A. (2020). Precrastination and the cognitive-load-reduction (CLEAR) hypothesis. *Memory*, *28*(1), 107–111. https://doi.org/10.1080/09658211.2019.1690001

Pellizzer, G., & Georgopoulos, ApostolosP. (1993). Common processing constraints for visuomotor and visual mental rotations. *Experimental Brain Research*, *93*(1). https://doi.org/10.1007/BF00227791

Pienciak-Siewert, A., Horan, D. P., & Ahmed, A. A. (2016). Trial-to-trial adaptation in control of arm reaching and standing posture. *Journal of Neurophysiology*, *116*(6), 2936–2949. https://doi.org/10.1152/jn.00537.2016

Pienciak-Siewert, A., Horan, D. P., & Ahmed, A. A. (2020). Role of muscle coactivation in adaptation of standing posture during arm reaching. *Journal of Neurophysiology*, *123*(2), 529–547. https://doi.org/10.1152/jn.00939.2017

Pimentel, R. E., Feldman, J. N., Lewek, M. D., & Franz, J. R. (2022). Quantifying mechanical and metabolic interdependence between speed and propulsive force during walking. *Frontiers in Sports and Active Living*, *4*. https://doi.org/10.3389/fspor.2022.942498

Prévost, C., Pessiglione, M., Météreau, E., Cléry-Melin, M.-L., & Dreher, J.-C. (2010). Separate Valuation Subsystems for Delay and Effort Decision Costs. *Journal of Neuroscience*, *30*(42), 14080–14090. https://doi.org/10.1523/JNEUROSCI.2752-10.2010

Raghunath, N., Fournier, L. R., & Kogan, C. (2021). Precrastination and individual differences in working memory capacity. *Psychological Research*, *85*(5), 1970–1985. https://doi.org/10.1007/s00426-020-01373-6

Ralston, H. J. (1958). Energy-speed relation and optimal speed during level walking. *Internationale Zeitschrift Für Angewandte Physiologie Einschließlich Arbeitsphysiologie*, *17*(4), 277–283. https://doi.org/10.1007/BF00698754

Redelmeier, D. A., & Kahneman, D. (1996). Patients' memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures. *Pain*, *66*(1), 3–8. https://doi.org/10.1016/0304-3959(96)02994-6

Reisman, D. S., Block, H. J., & Bastian, A. J. (2005). Interlimb Coordination During Locomotion: What Can be Adapted and Stored? *Journal of Neurophysiology*, *94*(4), 2403–2415. https://doi.org/10.1152/jn.00089.2005

Reuter, E.-M., Pearcey, G. E. P., & Carroll, T. J. (2018). Greater neural responses to trajectory errors are associated with superior force field adaptation in older adults. *Experimental Gerontology*, *110*, 105–117. https://doi.org/10.1016/j.exger.2018.05.020

Rigoux, L., & Guigon, E. (2012). A Model of Reward- and Effort-Based Optimal Decision Making and Motor Control. *PLOS Computational Biology*, *8*(10), e1002716. https://doi.org/10.1371/journal.pcbi.1002716

Rigoux, L., Stephan, K. E., Friston, K. J., & Daunizeau, J. (2014). Bayesian model selection for group studies—Revisited. *NeuroImage*, *84*, 971–985. https://doi.org/10.1016/j.neuroimage.2013.08.065

Rosenbaum, D. A. (1980). Human movement initiation: Specification of arm, direction, and extent. *Journal of Experimental Psychology: General*, *109*(4), 444–474. https://doi.org/10.1037/0096-3445.109.4.444

Rosenbaum, D. A., & Dettling, J. (2023). Carrying groceries: More items in early trips than in later trips or the reverse? Implications for pre-crastination. *Psychological Research*, *87*(2), 474–483. https://doi.org/10.1007/s00426-022-01681-z

Rosenbaum, D. A., Gong, L., & Potts, C. A. (2014). Pre-Crastination: Hastening Subgoal Completion at the Expense of Extra Physical Effort. *Psychological Science*, *25*(7), 1487–1496. https://doi.org/10.1177/0956797614532657

Rosenbaum, D. A., & Sauerberger, K. S. (2022). Deciding what to do: Observations from a psycho-motor laboratory, including the discovery of pre-crastination. *Behavioural Processes*, *199*, 104658. https://doi.org/10.1016/j.beproc.2022.104658

Rubaker, J. S. (2023). *A Virtual Environment to Probe the Effect of Reward on Preferred Walking Speed* [M.S., University of Colorado at Boulder]. https://www.proquest.com/docview/2868060074/abstract/6D0D7A7D4A554842PQ/1

Russ, D. W., Elliott, M. A., Vandenborne, K., Walter, G. A., & Binder-Macleod, S. A. (2002). Metabolic costs of isometric force generation and maintenance of human skeletal muscle. *American Journal of Physiology-Endocrinology and Metabolism*, *282*(2), E448–E457. https://doi.org/10.1152/ajpendo.00285.2001

Sackaloo, K., Strouse, E., & Rice, M. S. (2015). Degree of Preference and Its Influence on Motor Control When Reaching for Most Preferred, Neutrally Preferred, and Least Preferred Candy. *OTJR: Occupational Therapy Journal of Research*, *35*(2), 81–88. https://doi.org/10.1177/1539449214561763

Sauer, Y., Sipatchin, A., Wahl, S., & García García, M. (2022). Assessment of consumer VR-headsets' objective and subjective field of view (FoV) and its feasibility for visual field testing. *Virtual Reality*, *26*(3), 1089–1101. https://doi.org/10.1007/s10055-021-00619-x

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Schwob, N., Epping, A., Taglialatela, J., & Weiss, D. (2022). Animal Behavior and Cognition The Early Bonobo Gets the Juice? The Evolutionary Roots of Pre-crastination in Bonobos (Pan paniscus). *Animal Behavior and Cognition*, *9*, 3–13. https://doi.org/10.26451/abc.09.01.02.2022

Sedaghat-Nejad, E., Herzfeld, D. J., & Shadmehr, R. (2019). Reward Prediction Error Modulates Saccade Vigor. *Journal of Neuroscience*, *39*(25), 5010–5017. https://doi.org/10.1523/JNEUROSCI.0432-19.2019

Seidler-Dobrin, R. D., & Stelmach, G. E. (1998). Persistence in visual feedback control by the elderly. *Experimental Brain Research*, *119*(4), 467–474. https://doi.org/10.1007/s002210050362

Selinger, J. C., O'Connor, S. M., Wong, J. D., & Donelan, J. M. (2015). Humans Can Continuously Optimize Energetic Cost during Walking. *Current Biology*, *25*(18), 2452–2456. https://doi.org/10.1016/j.cub.2015.08.016

Shadmehr, R., & Ahmed, A. A. (2020). *Vigor: Neuroeconomics of Movement Control*. MIT Press.

Shadmehr, R., Huang, H. J., & Ahmed, A. A. (2016). A Representation of Effort in Decision-Making and Motor Control. *Current Biology*, *26*(14), 1929–1934. https://doi.org/10.1016/j.cub.2016.05.065

Shadmehr, R., & Mussa-Ivaldi, F. (1994). Adaptive representation of dynamics during learning of a motor task. *The Journal of Neuroscience*, *14*(5), 3208–3224. https://doi.org/10.1523/JNEUROSCI.14-05-03208.1994

Shadmehr, R., Orban de Xivry, J. J., Xu-Wilson, M., & Shih, T.-Y. (2010). Temporal Discounting of Reward and the Cost of Time in Motor Control. *The Journal of Neuroscience*, *30*(31), 10507–10516. https://doi.org/10.1523/JNEUROSCI.1343-10.2010

Shadmehr, R., Reppert, T. R., Summerside, E. M., Yoon, T., & Ahmed, A. A. (2019). Movement Vigor as a Reflection of Subjective Economic Utility. *Trends in Neurosciences*, *42*(5), 323–336. https://doi.org/10.1016/j.tins.2019.02.003

Shadmehr, R., Smith, M. A., & Krakauer, J. W. (2010). Error Correction, Sensory Prediction, and Adaptation in Motor Control. *Annual Review of Neuroscience*, *33*(1), 89–108. https://doi.org/10.1146/annurev-neuro-060909-153135

Sherback, M., Valero-Cuevas, F. J., & D'Andrea, R. (2010). Slower Visuomotor Corrections with Unchanged Latency are Consistent with Optimal Adaptation to Increased Endogenous Noise in the Elderly. *PLOS Computational Biology*, *6*(3), e1000708. https://doi.org/10.1371/journal.pcbi.1000708

Sidney, K. H., & Shephard, ROY. J. (1977). Perception of Exertion in the Elderly, Effects of Aging, Mode of Exercise and Physical Training. *Perceptual and Motor Skills*, *44*(3), 999–1010. https://doi.org/10.2466/pms.1977.44.3.999

Sloot, L. H., van der Krogt, M. M., & Harlaar, J. (2014). Self-paced versus fixed speed treadmill walking. *Gait & Posture*, *39*(1), 478–484. https://doi.org/10.1016/j.gaitpost.2013.08.022

Smith, E., & Peters, J. (2022). Motor response vigour and visual fixation patterns reflect subjective valuation during intertemporal choice. *PLOS Computational Biology*, *18*(6), e1010096. https://doi.org/10.1371/journal.pcbi.1010096

Smith, M. A., Ghazizadeh, A., & Shadmehr, R. (2006). Interacting Adaptive Processes with Different Timescales Underlie Short-Term Motor Learning. *PLOS Biology*, *4*(6), e179. https://doi.org/10.1371/journal.pbio.0040179

Song, S., Choi, H., & Collins, S. H. (2020). Using force data to self-pace an instrumented treadmill and measure self-selected walking speed. *Journal of NeuroEngineering and Rehabilitation*, *17*(1), 68. https://doi.org/10.1186/s12984-020-00683-5

Speakman, J. R. (1998). The history and theory of the doubly labeled water technique. *The American Journal of Clinical Nutrition*, *68*(4), 932S-938S. https://doi.org/10.1093/ajcn/68.4.932S

Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J., & Friston, K. J. (2009). Bayesian model selection for group studies. *NeuroImage*, *46*(4), 1004–1017. https://doi.org/10.1016/j.neuroimage.2009.03.025

Sugiwaka, H., & Okouchi, H. (2004). Reformative self-control and discounting of reward value by delay or effort1. *Japanese Psychological Research*, *46*(1), 1–9. https://doi.org/10.1111/j.1468-5884.2004.00231.x

Sukumar, S., Shadmehr, R., & Ahmed, A. A. (2024). Effects of reward and effort history on decision making and movement vigor during foraging. *Journal of Neurophysiology*, *131*(4), 638–651. https://doi.org/10.1152/jn.00092.2023

Summerside, E. M., & Ahmed, A. A. (2021). Using metabolic energy to quantify the subjective value of physical effort. *Journal of The Royal Society Interface*, *18*(180), 20210387. https://doi.org/10.1098/rsif.2021.0387

Summerside, E. M., Courter, R. J., Shadmehr, R., & Ahmed, A. A. (2024). Slowing of Movements in Healthy Aging as a Rational Economic Response to an Elevated Effort Landscape. *Journal of Neuroscience*, *44*(15). https://doi.org/10.1523/JNEUROSCI.1596-23.2024

Summerside, E. M., Kram, R., & Ahmed, A. A. (2018). Contributions of metabolic and temporal costs to human gait selection. *Journal of The Royal Society Interface*, *15*(143), 20180197. https://doi.org/10.1098/rsif.2018.0197

Summerside, E. M., Shadmehr, R., & Ahmed, A. A. (2018). Vigor of reaching movements: Reward discounts the cost of effort. *Journal of Neurophysiology*, *119*(6), 2347–2357. https://doi.org/10.1152/jn.00872.2017

Takikawa, Y., Kawagoe, R., Itoh, H., Nakahara, H., & Hikosaka, O. (2002). Modulation of saccadic eye movements by predicted reward outcome. *Experimental Brain Research*, *142*(2), 284–291. https://doi.org/10.1007/s00221-001-0928-1

Thoroughman, K. A., & Shadmehr, R. (2000). Learning of action through adaptive combination of motor primitives. *Nature*, *407*(6805), 742–747. https://doi.org/10.1038/35037588

Todorov, E. (2005). Stochastic Optimal Control and Estimation Methods Adapted to the Noise Characteristics of the Sensorimotor System. *Neural Computation*, *17*(5), 1084–1108. https://doi.org/10.1162/0899766053491887

Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, *5*(11), 1226–1235. https://doi.org/10.1038/nn963

Trent, M., & Ahmed, A. (2013). Learning from the value of your mistakes: Evidence for a risk-sensitive process in movement adaptation. *Frontiers in Computational Neuroscience*, *7*, 118. https://doi.org/10.3389/fncom.2013.00118

Trewartha, K. M., Garcia, A., Wolpert, D. M., & Flanagan, J. R. (2014). Fast But Fleeting: Adaptive Motor Learning Processes Associated with Aging and Cognitive Decline. *Journal of Neuroscience*, *34*(40), 13411–13421. https://doi.org/10.1523/JNEUROSCI.1489-14.2014

Tseng, Y., Diedrichsen, J., Krakauer, J. W., Shadmehr, R., & Bastian, A. J. (2007). Sensory Prediction Errors Drive Cerebellum-Dependent Adaptation of Reaching. *Journal of Neurophysiology*, *98*(1), 54–62. https://doi.org/10.1152/jn.00266.2007

Uno, Y., Kawato, M., & Suzuki, R. (1989). Formation and control of optimal trajectory in human multijoint arm movement. *Biological Cybernetics*, *61*(2), 89–101. https://doi.org/10.1007/BF00204593

Vandevoorde, K., & Orban de Xivry, J.-J. (2019). Internal model recalibration does not deteriorate with age while motor adaptation does. *Neurobiology of Aging*, *80*, 138–153. https://doi.org/10.1016/j.neurobiolaging.2019.03.020

Voelcker-Rehage, C. (2008). Motor-skill learning in older adults—A review of studies on age-related differences. *European Review of Aging and Physical Activity*, *5*(1), 5–16. https://doi.org/10.1007/s11556-008-0030-9

von Hofsten, C. (1991). Structuring of Early Reaching Movements: A Longitudinal Study. *Journal of Motor Behavior*, *23*(4), 280–292. https://doi.org/10.1080/00222895.1991.9942039

VonderHaar, R. L., McBride, D. M., & Rosenbaum, D. A. (2019). Task order choices in cognitive and perceptual-motor tasks: The cognitive-load-reduction (CLEAR) hypothesis. *Attention, Perception, & Psychophysics*, *81*(7), 2517–2525. https://doi.org/10.3758/s13414-019-01754-z

Warren, W. H., Kay, B. A., Zosh, W. D., Duchon, A. P., & Sahuc, S. (2001). Optic flow is used to control human walking. *Nature Neuroscience*, *4*(2), 213–216. https://doi.org/10.1038/84054

Wasserman, E. A., & Brzykcy, S. J. (2015). Pre-crastination in the pigeon. *Psychonomic Bulletin & Review*, *22*(4), 1130–1134. https://doi.org/10.3758/s13423-014-0758-3

Wei, K., & Koerding, K. (2010). Uncertainty of feedback and state estimation determines the speed of motor adaptation. *Frontiers in Computational Neuroscience*, *4*. https://doi.org/10.3389/fncom.2010.00011

Wei, K., & Körding, K. (2009). Relevance of Error: What Drives Motor Adaptation? *Journal of Neurophysiology*, *101*(2), 655–664. https://doi.org/10.1152/jn.90545.2008

Wickler, S. J., Hoyt, D. F., Cogger, E. A., & Hirschbein, M. H. (2000). Preferred speed and cost of transport: The effect of incline. *Journal of Experimental Biology*, *203*(14), 2195–2200.

Winter, C., Kern, F., Gall, D., Latoschik, M. E., Pauli, P., & Käthner, I. (2021). Immersive virtual reality during gait rehabilitation increases walking speed and motivation: A usability evaluation with healthy participants and patients with multiple sclerosis and stroke. *Journal of NeuroEngineering and Rehabilitation*, *18*(1), 68. https://doi.org/10.1186/s12984-021-00848-w

Wishart, L. R., Lee, T. D., Cunningham, S. J., & Murdoch, J. E. (2002). Age-related differences and the role of augmented visual feedback in learning a bimanual coordination pattern. *Acta Psychologica*, *110*(2), 247–263. https://doi.org/10.1016/S0001-6918(02)00036-7

Wodarski, P., Jurkojć, J., Polechoński, J., Bieniek, A., Chrzan, M., Michnik, R., & Gzik, M. (2020). Assessment of gait stability and preferred walking speed in virtual reality. *Acta of Bioengineering and Biomechanics*, *22*(1). https://doi.org/10.37190/ABB-01490-2019-03

Xu-Wilson, M., Zee, D. S., & Shadmehr, R. (2009). The intrinsic value of visual information affects saccade velocities. *Experimental Brain Research*, *196*(4), 475–481. https://doi.org/10.1007/s00221-009-1879-1

Yoon, T., Geary, R. B., Ahmed, A. A., & Shadmehr, R. (2018). Control of movement vigor and decision making during foraging. *Proceedings of the National Academy of Sciences*, *115*(44), E10476–E10485. https://doi.org/10.1073/pnas.1812979115

Yoon, T., Jaleel, A., Ahmed, A. A., & Shadmehr, R. (2020). Saccade vigor and the subjective economic value of visual stimuli. *Journal of Neurophysiology*, *123*(6), 2161–2172. https://doi.org/10.1152/jn.00700.2019