Journal of Neural Engineering

RECEIVED
1 December 2019

REVISED
8 April 2020

ACCEPTED FOR PUBLICATION
29 April 2020

PUBLISHED
28 May 2020

PAPER

Pavlovian control of intraspinal microstimulation to produce over-ground walking

Ashley N Dalrymple[1,4], **David A Roszko**[1,4], **Richard S Sutton**[2,4] **and Vivian K Mushahwar**[1,3,4,5]

[1] Neuroscience and Mental Health Institute, University of Alberta, Edmonton, AB, Canada
[2] Department of Computing Science, Faculty of Science, University of Alberta, Edmonton, AB, Canada
[3] Division of Physical Medicine and Rehabilitation, Department of Medicine, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada
[4] Sensory Motor Adaptive Rehabilitation Technology (SMART) Network, University of Alberta, Edmonton, AB, Canada
[5] Author to whom any correspondence should be addressed.

E-mail: vivian.mushahwar@ualberta.ca

Keywords: machine learning, neural prosthesis, spinal cord injury, walking, functional electrical stimulation, Pavlovian control, reinforcement learning

Abstract

Objective. Neuromodulation technologies are increasingly used for improving function after neural injury. To achieve a symbiotic relationship between device and user, the device must augment remaining function, and independently adapt to day-to-day changes in function. The goal of this study was to develop predictive control strategies to produce over-ground walking in a model of hemisection spinal cord injury (SCI) using intraspinal microstimulation (ISMS). *Approach.* Eight cats were anaesthetized and placed in a sling over a walkway. The residual function of a hemisection SCI was mimicked by manually moving one hind-limb through the walking cycle. ISMS targeted motor networks in the lumbosacral enlargement to activate muscles in the other, presumably 'paralyzed' limb, using low levels of current (<130 μA). Four people took turns to move the 'intact' limb, generating four different walking styles. Two control strategies, which used ground reaction force and angular velocity information about the manually moved 'intact' limb to control the timing of the transitions of the 'paralyzed' limb through the step cycle, were compared. The first strategy used thresholds on the raw sensor values to initiate transitions. The second strategy used reinforcement learning and Pavlovian control to learn predictions about the sensor values. Thresholds on the predictions were then used to initiate transitions. *Main results.* Both control strategies were able to produce alternating, over-ground walking. Transitions based on raw sensor values required manual tuning of thresholds for each person to produce walking, whereas Pavlovian control did not. Learning occurred quickly during walking: predictions of the sensor signals were learned rapidly, initiating correct transitions after $\leq$4 steps. Pavlovian control was resilient to different walking styles and different cats, and recovered from induced mistakes during walking. *Significance.* This work demonstrates, for the first time, that Pavlovian control can augment remaining function and facilitate personalized walking with minimal tuning requirements.

1. Introduction

After a spinal cord injury (SCI), people experience motor and sensory paralysis to varying degrees, depending on the severity and level of the injury. Two-thirds of all SCIs in the USA are incomplete ('Spinal Cord Injury (SCI) 2017 Facts and Figures at a Glance' 2017), yet 50% of people with SCI never recover the ability to walk again. For people with paraplegia, regaining the ability to walk is of high importance for improving quality of life, with a majority of people with paraplegia ranking the restoration of walking as a first or second priority (Collinger *et al* 2013). Currently, SCI has no cure; therefore, regaining the ability to walk has been pursued through other means such as rehabilitation (Musselman *et al* 2009, Lam *et al* 2015, Morrison *et al* 2018), neural technologies (Kobetic *et al* 1997, Stein and Mushahwar 2005, Hardin *et al* 2007, Moritz *et al* 2008, Holinski *et al* 2016), or a combinatorial

approach (Carhart *et al* 2004, Angeli *et al* 2018, Gill *et al* 2018).

The neural networks in the spinal cord below the SCI and their connections to the muscles remain intact (Hunter and Ashby 1994). These spinal networks can be targeted and activated using electrical stimulation (Mushahwar and Horch 2000b; Saigal *et al* 2004, Tator *et al* 2012, Hofstoetter *et al* 2015, Angeli *et al* 2018, Wagner *et al* 2018, Gill *et al* 2018). One electrical stimulation approach is intraspinal microstimulation (ISMS), in which fine, hair-like microwires are implanted in the ventral horn of the lumbosacral enlargement. Interestingly, stimulation in this region through a single microwire produces large, graded, single joint movements as well as coordinated multi-joint synergies (Mushahwar and Horch 1998, 2000a, 2000b, Mushahwar *et al* 2000, Saigal *et al* 2004, Lau *et al* 2007, Holinski *et al* 2016). Through targeted activation of hind-limb locomotor-related networks, ISMS has been used to restore walking in anaesthetized (Holinski *et al* 2013, 2016) and chronically spinalized cats (Saigal *et al* 2004). Nearly 1 km of over-ground, weight-bearing walking was produced by ISMS in cats (Holinski *et al* 2016). These distances were achieved immediately after implantation of the microdevice and without the need for extensive rehabilitation. Responses produced by ISMS remain consistent throughout the use of the implant (Mushahwar *et al* 2000), and long-term use of ISMS for walking will likely further improve walking distances achieved. Therefore, ISMS is poised to be a viable clinical approach to restoring walking after severe paralysis.

An important and clinically-relevant aspect of a successful neural prosthesis is the control of the device and how users interact with the control strategy. Current commercially available devices for restoring walking after SCI, such as those that use functional electrical stimulation (FES) and various exoskeletons, have limited control options. Walking is accomplished using open loop control with pre-defined timing of the limb movements (Chaplin 1996, Johnston *et al* 2005, Chang *et al* 2015, Ekelem and Goldfarb 2018). The users are expected to adapt their walking to accommodate the control strategy in the device. Control strategies developed for ISMS to date have primarily focused on restoring walking in models of complete SCI (Saigal *et al* 2004, Holinski *et al* 2011, 2016, Dalrymple and Mushahwar 2017). Feedback, such as ground reaction force, hip angle, or activity of sensory neurons from the dorsal root ganglia, were used to modify the inherent timing of the transitions between the phases of the gait cycle (Saigal *et al* 2004, Holinski *et al* 2011, 2013, 2016). A recent paper depicted the first control strategies developed for ISMS in a model of incomplete SCI (Dalrymple *et al* 2018). These strategies augmented the residual function in a model of hemisection SCI and, using supervised machine learning, adapted the control strategy for different speeds of walking. Supervised machine learning has also been used to control surface functional electrical stimulation (FES) systems in persons with SCI (Abbas and Triolo 1997, Popović *et al* 1999); however, few studies have focused on completing an entire walking cycle (Popović 1993). As people with SCI experience varying levels of paralysis, each person would require their own custom stimulation settings to restore walking. Moreover, incomplete injuries evolve over time requiring further updating of stimulation settings. Manual tuning of settings is burdensome; it is time-consuming and based qualitatively on trial and error. A recent machine learning approach demonstrated the feasibility of adaptive tuning of impedance parameters in a prosthetic knee (Wen *et al* 2019); however, to date, adaptive machine learning approaches have not been utilized in implanted neural prosthetic approaches for restoring over-ground walking.

Intuitive control of a neural prosthesis requires the device to know what the user wants to do pre-emptively with automatic adaptation to changes in the environment. Learning predictions of walking-relevant sensor signals for initiating control outputs may be a more reliable method to produce walking. Predictions allow for a timely response that can be modified with experience. In this study, we compared more traditional control methods with a new, prediction-based machine learning control method, called Pavlovian control, to produce over-ground, alternating walking in a model of hemisection SCI. Specifically, we assessed the need for manual tuning of control settings between reaction-based control and Pavlovian (or prediction-based) control over several cat experiments, with different people participating to move one limb through the walking cycle and after perturbations. This presents the first application of Pavlovian control to produce walking. It is also the first known application of reinforcement learning techniques in a spinal neural interface. Using Pavlovian control, we demonstrate that alternating over-ground walking can be achieved quickly using predictions of walking-relevant sensor signals, and that the thresholds for Pavlovian control do not require re-tuning across different conditions.

2. Methods

All experimental procedures were approved by the University of Alberta Animal Care and Use Committee under protocol AUP301. Eight adult male cats (3.96 to 5.22 kg) were individually housed in large cages and were provided with daily enrichment that included a large play pen, toys, human interaction, and soothing music.

2.1. Implant procedure

Investigations were conducted in acute, non-recovery experiments. Anaesthesia was initially induced with

isoflurane (5%), followed by sodium pentobarbital anaesthesia administered intravenously (induction: 25 mg kg^{-1}; maintenance: 1 in 10 dilutions in saline). All surgical procedures and data collection were performed under sodium pentobarbital (Toossi *et al* 2019). A laminectomy was performed to expose the lumbosacral enlargement. An array of 12 microwire electrodes made of Pt–Ir (80%–20%), 50 μm in diameter, insulated with 4 μm polyimide except for approximately 400 μm exposure at the tip, was implanted in one side of the spinal cord according to established procedures (Mushahwar *et al* 2000, Bamford *et al* 2017). The microwire tips targeted lamina IX in the ventral horn based on functional maps of the motoneuron pools (Vanderhorst and Holstege 1997, Mushahwar and Horch 1998, 2000b). In addition to motoneuronal pools, this region contains neural networks that, when stimulated, produce coordinated multi-joint synergistic movements of the leg (Engberg and Lundberg 1969, Mushahwar and Horch 2000b, Saigal *et al* 2004, Holinski *et al* 2016, Bhumbra and Beato 2018).

2.2. Stimulation protocol

Trains of stimuli were delivered using a customized current-controlled stimulator (Sigenics Inc. Chicago, IL, USA) and consisted of a trapezoidal waveform that ramped from threshold to chosen amplitude over 3 time-steps (time-step = 40 ms). The stimulus pulses in a train were 290 μs in duration, biphasic, charge-balanced and delivered at a rate of 50 Hz. Stimulation amplitudes ranged from threshold (<20 μA) to amplitudes that produced weight-bearing movements (60 to 80 μA) and did not exceed 130 μA through any electrode.

The movements elicited by stimulation through single electrodes were hip flexion, hip extension, knee extension, ankle dorsiflexion, ankle plantarflexion, and a backward extensor synergy, which were combined to construct a full walking cycle. Of the 12 electrodes implanted unilaterally, between 5 and 9 were needed to produce the desired walking movements and included redundancy in the functional targets. Stimulation channels were combined to construct the four phases of the walking cycle: F (early swing), E1 (late swing to paw-touch), E2 (mid-stance), and E3 (propulsion) (Engberg and Lundberg 1969, Goslow *et al* 1973). The phases F–E2 and E1–E3 were defined to be opposite phases of the walking cycle. During each phase of the walking cycle, the electrode combination and amplitudes remained constant.

2.3. Experimental setup

Following the implantation of the ISMS array, the anaesthetized cats were transferred to a custom-built instrumented walkway and placed in a sling that supported their trunk, head and forelimbs, while the hind-limbs where left to move freely over the walkway (figure 1) (Guevremont *et al* 2007, Mazurek *et al* 2012, Holinski *et al* 2016). The sling was suspended from a cart that moved with the cat over the walkway. The cart was partially unloaded to offset the weight of the recording and stimulating equipment and the mobile vital signs monitors that were placed on it.

Gyroscopes were placed on the tarsals of each hind-limb to measure angular velocity in real time. Force plates providing ground reaction force (GRF) in three dimensions were mounted underneath the walkway. The sensor signals were filtered using a Butterworth filter (f_C = 3 Hz, 2nd order), digitized at 1 kHz using the Grapevine Neural Interface Processor (Ripple, Salt Lake City, UT, USA) and streamed into Matlab (MathWorks, Inc., Natick, MA, USA) during walking.

Reflective markers were positioned on the iliac crest, hip, knee, ankle, and metatarsophalangeal (MTP) joints of the right hind-limb (moved using ISMS). Kinematics of this limb were recorded using a camera (120 fps, JVC Americas Corp., Wayne, NJ, USA) positioned 4.5 m away from the center of the walkway. Marker positions were tracked post-hoc using MotionTracker2D, a custom Matlab program written by Dr. Douglas Weber (University of Pittsburgh, Pittsburgh, PA, USA).

A hemisection SCI was modeled in the anaesthetized cats by having a person (naïve experimenter) manually move the left hind-limb through the walking cycle (person-moved limb; PML) to represent the 'intact' leg. The right hind-limb was moved using ISMS (stimulation-controlled limb; SCL) and represented the 'paralyzed' leg (figure 1). This hemisection SCI model is similar to Brown-Sequard syndrome in humans, where one leg is paralyzed and the other is motor-intact (Kunam *et al* 2018).

2.4. Control strategies

The goal of the control strategies was to transition the SCL through the walking cycle such that the phase of the SCL was opposite to the phase of the PML and the force produced by the SCL was enough to lift the animal's hindquarters (i.e. provide weight-support) and propel the animal across the walkway to produce over-ground walking. Both control strategies determined when the SCL transitioned from one phase to the next based on sensor information from the PML.

2.4.1. Reaction-based control strategy

For reaction-based control, thresholds were placed on the sensor signals recorded from the PML during walking to trigger transitions between the phases of the walking cycle in the SCL. The sensor signals used for defining the transitions between phases of the walking cycle were vertical GRF (vGRF) and angular velocity of the PML (figure 2(A)). The transitions were controlled by rules involving the current phase in the walking cycle, comparing the sensor values with threshold values, and the direction of the slope of the sensor values. Thresholds were placed on the

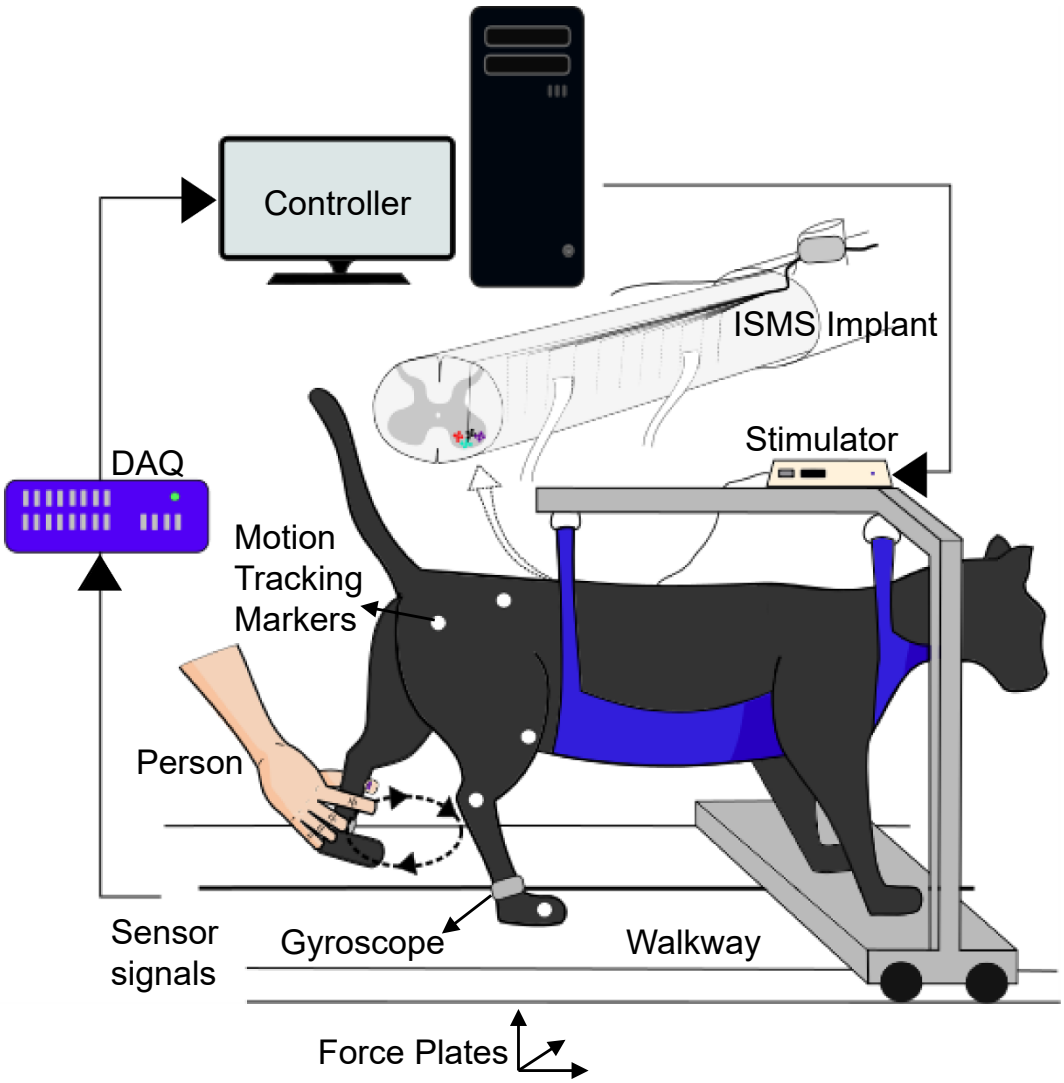

Figure 1. Experimental setup for over-ground walking. A naïve experimenter moved the left hind-limb through the walking cycle. Sensor signals from force plates under the walkway and a gyroscope on the tarsals from both hind-limbs were converted to digital signals by the DAQ and streamed into Matlab. In Matlab, a custom control algorithm was used to control the stimulation to the spinal cord to move the right hind-limb to the opposite phase of the walking cycle.

sensor signals at times that accounted for an electromechanical delay (the delay from the time a stimulus was delivered to the time movement occurred) of approximately 200 ms (Dalrymple *et al* 2018). Because the thresholds were placed on the raw sensor signals, this control strategy was reactive to these raw signals as opposed to predictive of the signals, which was the case in Pavlovian control (see below). Hence, this strategy was referred to as reaction-based control.

2.4.2. Pavlovian control strategy

Automatic prediction of the control output is required to produce personalized walking that augments remaining function after a SCI. We used reinforcement learning methods to predict three walking-relevant signals in real time: the vGRF, angular velocity, and unloading. Unloading was defined as the weight-bearing threshold (equal to 12.5% of the cat's body weight in this setup (Lau *et al* 2007)) minus the vGRF. Unloading differs from vGRF as it informs when the PML is below or above a weight-bearing threshold, which is specific to each animal, and indicates more precisely safe limb loading. Thresholds on the learned predictions, in addition to the slope of the predictions and knowledge of the current phase, were used to control ISMS to transition the SCL to the opposite phase of the PML (figure 2(B)). The use of a learned prediction of a sensory stimulus (reinforcement learning of sensor signals) to trigger a fixed response (ISMS transition to phase of the walking cycle) is referred to as Pavlovian control because it is modelled after Pavlovian conditioning (Modayil and Sutton 2014).

The transitions between the phases of the walking cycle were made either by the prediction of the sensor signals crossing the threshold, or the raw sensor signal crossing a threshold, whichever occurred first. Instances where the raw sensor signals were used to initiate a transition between the phases of the walking

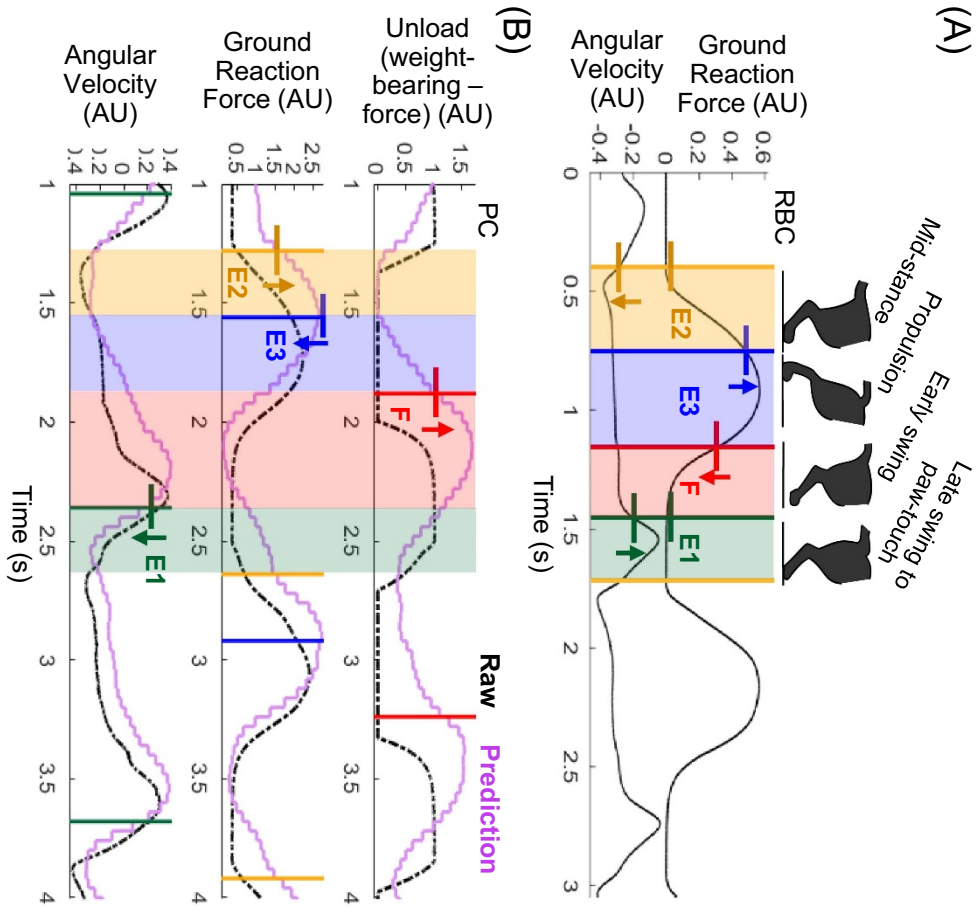

Figure 2. (A) Phases of the walking cycle (E2, E3, F, E1) and thresholds for one naïve experimenter (person A) on raw sensor signals from the PML (person-moved limb) for reaction-based control (RBC). (B) Threshold settings on predictions of sensor signals from the PML using Pavlovian control (PC). AU = arbitrary units due to normalization during acquisition. Shaded regions indicate the phase of the walking cycle detected on the PML, divided by vertical lines indicating timing of transitions. Horizontal lines mark the threshold values for corresponding phase. Arrows indicate the direction of the slope of the signal required by the algorithm.

cycle are referred to as back-up reactions. The relative thresholds for the back-up reactions were held constant throughout all walking trials for a given cat, as they were adjusted based on the weight of each cat.

2.4.3. Machine learning methods for Pavlovian control

2.4.3.1. State representation of sensor signals

Function approximation was used to represent the highly sampled, complex sensor signals as a binary vector representation of the state space, named the feature vector, **x**. Six sensor signals were chosen to form the state space: left ('intact leg') vGRF, right ('paralyzed leg') vGRF, the sum of left and right vGRFs, left angular velocity, right angular velocity, and the exponential moving average of the left vGRF. The exponential moving average gives a long-term history of the force signal and helps differentiate between the periodical increase and decrease of the other sensor signals. To create the feature vector **x,** the sensor values were first normalized to values between 0 and 1. These normalized values were then coded into a binary vector using Selective Kanerva coding (Travnik and Pilarski 2017).

To perform selective Kanerva coding, $K = 5000$ specific states, also referred to as prototypes, were randomly distributed over the entire normalized, 6-dimensional state space (6 sensors; figure 3). The prototype locations were held constant for all experiments. Hoare's quickselect was used to find the c closest prototypes to the current state according to their Euclidean distance (Travnik and Pilarski 2017). Three values of c, determined by choosing small ratios, η, such that $c = K\eta$ were used. The values of c were equal to 500, 125, and 25, corresponding to η values of 0.1, 0.025, and 0.005, respectively. Using multiple c values is similar to the use of overlapping tilings in tile coding (Sutton and Barto 2018); it allows for coarse and fine representation of the state in the feature vector.

During a walking trial, the combination of sensor signals acquired at each time-step corresponded to a state in the state space, termed the current state (Dalrymple 2019). The c-closest prototypes to the current state were activated in the feature vector (i.e. set equal to 1), while the rest were inactivated (i.e. set equal to 0). The total number of features in **x** was 3 K, where 650 ($c_1 + c_2 + c_3$) features

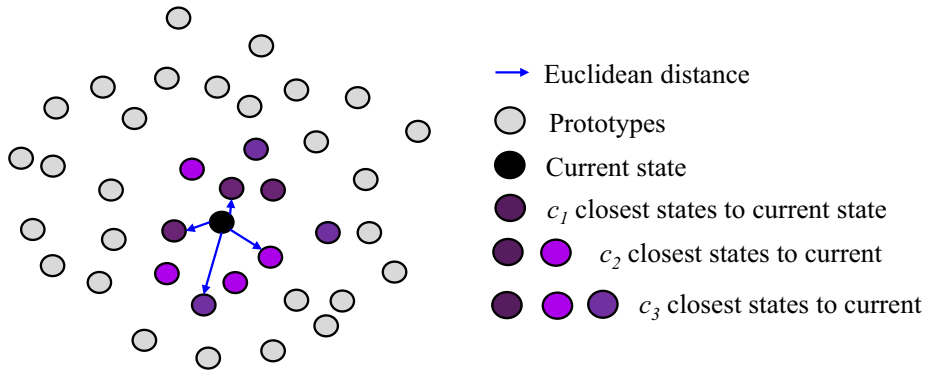

Figure 3. A depiction of selective Kanerva coding. The current state is 6-dimensional, one dimension per sensor signal used to represent the state space (left vGRF; right vGRF; sum of left and right vGRF; left angular velocity; right angular velocity; moving average of left vGRF). Prototypes closest to the current state within the state space are activated at each time-point. The result from selective Kanerva coding is a vector, where each point corresponds to a prototype. All points in the vector are either a 0 (unactivated) or a 1 (activated).

Algorithm 1. Selective Kanerva Coding as used in this work. Bolded variables refer to vectors or matrices; italicized variables refer to constants with values that pertain to this work. K = number of prototypes; n = number of sensors; c = closest prototypes to current state; **P** = prototype; **S** = state; **D** = distance vector; d = Euclidean distance; **x** = feature vector.

Selective Kanerva Coding

Parameters provided: K, n, c_1, c_2, c_3
Initialize prototypes **P** randomly once ever
Input new state **S**
Reset $\mathbf{D} = \text{zeros}(K,1)$
For i = 1 to K
For j = 1 to n
$\mathbf{D}_i \leftarrow d(\mathbf{P}_{i,j}, \mathbf{S}_j)$ d = Euclidean distance
$\mathbf{I} \leftarrow \text{Quickselect}(\mathbf{D})$ Indices of sorted distances
For m = 1 to 3
$\mathbf{ind}_m \leftarrow \mathbf{I}(1 \text{ to } c_m)$
$\mathbf{x}_{indm} \leftarrow 1$ Offset by $(m-1) \times K$
Output **x**

Algorithm 2. True online temporal difference learning. Reinforcement learning algorithm to estimate the discounted future values of sensor signals during walking. **w** = weight vector; **e** = eligibility trace; V = general value function; S = state; **x** = feature vector; Z = cumulant; δ = temporal difference; γ = termination signal/discounting factor; λ = eligibility trace parameter; α = learning step-size.

True Online TD(λ)

Initialize **w, e**, V_{old}, S, **x**
Repeat every time-step:
Generate next state S' and cumulant Z'
$\mathbf{x}' \leftarrow \text{SKC(S')}$
$V \leftarrow \mathbf{w}^T \mathbf{x}$
$V' \leftarrow \mathbf{w}^T \mathbf{x}'$
$\delta \leftarrow Z + \gamma V' - V$
$\mathbf{e} \leftarrow \gamma\lambda\mathbf{e} + \mathbf{x} - \alpha\gamma\lambda(\mathbf{e}^T\mathbf{x})\mathbf{x}$ Dutch trace
$\mathbf{w} \leftarrow \mathbf{w} + \alpha(\delta + V - V_{old})\mathbf{e} - \alpha(V - V_{old})\mathbf{x}$
$V_{old} \leftarrow V'$, $\mathbf{x} \leftarrow \mathbf{x}'$

were active at all times. The pseudocode for selective Kanerva coding used in this work is provided in algorithm 1.

2.4.4. True online temporal difference learning

True online temporal difference learning (TOTD), which is a reinforcement learning method, was used to learn the predictions of the sensor signals in real time during walking. Traditionally, reinforcement learning accomplishes a goal by maximizing future reward (Sutton and Barto 2018), but can also be used to estimate, or predict, the future values of signals other than reward. For example, general value functions (GVFs) can be learned to predict arbitrary signals of interest, called cumulants (Z) (White 2015). True online temporal difference learning (TOTD) was used to estimate the return, or future values of cumulants, using previously obtained estimates (algorithm 2). TOTD is an updated temporal difference learning method that has added terms to the eligibility trace and weight update equations (van Seijen *et al* 2015).

During walking, TOTD predicted the future values of three signals of interest that were recorded from the PML. Specifically, the returns of the cumulants were estimated in real time through the inner product of the weight vector (updated during TOTD) and the feature vector from function approximation (selective Kanerva coding), to produce the GVF for that cumulant (algorithm 2). The learning step-size (α), which determines the magnitude of the update, was set to 0.001, which was determined empirically. The bootstrapping parameter for the eligibility trace (λ) was set to 0.9 as is often standard. Different termination signals (γ) were determined for each cumulant empirically: 0.9 for unloading, 0.71 for vGRF, and 0.75 for the angular velocity of the PML. Because $\gamma = 1 - \frac{1}{T}$, where $T = 40$ ms (one time-step), these values of γ corresponded to timescales of 400 ms, 138 ms, and 160 ms, respectively.

2.5. Experimental protocol

A walking trial consisted of one trip across the walkway (~ 3 m). A naïve experimenter manually moved the PML through the walking cycle, and up to four different naïve experimenters moved the limb in each experiment. Since the cats were anaesthetized, the movements of the PML were entirely made by the experimenter, meaning that the signals from the vGRFs and angular velocity resulted from the experimenter's best approximation of the movements that would be made by an awake cat. The movements of the PML did not translate to the cat or cart due to the effects of the anaesthesia (i.e. did not produce weight-bearing or propulsive effects). The contractions of the muscles produced by ISMS in the SCL propelled the anaesthetized cat and cart along the walkway. The control method (reaction-based or Pavlovian) used for each walking trial was determined randomly by a different person than the one walking the PML, or by a random number generator. The person moving the limb was blinded to the control method driving ISMS for each trial.

For some trials, experimenters were told to purposefully make a mistake while walking the PML. A mistake was not explicitly defined; it was left to the discretion of the person walking the PML. Intentional mistakes included elongating the stance or the swing phase, shaking the limb in the air, or slipping forward or backward.

2.5.1. Reaction-based control trials

The first 2 consecutive walking trials for each naïve experimenter moving the PML were set up using thresholds for transitions through the phases of the walking cycle based on bench testing and knowledge from previous work (Dalrymple *et al* 2018). The thresholds were then revised based on the success of the transitions through the phases of the walking cycle and the collected sensor signals. The revised thresholds were then tested in the following walking trial and retuned until consistent walking was produced. After this initial tuning period, the person-specific thresholds remained constant throughout all cat experiments. An example of the thresholds specific to experimenter A is shown in figure 2(A). Optimizing initial thresholds for each person and holding them constant throughout all experiments allowed for a fair comparison of the control strategies with their best possible thresholds from start-up. Each person performed walking trials using the customized thresholds from the three other naïve experimenters walking the PML in addition to trials with their own thresholds.

2.5.2. Pavlovian control trials

Of the eight cat experiments conducted in this study, the first three had one set of thresholds on the predicted values of the sensor signals, while the remaining five had a different set of thresholds. As this was the first testing of these newly developed control strategies in an animal, initial setting of parameters was needed. By comparison, the rule-based control strategy had been employed in previous work (Dalrymple *et al* 2018), and was therefore tried and tested in an experimental setting already. The learning parameters and methods for Pavlovian control remained constant throughout the study. The initial thresholds for Pavlovian control were chosen based on testing on previously collected data from treadmill stepping (Dalrymple *et al* 2018) and bench testing on the walkway without a cat. These thresholds resulted in 55.4% of the steps triggered by back-up reactions and 2.0% of steps having missed phase transitions. Therefore, the thresholds were revised along with a change in which signal was used to predict some of the phases and held constant for the following five experiments. The back-up reaction thresholds were unchanged.

Several different trial types were conducted to investigate early learning, continued learning, and how the learning adapted or recovered after changes between cat experiments and people walking the PML. Early learning was evaluated by initializing the learning weights, eligibility trace, and GVFs to 0 at the beginning of a walking trial. In these trials, learning began anew with no prior knowledge. These early learning trials were repeated in every cat experiment with the different naïve experimenters walking the PML.

Learning also continued across several walking trials within each cat experiment. Throughout these trials within the experiment, multiple naïve experimenters took turns to walk the PML through the walking cycle. Furthermore, the carry-over of learning from one cat experiment to the next was tested over 5 cats. Repeating these carry-over trials in a new cat experiment allowed repeated investigation of the transfer of the learning algorithm between experiments with different cats and different experimenters walking the PML (i.e. different walking styles). A set of trials were also conducted whereby learning continued throughout 5 cat experiments, where multiple experimenters took turns to walk the PML within each experiment. These trials investigated the long-term learning and the adaptation to changes in cats and people walking the PML.

2.6. Data processing and analysis

2.6.1. Calculating alternation

The alternation of the two hind-limbs was calculated from the vGRFs using previously described methods (Dalrymple *et al* 2018). Briefly, the time spent in vertical loading per leg was converted into the degrees of a circle, with the onset of vertical loading of the PML defining the points of 0° and 360°. The half-way time of vertical loading for each limb was converted to degrees according to the step period. The difference of

the phase for each limb should equal 180º for perfect alternation.

2.6.2. Defining transitions as triggered by a prediction or a reaction

A step was considered to be entirely under Pavlovian control if transitions through all 4 phases of the gait cycle were achieved based on the thresholds placed on the predicted sensor values. If any of the phases required a back-up reaction (based on thresholds placed on the raw sensor values) to transition, then that entire step was counted as such.

2.6.3. Learning curves

The online prediction of the return (predicted discounted future values of the sensor signals) with the ideal return (actual discounted sum of future values of the sensor signals) was compared for each sensor value (Sutton and Barto 2018). During walking, TOTD estimated the return based on previous interaction and current sensor values. The ideal return was calculated post-hoc by summating the future raw sensor values discounted by the discount factor (γ) used for each sensor signal. The mean squared error between the online return and the ideal return for each sensor signal was calculated for early learning trials, and averaged the errors over the trial time (Modayil *et al* 2014).

2.7. Statistics

A one-sample t-test was used to compare the alternation phase differences with the target of 180°. A p-value ≤ 0.05 was considered to indicate significance. The effect size was determined using Cohen's d.

Chi-squared ($\mathbf{X}^2$) tests were conducted to compare the proportion of prediction-triggered phase transitions between different types of Pavlovian control walking trials (early, within one cat, carry-over, and continued learning), as well as for comparing the proportion of missed phase transitions across control methods. Cross-tabulations were generated for all pair-wise combinations. The $\mathbf{X}^2$ with the continuity correction for 2×2 contingency tables were reported, with the α-level adjusted using the modified Bonferroni correction for multiple comparisons.

3. Results

A total of 7943 steps from 770 trials were recorded from eight cats. On average, the step period was 1.32 s (SD = 0.26 s) and ranged from 0.44 s to 2.82 s.

3.1. Walking with reaction-based control

Reaction-based control was tested in all eight cats, resulting in 264 walking trials. The transferability of the tuned parameters for one walking pattern (by one naïve experimenter) to another and the need for retuning of parameters for reaction-based control of over-ground walking was documented. We found that one naïve experimenter's thresholds were not transferable to the other naïve experimenters. This was because there was high variability in the force production and movements produced by the 4 naïve experimenters walking the PML (figure 4(a)); therefore, customized thresholds for transitions between the phases of the walking cycle for each person were needed. The best performance of customized thresholds was 89.6% of steps successfully transitioning through the phases of the walking cycle. Overall, 18.7% (680/3645) of the total number of steps in all walking trials under reaction-based control had missed phase transitions due to the absence of threshold crossing by the raw sensors in these instances (table 1).

Table 1. Proportion of missed steps for combinations of people walking the PML (person-moved limb) using customized threshold settings for each person during reaction-based control.

		A	B	C	D
Threshold settings	A	11.0%	100.0%	94.0%	54.8%
	B	7.1%	10.4%	46.6%	12.5%
	C	5.9%	44.1%	12.0%	6.0%
	D	12.4%	12.8%	42.0%	11.5%

The alternation between the PML and the SCL was assessed using the phase difference between the two legs, where a phase difference of 180° indicated perfect alternation (Dalrymple *et al* 2018). Both hindlimb alternation and successful transitions through the phases of the walking cycle needed to occur for walking to be considered functionally effective. Overall, reaction-based control achieved a phase difference of 179.6° (SD = 19.0°). There were instances where the parameter settings for one naïve experimenter (person C) walking the PML resulted in fewer missed phase transitions when utilized for another experimenter walking the PML (persons A and D; table 1). However, poor PML-SCL alternation was encountered due to large variability in the walking patterns produced by the different experimenters walking the PML (figure 4(B)). This was because persons A and D made larger movements with larger sensor values than required for the settings tuned for person C, triggering phase transitions between the phases of the gait cycle earlier than needed to produce alternating walking. This produced a phase difference significantly less than 180° with very large effect sizes (phase difference for A = 155.0°; phase difference for D = 155.9°; $p < 0.0001$; df = 173, 46; one-sample t-test; Cohen's d = 0.64, 1.97). The inconsistent alternation and unsuccessful phase transitions across different people walking the PML (i.e. different walking styles) highlight the need for an automatically predictive control system.

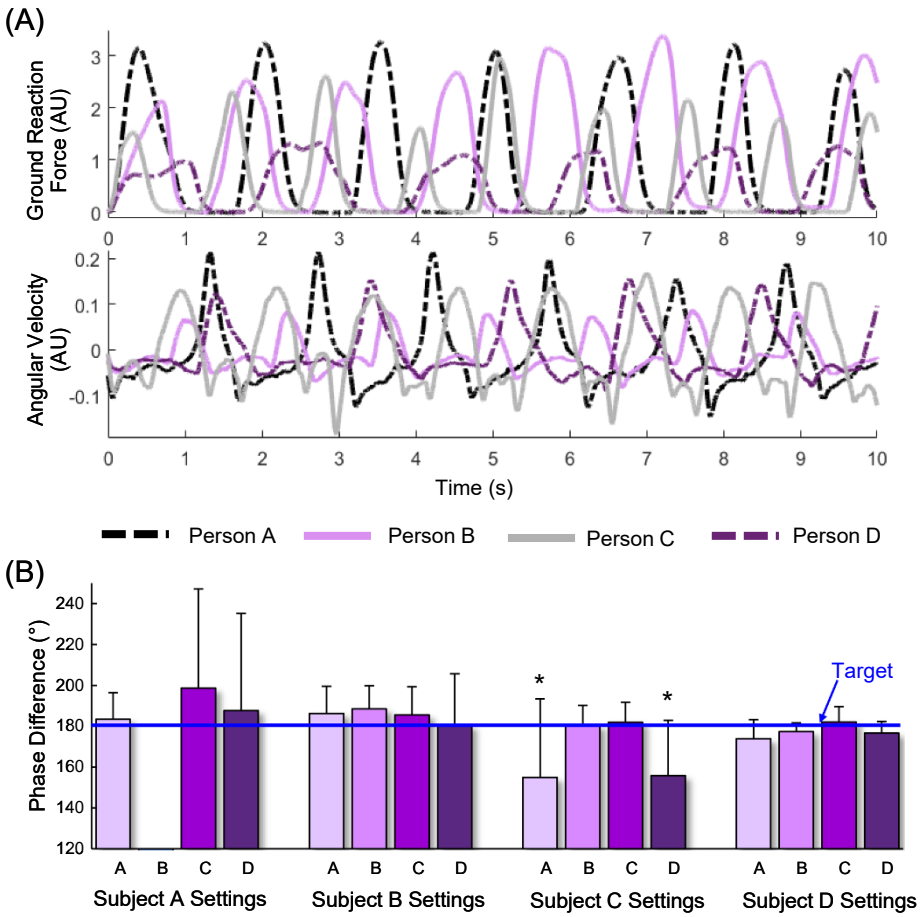

Figure 4. Walking using reaction-based control (n = 264 trials). (A) Ground reaction forces and angular velocities produced by each of the 4 naïve experimenters walking the PML (person-moved limb). (B) Alternation phase differences of the hind-limbs for each person walking the PML with threshold settings tuned for each person. Target alternation difference is 180°. *p < 0.0001.

3.2. Walking with Pavlovian control

3.2.1. Learning predictions of sensor signals occurs quickly

The Pavlovian controller learned predictions in real time during over-ground walking. Without prior learning, predictions became the only signals that initiated proper phase transitions within a maximum of four steps, which corresponded to approximately 4 s. Back-up reactions for phase transitions most commonly occurred within the first step, indicating that learning the predicted signals occurred quickly resulting in appropriate phase transitions (table 2). Fast learning is also demonstrated by the learning curves, where the mean squared error between the online and ideal returns decreased exponentially as learning continued within the trial (figure 5). Throughout all 1036 steps in the early learning trials (n = 88 trials), only 3 had failed phase transitions throughout the walking cycle. In 87.2% of the steps taken, the phase transitions were initiated by the predictions crossing the thresholds. Early learning trials had an average phase difference of 181.9° ± 7.8°. Therefore, the naïve learning algorithm was able to learn accurate predictions of walking-relevant sensor signals quickly, producing functional over-ground walking using ISMS.

Table 2. Back-up reactions in early learning trials using Pavlovian control. Within how many steps at the beginning of a walking trial was a back-up reaction triggered for each person walking the PML (person-moved limb).

	1 Step	2 Steps	3 Steps	More
A	92.3%	3.8%	3.8%	0.0%
B	97.1%	2.9%	0.0%	0.0%
C	69.6%	8.7%	17.4%	4.3%
D	20.0%	0.0%	40.0%	40.0%
All	84.1%	4.5%	8.0%	3.4%

3.2.2. Learning that continued within a cat experiment produced better Pavlovian control

As learning continued past one walking trial within a cat experiment, the predictions of the walking-relevant signals became smoother and more reliable as they accumulated more experience (figures 6(A) and B). The proportion of steps initiated by threshold crossings on the predicted sensor signals significantly increased compared to trials without prior learning (initialized to zero: 87.4% predictions; continued within one cat: 95.6% predictions; $p < 0.0001$, $\mathbf{X}^2$ test; figure 6(D)). The phase difference achieved in these trials was 181.1° and was not significantly

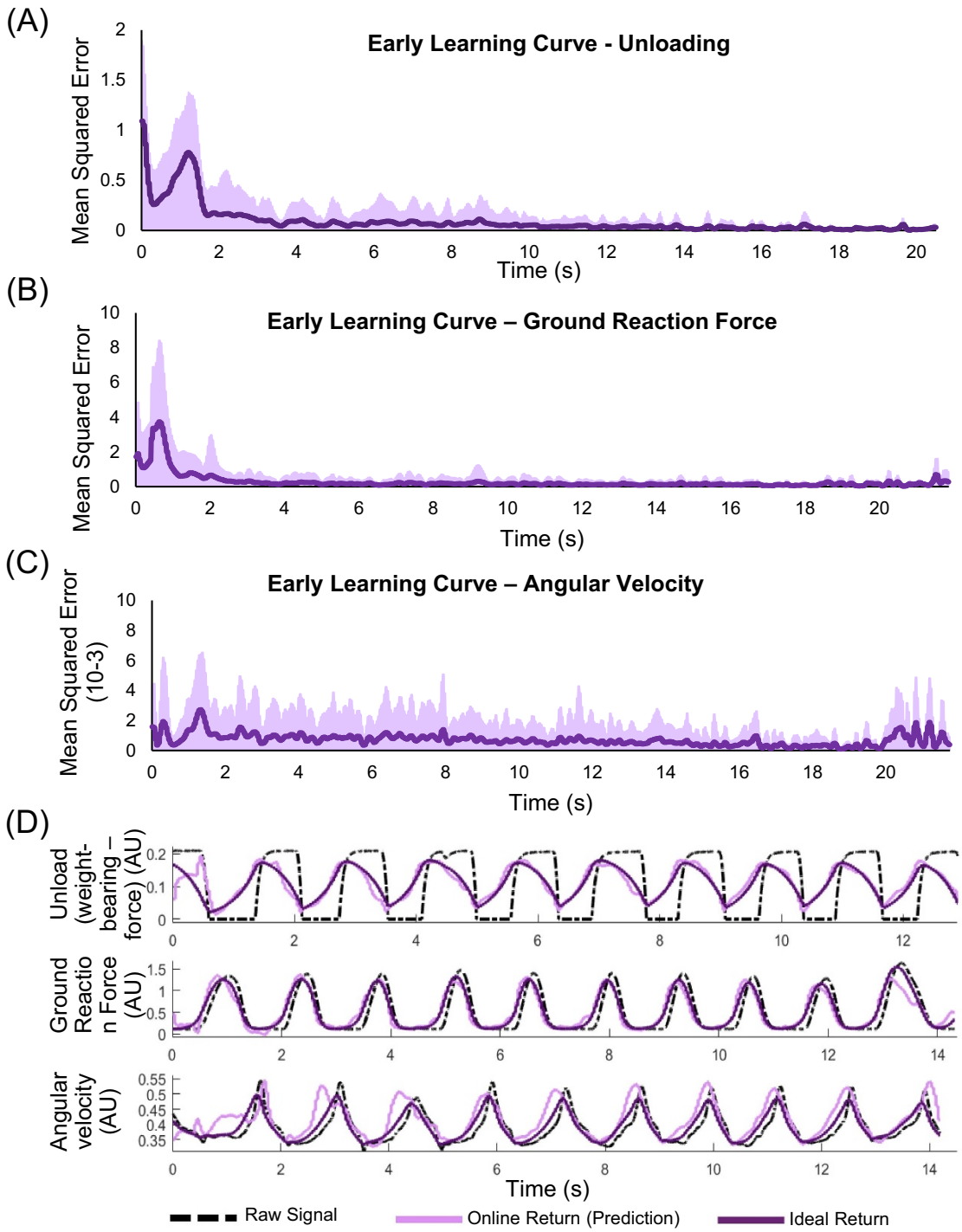

Figure 5. Speed of early learning (n = 88 trials). (A) Learning curve depicting the average (solid line) and standard deviation (shaded region) of the mean squared error of the unloading signal over time. (B) Learning curve for ground reaction force. (C) Learning curve for angular velocity. (D) Example of the actual sensor signals, online estimated return (prediction), and the ideal return (actual discounted sum of future values of the sensor signals).

different from the target of 180° ± 5.9° ($p = 0.077$; $df = 98$; one-sample t-test), demonstrating the ability to maintain alternation of the hind-limbs as online learning continued within a cat experiment for all experimenters walking the PML (figure 6(e)).

3.2.3. Learning continued to initiate prediction-based transitions across several cats and walking styles

The ability of the Pavlovian control to adapt to sudden changes in walking style was examined by evaluating the walking trials at the transition between different naïve experimenters walking the PML. As different naïve experimenters took turns to move the PML through the walking cycle, learning quickly acclimated to the new person and their style of walking. Of the 84 transition points between people, 64 did not require a back-up reaction to transition the SCL through the phases of the walking cycle (figure 7(A)). Only five trials required more than 1 step to adjust to the new naïve experimenter

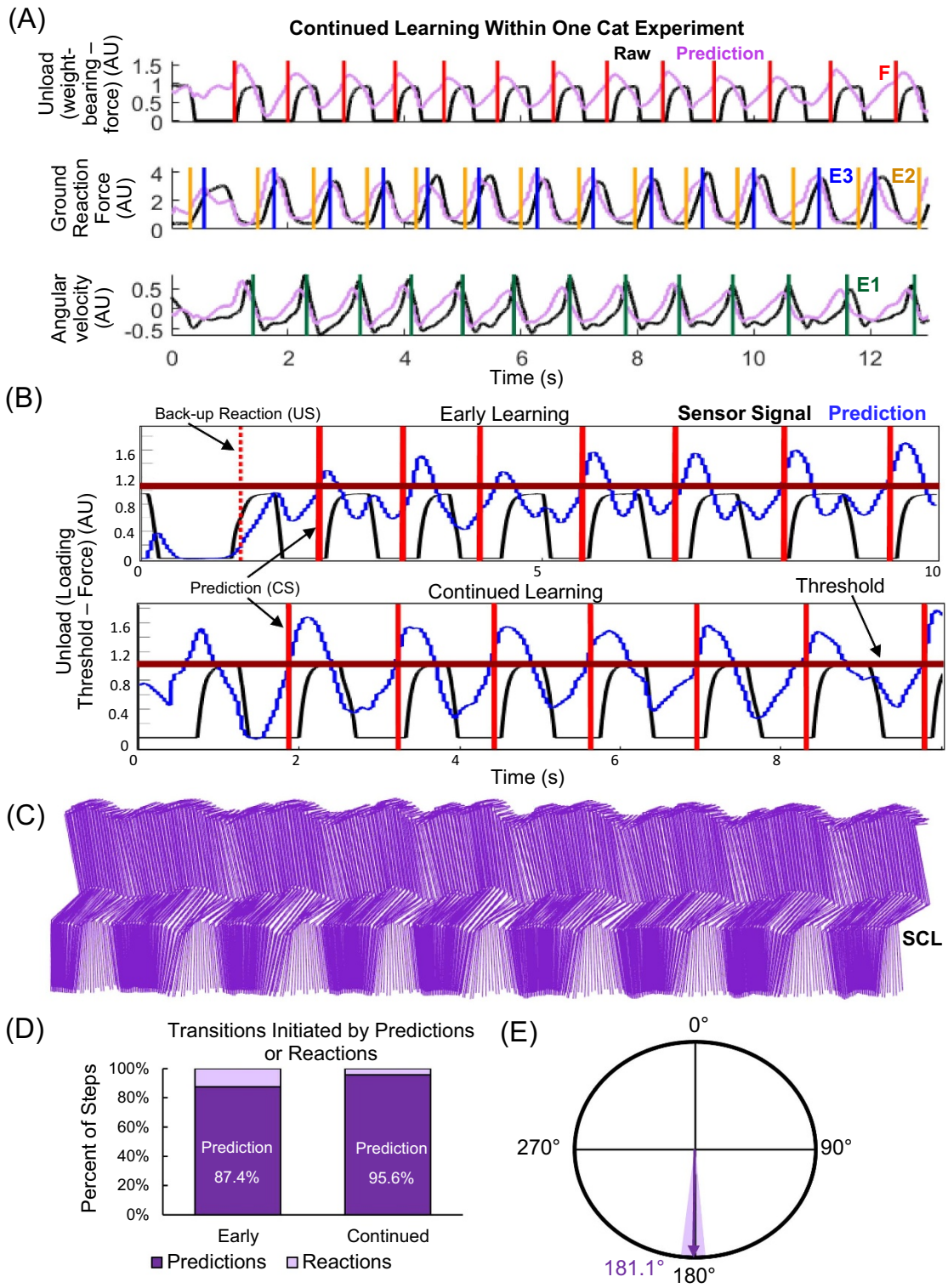

Figure 6. Effect of continued learning. (A) Raw sensor signals and learned predictions of unloading, ground reaction force, and angular velocity of the PML when learning continued between trials within a single cat experiment. The trigger times for the phases of the walking cycle are marked on their corresponding signal by a vertical line. (B) Representative examples of raw and predicted values of the unloading signal during early learning (initialized at 0) and continued learning (after several walking trials). Transitions for the F (early swing) phase made by back-up reactions (dashed vertical) and predictions (solid vertical) are marked. The solid horizontal line indicates the threshold value for predictions of unloading for Pavlovian control. (C) Movements produced by the SCL (stimulation-controlled limb) during a trial where learning continued within a single cat experiment. (D) Proportion of transitions initiated by predictions and reactions for early learning and continued learning walking trials. (E) Average (arrow) and standard deviation (shaded) alternation for trials where learning continued within a cat experiment.

walking the PML before only prediction-triggered transitions occurred. This impressively demonstrated fast adaptations to new environments, resulting in the first personalized and automatically predictive control strategy for a neural prosthesis. The next point of interest was to determine if the settings learned for

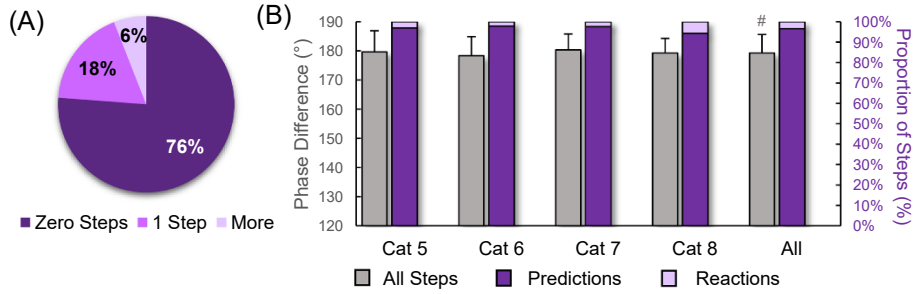

Figure 7. Learning adapts to different subjects during walking. (A) Proportion of steps with a back-up reaction following a transition to a new person walking the PML (person-moved limb; n = 84 transitions). (B) Alternation and proportion of transitions initiated by predictions and reactions when learning carried over between different cat experiments (n = 61 trials). $^{\#}p = 0.026$.

Pavlovian control in one animal could transfer and adapt to the next animal. This is analogous to having a new user of a clinical system have their initial settings mirror those of previous users instead of fine tuning the settings from scratch for the new user. Very interestingly, the learned predictions from previous experiments translated well to new cat experiments. In the first walking trial in the new animal, 83.3% of the steps taken did not require a back-up reaction for phase transitions, and 10.0% of steps requiring a back-up reaction were the first step in the trial. The steps in these walking trials were alternating, with an average phase difference of 179.1° ± 3.2°, which was significantly different from 180° but with a small Cohen's d effect size ($p = 0.026$; $df = 60$; one-sample t-test; Cohen's $d = 0.29$; figure 7(B)).

3.2.4. Learning continued to improve across several cats and walking styles

Long-term learning during walking was possible by continuing the learning over several cat experiments with different naïve experimenters taking turns to walk the PML. This provided an excellent representation of day-to-day changes that may occur in the walking styles produced by the users. The learned predictions triggered phase transitions in more than 91% of the steps taken for all naïve experimenters walking the PML, which was significantly higher than the proportion of prediction initiated transitions in early learning trials ($p < 0.0001$; $\mathbf{X}^2$ test figure 8(A)). Up to 98.7% of steps were transitioned using predictions (Person B; figure 8(B)). On average, these continuing walking trials had a phase difference of 180.8° ± 5.5° ($p = 0.113$; $df = 114$; one-sample t-test; figure 8(C)). Importantly, there were no missing phase transitions for walking in any trials where learning continued beyond the first learning trial.

3.2.5. Pavlovian control recovered from mistakes

Finally, we tested how the learning recovered from perturbations during walking. This is important because the end users of a neural prosthesis may have instances of instability. Different types of intentional mistakes were made by the naïve experimenters walking the PML throughout various stages of learning. The predictions were immune to the new and unexpected values of the sensor signals when walking was interrupted. Following a mistake, 94.4% (51/54) of the steps that followed had phase transitions triggered by the predicted sensor values (figure 9). Therefore, not only was the Pavlovian controller able to accommodate multiple users (i.e. cats) and multiple styles of walking (i.e. different people walking the PML), but it also was able to recover from mistakes made during walking.

4. Discussion

The goal of this study was to produce, for the first time, predictive, personalized, alternating, over-ground walking in a model of hemisection SCI using ISMS. The control strategy took advantage of 'residual function' to restore over-ground walking in anaesthetized cats. Reinforcement learning was used to learn predictions of walking-relevant sensor values. Pavlovian control used the predicted sensor values and threshold crossings on these predictions to control ISMS such that the 'paralyzed limb' is moved to the opposite phase of the walking cycle as the 'intact limb'. Pavlovian control was used across different people walking the 'intact limb' and throughout different cat experiments without requiring adjustments to the threshold settings. Learning occurred very quickly and consistently produced prediction-driven transitions between the phases of the gait cycle. The learned predictions were also resilient enough to recover quickly following a mistake during walking.

Personalized walking was possible for the first time because reinforcement learning acclimated to different people moving the 'intact limb' as well as to different cats. This comes in contrast to other approaches where the pattern of walking by the user (person) is dictated by the control algorithm. Adaptability and personalized walking are important because people with a SCI walk differently from each other, and there are changes in walking day-to-day for

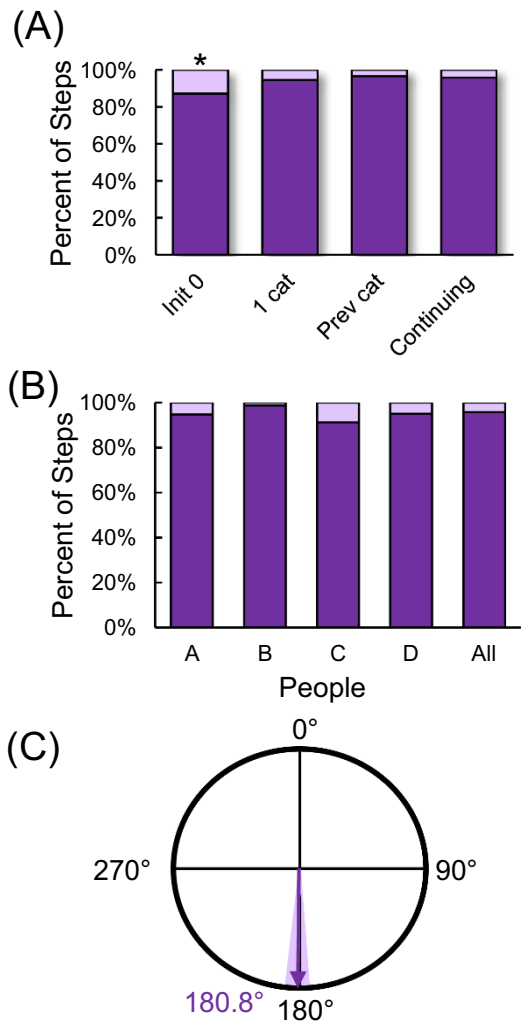

Figure 8. Predictions are reliable for the control of walking throughout learning trials for all people. (A) Proportion of steps with phase transitions initiated by threshold crossing on the learned predictions or by back-up reactions (threshold crossing on the raw sensor signals) at various stages of learning. *$p < 0.0001$. (B) Proportion of steps with phase transitions initiated by threshold crossing on the learned predictions or back-up reactions for all people walking the PML when learning continued across five cat experiments ($n = 115$ trials). (C) Average (arrow) and standard deviation (shaded) of the alternation phase difference of the hind-limbs when walking continued across five cat experiments and four people walking the PML.

a given person as well as with ongoing rehabilitation. Additionally, ISMS may facilitate plasticity and promote functional improvements as was demonstrated in the recovery of upper limb function of rats with a contusion SCI (Kasten *et al* 2013, Mcpherson *et al* 2015). Therefore, an automatically adaptable control system is highly desired to allow for minimal reprogramming.

4.1. Learning methods

This study used TOTD to learn GVFs for three cumulants during walking that were used for Pavlovian control. When a GVF crossed a pre-defined threshold, a stimulation response was delivered to move the SCL to the opposite phase of the walking cycle relative to the PML. The selective Kanerva function approximation method, predictions using GVFs, learning through TOTD, and Pavlovian control are relatively recent advancements made in the field of computing science (Sutton *et al* 2011, Modayil and Sutton 2014, van Seijen *et al* 2015, Travnik and Pilarski 2017). Selective Kanerva coding was chosen because it performs well online with a large number of sensors (Travnik and Pilarski 2017). It is also simple to implement and conceptualize. GVFs have proven to be a valuable tool in reinforcement learning: they allow the prediction of arbitrary signals; thus making reinforcement learning more powerful and applicable to more problems. In the field of rehabilitation, TD(λ) has been used to produce GVFs for upper-limb prostheses (Pilarski *et al* 2012, 2013a, 2013b, Sherstan and Pilarski 2014, Edwards *et al* 2016). TOTD offers an equivalence to the theoretical forward view of TD learning with negligible increase in computational cost (van Seijen *et al* 2015) and has been used to predict the shoulder angle of an upper-limb prosthesis (Travnik and Pilarski 2017). Pavlovian control has successively been used to control switching events of an upper-limb prosthesis in able-bodied study participants (Edwards *et al* 2013) and participants with amputations (Edwards *et al* 2016). It has also been used to control the turning off and spinning of a mobile robot (Modayil and Sutton 2014).

Pavlovian control is an appropriate approach to restoring walking after SCI using neural technology because learning the GVFs can occur very rapidly. Since the control strategy only requires the prediction to cross a threshold, online control can be initiated quickly. The learned predictions do not fluctuate nor are largely affected by sudden changes in the raw data, making them more reliable for threshold crossings in state control than the raw signals. Additionally, Pavlovian control does not require exploration of the state space, which is necessary in traditional reinforcement learning control methods. This is beneficial during walking because exploration of the state space could pose a danger to the user. For example, exploration may produce unsafe movement combinations such as double limb unloading. The state space could be restricted to avoid these dangerous situations, but this would limit the capacity of reinforcement learning and negate its usefulness. Therefore, Pavlovian control, which uses predictions to drive a fixed stimulation response, is suitable for a repetitive task such as walking.

Pavlovian control also allows for the knowledge of the expert designer to be incorporated into the rules that define the uses of the predictions and the output. This study, for the first time, combined all of these methods and used them to control a neural interface to produce over-ground walking *in vivo*.

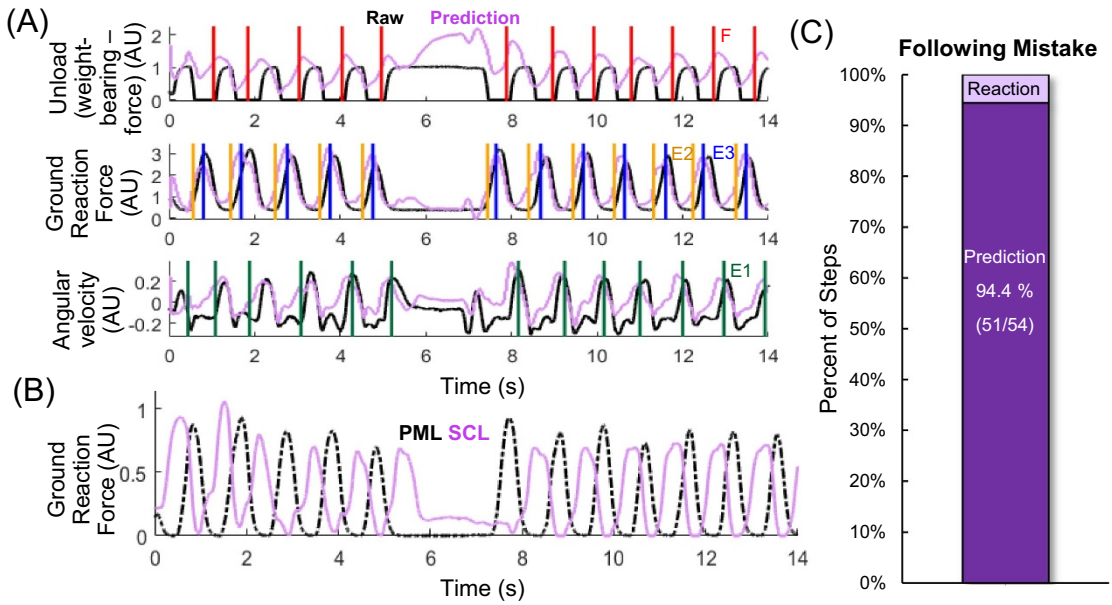

Figure 9. Example of an intentional mistake. (A) Raw signals for unloading, ground reaction force, and angular velocity of the PML (person-moved limb) and their corresponding learned predictions. Prediction-initiated transitions are indicated by solid vertical lines. The mistake begins at approximately 5 s and ends near 7.5 s. (B) Ground reaction forces produced by the PML (dashed trace) and SCL (stimulation-controlled limb; solid trace). (C) Proportion of steps with phase transitions initiated by threshold crossing on the learned predictions or by the back-up reaction following a mistake (n = 54 steps).

4.2. Biological parallels

Making predictions during a functional task is very useful and is commonly done naturally. For example, during walking, the central nervous system is continuously integrating sensory input from cutaneous receptors on the feet, stretch and loading sensors in the muscles and tendons, as well as visual and vestibular information to maneuver through the environment effectively and safely (Zehr *et al* 1997, Zehr and Stein 1999, Donelan and Pearson 2004, Marigold 2008, Mathews *et al* 2017). These sensory streams can be used to form short-term predictions that can be used in turn to adapt the gait pattern. Unexpected sensory stimuli result in reflexive changes, and with repetition, adaptation to the sensory stimuli occurs. For example, if an obstacle is placed in front of a cat's hind-limb during the swing phase causing activation of cutaneous receptors on the dorsum of the paw, the knee will flex further to clear the obstacle (Mcvea and Pearson 2007). This is a reflexive, or automatic response to the sensory stimulus, which is mediated by the spinal cord. If the obstacle is present for 20 continuous steps, the foot will begin lifting higher during swing in anticipation of the obstacle. Such effects, which are mediated by the cerebellum, can last for more than 24 h in some cases (Xu *et al* 2006). Although this is not exactly an example of Pavlovian control, it demonstrates the usefulness of predictions and how they can be utilized by the nervous system.

4.3. Relation to other control strategies

Current commercially available devices for restoring walking after SCI, such as the Parastep, Praxis, and various exoskeletons, have limited control options. The Parastep and Praxis systems use surface and implanted functional electrical stimulation (FES) electrodes, respectively (Chaplin 1996, Johnston *et al* 2005). Walking is accomplished using open loop alternation between stimulation of the quadriceps muscles and the peroneal nerve, with each step initiated using push-buttons on a walker. Powered exoskeletons initiate open-loop walking by the user leaning forward (Chang *et al* 2015, Ekelem and Goldfarb 2018). The users are expected to adapt their walking to accommodate the control strategy in the device. To restore meaningful and functional walking, especially after an incomplete SCI, the control strategy needs to adapt to the user, utilize residual function, and deliver stimulation to compensate for the deficits as needed.

Both the Pavlovian and reaction-based controllers were finite state controllers, which is a concept that has been used previously to produce walking in models of SCI. Finite state control has the advantage of incorporating expert knowledge in a straight-forward manner to define the rules for walking (Popović 1993, Sweeney *et al* 2000). Finite state control of surface (Andrews *et al* 1988) and intramuscular (Guevremont *et al* 2007, Mazurek *et al* 2012) FES of the leg muscles used information from ground reaction forces and hip angle to control the transition between the phases of the gait cycle. Previous controllers for ISMS in a model of complete SCI used ground reaction forces and hip angle (Saigal *et al* 2004, Holinski *et al* 2016) or recordings from the dorsal root ganglia (Holinski *et al* 2013) to transition the hind-limbs

through the different phases of the walking cycle (Dalrymple and Mushahwar 2017).

The current study demonstrated that predictions can be learned to initiate transitions between the phases of the gait cycle using only two sensor signals: ground reaction force and angular velocity. These sensors can easily be integrated into a wearable system; gyroscopes are small microchips and force sensitive resistors can be placed in the soles of shoes (Kirkwood *et al* 1989, Kostov *et al* 1992). Recent work has demonstrated that kinematic data can be used to identify the phases of the gait cycle during walking (Drnach *et al* 2018). Switched linear dynamical systems (SLDS) was used to model the joint angle kinematics in neurologically intact people walking on a treadmill. The offline SLDS models were able to label the correct phase of the gait cycle with 84% precision. Future work may incorporate more portable sensors such as goniometers along with online models to build predictions of gait phases.

Control strategies utilizing machine learning are needed for automatic adaptation of stimulation settings to restore walking. Supervised machine learning has been used to control surface functional electrical stimulation (FES) systems in persons with SCI to track joint angles (Abbas and Triolo 1997, Popović *et al* 1999, Qi *et al* 1999), initiate the swing phase (Kirkwood and Andrews 1989, Kostov *et al* 1992, 1995, Sepulveda *et al* 1997, Tong and Granat 1999), control FES over multiple joints (Fisekovic and Popovic 2001), predict different phases of the gait cycle in neurologically-intact subjects (Kirkwood and Andrews 1989, Williamson and Andrews 2000), and in finite control of FES walking after complete SCI (Popović 1993). However, supervised learning requires manual labelling of data and is limited by the data set used for training. Many examples with sufficient variability are needed in the training data set to obtain an accurate generalization. Ideally, stimulation settings would be tuned once during the initial set-up for each person, and thereafter automatically adjust to any changes in daily gait patterns.

4.4. Experimental limitations

The model of a hemisection SCI used in this study enabled thorough testing of the control strategies while avoiding the need for inducing SCIs. It allowed testing of the ability of the control strategies to augment residual function in a controlled manner. This necessitated voluntary control of one hind-limb to be mimicked by a person moving the limb through the walking cycle. This was the first testing of these control strategies, and the outcomes served as a proof-of-concept implementation. Further work may test these control strategies in chronically injured cats.

A hemisection SCI has more stereotypic functional deficits compared to other injuries such as bilateral contusion SCIs. Although these SCIs are rare, e.g. Brown-Sequard syndrome (Roth *et al* 1991, Wirz *et al* 2010), the control strategies may be extended to hemiplegia in general, which includes stroke and traumatic brain injury.

The thresholds for Pavlovian control were finalized after initial testing in early experiments. They were chosen based on testing on previously collected data from treadmill stepping (Dalrymple *et al* 2018) and bench testing on the walkway without a cat. Moderate performance of walking was achieved; however, transitions were improved with changes to the thresholds and the signals on which the thresholds were placed. It is important to note that the learning parameters of the predictions were never changed as they were consistently accurate. Additionally, once the new thresholds were set, they were never modified thereafter. This demonstrates that the initial design decision of where to place the thresholds was important, but once it was finalized, no further changes were necessary. The thresholds for Pavlovian control did not require tuning for different people and cats, because the learned predictions acclimated to the changes. However, it may be beneficial to introduce adaptive thresholds in the future, especially if these strategies were to be employed in more variable injury models. Furthermore, the stimulation amplitudes and channels that produced the functional responses remained constant during a walking trial. Future work may introduce a learning strategy that aims to optimize and adapt the stimulation channels and amplitudes in addition to current strategy which controls the timing.

4.5. Future considerations

Pavlovian control learned predictions for ground reaction force and angular velocity signals; however, other sensor signals could also be used to provide more information about the environment. For example, muscle activity recorded using EMG, joint angles provided by goniometers, or visual information through cameras or infrared sensors could all be recorded and used to acquire more predictions. The addition of sensors (e.g. EMG, goniometers) could be useful to restore walking after variable injuries or to provide information regarding the walking terrain (visual, infrared) to adapt the control strategy. For example, if a user needed to step up a curb, visual or infrared sensors as well as goniometers and EMG activity could be used by a Pavlovian controller to predict the change in gait and adjust the stimulation output accordingly to facilitate curb-stepping. Additional sensors could also be used to provide stability information such as loss of balance, fatigue, and the reliance on the upper body for support. More control rules could be incorporated to predict and correct these unsafe situations. Furthermore, the addition of sensors is feasible if a state representation method such as selective Kanerva coding is used, as was the case in this work, because it is not affected

by the increase in dimensions (Travnik and Pilarski 2017).

Pavlovian control can easily be expanded to neuromodulation systems such as deep brain stimulation for various conditions including Parkinson's disease and depression, neuroprosthetic systems for restoring function after stroke or traumatic brain injury, and exoskeletons and artificial limbs.

5. Conclusion

Pavlovian control of walking augmented function in a model of hemisection SCI. Using predictions of sensor signals during walking, Pavlovian control was resilient to transitions between different walking styles, between cat experiments, and recovered from mistakes made during walking.

Pavlovian control of ISMS has the potential to enhance ambulation capacity greatly, generating functional over-ground walking. Very importantly, we have demonstrated, for the first time, that control strategies using intelligent machine learning approaches such as Pavlovian control can reduce the burden of tuning stimulation parameters for controlling a neuroprosthesis. This increases the ease of translation of innovative neural technologies to clinical settings. The control strategy can also be extended to other injury models and interventions such as peripheral FES, lower-limb prostheses, and exoskeletons.

Acknowledgments

The authors thank Adrian Lopera Valle and Amirali Toossi for their assistance with the experiments and Rod Gramlich for building the instrumented walkway used in this study. This work was funded by the Canadian Institutes of Health Research. AND was supported by a Queen Elizabeth II Graduate Student Scholarship and DAR was supported by the University of Alberta Undergraduate Research Initiative. VKM is a Canada Research Chair (Tier 1) in Functional Restoration and a Killam Annual Professor.

Author contributions

VKM secured funding, conceived, and supervised the study. AND and RSS formulated the machine learning approach. AND developed and programmed the control strategies. AND, RSS, and VKM designed the experimental protocols. AND manufactured the implant arrays. AND and DAR performed the surgeries. VKM performed the procedures associated with implantation of the intraspinal microstimulation electrodes. AND and DAR collected and analyzed the data. AND wrote the initial version of the manuscript and made the figures; all authors contributed to its editing. VKM approved the final version of the manuscript on behalf of the co-authors.

ORCID iDs

Ashley N Dalrymple https://orcid.org/0000-0001-8566-7178
David A Roszko https://orcid.org/0000-0002-7128-3517
Richard S Sutton https://orcid.org/0000-0002-3679-3415
Vivian K Mushahwar https://orcid.org/0000-0001-9873-611X

References

Abbas J J and Triolo R J 1997 Experimental evaluation of an adaptive feedforward controller for use in functional neuromuscular stimulation systems *IEEE Trans. Rehabil. Eng.* **5** 12–22

Andrews B J, Baxendale R H, Barnett R, Phillips G F, Yamazaki T, Paul J P and Freeman P A 1988 Hybrid FES orthosis incorporating closed loop control and sensory feedback *J. Biomed. Eng.* **10** 189–95

Angeli C A, Boakye M, Morton R A, Vogt J, Benton K, Chen Y, Ferreira C K and Harkema S J 2018 Recovery of over-ground walking after chronic motor complete spinal cord injury *New Engl. J. Med.* **379** 1244–50

Bamford J A, Marc Lebel R, Parseyan K and Mushahwar V K 2017 The fabrication, implantation, and stability of intraspinal microwire arrays in the spinal cord of cat and rat *IEEE J. Mag.* **25** 287–96

Bhumbra G S and Beato M 2018 Recurrent excitation between motoneurones propagates across segments and is purely glutamatergic *PLoS Biol.* **16** e2003586

Carhart M R, He J, Herman R, D'Luzansky S and Willis W T 2004 Epidural spinal-cord stimulation facilitates recovery of functional walking following incomplete spinal-cord injury *IEEE Trans. Neural Syst. Rehabil. Eng.* **12** 32–42

Chang S R, Kobetic R, Audu M L, Quinn R D and Triolo R J 2015 Powered lower-limb exoskeletons to restore gait for individuals with paraplegia—a review *Case Orthop. J.* **12** 75–80

Chaplin E 1996 Functional neuromuscular stimulation for mobility in people with spinal cord injuries. The parastep I system *J. Spinal Cord Med.* **19** 99–105

Collinger J L, Boninger M L, Bruns T M, Curley K, Wang W and Weber D J 2013 Functional priorities, assistive technology, and brain–computer interfaces after spinal cord injury *J. Rehabil. Res. Dev.* **50** 145

Dalrymple A 2019 *Machine Learning to Characterize Motor Patterns and Restore Walking After Neural Injury* (Edmonton, AB, Canada: University of Alberta) (https://doi.org/10.7939/r3-6q2s-s362)

Dalrymple A N, Everaert D G, Hu D S and Mushahwar V K 2018 A speed-adaptive intraspinal microstimulation controller to restore weight-bearing stepping in a spinal cord hemisection model *J. Neural. Eng.* **15** 056023

Dalrymple A N and Mushahwar V K 2017 Stimulation of the spinal cord for the control of walking *Neuroprosthetics* vol 8 Series on Bioengineering and Biomedical Engineering (Singapore: World Scientific) pp 811–49

Donelan J M and Pearson K G 2004 Contribution of sensory feedback to ongoing ankle extensor activity during the stance phase of walking *Can. J. Physiol. Pharmacol.* **82** 589–98

Drnach L, Essa I and Ting L H 2018 Identifying gait phases from joint kinematics during walking with switched linear dynamical systems *BioRxiv* 378380

Edwards A L, Dawson M R, Hebert J S, Sherstan C, Sutton R S, Chan K M and Pilarski P M 2016 Application of real-time machine learning to myoelectric prosthesis control: a case series in adaptive switching *Prosthet. Orthot. Int.* **40** 573–81

Edwards A L, Kearney A, Dawson M R, Sutton R S and Pilarski P M 2013 Temporal-difference learning to assist human decision making during the control of an artificial limb arXiv:1309.4714 [Cs] (http://arxiv.org/abs/1309.4714).

Ekelem A and Goldfarb M 2018 Supplemental stimulation improves swing phase kinematics during exoskeleton assisted gait of SCI subjects with severe muscle spasticity *Front. Neurosci.* **12** 374

Engberg I and Lundberg A 1969 An electromyographic analysis of muscular activity in the hindlimb of the cat during unrestrained locomotion *Acta. Physiol. Scand.* **75** 614–30

Fisekovic N and Popovic D B 2001 New controller for functional electrical stimulation systems *Med. Eng. Phys.* **23** 391–9

Gill M L *et al* 2018 Neuromodulation of lumbosacral spinal networks enables independent stepping after complete paraplegia *Nat. Med.* **24** 1677–82

Goslow G E, Reinking R M and Stuart D G 1973 The cat step cycle: hind limb joint angles and muscle lengths during unrestrained locomotion *J. Morphol.* **141** 1–41

Guevremont L, Norton J A and Mushahwar V K 2007 Physiologically based controller for generating overground locomotion using functional electrical stimulation *J. Neurophysiol.* **97** 2499–510

Hardin E, Kobetic R, Murray L, Corado-Ahmed M, Pinault G, Sakai J, Bailey S N, Ho C and Triolo R J 2007 Walking after incomplete spinal cord injury using an implanted FES system: a case report *J. Rehabil. Res. Dev.* **44** 333–46

Hofstoetter U S, Krenn M, Danner S M, Hofer C, Kern H, Mckay W B, Mayr W and Minassian K 2015 Augmentation of voluntary locomotor activity by transcutaneous spinal cord stimulation in motor-incomplete spinal cord-injured individuals *Artif. Organs* **39** E176-186

Holinski B J, Everaert D G, Mushahwar V K and Stein R B 2013 Real-time control of walking using recordings from dorsal root ganglia *J. Neural. Eng.* **10** 056008

Holinski B J, Mazurek K A, Everaert D G, Stein R B and Mushahwar V K. 2011 Restoring stepping after spinal cord injury using intraspinal microstimulation and novel control strategies *Conf. Proc.: Annual Int. Conf. of the IEEE Eng. Med. Biol. Soc.* pp 5798–801

Holinski B J, Mazurek K A, Everaert D G, Toossi A, Lucas-Osma A M, Troyk P, Etienne-Cummings R, Stein R B and Mushahwar V K 2016 Intraspinal microstimulation produces over-ground walking in anesthetized cats *J. Neural. Eng.* **13** 056016

Hunter J P and Ashby P 1994 Segmental effects of epidural spinal cord stimulation in humans *J. Physiol.* **474** 407–19

Johnston T E, Betz R R, Smith B T, Benda B J, Mulcahey M J, Davis R, Houdayer T P, Pontari M A, Barriskill A and Creasey G H 2005 Implantable FES system for upright mobility and bladder and bowel function for individuals with spinal cord injury *Spinal Cord* **43** 713–23

Kasten M R, Sunshine M D, Secrist E S, Horner P J and Moritz C T 2013 Therapeutic intraspinal microstimulation improves forelimb function after cervical contusion injury *J. Neural. Eng.* **10** 044001

Kirkwood C A and Andrews B J 1989 Finite state control of FES systems: application of AI inductive learning techniques *Images of the Twenty-First Century. Proc. Annu. Int. Eng. Med. Biol. Soc.* **3** 1020–1

Kirkwood C A, Andrews B J and Mowforth P 1989 Automatic detection of gait events: a case study using inductive learning techniques *J. Biomed. Eng.* **11** 511–16

Kobetic R, Triolo R J and Marsolais E B 1997 Muscle selection and walking performance of multichannel FES systems for ambulation in paraplegia *IEEE Trans. Rehabil. Eng.* **5** 23–29

Kostov A, Andrews B J, Popović D B, Stein R B and Armstrong W W 1995 Machine learning in control of functional electrical stimulation systems for locomotion *IEEE Trans. Biomed. Eng.* **42** 541–51

Kostov A, Stein R B, Armstrong W W and Thomas M 1992 Evaluation of adaptive logic networks for control of walking in paralyzed patients *1992 14th Annual Int. Conf. IEEE Eng. Med. Biol. Soc.* **4** 1332–4

Kunam V K, Velayudhan V, Chaudhry Z A, Bobinski M, Smoker W R K and Reede D L 2018 Incomplete cord syndromes: clinical and imaging review *Radiographics* **38** 1201–22

Lam T, Pauhl K, Ferguson A, Malik R N, Krassioukov A, Eng J J and Eng J J 2015 Training with robot-applied resistance in people with motor-incomplete spinal cord injury: pilot study *J. Rehabil. Res. Dev.* **52** 113–29

Lau B, Guevremont L and Mushahwar V K 2007 Strategies for generating prolonged functional standing using intramuscular stimulation or intraspinal microstimulation *IEEE Trans. Neural Syst. Rehabil. Eng.* **15** 273–85

Marigold D S 2008 Role of peripheral visual cues in online visual guidance of locomotion *Exerc. Sport Sci. Rev.* **36** 145–51

Mathews M A, Camp A J and Murray A J 2017 Reviewing the role of the efferent vestibular system in motor and vestibular circuits *Front. Physiol.* **8** 552

Mazurek K A, Holinski B J, Everaert D G, Stein R B, Etienne-Cummings R and Mushahwar V K 2012 Feed Forward and feedback control for over-ground locomotion in anaesthetized cats *J. Neural. Eng.* **9** 026003

Mcpherson J G, Miller R R and Perlmutter S I 2015 Targeted, activity-dependent spinal stimulation produces long-lasting motor recovery in chronic cervical spinal cord injury *Proc. Natl. Acad. Sci. USA* **112** 12193–8

Mcvea D A and Pearson K G 2007 Long-Lasting, context-dependent modification of stepping in the cat after repeated stumbling-corrective responses *J. Neurophysiol.* **97** 659–69

Modayil J and Sutton R S 2014 Prediction driven behavior: learning predictions that drive fixed responses (https://www.aaai.org/ocs/index.php/WS/AAAIW14/paper/view/8740)

Modayil J, White A and Sutton R S 2014 Multi-timescale nexting in a reinforcement learning robot *Adapt. Behav.* **22** 146–60

Moritz C T, Perlmutter S I and Fetz E E 2008 Direct control of paralysed muscles by cortical neurons *Nature* **456** 639–42

Morrison S A, Lorenz D, Eskay C P, Forrest G F and Basso D M 2018 Longitudinal recovery and reduced costs after 120 sessions of locomotor training for motor incomplete spinal cord injury *Arch. Phys. Med. Rehabil.* **99** 555–62

Mushahwar V K, Collins D F and Prochazka A 2000 Spinal cord microstimulation generates functional limb movements in chronically implanted cats *Exp. Neurol.* **163** 422–9

Mushahwar V K and Horch K W 1998 Selective activation and graded recruitment of functional muscle groups through spinal cord stimulation *Ann. N. Y. Acad. Sci.* **860** 531–5

Mushahwar V K and Horch K W 2000a Muscle recruitment through electrical stimulation of the lumbo-sacral spinal cord *IEEE Trans. Rehabil. Eng.* **8** 22–29

Mushahwar V K and Horch K W 2000b Selective activation of muscle groups in the feline hindlimb through electrical microstimulation of the ventral lumbo-sacral spinal cord *IEEE Trans. Rehabil. Eng.* **8** 11–21

Musselman K E, Fouad K, Misiaszek J E and Yang J F 2009 Training of walking skills overground and on the treadmill: case series on individuals with incomplete spinal cord injury *Phys. Ther.* **89** 601–11

Pilarski P M, Dawson M R, Degris T, Carey J P, Chan K M, Hebert J S and Sutton R S 2013a Adaptive artificial limbs: a real-time approach to prediction and anticipation *IEEE Robot. Autom. Mag.* **20** 53–64

Pilarski P M, Dawson M R, Degris T, Carey J P and Sutton R S 2012 Dynamic switching and real-time machine learning for improved human control of assistive biomedical robots *2012 4th IEEE RAS EMBS Int. Conf. on Biomedical Robotics and Biomechatronics* pp 296–302

Pilarski P M, Dick T B and Sutton R S 2013b Real-time prediction learning for the simultaneous actuation of multiple prosthetic joints *IEEE Int. Conf. on Rehabilitation Robotics* pp 6650435

Popović D, Stein R B, Oğuztöreli N, Lebiedowska M and Jonić S 1999 Optimal control of walking with functional electrical stimulation: a computer simulation study *IEEE Trans. Rehabil. Eng.* **7** 69–79

Popović D B 1993 Finite state model of locomotion for functional electrical stimulation systems *Prog. Brain Res.* **97** 397–407

Qi H, Tyler D J and Durand D M 1999 Neurofuzzy adaptive controlling of selective stimulation for FES: a case study *IEEE Trans. Rehabil. Eng.* **7** 183–92

Roth E J, Park T, Pang T, Yarkony G M and Lee M Y 1991 Traumatic cervical Brown-Sequard and Brown-Sequard-plus syndromes: the spectrum of presentations and outcomes *Paraplegia* **29** 582–9

Saigal R, Renzi C and Mushahwar V K 2004 Intraspinal microstimulation generates functional movements after spinal-cord injury *IEEE Trans. Neural. Syst. Rehabil. Eng.* **12** 430–40

Sepulveda F, Granat M H and Cliquet A 1997 Two artificial neural systems for generation of gait swing by means of neuromuscular electrical stimulation *Med. Eng. Phys.* **19** 21–28

Sherstan C and Pilarski P 2014 Multilayer general value functions for robotic prediction and control *IROS 2014 Workshop on AI and Robotics, IEEE/RSJ Int. Conf. Intelligent Robot. Syst.* vol 6 (Chicago, IL)

Spinal Cord Injury (SCI) 2017 SCI Facts and Figures *J. Spinal Cord Med.* **40** 872–3

Stein R B and Mushahwar V 2005 Reanimating limbs after injury or disease *Trends Neurosci.* **28** 518–24

Sutton R S and Barto A G 2018 Reinforcement learning: an introduction *Adaptive Computation and Machine Learning Series* 2nd edn (Cambridge, MA: MIT Press) pp 53–221

Sutton R S, Modayil J, Delp M, Degris T, Pilarski P M, White A and Precup D 2011 Horde: a scalable real-time architecture for learning knowledge from unsupervised sensorimotor interaction *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Syst.* (May 2–6, 2011, Taipei, Taiwan) pp 761–68

Sweeney P C, Lyons G M and Veltink P H 2000 Finite state control of functional electrical stimulation for the rehabilitation of gait *Med. Biol. Eng. Comput.* **38** 121–6

Tator C H, Minassian K and Mushahwar V K 2012 Spinal cord stimulation: therapeutic benefits and movement generation after spinal cord injury *Handb. Clin. Neurol.* **109** 283–96

Tong K Y and Granat M H 1999 Gait control system for functional electrical stimulation using neural networks *Med. Biol. Eng. Comput.* **37** 35–41

Toossi A, Everaert D G, Uwiera R R E, Hu D S, Robinson K, Gragasin F S and Mushahwar V K 2019 Effect of anesthesia on motor responses evoked by spinal neural prostheses during intraoperative procedures *J. Neural. Eng.* **16** 036003

Travnik J B and Pilarski P M 2017 Representing high-dimensional data to intelligent prostheses and other wearable assistive robots: a first comparison of Tile Coding and selective Kanerva Coding *IEEE Int. Conf. on Rehabilitation Robotics* pp 1443–50

van Seijen H, Mahmood A R, Pilarski P M, Machado M C and Sutton R S 2015 True online temporal-difference learning *J. Mach. Learn. Res.* **17** 1–40

Vanderhorst V G and Holstege G 1997 Organization of lumbosacral motoneuronal cell groups innervating hindlimb, pelvic floor, and axial muscles in the cat *J. Comp. Neurol.* **382** 46–76

Wagner F B, Mignardot J-B, Le Goff-Mignardot C G, Demesmaeker R, Komi S, Capogrosso M and Rowald A *et al* 2018 Targeted neurotechnology restores walking in humans with spinal cord injury *Nature* **563** 65–71

Wen Y, Si J, Brandt A, Gao X and Huang H 2019 Online reinforcement learning control for the personalization of a robotic knee prosthesis *IEEE Trans. Cybern.* **50** 2346–56

White A 2015 *Developing a Predictive Approach to Knowledge* (Edmonton, AB, Canada: University of Alberta)

Williamson R and Andrews B J 2000 Gait event detection for FES using accelerometers and supervised machine learning *IEEE Trans. Rehabil. Eng.* **8** 312–19

Wirz M, Zörner B, Rupp R and Dietz V 2010 Outcome after incomplete spinal cord injury: central cord versus Brown-Sequard syndrome *Spinal Cord* **48** 407–14

Xu D, Liu T, Ashe J and Bushara K O 2006 Role of the olivo-cerebellar system in timing *J. Neurosci.* **26** 5990–5

Zehr E P, Komiyama T and Stein R B 1997 Cutaneous reflexes during human gait: electromyographic and kinematic responses to electrical stimulation *J. Neurophysiol.* **77** 3311–25

Zehr E P and Stein R B 1999 What functions do reflexes serve during human locomotion? *Prog. Neurobiol.* **58** 185–205