Mechanisms of energy optimization in human walking

by

Sabrina J. Abram

B.Sc. (Hons.), Cleveland State University, 2015

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
School of Engineering Science
Faculty of Applied Sciences

SIMON FRASER UNIVERSITY
Fall 2021

Declaration of Committee

Name:	**Sabrina J. Abram**
Degree:	**Doctor of Philosophy**
Title:	**Mechanisms of energy optimization in human walking**
Committee:	**Chair:** **Pierre M. Lane** Associate Professor, Engineering Science
	J. Maxwell Donelan Supervisor Professor, Biomedical Physiology and Kinesiology
	M. Faisal Beg Committee Member Professor, Engineering Science
	Ivan V. Bajić Examiner Professor, Engineering Science
	Amy J. Bastian External Examiner Professor, Neuroscience Johns Hopkins University

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University

or has conducted the research

c. as a co-investigator, collaborator, or research assistant in a research project approved in advance.

A copy of the approval letter has been filed with the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

Update Spring 2016

Abstract

Humans can learn to move optimally. For many movements, we have a control strategy—or control policy—that optimizes some objective. In walking, we prefer the combination of step widths, step lengths, and speeds that optimizes the amount of energy we need. In familiar contexts, we have had many opportunities to establish this optimal control policy. But in new contexts, the nervous system must quickly learn new control policies in order to continue to move optimally. Our lab has recently demonstrated that humans can continuously optimize energetic cost during walking. This is an impressive feat given that the nervous system has tens of thousands of motor units at its disposal, and it can coordinate these motor units over millisecond timescales, which results in countless combinations of motor unit coordination. The goal of this thesis is to determine how the nervous system navigates this combinatorial problem to learn new energy optimal control policies in new walking contexts. I used three distinct studies to accomplish this goal. For the first two studies, I designed and implemented a simple mechatronic system that applies energetic penalties in the form of walking incline as a function of gait. This creates a new relationship between gait and energetic cost—or new cost landscape—that shifts the energy optimal gait. For the third study, I used exoskeletons that apply assistive torques to each ankle at each walking step to shift the energy optimal gait. The first study tested whether previous findings that people can learn to adapt their control policy when the energy optimum is shifted along step frequency generalize to a different gait parameter and to a different experimental setup. I found that, like step frequency, people can learn to adapt their control policy when the energy optimum is shifted along step width. The second study tested if and how energy optimization extends to multiple gait parameters at the same time. I found that, when the energy optimum is shifted along step width and step frequency, people are limited in their ability to optimize both gait parameters. The third study asked how people learn in which ways to optimize their policy. I found that general variability leads to specific adaptation toward optimal policies. Taken together, these findings provide insight into the mechanisms that underlie energy optimization in walking, as well as the limitations of this optimization.

Keywords: energetics; optimization; locomotion; motor learning; motor variability; reinforcement learning

Acknowledgements

I would first like to thank my supervisor, Dr. Max Donelan. Max, thank you for your support and guidance over the years. I am grateful for the countless hours of your time that you have devoted to the projects in this thesis and to my development as a scientist. I am continually inspired by your passion and curiosity. You have taught me to ask thoughtful questions and to clearly communicate these questions. More importantly, you have taught me to prioritize lifelong learning, as well as experiences and relationships with colleagues, friends, and family.

I would next like to thank my committee members: Dr. Faisal Beg and Dr. Jessica Selinger. Faisal, thank you for showing me the beauty of linear systems theory years ago, and for your encouragement in the time since. Jess, thank you for your friendship and mentorship. You are an empowering presence in my life and career. You have taught me about breaking biases and stereotypes and have shown me what it looks like to be a strong woman in science. I am grateful for our many stimulating conversations over the years.

To my friends and labmates, thank you for the memories—the highs and lows of graduate school make for a special bond. I am grateful for our support system. I have enjoyed our regular coffees, as well as the occasional conference abroad.

A special thank you to the volunteers that participated in my experiments.

Thank you to my funding sources for providing me with many opportunities over the years. These funding sources include the National Sciences and Engineering Research Council of Canada, the Vanier Canada Graduate Scholarship, Simon Fraser University, and the Michael Smith Foreign Study Supplement. The Michael Smith Foreign Study Supplement provided me the opportunity to travel to the Biomechatronics Lab at Stanford University as a visiting researcher. I would like to thank Dr. Steve Collins for welcoming me into the Biomechatronics Lab, and Dr. Katherine Poggensee for kindly sharing her dataset with me.

Most importantly, I would like to thank my family. To my parents, Marion and Dan, you have given me the tools for life and have been there every step of the way, despite the distance. To my sister and brother, Vanessa and Derek, you have set high standards and shown me that they can be achieved. To my partner, Ian, you have been my light over the years, and I cannot wait for our future. Thank you all for your endless love and support.

Table of Contents

List of Figures

Glossary

Action	The way in which the nervous system interacts with the environment. The nervous system sends motor commands to the motor units of muscles, which are the actuators of the body. It has hundreds of muscles and even more motor units at its disposal, and it can vary each motor unit's activity many times per second.
Adaptation	The process of modifying a control strategy in response to a new situation. A new situation can result from changes to the body, task, or environment.
Control policy	A mapping between an estimate of the state of the body at a given time and an action for the next moment in time. This term is used interchangeably with the term *control strategy*.
Cost landscape	The relationship between gait and energetic cost. In this thesis, we alter this relationship to achieve a new relationship that shifts the energetic cost optimum.
Dimension	A parameter that the nervous system considers in its control policy. The total number of parameters is the dimensionality of the control policy.
Energetic cost	The amount of metabolic energy that the body consumes. This can be approximated by measuring rates of oxygen consumption and carbon dioxide production. This term is used interchangeably with the term *metabolic cost*.
Exoskeletons	Exoskeletons are wearable devices that seek to emulate and assist the body. In this thesis, exoskeletons apply assistive torques to each ankle at each walking step.
Exploration	Searching among candidate control policies to discover new optimal policies. In this thesis, exploration is approximated by the variability in the signal that is not considered noise.
Objective function	A function that assigns a cost to each movement. This function may consist of multiple terms, and the relative importance of these terms may depend on the situation. The optimal movement is one that minimizes the objective function.
Optimization	The iterative process of minimizing some objective function. How the nervous system converges on the optimum depends on the nature of its algorithms.
Reinforcement learning	An algorithm in which the nervous system interacts with its environment to learn an optimal or near optimal policy that maximizes the reward signal.

Reward	The signal that is maximized in reinforcement learning. An action selected by a control policy is considered good or bad according to the measured reward. The control policy can then be modified to improve future reward.
Reward prediction error	The difference between the predicted and the measured reward signal. A positive reward prediction error is when the reward signal is greater than predicted, and negative otherwise.
State	The nervous system's estimate of the state of the body or environment given its sensory inputs. The nervous system can take actions that affect its state.

Published Studies

Research contained in this thesis has been published, or is in revision, in the following peer-reviewed journals:

Chapter 2	Abram S.J., Selinger J.C., and Donelan J.M. (2019). Energy optimization is a major objective in the real-time control of step width in human walking. Journal of Biomechanics, 91, 85-91.
Chapter 4	Abram S.J., Poggensee K.L., Sánchez N., Simha S.N., Finley J.M., Collins S.H., and Donelan J.M. General variability leads to specific adaptation toward optimal movement policies. (in revision). Current Biology.

Chapter 1.

Introduction

The control of movement is complex, yet humans can learn to adapt their control strategies in various situations. For example, we adapt our walking speed according to the distance we need to travel and the time we have to do so, and we transition from walking to running to sprinting as our speed increases. In familiar situations, we have had many opportunities to learn which control strategies are successful. Perhaps we walk down the same driveway each morning to get the mail—we would do so roughly ten thousand times from age twenty to age fifty. But we also face new situations with changes to our environment, such as icy terrain, as well as changes to our body, such as muscle fatigue, in which we should adapt our control strategy to remain successful. Adapting to new situations becomes even more difficult when we consider the complexity of our nervous system's motor control. We have great flexibility in how we coordinate tens of thousands of motor units to walk from one point to another—there are many more combinations of motor unit coordination than there are moves on a chessboard. How does the nervous system navigate this combinatorial complexity to learn new control strategies? This can be viewed through the lens of optimization—the nervous system has an objective function for movement and optimizes this objective function in response to changes to our body, task, and environment.

One framework for motor learning is the combination of a fast process that reflects prediction and a slow process that reflects optimization. This framework can explain the coordination of many different movements such as reaching and walking [1–4]. A fast predictive process relies on feedforward control to rapidly select near optimal solutions, whereas a slow optimization process relies on feedback control to iteratively converge on more optimal solutions. Optimization may be ineffective given its relatively slow response times and large computational costs. But on the other hand, it may lead to the discovery of new optimal solutions in new contexts. The focus of this thesis is to better understand the nervous system's optimization algorithms.

The question of if and how humans learn new optimal control policies has been the focus of many decades of research. In this thesis, I use the term *control policy* to refer to the mapping between states and actions taken in those states. The nervous system's actions are the motor

commands it sends to muscle motor units, whereas its states are more ambiguous—it may estimate states ranging from lengths of individual muscles to walking speeds. A control policy can be considered optimal in terms of an objective function, where this objective function can include many terms that differ in their relative importance. For example, one term in the objective function for reaching might be target error—reaching for a cup of coffee with large error might result in spilling the coffee [5]. For walking, one term might be stability—walking with decreased stability can result in falling, which can have serious consequences especially for older adults [6]. Another term might be energy—for our human ancestors, calories were often scarce and conserving energy was important for survival [7–10]. These and many other candidate terms can combine to form an objective function that the nervous system optimizes for movement. To determine whether a control policy is optimal in some sense, we can begin by investigating one movement objective for one movement variable. We can vary this variable while measuring this objective to map this relationship, and then determine where people's preferences lie on this relationship. To determine whether the nervous system can continuously optimize this movement variable for this objective, we can artificially change this relationship and measure how people adapt their preferences.

The general aim of this thesis is to determine how the nervous system learns new optimal control policies. In this introduction, I will first provide a background of our understanding of the nervous system's control of movement. I will then discuss our understanding of how the nervous system adapts its control of movement and how we can develop computational models that predict this adaptation. Lastly, I will reframe these concepts in terms of a model system of energy optimization in human walking.

1.1. Human motor control

Humans use skeletal muscles to perform movements. Muscles can generate forces and these forces can move our limbs. In this section, I will briefly describe how muscle generates force, and then how force can be sensed and controlled. A muscle's architecture is the basis of force generation. A muscle is composed of many muscle fascicles; a muscle fascicle is composed of many muscle fibres; and a motor unit is an ensemble of muscle fibres that are innervated by the same motor neuron. Any given muscle fibre is innervated by only one motor neuron. An average sized muscle can be controlled by roughly 100 motor neurons [11].

Muscles are diverse and specialized in both structure and function. They can function as struts, brakes, or motors depending on the task [12]. They are classified as brakes when they produce negative power and as motors when they produce positive power. Here, I will describe how a muscle contracts to act as a motor because this is most relevant to motor adaptation. A muscle fibre is composed of many sarcomeres and each sarcomere is composed of thick (myosin) and thin (actin) filaments that attach by way of cross bridges [13]. A cross bridge is formed when the myosin head binds to the actin filament. The sliding filament theory proposes that actin filaments slide past myosin filaments in an attach, pull, detach process to shorten the muscle fibres and generate a contractile force—a process powered by energy from the hydrolysis of adenosine triphosphate (ATP) to adenosine diphosphate (ADP) [14,15]. Muscle structure can also vary between muscles and across tasks. For example, muscle fibres are composed of organelles such as mitochondria, i.e., energy supply, and their proportion changes between slow and fast muscle fibre types as well as between species [16]. These slow and fast muscle fibres exhibit low and high maximal shortening velocities and can be recruited for different tasks where they operate at their optimal velocity—for instance, faster fibres in the medial gastrocnemius are preferentially recruited as pedaling cadence increases in cycling [17,18].

The diverse and specialized nature of muscle lies not only in its structure and function, but also in how it is controlled. The nervous system interacts with muscle at the neuromuscular junction, where motor neurons and their axons meet muscle fibres [11]. The nervous system can send a signal that travels along an axon to the neuromuscular junction to generate an action potential, which causes an increase in calcium levels in a muscle fibre. This calcium release is needed for the myosin heads to bind to the actin filaments and generate a muscle contraction. The nervous system's control can differ in which muscle fibres are recruited, how this recruitment is modulated over time, and how this recruitment is coordinated with other muscles, all of which likely depend on the task at hand [19]. The nervous system faces a difficult problem—it has tens of thousands of motor units at its disposal, and it can vary each motor unit's activity many times per second. Several studies have proposed that muscle synergies may simplify this problem. They found that muscle activation patterns can be explained by a limited set of muscle synergies—muscle activation patterns with consistent spatial and temporal characteristics [20–23]. However, it is still unclear whether the nervous system considers motor units, muscles, groups of muscles, or even whole-body parameters in its control of movement.

The nervous system must be able to sense movement to control movement. The nervous system has sensors that estimate the state of our muscles, joints, and limbs, and receives many different sensory signals many times per second [11]. For example, visual signals make measurements that help estimate the location of our limbs in our environment [24]. Muscle spindles can help estimate a muscle's length and velocity, and Golgi tendon organs can help estimate the force generated by a muscle [25]. These and other sensory signals are carried by sensory neuron axons to the central nervous system to be processed [11]. The nervous system may need to combine sensory signals to form an estimate of a particular state, or even estimate a state that is not directly sensed [26,27]. The process of combining sensory signals with a copy of motor commands to form an estimate of how motor commands affect states is known as state estimation. Over the past 50 years, researchers have used both computational and experimental approaches to help understand how the nervous system implements state estimation. A well-known model for this is the Kalman filter, which is well-suited for time-varying systems like human movement [28]. A Kalman filter seeks to reduce the contribution of two features of sensory signals—noise and delay. It uses a forward model of body dynamics to predict how motor commands might affect states, compensating for time delays in sensory feedback signals. It also uses a forward model of sensors to predict sensory feedback and then compare this to measured sensory feedback. It compensates for signal noise using a weighted combination of predictions and measurements [29]. This model is supported by evidence of state estimation of hand position [30] and human postural balance [31]. The nervous system can estimate any given state using these methods—as previously mentioned, this could range from lengths of muscles to walking speeds—to then select the appropriate motor commands.

The process of sensing a particular state and commanding the appropriate action requires the interaction of signals from our sensory system and motor system. This interaction is referred to as sensorimotor control. The nervous system can process signals in two ways—some signals travel only to the spinal cord, whereas others travel through the spinal cord to the brain [11]. The sensory signals that travel to the spinal cord but not the brain are called short-latency reflexes, and the sensory signals that travel through the spinal cord to the brain are called long-latency reflexes [32–34]. A short-latency reflex is useful because it provides a fast response to a stimulus—the monosynaptic stretch reflex provides direct communication between the sensory neuron and motor neuron. Still, the time between a stimulus and peak muscle force is relatively long—about 120 milliseconds for muscles at the ankle, which is about the total amount of time the foot is in contact with the ground during sprinting [35].

Long-latency reflexes add even more time [34]. However, with this added time, the nervous system can process feedback signals for adaptive control.

1.2. Motor learning in new contexts

The nervous system may have to modify motor behaviour in response to changes to our body, task, and environment. Changes to our body, for example, can occur over long timescales such as evolution and development, as well as over short timescales such as the onset of fatigue. In this thesis, I will focus on changes that occur over short timescales—that is, over seconds, minutes, hours, or days—that require we adapt our previously learned control policy to learn new optimal policies. The process by which the nervous system learns new optimal policies is of great interest for facilitating and accelerating learning, and it is still being understood. In this section, I will present candidate processes that drive the nervous system's learning, as well as possible challenges faced by the nervous system.

Adaptation refers to the process of modifying motor behaviour [36]. This has been studied in tasks such as eye movements [37], reaching movements [38,39], and locomotion [40]. We can experimentally alter these tasks in such a way that requires people to adapt their movement. For example, we can study adaptation by applying forces that distort reaching or walking [38,41]. We can also apply forces to the knee [42], or alter the relative speeds at which the left and right legs travel when in contact with the ground [40,43]. We can then investigate the mechanisms that underlie the observed adaptation. One theory is that adaptation arises from the nervous system updating an internal model. Wolpert et al. (1995) define an internal model as a relationship between input and output that mimics a natural process [30]. It is generally accepted that internal models take two forms: forward models and inverse models. Forward models predict the sensory consequences of motor commands [44], and the nervous system may use these models when estimating the state of its body and environment [26,45]. It may also use forward models for adaptation, updating its predicted sensory consequences in response to movement errors [39]. However, new evidence challenges this view and proposes that forward models are not responsible for adaptation as they do not plan future motor behaviour [46]. Inverse models, on the other hand, do plan future motor behaviour by selecting actions, or motor commands, that achieve the desired state from the current state, which is also the function of a control policy [44]. It is important to note that this thesis does not seek to study learning in terms of forward models and inverse models. It also does not

attempt to dissociate the two. Rather, the aim of this thesis is to determine how the nervous system adapts its control policy—a mapping between states and actions—to learn new optimal policies. It is useful to understand how this approach fits with prominent theories such as internal models.

Our understanding of the signal that drives adaptation in certain situations continues to be developed. Here, I will focus on two theories of motor adaptation that involve two different signals: error-based learning and reward-based learning. Hermann von Helmholtz (1867) was one of the first to study error-based learning [47]. He used prism glasses to shift the visual input by a certain degree, which gives rise to movement errors for basic tasks—when reaching to a target with the shift in the leftward direction, we overshoot reaches in that same direction. But we can learn to adapt our reach direction to compensate for movement errors with practice. And when the prism glasses are removed, we produce movement errors in the opposite direction [48]. This behaviour is indicative of sensorimotor remapping—we don't immediately recall past movement strategies when the prism glasses are removed. Further evidence of error-based learning came from the task of walking on a split-belt treadmill [40]. The split-belt treadmill creates a difference between the speeds at which the left and right legs travel when in contact with the ground. When initially walking on a split-belt treadmill, we exhibit longer steps on the slow-moving belt and shorter steps on the fast-moving belt—the comparison between the two step lengths is referred to as *step length asymmetry*. The nervous system selects its preferred motor commands and expects a symmetric gait but receives feedback that the gait is asymmetric. Like the task of reaching with prism glasses, we can learn to adapt our step lengths to near zero step length asymmetries with practice. And when the belt speeds are again matched, we produce step length asymmetries in the opposite direction. This evidence is consistent with the hypothesis that the nervous system uses a sensory prediction error between predicted and measured sensory signals to adapt the motor commands [39,43].

Reward-based learning relies on a signal that indicates whether a movement is good or bad. The notion of good or bad can be formalized in many different ways such as success or failure; pleasure or pain; or more or less economical. An alternative view of split-belt walking is that people adapt to be more economical, i.e., reduce metabolic cost [49]. In terms of error-based learning, it is not immediately clear why the nervous system would seek to reduce step length asymmetry when this increases step time asymmetry [50]. New evidence on prolonged split-

belt walking suggests that people learn to adopt positive asymmetries with foot placement further forward on the fast belt rather than near zero asymmetries [51,52]. Such a strategy takes advantage of the positive work performed by the treadmill—due to one belt moving faster than the other—and reduces the work performed by the legs as well as metabolic cost. The collective findings of adaptation in split-belt walking provide a more unified framework for understanding the signal that drives this adaptation. For many other real-world tasks, there is not a clear error signal that indicates in which direction we should adapt. The nervous system may instead use a reward prediction error, or difference between predicted and measured reward, as a teaching signal for learning new movement strategies [53]. We can experimentally manipulate the reward associated with a particular feature of movement, such as reach direction and curvature, and then study how people adapt this feature [54]. For example, people can learn to adapt their movements in response to binary reward feedback that indicates success or failure [55]. People can also learn to adapt their movements in order to maximize variable feedback in the form of monetary reward [56]. The nervous system may have some internal representation of reward to adapt its movement in this way. For example, it may have a subjective measure of utility or usefulness of motor commands [57]. It may also use the same processes but consider the opposite of reward, i.e., penalty or cost [56]. In this thesis, I study reward-based learning in terms of metabolic energy cost. The nervous system may have some measure of energy cost, compare this against its predicted cost, and then adapt its movement to minimize this cost.

While it is not the focus of this thesis, it is essential to probe neural mechanisms of processes such as error-based learning and reward-based learning. In error-based learning, there is evidence that supports the involvement of the cerebellum in computing sensory prediction errors. This has been shown in humans and other animals, as well as for various tasks [26,39]. A useful approach is to study adaptation in humans with cerebellar damage. For example, Morton and Bastian (2006) showed that, in the task of split-belt walking, cerebellar damage disrupts predictive feedforward processes but not reactive feedback processes [40]. This is also the case in reaching tasks [58]. In reward-based learning, there is evidence that supports the involvement of the basal ganglia in computing reward prediction errors. Dopamine is a neurotransmitter that is involved in reward processing, and the axons of dopamine neurons appear to convey signals to brain areas involved in learning and decision making such as the basal ganglia [53]. This has been shown in the response of dopamine neurons when monkeys reach to a bin that does or does not contain food—they experienced a reward prediction error

when the food outcome was different from what they expected [59]. Another useful approach for studying reward-based learning is in Parkinson's patients where there is degeneration of dopamine neurons. It is important to note that these studies point to the involvement rather than the role of specific brain areas in these processes. Even with sophisticated techniques such as neuroimaging, it is difficult to probe neural mechanisms [60].

The process of optimizing a control policy faces several challenges. This is in large part due to complexity of the nervous system's control. This complexity was realized almost a century ago by Nikolai Bernstein—one of the first to study the control of multi-joint movements [61]. To understand this complexity, consider musculoskeletal models such as OpenSim [62]. Even with high computing power, sophisticated algorithms, and fewer muscles than humans, it is difficult for these models to recreate human movement, and even more difficult to learn to adapt movement [63]. One challenge associated with learning new optimal control policies is the explore-exploit dilemma [64]. Exploration is risky but can lead to more optimal solutions, while exploitation is safe but can be suboptimal. Such trade-offs are ubiquitous in learning and decision-making problems. A second challenge is the curse of dimensionality [53]. As the size of the control policy increases incrementally, the space in which we explore for new optimal policies increases rapidly. A third challenge is the credit assignment problem [65]. When the nervous system receives some change in a reward signal, it must determine which of countless parameters is responsible for this change and then adapt this parameter in order to achieve larger reward in the future.

How does the nervous system overcome such challenges to learn new optimal policies? I will focus on two mechanisms that the nervous system may rely on: structural learning and use-dependent learning. Structural learning refers to accelerated learning in a new situation with shared features of previously experienced situations. For example, in a reaching task where the visual input is rotated to some variable angle at each reach, participants more quickly learn a novel rotation [66]. They also appear to use prior assumptions about the structure of the task, and rely on explicit strategic resources to do so [1,67]. This can narrow the nervous system's search space in new situations. Use-dependent learning refers to biases in future movements that are due to previous movements. For example, when participants passively experience a strategy, i.e., initial reach angle of some degree, while achieving the task goal of reaching to the target, they adopt this arbitrary strategy in future movements [68]. This can also narrow the nervous system's search space in new situations. The nervous system may

also rely on contextual cues. One study showed that, in learning two opposite visual rotations, an implicit contextual cue, such as the wrist performing the counter-clockwise learning and the arm performing the clockwise learning, may be more effective than an explicit one, such as a verbal or color cue [69]. The nervous system has more degrees of freedom than needed for most tasks, and it may need to use these and other strategies to manage the complexity and redundancy in its control of movement.

1.3. Computational models of motor learning

We can use computational models to develop a more unified theoretical framework for motor control. The goal of any model is to make predictions about the behaviour of a given system. Optimal control models have had great success in predicting the control of human movement [70]. These models are attractive because they can predict various movements using a cost function, where the optimal control policy minimizes this function. However, there is no one cost function for movement—the specific costs that govern movement differ between tasks, and their relative importance change between conditions. Optimal control models have two categories: feedforward and feedback (Figure 1.1). Optimal feedforward control models plan the optimal sequence of motor commands while ignoring online sensory feedback. Optimal feedback control models take into account online sensory feedback to learn optimal motor commands, allowing for flexible behaviour. One theory of motor control that is based on optimal feedback control navigates the abundance of possible solutions by choosing the best possible solution while making little effort to correct for behaviour that does not affect task performance. This hypothesis has been quantified using a method known as the uncontrolled manifold and directly addresses the challenges highlighted above in Section 1.2. It has been supported with evidence in movements such as reaching [71], and can explain coordination among groups of muscles, or muscle synergies [20–22]. But it does not speak to the *process* by which the nervous system learns new optimal control policies.

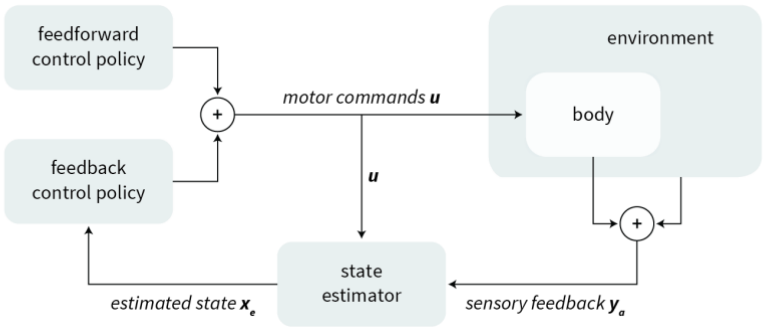

Figure 1.1: Motor control
A conceptual diagram illustrating the nervous system's motor control. Here, the controller combines a feedforward control policy with a feedback control policy. Both control policies must be learned by the nervous system. The feedforward control policy generates motor commands as a function of time. The feedback control policy is a mapping between the estimated state at a given time and the action, or motor command, for the next moment in time. The actual state likely differs from the desired state due to imperfections in the nervous system's control. The nervous system can estimate the actual state by combining sensory signals with a copy of motor commands, which can help compensate for time delays and filter sensory noise.

Reinforcement learning is one possible way in which the nervous system can interact with the environment to learn new optimal control policies. This framework is built on principles of trial-and-error learning and decision-making [53]. The main components of reinforcement learning are the policy, value function, and reward signal (Figure 1.2). A policy determines how the learner behaves, a value function specifies how good this behaviour is in the long run, and a reward signal is that which the learner tries to maximize [53]. Here, I will focus on one category of reinforcement learning—value iteration—and describe this algorithm as it relates to the work in this thesis—energy optimization in walking. In this case, the nervous system's value function is some predicted relationship between gait and energy cost. At each step, it selects an action that obeys its policy to choose the energy optimal gait. After executing this action with some execution noise, the nervous system measures the reward by sensing energy cost with some measurement noise, and then uses this reward to iteratively update its predicted value function. Reinforcement learning models are well-suited for studying the process of energy optimization in walking—they can learn new optimal movement policies in simulation [72], they can predict adaptation toward new energy optimal policies in human walking experiments [73], and the brain appears to have the systems needed to compute the difference between predicted and measured reward, or the reward prediction error [74].

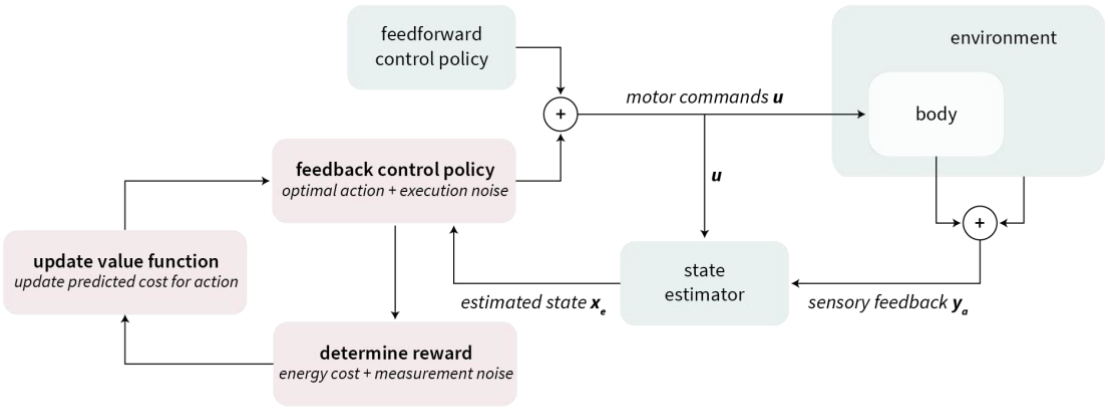

Figure 1.2: Motor learning
A conceptual diagram illustrating how the nervous system can learn new optimal control policies in new contexts. Blue represents processes associated with motor control and red represents processes associated with motor learning in terms of reinforcement learning. At each step, the reinforcement learner selects an action in accordance with its control policy and executes this action with some noise. It then measures the reward, or energy cost, with some noise and updates its predicted value function.

1.4. Energy optimization in human locomotion

Our understanding of the nervous system's objective function, or cost function, for movement continues to be developed. The specific terms of this function, and their relative importance, change depending on the context. Possible terms include stability, maneuverability, effort, or error, to name a few. Several studies provide evidence for error and effort as terms in the nervous system's objective function [41,75]. For example, people learn to adapt their control policy in a new context where a robotic device applies upward forces during the swing phase of gait, and this adaptation can be modeled as the optimization of an objective function that includes muscle effort and kinematic error [41]. A large body of evidence suggests that effort in terms of metabolic cost is a dominant term in the nervous system's objective function. For example, when reaching to a target with forces being applied to the arm, people learn to adapt their reach to reduce error, and this adaptation is accompanied by decreases in metabolic cost. Reaching movements are also stereotypically smooth, where smoothness has in the past been explained by minimizing jerk—the rate at which acceleration changes with time [76]. But new evidence suggests that this smoothness can also be explained by minimizing energy [77]. We can gain insight into the nervous system's objective function by measuring changes in behaviour in response to changes in this function.

A central objective in human walking is to minimize metabolic energetic cost. Walking can be varied in a number of ways that influence this cost. For example, energetic cost increases with increases in both walking speed and running speed [78,79]. When considering the energetic cost to travel a unit distance, or cost of transport, the relationship with speed is bowl-shaped and people prefer to walk at an intermediate speed that minimizes this cost of transport [7]. Energetic cost is also dependent upon other gait parameters, such as step frequency and step width, where these relationships are similarly bowl-shaped. For example, at any given speed, people prefer to walk with a step frequency that minimizes energetic cost [10]. People also prefer to walk near the energetically optimal step width, avoiding high step-to-step transition costs at wide widths, as well as high lateral limb swing at narrow widths [80]. This previous understanding of preferred gait was that it is energetically optimal in familiar contexts, but it did not provide insight into the timescale by which this preference is established. To test this, Selinger and colleagues used robotic knee exoskeletons to reshape the relationship between gait and energetic cost, making abnormal gaits energetically optimal, and found that people discover these gaits within seconds and prefer to walk with them [42]. They demonstrated that preferred gait arises not just through evolution and development, but that the nervous system can continuously optimize energetic cost during walking [42].

The nervous system appears to have the necessary physiological systems to perform energy optimization. Humans and other animals implement algorithms that are strikingly similar to reinforcement learning algorithms [53]. In fact, many aspects of reinforcement learning are influenced by neuroscience. To implement these algorithms, the nervous system must have some representation of states, rewards, and actions. To estimate the state of the body, the nervous system may need to combine many sensory signals that are both noisy and delayed with a copy of motor commands [28]. There is evidence that the nervous system is capable of this sort of state estimation of hand position [30] and human postural balance [31]. But it is unclear exactly which states the nervous system estimates. For example, it may estimate a muscle's length and velocity from muscle spindles, or the force that is generated by a muscle from Golgi tendon organs [25]. The nervous system also needs to estimate reward, or in this case energetic cost. To do so, it may use sensory signals from blood gas receptors—which are sensitive to oxygen and carbon dioxide [81]—or group III and IV muscle afferents—which are sensitive to the byproducts of muscle metabolism [82]. Like estimating the state of the body, the nervous system may need to combine many sensory signals to estimate energetic cost. Lastly, the nervous system needs to evaluate these states and rewards in order to select

the appropriate actions. According to reinforcement learning theory, reward prediction error can drive action selection for learning. There is compelling evidence that the nervous system tracks reward prediction error, or difference between predicted and measured reward [74]. The first evidence of this came in the 1990s from Wolfram Schultz, where he measured the response of dopamine neurons in monkeys while they reached to a bin that either did or did not contain food—they experienced an error when the food outcome was different from what they predicted [59]. The axons of these dopamine neurons appear to convey this error signal to brain areas involved in learning and decision making such as the basal ganglia [53]. New research continues to shape our understanding of these signals and their functions, as well as where they are implemented in the brain.

While we have evidence that people can continuously optimize energetic cost during walking, it is unclear exactly how the nervous system achieves this. To further study specific aspects of the nervous system's optimization algorithms, Selinger and colleagues again used knee exoskeletons to alter the relationship between gait and energetic cost and create a new cost landscape [73]. They found that people with high natural gait variability can spontaneously initiate optimization, whereas people with low natural gait variability required experience with a low energetic cost to initiate optimization. Once optimization was initiated, people converged on a new energy optimal control policy and learned to predict this policy. Lastly, they modeled energy optimization as reinforcement learning and found that this model can predict people's behaviour. A different study used the same knee exoskeletons to create new cost landscapes and test for energy optimization in overground walking rather than treadmill walking [83]. Contrary to the hypothesis that increased natural gait variability in overground walking can spontaneously initiate optimization, they found that people did not adapt their gait in the new cost landscape. One reason for this may be that it is difficult for the nervous system to determine which aspects of gait to adapt in overground walking as it faces a credit assignment problem—changes in energetic cost can be due to many aspects of gait such as step frequency or speed, which is not the case in other studies where speed is constrained in treadmill walking. The initiation of optimization depends on several factors, and these factors do not act in isolation. One factor is the variability about the preferred gait. A second factor is the cost gradient about the preferred gait. Both influence the range of cost savings that the nervous system experiences. In another study, Simha and Donelan built a new system to test whether the nervous system relies on the energy cost gradient to initiate optimization [84]. This system applies forces to the user's torso to create new cost landscapes with steep or

shallow cost gradients. Contrary to the hypothesis that a steeper cost gradient can initiate optimization, we found that people did not spontaneously initiate adaptation in response to any of the gradients [85]. Instead, they all required enforced experience with a lower cost gait to initiate adaptation. These studies use a number of experimental setups to ask if and how the nervous system optimizes energetic cost during walking. They seek to uncover features of the nervous algorithms and provide the foundation for others to test additional hypotheses about energy optimization.

1.5. Aims

The overall aim of this thesis is to better understand the principles that govern human locomotion. To accomplish this, I use energy optimization in walking as a model system for learning. Like other model systems, the findings are meant to generalize to other situations: objective functions other than simply energetic cost and movements other than walking. Decades of research have provided evidence that humans prefer to move in energy optimal ways [7,8]. More recently, our lab has demonstrated that this preference arises not just over evolutionary and developmental timescales, but that the nervous system can continuously optimize energetic cost during walking [42,84,86]. Optimizing a control policy faces several challenges such as the explore-exploit dilemma and curse of dimensionality [53]. The specific aims of this thesis seek to determine how the nervous system overcomes such challenges to learn new optimal control policies in new contexts.

Aim 1: Determine if the nervous system's optimization algorithms generalize. Our lab previously found that people can continuously optimize step frequency in response to new energetic cost landscapes. The purpose of this first study was to generalize these findings to another gait parameter—step width—and another experimental setup. I selected step width because this gait parameter influences energetic cost and people prefer to walk with a step width that is near the energy optimal width in familiar contexts. But this does not necessarily mean that energy drives the nervous system's real-time control of step width in new contexts. I hypothesized that, like step frequency, the nervous system controls step width to optimize energetic cost. To test this hypothesis, I built a device that applies energetic penalties in the form of walking incline as a function of step width, creating new cost landscapes that shift the energy optimal step width to wider widths. I determined whether participants adapted their

step width toward the new energy optimal width spontaneously, and then after being given enforced experience in the new cost landscape.

Aim 2: Determine if and how the nervous system's optimization algorithms scale. How its algorithms scale from optimizing one gait parameter to optimizing two gait parameters may provide insight into how it defines the control policy search space. The nervous system may need to define a low dimensional search space to overcome challenges associated with searching a high dimensional space. In this second study, I combined computational models with human experiments to gain insight into how the nervous system defines its search space. I used simple reinforcement learning models to quantify how shifting the energy optimum in one vs. two gait parameters influences the search space and timescale of adaptation. I next compared these model predictions to human experiments. I used the device from Aim 1 to test if people learn to adapt their control policy when the energy optimum is shifted along both step width and step frequency, and how the timescale of adaptation changes when the energy optimum is shifted along one gait parameter compared to two gait parameters.

Aim 3: Determine how the nervous system learns which ways to optimize. One strategy for learning which ways to optimize is to explore. In this third study, I collaborated with the Stanford Biomechatronics Lab to ask how the nervous system explores through variations in its control policy to identify more optimal policies in new contexts. I created new contexts using exoskeletons that apply assistive torques to each ankle at each walking step and then gave participants experience with this context over multiple days. I tested three hypotheses. First, the nervous system initially explores by increasing general variability—where general variability refers to variability across many or all aspects of gait—to identify variables that improve its objective. Second, the nervous system selectively decreases variability to refine its control policy search space with experience. And third, the nervous system learns to adapt the magnitude of specific variables and exploit a new control policy that reduces energetic cost.

These findings provide insight into how the nervous system navigates the space of control policies to learn new optimal policies in new contexts.

Chapter 2.

Energy optimization is a major objective in the real-time control of step width in human walking

2.1. Abstract

People prefer to move in energetically optimal ways during walking. We recently found that this preference arises not just through evolution and development, but that the nervous system can continuously optimize step frequency in response to new energetic cost landscapes. Here we tested whether energy optimization is also a major objective in the nervous system's real-time control of step width using a device that can reshape the relationship between step width and energetic cost, shifting people's energy optimal step width. We accomplished this by changing the walking incline to apply an energetic penalty as a function of step width. We found that the people did not spontaneously initiate energy optimization, but instead required experience with a lower energetic cost step width. After initiating optimization, people adapted, on average, 3.5 standard deviations of their natural step width variability towards the new energy optimal width. Within hundreds of steps, they updated this as their new preferred width and rapidly returned to it when perturbed away. This new preferred width reduced energetic cost by roughly 14%, however, it was slightly narrower than the energetically optimal width, possibly due to non-energy objectives that may contribute to the nervous system's control of step width. Collectively, these findings suggest that the nervous systems of able-bodied people can continuously optimize energetic cost to determine preferred step width.

2.2. Introduction

When we walk, we tend to prefer a particular step width and execute this preference with remarkably small variability. For young healthy walkers, preferred step width is about 12 cm and varies between steps by less than two centimeters [80,87]. However, this preference can be influenced by walking context. For example, increases in speed result in decreases in step width and increases in its variability [88]. Another context is stability—preferred width and its variability decrease with external lateral stabilization [87,89] and increase with visual field perturbations [6,90,91]. Step width has also been found to increase with both age and obesity [92–94]. One way to understand how the nervous system controls step width is to identify the relative importance of different walking objectives in response to changes in walking contexts.

One well established objective of the nervous system is to optimize metabolic energetic cost. Decades of research show that people prefer to move in energetically optimal ways [7,8,95]. In the case of step width, people prefer to walk near the energetically optimal width, avoiding high step-to-step transition costs at wide widths, as well as high lateral limb swing and active lateral stabilization costs at narrow widths [80,87,96–98]. When steps become wider, step-to-step transition costs increase because of increases in the mechanical work required to redirect the center of mass velocity from one step to the next [96]. Whereas when steps become narrower, active lateral stabilization costs increase because of increases in the control required to remain stable [87]. And when steps become narrower than the width of the foot, lateral limb swing costs increase because of the mechanical work required to laterally move the swing leg to avoid the stance leg [98]. These individual costs, and as a consequence, the energetically optimal step width, depend upon many factors including some that change with everyday circumstances. For example, the energetically optimal step width may differ with each shoe change as a result of the shoe compliance reshaping the transition cost, and the shoe width and mass reshaping the limb swing cost.

Although the preferred width is near the energy optimal width in familiar walking conditions, this does not necessarily mean that energy optimality drives the nervous system's real-time control of width. First, this preference may arise from long time-scale processes such as evolution and development [95,99–102]. For example, recent evidence suggests that narrow step widths are more feasible in humans compared to our closest living relatives,

chimpanzees, as a result of adaptations to our hips and knees [102]. Second, the nervous system likely considers objectives other than energy in determining preferred step width. As one example of a step width objective function, the nervous system may seek to simultaneously optimize energy, stability, and maneuverability [103]. These objectives may be independently weighted, and their weightings may depend upon walking contexts. For example, the priority for stability may be increased when the consequences of falling are more severe [104,105]. Some objectives may also be treated as constraints, such as walking on stepping stones where only some widths result in stepping on the stones. We can gain insight into if and how the nervous system represents the real-time control of step width as a multi-term optimization by determining how changes in individual objectives, relative to all other objectives, influence the control of width. If the nervous system weights other objectives higher than energy in its control of step width, then we expect people to persist at or near their original preferred width when the energetic context changes.

The purpose of this study was to test if energy optimization is a major objective in the real-time control of step width. Our lab previously found that people adapt their step frequency to converge on new energy optima, demonstrating that energy is indeed a major objective in the real-time control of some aspects of walking [42]. It is not clear, however, that this finding will apply to step width as the nervous system's control of these two gait parameters is typically thought of differently [6,97]. And while there is evidence that both the preferred frequency and width are energetically optimal [80,106], a common alternative view is that step width is primarily determined by stability [107]. Here we hypothesize that step frequency and width are similarly controlled to optimize energy. To test this hypothesis, we shifted the energy optimal width and observed the nervous system's response. To accomplish this, we built a custom device that applies energetic penalties, in real-time, as a function of measured step width. With this study, we aim to test the generality of energy optimization by using a second method of applying energetic penalties and by studying a different gait parameter.

2.3. Methods

2.3.1. Experimental design

After Simha and Donelan [84], we built a simple mechatronic system to shift people's energy optimal step width. In this system, participants walked on an instrumented split-belt treadmill at 1.25 m/s (FIT, Bertec Corporation, Columbus, OH, USA). To shift people's energy optimal width, we commanded the treadmill incline based on the desired energetic penalty for the step width measured from the previous step (Figure 2.1A). We implemented this closed loop control of incline based on measured step width using Simulink Real-Time Workshop running at 200 Hz (Simulink Real-Time Workshop, MathWorks Inc., Natick, MA, USA). We first filtered measured forces and moments using a second-order, low-pass, one-way, digital Butterworth filter (15 Hz cut-off), and then calculated the left and right lateral centers of pressure by dividing the left and right lateral moments by their vertical forces [108]. We identified foot contact events from the measured ground reaction forces and moments at the beginning of double support, and estimated step width by taking the difference between the lateral centers of pressure for consecutive steps. The incline changes were limited to 0.5 degrees per second, equating to a shift in vertical position under the participant of approximately 8 mm per second. This change was subtle, and no participants reported feeling destabilized.

To eliminate the concern that participants may be adapting their width to achieve level walking rather than to optimize energy, we first modified our design to have them always walking at an incline, and then decided to compensate for this additional energetic penalty above level walking with a forward horizontal force that acted as an energetic reward. The combination of controllable walking incline with a near constant forward horizontal force added a degree of novelty to the experimental design, perhaps making it more likely that the nervous system would search for a new optimal width rather than rely on a prediction of the preferred width. To apply the forward horizontal force, we connected a tensioned cable in series with long rubber tubing to a hip belt worn by the user—the long and compliant rubber tubing allowed for small changes in the walking position on the treadmill without large changes in the applied horizontal force. We adjusted the tension to apply a 10% body weight force, on average, and monitored the applied horizontal force with a load cell mounted on a

hip-belt (LCM201, Omega Engineering, Norwalk, CT, USA). This forward horizontal force did not depend upon step width or incline and remained nearly perpendicular to the ground.

We used this system to reshape the relationship between step width and energetic cost by providing energetic penalties as a function of step width. We define the relationship between gait and total energetic cost as *cost landscape* and the relationship between gait and energetic penalties as *control function*. We simulated our new cost landscape (red; Figure 2.1C) by adding literature values for energetic costs associated with walking at different step widths (grey; Figure 2.1C) with our control function (Figure 2.1D). We designed our control function (Figure 2.1D) by solving for the walking incline (Figure 2.1E) [109], combined with a 10% body weight forward horizontal force (Figure 2.1F) [110], to achieve an energetic penalty of 1 W/kg at a step width of 3 standard deviations (SD) narrower than initial preferred, decreasing with a constant slope to 0 W/kg at the new energy optimal width of 3 SD wider than initial preferred:

$$Incline = \begin{cases} -0.49 * sw + 5.16 & sw \leq 3 \\ -0.49 * 3 + 5.16 & sw > 3 \end{cases} \quad (2.1)$$

where step width (sw) is in units of standard deviations from preferred. We normalized step width by the variability in width to allow us to distinguish between a shift in step width occurring as a result of energy optimization, and a shift occurring by random chance. We chose to shift the energy optimal width wider than that initially preferred in order to be in the opposite direction of people's tendency to narrow their step width as they become more comfortable walking on a treadmill [111].

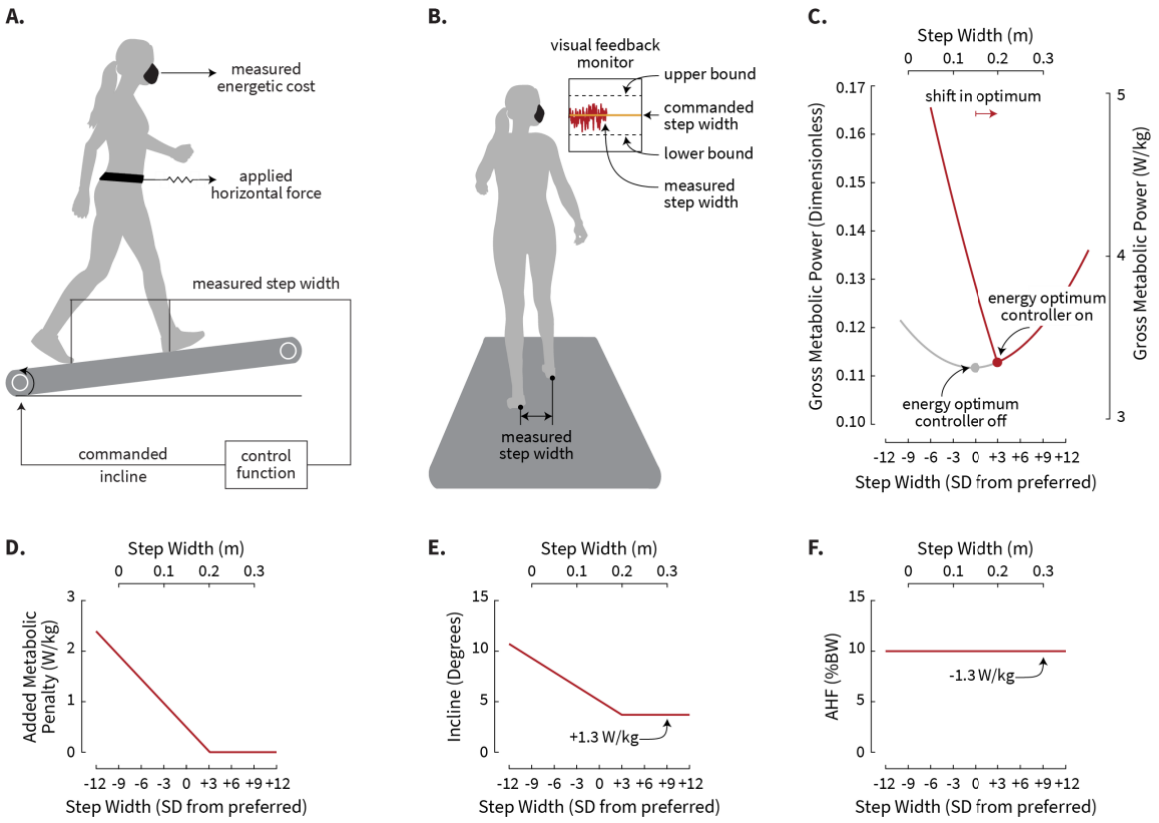

Figure 2.1: Experimental setup and design
(A) Our real-time controller uses forces and moments measured by the treadmill to identify foot contact events, calculate step width, and then command the appropriate treadmill incline based on the desired energetic penalty for the measured step width. (B) Our real-time visual feedback system allows us to enforce specific widths at times during the protocol by instructing participants to keep their measured step width signal within small bounds about the commanded width. (C) We simulated our new cost landscape (red) by adding literature values for energetic costs of walking at different step widths (grey) with (D) our control function that applies energetic penalties as a function of step width. To achieve this control function, we applied (E) an energetic penalty of walking incline as a function of step width and combined this with (F) a near constant energetic reward of a forward horizontal force applied to the user.

2.3.2. Experimental protocol

Eight participants (mean ± SD; body mass: 62.6 ± 9.0 kg; height: 165.3 ± 9.5 cm; sex: 5 females, 3 males) participated in the study. All participants were healthy and had no known gait or cardiopulmonary abnormalities. The Simon Fraser University Research Ethics board approved the protocol and the participants gave their written, informed consent before completing the experiment.

First, participants completed a baseline trial to measure their initial preferred width and width variability (Figure 2.2A). They completed this while walking on the level, without

treadmill incline control or forward pulling force. We calculated initial preferred width as the average width during the final three minutes of this 12-minute baseline trial. We calculated each participant's width variability as the standard deviation during the same averaging window. For each participant, we used their initial preferred width and width variability to design the control function.

Next, we turned the incline controller on, applied the forward horizontal force, and measured whether participants adapted their step width towards the new energy optimal width. First, we measured whether participants would adapt their width spontaneously, calculated as the average width during the final three minutes of the first six minutes after the controller was turned on (pre; Figure 2.2B). We estimated that six minutes would be sufficient to test for spontaneous energy optimization of step width as we have previously observed a time constant of roughly 66 seconds in step frequency optimization [73]. Second, we measured whether participants would adapt their width when given enforced experience in the new cost landscape. This enforced experience consisted of eight perturbations—each perturbation included a five-minute hold at either higher costs (6 SD narrower than initial preferred) or lower costs (6 SD wider than initial preferred), followed by a five-minute release, during which participants self-selected their widths (Figure 2.2B). During each hold, we instructed participants to keep the real-time calculated step width, normalized to the commanded width, within small bounds about 1 presented visually on a computer monitor. When the visual feedback targets disappeared, we instructed them to walk however they liked, with any step width. We did not provide participants with any information about how the controller worked. We calculated the change in preferred width after each perturbation as the average width during the final three minutes of the five-minute release. We calculated the final preferred width as the average width during the final three minutes of this experience period (post; Figure 2.2B).

Lastly, participants completed a cost mapping trial with the incline controller on and forward horizontal force applied. We used visual feedback to enforce steady-state walking at different widths (in random order) for five minutes each (Figure 2.2C). We used respiratory gas analysis (Vmax Encore Metabolic Cart, ViaSys, Conshohocken, PA, USA) to measure rates of oxygen consumption and carbon dioxide production. Participants walked at widths both about the initial preferred width (-2, 0, +2 SD from initial preferred) and about the final preferred width (-2, 0, +2, +6 SD from final preferred). We measured energetic cost at these

particular step widths to give an expansive estimate of the new cost landscape, while measuring the energetic cost about both initial and final preferred widths. We calculated metabolic power using the Brockway equation [112] and determined metabolic power at specific step widths by averaging over the final three minutes of each five-minute walking period.

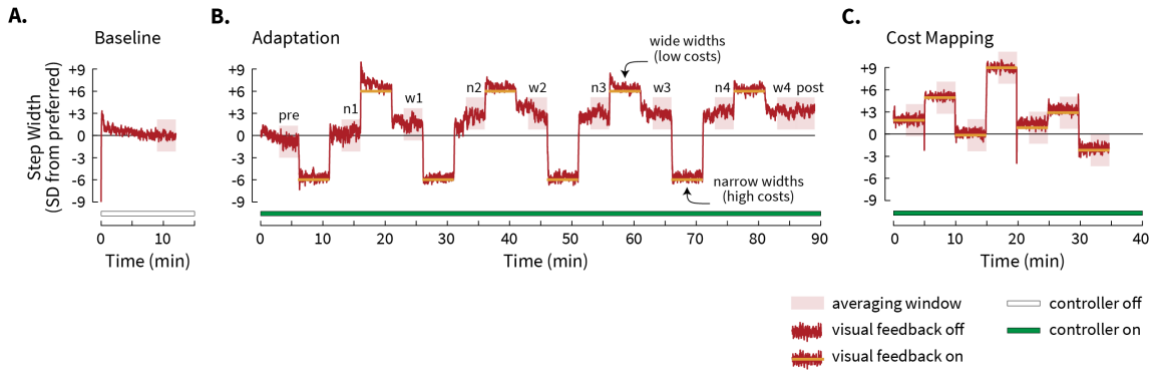

Figure 2.2: Experimental protocol
Each participant completed all three trials. We averaged measured step width across participants. (A) First, before we engaged the controller, each participant completed a baseline trial where we determined their preferred width and width variability. (B) Next, each participant completed the experience period where we measured whether they would adapt their preferred width towards the new energy optimal width spontaneously, and when given enforced experience with the new cost landscape through perturbations to different widths. During the periods denoted by the red horizontal lines, we used visual feedback to command either narrow widths (higher costs) or wider widths (lower costs). During the periods without the red horizontal lines, participants walked at self-selected widths. (C) Lastly, each participant completed a cost mapping trial where we measured their energetic cost at different widths in the new cost landscape. The regions of red shading illustrate the averaging window for steady-state step widths.

2.3.3. Analysis

We compared the new cost landscape to an original cost landscape to verify that we had shifted the energy optimal width wider than that initially preferred. For the original cost landscape, we used existing data collected in different participants (n=10) [80]. We performed all energetic cost analysis in dimensionless units. We made energetic cost, measured in W/kg, dimensionless using by multiplying by a normalization factor of $g^{3/2}L^{1/2}$ ($g = 9.81\ ms^{-2}$; $L = 0.53 \times H$, where H = height) [113]. We calculated each participant's energy optimum by averaging step width and energetic cost over each steady-state walking period during the cost mapping trial, fitting a second-order polynomial to these points, and then calculating the minimum of this polynomial. We calculated the original and new energy optimal widths as the average fitted minima across participants in the original and new cost

landscapes. We calculated the uncertainty in these minima as the standard deviation of the fitted minima across participants. We used a one-tailed paired Student's t-test to test if the new energy optimal widths were wider than initial preferred widths in each cost landscape.

We tested for adaptation towards the energy optimal width throughout the protocol. We used one-tailed paired Student's t-tests to determine if each participant's self-selected width during each release was wider than their initial preferred width. To determine the rate of adaptation, we averaged self-selected widths across participants during the release periods after optimization was initiated and then used a first-order exponential to determine the time constant. We used two-tailed paired Student's t-tests to test for differences in the final preferred widths compared to the new energy optimal widths, as well as for differences in the costs of these widths. We calculated the cost of each participant's final preferred width by commanding this width during the cost mapping trial, and the cost of each participant's energy optimal width as the minimum metabolic power of the fitted new cost landscape.

2.4. Results

Our control system successfully shifted the energy optimal step width wider than that initially preferred ($p = 1.3 \times 10^{-5}$; Figure 2.3). In the previously collected original cost landscape, the average energy optimal width (0.10 ± 0.05 m; mean±SD) was 1.1 ± 3.5 SD narrower than the initial preferred width (0.12 ± 0.01 m) [80]. In our new cost landscape, the average energy optimal width (0.24 ± 0.03 m) was 4.9 ± 1.1 SD wider than the initial preferred width (0.15 ± 0.04 m).

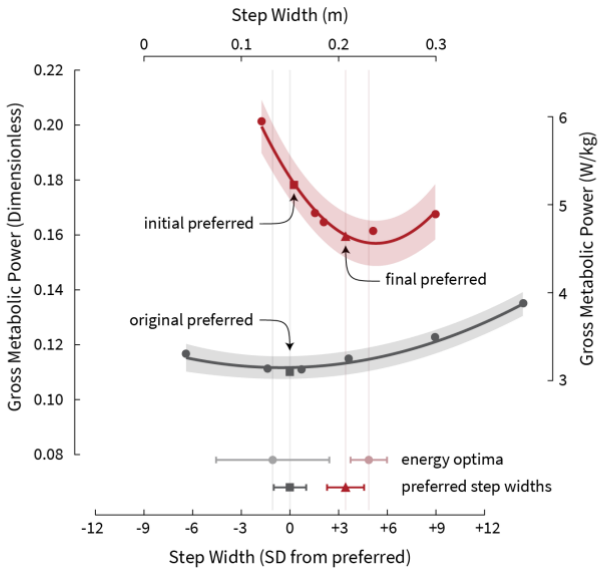

Figure 2.3: Original and new energetic cost landscapes
Original cost landscape (grey) from existing data in different participants [80] and measured new cost landscape (red). In each cost landscape, we averaged measured step width and energetic cost across participants. We fit second-order polynomial curves to both the original and new cost landscapes. The shading shows their 95% confidence intervals. We calculated the original preferred width (grey square) as the average width across participants in familiar walking conditions [80], the initial preferred width (red square) as the average width across participants when the controller is off, and the final preferred width (red triangle) as the average width across participants when the controller is on and following enforced experience. We calculated the variability in these preferred widths as the standard deviation of the original and final preferred widths across participants. We calculated the original (grey circle) and new (red circle) energy optimal widths as the average fitted minima across participants in the original and new cost landscapes. We calculated the variability in these energy optimal widths as the standard deviation of the fitted minima across participants in the original and new cost landscapes. Error bars represent 1 standard deviation.

We found that participants needed experience with a lower cost width in the new cost landscape to initiate optimization and adapt towards the new energy optimal width. They did not spontaneously adapt towards the new energy optimal width ($p = 0.9$; pre; Figure 2.4A) and returned to their initial preferred width in self-selected steps after being held at a higher cost width (narrower width) ($p = 0.2$; n1; Figure 2.4A). Participants only initiated optimization in self-selected steps after being held at a lower cost width (wider width), where they walked at an average width of 1.6 ± 0.6 SD from initial preferred ($p = 9.2 \times 10^{-5}$; w1; Figure 2.4A). After optimization was initiated, participants gradually converged on their final preferred width with an average time constant of 248 seconds (95% CI [247 249]), or about 468 steps (Figure 2.4B). And after converging on the final preferred width that was towards the energy optimal width, participants quickly returned to it when perturbed away (w1-post; Figure 2.4A).

Participants adapted their step width to a width that reduced energetic cost. On average, participants adapted their width 3.5 ± 0.8 SD wider than that initially preferred ($p = 3.1 \times 10^{-6}$; post; Figure 2.4A) and reduced energetic cost by 14.4% ± 6.1% relative to the cost of the initial preferred width in the new cost landscape ($p = 2.4 \times 10^{-4}$; Figure 2.3). This adaptation is not by random chance—a step width 3.3 standard deviations wider than preferred is likely to happen only once in every 1000 steps during normal walking. And although this final preferred width is significantly narrower than the new energy optimal width ($p = 2.2 \times 10^{-4}$), the costs of these widths are not significantly different ($p = 0.6$).

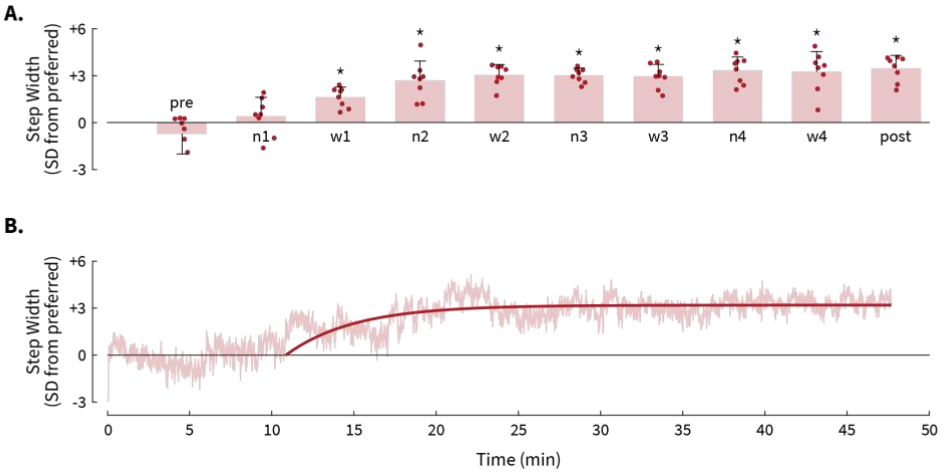

Figure 2.4: Timescale of adaptation
(A) Self-selected step width prior to enforced experience (pre), during narrow (n1-n4) and wide (w1-w4) releases and following enforced experience (post). Bar height is average across participants, error bars represent one standard deviation, and asterisks indicate statistically significant differences in width compared to initial preferred width. (B) Self-selected step widths averaged across participants. We determined the rate of adaptation after optimization was initiated (w1-post) by first removing steps during holds and only considering self-selected steps during releases, fitting an exponential model to these steps, and then calculating the time constant of this fitted exponential. The regions of red shading illustrate the averaging windows used to calculate the average step width during specific periods throughout the protocol.

2.5. Discussion

We tested the hypothesis that energy optimization is a major objective in the real-time control of step width by creating a new cost landscape and then measuring whether people adapt towards the new energy optimal width. Our previous understanding of preferred width was that it is energetically optimal in familiar walking conditions, yet we did not know the timescale by which this preference is established [80]. Here, we find that the nervous system establishes this preferred width by optimizing energetic cost in real-time.

Our experiment has several limitations. One concern is that, while walking at an incline or with a forward horizontal force, people may naturally prefer to walk with a wider width and thus may be adapting in response to our setup rather than to minimize energy. However, others have found that step width changes minimally within the range of inclines that participants experience in our study [114]. And in pilot experiments, we found no clear adjustments to step width in response to either constant incline or constant forward horizontal force. Another concern may be that the incline controller perturbs walking participants and that this causes participants to increase width. Indeed, some studies that perturb level walking with changes in terrain do find that people spontaneously widen their step widths [115]. However, we suspect that this does not explain our results for several reasons. First, this is not a universal finding as others find no systematic effect of changing terrain on step width [116,117]. Second, as we describe in the methods section, our incline control is not a large perturbation—the incline changed slowly, and no participants reported feeling destabilized. Finally, our participants did not spontaneously widen their step width in response to the incline control—they first required experience with lower cost step widths. Another concern is that participants may be adapting their width to achieve level walking rather than to optimize energy. We partially addressed this by making the minimum incline occur at a non-zero treadmill slope (~4°), however, the minimum incline is still the energy optimal incline. Thus, we cannot distinguish between a nervous system objective of walking at the shallowest incline from an objective of minimizing energy. Experiments like ours can never entirely rule out other possible explanations—it is always possible that the nervous system has some other objective that happens to produce a gait that is the same as the new energy optimal gait.

The nervous system's criteria for initiating step width optimization in this experiment, and the process used to converge towards energy optimal widths, is consistent with what we have previously observed for step frequency optimization using knee exoskeletons. The first similarity between studies is that perturbations towards lower cost gaits initiated the optimization process [42,73]. Experience with this lower cost gait may cue the nervous system to explore the new cost landscape by indicating that its preferred gait is now energetically suboptimal. The second similarity between studies is that the nervous system learned to predict the energetically optimal width and rapidly returned to it when perturbed away [42,73]. These similarities suggest that continuous energy optimization is a dominant

and general objective of nervous systems in young healthy walking, and that our observed behaviors are common characteristics of the nervous system's energy optimization.

There are, however, some differences in experiment findings. One difference is that no participants in our current experiment spontaneously optimized their step width, whereas a small subset of people spontaneously optimize step frequency [73]. One candidate explanation for this difference is that the general population does have spontaneous step width optimizers, but by chance, we did not have any in our random sample. A more mechanistic candidate explanation is that the nervous system is not primed to identify differences in energetic cost with step width because they are normally relatively small due to the shallowness of the cost landscape around the preferred width. The second difference is that participants converged on the energy optimum slower in our step width study compared to our step frequency study. In our step width study, people gradually converged on the energy optimal width with a time constant of 248 seconds (Figure 2.4B). Others have found similar rates of adaptation (i.e. hundreds of seconds) in converging on energy optimal movements in both split-belt walking and reaching paradigms [49,118]. However, in our step frequency study, people converged on the new energy optimal frequency with a time constant of 11 seconds once optimization was initiated [42]. One explanation may be that the nervous system takes longer to learn the new step width cost landscape because of the slow rate of change of our incline controller. This can result in an error between what we design the new cost landscape to be and what the nervous system senses. Simple reinforcement learning models also predict these different rates of adaptation, ranging from tens to hundreds of seconds, given a change to the learning rate [73]. Given this flexibility, we do not consider the differences in rates to be evidence against a shared adaptation process.

The nervous system likely determines step width by optimizing energy simultaneously with other objectives such as stability and maneuverability [6,89,107,119–122]. Our study suggests that, in the present walking conditions, the relative contributions of other objectives are small as people converged on a final preferred width that was near the new energy optimal width. However, it is possible that differences in the adaptation rates we observed between our step frequency and step width experiments reflects differences in the complexity of the optimization problem—the addition of other objectives in determining step width may slow the nervous system's refinement of energy. And although the preferred widths are near their energy optimal widths, they do not perfectly coincide, and the influence

of other objectives may explain these small differences. In everyday walking, the contribution from energy and non-energy objectives to the real-time control of gait will depend not only on biomechanics, but also on how heavily the nervous system weights their individual importance.

Chapter 3.

Limited energy optimization along multiple gait parameters in human walking

3.1. Abstract

Humans can learn new optimal control policies for complex movements. In walking, they can learn to adapt their control policy when the energy optimum is shifted along a particular gait parameter such as step width or step frequency. But it is unknown how the nervous system searches through and discovers these new optimal policies—it may recognize what needs to be learned and search for new optimal policies in this way, or it may search for new optimal policies in many ways irrespective of what needs to be learned. How the nervous system's optimization algorithms extend to multiple gait parameters may reveal how it represents its search space. Here, we tested if and how humans adapt their control policy when the energy optimum is shifted along both step width and step frequency. We previously demonstrated that they can do so when the energy optimum is shifted along either gait parameter. Now, we shift the energy optimum using a device that applies energetic penalties of walking incline as a function of both of these gait parameters. We find that participants spontaneously increase variability and initiate adaptation along step frequency but not step width. They adapt step frequency for cost savings of about 7%. They do not adapt step width despite potential for larger cost savings and experiencing conditions previously shown to initiate its adaptation. That exploratory variability and adaptive changes are limited to one gait parameter suggests that the nervous system recognizes to some extent what needs to be learned and searches for new optimal policies in this way.

3.2. Introduction

Humans can learn new optimal movement coordination strategies. This is an impressive feat given that the nervous system has tens of thousands of motor units that it can coordinate over millisecond timescales—it faces a high dimensional learning problem. In many situations, it need not search within this entire space of coordination strategies for new optimal strategies. And it may not be realistic to do so within a reasonable amount of time. One way to mitigate this high dimensional problem is to represent only a low dimensional search space of control policies. The term *control policy* refers to a mapping between states and actions taken in those states, and the term *search space* refers to a space within which the nervous system searches through and evaluates new control policies. Here we seek to gain insight into how the nervous system represents its control policy search space.

Humans can learn to adapt their control policy for walking to optimize metabolic energy cost. That is, they prefer to walk with the combination of step widths, step frequencies, and speeds that optimizes energy cost [7,10,80,106]. And they can learn to adapt this control policy when the energy optimum is shifted to a new step width or new step frequency [42,84,86]. Energy optimization in walking can serve as a model system for learning because metabolic energy cost is one of the nervous system's major objectives that we can both directly measure and manipulate. It also has several hallmarks that can provide insight into the nervous system's mechanisms for learning [73]. First, high natural gait variability can spontaneously initiate adaptation. Second, enforced experience with lower energy cost gaits can initiate adaptation. And third, adaptation leads to new predictions of energy optimal gaits. But there are some conditions where people do not initiate adaptation—increasing the saliency of energy cost savings through the cost gradient does not initiate adaptation [85], nor does increasing the natural gait variability through overground walking [83].

The size of the search space can affect learning. As the number of states and actions increases, the number of candidate control policies increases exponentially. This expansion of the space of control policies can affect learning as searching through and evaluating many policies takes time—a challenge known as the 'curse of dimensionality' [123]. The way in which the nervous system overcomes this challenge to learn new energy optimal policies is presently unclear. One possibility is that it identifies a lower dimensional search space. In past studies, we found that enforced experience with lower energy cost gaits can initiate adaptation [73,85,86]. This

experience may inform a lower dimensional search space by indicating to the nervous system that the energy optimal control policy has shifted, and along which dimensions these shifts have occurred.

Humans appear to identify the relevant dimensions by learning the structure of the task. One example of structural learning is in a reaching task where the visual input is rotated to some random angle every few reaches, and participants learn the structure (rotation) rather than the parameter (angle) [66]. When later presented with a new context, they explore and adapt along this structure. In a different reaching task, participants learn to adapt their movements to compensate force fields of varying temporal structure [1]. When later presented with a new context, they rely on previous experience to make assumptions about its structure. In both of these tasks, participants learn faster when presented with the new context. Structural learning is one instance where the nervous system may use a lower dimensional search space.

How does the nervous system represent its search space? To generate competing hypotheses and design experiments that distinguish between possibilities, we began by modeling energy optimization as reinforcement learning. Our previous work shows that simple reinforcement learning models can predict hallmarks of energy optimization [73]. We built on this work by simulating two possible cases: a low dimensional search space and a high dimensional search space (See Methods). In a low dimensional search space, the learner recognizes what needs to be learned and attempts to adapt its control policy along these dimensions. Alternatively, in a high dimensional search space, the learner attempts to adapt its control policy along many dimensions irrespective of what needs to be learned. How it optimizes its control policy for tasks of increasing complexity (where the energy optimum is shifted along one and two dimensions) can reveal the nature of its search space. When a task increases in complexity, there are at least two possibilities that would suggest a low dimensional search space. The first possibility is that the learner optimizes more slowly (red vs. green case 1; Figure 3.1A). The second possibility is that it does not optimize completely—its low dimensional search space is insufficient to converge on the new optimum (red vs. green case 2; Figure 3.1A). And there is at least one possibility that suggests a high dimensional search space—it optimizes tasks of increasing complexity with similar time constants of adaptation (red vs. green; Figure 3.1B). We used energy optimization in walking to test these model predictions.

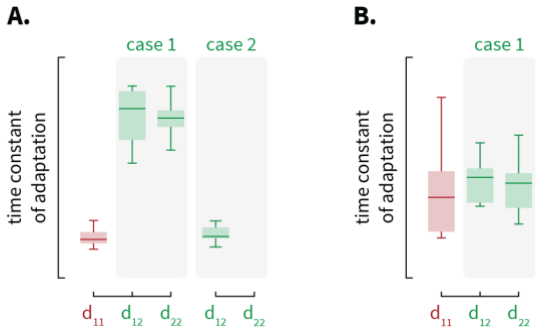

Figure 3.1: How the nervous system can represent its search space
(A) Low dimensional search space and (B) high dimensional search space with low task complexity (red) and high task complexity (green). We modeled energy optimization as reinforcement learning to ask how the search space might influence the time constant of adaptation. For each search space, we compared the time constant of adaptation for low task complexity, i.e., when the energy optimum is shifted along only one gait parameter (d_{11}), to the time constant of adaptation for high task complexity, i.e., when the energy optimum is shifted along two gait parameters (d_{12} and d_{22}).

The aim of the present study was to test how the nervous system represents its search space. To accomplish this, we determined (1) if people learn to adapt their control policy when the energy optimum is shifted along two gait parameters, and (2) how this compares to when the energy optimum is shifted along one gait parameter. In our past studies, we shifted the energy optimal step width or the energy optimal step frequency and observed energy optimization along each gait parameter [42,84,86]. In the present study, we shifted the energy optimum along both gait parameters using a device that applies energetic penalties in the form of walking incline and does so as a function of people's real-time measured step width and step frequency. To gain insight into how the nervous system represents its search space, we used reinforcement learning models to predict how adaptation compares with a low dimensional and high dimensional search space, and then compared model predictions against human experiments. We hypothesized that the nervous system identifies a low dimensional search space to quickly learn new optimal control policies in new contexts.

3.3. Methods

3.3.1. Experimental design

We used a simple mechatronic system to study energy optimization in human walking. This system is described in more detail in our previous study [86]. In brief, we used this system to shift the energy optimal step width and step frequency by commanding an energetic penalty of walking incline based on real-time measurements of these gait parameters (Figure 3.2A).

We measured ground reaction forces and moments while participants walked on an instrumented split-belt treadmill at a constant speed of 1.25 m/s (FIT, Bertec Corporation, Columbus, OH, USA). We used these ground reaction forces and moments to calculate the center of pressures by dividing the left and right moments by their vertical forces. We identified foot contact events as the rapid fore-aft translation in center of pressure during double support, and then quantified step frequency as the inverse of time between consecutive foot contact events. We quantified step width as the difference in lateral center of pressures between consecutive foot contact events. The closed-loop control system calculated the desired energetic penalty based on the step width and step frequency from the previous step, and then commanded the appropriate walking incline. We implemented this system at 200 Hz using MATLAB Simulink (Simulink Real-Time Workshop, MathWorks Inc., Natick, MA, USA). This system was inherently limited by the treadmill's ability to change incline, which was about 0.5 degrees per second. This meant that participants experienced the appropriate energetic penalty with small step-to-step changes in step width or step frequency but did not immediately experience the appropriate energetic penalty with large step-to-step changes. This apparent limitation may have prevented participants from feeling perturbed or identifying the behaviour of the control system. In a brief survey collected after the experimental protocol, we asked participants whether they realized the behaviour of the control system—that the way they walked had control over the treadmill incline.

We created a cost landscape that shifts the energy optimal step width and step frequency. We use the term *cost landscape* to refer to a relationship between gait parameters and energetic cost, and *control function* to refer to a relationship between gait parameters and energetic penalties of walking incline:

$$Incline(sw, sf) = \begin{cases} -0.92(sw) + 5.3 & sw \leq +3 \\ -0.06(sw) + 2.4 & sw > +3 \end{cases}' \begin{cases} 0.91(sf) + 5.3 & sf \geq -3 \\ 0.12(sf) + 2.6 & sf < -3 \end{cases} \quad (3.1)$$

where sw is step width and sf is step frequency. We simulated the new cost landscape by adding the original cost landscape to the energetic costs associated with the control function. The original cost landscape is the relationship between gait parameters and energetic cost when the controller is turned off, and the new cost landscape is this relationship when the controller is turned on. We designed the new cost landscape to shift the energy optimal step width to three standard deviations wider than the initial preferred width and energy optimal step frequency to three standard deviations lower than the initial preferred frequency. We

parameterized the new cost landscape by step width and step frequency in units of standard deviations (SD) from the initial preferred in the original cost landscape. We calculated these units by subtracting each participant's initial preferred value from each step's measured value and then dividing by the standard deviation about their initial preferred value. In past studies, we demonstrated that people can learn to adapt their control policy to converge on the new energy optimum when it is shifted in this same way along either step width or step frequency [42,84,86]. In the present study, we expanded on this by testing how people learn to adapt their control policy when the energy optimum is shifted along both gait parameters.

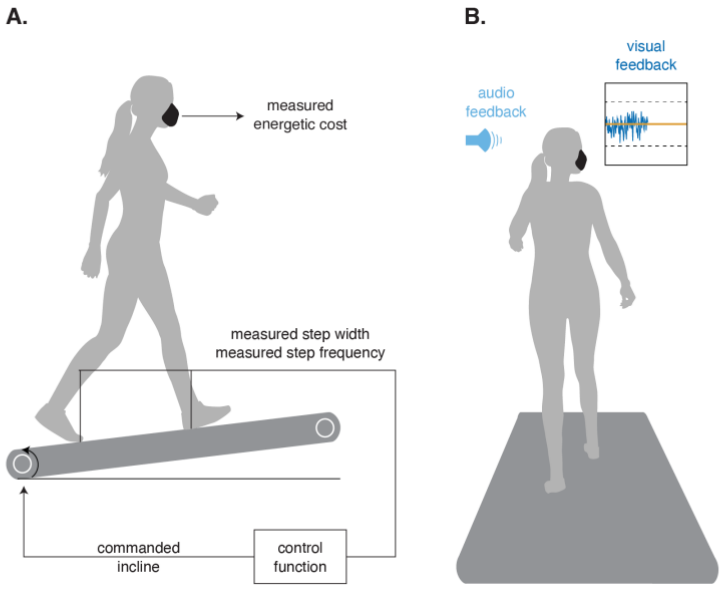

Figure 3.2: Experimental setup and design
(A) We used a simple mechatronic device that commanded the walking incline based on the desired energetic penalty for the real-time measured step width and step frequency. (B) During certain periods of the protocol, we used visual feedback to enforce specific step widths and audio feedback to enforce specific step frequencies.

3.3.2. Experimental protocol

We collected data from 12 participants (sex: 7 females, 5 males; age: 25 ± 3 years; body mass: 67.6 ± 10.9 kg; mean ± SD). All participants were healthy and had no known cardiopulmonary or gait abnormalities. The Simon Fraser University Research Ethics board approved the study protocol, and all participants gave their written, informed consent before participating in the study.

We used a baseline period to determine each participant's initial preferred step width and step width variability, as well as their initial preferred step frequency and step frequency variability. We measured these baseline values while participants walked on a treadmill at a constant speed of 1.25 m/s and at a constant incline of zero. In this period, they experienced the original cost landscape with the controller turned off. We calculated each participant's initial preferred step width and initial preferred step frequency as the average of each signal during the final three minutes of a 20-minute baseline period (Figure 3.3A). We calculated baseline variability as the standard deviation of each signal during the same time window.

We used an experience period to determine whether participants adapt step width and step frequency in the new cost landscape. We turned the controller on at the beginning of the 90-minute experience period, and it remained on for all but the final five minutes of this period (Figure 3.3B). Prior to this period, we informed participants that they would need to "match step width to visual feedback and step frequency to audio feedback when the feedback was turned on". We also informed them that they "did not need to continue to walk in the same way when the feedback was turned off". We did not give them any further instructions about how to walk and did not inform them of the behaviour of the control system. Participants self-selected steps for the first five minutes of this period, and we used this time window to test for spontaneous increases in variability as well as changes in magnitude of gait parameters. Participants then alternated between five-minute enforced experience periods with feedback and five-minute releases without feedback. During the first enforced experience period, we commanded a narrower step width (-6 SD) and higher step frequency (+6 SD) than initially preferred, resulting in a higher energetic cost. During the second enforced experience period, we commanded a wider step width (+6 SD) and lower step frequency (-6 SD) than initially preferred, resulting in a lower energetic cost. We alternated between enforcing high and low energetic costs for a total of eight enforced experience periods. During enforced experience periods, participants matched their real-time step width signal to be within small bounds of a steady-state step width signal displayed on a screen, and their real-time step frequency to a metronome playing a steady-state tempo (Figure 3.2B). Participants successfully matched commanded step widths with an error of 0.012 ± 0.0088 m and commanded step frequencies with an error of 1.5 ± 1.3 steps per minute. After each enforced experience period, they self-selected steps during a five-minute release where we tested for changes in magnitude of gait parameters.

We used cost mapping to determine how our controller shifted the energetic cost optimum. Because participants walked for nearly two hours to complete the baseline period and the experience period, we had participants complete the cost mapping periods on a different day. Only a subset of six participants were able to return for this day. In one cost mapping period, we measured the original cost landscape with the controller turned off (Figure 3.3C). In the second cost mapping period, we measured the new cost landscape with the controller turned on (Figure 3.3D). We used respiratory gas analysis to measure rates of oxygen consumption and carbon dioxide production (Vmax Encore Metabolic Cart, ViaSys, Conshohocken, PA, US), and then calculated metabolic power using the Brockway equation [112]. We measured each participant's resting metabolic power prior to completing these cost mapping periods during a six-minute quiet standing period. We then measured each participant's metabolic power at 17 combinations of step width and step frequency that nearly formed a 5x5 grid about a step width of 0 SD and a step frequency of 0 SD in the original cost landscape, and a step width of +3 SD and a step frequency of -3 SD in the new cost landscape— the expected locations of the optima in the original and new cost landscapes. We designed these combinations to estimate the energy optimum while limiting each participant's total walking time, where they walked for two minutes at each combination of step width and step frequency. We commanded these combinations using visual feedback for step width and audio feedback for step frequency, and randomly presented these combinations to each participant to mitigate the effects of order.

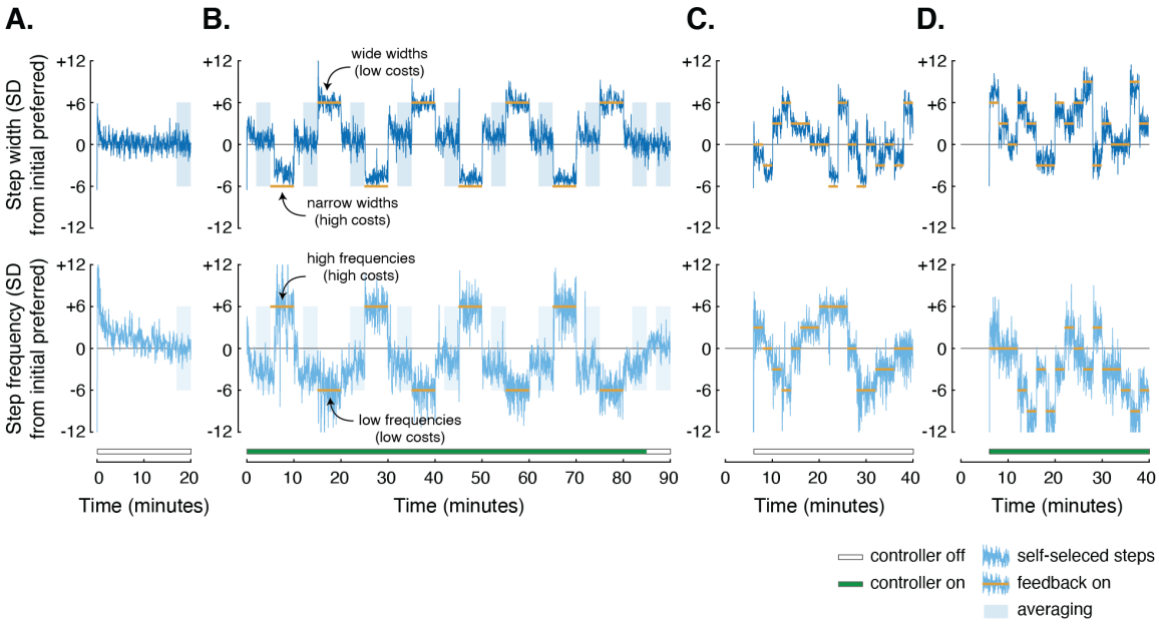

Figure 3.3: Experimental protocol

Each participant completed the baseline period and the experience period. Some of these participants also completed the cost mapping periods with the controller turned off and controller turned on. Here we present a representative participant's time series data for step width (top row) and step frequency (bottom row). The shaded regions illustrate time windows during which we analyzed step width and step frequency. (A) In the baseline period, we determined each participant's preferred step width and step width variability, as well as their preferred step frequency and step frequency variability, with the controller turned off. (B) In the experience period, we turned the controller on and determined whether participants adapted toward the new energy optimum spontaneously and then after enforced experience with different combinations of step width and step frequency in the new cost landscape. Participants experienced either high energetic costs (narrow step widths and high step frequencies) or low energetic costs (wide step widths and low step frequencies) during periods with the yellow horizontal lines, and self-selected steps during periods without the yellow horizontal lines. (C) In cost mapping, we measured each participant's energetic cost while they walked at combinations of step widths and step frequencies. We turned the controller off to measure the original cost landscape and (D) turned the controller on to measure the new cost landscape.

3.3.3. Analysis

To determine whether our controller effectively shifted the energy optimum, we estimated the energetic cost of combinations of step width and step frequency for each cost landscape. For each these combinations, we fit an exponential model to two minutes of breath-by-breath metabolic power measurements and calculated the steady-state value as the asymptote of the model [124]. We calculated each participant's resting metabolic power by averaging their metabolic power across the final three minutes of the six-minute quiet standing period. We calculated net metabolic power by subtracting resting metabolic power from the steady-state value and dividing by each participant's mass. We then used these combinations to fit a two-dimensional second-order polynomial of the form:

$$Y(sw, sf) = a(sw)^2 + b(sf)^2 + c(sw)(sf) + d(sw) + e(sf) + f \quad (3.2)$$

where sw is step width, sf is step frequency, and model output $Y(sw, sf)$ is energetic cost. For each participant, we estimated coefficients $\{a, b, c, e, d, f\}$ using least squares regression. We then estimated their energy optimal step width and energy optimal step frequency as the minimum of the two-dimensional polynomial. For each cost landscape, we used these values to calculate the group mean and standard deviation of the energy optimal step width and energy optimal step frequency. We used a one-tailed paired Student's t-test to determine whether the energy optimum was shifted in the right direction, i.e., the energy optimal step width shifted wider, and the energy optimal step frequency shifted lower, in the new cost landscape compared to the original cost landscape. We wrote custom MATLAB scripts to

process the data, perform the statistical comparisons, and generate the figures included in this manuscript.

We analyzed changes in variability as well as changes in magnitude of each gait parameter. We quantified the variability of each gait parameter by high-pass filtering its signal to include changes over less than 30 steps, and then calculating the standard deviation of this filtered signal. This method filtered out the relatively slow signal changes that we associate with adaptive changes in magnitude, but not the relatively rapid signal changes that occur from step to step and over several steps that we associate with exploratory variability [125]. We calculated variability during the last three minutes of the 20-minute baseline period in the original cost landscape, and during the last three minutes of the first five minutes of the experience period in the new cost landscape. We used a one-tailed paired Student's t-test to determine whether each participant's variability in the new cost landscape was higher than their baseline variability in the original cost landscape. The reason for this comparison is that baseline variability can provide an estimate of variability that reflects noise rather than exploration—people likely have an established control policy for walking in the original cost landscape and need not explore. Any increase in variability relative to baseline variability might therefore reflect exploration of the new cost landscape rather than noise. We next quantified the magnitude of each gait parameter by averaging its signal during the last three minutes of each five-minute release throughout the experience period. For each release, we used a one-tailed paired Student's t-test to determine whether participants adapted their gait in the direction that we shifted the energy optimum. We tested whether step width was wider than their initial preferred width and step frequency was lower than their initial preferred frequency.

We compared the time constants of adaptation in new cost landscapes where we shifted the energy optimum along one or two gait parameters. In our past study (n = 8), we shifted the energy optimum along step width [86]. In a pilot study (n = 5), we shifted the energy optimum along step frequency using the same methods as in our past studies [42,84]. In our present study, we shifted the energy optimum along both step width and step frequency. We analyzed the time constant of adaptation for gait parameters along which we observed adaptation. For each of these gait parameters, we modeled the timing of adaptation as an exponential change from an initial value to a final steady-state value:

$$Y(t) = a(e^{-t/\tau}) + b \tag{3.3}$$

where t is the amount of time during which participants self-selected steps and model output $Y(t)$ is either step width or step frequency. We used a mixed-effects model with a single time constant (τ) shared between participants and individual participant offsets (b). We then used nonlinear optimization to estimate these coefficients [126]. We calculated the median and the interquartile range, which we define as the difference between the 75th and 25th percentiles, for each time constant of adaptation. We tested for differences in time constants of adaptation between studies where we shifted the energy optimum along one gait parameter and studies where we shifted the energy optimum along two gait parameters. We used a Kruskal–Wallis one-way ANOVA to test for these differences because these time constants did not follow a normal distribution (Anderson–Darling test: $p = 5.0 \times 10^{-4}$).

We used simple reinforcement learning models to generate competing hypotheses about how the nervous represents its control policy search space. This builds off previous work showing that reinforcement learning can predict experimental behaviour in energy optimization [73]. In these models (Figure 1.2), the learner iteratively updates a value function (Q), which we define to be the predicted relationship between gait and energetic cost, or the cost landscape. At each step (i), the learner chooses an action (a) from all possible actions (A) in agreement with its control policy (π). That is, it chooses the action that optimizes its value function—the energy optimal action. We define the set of all possible actions to range between -15 and +15 standard deviations from initial preferred gait. We assume that the nervous system cannot execute this action perfectly and therefore add execution noise to each action. We determine the execution noise for each action by sampling from a Gaussian distribution with zero mean and standard deviation of 1.0, similar to that which we experimentally observed. The learner then measures a reward (r), or in this case energetic cost. We assume that the nervous system cannot measure energetic cost perfectly and add measurement noise to each reward, again sampling from a Gaussian distribution with zero mean and standard deviation of 0.02, which we determined in our previous work [73]. And because these measurements are imperfect, the learner updates its value function with only a fraction ($\alpha = 0.5$) of this reward:

$$Q(a_i) = Q(a_i) + \alpha\big(r - Q(a_i)\big) \tag{3.4}$$

We initialized the value function to be the original cost landscape with the controller turned off and the optimum at 0 standard deviations from initial preferred gait:

$$Q_{off}(A_d) = 10\left(\frac{A_d}{100}\right)^2 + 1 \tag{3.5}$$

where A is the set of all possible combinations of actions across all dimensions, and A_d is the subset of actions along each dimension (d). We simulated 1200 steps, or roughly 10 minutes of walking, in the new cost landscape with the controller turned on and the optimum at +3 standard deviations from initial preferred gait:

$$Q_{on}(A_d) = 10\left(\frac{A_d}{100} - \frac{3}{100}\right)^2 + 1 \tag{3.6}$$

We simulated two possible cases: low dimensional search space and high dimensional search space. The dimensionality of the search space is defined by the size of the value function. In a low dimensional search space, the learner recognizes what needs to be learned and then only attempts to adapt the control policy along these dimensions—the value function is Q_{Ax1} when the energy optimum is shifted along one dimension, and Q_{Ax2} when shifted along two. In a high dimensional search space, the learner attempts to adapt the control policy along many dimensions irrespective of what needs to be learned—the value function is always Q_{Ax5}. We selected only 5 dimensions because of the computational resources needed to search among all possible combinations of these dimensions. We tested how increasing the task complexity (shifting the energy optimum along one and two dimensions) influences the search space and time constant of adaptation in each case. We quantified the search space as the actions that the learner selects. We quantified the time constant of adaptation using Equation 3.3 to fit an exponential to actions and calculating the time constant of this fitted exponential. How the nervous system optimizes its control policy for tasks of increasing complexity, and how this compares to model predictions, may reveal the nature of its search space.

3.4. Results

Our device successfully shifted the energy optimum along two gait parameters. We calculated each participant's initial preferred step width (0.16 ± 0.043 m) and step width variability (0.013 ± 0.0037 m), as well as initial preferred step frequency (109.6 ± 5.6 steps per minute) and step frequency variability (1.1 ± 0.28 steps per minute), during the baseline period. We used each participant's baseline values to shift the energy optimal step width to wider widths and energy optimal step frequency to lower frequencies. In the original cost landscape, the energy optimal step width was +1.8 SD from initial preferred (IQR [+0.42, +1.9]; Figure 3.4A

and C) and the energy optimal step frequency was -0.98 SD from initial preferred (IQR [-1.6, +3.6]; Figure 3.4A and C). In the new cost landscape, the energy optimal step width was +7.4 SD from initial preferred (IQR [+6.1, +7.5]; Figure 3.4B and D) and the energy optimal step frequency was -9.2 SD from initial preferred (IQR [-10.2, -8.4]; Figure 3.4B and D).

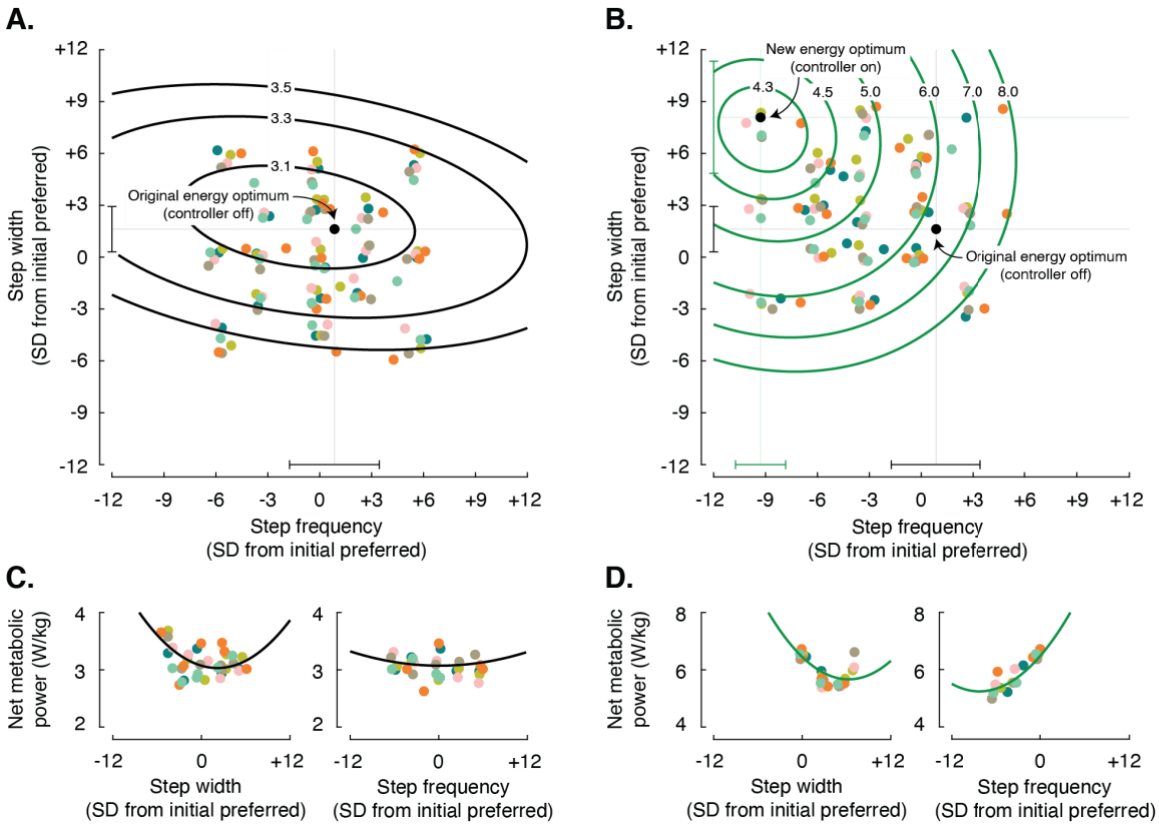

Figure 3.4: Original and new energetic cost landscapes
Energetic cost landscapes measured from six participants walking at a range of combinations of step width and step frequency in (A) the original cost landscape and (B) the new cost landscape. Distinct colours represent individual participants and solid lines represent the two-dimensional second-order polynomial model with mixed-effects. The contours describe the model fit to energetic cost data in units of W/kg, where each contour indicates a higher cost than the one inside. We determined the energy optimum (black dot) and uncertainty in this optimum (error bar) by calculating the mean and standard deviation across participants. Slices of (C) the original cost landscape and (D) the new cost landscape. On the left, we show participants' energetic cost data (colour dots) and polynomial model (solid line) at a step frequency of zero and a range of step widths. On the right, we show participants' energetic cost data at a step width of zero and a range of step frequencies. To better illustrate the relationship with energetic cost as determined by the mixed-effects model, we subtracted each participant's random-effects offset term from their energetic cost data.

Participants spontaneously increased variability and initiated adaptation in step frequency but not step width. We analyzed spontaneous changes in step width and step frequency prior to enforced experience in the new cost landscape—during the last three minutes of the first

five minutes of the experience period. We found that participants did not spontaneously increase step width variability compared to baseline variability (-0.76 ± 12.8%; one-tailed paired t-test: p = 0.58; Figure 3.5A), whereas they slightly increased step frequency variability compared to baseline variability (12.6 ± 25.9%; one-tailed paired t-test: p = 0.061; Figure 3.5A). We found that participants spontaneously adapted step frequency lower than initial preferred step frequency (-2.4 ± 1.7 SD from initial preferred; one-tailed paired t-test: p = 2.1 x 10^{-4}; Figure 3.5B), whereas they did not adapt step width wider than initial preferred step width (+0.18 ± 0.77 SD from initial preferred; one-tailed paired t-test: p = 0.22; Figure 3.5B).

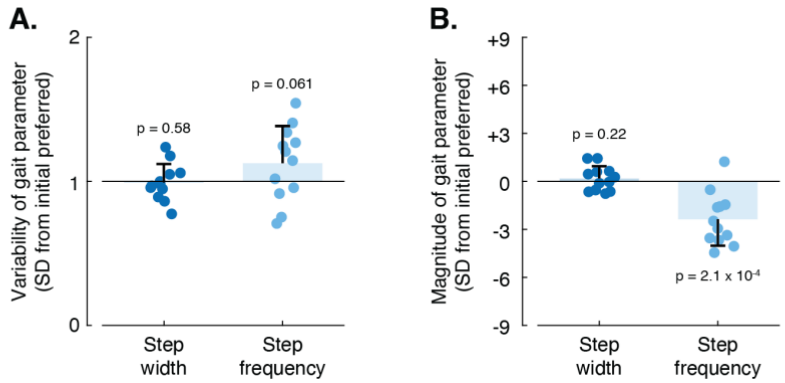

Figure 3.5: Behaviour upon first exposure to the new energetic cost landscape
(A) Variability and (B) magnitude of step width (dark blue) and step frequency (light blue) during the last three minutes of the first five minutes of walking in the new cost landscape. We compared changes in variability to baseline variability, i.e., value of 1, and changes in magnitude to initial preferred, i.e., value of 0. Each circle represents each participant's variability or magnitude, bar height is average across participants, and error bars represent one standard deviation.

Participants learned to adopt a new policy that reduced energy cost at step frequencies lower than initial preferred and step widths similar to initial preferred. During enforced experience periods, participants matched step widths and step frequencies that resulted in either lower or higher energetic costs than initial preferred. When self-selecting steps after each enforced experience period, we found that participants adopted a control policy similar to that which they spontaneously learned prior to enforced experience. They did not adapt step width after enforced experience with high costs (+0.16 ± 1.6 SD from initial preferred; one-tailed paired t-test: p = 0.36; Figure 3.6A release 1) or low costs (+0.92 ± 2.5 SD from initial preferred; one-tailed paired t-test: p = 0.11; Figure 3.6A release 2). However, they did adapt step frequency after enforced experience with high costs (-2.7 ± 2.3 SD from initial preferred; one-tailed paired t-test: p = 0.0010; Figure 3.6B release 1) and low costs (-3.0 ± 2.2 SD from initial preferred; one-tailed paired t-test: p = 3.5 x 10^{-4}; Figure 3.6B release 2). After 85 minutes of

walking in the new cost landscape, participants preferred a step width of -0.16 ± 1.7 SD from initial preferred width (one-tailed paired t-test: p = 0.62; Figure 3.6A post-experience) and a step frequency of -2.7 ± 2.6 SD from initial preferred frequency (one-tailed paired t-test: p = 0.0024; Figure 3.6B post-experience). This combination of step width and step frequency reduced energetic cost by 6.8 ± 10.5% compared to the energetic cost of the initial preferred in the new cost landscape (one-tailed paired t-test: p = 0.023). After we turned the controller off and participants again walked in the original cost landscape, they adopted a narrower step width than their initial preferred width (-1.0 ± 1.1 SD from initial preferred; two-tailed paired t-test: p = 0.013; Figure 3.6A re-adaptation) and converged on their initial preferred step frequency (-0.45 ± 1.5 SD from initial preferred; two-tailed paired t-test: p = 0.33; Figure 3.6B re-adaptation).

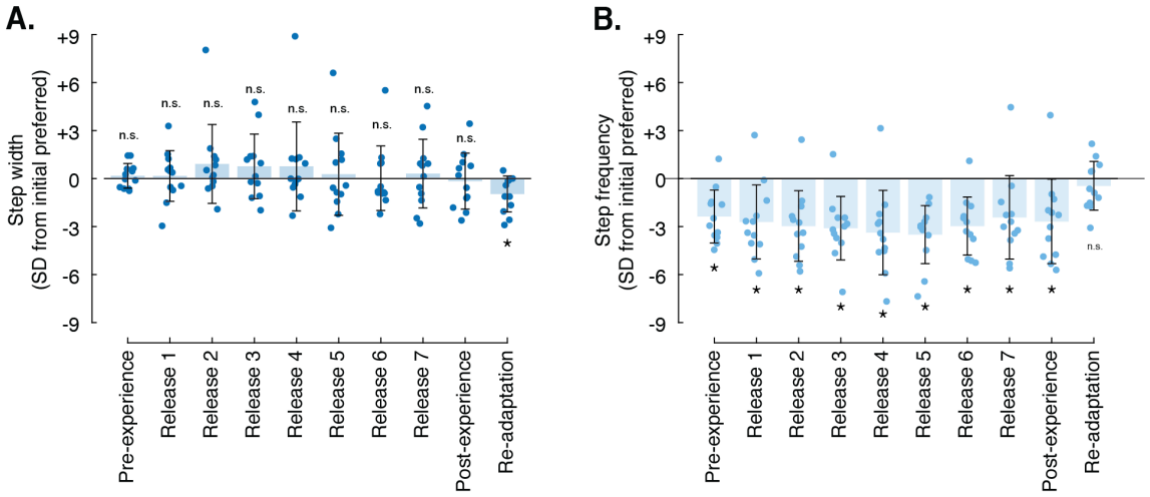

Figure 3.6: Behaviour with experience in the new energetic cost landscape
(A) Step width and (B) step frequency when participants self-selected steps throughout the experience period. They self-selected steps prior to enforced experience (pre-experience), after experience with high energetic costs (release 1, 3, 5, 7), after enforced experience with low energetic costs (release 2, 4, 6, post-experience), and after the controller was turned off (re-adaptation). The circles represent participant averages, the bar height represents average across participants, the error bars represent one standard deviation, and the asterisks indicate statistically significant differences from the initial preferred using the notation * for p<0.05 and n.s. for not significant.

Adaptation progressed over similar timescales when the energy optimum was shifted along one or two gait parameters. We determined the timescale of adaptation by removing steps during enforced experience periods and considering only self-selected steps. We then fit an exponential model to these steps and calculated the time constant of this fitted exponential. We compared the timescale of adaptation in the present study where we shifted the energy optimum along two gait parameters (2d) to the timescales of adaptation in past studies where

we shifted the energy optimum along one gait parameter (1d). In the present study, we found that participants did not learn to adapt step width (Figure 3.7A). They did, however, learn to adapt step frequency with a time constant of 4.1 minutes (IQR [2.3, 8.5]; Figure 3.7B). And this timescale of adaptation was similar to that which we have previously observed when we shifted the energy optimum along one gait parameter (one-way ANOVA; step width 1d vs. step frequency 2d: p = 0.96; step frequency 1d vs. step frequency 2d: p = 0.82; step width 1d vs. step frequency 1d: p = 0.97; Figure 3.7C).

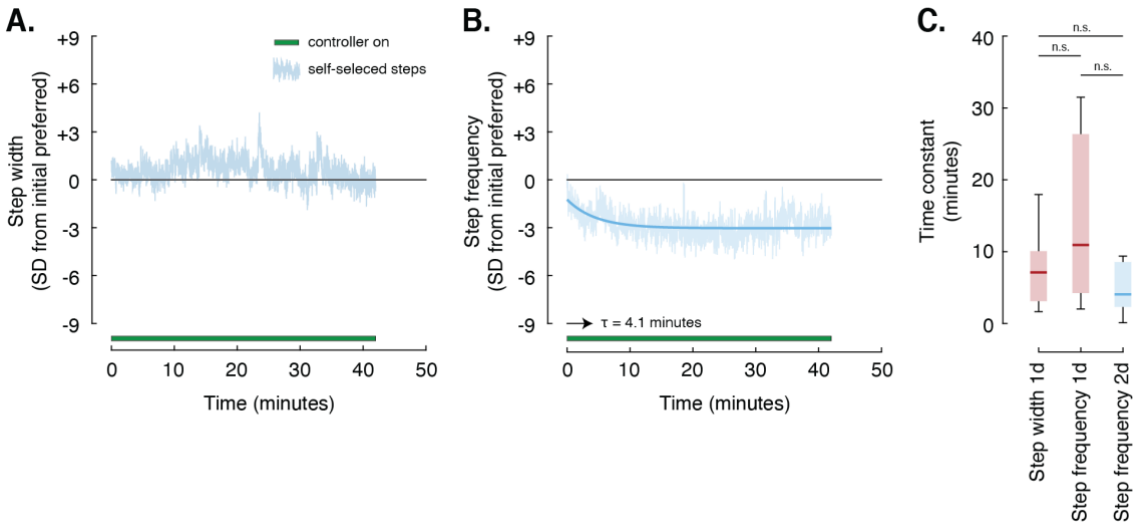

Figure 3.7: Timescale of adaptation
Self-selected (A) step widths and (B) step frequencies averaged across participants. The elapsed time consists of windows throughout the experience period where the controller was turned on and where participants self-selected steps. We normalized the magnitude of each gait parameter to participants' initial preferred (black horizontal line). For step width, we did not calculate its time constant because we did not observe changes in its magnitude. For step frequency, we calculated the time constant of the fitted exponential model (blue line). (C) The time constants of adaptation for step width and step frequency when the energy optimum was shifted along one parameter (1d) and two parameters (2d). We determined time constants and uncertainty in these time constants by calculating the median and interquartile range across participants. The central mark indicates median, bottom edge of the box indicates lower quartile (25th percentile), and top edge of the box indicates upper quartile (75th percentile). Error bars extend to the most extreme data points not considered outliers, which we define as more than 1.5x the interquartile range away from the edges of the box. We use the notation n.s. to indicate not statistically significant differences between time constants.

3.5. Discussion

The nervous system appears to be limited in its ability to optimize energy along multiple gait parameters. We shifted the energy optimum along two gait parameters—step width and step frequency—and created the conditions previously shown to initiate adaptation along each

parameter. That is, we created a new cost landscape with an energy optimum shifted to wider step widths and lower step frequencies and tested for spontaneous initiation of adaptation as well as initiation of adaptation after enforced experience with combinations of step width and step frequency that resulted in high or low energetic costs. We found that participants did not learn to adapt step width despite experiencing the conditions previously shown to initiate its adaptation. However, they did learn to adapt step frequency—they spontaneously increased variability and initiated adaptation along this gait parameter, where adaptation resulted in a new control policy that reduced energetic cost. Participants gradually adapted one of two gait parameters over several minutes, similar to that which we observed when the energy optimum was shifted along only one gait parameter. That changes were limited to one gait parameter suggests that the nervous system's search space is low dimensional.

Our study has three main limitations. One limitation is that we cannot distinguish between adaptation that is due to minimizing the objective of energetic cost and adaptation that is due to minimizing the objective of walking incline. To partly address this limitation, we designed the new energy optimum to be at a non-zero walking incline so that participants did not adapt to reach zero walking incline. We also designed the incline controller so that participants did not become aware of the relationship between gait and incline. The rate of change of incline was small—about 0.5 degrees per second. And the range of incline given participants' step-to-step variability was also small—about one degree. Participants did, however, experience large walking inclines of about 10 degrees and small walking inclines of about 2 degrees during enforced experience periods, but survey responses suggest that they were unaware of the relationship between changes in gait and incline. This is in line with a previous study that used a dual-task to demonstrate that energy optimization can be a primarily implicit process that requires minimal attention and strategy [127]. A second limitation is that preferred step width and preferred step frequency may change with changes in walking incline—adaptation may be due to our experimental setup rather than due to the nervous system optimizing energetic cost. However, previous studies have found that changes in step width and step frequency are minimal within the range of walking inclines that participants experience in our study [114,128]. A third limitation is that participants may not be able to adapt step width while adapting step frequency due to biomechanical coupling or constraints. That they can walk with the combinations of step width and step frequency during enforced experience periods suggests that they can walk with the step width and step frequency at the new energy optimum. However, because participants use visual feedback and audio feedback to achieve

the enforced combinations, we cannot completely rule out the possibility that they cannot freely adapt step width and step frequency at the same time.

The nervous system appears to define a control policy search space that is low dimensional. Depending on the nature of its algorithms, searching within a low dimensional space could take seconds, minutes, or even hours when the optimum is shifted along one gait parameter. This could also be the case when searching within a high dimensional space. To gain insight into how the nervous system defines its control policy search space, we determined how its energy optimization algorithms extend to multiple gait parameters. If the nervous system were to search within a high dimensional space, its optimization would likely not be affected by whether the optimum is shifted along one or two gait parameters. On the other hand, if the nervous system were to search within a low dimensional space, it might (1) optimize two gait parameters more slowly than one gait parameter, or (2) optimize only a subset of these two gait parameters (Figure 3.1). Our finding that people adapt one of two gait parameters is in line with the hypothesis that the nervous system searches within a low dimensional space. This can help reduce the complexity of the optimization problem to learn new energy optimal policies within a reasonable amount of time—time duration appears to be included in the nervous system's processes [129,130]. However, this can also restrict learning—the nervous system may not recognize the extent of what can be learned. Here, searching within a low dimensional space may help reduce the complexity of the problem but at the expense of cost savings.

Why does the nervous system initiate adaptation along step frequency but not step width? One candidate possibility is that the nervous system is primed to vary step frequency given our past walking experiences [131,132]. We walk at different speeds in different situations. And for each speed, we select a different step frequency that minimizes energetic cost [133]. Unlike step frequency, step width does not change considerably with speed, and thus the nervous system may not be primed to vary this gait parameter [134]. This is in line with our finding that people increase variability and initiate adaptation along step frequency but do not increase variability or initiate adaptation along step width. Another candidate possibility is that the nervous system cannot distinguish between energy cost savings due to step width and due to step frequency, and thus arbitrarily chooses one gait parameter to optimize. Two features of our experimental design and protocol may have added to this problem. First, we gave enforced experience using different forms of feedback for different gait parameters—

visual feedback for step width and audio feedback for step frequency. Future work should identify the type of enforced experience, both in terms of statistical structure and origin, that the nervous system can use for motor learning. Second, we gave enforced experience along two gait parameters at the same time. That the nervous system learns to adapt one gait parameter suggests that it can solve this problem to some extent. That it did not learn to adapt the other gait parameter suggests that it did not solve it completely. Future work should also determine how we can tailor this experience to facilitate learning.

How the nervous system learns new energy optimal control policies in the face of increasingly complex problems can advance our understanding of its algorithms for movement. Our intent was to determine if and how the nervous system's algorithms extend to optimizing energetic cost along multiple gait parameters—this may reveal how it defines its control policy search space. We found that the nervous system is limited in the number of gait parameters that it can simultaneously optimize, suggesting that it defines a low dimensional search space. This new understanding of the nervous system's control policy search space can benefit coaches who seek to design effective training protocols, as well as therapists who seek to design effective rehabilitation protocols. It is also important to understand the limitations of the nervous system's algorithms for applications in both healthy and impaired individuals.

Chapter 4.

General variability leads to specific adaptation toward optimal movement policies

4.1. Abstract

Our nervous systems can learn optimal control policies in response to changes to our bodies, tasks, and movement contexts. For example, humans can learn to adapt their control policy in walking contexts where the energy optimal policy is shifted along variables such as step frequency or step width. But it is unclear how the nervous system determines which ways to adapt its control policy. Here we asked how human participants explore through variations in their control policy to identify more optimal policies in new contexts. We created new contexts using exoskeletons that apply assistive torques to each ankle at each walking step. We analyzed four variables that spanned the levels of the whole movement, joint, and muscle: step frequency, ankle angle range, total soleus activity, and total medial gastrocnemius activity. We found that, across all of these analyzed variables, variability increased upon initial exposure to new contexts, and then decreased with experience. This led to adaptive changes in the magnitude of specific variables and these changes were correlated with reduced energetic cost. The timescales by which adaptive changes progressed and variability decreased were faster for some variables than others, suggesting a reduced search space within which the nervous system continues to optimize its policy. These collective findings support the principle that exploration through general variability leads to specific adaptation toward optimal movement policies.

4.2. Introduction

Humans are adept at learning optimal control policies. We use the term control policy to refer to the nervous system's mapping from states to the actions taken in those states [53]. The nervous system's actions are realized through motor commands, and its perceived states may range from the lengths of individual muscles to its estimate of the unevenness of the terrain. A control policy can be adapted to optimize an objective function, and when adaptation can no longer improve the objective, we refer to this as the optimal control policy. The objective function may consist of multiple terms, and the relative importance of these terms may depend on task and context. For example, in the task of reaching, people can adapt their control policy to optimize an objective function consisting of error and effort [5]. In walking, studies have shown that the nervous system's objective function includes metabolic energetic cost among other terms and constraints such as stability or risk of falling [6,105]. When we measure the relationship between energetic cost and step frequency, we find that it is bowl-shaped and that people prefer to walk with a step frequency that coincides with the minimum of this bowl [10,106]. And when we reshape this relationship to shift the energy optimal step frequency, we find that people adapt their step frequency to optimize energetic cost [42]. This same objective appears to influence the nervous system's control of other gait parameters such as step width, suggesting that the nervous system's continuous learning of optimal control policies for walking is a general phenomenon [80,86].

Optimizing a control policy involves several steps. One step is to select a more optimal control policy from amongst candidate policies. Studies suggest that the nervous system greedily selects local solutions that improve on the underlying objective function [41,75]. Prior to selecting the next policy, the nervous system must first evaluate a set of candidate policies. One way to do so is to locally explore. Songbirds, for example, exhibit variability in vocal control which appears to enable optimization of song performance [135–137]. Similar mechanisms underlie human vocal control [138]. That is, variability is not just an undesirable outcome of noise in the nervous system's control, but also a means to explore and discover better outcomes [64,131]. However, variability can be costly—deviating from the previously optimal policy in contexts where the optimal policy has not shifted yields suboptimal behaviour. Even in contexts where the optimal policy has shifted, variability that is not in the same direction as the shift is suboptimal. It seems conceivable that the nervous system has some understanding of which aspects of its control policy to explore and adapt given prior

experience. For example, in reaching arm movements, some people exhibit increases in baseline variability along aspects that are relevant to the task, and these selective increases in variability appear to enable faster learning [132]. But it is useful to consider an alternative perspective where the nervous system has minimal prior experience that it can draw from and may need to vary many aspects of its control policy in order to learn which aspects give rise to better outcomes. When learning to walk with an assistive device, for example, the nervous system may need to determine which aspects of its control policy to adapt—it is a new context that introduces a control system external to that of the nervous system.

How the nervous system explores the space of control policies may influence how quickly it learns new optimal policies. Exploration faces a challenge—as the number of possible states and actions increases, the number of candidate control policies increases rapidly. This expansion of the space of control policies can impede learning as searching through and evaluating many policies takes time—a challenge known as the 'curse of dimensionality' [53]. The nervous system might overcome this challenge of the combinatorial complexity of the policy space by reducing the dimensionality of the policy space that it searches so it has fewer dimensions along which to explore for new optimal control policies [139]. How the nervous system does this is presently unknown.

Here we determined how the nervous system explores through variations in its control policy to learn new optimal policies in new contexts. To accomplish this, we performed a post-hoc analysis of data from our recent study where we gave participants experience with ankle exoskeleton assistance over six non-consecutive days [140]. This recent study found that people arrive at new energy optimal policies when given sufficient experience with exoskeleton assistance [140]. But the nervous system's mechanisms for converging on this new optimal policy are still unclear. Here we tested three hypotheses about learning in new contexts: (1) the nervous system first explores by increasing general variability—where general variability refers to variability across many or all aspects of gait—to identify variables that improve its objective, (2) with experience, the nervous system selectively decreases variability to refine its control policy search space, and (3) the nervous system learns to adapt the magnitude of specific variables and exploit a new control policy that reduces energetic cost.

4.3. Results

We created new contexts using ankle exoskeletons. We applied assistive torques to each ankle at each walking step by transmitting forces through a Bowden cable that was attached to an ankle lever on an ankle exoskeleton. High-powered off-board motors generated the forces that were transmitted through the cables, and high-frequency controllers commanded the motors to generate the desired torques. All participants experienced two main contexts while walking on an instrumented treadmill: without-assistance and with-assistance. In the without-assistance context, participants walked without the ankle exoskeletons (Figure 4.1A) or while wearing the ankle exoskeletons but with slack cables and minimal applied torques (Figure 4.1B). In the with-assistance context, participants walked with the ankle exoskeletons while they generated torques that acted to extend each ankle during its stance phase. Figure 4.1C illustrates the general pattern of the applied torque, which was determined by a predefined control law that applied a constant magnitude of peak torque, while rise time, peak torque time, and fall time were constant percentages of the stride time.

To study the nervous system's learning mechanisms in new contexts, we gave participants experience with walking with exoskeleton assistance. On each day over six non-consecutive days, all participants walked without the ankle exoskeletons, with the ankle exoskeletons providing minimal applied torques, and with the ankle exoskeletons providing assistive torques. Participants completed these three conditions twice for six minutes each. The three conditions were first completed in random order, and then again in that same order, but reversed (e.g., CABBAC; Figure 4.1D and E). Participants completed additional exoskeleton assistance trials from the second day onwards (Figure 4.1E), but we did not include these trials in our current analyses as they were designed to test the effect of different training protocols [140]. In these additional trials, some participants (n = 5) repeatedly experienced the predefined control law described above. Other participants (n = 5) experienced this predefined control law interspersed between human-in-the-loop optimization, where we used real-time measures of energetic cost to customize the control law parameters (Figure A.1). Despite differences in these additional trials, all participants achieved similar reductions in energetic cost in response to the predefined control law that they repeatedly experienced on each day for six days ($p = 0.62$) [140]. We therefore grouped participants (n = 10) and restricted our analyses to changes in response to the predefined control law that we define above as the with-assistance context. When accounting for the amount of experience with

exoskeleton assistance, we include not only the time spent walking in the with-assistance trials that we analyzed, but also the time spent walking in the additional trials because participants also experienced assistive ankle torques in human-in-the-loop optimization.

We studied exploration and adaptation in this new context by measuring changes at the levels of the whole movement, the joint, and the muscle. We focused our analysis on four variables: step frequency, ankle angle range during stance, total soleus activity, and total medial gastrocnemius activity (see Methods). Step frequency can influence exoskeleton assistance through the timing of the assistive torque pattern because rise time, peak torque time, and fall time were all expressed as percentages of stride time in our control law parameterization. Ankle angle range may influence the power and work that the exoskeleton applies to the ankle by changing the angular displacement over which the assistive torque is applied. Lastly, the nervous system may learn to accept assistive torques at the ankle by lowering the contribution to the total ankle torque provided by the extensor muscles. The two primary ankle extensor muscles are soleus and gastrocnemius, and here we analyzed both of their activities. We selected these variables *a priori*, based on preliminary evidence of how walking can take advantage of ankle exoskeleton assistance [141–145]. These variables reflect only some of the nervous system's control policy parameters—there are likely many other parameters that allow the nervous system to, for example, achieve a particular step frequency using many combinations of ankle angle range, total soleus activity, and total medial gastrocnemius activity.

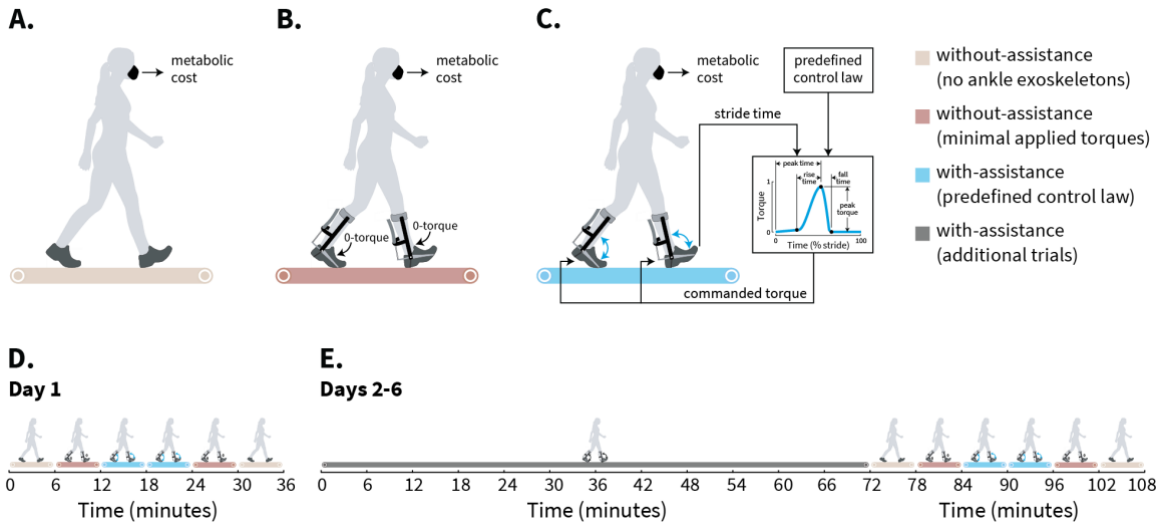

Figure 4.1: Experimental design and protocol

All participants experienced two main contexts: without-assistance and with-assistance. In the without-assistance context, participants (A) walked without the ankle exoskeletons or (B) walked while wearing the ankle exoskeletons with minimal applied torques via slack cables. In the with-assistance context, participants (C) walked while wearing the ankle exoskeletons which used a predefined control law to generate assistive ankle torques at each walking step. (D) On day 1, all participants completed these three conditions twice. (E) On days 2-6, all participants completed additional trials which were followed by the original three conditions twice.

4.3.1. A general increase in variability upon initial exposure to new contexts

We quantified variability within each variable by high-pass filtering its signal to include timescales of 30 steps or less, and then calculating the standard deviation of this filtered signal during the last three minutes of each six-minute trial (see Methods). This method filtered out the relatively slow signal changes that we associate with adaptive changes, but not the relatively rapid changes that occur from step to step and over several steps that we associate with exploration. For most variables, we quantified each participant's without-assistance variability from the condition where they were not wearing the ankle exoskeletons. The exception was ankle angle range where we applied identical calculations but to the condition where participants were walking while wearing the ankle exoskeletons but with the devices applying minimal torques. The reason for this exception was that we needed the exoskeleton sensors to calculate ankle angle. We refer to the without-assistance variability averaged across the two trials on the first day as *baseline variability*. We use this baseline variability to normalize with-assistance variability.

Upon initial exposure to the exoskeleton assistance context, participants walked with increased variability across all variables that we analyzed. This was evident when comparing with-assistance variability averaged across the two first day trials to without-assistance baseline variability, with increases ranging from 21% to 279% (mean ± standard deviation, paired t-test; step frequency: +56.2 ± 27.6%, $p = 2.0 \times 10^{-4}$, Figure 4.2A; ankle angle range: +278.8 ± 85.0%, $p = 6.8 \times 10^{-7}$, Figure 4.2B; total soleus activity: +21.1 ± 25.3%, $p = 0.026$, Figure 4.2C; total medial gastrocnemius activity: +46.0 ± 19.6%, $p = 1.7 \times 10^{-5}$, Figure 4.2D; representative participant: Figure 4.2E).

4.3.2. A general decrease in variability with increased experience

As participants walked with exoskeleton assistance over multiple days, we determined how variability changed with this increased experience. We found that, as experience increased, participants walked with decreased variability across all variables that we analyzed. This was evident when comparing with-assistance variability measured on the last day to that measured on the first day, with decreases ranging from -18% to -39% (mean ± standard deviation, paired t-test; step frequency: -38.5 ± 9.8%, $p = 4.9 \times 10^{-5}$, Figure 4.2A; ankle angle range: -26.3 ± 22.8%, $p = 6.5 \times 10^{-3}$, Figure 4.2B; total soleus activity: -18.3 ± 21.2%, p = 0.013, Figure 4.2C; total medial gastrocnemius activity: -19.7 ± 13.2%, $p = 1.5 \times 10^{-3}$, Figure 4.2D). By the last day, participants' with-assistance variability was indistinguishable from their baseline variability for step frequency (-5.1 ± 17.3%, paired t-test: p = 0.32) and total soleus activity (-3.3 ± 26.9%, paired t-test: p = 0.39), but remained elevated for ankle angle range (+166.2 ± 52.5%, paired t-test: $p = 6.7 \times 10^{-6}$) and total medial gastrocnemius activity (+15.8 ± 15.8%, paired t-test: p = 0.020). That is, the nervous system returned variability towards, and in some cases to, baseline variability with increased experience.

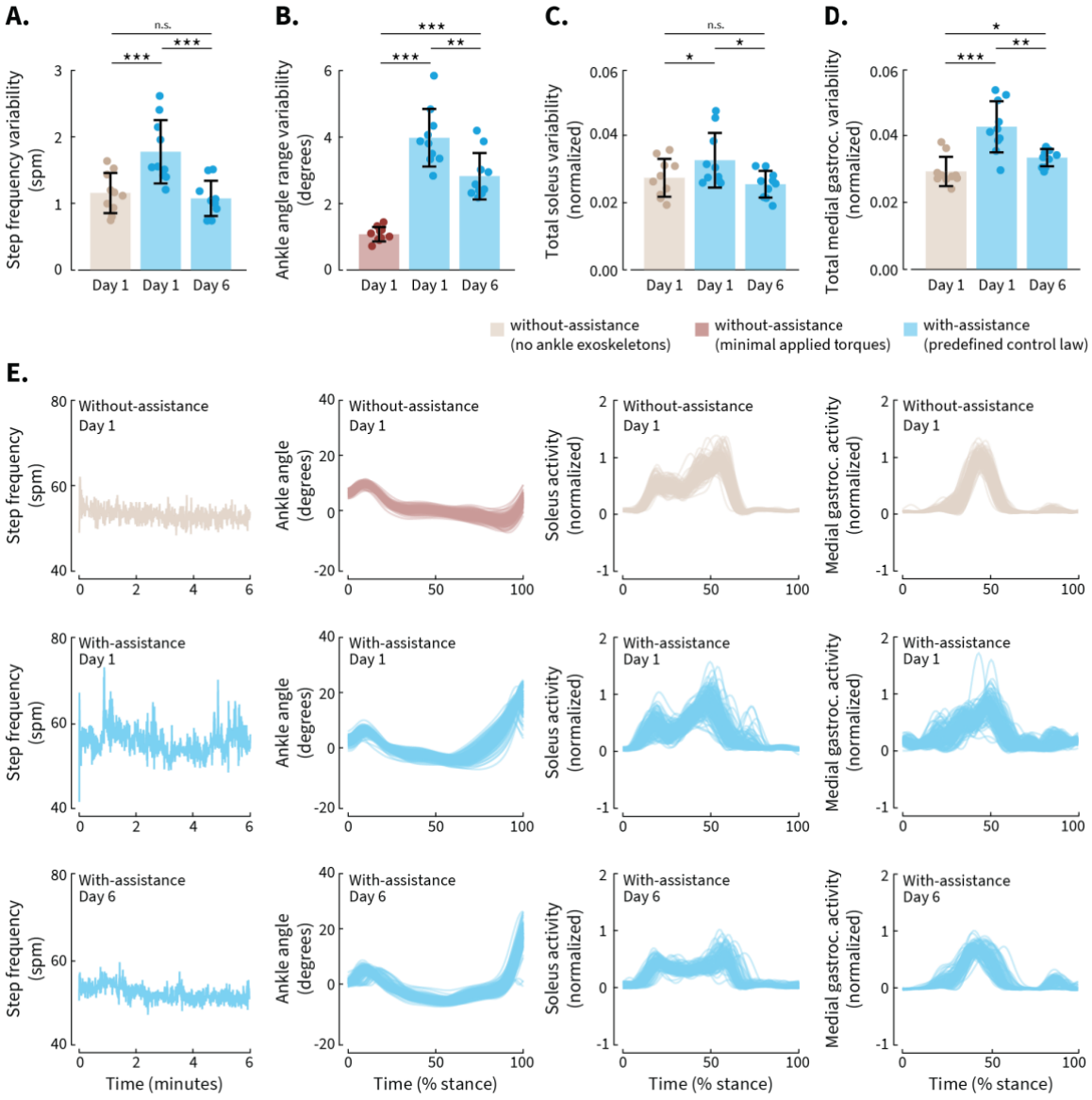

Figure 4.2: Changes in variability with experience with exoskeleton assistance
(A) Step frequency variability, (B) ankle angle range variability, (C) total soleus variability, and (D) total medial gastrocnemius variability. Variability in muscle activity is expressed as a fraction of the peak activation (see Methods). Walking without the ankle exoskeletons is shown in beige, walking with the ankle exoskeletons with minimal applied torques is shown in red, and walking with the ankle exoskeletons with the predefined control law generating assistive torques at each walking step is shown in blue. Circular markers are participant averages, which we calculated by averaging each variable's variability across two six-minute trials on a given day. Bar height is average across participants, error bars represent one standard deviation, and asterisks indicate statistically significant differences between conditions or days using the notation: *** for $p<0.001$, ** for $p<0.01$, * for $p<0.05$, and n.s. for not significant. (E) Measured step frequency in strides per minute (spm) as a function of walking time in a given six-minute trial, and measured ankle angle, soleus activity, and medial gastrocnemius activity for the left leg as a function of stance phase at each walking step in a given six-minute trial. Data are from one representative participant.

4.3.3. Adaptation occurs along specific variables and these changes correlate with reduced energetic cost

We use the term *adaptation* to refer to changes in the magnitude of a variable that occurs with experience. We quantified each variable's magnitude by averaging its signal during the last three minutes of each six-minute trial. People appear to have an established policy for without-assistance contexts—we observed minimal changes in the magnitude of variables when comparing the first and last days of without-assistance walking (mean ± standard deviation, paired t-test; step frequency: -0.78 ± 2.6%, p = 0.38; ankle angle range: 5.6 ± 7.2%, p = 0.046; total soleus activity: -2.7±8.1%, p = 0.22; total medial gastrocnemius activity: -4.0 ± 6.5%, p = 0.074). We refer to the without-assistance magnitude averaged over the two trials on the first day as the baseline value. We use this baseline value to normalize the measured with-assistance magnitudes. We estimated the energetic cost of each trial in the standard manner—using respiratory gas analysis of the last three minutes of each six-minute trial (see Methods). We normalized energetic cost during all with-assistance trials to each participant's energetic cost averaged over two without-assistance trials—when walking with the ankle exoskeletons applying minimal torques—on the first day. We used linear mixed-effects regression to estimate the slope of the relationship between a variable and energetic cost. This mixed-effects model used a single slope for each variable to estimate the relationship that is shared between participants while allowing for individual energetic cost intercepts (see Methods).

Participants learned to adapt three of four variables and these three variables correlate with energetic cost. We found relatively strong and significant relationships between energetic cost and step frequency (slope = 2.2, 95% CI [1.5, 2.8], $p = 3.4 \times 10^{-9}$; Figure 4.3A), total soleus activity (slope = 0.69, 95% CI [0.44, 0.94], $p = 2.9 \times 10^{-7}$; Figure 4.3C), and total medial gastrocnemius activity (slope = 0.36, 95% CI [0.18, 0.54], $p = 1.7 \times 10^{-4}$; Figure 4.3D). The relationship between ankle angle range and energetic cost was weaker and not significant (slope = 0.069, 95% CI [-0.030, 0.17], p = 0.17; Figure 4.3B). As participants gained experience with walking with exoskeleton assistance, variables that correlated with energetic cost adapted in the direction that reduced cost. Comparing the first and last days of with-assistance trials, we found changes in step frequency (-3.5 ± 3.4%, paired t-test: $p = 9.8 \times 10^{-3}$, Figure 4.3E), total soleus activity (-10.4 ± 11.0%, paired t-test: p = 0.023, Figure 4.3G), and total medial gastrocnemius activity (-13.4 ± 12.4%, paired t-test: $p = 5.4 \times 10^{-3}$, Figure 4.3H),

but not in ankle angle range (-0.32 ± 11.5%, paired t-test: p = 0.78, Figure 4.3F). Participants learned to exploit a new control policy that reduced energetic cost by -25.0 ± 9.9% (p = 3.0 x 10^{-5}, Figure 4.3I) when walking with ankle exoskeleton assistance on the last day compared to the first day. To be clear, adaptation along a specific variable that correlates with reductions in energetic cost suggests but does not prove that energetic cost drives this adaptation. We cannot rule out, for example, that the observed adaptation is due to optimization of different objectives that also correlate with changes in the observed variables.

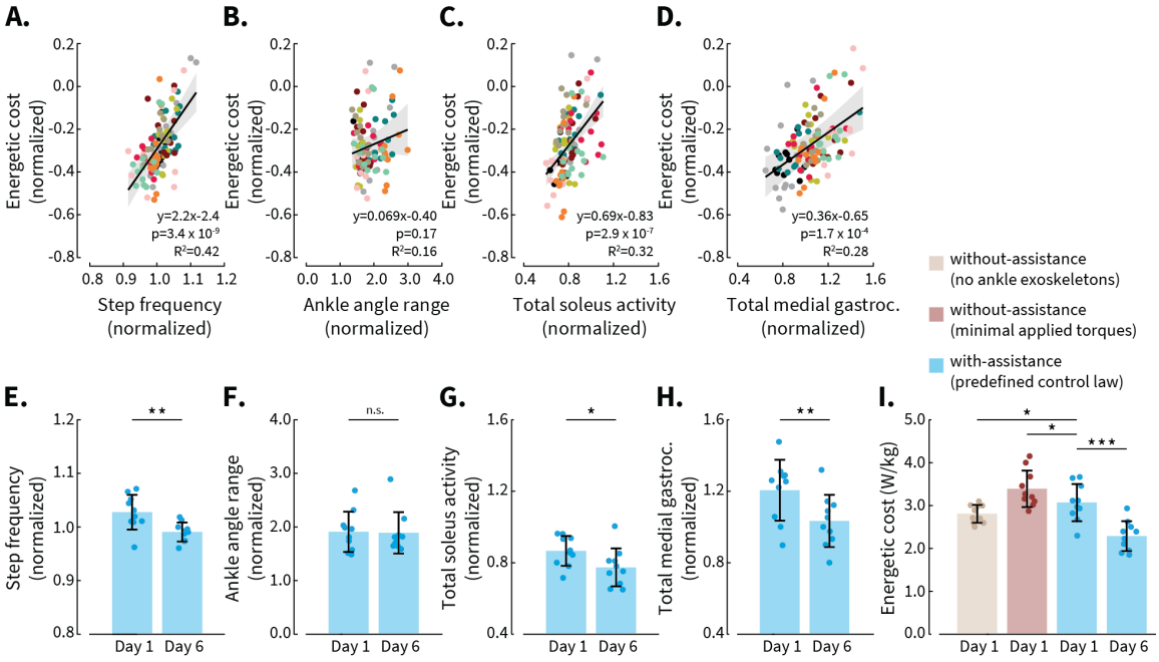

Figure 4.3: Changes in magnitude of variables that reduce energetic cost
Estimated relationships between energetic cost and (A) step frequency, (B) ankle angle range during stance, (C) total soleus activity, and (D) total medial gastrocnemius activity during with-assistance trials across all days. We normalized energetic cost during all with-assistance trials to each participant's energetic cost during without-assistance trials—when walking with ankle exoskeletons with minimal applied torques—on the first day. We normalized each variable's with-assistance magnitude to each participant's without-assistance baseline value. The solid black lines are linear mixed-effects models with 95% confidence intervals (grey shading). Individual participants are represented by distinct coloured circles. Each participant experienced the same number of with-assistance trials—two per day for six days. To better illustrate a variable's relationship with energetic cost as determined by the linear mixed-effects model, we subtracted each participant's random-effects intercept term from their energetic cost data, which changes their cost data between variables (see Methods). Magnitude of (E) step frequency, (F) ankle angle range, (G) total soleus activity, and (H) total medial gastrocnemius activity during with-assistance trials on the first day and on the last day. (I) Energetic cost during without-assistance trials on the first day, as well as during with-assistance trials on the first day and the last day. Without-assistance trials include walking without the ankle exoskeletons (beige) and walking with the ankle exoskeletons with minimal applied torques (red). With-assistance trials are walking with the ankle exoskeletons generating assistive torques at each step (blue). Bar height is average across participants, error bars represent one standard deviation, and

asterisks indicate statistically significant differences between days using the notation: *** for p<0.001, ** for p<0.01, * for p<0.05, and n.s. for not significant.

4.3.4. Variability decreases quickly for quickly-adapting variables

We modeled the timing of adaptive changes, as well as the timing of decreases in variability, as an exponential decrease from an initial value to a final steady-state value. We used nonlinear mixed-effects regression with a single time constant to estimate the time constant that is shared between participants while allowing for individual participant offsets (Equation 4.2 in Methods). We then used bootstrapping to determine the dispersion of each time constant (see Methods). We found that participants adapted step frequency with a time constant of 208.9 minutes (IQR [150.0, 311.7], R-squared = 0.52; Figure 4.4A), total soleus activity with a time constant of 5.7 minutes (IQR [3.9, 8.6], R-squared = 0.47; Figure 4.4C), and total medial gastrocnemius activity with a time constant of 7.8 minutes (IQR [6.3, 10.2], R-squared = 0.77; Figure 4.4D). We found that variability decreased along step frequency with a time constant of 112.4 minutes (IQR [97.5, 129.9], R-squared = 0.70; Figure 4.4E), ankle angle range with a time constant of 87.4 minutes (IQR [70.5, 108.7], R-squared = 0.48; Figure 4.4F), total soleus activity with a time constant of 4.6 minutes (IQR [3.4, 6.3], R-squared = 0.58; Figure 4.4G), and total medial gastrocnemius activity with a time constant of 5.6 minutes (IQR [4.6, 6.9], R-squared = 0.63; Figure 4.4H).

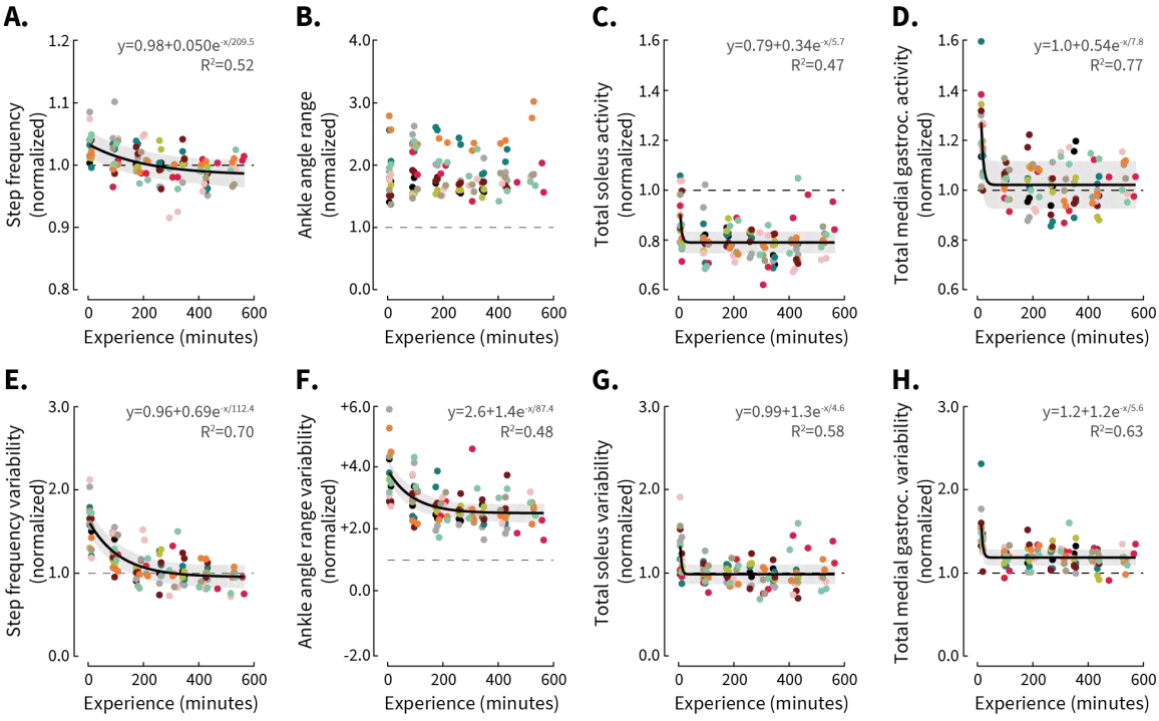

Figure 4.4: Timescales of changes in variability and changes in magnitude
Changes in magnitude of (A) step frequency, (B) ankle angle range during stance, (C) total soleus activity, and (D) total medial gastrocnemius activity as experience increased. Changes in variability of (E) step frequency, (F) ankle angle range during stance, (G) total soleus activity, and (H) total medial gastrocnemius activity as experience increased. For each variable, we normalized with-assistance magnitude and variability to each participant's without-assistance baseline levels (dashed horizontal lines). The solid black lines are exponential model fits with 95% confidence intervals (grey shading). We did not fit an exponential model to (B) as we did not observe changes in magnitude of ankle angle range during stance. Individual participants—represented by distinct coloured circles—experienced the same number of with-assistance trials—two per day for six days—but at different experience times depending on the design of their additional trials. To better illustrate the exponential fit, we subtracted each participant's random-effects offset from the plotted data. The elapsed experience time includes not only the time spent walking in the with-assistance trials that we analyzed here, but also the time spent walking in the additional exoskeleton training trials.

Variables that adapted quickly also had rapid decreases in variability. We used bootstrapping to determine the dispersion of the time constant of adaptation for each variable, and then tested for differences in time constants between variables (see Methods). We found that total soleus activity and total medial gastrocnemius activity adapted with faster time constants than step frequency (ANOVA; total soleus activity vs. step frequency: $p < 0.001$; total medial gastrocnemius activity vs. step frequency: $p < 0.001$; total soleus activity vs. total medial gastrocnemius activity: $p < 0.001$; Figure 4.5A). Next, we performed the same analysis but for time constants of variability. We found that variability in a similar way decreased faster for total soleus activity and total medial gastrocnemius activity than for step frequency (ANOVA; total soleus activity vs. step frequency: $p < 0.001$; total medial gastrocnemius activity vs. step frequency: $p < 0.001$; total soleus activity vs. total medial gastrocnemius activity: $p < 0.001$; Figure 4.5B).

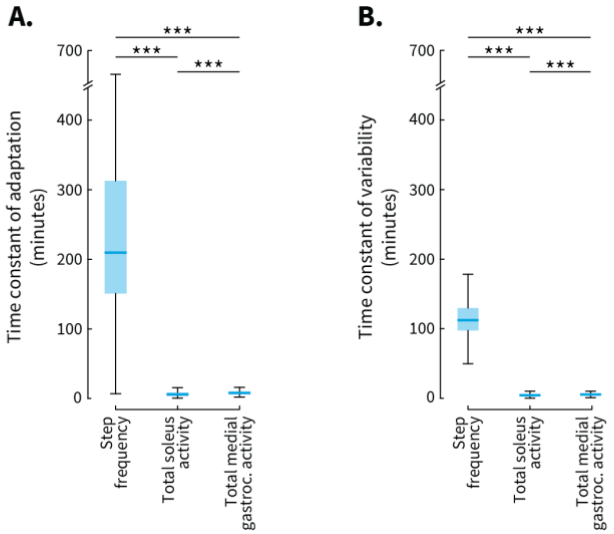

Figure 4.5: Differences in timescales between variables

(A) Time constants of adaptation and (B) time constants of variability for step frequency, total soleus activity, and total medial gastrocnemius activity during with-assistance trials (blue). We did not observe adaptation in ankle angle range during stance and therefore excluded this variable from this analysis. The central mark indicates the median, the bottom edge of the box indicates the lower quartile (25th percentile), and the top edge of the box indicates the upper quartile (75th percentile). Error bars extend to the most extreme data points not considered outliers, which we define as more than 1.5 times the interquartile range away from the edges of the box. Asterisks indicate statistically significant differences between variables using the notation: *** for $p<0.001$.

4.4. Discussion

We provide insight into how the nervous system navigates a space of control policies to learn new optimal policies in new contexts. We created new contexts using ankle exoskeleton assistance and studied learning as energy optimization in human walking. We analyzed two processes—variability and adaptation—across four variables—step frequency, ankle angle range, total soleus activity, and total medial gastrocnemius activity. We found that, with minimal experience in new contexts, variability increased across all variables that we analyzed. And with increased experience, variability decreased across all variables. This appeared to lead to adaptive changes in the magnitude of specific variables—those which were associated with reductions in energy cost. Variability decreased and adaptation progressed quickly for some variables but slowly for others, suggesting that the nervous system can independently control these variables, and that it may optimize in a manner that reduces its control policy search space over time.

These findings generalize to other movement variables. We *a priori* selected four variables based on our understanding of how walking can take advantage of ankle exoskeleton assistance and without knowledge of how these variables changed over time or how these changes were associated with energetic cost. These four variables are only a subset of our measured dataset—which includes stride parameters, ground reaction forces, joint kinematics, and muscle activity—enabling us to test whether the conclusions we arrive at from our first analysis generalize to other variables. Toward this, we sampled four additional variables: step width, peak ankle extension angle (which occurs during swing), total rectus femoris activity (a knee extensor and hip flexor muscle), and total biceps femoris activity (a knee flexor and hip extensor muscle). These four additional variables do not directly take advantage of ankle exoskeleton assistance but in the same way capture learning at the levels of the whole movement, the joint, and the muscle. Our additional analysis revealed that participants first increased and then decreased variability across three of these four

variables: peak ankle extension angle, total rectus femoris activity, and total biceps femoris activity (Figure A.2). Participants learned to adapt along specific variables and these variables were associated with reductions in energetic cost (Figure A.3). Some variables adapted faster than others, and this adaptation was accompanied by decreases in variability (Figure A.4). We interpret these collective findings as supporting the principle that general variability leads to specific adaptation toward optimal control policies. Our data and code are open access, allowing others to test additional variables for the generalizability of these conclusions.

These findings generalize to other movement contexts. Split-belt walking is another movement context in which people learn to take advantage of external assistance and reduce their energetic cost. That is, people learn to adapt their foot placement to take advantage of the positive work performed by the treadmill—due to one belt moving faster than the other—and reduce the work performed by their legs. Learning to take advantage of external assistance from a split-belt treadmill is different from learning to take advantage of external assistance from ankle exoskeletons—the difference in belt speeds is constant whereas the assistive ankle torque is controlled as a function of stride time at each walking step. Here we used a split-belt walking dataset to ask how our findings generalize to other movement contexts [52]. We analyzed step length asymmetry, which is commonly reported in split-belt walking studies. We found that variability in step length asymmetry increased upon initial exposure to the new context of split-belt walking compared to tied-belt walking ($p = 9.2 \times 10^{-10}$; Figure 4.6). Variability then decreased by the end of 45 minutes of split-belt walking compared to initial variability ($p = 1.2 \times 10^{-4}$; Figure 4.6). During this time, participants learned to adopt positive asymmetries—with foot placement further forward on the fast belt—and reduce metabolic cost [52]. Thus, our conclusions that general variability leads to specific adaptation toward optimal control policies generalize not only to other movement variables but also to other movement contexts.

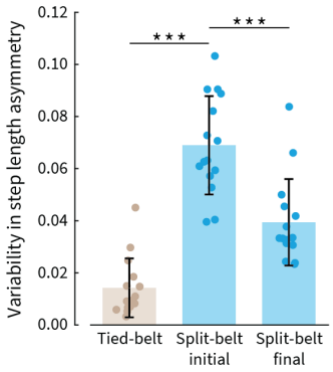

Figure 4.6: Changes in variability in response to the new context of split-belt walking

Variability in step length asymmetry during the first 100 strides of tied-belt walking (beige), during the first 100 strides of split-belt (blue), and during the last 100 strides of split-belt (blue). Bar height is average across participants (n = 15), error bars represent one standard deviation, and asterisks indicate statistically significant differences between conditions or days using the notation: *** for $p<0.001$.

The nervous system appears to reduce the dimensionality of its search space over time. Here we infer the extent of the space in which the nervous system searches for new policies from the extent of our measured gait, joint, and muscle variability. Variability first increases across all biomechanical variables, suggesting a large search space. Variability then quickly decreases for some biomechanical variables as their adaptation progresses, suggesting a reduced space in which the nervous system optimizes fewer control policy parameters. The biomechanical variables we have chosen to measure do not map perfectly onto the nervous system's control policy parameters—it is unlikely that the nervous system's control policy has a parameter specifically for step frequency or total soleus activity. Consequently, if the nervous system were to stop exploring and start exploiting just one of its control parameters, we would likely measure this as a small decrease in variability across many biomechanical variables rather than a return to baseline variability in any single biomechanical variable. And because biomechanical variability is a consequence of many control parameters, that some biomechanical variables do return to baseline variability suggests a large reduction in the nervous system's control policy search space.

A reduced search space is superior to continuing with general exploration of learned policies. One reason for this is that continuing to explore variables that are already optimized results in needlessly increased costs. A second reason is that searching by general exploration may take a long time. That is, exploring more variables results in larger search spaces, and these search spaces fall victim to the curse of dimensionality where there are exponentially more

candidate control policies to evaluate [123]. If the nervous system reduces the number of variables along which it explores, it has fewer combinations of states and actions to try and can therefore employ a more directed optimization for selecting more optimal control policies from amongst candidate policies [41,75]. This understanding can be applied to the optimization of any cost function. For example, a multi-objective cost function might provide a more general view of optimization during non-steady-state walking by including terms such as energy, or stability, or some weighted combination of these and other terms [6,86]. A recent study has provided a theoretical basis for how people learn new optimal policies in new contexts such as walking with exoskeleton assistance and walking on a split-belt treadmill—they found that prioritizing stability over short timescales and improving energy expenditure over long timescales, as well as using exploratory variability to estimate gradients, can explain learning in these contexts [146]. This notion that the nervous system uses exploration to first evaluate a set of candidate policies is similar to what we find in our current study. In contrast to general exploration during learning of a new policy, learned policies often appear to have a low-dimensional structure. In motor coordination, for example, muscle activation patterns can be explained by a limited set of muscle synergies—or muscle activation patterns with consistent spatial and temporal characteristics [20,21,23,147]. And in neural systems such as the motor cortex of a monkey, relatively complex responses from individual neurons can be explained by relatively simple responses from a population of neurons [148,149]. Our findings show that, in new contexts, the nervous system can arrive at such a low-dimensional structure of control after first benefiting from general exploration. An interesting and open question is what elicits adaptation and why adaptation occurs quickly for some variables, and more slowly for others. There are several factors—which do not act in isolation—that can influence adaptation. Two factors may be the level of exploratory variability and the cost gradient—both can influence the range of cost savings that the nervous system experiences. Another factor may be how the biomechanical variables relate to the nervous system's control policy parameters—a variable that we measure may reflect one or many parameters that the nervous system optimizes, and the level of complexity may influence the timescale of adaptation that we observe.

Not all variability is exploration. We must first differentiate exploration from other contributors to measured biomechanical variability, such as that arising from unintentional noise in the nervous system's control, variability in the forces that muscles produce, or unpredicted changes in the environment [70]. Müller and Sternad sought to do so in a

throwing task, where participants could influence their performance of hitting a target by varying two parameters: angle and velocity at release [150]. They decomposed variability along these two parameters into components that differentiated task-relevant variability from stochastic noise. Here we propose that baseline levels of variability in familiar contexts may mostly represent unintentional noise, either from the nervous system's control or some other source (i.e., stochastic noise in Müller and Sternad's terminology). This can then be used as a benchmark against which variability in new contexts is compared. We found that, upon first exposure to a new context, variability increased above baseline levels across all variables that we analyzed. With experience in the new context, variability gradually decreased and then plateaued across all variables that we analyzed, converging on baseline levels for some variables. We interpret variability as exploration when it is information seeking. That variability decreased as people learned to adapt the magnitude of specific variables—and that these two processes occurred over similar timescales for each variable—suggests that variability may reflect exploration. As in other studies that aim to determine how variability relates to intentional exploration, we recognize that we can never entirely rule out the alternative hypothesis that variability is unintentional. However, this alternative hypothesis seems less likely based on theory—a key concept in reinforcement learning is that exploration improves learning of new optimal policies—and evidence—higher levels of motor variability appear to enable faster learning of new optimal policies [53,132]. Plateauing of variability may indicate that the nervous system has settled on a new policy and has shifted from exploring candidate policies to exploiting the new preferred one. That variability remained elevated above baseline levels for some variables may reflect that the nervous system is still refining aspects of the control policy or, as we suspect, it may simply be additional variability introduced by imperfect torque control by the exoskeleton, which was not present when we established baseline levels.

A deeper understanding of the nervous system's mechanisms for learning can be used to both facilitate learning and customize training. We might facilitate learning by encouraging general exploration. That is, we might increase variability across many variables—through strategies such as biofeedback—and then decrease variability along specific variables as people learn to adapt, reducing the nervous system's policy search space. We might also give experience with specific variables that affect energetic cost, indicating to the nervous system which variables are relevant to optimize [73]. To improve on the design of the present study, future studies should seek to determine the energetic cost landscape of walking with ankle

exoskeleton assistance by mapping the relationship between energetic cost and biomechanical variables. Future studies should also seek to develop methods for estimating energetic cost with increased time resolution to determine how variability in biomechanical variables relates to changes in energetic cost for energetic cost optimization. This can benefit those who seek to design wearable systems—such as orthoses, exoskeletons, and prosthetics—by facilitating learning, and then evaluating people's optimal responses to a range of designs. We also might customize training time by using baseline levels of variability as a benchmark to indicate when the nervous system initiates exploration, and when it shifts to exploiting a new policy. Participants in this study had roughly nine hours of experience walking in this new context [140]. Determining the onset and termination of learning would be useful for coaching athletes who would benefit from knowing at which point they should transition to learning new skills in order to maximize their high capacity for training. And knowing when to terminate experience is especially important for rehabilitating those with mobility disorders who have a limited capacity for training.

4.5. Methods

4.5.1. Experimental design

Participants walked on an instrumented treadmill while wearing a bilateral, tethered ankle exoskeleton emulator that applied assistive torques to each ankle at each walking step. This system is described in more detail in our previous work [140,151]. In brief, we used an off-board controller to command the desired force to an off-board electric motor via a motor driver [152]. This force was then transmitted through a Bowden cable to the end of the ankle lever on the exoskeleton, which in turn applied an ankle plantarflexion torque to the participant. This ankle plantarflexion torque was achieved by the tension in the cable, combined with the moment arm of the ankle lever on the exoskeleton, producing forces on the body where it interfaces with the exoskeleton—the heel, the shank, and the toe [151].

Our control system produced desired torque patterns as a function of the user's stride time. As detailed in [140], we used a high-speed controller running at 1000 Hz to achieve the predefined torque pattern (Speedgoat, Liebefeld, Switzerland). This controller sampled from sensors, calculated time since heel strike as well as desired force at that time, and then commanded that force to the motor. We used contact switches on each heel to measure heel strikes and calculate stride time as the difference in time between consecutive heel strikes with the same leg (Pololu, NV, USA; McMaster-Carr, IL, USA). We used strain gauges to measure the tension in each cable and calculate torques by multiplying tension with the moment arm of the ankle lever on the exoskeleton (Omega Engineering, CT, USA). The combination of real-time measured stride times and torques allowed for accurate torque tracking that was a function of time since heel strike, normalized to average stride time. We used average stride time as a filter for large, perhaps inaccurate, changes in stride time that may result in undesirable torques. We calculated average stride time as:

$$\tau_{avg} = t_{avg}(1 - \mu) + t_{stride}\mu \tag{4.1}$$

where $\mu = 0.9$. Our controller was also designed to include many features that prioritized the user's comfort [140]. For example, at the beginning of all walking periods with exoskeleton assistance, the peak torque was slowly increased from zero to that of the predefined torque pattern to ensure that participants were not perturbed by the assistance. During each swing phase, the controller initiated a 'swing mode' where the cable tracked the ankle angle with

added slack. And during each stance phase, the torque was slowly increased from zero after heel strike and decreased to zero before toe-off as prescribed by the parameterization of the control law.

We parameterized the control law to achieve a range of customized torque profiles, as well as one predefined torque profile. Because our original study was designed to test the relative benefits of training with a predefined control law and a customized control law [140], some participants received training with repeated exposure to the predefined control law, whereas other participants received training with human-in-the-loop optimization of the control law and only periodic exposure to the predefined control law. All control laws were defined by four parameters: magnitude of peak torque (Nm), timing of peak torque (% stride), rise time (% stride), and fall time (% stride). We defined the magnitude of peak torque to be a function of each participant's body mass. We determined the predefined control law in a pilot experiment prior to the main experiment [140]. In this pilot experiment, ten participants—which were not the participants in our main experiment—completed one day of habituation with the bilateral ankle exoskeletons followed by one day of human-in-the-loop optimization of the control law. The predefined control law for the main experiment was defined as the average optimized parameters from this pilot study: magnitude of peak torque was 0.54M where M is participant's mass in kilograms, timing of peak torque was 52.9% of stride, rise time was 26.2% of stride, and fall time was 9.8% of stride.

We measured ground reaction forces, joint kinematics, and muscle activity to quantify variables that people may optimize. First, we measured ground reaction forces and moments while participants walked on an instrumented split-belt treadmill (Bertec, Columbus, OH, USA). We used ground reaction forces and moments to calculate the center of pressure and identify foot contact events as the rapid fore-aft translation in center of pressure during double support. Second, we measured ankle angle in all trials with the exoskeleton using a rotary magnetic encoder mounted on the ankle joint of the exoskeleton (Renishaw, Gloucestershire, UK). We zeroed the encoder during standing on each day, and calculated ankle flexion and extension as the angle from neutral position in the sagittal plane. Third, we collected electromyography (EMG) data from medial gastrocnemius, lateral gastrocnemius, soleus, tibialis anterior, rectus femoris, vastus medialis, biceps femoris, and semitendinosus using surface electrodes on each muscle for both legs, and during all walking trials (Delsys, Boston, MA, USA). For the muscles that we considered in our analyses, we inspected EMG data

to exclude channels with poor signal quality. Lastly, we used a respiratory gas analysis system to measure rates of oxygen consumption and carbon dioxide production (Cosmed Quark CPET, Rome, Italy). We calculated gross metabolic power using the standard Brockway equation, and then subtracted each participant's resting metabolic power measured on the same day to obtain net metabolic power [112]. We calculated resting metabolic power as the average metabolic power during the final three minutes of a six-minute standing resting period at the beginning of each day of testing. We collected EMG at 1000 Hz and all other measures at 500 Hz.

4.5.2. Experimental protocol

We included a total of ten participants (mean ± SD; age: 24 ± 2 years; body mass: 68.3 ± 11.1 kg; height: 1.7 ± 0.094 m; sex: 4 females, 6 males) in our study. To investigate the nervous system's learning mechanisms, we required that participants included in our study learned to reduce their energetic cost of walking with exoskeleton assistance. These ten participants consisted of two groups, which we randomly assigned participants to prior to the experiment. Each group of five participants completed similar protocols but with slight differences in additional trials that they completed from the second day onwards. Our previous study found that, despite these slight differences, these two groups learned to reduce their energetic cost of walking in response to a general pattern of assistive ankle torque on the last day compared to the first day (group 1: $p = 0.016$; group 2: $p = 0.005$). They also achieved similar reductions in energetic cost ($p = 0.62$) [140]. We therefore grouped participants ($n = 10$) and restricted our analyses to changes in response to this general pattern of assistive ankle torque. We excluded a third group of five participants as they did not meet our requirement of learning to reduce their energetic cost of walking on the last day compared to the first day (group 3: $p = 0.73$) [140]. This was perhaps due to the nature of their additional trials, where they experienced many different patterns of assistive ankle torques. All participants were healthy and had no known gait or cardiopulmonary abnormalities. The Stanford Institutional Review Board approved the study protocol, and all participants gave their written, informed consent before participating in the study.

The protocol consisted of a testing session on each day for a total of 6 days. During all testing sessions, participants walked on an instrumented treadmill at a constant speed of 1.25 m/s.

For conditions that involved exoskeleton assistance, we instructed participants to "walk comfortably" and to "let the device do the work for you".

We gave participants experience with walking with exoskeleton assistance over multiple days. On each day, participants experienced at least 3 conditions: walking without the ankle exoskeletons (Figure 4.1A), walking with the ankle exoskeletons providing minimal applied torques via slack cables (Figure 4.1B), and walking with the ankle exoskeletons with the predefined control law generating assistive ankle torques at each walking step (Figure 4.1C). Some participants had an additional walking condition from the second day onwards, where they experienced the final control law from human-in-the-loop optimization during additional trials on that same day. Participants completed these conditions twice for six minutes each. The conditions were first completed in random order, and then again in that same order, but reversed (e.g., CABBAC; Figure 4.1D and E). When comparing between conditions or days, we averaged a given variable over the final three minutes of each six-minute trial, and then across the two repeated six-minute trials on each day.

Participants completed additional trials which altered the training that each group received. We did not include these trials in our analyses as they were designed primarily to test the effect of different training protocols in our previous study [140]. They were also designed to increase participants' experience with the predefined control law. Participants completed these trials from the second day onward, prior to the main six-minute trials (Figure 4.1E). In brief, one group (n = 5) repeatedly experienced the predefined control law at each walking step over 72 minutes (Figure A.1). A second group (n = 5) experienced two minutes of the predefined control law followed by 16 minutes of human-in-the-loop optimization and repeated this four times resulting in 72 minutes of training (Figure A.1). During periods of human-in-the-loop optimization, participants experienced a series of eight control laws for two minutes each. We selected these control laws based on our estimate of the optimal control law, which was determined by an algorithm that ranked previously experienced control laws by their respective energetic cost measurements. The process of human-in-the-loop optimization is described in more detail in our previous work [140,151].

4.5.3. Analysis

We wrote custom MATLAB scripts to process and analyze the data, as well as perform statistical comparisons and generate figures included in this manuscript.

We quantified the variability and magnitude of step frequency. This variable can influence exoskeleton assistance through the timing of the assistive torque pattern because rise time, peak torque time, and fall time were all expressed as percentages of stride time in our control law parameterization. We calculated step frequency by identifying foot contact events and then taking the inverse time difference between consecutive steps. We determined the variability within step frequency by applying a third-order, high-pass, bi-directional digital Butterworth filter with a cut-off frequency of 0.033 $steps^{-1}$ (period of 30 steps), and then calculated the standard deviation of this filtered signal during the last three minutes of each six-minute trial. We used MATLAB's filtfilt command to perform zero-phase digital filtering. We determined the cut-off frequency to be 0.033 $steps^{-1}$ based on previous studies of variability in human walking [125], as well as visual inspection of the power spectrum. We calculated the magnitude of step frequency as the average of this signal during the last three minutes of each six-minute trial.

We quantified the variability and magnitude of ankle angle range during stance. This variable can influence the power and work that the exoskeleton applies to the ankle by changing the angle over which the torque is applied. We calculated ankle angle range during stance by first time-locking ankle angle to heel strike events, and then calculating the difference between the maximum and minimum ankle angles during stance. We determined the variability within ankle angle range by applying a high-pass filter with a cut-off frequency of 0.033 $steps^{-1}$ (third-order Butterworth), and then calculated the standard deviation of this filtered signal during the last three minutes of each six-minute trial. We calculated the magnitude of ankle angle range as the average of this signal during the last three minutes of each six-minute trial.

We quantified the variability and magnitude of total ankle extensor muscle activity. The nervous system may learn to accept assistive torques at the ankle by lowering the contribution to the total ankle torque provided by the extensor muscles. We selected two variables at the level of the muscle as there are two primary ankle extensor muscles—soleus and gastrocnemius. For each muscle on each leg, we applied a high-pass filter with a 20 Hz cut-off (third-order Butterworth), rectified the signal, and then applied a low-pass filter with

a 6 Hz cut-off (third-order Butterworth) [153]. We next time-locked the signal to heel strike events, divided each stance phase into 100 evenly spaced segments, and normalized each muscle's activity for each participant to their same day average peak activation while walking without the ankle exoskeletons. We calculated total soleus activity and total medial gastrocnemius activity by integrating each muscle's activity during stance at each walking step. Similar to our previous analyses, we quantified the variability of total soleus activity and total medial gastrocnemius activity by applying a high-pass filter with a cut-off frequency of 0.033 $steps^{-1}$ (third-order Butterworth), and then calculated the standard deviation of each filtered signal during the last three minutes of each six-minute trial. We calculated the magnitude of each muscle's total activity as the average of its signal during the last three minutes of each six-minute trial.

We analyzed four additional variables: step width, peak ankle extension angle, total rectus femoris activity, and total biceps femoris activity. We calculated step width by identifying foot contact events and then taking the difference between the lateral centers of pressures—which we determined by dividing the left and right lateral moments by their vertical forces—for consecutive steps. We calculated peak ankle extension angle by first time-locking ankle angle to heel strike events, and then calculating the peak angle during the stride. Similar to our previous analysis of muscle activity, we calculated total rectus femoris activity and total biceps femoris activity by first applying a high-pass filter, rectifying the signals, and then applying a low-pass filter. We next time-locked the signals to heel strike events and normalized each muscle's activity to its average peak activation while walking without the ankle exoskeletons on the same day. We calculated total muscle activity by integrating each muscle's activity during stance at each walking step. Lastly, we quantified the variability and magnitude of each additional variable in the same way as our original four variables. That is, we calculated the variability of each variable by applying a high-pass filter with a cut-off frequency of 0.033 $steps^{-1}$ (third-order Butterworth), and then calculated the standard deviation of this filtered signal during the last three minutes of each six-minute trial. We calculated the magnitude of each variable by averaging its signal during the last three minutes of each six-minute trial. We analyzed these additional variables and reported these additional results in Figures A.1, A.2, and A.3.

Our choice of high-pass filter cut-off frequency makes an assumption about the timescale of changes in variability. We performed a sensitivity analysis of this high-pass filter cut-off

frequency to determine the effect of this assumption. In our main analysis, we used a high-pass filter cut-off frequency of 0.033 steps^{-1} (period of 30 steps). Here, we performed the same analysis but with high-pass filter cut-off frequencies of 0.1 steps^{-1} (period of 10 steps) and 0.02 steps^{-1} (period of 50 steps). This sensitivity analysis revealed that changes in variability over timescales ranging from 10-50 steps do not impact our findings of general variability. That is, across all variables that we analyzed, we observed increases in with-assistance variability on the first day compared to without-assistance baseline, and then decreases in with-assistance variability on the last day compared to the first day. We summarized these results in Table A.1.

We compared variability in the with-assistance condition to variability in the without-assistance condition for each variable. We refer to variability in the without-assistance condition averaged across the two six-minute trials on the first day as *baseline variability*. For most variables, we quantified each participant's baseline from the condition where they walked without the ankle exoskeletons. For ankle angle range, we instead used the condition where participants walked with the ankle exoskeletons but with the devices applying minimal torques as we needed the exoskeleton sensors to calculate ankle angle. We calculated with-assistance variability on the first day by averaging variability across the two six-minute trials where participants walked with the ankle exoskeletons with the predefined control law generating assistive ankle torques at each walking step. For each variable, we used a one-tailed paired Student's t-test to determine whether with-assistance variability on the first day was higher than baseline variability. In all statistical analyses, we used a significance level of 0.05.

We determined how participants modified with-assistance variability along each variable as they gained experience walking with exoskeleton assistance. We calculated with-assistance variability on the last day by averaging variability across the two six-minute trials. For each variable, we used a one-tailed paired Student's t-test to determine whether with-assistance variability on the last day was lower than the first day. We used a two-tailed paired Student's t-test to determine if with-assistance variability on the last day had converged on baseline variability. Lastly, we normalized each participant's with-assistance variability in each six-minute trial to their baseline variability.

We determined how participants adapted the magnitude of each variable as they gained experience walking with exoskeleton assistance. We calculated the magnitude of each variable on the first day and the last day by averaging magnitude across the two six-minute with-assistance trials on each day. We used a two-tailed paired Student's t-test to determine whether participants adapted the magnitude of each variable on the last day compared to the first day. We normalized the magnitude of each variable in each six-minute with-assistance trial to each participant's baseline value. We calculated the baseline value by averaging the magnitude of each variable across the two six-minute without-assistance trials on the first day.

We estimated the slope of the relationship between a variable and energetic cost. We used each variable's normalized values during all with-assistance trials and their respective metabolic costs. For each six-minute trial, we calculated steady-state metabolic cost by averaging net metabolic power during the final three minutes. We normalized metabolic cost during all with-assistance trials to each participant's metabolic cost during without-assistance trials on the first day, which we calculated by averaging metabolic cost over the two six-minute trials of walking with ankle exoskeletons applying minimal torques. We used a linear mixed-effects regression model to estimate participants' shared relationship between a given variable and energetic cost, while allowing for individual differences in their energetic cost intercepts. When plotting this linear model with individual participant data, we subtracted each participant's random effects term (offset) from their data to better illustrate the fixed effects term (slope) that was the focus of this analysis. For each variable, we used a Student's t-test to test if the slope of the linear model was different from zero.

We modeled the timing of adaptive changes, as well as the timing of decreases in variability, for each variable as an exponential decrease from an initial value to a final steady-state value. We used nonlinear mixed-effects regression of the form:

$$Y(t) = a \times e^{\frac{-t}{\tau}} + b \tag{4.2}$$

where t is the experience calculated as the total amount of walking time with exoskeleton assistance and $Y(t)$ is the model output, which is the magnitude or variability of a given variable. We determined the total amount of walking time with exoskeleton assistance as the time spent walking in the with-assistance trials that we analyzed, as well as the time spent walking in the additional trials where participants experienced assistive ankle torques in

human-in-the-loop optimization. We estimated the time constant (τ), amplitude (a), and offset (b) model parameters using nonlinear optimization. We used a mixed-effects model to estimate a single time constant (τ) that is shared between participants while allowing for individual participant offsets (b). When plotting this exponential model with individual participant data, we subtracted each participant's random effects term (offset) from their data to better illustrate the fixed effects term (time constant) that was the focus of this analysis.

We compared time constants between variables. First, we used bootstrapping to estimate the dispersion of each time constant [154,155]. We used the model output from Equation 4.2 to calculate each participant's residuals as the difference between their data points and the model output at these time points. We sampled from each participant's residuals with replacement, and then added each participant's residuals to the model output at these time points to simulate 10 new participants. We fit the exponential model to 10 new participants to simulate a new experiment and estimate a new time constant, and then repeated this process 10000 times for the time constants of adaptation, as well as for the time constants of variability, for each variable. For each time constant, we report the median and interquartile range (IQR), calculated as the difference between 75th and 25th percentiles. We did not observe adaptation in ankle angle range during stance and therefore excluded this variable from this analysis. Next, we tested for differences in time constants of adaptation between step frequency, total soleus activity, and total medial gastrocnemius activity. We used a Kruskal–Wallis one-way ANOVA to test for differences in time constants of adaptation between variables—the bootstrapped time constants did not follow a normal distribution (Anderson–Darling test; $p = 5.0 \times 10^{-4}$)—and then performed a multiple comparison test (Dunn–Šidák correction) of the time constants. We repeated the same analysis but for time constants of variability. We report p-values for total soleus activity versus total medial gastrocnemius activity, total soleus activity versus step frequency, and total medial gastrocnemius activity versus step frequency.

Chapter 5.

Discussion

5.1. Summary

In this thesis, I used energy optimization in walking to investigate how the nervous system learns new optimal control policies in new contexts. I selected this model system because metabolic cost of walking is one of the nervous system's major objectives that I could both directly measure and manipulate. **Chapter 1** and **Chapter 2** used the same experimental setup to study energy optimization in the face of increasingly complex problems. In **Chapter 2**, I tested the generality of previous findings from our lab—that people can continuously optimize step frequency in response to new energetic cost landscapes. I designed and implemented a device that creates new energetic cost landscapes by applying energetic penalties in the form of walking incline as a function of people's real-time measured step width. This device applied a different form of energetic penalties and studied a different gait parameter than that in previous work [42,84]. I found that, like step frequency, people can continuously optimize step width, demonstrating that energy optimization is one of the general principles in the nervous system's control of walking. In **Chapter 3**, I tested how energy optimization extends to multiple gait parameters. Given that people can learn to adapt step frequency and step width when the energy optimum is shifted along either parameter, I next used this device to shift the energy optimum along both parameters. I found that people increase variability and initiate adaptation along step frequency but not step width, suggesting that the nervous system searches within a low dimensional space and that this search space limits energy optimization.

In **Chapter 4**, I used two different experimental setups to ask how the nervous system learns which ways to optimize. I tested the hypothesis that people explore through variations in their control policy to learn more optimal policies in new contexts. I created new contexts using exoskeletons that apply assistive torques to each ankle at each walking step. I found that people initially explore by increasing variability across many variables that span the levels of the whole movement, the joint, and the muscle. With experience, they learn to adapt some variables and decrease variability along of these variables. Variability decreases quickly for quickly-adapting variables, suggesting a reduced space in which the nervous system

optimizes its control policy. People in this way learn to exploit a new control policy that reduces energetic cost. These findings support the principle that general variability leads to specific adaptation toward optimal movement policies. I next asked whether these findings extend to different movement contexts such as split-belt walking. I found that people similarly increase and then decrease variability as they learn to take advantage of the external assistance from the split-belt treadmill and reduce energetic cost. Overall, the findings of this thesis provide insight into the algorithms by which the nervous system learns new optimal control policies, as well as the limitations of these algorithms.

How the nervous system searches through and evaluates control policies can influence how it learns new optimal policies. My second study built on my first study to test how the nervous system represents its search space of control policies. My first study showed that people can learn to adapt their control policy when the energy optimum is shifted along step width. My second study showed that, when presented with shifts in both step width and step frequency, people spontaneously increased variability and initiated adaptation along step frequency but not step width. That changes were limited to one gait parameter suggests that the nervous system's search space is low dimensional—the nervous system recognizes one parameter that needs to be learned and only attempts to adapt its policy in this way. While such a low dimensional search space can accelerate learning along this parameter, it can also restrict it—the nervous system may not recognize all that needs to be learned in new contexts. How does the nervous system identify which ways to adapt its policy? My third study built on my second study to test how the nervous system searches the space of control policies to learn which ways to adapt its policy. I found that people spontaneously explore through increases in variability across many aspects of gait, and then learn to adapt their control policy along specific aspects. Variability decreased as adaptation progressed, and this was rapid for some aspects, suggesting that the nervous system reduces its control policy search space over time. The findings of my third study differ from the findings of my second study. In my third study, the nervous system appears to adopt a high dimensional search space in order to learn a low dimensional search space, whereas in my second study, it appears to adopt a low dimensional search space that results in limited energy optimization. These differences may be due to many factors which do not act in isolation. One factor may be whether the nervous system recognizes a change in context and how it interprets this new context. Another factor may be the shape of the new energetic cost landscape in this new context. There is opportunity for future work to identify conditions in which the nervous system refines its search space for

complete optimization. This is relevant to other situations in which there is incomplete optimization or even lack of initiation of optimization, such as learning to walk with assistive devices or adapt to opposite visuomotor mappings with interference [156]. Together, these findings provide insight into the nature of the nervous system's search space.

5.2. Limitations

While these three studies sought to rule out alternative hypotheses in their systematic design, they are not without limitations. It is important to note that the limitations of these studies are common to many studies in neuroscience that attempt to gain a more mechanistic view of the nervous system [60,157]. One such limitation is that correlations between observed behaviour and outcome variables do not prove causation. An example of this limitation is in studying energetic cost optimization—it is possible that the nervous system optimizes some other objective that happens to correlate with behaviour in the same way as energetic cost. We can carefully design the new cost landscape to reduce the contribution of these other objectives—for example, energetic penalties should not be destabilizing. And we can measure the new cost landscape to determine where the new energy cost optimum lies relative to the observed behaviour. But we cannot rule out the alternative hypothesis that this behaviour is due to the nervous system optimizing some other objective. However, we can consider the rationale for optimizing energetic cost and the evidence that supports it—an approach used in other studies that came before this thesis [42,84]. First, there is an evolutionary incentive to conserve energy—for our human ancestors, food was scarce and movement was required for hunting and gathering [9]. Second, a well-established objective of the nervous system is optimizing energetic cost—several studies have demonstrated that people prefer to walk with the combination of step widths, step lengths, and speeds that optimizes energetic cost [7,10,80,106], and that they continuously optimize this cost [42,84,86]. It is useful to consider these factors when developing hypotheses as well as when interpreting results.

There are additional limitations that are ubiquitous in human behavioural experiments. For example, we cannot directly measure from the nervous system and thus we cannot determine the intent of the nervous system. This limitation is relevant to the third study of my thesis, which asks whether motor variability reflects intentional exploration for adaptation. While I found that changes in variability progressed over similar timescales as adaptation, it is also possible that the nervous system uses some other process to adapt, and that this process

happens to occur at the same time as the observed changes in variability. As when reasoning whether energetic cost is one of the nervous system's objectives for movement, I consider the theory and the evidence that supports variability as intentional exploration. In reinforcement learning theory, a key concept is that exploratory variability can improve learning of new optimal control policies [53]. In humans and other animals, motor variability appears to be regulated in such a way that enables motor learning [64,131]. In this thesis, I not only tested my hypotheses in multiple aspects of gait, but also generalized these findings across multiple experimental paradigms. These findings provide evidence supported by strong rationale and theory.

5.3. Implications and future directions

The findings of this thesis are of fundamental importance and have broad applications. The fundamental importance lies in the better understanding the nervous system's objective for walking and how it optimizes this objective. These findings provide insight into the capability as well as the limitations of the nervous system's optimization algorithms. The process by which the nervous system optimizes energetic cost can provide insight into how it might overcome such limitations. For example, the final study of this thesis demonstrates that the nervous system can use general exploration to learn specific aspects of walking to optimize. The broad applications extend from training elite athletes to rehabilitating stroke patients, and from designing autonomous robots to controlling assistive devices. For elite athletes, they devote their lifetime to mastering the motor skills associated with their sport, which is typified by a decrease in motor variability [131]. Coaches can monitor this variability to determine the onset and termination of learning, which is useful in identifying at which point athletes should transition to learning new skills to maximize their high capacity for training. Coaches can also routinely encourage athletes to increase general variability in search of new optimal solutions—a slight change in context might result in a new cost landscape that shifts the optimum. For stroke patients, they may experience an entirely new cost landscape that presents challenges for identifying which ways to optimize. In this case, therapists should attempt to measure this new cost landscape, and then use bio-feedback to reduce variability along aspects of gait that need not be adjusted, narrowing the nervous system's search space. For designing walking robots and assistive exoskeletons, we must understand the nervous system's algorithms for movement to emulate or assist it. Lastly, and more speculatively, how our nervous system performs optimization may inspire new approaches to how we optimize

businesses, school systems, or even our personal life—we may benefit from purposefully exploring many aspects of these situations to learn which aspects to optimize. After all, our nervous system appears to solve problems of even greater complexity in how it optimizes movement coordination.

There are situations in everyday walking that can alter the cost landscape and shift the energy optimal gait. In my first study, for example, people learned to adapt their control policy when the energy optimum was shifted along step width. In order to realize situations in which this can occur in everyday walking, we must first consider the determinants of the energetic cost of preferred step width [80]. The energetic cost landscape as it relates to step width has large contributions from step-to-step transition costs and lateral limb swing costs. Transition costs dominate at wide widths and push the optimal step width to narrow step widths [96], while swing costs dominate at narrow widths and push the optimal step width to step wider widths [98]. Changes in the shape or magnitude of either of these costs can change the cost landscape and shift the energy optimal width. One real world situation that can change both of these costs is the shoe type. The compliance and damping in the midsole of a shoe may change the shape of the transition cost contribution, whereas the width and mass of a shoe may change the limb swing cost contribution. Thus, the optimal step width may change with each shoe change. Another real-world situation comes not from changing the cost landscape, but from constraining the allowable step widths. For example, stepping stones allow some flexibility in step width, but not all widths will result in stepping on the stones. If this constraint doesn't allow people to walk with their previously optimal width, there will be a new energetically optimal width. This same argument can be made for many other aspects of gait, and future studies should seek to further generalize energy optimization in walking.

Adaptation is needed to recover from changes to our body due to injury or disease. A better understanding of adaptation in healthy individuals may be used to inform training protocols that facilitate adaptation in impaired individuals. The findings of this thesis suggest that the search space of control policies is crucial for learning new energy optimal policies. Thus, future work should consider the nervous system's desire for energy optimal policies and examine the search space of patients that do not initiate gait adaptation or exhibit incomplete adaptation—it is still unknown why some patients adapt while others do not. Poststroke survivors typically exhibit gait abnormalities known as ataxia. A well-established paradigm for rehabilitation is split-belt treadmill training, where they learn to improve gait symmetry,

and these improvements not only apply to treadmill walking, but also transfer to natural situations of overground walking [158]. However, there are also 'non-responders' that do not adapt gait symmetry in response to this training [158]. Our finding that, in some cases, the nervous system uses a low dimensional search space that limits its adaptation may provide insight into the lack of initiation of adaptation in these non-responders. Based on our finding that general variability leads to specific adaptation, we might use bio-feedback to initially increase their search space and elicit their initiation of adaptation. Lastly, our finding that variability decreases and then plateaus as adaptation progresses provides a real-time signal that therapists may use to design training protocols by identifying the onset and termination of learning. This is especially important for those with mobility disorders who have a limited capacity for training.

We can begin to improve human-robot systems by understanding the human nervous system. For example, the designs of assistive exoskeletons have been recently improved by drawing inspiration from the control strategies of the user's nervous system—researchers optimize exoskeleton assistance, or adjust control parameters, to reduce the user's metabolic cost in real-time, which is referred to as human-in-the-loop optimization [159,160]. The observed reductions in metabolic cost may be in large part due to the users adapting to the device. A recent study showed that people can learn to further reduce metabolic cost given proper training and sufficient time to adapt [140]. An effective strategy for designing exoskeletons is to facilitate the nervous system's learning of new gaits, and then evaluate people's optimal response to a range of designs. In my third study, I showed that people can learn how to adapt by initially increasing general variability. One avenue of future research will be to develop protocols that intelligently guide people through the new context, encouraging exploration and faster learning. Previous research has shown that certain properties of the environment, such as how consistent or variable it is, can influence how fast people learn [161–164]. One hypothesis is that the nervous system learns faster in more variable environments [161,162]. And while some studies provide evidence of this, others have found that more consistent environments result in faster learning [163,164]. Human-in-the-loop optimization provides a new paradigm to test how the properties of the environment influence the rate of learning—the nervous system may consider changes in exoskeleton control parameters as variability in the environment.

More fundamental questions remain about the sensory and neural circuit mechanisms that the nervous system uses to implement energy optimization. To gain a more unified view of the nervous system's energy optimization, it is useful to analyze the system at three levels by identifying: (1) the objective, (2) the algorithms used to achieve this objective, and (3) the systems used to implement these algorithms [165]. The studies in this thesis were designed to address the first two levels of analysis [73,85]. However, it is still unclear what sensory mechanisms implement energy optimization—the nervous system must be able to sense its objective to optimize it. A previous study tested the role of one sensory signal—blood gas receptors, which are sensitive to oxygen and carbon dioxide—in energy optimization during walking [81]. This study manipulated the concentration of these blood gases as a function of step frequency to simulate a new cost landscape. They found that people did not adapt their step frequency in response to this new cost landscape, suggesting that blood gas receptors are not the primary sensory signal in the nervous system's energy optimization. Another candidate sensory signal that the nervous system may use to estimate energetic cost is group III and IV muscle afferents, which are sensitive to the byproducts of muscle metabolism [82]. Future research should seek to employ more sophisticated techniques to inhibit specific sensory signals [166]. The nervous system may also need to combine many sensory signals to form a proxy estimate of energetic cost—there is evidence that the nervous system is capable of this sort of state estimation [30]. It is also unclear what neural circuit mechanisms implement energy optimization. Energy optimization in walking appears to involve reward-based learning [73]. One way to probe neural circuit mechanisms is to conduct experiments similar to those in this thesis but with impaired populations. That is, we can test for deficits in adaptation in individuals with and without impairments to shed light on the signals and brain regions involved in energy optimization. For example, we might test for adaptation in Parkinson's patients, where there is degeneration of dopamine neurons that appear to be fundamental to reward processing in the brain. It will be critical to identify the systems that implement energy optimization algorithms to further understand behaviour, both in healthy and impaired individuals.

5.4. Concluding remarks

This collective work advances our understanding of how humans learn to move optimally. It demonstrates the generality of energy optimization by studying multiple aspects of walking using multiple experimental paradigms. It also demonstrates that energy optimization can be limited when the nervous system adopts a lower dimensional search space. To overcome such limitations, the nervous system can initially adopt a higher dimensional search space to refine its lower dimensional search space. The findings of this thesis provide new insight into the structure and role of motor variability for motor learning. They also have applications in the design of assistive devices, as well as the design of training and rehabilitation protocols. Lastly, optimization is a general and dominant principle of human movement. These findings have the potential to apply to objectives other than energetic cost and movements other than walking.

References

1. Yousif, N., and Diedrichsen, J. (2012). Structural learning in feedforward and feedback control. J. Neurophysiol. *108*, 2373–2382.

2. Kuo, A.D. (2002). The relative roles of feedforward and feedback in the control of rhythmic movements. Motor Control *6*, 129–145.

3. Snaterse, M., Ton, R., Kuo, A.D., and Donelan, J.M. (2011). Distinct fast and slow processes contribute to the selection of preferred step frequency during human walking. J. Appl. Physiol. *110*, 1682–1690.

4. Pagliara, R., Snaterse, M., and Donelan, J.M. (2014). Fast and slow processes underlie the selection of both step frequency and walking speed. J. Exp. Biol. *217*, 2939–2946.

5. Block, H.J., and Bastian, A.J. (2010). Sensory reweighting in targeted reaching: effects of conscious effort, error history, and target salience. J. Neurophysiol. *103*, 206–217.

6. Bauby, C.E., and Kuo, A.D. (2000). Active control of lateral balance in human walking. J. Biomech. *33*, 1433–1440.

7. Ralston, H.J. (1958). Energy-speed relation and optimal speed during level walking. Int. Z. Angew. Physiol. *17*, 277–283.

8. Zarrugh, M.Y., Todd, F.N., and Ralston, H.J. (1974). Optimization of energy expenditure during level walking. Eur. J. Appl. Physiol. Occup. Physiol. *33*, 293–306.

9. Brown, P.J. (1991). Culture and the evolution of obesity. Hum. Nat. *2*, 31–57.

10. Minetti, A.E., Capelli, C., Zamparo, P., Di Prampero, P.E., and Saibene, F. (1995). Effects of stride frequency on mechanical power and energy expenditure of walking. Med. Sci. Sports Exercise *27*, 1194–1202.

11. Kandel, E.R., Schwartz, J.H., Jessell, T.M., Siegelbaum, S., Hudspeth, A.J., and Mack, S. (2000). Principles of neural science (McGraw-hill New York).

12. Dickinson, M.H., Farley, C.T., Full, R.J., Koehl, M.A., Kram, R., and Lehman, S. (2000). How animals move: an integrative view. Science *288*, 100–106.

13. Llewellyn, M.E., Barretto, R.P.J., Delp, S.L., and Schnitzer, M.J. (2008). Minimally invasive high-speed imaging of sarcomere contractile dynamics in mice and humans. Nature *454*, 784–788.

14. Goody, R.S. (2003). The missing link in the muscle cross-bridge cycle. Nat. Struct. Biol. *10*, 773–775.

15. Huxley, H., and Hanson, J. (1954). Changes in the cross-striations of muscle during contraction and stretch and their structural interpretation. Nature *173*, 973–976.

16. McNeill Alexander, R. (2003). Principles of Animal Locomotion (Princeton University Press).

17. Rome, L.C., Funke, R.P., Alexander, R.M., Lutz, G., Aldridge, H., Scott, F., and Freadman, M. (1988). Why animals have different muscle fibre types. Nature *335*, 824–827.

18. Wakeling, J.M., Uehli, K., and Rozitis, A.I. (2006). Muscle fibre recruitment can respond to the mechanics of the muscle contraction. J. R. Soc. Interface *3*, 533–544.

19. Hodson-Tole, E.F., and Wakeling, J.M. (2007). Variations in motor unit recruitment patterns occur within and between muscles in the running rat (Rattus norvegicus). J. Exp. Biol. *210*, 2333–2345.

20. Ting, L.H., and Macpherson, J.M. (2005). A limited set of muscle synergies for force control during a postural task. J. Neurophysiol. *93*, 609–613.

21. Ting, L.H. (2007). Dimensional reduction in sensorimotor systems: a framework for understanding muscle coordination of posture. Prog. Brain Res. *165*, 299–321.

22. Wakeling, J.M., and Horn, T. (2009). Neuromechanics of muscle synergies during cycling. J. Neurophysiol. *101*, 843–854.

23. Safavynia, S.A., and Ting, L.H. (2013). Sensorimotor feedback based on task-relevant error robustly predicts temporal recruitment and multidirectional tuning of muscle synergies. J. Neurophysiol. *109*, 31–45.

24. Kravitz, D.J., Saleem, K.S., Baker, C.I., and Mishkin, M. (2011). A new neural framework for visuospatial processing. Nat. Rev. Neurosci. *12*, 217–230.

25. Tuthill, J.C., and Azim, E. (2018). Proprioception. Curr. Biol. *28*, R194–R203.

26. Franklin, D.W., and Wolpert, D.M. (2011). Computational mechanisms of sensorimotor control. Neuron *72*, 425–442.

27. Sober, S.J., and Sabes, P.N. (2005). Flexible strategies for sensory integration during motor planning. Nat. Neurosci. *8*, 490–497.

28. Kalman, R.E. (1960). A New Approach to Linear Filtering and Prediction Problems. J. Basic Eng *82*, 35–45.

29. Wolpert, D.M. (2007). Probabilistic models in human sensorimotor control. Hum. Mov. Sci. *26*, 511–524.

30. Wolpert, D.M., Ghahramani, Z., and Jordan, M.I. (1995). An internal model for sensorimotor integration. Science *269*, 1880–1882.

31. Kuo, A.D. (2005). An optimal state estimation model of sensory integration in human postural balance. J. Neural Eng. *2*, S235-49.

32. McMahon, T.A. (1984). Muscles, Reflexes, and Locomotion (Princeton University Press).

33. Sherrington, C.S. (1910). Flexion-reflex of the limb, crossed extension-reflex, and reflex stepping and standing. J. Physiol. *40*, 28–121.

34. Pruszynski, J.A., Kurtzer, I., and Scott, S.H. (2011). The long-latency reflex is composed of at least two functionally independent processes. J. Neurophysiol. *106*, 449–459.

35. Wong, J.D., and Donelan, J.M. (2017). Principles of energetics and stability in human locomotion. In Humanoid Robotics: A Reference (Dordrecht: Springer).

36. Bastian, A.J. (2008). Understanding sensorimotor adaptation and learning for rehabilitation. Curr. Opin. Neurol. *21*, 628–633.

37. Wong, A.L., and Shelhamer, M. (2011). Saccade adaptation improves in response to a gradually introduced stimulus perturbation. Neurosci. Lett. *500*, 207–211.

38. Shadmehr, R., and Mussa-Ivaldi, F.A. (1994). Adaptive representation of dynamics during learning of a motor task. J. Neurosci. *14*, 3208–3224.

39. Shadmehr, R., Smith, M.A., and Krakauer, J.W. (2010). Error correction, sensory prediction, and adaptation in motor control. Annu. Rev. Neurosci. *33*, 89–108.

40. Morton, S.M., and Bastian, A.J. (2006). Cerebellar contributions to locomotor adaptations during splitbelt treadmill walking. J. Neurosci. *26*, 9107–9116.

41. Emken, J.L., Benitez, R., Sideris, A., Bobrow, J.E., and Reinkensmeyer, D.J. (2007). Motor adaptation as a greedy optimization of error and effort. J. Neurophysiol. *97*, 3997–4006.

42. Selinger, J.C., O'Connor, S.M., Wong, J.D., and Donelan, J.M. (2015). Humans Can Continuously Optimize Energetic Cost during Walking. Curr. Biol. *25*, 2452–2456.

43. Donelan, J.M. (2016). Motor Control: No Constant but Change. Curr. Biol. *26*, R915–R918.

44. Jordan, M.I., and Rumelhart, D.E. (1992). Forward models: Supervised learning with a distal teacher. Cogn. Sci. *16*, 307–354.

45. Shadmehr, R., and Krakauer, J.W. (2008). A computational neuroanatomy for motor control. Exp. Brain Res. *185*, 359–381.

46. Hadjiosif, A.M., Krakauer, J.W., and Haith, A.M. (2021). Did we get sensorimotor adaptation wrong? Implicit adaptation as direct policy updating rather than forward-model-based learning. Journal of Neuroscience *41*, 2747–2761.

47. von Helmholtz, H. (1867). Handbuch der physiologischen Optik (Voss).

48. Martin, T.A., Keating, J.G., Goodkin, H.P., Bastian, A.J., and Thach, W.T. (1996). Throwing while looking through prisms: II. Specificity and storage of multiple gaze—throw calibrations. Brain *119*, 1199–1211.

49. Finley, J.M., Bastian, A.J., and Gottschall, J.S. (2013). Learning to be economical: the energy cost of walking tracks motor adaptation. J. Physiol. *591*, 1081–1095.

50. Finley, J.M., Long, A., Bastian, A.J., and Torres-Oviedo, G. (2015). Spatial and Temporal Control Contribute to Step Length Asymmetry During Split-Belt Adaptation and Hemiparetic Gait. Neurorehabil. Neural Repair *29*, 786–795.

51. Sánchez, N., Simha, S.N., Donelan, J.M., and Finley, J.M. (2019). Taking advantage of external mechanical work to reduce metabolic cost: the mechanics and energetics of split-belt treadmill walking. J. Physiol. *597*, 4053–4068.

52. Sánchez, N., Simha, S.N., Donelan, J.M., and Finley, J.M. (2021). Using asymmetry to your advantage: learning to acquire and accept external assistance during prolonged split-belt walking. J. Neurophysiol. *125*, 344–357.

53. Sutton, R.S., and Barto, A.G. (2018). Reinforcement Learning: An Introduction (MIT Press).

54. Dam, G., Kording, K., and Wei, K. (2013). Credit assignment during movement reinforcement learning. PLoS One *8*, e55352.

55. Izawa, J., and Shadmehr, R. (2011). Learning from sensory and reward prediction errors during motor adaptation. PLoS Comput. Biol. *7*, e1002012.

56. Galea, J.M., Mallia, E., Rothwell, J., and Diedrichsen, J. (2015). The dissociable effects of punishment and reward on motor learning. Nat. Neurosci. *18*, 597–602.

57. Shadmehr, R., Reppert, T.R., Summerside, E.M., Yoon, T., and Ahmed, A.A. (2019). Movement Vigor as a Reflection of Subjective Economic Utility. Trends Neurosci. *42*, 323–336.

58. Zimmet, A.M., Cao, D., Bastian, A.J., and Cowan, N.J. (2020). Cerebellar patients have intact feedback control that can be leveraged to improve reaching. Elife *9*, e53246.

59. Romo, R., and Schultz, W. (1990). Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-initiated arm movements. J. Neurophysiol. *63*, 592–606.

60. Mehler, D.M.A., and Kording, K.P. (2018). The lure of misleading causal statements in functional connectivity research. arXiv.

61. Bernstein, N.A. (1947). On the construction of movements (Pedagogy Publ).

62. Reinbolt, J.A., Seth, A., and Delp, S.L. (2011). Simulation of human movement: applications using OpenSim. Procedia IUTAM *2*, 186–198.

63. Falisse, A., Serrancolí, G., Dembia, C.L., Gillis, J., and De Groote, F. (2019). Algorithmic differentiation improves the computational efficiency of OpenSim-based trajectory optimization of human movement. PLoS One *14*, e0217730.

64. Dhawale, A.K., Smith, M.A., and Ölveczky, B.P. (2017). The Role of Variability in Motor Learning. Annu. Rev. Neurosci. *40*, 479–498.

65. Lansdell, B.J., Prakash, P.R., and Kording, K.P. (2019). Learning to solve the credit assignment problem. arXiv. Available at: http://arxiv.org/abs/1906.00889.

66. Braun, D.A., Aertsen, A., Wolpert, D.M., and Mehring, C. (2009). Motor task variation induces structural learning. Curr. Biol. *19*, 352–357.

67. Bond, K.M., and Taylor, J.A. (2017). Structural Learning in a Visuomotor Adaptation Task Is Explicitly Accessible. eneuro *4*.

68. Diedrichsen, J., White, O., Newman, D., and Lally, N. (2010). Use-dependent and error-based learning of motor behaviors. J. Neurosci. *30*, 5159–5166.

69. Krakauer, J.W., Mazzoni, P., Ghazizadeh, A., Ravindran, R., and Shadmehr, R. (2006). Generalization of motor learning depends on the history of prior action. PLoS Biol. *4*, e316.

70. Todorov, E., and Jordan, M.I. (2002). Optimal feedback control as a theory of motor coordination. Nat. Neurosci. *5*, 1226–1235.

71. Scholz, J.P., and Schöner, G. (1999). The uncontrolled manifold concept: identifying control variables for a functional task. Exp. Brain Res. *126*, 289–306.

72. Peters, J., and Schaal, S. (2008). Reinforcement learning of motor skills with policy gradients. Neural Netw. *21*, 682–697.

73. Selinger, J.C., Wong, J.D., Simha, S.N., and Donelan, J.M. (2019). How humans initiate energy optimization and converge on their optimal gaits. J. Exp. Biol. *222*.

74. Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. Science *275*, 1593–1599.

75. Ganesh, G., Haruno, M., Kawato, M., and Burdet, E. (2010). Motor memory and local minimization of error and effort, not global optimization, determine motor behavior. J. Neurophysiol. *104*, 382–390.

76. Flash, T., and Hogan, N. (1985). The coordination of arm movements: an experimentally confirmed mathematical model. J. Neurosci. *5*, 1688–1703.

77. Wong, J.D., Cluff, T., and Kuo, A.D. (2020). The energetic basis for smooth human arm movements. BioRxiv.

78. Cavagna, G.A., Heglund, N.C., and Taylor, C.R. (1977). Mechanical work in terrestrial locomotion: two basic mechanisms for minimizing energy expenditure. Am. J. Physiol. *233*, R243-61.

79. Margaria, R., Cerretelli, P., Aghemo, P., and Sassi, G. (1963). Energy cost of running. J. Appl. Physiol. *18*, 367–370.

80. Donelan, M.J., Kram, R., and D., K.A. (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*, 1985–1992.

81. Wong, J.D., O'Connor, S.M., Selinger, J.C., and Donelan, J.M. (2017). Contribution of blood oxygen and carbon dioxide sensing to the energetic optimization of human walking. J. Neurophysiol. *118*, 1425–1433.

82. Amann, M. (2012). Significance of Group III and IV muscle afferents for the endurance exercising human. Clin. Exp. Pharmacol. Physiol. *39*, 831–835.

83. Wong, J.D., Selinger, J.C., and Donelan, J.M. (2019). Is natural variability in gait sufficient to initiate spontaneous energy optimization in human walking? J. Neurophysiol. *121*, 1848–1855.

84. Simha, S.N., Wong, J.D., Selinger, J.C., and Donelan, J.M. (2019). A Mechatronic System for Studying Energy Optimization During Walking. IEEE Trans. Neural Syst. Rehabil. Eng. *27*, 1416–1425.

85. Simha, S.N., Wong, J.D., Selinger, J.C., Abram, S.J., and Donelan, J.M. (2021). Increasing the gradient of energetic cost does not initiate adaptation in human walking. J. Neurophysiol. *126*, 440–450.

86. Abram, S.J., Selinger, J.C., and Donelan, J.M. (2019). Energy optimization is a major objective in the real-time control of step width in human walking. J. Biomech. *91*, 85–91.

87. Donelan, J.M.J.M., Shipman, D.W.D.W., Kram, R., and Kuo, A.D.A.D. (2004). Mechanical and metabolic requirements for active lateral stabilization in human walking. J. Biomech. *37*, 827–835.

88. Stimpson, K.H., Heitkamp, L.N., Horne, J.S., and Dean, J.C. (2018). Effects of walking speed on the step-by-step control of step width. J. Biomech. *68*, 78–83.

89. Dean, J.C., Alexander, N.B., and Kuo, A.D. (2007). The effect of lateral stabilization on walking in young and old adults. IEEE Trans. Biomed. Eng. *54*, 1919–1926.

90. McAndrew, P.M., Dingwell, J.B., and Wilken, J.M. (2010). Walking variability during continuous pseudo-random oscillations of the support surface and visual field. J. Biomech. *43*, 1470–1475.

91. O'Connor, S.M., and Kuo, A.D. (2009). Direction-dependent control of balance during walking and standing. J. Neurophysiol. *102*, 1411–1419.

92. Browning, R.C., and Kram, R. (2007). Effects of obesity on the biomechanics of walking at different speeds. Med. Sci. Sports Exerc. *39*, 1632–1641.

93. Owings, T.M., and Grabiner, M.D. (2004). Step width variability, but not step length variability or step time variability, discriminates gait of healthy young and older adults during treadmill locomotion. J. Biomech. *37*, 935–938.

94. Owings, T.M., and Grabiner, M.D. (2004). Variability of step kinematics in young and older adults. Gait Posture *20*, 26–29.

95. Alexander, R.M. (1996). Optima for animals (Princeton University Press).

96. Donelan, J.M., Kram, R., and Kuo, A.D. (2002). Mechanical work for step-to-step transitions is a major determinant of the metabolic cost of human walking. J. Exp. Biol. *205*, 3717–3727.

97. Shipman, D.W., Donelan, J.M., Kram, R., and Kuo, A.D. (2002). Metabolic cost of lateral leg swing in human walking. In World Congress of Biomechanics (American Society of Biomechanics).

98. Shorter, K.A., Wu, A., and Kuo, A.D. (2017). The high cost of swing leg circumduction during human walking. Gait Posture *54*, 265–270.

99. Alexander, R.M. (2001). Design by numbers. Nature *412*, 591.

100. Rodman, P.S., and McHenry, H.M. (1980). Bioenergetics and the origin of hominid bipedalism. Am. J. Phys. Anthropol. *52*, 103–106.

101. Sockol, M.D., Raichlen, D.A., and Pontzer, H. (2007). Chimpanzee locomotor energetics and the origin of human bipedalism. Proc. Natl. Acad. Sci. U. S. A. *104*, 12265–12269.

102. Thompson, N.E., O'Neill, M.C., Holowka, N.B., and Demes, B. (2018). Step width and frontal plane trunk motion in bipedal chimpanzee and human walking. J. Hum. Evol. *125*, 27–37.

103. Huang, H.J., and Ahmed, A.A. (2011). Tradeoff between stability and maneuverability during whole-body movements. PLoS One *6*, e21815.

104. Rogers, M.W., Hedman, L.D., Johnson, M.E., Cain, T.D., and Hanke, T.A. (2001). Lateral stability during forward-induced stepping for dynamic balance recovery in young and older adults. J. Gerontol. A Biol. Sci. Med. Sci. *56*, M589-94.

105. Rogers, M.W., and Mille, M.-L. (2003). Lateral stability and falls in older people. Exerc. Sport Sci. Rev. *31*, 182–187.

106. Umberger, B.R., and Martin, P.E. (2007). Mechanical power and efficiency of level walking with different stride rates. J. Exp. Biol. *210*, 3255–3265.

107. Townsend, M.A. (1985). Biped gait stabilization via foot placement. J. Biomech. *18*, 21–38.

108. Verkerke, G.J., Hof, A.L., Zijlstra, W., Ament, W., and Rakhorst, G. (2005). Determining the centre of pressure during walking and running using an instrumented treadmill. J. Biomech. *38*, 1881–1885.

109. Margaria, R. (1968). Positive and negative work performances and their efficiencies in human locomotion. Int. Z. Angew. Physiol. *25*, 339–351.

110. Gottschall, J.S., and Kram, R. (2005). Energy cost and muscular activity required for leg swing during walking. J. Appl. Physiol. *99*, 23–30.

111. Zeni, J.A., Jr, and Higginson, J.S. (2010). Gait parameters and stride-to-stride variability during familiarization to walking on a split-belt treadmill. Clin. Biomech. *25*, 383–386.

112. Brockway, J.M. (1987). Derivation of formulae used to calculate energy expenditure in man. Hum. Nutr. Clin. Nutr. *41*, 463–471.

113. Drillis, R., Contini, R., and Bluestein, M. (1964). Body segment parameters. Artif. Limbs *8.1*, 44–66.

114. Kawamura, K., Tokuhiro, A., and Takechi, H. (1991). Gait analysis of slope walking: a study on step length, stride width, time factors and deviation in the center of pressure. Acta Med. Okayama *45*, 179–184.

115. Gottschall, J.S., Okorokov, D.Y., Okita, N., and Stern, K.A. (2011). Walking Strategies During the Transition Between Level and Hill Surfaces. Journal of Applied Biomechanics *27*, 355–361.

116. Gates, D.H., Wilken, J.M., Scott, S.J., Sinitski, E.H., and Dingwell, J.B. (2012). Kinematic strategies for walking across a destabilizing rock surface. Gait Posture *35*, 36–42.

117. Voloshina, A.S., Kuo, A.D., Daley, M.A., and Ferris, D.P. (2013). Biomechanics and energetics of walking on uneven terrain. J. Exp. Biol. *216*, 3963–3970.

118. Huang, H.J., Kram, R., and Ahmed, A.A. (2012). Reduction of metabolic cost during motor learning of arm reaching dynamics. J. Neurosci. *32*, 2182–2190.

119. Acasio, J., Wu, M., Fey, N.P., and Gordon, K.E. (2017). Stability-maneuverability trade-offs during lateral steps. Gait Posture *52*, 171–177.

120. Jindrich, D.L., and Qiao, M. (2009). Maneuvers during legged locomotion. Chaos: An Interdisciplinary Journal of Nonlinear Science *19*, 026105.

121. Patla, A.E., Prentice, S.D., Robinson, C., and Neufeld, J. (1991). Visual control of locomotion: Strategies for changing direction and for going over obstacles. Journal of Experimental Psychology: Human Perception and Performance *17*, 603–634.

122. Wu, M., Matsubara, J.H., and Gordon, K.E. (2015). General and Specific Strategies Used to Facilitate Locomotor Maneuvers. PLoS One *10*, e0132707.

123. Bellman, R. (1957). Dynamic programming (Princeton University Press).

124. Selinger, J.C., and Donelan, J.M. (2014). Estimating instantaneous energetic cost during non-steady-state gait. J. Appl. Physiol. *117*, 1406–1415.

125. Collins, S.H., and Kuo, A.D. (2013). Two independent contributions to step variability during over-ground human walking. PLoS One *8*, e73597.

126. Seber, G.A.F., and Wild, C.J. (2003). Nonlinear regression (New Jersey: John Wiley & Sons).

127. McAllister, M.J., Blair, R.L., Maxwell Donelan, J., and Selinger, J.C. (2021). Energy optimization during walking involves implicit processing. J. Exp. Biol. *224*.

128. Silder, A., Besier, T., and Delp, S.L. (2012). Predicting the metabolic cost of incline walking from muscle activity and walking mechanics. J. Biomech. *45*, 1842–1849.

129. Ivry, R.B. (1996). The representation of temporal information in perception and motor control. Curr. Opin. Neurobiol. *6*, 851–857.

130. Ivry, R.B., and Schlerf, J.E. (2008). Dedicated and intrinsic models of time perception. Trends Cogn. Sci. *12*, 273–280.

131. Sternad, D. (2018). It's Not (Only) the Mean that Matters: Variability, Noise and Exploration in Skill Learning. Curr Opin Behav Sci *20*, 183–195.

132. Wu, H.G., Miyamoto, Y.R., Castro, L.N.G., Ölveczky, B.P., and Smith, M.A. (2014). Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. Nat. Neurosci. *17*, 312–321.

133. Bertram, J.E.A. (2005). Constrained optimization in human walking: cost minimization and gait plasticity. J. Exp. Biol. *208*, 979–991.

134. Ortega, J.D., and Farley, C.T. (2007). Individual limb work does not explain the greater metabolic cost of walking in elderly adults. J. Appl. Physiol. *102*, 2266–2273.

135. Kuebrich, B.D., and Sober, S.J. (2015). Variations on a theme: Songbirds, variability, and sensorimotor error correction. Neuroscience *296*, 48–54.

136. Sober, S.J., and Brainard, M.S. (2012). Vocal learning is constrained by the statistics of sensorimotor experience. Proc. Natl. Acad. Sci. U. S. A. *109*, 21099–21103.

137. Tumer, E.C., and Brainard, M.S. (2007). Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. Nature *450*, 1240–1244.

138. Tachibana, R.O., Xu, M., Hashimoto, R.-I., Homae, F., and Okanoya, K. (2020). Spontaneous variability predicts adaptive motor response in vocal pitch control. bioRxiv.

139. Niv, Y., Daniel, R., Geana, A., Gershman, S.J., Leong, Y.C., Radulescu, A., and Wilson, R.C. (2015). Reinforcement learning in multidimensional environments relies on attention mechanisms. J. Neurosci. *35*, 8145–8157.

140. Poggensee, K.L., and Collins, S.H. (2021). How adaptation, training, and customization contribute to benefits from exoskeleton assistance. Sci Robot *6*, eabf1078.

141. Gordon, K.E., Kinnaird, C.R., and Ferris, D.P. (2013). Locomotor adaptation to a soleus EMG-controlled antagonistic exoskeleton. J. Neurophysiol. *109*, 1804–1814.

142. Gordon, K.E., and Ferris, D.P. (2007). Learning to walk with a robotic ankle exoskeleton. J. Biomech. *40*, 2636–2644.

143. Kao, P.-C., Lewis, C.L., and Ferris, D.P. (2010). Invariant ankle moment patterns when walking with and without a robotic ankle exoskeleton. J. Biomech. *43*, 203–209.

144. Sawicki, G.S., and Ferris, D.P. (2008). Mechanics and energetics of level walking with powered ankle exoskeletons. J. Exp. Biol. *211*, 1402–1413.

145. Sawicki, G.S., and Ferris, D.P. (2009). Powered ankle exoskeletons reveal the metabolic cost of plantar flexor mechanical work during walking with longer steps at constant step frequency. J. Exp. Biol. *212*, 21–31.

146. Seethapathi, N., Clark, B., and Srinivasan, M. (2021). Exploration-based learning of a step to step controller predicts locomotor adaptation. bioRxiv.

147. Steele, K.M., Jackson, R.W., Shuman, B.R., and Collins, S.H. (2017). Muscle recruitment and coordination with an ankle exoskeleton. J. Biomech. *59*, 50–58.

148. Churchland, M.M., Cunningham, J.P., Kaufman, M.T., Foster, J.D., Nuyujukian, P., Ryu, S.I., and Shenoy, K.V. (2012). Neural population dynamics during reaching. Nature *487*, 51–56.

149. Cunningham, J.P., and Yu, B.M. (2014). Dimensionality reduction for large-scale neural recordings. Nat. Neurosci. *17*, 1500–1509.

150. Müller, H., and Sternad, D. (2004). Decomposition of variability in the execution of goal-oriented tasks: three components of skill improvement. J. Exp. Psychol. Hum. Percept. Perform. *30*, 212–233.

151. Witte, K.A., and Collins, S.H. (2020). Chapter 13 - Design of Lower-Limb Exoskeletons and Emulator Systems. In Wearable Robotics, J. Rosen and P. W. Ferguson, eds. (Academic Press), pp. 251–274.

152. Zhang, J., Cheah, C.C., and Collins, S.H. (2015). Experimental comparison of torque control methods on an ankle exoskeleton during human walking. In 2015 IEEE International Conference on Robotics and Automation (ICRA), pp. 5584–5589.

153. Jackson, R.W., and Collins, S.H. (2015). An experimental comparison of the relative benefits of work and torque assistance in ankle exoskeletons. J. Appl. Physiol. *119*, 541–557.

154. Duhamel, A., Bourriez, J.L., Devos, P., Krystkowiak, P., Destée, A., Derambure, P., and Defebvre, L. (2004). Statistical tools for clinical gait analysis. Gait Posture *20*, 204–212.

155. Pataky, T.C., Vanrenterghem, J., and Robinson, M.A. (2015). Zero- vs. one-dimensional, parametric vs. non-parametric, and confidence interval vs. hypothesis testing procedures in one-dimensional biomechanical trajectory analysis. Journal of Biomechanics *48*, 1277–1285.

156. Krakauer, J.W., Ghez, C., and Ghilardi, M.F. (2005). Adaptation to visuomotor transformations: consolidation, interference, and forgetting. J. Neurosci. *25*, 473–478.

157. Stanford, K. (2017). Underdetermination of Scientific Theory. The Stanford Encyclopedia of Philosophy.

158. Reisman, D.S., McLean, H., Keller, J., Danks, K.A., and Bastian, A.J. (2013). Repeated split-belt treadmill training improves poststroke step length asymmetry. Neurorehabil. Neural Repair *27*, 460–468.

159. Felt, W., Selinger, J.C., Donelan, J.M., and Remy, C.D. (2015). "Body-In-The-Loop": Optimizing Device Parameters Using Measures of Instantaneous Energetic Cost. PLoS One *10*, e0135342.

160. Zhang, J., Fiers, P., Witte, K.A., Jackson, R.W., Poggensee, K.L., Atkeson, C.G., and Collins, S.H. (2017). Human-in-the-loop optimization of exoskeleton assistance during walking. Science *356*, 1280–1284.

161. Burge, J., Ernst, M.O., and Banks, M.S. (2008). The statistical determinants of adaptation rate in human reaching. Journal of Vision *8*, 20.

162. Wei, K., and Körding, K. (2010). Uncertainty of feedback and state estimation determines the speed of motor adaptation. Front. Comput. Neurosci. *4*, 11.

163. Gonzalez Castro, L.N., Hadjiosif, A.M., Hemphill, M.A., and Smith, M.A. (2014). Environmental Consistency Determines the Rate of Motor Adaptation. Curr. Biol. *24*, 1050–1061.

164. van Beers, R.J. (2012). How does our motor system determine its learning rate? PLoS One *7*, e49373.

165. Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information. San Francisco, CA: W H Freeman.

166. Grant, G.J., Susser, L., Cascio, M., Moses, M., and Zakowski, M.I. (1996). Hemodynamic effects of intrathecal fentanyl in nonlaboring term parturients. J. Clin. Anesth. *8*, 99–103.

Appendix.

Supplementary figures

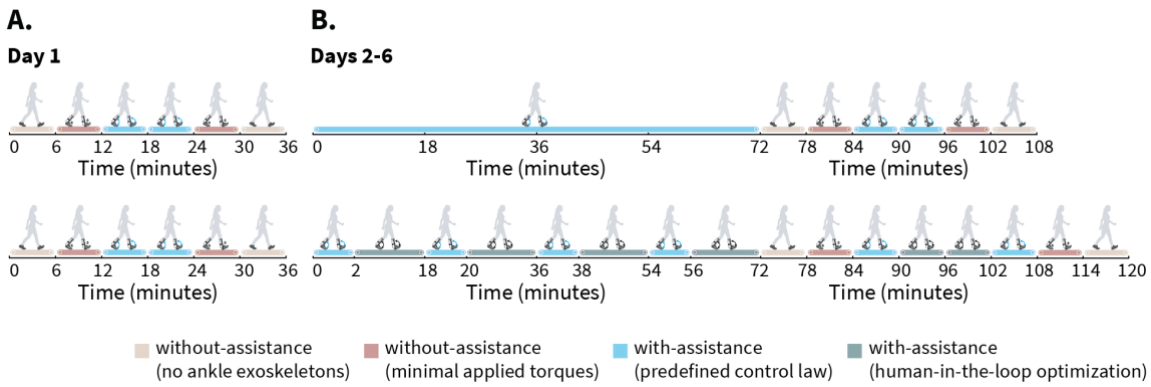

Figure A.1: Detailed experimental protocol
(A) On day 1, all participants completed 3 conditions twice: walking without the ankle exoskeletons, walking while wearing the ankle exoskeletons but with minimal applied torques, and walking while wearing the ankle exoskeletons which used a predefined control law to generate assistive ankle torques at each walking step. On days 2-6, all participants completed (B) additional trials which were followed by the original three conditions twice. In the additional trials, some participants repeatedly experienced the predefined control law (top row), whereas other participants experienced the predefined control law interspersed between human-in-the-loop optimization of the control law (bottom row).

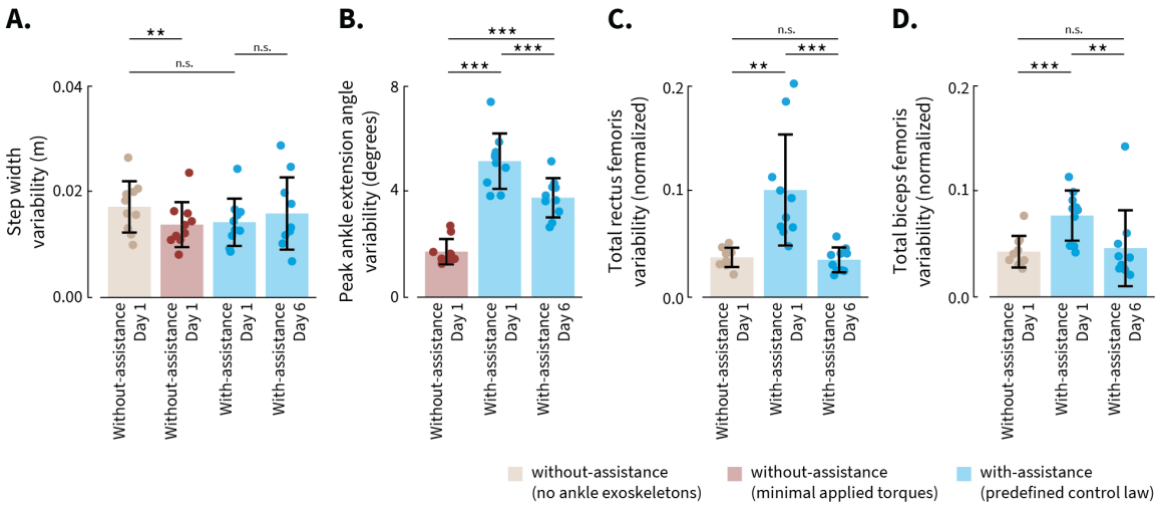

Figure A.2: Changes in variability as participants gain experience with exoskeleton assistance.

(A) Step width variability, (B) peak ankle extension angle variability, (C) total rectus femoris variability, and (D) total biceps femoris variability. Circles are participant averages, bar height is average across participants, error bars represent one standard deviation, and asterisks indicate statistically significant differences between conditions or days using the notation: *** for $p<0.001$, ** for $p<0.01$, * for $p<0.05$, and n.s. for not significant. We found that, for three of these four additional variables, participants increased with-assistance variability on the first day compared to without-assistance baseline variability (mean ± standard deviation, one-tailed paired t-test; peak ankle extension angle: +212.7 ± 84.5%, $p = 1.8 \times 10^{-6}$; total rectus femoris activity: +176.7 ± 144.3%, $p = 1.6 \times 10^{-3}$; total biceps femoris activity: +93.5 ± 70.6%, $p = 7.2 \times 10^{-4}$). Participants then decreased with-assistance variability on the last day compared to the first day (mean ± standard deviation, one-tailed paired t-test; peak ankle extension angle: -25.4 ± 15.5%, $p = 9.5 \times 10^{-4}$; total rectus femoris activity: -61.7 ± 11.3%, $p = 5.3 \times 10^{-4}$; total biceps femoris activity: -42.5 ± 27.4%, $p = 2.4 \times 10^{-3}$). We suspect that the explanation for why step width variability did not initially increase with ankle exoskeleton assistance compared to without the ankle exoskeletons (-16.5 ± 12.5%, one-tailed paired t-test: $p = 0.99$), and therefore also did not decrease with experience (+11.3 ± 34.7%, one-tailed paired t-test: $p = 0.86$), was that the added mass of the ankle exoskeletons introduced a large energetic penalty to step width variability. Indeed, step width variability decreased when walking with the ankle exoskeletons with minimal applied torques compared to walking without the ankle exoskeletons (-19.1 ± 10.5%, two-tailed paired t-test: $p = 1.2 \times 10^{-3}$).

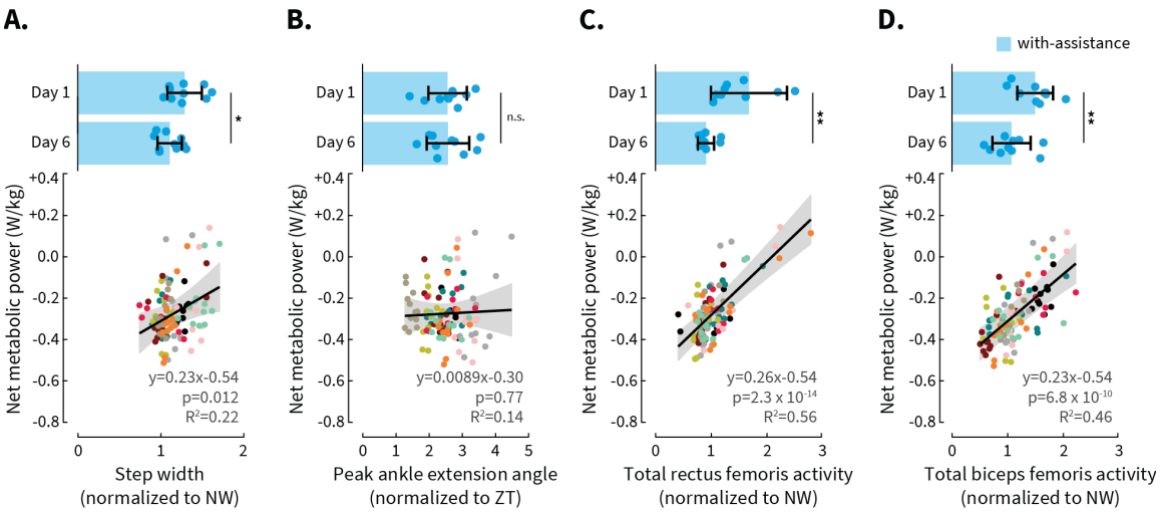

Figure A.3: Changes in magnitude of variables that reduce energetic cost

Estimated relationships between energetic cost and (A) step width, (B) peak ankle extension angle, (C) total rectus femoris activity, and (D) total biceps femoris activity during with-assistance trials across all days. We normalized energetic cost during all with-assistance trials to each participant's energetic cost during without-assistance trials—walking with ankle exoskeletons with minimal applied torques—on the first day. We normalized each variable's magnitude with-assistance to each participant's baseline value without-assistance. The solid black lines are linear mixed-effects models with 95% confidence intervals (grey shading). Individual participants are represented by distinct colours with their random-effects intercept term subtracted from their energetic cost data. We found relatively strong and significant relationships between energetic cost and step width (slope = 0.23, 95% CI [0.053, 0.41], p = 0.012), total rectus femoris activity (slope = 0.26, 95% CI [0.20, 0.31], p =2.3 x 10^{-14}), and total biceps femoris activity (slope = 0.23, 95% CI [0.16, 0.29], p = 6.8 x 10^{-10}). The relationship between peak ankle extension angle and energetic cost was weaker and not significant (slope = 0.0089, 95% CI [-0.051, 0.069], p = 0.77). The bar graph shows the with-assistance magnitude of each variable on the first and last day. Circles are participant averages, bar height is average across participants, error bars represent one standard deviation, and asterisks indicate statistically significant differences between days using the notation: *** for p<0.001, ** for p<0.01, * for p<0.05, and n.s. for not significant. We found that participants learned to adapt variables that were associated with reductions in energetic cost—we observed changes in magnitude between the first and last days of with-assistance trials for step width (mean ± standard deviation, paired t-test; -12.4 ± 15.5%, p = 0.028), total rectus femoris activity (-40.7 ± 16.4%, p = 2.9 x 10^{-3}), and total biceps femoris activity (-28.0 ± 18.2%, p = 1.6 x 10^{-3}), but not for peak ankle extension angle (2.5 ± 22.2%, p = 0.94).

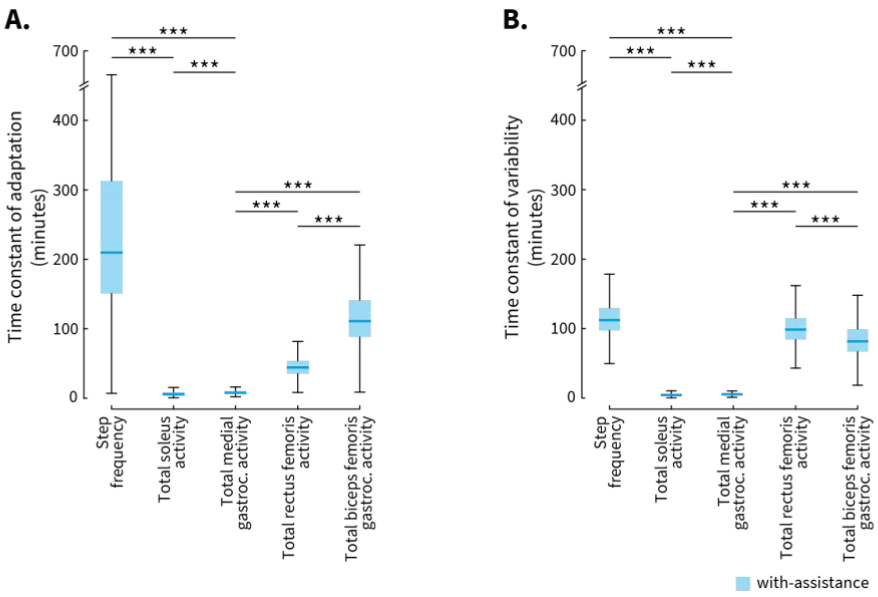

Figure A.4: Differences in timescales between variables

(A) Time constants of adaptation and (B) time constants of variability for three original variables—step frequency, total soleus activity, and total medial gastrocnemius activity—and two additional variables—total rectus femoris activity and total biceps femoris activity. We did not observe changes in variability for step width or changes in magnitude for peak ankle extension angle and therefore excluded these additional variables from this analysis. The central mark indicates the median, the bottom edge of the box indicates the lower quartile (25th percentile), and the top edge of the box indicates the upper quartile (75th percentile). Error bars extend to the most extreme data points not considered outliers, which we define as more than 1.5 times the interquartile range away from the edges of the box. Asterisks indicate statistically significant differences between variables using the notation: *** for $p<0.001$. We found that total rectus femoris activity and total biceps femoris activity had different time constants of adaptation, and that these time constants were different from original variables (ANOVA; total rectus femoris activity vs. total biceps femoris activity: $p < 0.001$; total rectus femoris activity vs. total medial gastrocnemius activity: $p < 0.001$; total biceps femoris activity vs. total medial gastrocnemius activity: $p < 0.001$). We also found similar results for time constants of variability (ANOVA; total rectus femoris activity vs. total biceps femoris activity: $p < 0.001$; total rectus femoris activity vs. total medial gastrocnemius activity: $p < 0.001$; total biceps femoris activity vs. total medial gastrocnemius activity: $p < 0.001$).

Variable	Comparison	10-step cut-off	30-step cut-off	50-step cut-off
Step frequency variability (mean ± SD, paired t-test)	baseline vs. day 1	+57.0 ± 27.2% $p = 1.9 \times 10^{-4}$	+56.2 ± 27.6% $p = 2.0 \times 10^{-4}$	+56.4 ± 28.1% $p = 1.9 \times 10^{-4}$
	day 1 vs. day 6	-39.5 ± 9.7% $p = 3.1 \times 10^{-5}$	-38.5 ± 9.8% $p = 4.9 \times 10^{-5}$	-38.3 ± 10.0% $p = 5.7 \times 10^{-5}$
	baseline vs. day 6	-6.0 ± 18.3% p = 0.31	-5.1 ± 17.3% p = 0.32	-5.0 ± 17.0% p = 0.33
Ankle angle range variability (mean ± SD, paired t-test)	baseline vs. day 1	+295.5 ± 89.0% $p = 8.9 \times 10^{-7}$	+278.8 ± 85.0% $p = 6.8 \times 10^{-7}$	+277.2 ± 81.8 $p = 5.1 \times 10^{-7}$
	day 1 vs. day 6	-29.1 ± 19.7% $p = 3.1 \times 10^{-3}$	-26.3 ± 22.8% $p = 6.5 \times 10^{-3}$	-25.5 ± 24.4% $p = 7.8 \times 10^{-3}$
	baseline vs. day 6	+169.0 ± 53.8% $p = 1.8 \times 10^{-6}$	+166.2 ± 52.5% $p = 6.7 \times 10^{-6}$	+167.5 ± 54.2% $p = 1.2 \times 10^{-5}$
Total soleus variability (mean ± SD, paired t-test)	baseline vs. day 1	+15.4 ± 20.6% p = 0.040	+21.1 ± 25.3% p = 0.026	+22.9 ± 26.8% p = 0.023
	day 1 vs. day 6	-16.0 ± 18.5% p = 0.013	-18.3 ± 21.2% p = 0.013	-19.0 ± 21.2% p = 0.012
	baseline vs. day 6	-4.3 ± 24.3% p = 0.33	-3.3 ± 26.9% p = 0.39	-3.0 ± 27.2% p = 0.40
Total medial gastrocnemius variability (mean ± SD, paired t-test)	baseline vs. day 1	+46.2 ± 21.5% $p = 3.0 \times 10^{-5}$	+46.0 ± 19.6% $p = 1.7 \times 10^{-5}$	+46.5 ± 19.8% $p = 1.9 \times 10^{-5}$
	day 1 vs. day 6	-17.7 ± 14.3% $p = 3.1 \times 10^{-3}$	-19.7 ± 13.2% $p = 1.5 \times 10^{-3}$	-20.5 ± 13.2 $p = 1.3 \times 10^{-3}$
	baseline vs. day 6	+18.6 ± 15.9% p = 0.0073	+15.8 ± 15.8% p = 0.020	+15.0 ± 15.1% p = 0.021

Table A.1: Sensitivity analysis

Results from sensitivity analysis of high-pass filter cut-off frequency. We recomputed variability for each variable with filter cut-off frequencies of 0.1 steps-1 (period of 10 steps), 0.033 steps-1 (period of 30 steps) used in main analysis, and 0.02 steps-1 (period of 50 steps). We recalculated changes in variability between conditions and days—baseline vs. day 1, day 1 vs. day 6, and baseline vs. day 6—and found that statistically significant results remained unchanged.