TITLE

Cerebellar patients have intact feedback control that can be leveraged to improve reaching

AUTHORS

Amanda M. Zimmet[a,b]
amandazimmet@gmail.com
ORCID ID: 0000-0003-1457-3072

Di Cao[d]
dcao9@jhu.edu
ORCID ID: 0000-0002-1547-9929

Amy J. Bastian[b,c*]
bastian@kennedykrieger.org
ORCID ID: 0000-0001-6079-0997

Noah J. Cowan[d*]
ncowan@jhu.edu
ORCID ID: 0000-0003-2502-3770

*these authors jointly supervised this work

AUTHOR AFFILIATIONS

a. Johns Hopkins University Department of Biomedical Engineering
School of Medicine
720 Rutland Avenue / Ross 720
Baltimore, MD 21205, USA

b. Kennedy Krieger Institute
707 N. Broadway St.
Baltimore, MD 21205, USA

c. Johns Hopkins University Department of Neuroscience
School of Medicine
1003 Wood Basic Science Building
725 N. Wolfe Street
Baltimore, MD 21205, USA

d. Johns Hopkins University Department of Mechanical Engineering
223 Latrobe Hall
3400 North Charles Street
Baltimore, MD 21218, USA

CORRESPONDING AUTHOR

Noah J. Cowan
ncowan@jhu.edu
410-516-5301

Laboratory for Computational Sensing and Robotics
121 Hackerman Hall
3400 N. Charles Street
Baltimore, MD 21218, USA

ABSTRACT

It is thought that the brain does not simply react to sensory feedback, but rather uses an internal model of the body to predict the consequences of motor commands before sensory feedback arrives. Time-delayed sensory feedback can then be used to correct for the unexpected—perturbations, motor noise, or a moving target. The cerebellum has been implicated in this predictive control process. Here we show that the feedback gain in patients with cerebellar ataxia matches that of healthy subjects, but that patients exhibit substantially more phase lag. This difference is captured by a computational model incorporating a Smith predictor in healthy subjects that is missing in patients, supporting the predictive role of the cerebellum in feedback control. Lastly, we improve cerebellar patients' movement control by altering (phase advancing) the visual feedback they receive from their own self movement in a simplified virtual reality setup.

INTRODUCTION

Humans normally rely on a balance of feedback and predictive control mechanisms to make smooth and accurate movements. Proprioceptive and visual feedback are necessary for determining body postures at the beginning and end of a movement and can be used to guide slow movements accurately. However, feedback is time delayed, and thus it never represents the current state of the body during movement. Because of this, it is thought that we depend on internal models of the body that are built based on prior experience. These models can be rapidly accessed and thus provide a fast internal prediction system to estimate how a movement will unfold, enabling us to better understand where our limbs are at any given moment. This allows us to make fast and accurate movements despite long-latency feedback.

People with cerebellar damage show a characteristic pattern of incoordination during movement that is referred to as ataxia. When reaching, they make curved movements that miss intended targets and require multiple corrections. This pattern of over- and undershooting a target (dysmetria) and oscillatory corrections (intention tremor) are hallmarks of cerebellar ataxia. One hypothesis that might explain ataxia is that the predictive estimation and control provided by cerebellar circuits is dysfunctional or lost (Miall et al., 2007; Wolpert et al., 1998).

Normally, the estimation of limb state (e.g., position and velocity) benefits from integrating proprioceptive measurements with an internal predictive control model during a movement (Adamovich et al., 1998; Fuentes and Bastian, 2010; Paillard and Brouchon, 1974). However, patients with cerebellar damage do not seem to receive this benefit (Bhanpuri et al., 2013; Weeks et al., 2017). Worse, it is possible that their predictive model actually conveys incorrect state information during active movements, which could corrupt rather than enhance proprioceptive estimation of limb state. This difficulty of predicting the future state of limbs during active movement leads to movements that are poorly directed and scaled, requiring ongoing corrections to reach a goal location.

Patients with cerebellar ataxia may rely more heavily on visual feedback to correct dysmetric movements (Beppu et al., 1987, 1984; Day et al., 1998). A drawback of visual feedback is that it is slower than proprioception; intention tremor in these patients' movements is thought to stem from dependence on time-delayed visual feedback to make corrections (Day et al. 1998). However, it is not known how well they incorporate this visual feedback into their movements. Is visual feedback control impaired? Or does the dysmetria stem solely from errors in predictive (i.e., feedforward) control (Bhanpuri et al., 2014; Manto et al., 1994; Smith and Shadmehr, 2005)? A gait study by Morton and Bastian suggested that cerebellar patients could

use feedback information to make reactive corrections during split-belt treadmill walking in order to maintain stability, but lacked the ability to adapt their predictive motor patterns (Morton and Bastian, 2006). Other studies have indirectly supported the notion that patients can incorporate some level of visual feedback control in reaching (Day et al., 1998; Smith et al., 2000).

In this study, we used behavioral experiments and computational modeling to test for any impairment in visual feedback control (correction based on measured error) and disambiguate it from previously described impairments in feedforward control (Bares et al., 2007; Broersen et al., 2016). Subjects performed a visuomotor task which required them to track an unpredictable target (Roth et al., 2011). This method allowed us to determine that cerebellar patients can integrate visual feedback control similarly to healthy, age-matched control subjects. We then hypothesized that we could exploit intact feedback control to reduce dysmetria. Specifically, we provided an acceleration-dependent alteration to visual feedback of a hand cursor. This alteration serves to compensate for the phase lags introduced by time delay, allowing patients to use visual feedback more effectively. This real time controller successfully reduced patients' dysmetria.

METHODS

EXPERIMENTAL MODEL AND SUBJECT DETAILS

A total of 17 patients with cerebellar deficits and 14 age-matched controls were tested in one or more of the following experiments. Patients were excluded if they had any clinical or MRI evidence of damage to extra-cerebellar brain structures, or clinical evidence of dementia, aphasia, peripheral vestibular loss, or sensory neuropathy. The age-matched controls were clinically screened for any neurological impairments. Experiments 1-2 tested 11 cerebellar

patients (7 male, 4 female) and 11 age-matched controls (4 male, 7 female). Patient and control ages were within +/-3 years of age. The one exception is a pairing of a 78 year old patient with a 71 year old control. Experiment 3 tested 12 cerebellar patients (8 male, 4 female) and 12 age-matched controls (3 male, 9 female), and all patient and control ages were within +/-3 years of age. Subjects gave informed consent according to the Declaration of Helsinki. The experimental protocols were approved by the Institutional Review Board at Johns Hopkins University School of Medicine. Each subject used his or her dominant arm for all tasks. The only exception was one unilaterally affected cerebellar patient who was instructed to use her affected, non-dominant, hand. We quantified each subject's ataxia using the International Cooperative Ataxia Rating Scale (ICARS) (Trouillas et al., 1997). Because this study involves upper limb reaching behavior, we calculated an Upper Limb ICARS sub-score comprising the sum of the upper-limb kinetic function elements of the test. The results of the ICARS as well as demographic information for all patients are shown in Table 1.

EXPERIMENTAL APPARATUS, EXPERIMENTS, AND TASK INSTRUCTIONS

For all experiments, subjects were seated in a KINARM exoskeleton robot (BKIN Technologies Ltd., Kingston, Ontario, Canada), shown in Figure 1A, which provided gravitational arm support while allowing movement in the horizontal plane. A black video screen occluded the subjects' view of their arm movements. The shoulder position was fixed at a 75 degree angle, as shown in Figure 1C. The wrist joint was also fixed; therefore, the elbow joint was the only freely mobile joint. Data were recorded at 1 kHz. The KINARM system exhibits a cursor delay. Using a high-speed camera, we measured this delay (0.0458s) and took the delay into account in modeling the human control system.

144 A 1 cm diameter white dot (cursor) was projected onto the display to indicate the
145 subject's index fingertip position. The display was calibrated so that the projected dot position
146 aligned with the subject's fingertip position. Subjects performed three experiments:

147 • **Experiment 1:** sum-of-sines tracking task.

148 • **Experiment 2:** single-sines tracking task.

149 • **Experiment 3:** discrete reaches with acceleration-dependent feedback.

150

151 For Experiments 1 and 2 (tracking tasks), subjects were instructed to try to keep the
152 cursor in the center of the target, a 1.5 cm diameter green dot (Figure 1C). The target angle
153 followed a pseudorandom sum-of-sines (Figure 1D) or single-sine pattern. Each trial was 100
154 seconds long. We tested three different Feedback Gain conditions in the sum-of-sines
155 (Experiment 1) task: 1.35 (cursor moves 35% farther than hand), 0.65 (cursor motion is
156 attenuated by 35%) and 1.0 (veridical feedback; Figure 1B). Subjects were not informed that
157 there was a gain change.

158 For Experiment 3 (discrete reaches), the robot moved the arm to align the cursor with the
159 red dot start position (55 degree elbow angle) before each reach. After approximately 2 seconds,
160 the red dot disappeared, and a green target dot (1.5 cm diameter, at 85 degree elbow angle)
161 appeared elsewhere on the screen. Subjects were instructed to bring the white cursor dot to the
162 center of the green target dot (a 30 degree flexion movement) as smoothly and accurately as
163 possible without over or undershooting the target. If the fingertip did not reach within 1.75 cm
164 of the target within 800 ms, the green dot changed to blue and subjects were instructed to
165 complete the reach (trials were not repeated); the subjects were also asked to try to move faster

on the next trial. The next trial started after the subject kept the fingertip in the target dot for two continuous seconds.

Healthy control subjects performed five practice trials and patients performed 10 practice trials to familiarize themselves with the procedure. If subjects were still confused about the procedure after the practice trials, they were given more practice trials until they were comfortable with the procedure.

Some subjects completed only a subset of the experiments due to time constraints (see Table 1). For every experiment, between trials, subjects were told to relax while the robot moved the subject's hand to bring the cursor to the start position.

EXPERIMENTAL DESIGN AND STATISTICAL ANALYSIS

Generating Target Trajectories (Experiments 1 & 2)

The goal of Experiment 1 was to determine how well patients and controls could track an unpredictable sum-of-sines stimulus and probe their ability to use visual feedback for movement control. The target trajectory comprised 15 sine waves whose frequencies were prime multiples of 0.05 Hz, namely 0.1, 0.15, 0.25, 0.35, 0.55, 0.65, 0.85, 0.95, 1.15, 1.45, 1.55, 1.85, 2.05, 2.15, and 2.35 Hz. Phases of each sinusoid were randomized. Each trial was 100 s long, comprising five replicates of the sum-of-sines trajectory. The sinusoids were scaled so that the *angular velocity* of each sinusoid never exceeded 720 degrees per second while the *positional* amplitude of each sine wave never exceeded 2 degrees. The non-harmonic relation of the component sinusoids created a temporally complex target motion with 45 degrees of freedom (amplitude, phase, and frequency for 15 sinusoids) that repeats every 20 s; a portion of the trajectory is shown in Figure 1D, wherein it can be observed that the signal trajectory includes multiple

seemingly random turn-arounds within even very short timescales. This complexity, coupled with the long period, makes this target motion far less predictable than a simple single-sine motion (Miall and Jackson, 2006; Poulton, 1974; Roth et al., 2011). Lastly, the dramatic increase in tracking phase lag presented in this paper (see Results) for sum-of-sine responses recapitulates the main finding of (Roth et al., 2011), confirming that from a practical point of view subjects were less able to predict the movement trajectories. Nevertheless, this signal design allowed good signal-to-noise ratio at each frequency while avoiding excessively large or fast motions of the target, making it well suited as a probe signal for tracking behavior (Roth et al., 2011). The normalized magnitude of the Fourier spectrum of our input signal is shown in the top half of in Figure 1E. Five trials were conducted for each condition tested.

In Experiment 2, we wanted to determine how well patients and controls could track a predictable single sine wave. Every other frequency tested in the sum-of-sines condition was tested as a single, standalone sine wave: 0.1, 0.25, 0.55, 0.85, 1.15, 1.55, 2.05, and 2.35 Hz. The amplitude of each sine wave matched (component-wise) the sum-of-sines experiments. Subjects were once again instructed to keep the cursor in the target dot as much as possible. The order of presentation of the stimuli was randomized for each subject.

Estimating Frequency Responses (Experiments 1 & 2)

All analysis was performed using custom scripts (Zimmet, Cao, Bastian, and Cowan 2020) in MATLAB (The Mathworks Inc., Natick, MA, USA). To obtain the steady-state frequency response of the subject, the first period (20 seconds) was discarded as transient, leaving four periods per trial. The data were visually inspected for unusual activities that were not representative of the subject's typical behavior. For example, if the subject turned away from

the screen to cough or talk, movement would cease for a few seconds and the data from those few seconds were removed from further analysis. The movement trajectories for a given subject were then averaged across trials at each time instant, excluding any deleted data. The averaged data were linearly detrended and converted to the frequency domain via the discrete Fourier transform (DFT). An example of the result of this process for a cerebellar patient is shown in Figure 1F (the frequency domain data is complex valued at each frequency so only the magnitude is shown).

For Experiment 1 (sum of sines), to estimate a subject's frequency response, we calculated the ratio of the DFT of the subject's movement to the DFT of the underlying target trajectory at each of the 15 frequencies in the sum-of-sines. The result of this calculation is 15 complex numbers, called *phasors*, representing the frequency response estimate for a given subject. The magnitude of each phasor is the *gain* and the angle is the *phase* of the subject's response at that frequency.

The processing of the data for Experiment 2 (single sines) was identical to the processing of the sum-of-sines data, except that there was no averaging across trials and no manual inspection/filtering of the data for unusual activities. This is because we were only looking at one frequency from each sine wave trial, and any error caused by an unusual, aperiodic activity (such as pausing) would introduce very little power at any individual frequency.

Phasor Plot (Experiments 1 & 2)

To visualize the frequency response, we used Phasor Plots as shown in Figure 2A. In these plots, the gain is the radial distance from the origin, and phase lag is the clockwise angle of the data with respect to the horizontal axis. Perfect tracking at a given frequency corresponds to

unity gain and zero phase lag, as illustrated by the large dot in Figure 2B. The distance between any point on the phasor plot and this dot provides a measure of the amplitude of the (sinusoidal) error signal at that particular frequency as a proportion of the input signal amplitude.

Additionally, this visualization can be enhanced with a circle with a 0.5 radius centered at +0.5 along the horizontal axis as shown in Figure 2B. Points on this circle correspond to the ideal response gain for minimizing error given a particular phase lag (Roth et al., 2011). Points on this circle represent the gain values that minimize error assuming the subject cannot further minimize phase lag. Points outside the circle indicate the subject is moving too much for a given phase lag; reducing the amplitude of movement would achieve a lower tracking error and would do so with less effort. Points within the circle may represent striking a balance between effort and error in that the subject is not moving enough to fully minimize their error given a certain phase lag. Lastly, points that lie on the unit circle, traced in bold black in Figure 2C, would indicate that the subject is replicating the sinusoidal component exactly, albeit at a phase shift.

Phasor Plot Scaling (Experiment 1)

If a subject were to "fully" respond to an applied visual gain, the subject would need to scale his or her movement so that the visual output appeared the same as it did in the veridical feedback condition, although there is no constraint (or instruction) to do so. In reality a subject may scale his or her movements more than this, less than this, or even not at all. To quantify this, we fit a Scaling Factor that, when multiplied by the frequency response phasors for a given Feedback Gain condition, best aligned the phasors with those of the veridical condition. We only included the five lowest frequencies in this calculation because at these frequencies subjects

generally appeared to have robust responses (based on visual inspection of the data). At high frequencies, however, response magnitudes were low and therefore fitting the Scaling Factor could be dominated by noise.

Scaling Factor (dependent measure) was analyzed using a 2-way ANOVA, testing for the effects of two independent variables, group (patient vs. control) and Feedback Gain (0.65 and 1.35). If a main effect was found to be significant, we calculated an effect measure (η_p^2) using the Measures of Effect Size toolbox for Matlab (Hentschke and Stüttgen, 2011).

Modeling (Using Data from Experiment 1)

To capture the visuomotor tracking response of both cerebellar patients and age-matched healthy controls, we adapted the classic McRuer gain-crossover model: a delayed, scaled integrator which specifically assumes that the subject is responding to an unpredictable stimulus (McRuer and Krendel, 1974). The McRuer gain-crossover model is highly simplified, and in its original incarnation was meant to capture the sensorimotor frequency response near the gain crossover frequency (i.e. where the open-loop gain of the combined plant and controller has unit magnitude). The idea is that near this frequency, closed-loop robustness is governed by the amount of phase lag; if that phase lag exceeds 180 degrees, the closed-loop system becomes unstable. In a neighborhood of the gain-crossover frequency, McRuer and Krendel (1974) recognized that subjects tended to have an open-loop gain whose magnitude dropped off inversely with frequency (like an integrator, or k/s), and the phase lag was a bit more than 90 degrees (the integrator adds 90 degrees, and a delay in their model increases the phase lag as a function of frequency). We adopt a version of this simplified model to facilitate the interpretation of data (Figure 4).

We hypothesized that the patient and control models would be equivalent except for the magnitude of the feedback delay (Figure 4A). To test this, we considered a more general class of models that allows for distinct parameters for the patients and healthy control subjects, and includes the possibility of distinct delays on the visual measurement of target motion from that of self-movement feedback (which can incorporate for example, proprioception, which generally incurs a lower latency than visual). To determine the simplest combined model of patients and controls which provided both a good and consistent fit without overfitting, we used a model selection process similar to that described in (Madhav et al., 2013) based on Experiment 1 data (sum-of-sines). In the most flexible model, all patient and control parameters were allowed to vary independently with eight free parameters: $k_{patient}$, $k_{control}$, Visual $Gain_{patient}$, Visual $Gain_{control}$, Visual $Delay_{patient}$, Visual $Delay_{control}$, Feedback $Delay_{patient}$, and Feedback $Delay_{control}$ (Figure 4S1). The following model selection procedure aimed to determine which of those eight parameters should be free and which should be yoked together to most parsimoniously fit the data. All combinations of yoked parameters were tested, subject to yoking gains with gains, and delays with delays. We also tested a few degenerate model structures (such as a pure gain or delay, with no feedback); some of these model structures allowed for the elimination of one or more of the blocks within the model structure.

The following procedure was repeated for each possible model (note: here a "model" includes both patient and control submodels). For each subject, we represented their frequency response as an array of 15 complex numbers (one for each of the frequencies tested in the sum-of-sines experiment). We began by pulling 10 of the 11 age-matched patient–control pairs from the full dataset. Then, we took the average of those 10 patients' (and, separately, controls) frequency responses at each tested frequency. Using the MATLAB function 'fminsearch,' we

used these average frequency response functions to fit the current model parameters to the average data by minimizing the frequency-domain (FD) error:

$$\text{FD-error} = \left\|\text{control}_{\text{avg}} - \text{control}_{\text{model fit}}\right\|^2 + \left\|\text{patient}_{\text{avg}} - \text{patient}_{\text{model fit}}\right\|^2 \quad (1)$$

where *control* and *patient* in the above equation were arrays with 15 complex values, corresponding respectively to frequency responses of healthy control subjects and patients at the 15 frequencies tested. This was repeated 100 times with different initial parameter values selected for the 'fminsearch' function, increasing the likelihood of finding a global minimum. The model parameters that generated the lowest overall error were kept (as well as the magnitude of the error). This was repeated for each patient–control pair, yielding 11 sets of model parameters and 11 leave one out error (loo-error) values, defined as:

$$\text{loo-error} = \left\|\text{control}_{\text{left out}} - \text{control}_{\text{model fit}}\right\|^2 + \left\|\text{patient}_{\text{left out}} - \text{patient}_{\text{model fit}}\right\|^2 \quad (2)$$

The leave one out error values were averaged together to create the overall leave one out error for that model structure. This was our gauge for the ability of the model to capture the responses. Using the 11 sets of the model parameters for that particular model (i.e. one set of parameters for each subject pair that was left out), we created an 11 x 8 matrix, each row of which corresponding to the parameters for a given leave-one-out fit. The elements of the matrix were the residuals between each of the 8 parameters and their average value. We calculated the maximum singular value of this matrix as a gauge for model consistency. This was repeated for all potential models. Finally, by modifying our fitting procedure above, we used bootstrapping to generate confidence intervals for the parameter values in our best model. Specifically, we

randomly sampled eleven times with replacement from the eleven subjects. This was repeated 1000 times to generate 1000 sets of parameter values. From these values, we found the 95% confidence intervals.

"Essentially all models are wrong, but some are useful" (Box and Draper, 1987) so rather than choosing a single "correct" model, we scrutinized all models that exhibit a good tradeoff of model fit and model consistency. We also considered the number of free parameters (parsimony). Lastly, we examined whether a given model produced physiologically realistic parameters; models with nonphysiological parameters would suggest inadequate model structure leading to parameter bias. We determined the common features and subtle differences between these good models and the outcome of these meta analyses are described in Results.

We converted each frequency-domain model (which includes patient and control submodels) to state space in Matlab, and simulated the model pair using the lsim command (Figure 4S1D). The input was the sum-of-sines target angle trajectory used in experiments. The model output produced a distinct prediction for the elbow-angle trajectory for each group (patients and controls; Figure 4C). These simulated elbow-angle responses were compared to subject responses as follows. We averaged across trials for each subject to compute a single time-domain response for each patient and each age-matched control; we then averaged across subjects in each group to compute a mean elbow-angle response for patients and a mean elbow-angle response for controls. Given the simulated model responses and actual patient and control mean responses, the time-domain (TD) error was calculated as follows:

$$\text{TD-error} = \sqrt{\frac{1}{T}\int_0^T (\text{control}(t)_{\text{mean}} - \text{control}(t)_{\text{model}})^2 dt} + \sqrt{\frac{1}{T}\int_0^T (\text{patient}(t)_{\text{mean}} - \text{patient}(t)_{\text{model}})^2 \, dt} \quad (3)$$

where the integration is computed as a Riemann sum over the last $T = 80$ seconds of the 100s trial period, with a sampling time of 1ms. The TD-error was normalized so that the maximum error was 1 for visualization purposes.

Comparing Phase Lags across Cohorts and Conditions (Experiments 1 & 2)

We tested whether the cerebellar group showed a different pattern of phase lag across frequencies compared to controls. Specifically, we compared the phase lag for the lowest common frequencies (0.10, 0.25, 0.55, 0.85, 1.15) in the sum-of-sines and single-sine conditions across the two groups. We hypothesized that the control group would be able to use prediction and follow the single sine wave with little phase lag compared to the cerebellar patients. We expected that there would be less of a difference between groups when they followed the sum-of-sines. We used a parametric two-way ANOVA for circular data called the Harrison-Kanji test for this analysis (Berens, 2009).

Acceleration-Dependent Feedback (Experiment 3)

To examine the effect of modified visual cursor feedback in VR, we implemented Acceleration-Dependent Feedback, where the cursor angle was set to follow the elbow angle plus an acceleration-dependent term:

$$\mathrm{CursorAngle(t)} = \mathrm{ElbowAngle(t)} + k_a \frac{d^2}{dt^2}\mathrm{ElbowAngle(t)} \quad (4)$$

To implement this acceleration-dependent feedback, we used the KINARM's real-time computer to calculate the average of the previous 100 elbow angular acceleration values as an approximation of the instantaneous acceleration (at 1 kHz sampling, this resulted in a $\sim$ 50 ms delay). This acceleration estimate was multiplied by the Acceleration-Dependent Feedback Gain (k_a), added to the real elbow angle, and the position of the cursor was displayed on the Kinarm

screen at this new, slightly shifted angle. In practical terms, this moving-average-filtered acceleration-dependent feedback amounts to a high-pass filter, providing anticipatory (i.e. phase-leading) feedback to the user when $k_a > 0$, and providing phase-lagging feedback when $k_a < 0$.

To find a patient-specific feedback gain k_a, cerebellar patients performed blocks of 10 reaches where this gain was held constant. Each patient started with $k_a = 0$ (veridical feedback) for their first block of reaches. Individual reaches were categorized as hypometric, on-target, or hypermetric. Based on the mode of this categorization for the 10 reaches, the gain was increased, remained the same, or decreased, as determined by the Parameter Estimation by Sequential Testing (PEST) algorithm (see Implementation of the PEST Algorithm (Experiment 3), below) (Taylor and Creelman, 1967).

In addition, control and patient participants performed blocks of 5 reaches at 9 specific gains in this order: k_a = 0 (veridical), 0.005, 0.010, 0.015, 0.020, −0.005, −0.010, −0.015, and −0.020. Participants were given five trials of one Feedback Gain before being exposed to the next Feedback Gain on the list. Note that only six of the twelve patients completed this experiment due to time constraints.

Quantifying Dysmetria for Discrete Reaches (Experiment 3)

We quantified dysmetria in this single-joint task by measuring the elbow angle at which the subject made his or her first correction. We determined the angle of the first correction by finding when the velocity crossed zero after the initiation of the reach. The angle of the elbow at this time point is the angle of the first correction. This angle was divided by the goal angle of 30 degrees. In a smooth and accurate reach, no correction would be needed and the "angle of first correction" would be the target angle; in this scenario the ratio between the "angle of first

correction" and the goal angle would be 1. If the result was greater than 1.03, the reach was categorized as hypermetric, and likewise, if the result was less than 0.97, the reach would be categorized as hypometric. Results between 0.97 and 1.03 were classified as "on target" reaches.

Implementation of the PEST Algorithm (Experiment 3)

The Parameter Estimation by Sequential Testing (PEST) algorithm was used to iteratively determine the best Acceleration-Dependent Feedback Gain values to apply in Experiment 3 by analyzing the history of the responses to different applied values (Taylor and Creelman, 1967). If the mode of the initial set of reaches was hypermetric, the Acceleration-Dependent Feedback Gain would be decreased. Likewise, it would be increased if the mode of the set was hypometric. The initial step size used for the PEST algorithm was 0.01. The maximum step size for the PEST algorithm was limited to 0.01 and the minimum step size was 0.0035. The maximum Acceleration-Dependent Feedback Gain was set to +/-0.02. The PEST algorithm was terminated when two blocks of the same Acceleration-Dependent Feedback Gain yielded a mode of reaches classified as "on target." Alternatively, if this portion of the experiment took longer than approximately 20 minutes or the subject was experiencing fatigue, the most successful Acceleration-Dependent Feedback Gain was selected based on those which had already been tested.

RESULTS

Experiment 1: Cerebellar Patients and Age-Matched Controls Respond Similarly to Rescaling of Visual Self-Motion Feedback

To probe participants' ability to use visual feedback, we challenged cerebellar patients and age-matched controls to keep a cursor within a target dot that was following an unpredictable trajectory. The normalized Fourier spectrum (magnitude only) of the target trajectory is shown in Figure 1E and the spectrum of a single subject's response is shown in Figure 1F. Note that this subject had clear peaks at all of the frequencies of target movement in the veridical condition (1x).

Subjects also were exposed to two Feedback Gain conditions where the cursor was presented at 0.65x or 1.35x. A phasor plot for each Feedback Gain condition is shown for a single cerebellar patient in Figure 3A. Points with the smallest phase lag (~30°) are responses from the lowest frequencies (0.1 Hz). Responses at the five lowest frequencies are marked with solid dots. As a function of increasing input frequency, the phase lag increases and the plotted points move progressively clockwise around the origin. Note in Figure 3A that the example subject scales his movement, especially at the lowest frequencies, in response to the Feedback Gain applied. For the 0.65 Feedback Gain condition, he increases the scaling of his movement in comparison to the veridical condition, as expected. Similarly, in the 1.35 Feedback Gain condition, he decreases the scaling of his movement in comparison to the veridical condition.

For each patient, we computed a Scaling Factor that best scaled responses at the lowest 5 frequencies (Methods); to visualize how well this Scaling Factor represents our data, the phasor

437 data from an example cerebellar patient in the 1.35 Feedback Gain condition shown in Figure 3A
438 is multiplied by its scaling factor to yield the corresponding line in Figure 3B. If the scaling
439 factor is a good fit for the data, the data should rest on top of the veridical line. Note that the
440 phase is not modified by this scaling computation, so we should not expect to see an
441 improvement in phase alignment in this second plot.
442
443 Figure 3C shows the Scaling Factors for all subjects for each Feedback Gain condition.
444 We conducted a one-sided t-test to determine whether patients scaled their gain in the
445 hypothesized direction against the null hypothesis that patients' scaling factors would be 1 (i.e.,
446 they would not respond to the applied feedback gain). Both groups scaled up (patients: $p=0.03$,
447 $t=-2.1$, DOF=10; controls: $p=4\times10^{-6}$, $t=-8.3$, DOF=10) or scaled down (patients: $p=0.005$, $t=3.23$,
448 DOF=10; controls $p=0.0002$, $t=5.3$, DOF=10) their movement, commensurate with the
449 respective decrease or increase of the cursor Feedback Gain. We performed a two-way ANOVA,
450 confirming that the effect of Feedback Gain (0.65 vs. 1.35) on Scaling Factor was significant
451 ($F(1,40)=59.7$, $p=2\times10^{-9}$, $\eta_p^2=0.7$), whereas the effect of group (patient vs. control) on Scaling
452 Factor was not significant ($F(1,40)=0.05$, $p=0.83$), with no significant interaction between group
453 and gain ($F(1,40)=3.57$, $p=0.07$). We accounted for the samples being dependent on one another
454 because each subject was exposed to both the 0.65 and 1.35 gain conditions. The effect size was
455 computed using the Hentschke and Stüttgen toolbox (Hentschke and Stüttgen, 2011).
456
457 Note that the subjects typically exhibited a Scaling Factor of around 1.2 in the 1.35
458 Feedback Gain condition. Intuitively, this indicates that subjects rescaled their movements as if
459 to compensate perfectly for an experimentally applied Feedback Gain of 1.2; had they

compensated perfectly, their scaling factor would have been 1.35. Likewise, for a Feedback Gain of 0.65, the Scaling Factor was typically approximately 0.8, again implying that subjects compensated, albeit not fully, for the visual rescaling (Figure 3). Critically, the patients' responses to the feedback scaling were comparable to that of age-matched controls, suggesting that they were able to use visual feedback in a substantively similar manner.

Cerebellar Patient Performance Best Captured by Long Latency Closed-Loop Model

In computational terms, one can interpret the data from Experiment 1 to mean that cerebellar patients had a functional feedback loop in their control system. Here we asked if they are using the same "control structure" (i.e., a model of the interplay between sensory feedback, external sensory input, and motor output) as age-matched controls. Because the cerebellar and age-matched control groups scaled their movements similarly in Experiment 1, one might hypothesize that they were using a similar control structure. However, the scaling was measured relative to each subject's baseline movement in the veridical condition and was not a measure of their overall error performance on the task; for example, issues of phase lag would not be captured by the scaling analysis. Thus, here we use our findings from Experiment 1 along with the model selection procedure described by Madhav et al. (2013) to determine any differences in control structures used by patients and age-matched controls.

Consider the hypothesized model depicted in Figure 4A. In this model, the brain calculates the error between perceived elbow angle and target angle, as shown in the subtraction calculation in the feedback diagram. The model merges the cascade of the subject's internal controller and mechanical plant (arm and robot dynamics); the combined plant and controller is

treated as a classic McRuer gain-crossover model—a scaled integrator—a model based on the assumption that the subject is responding to an unpredictable stimulus (McRuer and Krendel, 1974). Our hypothesis is that all model parameters would be equivalent between patients and age-matched controls, except that patients would have a feedback delay commensurate with their visual processing time and controls would have a lower-latency feedback delay commensurate with their proprioceptive feedback processing time (Bhanpuri et al., 2013; Cameron et al., 2014; Crevecoeur and Scott, 2013; Izawa and Shadmehr, 2008). This is because we expect that patients would rely more on (slower) visual feedback to compensate for their deficient estimation of limb state. Visuomotor delay during smooth pursuit tracking is generally much faster than the time required for movement initiation, and estimates for such visuomotor tracking delay vary, but would be expected in the range of 110 to 160ms (Brenner and Smeets, 1997; Day and Lyon, 2000; Franklin and Wolpert, 2008; Haith et al., 2016; Pruszynski et al., 2016; Saunders and Knill, 2003).

To test this hypothesized model, we performed model fitting and model selection on all possible model configurations (Figure 4S1A, B). Our hypothesized model was just one possibility and the parameter values of the models were not constrained *a priori*. The models were compared based on joint consideration of model consistency and model fit. Model fitting and selection produced five models, called Best 4 (Lowest Variance), Best 4 (Lowest Error), Best 5 (Lowest Variance), Best 5 (Lowest Error), and Best 6, that were nearly equivalent in their trade-off between model inconsistency and model fitting error (Figure 4S1C). Model Best 6 added one more free parameter compared to Best 5 (Lowest Variance) and was the only one of the top 5 models that allowed for variation in the feedback delay for controls; this addition

yielded a small delay (~39ms) and while only providing a small enhancement in model fit-error, shown in both frequency domain plot and time domain validation (Figure 4S1C and Figure 4S1D). Thus, out of parsimony, we eliminated model Best 6 considered both Best 4 and both Best 5 models as the top models. These top models were extremely similar in their structure and parameters (Figure 4S1B and Table 2). Indeed, in all four top models, the feedback delay was zero for control subjects, substantially shorter than cerebellar patients. This suggests that controls can rely on an internal model of their hand position and/or proprioceptive feedback to make corrections for their future movements while cerebellar patients must rely on delayed visual feedback (see Discussion).

Rather than picking up a single best model, we drew the following general conclusions from the modeling. The Final Model (Figure 4B) was distilled from the top four models to show the general features among them, together with the range of parameters derived from these models. In addition to the parameter values having consistent values across the top four models, they also have intuitively plausible values (Table 2). While the true parameters have biological limitations, we did not restrict their values during the fitting procedure. This provides a useful diagnostic tool, since unrealistic parameter fits would indicate inadequate model structure: parameter bias is a hallmark of model deficiency. In the four top models, Visual Delays for controls were found to be 141-147 ms, which is physiologically plausible. Interestingly, this delay was much shorter than for patients, whose Visual Delays were 181-211 ms. Critically, patients exhibited longer response delay both on the visual measurement of target motion and that of self-movement feedback, compared to controls. The top models also suggest that Feedback Delay is shorter than Visual Delay for both patients and controls, although for two of top models, the patients exhibited equivalent Visual and Feedback Delays. Visual Gain values

529 are all approximately 0.40 which also seems reasonable given a visual inspection of the data:
530 subjects do not appear to be attempting to replicate the full magnitude of the signal, but some
531 smaller portion of the signal. We further examined the frequency responses of patients and
532 controls (Figure 4S2A). Patients exhibited substantially greater phase lag at high frequencies
533 than controls; at low frequencies, both populations exhibited very little phase lag, but,
534 surprisingly, controls exhibited slightly more phase lag than patients in this frequency range.
535 Thus, there was a phase "cross over" frequency between patients and controls, which was also
536 captured by our top models (Figure 4S2B).
537
538 As a final test of our modeling approach, we examined how our Best 4 (Lowest Err)
539 Model's parameters changed when we applied variations in Feedback Gain as shown in Table 3.
540 Recall that in the 1.35 Feedback Gain condition, the dot moves more than the person's hand
541 position; therefore, we expect the subject to move slightly less than they do in the veridical
542 feedback condition because they do not need to move as much to get the same visual output (i.e.,
543 we expect k to decrease). Similarly, we expect the subject to attempt to replicate a greater portion
544 of the signal because it is easier to do so. Thus, we expect the Visual Gain (the amount of the
545 input signal that the subject is trying to reproduce) to be greater in the 1.35 Feedback Gain
546 Condition than in the veridical condition. Recall that the Visual Gain affects the *input* to the
547 visual error computation. We expect the opposite trend for the 0.65 Feedback Gain condition, as
548 is detailed in Table 3. Given that we believe the delays are biologically limited based on
549 transmission time of visual information, we expect the delay magnitudes to stay the same
550 between Feedback Gain conditions. Indeed, the model parameters change in the hypothesized
551 manner when different Feedback Gains were applied. The results from fitting the Best 4 (Lowest

552 Err) Model to the data from each of the different Feedback Gains yields the parameter values
553 shown in Table 4. The visual feedback delays are nearly identical between Feedback Gain
554 conditions, and the k and Visual Gain values increase and decrease as hypothesized in Table 3.
555 The one difference between the hypothesis and these model values is in the lack of a feedback
556 delay for the control subjects. This discrepancy is addressed in the Discussion section.
557
558 Lastly, we used bootstrapping to generate confidence intervals for the parameter values in
559 our Best 4 (Lowest Err) Model. Confidence intervals (see Methods) for our best model to be:
560 2.1-3.4 for $k_{patient}$ and $k_{control}$; 0.32-0.46 for Visual $Gain_{patient}$ and Visual $Gain_{control}$; 174-261 ms
561 for Visual $Delay_{patient}$; and 122-161 ms for Visual $Delay_{control}$, and Feedback $Delay_{patient}$.
562
563 ## Experiment 2: Poor Tracking of Simple Oscillatory Trajectories Highlights Cerebellar
564 ## Patients' Predictive Deficit
565 We validated the sinusoidal tracking paradigm by quantifying a known behavioral deficit
566 in cerebellar visuomotor control. In this predictable task, we expected healthy participants would
567 be efficient, exerting less than or equal to the amount of effort required to minimize error for a
568 given phase lag. Where subjects lie relative to this tradeoff can be visualized on a phasor plot;
569 efficient tracking would yield points on or inside the effort/error tradeoff circle (Figure 2B,
570 Methods). We hypothesized that patients with impaired prediction may in some cases exert more
571 effort than needed to minimize error, resulting in phasor points outside the effort/error tradeoff
572 circle. We also hypothesized that patients would generate larger tracking errors.
573

At the five lowest frequencies tested, all of the control responses landed in/on the circle, whereas only 31% of patient responses landed in/on the circle (Figure 5). Additionally, patients generated larger tracking errors: 70% of patient responses exhibited larger error magnitude than the respective age-matched control responses. Furthermore, we created an aggregate tracking error for the single-sines experiment by calculating the sum of the tracking errors for the lowest five frequencies. We compared the aggregate tracking error between patients and age-matched controls by calculating the difference for each pair and testing those differences against 0 using a two-sided sign test. The null hypothesis for a two-sided sign test is that there was no difference between patients and their age-matched controls. The alternative hypothesis is that the controls may be either better or worse than patients. When we compared the difference between the aggregate tracking errors for each patient/control pair to 0, controls were significantly better than patients at single-sines tracking with 10 out of 11 controls performing better than their age-matched counterpart ($p=0.012$). These results illustrate cerebellar patients' poor predictive ability. Importantly, note that this analysis does not distinguish what portion of their errors stemmed from poor prediction of the stimulus trajectory versus poor control of the arm.

For comparison, we also included the results from the task where the stimulus was unpredictable (sum-of-sines from Experiment 1) in Figure 5. For this experiment, because neither group would be able to predict the target trajectory, we expected that patients would be less impaired relative to controls. Indeed, patient and control behaviors were more similar in this condition: for the five lowest frequencies tested in the sum-of-sines condition, 72% of control responses and 64% of patient responses were in/on the circle. Note, however, that there is an increased phase lag of the cerebellar patients (relative to controls), particularly as the frequency

of the stimulus increases, which would be expected if the patients were dependent primarily on time-delayed visual feedback (a pure feedback delay introduces greater phase lag at higher frequencies). Specifically, 73% of individual patient frequency responses were more phase lagged than the responses of the age-matched control (as computed on an individual frequency basis).

As expected, we also observed a reduction in tracking error from the sum-of-sines task to the single-sines task. We calculated the magnitude of the tracking error for each subject for each frequency (i.e., for the frequencies that were tested in both conditions) and then compared the errors between the conditions. 93% of controls' and 75% of patients' tracking errors were reduced in the single-sine task in comparison to his or her performance in the sum-of-sines task. Again, we computed the aggregate tracking error for the sum-of-sines task, as described earlier, using the same five frequencies used in the single-sines aggregate tracking error calculation. We then used a two-sided sign test to determine that patients ($p=9x10^{-4}$) and controls ($p=9x10^{-4}$) were better at tracking single-sines than sum-of-sines, with all patients and all controls having less tracking error for single-sines than sum-of-sines.

To specifically look at phase lag between groups (patients vs. controls) and conditions (single- vs. sum-of-sines), we used a circular statistical approach. Figure 6 shows polar representations of the phase lags from the single-sine and sum-of-sines conditions (Figures 6A, C) and an example time series from a control and cerebellar subject tracking the 0.85 Hz sine wave (Figures 6B, D). The example control subject showed little phase lag suggesting that this individual could make use of an internal prediction of the dot movement and their arm movement. In contrast, the example cerebellar subject showed a systematic phase lag suggesting

that their ability to use prediction is impaired. The pattern of lag can be visualized across frequencies in the polar plots. As a group, controls show a small increase in lag as the frequency increases in the single-sine condition, whereas the cerebellar group shows large lags (Figures 6A and C, compare purple vectors). Circular ANOVA for the single-sine condition showed that the cerebellar group had greater lags compared to controls (Group effect, $p=1x10^{-8}$), that the lags increased with frequency (Frequency effect, $p=7x10^{-11}$) and an interaction such that the cerebellar group lags increased more than controls across frequency (Interaction effect, $p=2x10^{-6}$). In the sum-of-sines condition, the difference in lags across frequency was qualitatively greater when comparing cerebellar and control groups (Figures 6A and C, compare black vectors). Overall the cerebellar group showed greater phase lags compared to controls (Group effect, $p=0.02$), and there was a statistically significant effect of Frequency (Frequency effect, $p=7x10^{-30}$). Group x Frequency interaction (Interaction effect, $p=0.03$), was less significant than that seen in the single-sine condition.

Experiment 3: Virtual Reality Feedback Can Reduce Dysmetria

The results from Experiment 1 indicated that cerebellar patients' feedback control may be largely intact. In Experiment 3, we leveraged this intact element of their control system to reduce dysmetria. Our goal was to provide each cerebellar patient with customized visual position feedback information that took into account their controller's mismatched feedforward model. This altered visual feedback would help them generate the correct motor command for a simple elbow movement (30 degrees flexion) by accounting for their actual arm dynamics.

Here we used a motor task that requires subjects make a discrete movement to a target that is stepped to a new position. This was chosen so that we could compare our data to

643 previously published results using the same task with cerebellar patients. Subjects were asked to
644 make a 30 degree elbow flexion movement from a home target to a stepped target. The task was
645 completed with both veridical visual feedback and with altered visual feedback that was
646 designed to reduce their dysmetria.
647
648 Bhanpuri et al. theorized that cerebellar patients have a static mismatch between their
649 controller's internal model of their limb inertia and the actual limb inertia for elbow flexion
650 movements (Bhanpuri et al., 2014). Critically, for single-joint elbow flexion or extension, an
651 inertial mismatch causes an acceleration-dependent error in the internal dynamic model because
652 inertia and acceleration are kinematically yoked. Thus, we predicted that an Acceleration-
653 Dependent Feedback Gain, k_a (block diagram inset in Figure 7A) could provide corrective
654 feedback to enhance reaching performance of cerebellar patients. Specifically, we took a
655 subject's estimated acceleration (see Methods), multiplied this by the Feedback Gain, and then
656 added it to the position of the subject's cursor:

$$\text{Shown Elbow Angle} = (\text{Actual Elbow Angle}) + k_a \cdot (\text{Elbow Angular Velocity})$$

657 We predicted that positive value for k_a would make a subject more hypermetric and
658 negative value would make a subject more hypometric. Thus, we expected hypometric patients
659 would experience a reduction in dysmetria with a positive gain. Similarly, we expected
660 hypermetric patients would experience a reduction in dysmetria when a negative gain was
661 applied. For a given value for the gain for k_a, subjects completed 30 degree elbow flexion
662 reaching movements in the Kinarm exoskeleton robot. Reaches were categorized as hypometric,
663 hypermetric, or on-target based on the angle where they made their first correction to their
664 movement. The Acceleration-Dependent Feedback Gain k_a was applied to the visual feedback

665 provided on the Kinarm screen. For more details on the implementation and instructions of this
666 task, see Methods.
667
668 When an appropriate Acceleration-Dependent Feedback Gain k_a was applied, dysmetria
669 was reduced. The effect of this gain on the displayed trajectory is shown in Figure 7A, where the
670 dashed green line shows the altered visual feedback given to the subject and the solid green line
671 shows the subject's actual reach to the target for a single Feedback Gain condition. Increasing
672 the magnitude of k_a had a graded effect on the trajectory profile, as shown by the decreasing
673 amplitude of the solid traces in Figure 7A. To find the "best" k_a we applied the PEST algorithm
674 (see methods). Applying this gain shifted the angle of first correction in the way we predicted, as
675 shown in both Figure 7B. Figure 7C shows all of the data collected during the implementation of
676 PEST as well as additional trials across a range of gains k_a (see Methods) that were collected for
677 a subset of patients (and all control subjects; see Figure 8). Critically, positive Acceleration-
678 Dependent Feedback Gain made patients more hypermetric and negative Acceleration-
679 Dependent Feedback Gain made patients more hypometric; likewise for control subjects. Thus,
680 subject-specific Acceleration-Dependent Feedback Gain was needed to best reduce dysmetria. In
681 general, increased Acceleration-Dependent Feedback Gain magnitude was needed to ameliorate
682 greater dysmetria.
683

684 DISCUSSION

685 We found that cerebellar patients can use visual feedback control in a manner similar to
686 control subjects. This was assessed during a sum-of-sines visuomotor tracking task by subtly
687 changing the gain of the visual cursor feedback (i.e. hand) and measuring the degree to which an

688 individual subject could modulate the gain of his or her corrective movements. Our modeling
689 corroborated this finding: the structure and parameters between patients and controls were
690 generally the same, except that the patient and control models differed substantively in their
691 feedback delay. The patients relied on time delayed cursor feedback of their hand position,
692 whereas controls appeared to be able to generate predictions of their hand position.
693 We also saw increases in phase lag for patients versus controls when following
694 predictable sine waves. While patients showed large phase lags that increased with frequency in
695 both conditions, controls only showed large phase lags in the sum-of-sines condition when the
696 stimulus was unpredictable. We interpret this to mean that the control subjects could predict the
697 movement of *both* the hand *and* stimulus in the single-sine condition, but the cerebellar patients
698 could not predict *either* as reliably as controls, relying more heavily on visual, moment-to-
699 moment feedback for tracking.
700 The above results are significant as it has previously been difficult to assess visual
701 feedback control in these patients due to their clear impairments in feedforward control of
702 movement. In Experiment 3, we tested whether patients' largely intact visual feedback could be
703 exploited to benefit their targeting performance in a reaching task. Indeed, we found that
704 cerebellar patients could leverage their intact, albeit delayed, feedback control ability to improve
705 their movement when we manipulate their visual feedback to assist them.
706

707 ## Delay in Patients but not in Healthy Controls

708 Cerebellar patients are thought to make faulty state estimates of arm motion due to a
709 damaged internal model, which adversely affects their proprioceptive estimates during active
710 movement (Bhanpuri et al., 2013). Thus, we expect them to increase their reliance on visual

feedback, which is time delayed compared to proprioceptive feedback. Here we hypothesized that our patients' feedback control system would have an increased delay in comparison to controls. This hypothesis is supported by work that showed that healthy subjects normally rely more on proprioception than vision for feedback control (Crevecoeur et al., 2016). Based on the Crevecoeur et al. (2016) model, increased variability of the proprioceptive information would lead to increased weighting of visual feedback (Crevecoeur et al., 2016; Weeks et al., 2017), which comes at the cost of increased delay for patients compared to controls.

Our modeling results reveal that patients have a delay in their feedback loop that approximates what one would expect from visual feedback, while controls appear to have negligible feedback delay. One interpretation of the control data is that controls are relying on a predictive internal model of their arm dynamics. This would allow them to make corrections even faster than they could using purely proprioceptive feedback. In other words, healthy controls can make corrections for current target movements based on a prediction of the present limb state from outdated measurements and known motor commands, rather than waiting on delayed sensory feedback.

We think that the cerebellar patients' deficits are best explained by their dependence on time-delayed visual feedback, though it is theoretically possible that delays in motor execution could contribute. Previous studies have shown that some patients with cerebellar damage exhibit longer than usual reaction time delays when asked to make discrete, rapid movements (Beppu et al., 1987, 1984; Day et al., 1998; Holmes, 1917; Vilis and Hore, 1980). These reaction time studies aimed to quantify a movement initiation delay as opposed to a feedback delay. Here, we

734 specifically assessed the timing of the in-flight trajectory alterations made while tracking an
735 unpredictable stimulus. While our feedback delay is a type of "reaction time" because this task
736 requires making corrections on the fly, the neural mechanisms underlying these corrections may
737 differ from those responsible for planning and initiation of a rapid movement.

738 Cerebellar patients can also show a reduced feedback gain in situations where responses
739 are driven by proprioception more than vision. Kurtzer et al. (Kurtzer et al., 2013) studied
740 cerebellar patients' ability to modulate long latency stretch responses (i.e. EMG responses within
741 45-100 ms of a muscle stretch) that depend on knowledge of intersegmental dynamics of the
742 arm. This was assessed by mechanically perturbing the elbow joint and measuring long latency
743 responses from elbow muscles and from shoulder muscles in anticipation of the passive shoulder
744 motion induced by elbow movement. Cerebellar patient responses were found to have a lower
745 gain and a slight delay in timing compared with controls. We interpret this in the context of the
746 different task demands between studies-- in that study, proprioception drove the response; here
747 vision appears to be more important to patients. These results combined suggest that cerebellar
748 subjects can reweight which feedback modality to rely on (proprioception versus vision)
749 depending on task demands, consistent with Block et al. (Block and Bastian, 2012), and may
750 reduce the sensorimotor gain on either modality to minimize the negative effects of delay.

751 Lastly, while the confluence of data and modeling presented here seems to indicate that
752 patients rely disproportionately on delayed feedback from vision for this visuomotor tracking
753 task, two of our top models (Model Best 5, Lowest Err and Best 4, Lowest Err; see Table 2)
754 actually exhibited shorter feedback delays than visual delays for patients. This suggests that the
755 patients' control systems in this task may include some reliance, albeit weaker, on prediction and
756 proprioception.

Modeling Results Support the Cerebellum as a Smith Predictor

Sensorimotor tasks are naturally analyzed using control systems theory and system identification (Roth et al., 2014), and because cerebellar patients' dysmetria resembles a poorly tuned control system, many researchers have analyzed cerebellar behavior using control theory (Luque et al., 2011; Manto, 2009; Miall et al., 1993; Porrill et al., 2013).

Our modeling results indicate that control subjects make corrections based on prediction of a future limb state measurement. While there is some controversy on this (Pelisson et al., 1986; Wolpert et al., 1998), previous work has proposed that the cerebellum might function in a manner consistent with a Smith Predictor (Miall et al., 1993). Essentially, a Smith Predictor is a control architecture that compensates for self-movement feedback delay. To achieve this requires a simulation of the plant in order to predict sensory consequences of motor output, and a simulated delay that "stores" the plant simulation for comparison with delayed feedback. In order to understand the consequences of incorporating a Smith Predictor, consider the generic feedback block diagram of sensorimotor control depicted in Figure 9 (top) and model the brain as either a controller *with* (Figure 9, left) or *without* (Figure 9, right) a Smith Predictor. Interestingly, upon simplification, we find the exact structures yielded by two of the top models (*Best 5 Lowest Var* and *Best 4 Lowest Var*): the patient model matches the simplified block diagram without a Smith Predictor and the age-matched control model matches the simplified block diagram which one. This suggests that an intact cerebellum may act in a manner qualitatively similar to a Smith Predictor and that patients ability to perform such prediction is impaired.

Perhaps the increased feedback delay in the cerebellar patients is unavoidable if one assumes the patients' control systems lack the essential capabilities of a Smith Predictor (making and storing predictive plant simulations). Thus, the apparent increased reliance on visual feedback seen in cerebellar patients may not be a compensatory mechanism at all—perhaps all subjects rely on delayed visual feedback, but healthy subjects disguise the delay using predictions of future measurements of limb state. Patients may simply lack this capability and it is not that they *compensate* for their deficit by relying *more* on visual feedback but instead that the loss of the sensory prediction and storage reveals the visual feedback delay inescapably present in both groups. This would provide some clarity on how cerebellar patients could "learn" to "compensate" in this way even when their cerebellum was damaged.

At its core, the Smith Predictor has to do with simulating the mechanical plant (and appropriate delay) in order to mitigate the effect of *feedback* delay. But notice that in our model there is a large visuomotor delay on the reference input even with the Smith Predictor in place. In order to compensate for the delayed visual measurement of the stimulus, the brain would need to be able to predict reference motions. To avoid this confound of stimulus prediction, and focus on feedback delay compensation, we fit our models using responses to unpredictable (sum-of-sine) target motions.

But, what if the target motion were predictable? Previously, it was shown in another species (weakly electric fish) that target predictability can indeed improve tracking performance, and it was hypothesized that this was based on an internal model of predictable reference motion

(Roth et al., 2011). Here, we extend this to healthy human subjects who show marked improvement in the single-sine (predictable) tracking task. Cerebellar patients exhibit much more modest improvement in tracking predictable stimuli, suggesting that they may have impairment in both their own plant model predictions and predictions of external sensory references (see Results, Experiment 2). Further data and experiments would be needed to better isolate the effects of cerebellar damage on these potentially disparate functions of the cerebellum.

Lastly, even if a healthy cerebellum helps "cancel" expected cursor feedback, this does not imply that visual cursor feedback goes unused: any errors in visual feedback, i.e., cursor self-motion feedback that is not precisely anticipated by the cerebellum, could be highly informative, and recent evidence (Yon et al., 2018) suggests that a predictive model of action–perception could heighten perceptual sensitivity, making self-motion (in this case, cursor) feedback more nuanced and precise.

There are potential alterative modeling interpretations of the data presented in this paper. For example, a state predictor with optimal state feedback (Crevecoeur and Gevers, 2019) is consistent with our observation that healthy subjects compensate for delay and is an avenue for further investigation. Also, our observation that visual target delay was longer in the patient model than the control-subject could conceivably be due to a different source of phase lag that our simplified modeling structure was not able to capture. Future experiments and analysis will investigate this further.

Potential for Real-World Application

In Experiment 3, we were successfully able to reduce the dysmetria of both the hypermetric and hypometric cerebellar patients in an elbow flexion task. Our visual feedback-control-based intervention was developed based on a control model which represented dysmetria as a mismatch between the patient's internal model of their limb inertia and the actual limb inertia (Bhanpuri et al., 2014). Thus, the success of our intervention provides support to this inertial mismatch theory. However, recent evidence based on the application of limb weights suggest that the inertial mismatch theory may be insufficient to explain dysmetria in multi-joint movements (Zimmet et al., 2018). Thus, feedback enhancement of multi-joint movements will likely require modifications of the simple acceleration-based controller tested in this paper. Moreover, the "best" acceleration-dependent feedback gain was patient specific even for the single-degree-of-freedom reaches tested for this study, suggesting that idiosyncrasies in each patient's motor control system (Kuling et al., 2017; Rincon-Gonzalez et al., 2011) may mandate patient-specific tuning. Such tuning will be made more complex for 2D and 3D arm movements.

Though our intervention successfully reduces the initial over or undershooting component of the reach, our method does not attempt to correct all aspects of dysmetria. Patients with cerebellar ataxia also typically exhibit increased movement variability. Application of this visual feedback gain does not attempt to reduce the variability of patient reaches. Secondly, our intervention does not completely eliminate the oscillations experienced by the subject after this first correction is made. Moreover, our modeling work does not characterize or explain this increased variability directly. Mild oscillations after the first correction can be seen in Figure 7A.

848 It is possible that adding a damping term to the feedback intervention might be able to address
849 this oscillatory behavior.
850
851 Our results provide an encouraging foundation for future studies because they show that
852 cerebellar patients are capable of using their own intact visual feedback control system to make
853 accurate reaches. Further study is needed to determine whether this intact feedback control
854 system could be leveraged therapeutically to reduce dysmetria without need for a virtual reality
855 system. Alternatively, it is possible that wearable sensory augmentation (for example, tendon
856 vibration or skin stretch) could be used as a surrogate for the vision-based sensory shift provided
857 here. These alternatives are especially important because drug therapy for cerebellar ataxia is not
858 currently viable, leaving rehabilitative and assistive therapies alone as the primary means of
859 treatment for these patients. In any case, we hope that by identifying these intact movement
860 control mechanisms we might help move treatment possibilities forward.
861

ACKNOWLEDGEMENTS

National Institute of Child Health and Human Development HD040289 to AJB and NJC

Applied Physics Lab Graduate Fellowship FNACCCX19 to AMZ

National Science Foundation 1825489 to NJC and AJB

Thanks to James Freudenberg, Brent Gillespie, Adrian Haith, Matthew Statton, and Sarah Stamper for helpful comments and discussion. Thanks to Amanda Therrien for recruitment of patients, clinical assessments, and feedback throughout this process.

COMPETING INTERESTS

We have no competing financial interests to disclose.

REFERENCES

Adamovich S V., Berkinblit MB, Fookson O, Poizner H. 1998. Pointing in 3D Space to Remembered Targets. I. Kinesthetic Versus Visual Target Presentation. *J Neurophysiol* **79**:2833–2846.

Bares M, Lungu O, Liu T, Waechter T, Gomez CM, Ashe J. 2007. Impaired predictive motor timing in patients with cerebellar disorders. *Exp Brain Res* **180**:355–365. doi:10.1007/s00221-007-0857-8

Beppu H, Nagaoka M, Tanaka R. 1987. Analysis of Cerebellar Motor Disorders By Visually-Guided Elbow Tracking Movement 2. Contribution of the Visual Cues on Slow Ramp Pursuit. *Brain*.

Beppu H, Suda M, Tanaka R. 1984. Analysis of Cerebellar Motor Disorders by Visually Guided Elbow Tracking Movement. *Brain* **107**:787–809.

Berens P. 2009. CircStat : A MATLAB Toolbox for Circular Statistics. *J Stat Softw* **31**. doi:10.18637/jss.v031.i10

Bhanpuri NH, Okamura AM, Bastian AJ. 2014. Predicting and correcting ataxia using a model of cerebellar function. *Brain* **137**:1931–1944. doi:10.1093/brain/awu115

Bhanpuri NH, Okamura AM, Bastian AJ. 2013. Predictive Modeling by the Cerebellum Improves Proprioception. *J Neurosci* **33**:14301–14306. doi:10.1523/JNEUROSCI.0784-13.2013

Block HJ, Bastian AJ. 2012. Cerebellar involvement in motor but not sensory adaptation. *Neuropsychologia* **50**:1766–1775. doi:10.1016/j.neuropsychologia.2012.03.034

Box GEP, Draper NR. 1987. Empirical Model-Building and Response Surfaces. John Wiley & Sons, Inc.

Brenner E, Smeets JBJ. 1997. Fast responses of the human hand to changes in target position. *J Mot Behav* **29**:297–310. doi:10.1080/00222899709600017

Broersen R, Onuki Y, Abdelgabar AR, Owens CB, Picard S, Willems J, Boele H-J, Gazzola V, Van der Werf YD, De Zeeuw CI. 2016. Impaired Spatio-Temporal Predictive Motor Timing Associated with Spinocerebellar Ataxia Type 6. *PLoS One* **11**. doi:10.1371/journal.pone.0162042

Cameron BD, de la Malla C, López-Moliner J. 2014. The role of differential delays in integrating transient visual and proprioceptive information. *Front Psychol* **5**. doi:10.3389/fpsyg.2014.00050

Crevecoeur F, Gevers M. 2019. Filtering Compensation for Delays and Prediction Errors during Sensorimotor Control. *Neural Comput* **31**:738–764. doi:10.1162/neco_a_01170

Crevecoeur F, Munoz DP, Scott SH. 2016. Dynamic Multisensory Integration : Somatosensory Speed Trumps Visual Accuracy during Feedback Control. *J Neurosci* **36**:8598–8611. doi:10.1523/JNEUROSCI.0184-16.2016

Crevecoeur F, Scott SH. 2013. Priors Engaged in Long-Latency Responses to Mechanical

910 Perturbations Suggest a Rapid Update in State Estimation. *PLoS Comput Biol* **9**:1003177.
911 doi:10.1371/journal.pcbi.1003177

912 Day BL, Lyon IN. 2000. Voluntary modification of automatic arm movements evoked by motion
913 of a visual target. *Exp Brain Res* **130**:159–168. doi:10.1007/s002219900218

914 Day BL, Thompson PD, Harding a. E, Marsden CD. 1998. Influence of vision on upper limb
915 reaching movements in patients with cerebellar ataxia. *Brain* **121**:357–372.
916 doi:10.1093/brain/121.2.357

917 Franklin DW, Wolpert DM. 2008. Specificity of reflex adaptation for task-relevant variability. *J*
918 *Neurosci* **28**:14165–14175. doi:10.1523/JNEUROSCI.4406-08.2008

919 Fuentes CT, Bastian AJ. 2010. Where Is Your Arm? Variations in Proprioception Across Space
920 and Tasks. *J Neurophysiol* **103**:164–171. doi:10.1152/jn.00494.2009

921 Haith AM, Pakpoor J, Krakauer JW. 2016. Independence of movement preparation and
922 movement initiation. *J Neurosci* **36**:3007–3015. doi:10.1523/JNEUROSCI.3245-15.2016

923 Hentschke H, Stüttgen MC. 2011. Computation of measures of effect size for neuroscience data
924 sets. *Eur J Neurosci*. doi:10.1111/j.1460-9568.2011.07902.x

925 Holmes G. 1917. The Symptoms of Acute Cerebellar Injuries Due to Gunshot Injuries **40**:461–
926 535. doi:http://dx.doi.org/10.1093/brain/40.4.461

927 Izawa J, Shadmehr R. 2008. On-line processing of uncertain information in visuomotor control. *J*
928 *Neurosci* **28**:11360–11368. doi:10.1523/JNEUROSCI.3063-08.2008

929 Kuling IA, Marieke ·, Van Der Graaff CW, Brenner E, Jeroen ·, Smeets BJ. 2017. Matching
930 locations is not just matching sensory representations. *Exp Brain Res* **235**:533–545.
931 doi:10.1007/s00221-016-4815-1

932 Kurtzer I, Trautman P, Rasquinha RJ, Bhanpuri NH, Scott SH, Bastian AJ. 2013. Cerebellar
933 damage diminishes long-latency responses to multijoint perturbations. *J Neurophysiol*
934 **109**:2228–41. doi:10.1152/jn.00145.2012

935 Luque NR, Garrido JA, Carrillo RR, Coenen OJMD, Ros E. 2011. Cerebellarlike corrective
936 model inference engine for manipulation tasks. *IEEE Trans Syst Man, Cybern Part B*
937 *Cybern* **41**:1299–1312. doi:10.1109/TSMCB.2011.2138693

938 Madhav MS, Stamper S a, Fortune ES, Cowan NJ. 2013. Closed-loop stabilization of the
939 Jamming Avoidance Response reveals its locally unstable and globally nonlinear dynamics.
940 *J Exp Biol* **216**:4272–4284. doi:10.1242/jeb.088922

941 Manto M. 2009. Mechanisms of human cerebellar dysmetria: experimental evidence and current
942 conceptual bases. *J Neuroeng Rehabil* **6**. doi:10.1186/1743-0003-6-10

943 Manto M, Godaux E, Jacquy J. 1994. Cerebellar hypermetria is larger when the inertial load is
944 artificially increased. *Ann Neurol* **35**:45–52. doi:10.1002/ana.410350108

945 McRuer DT, Krendel ES. 1974. Mathematical Models of Human Pilot Behavior. Paris.

946 Miall RC, Christensen LOD, Cain O, Stanley J, Rothwell J. 2007. Disruption of State Estimation

947 in the Human Lateral Cerebellum. *PLoS Biol* **5**:2733–2744.
948 doi:10.1371/journal.pbio.0050316

949 Miall RC, Jackson JK. 2006. Adaptation to visual feedback delays in manual tracking: Evidence
950 against the Smith Predictor model of human visually guided action. *Exp Brain Res* **172**:77–
951 84. doi:10.1007/s00221-005-0306-5

952 Miall RC, Weir DJ, Wolpert DM, Stein JF. 1993. Is the cerebellum a smith predictor? *J Mot*
953 *Behav* **25**:203–216. doi:10.1080/00222895.1993.9942050

954 Morton SM, Bastian AJ. 2006. Cerebellar Contributions to Locomotor Adaptations during
955 Splitbelt Treadmill Walking. *J Neurosci* **26**:9107–9116. doi:10.1523/JNEUROSCI.2622-
956 06.2006

957 Paillard J, Brouchon M. 1974. A proprioceptive contribution to the spatial encoding of position
958 cues for ballistic movements. *Brain Res* **71**:273–284. doi:10.1016/0006-8993(74)90971-8

959 Pelisson D, Prablanc C, Goodale MAA, Jeannerod M. 1986. Visual control of reaching
960 movements without vision of the limb II. Evidence of fast unconscious process correcting
961 the trajectory of the hand to the final position of a double-step stimulus. *Exp Brain Res*
962 **62**:303–311. doi:10.1007/BF00238849

963 Porrill J, Dean P, Anderson SR. 2013. Adaptive filters and internal models: Multilevel
964 description of cerebellar function. *Neural Networks* **47**:134–149.
965 doi:10.1016/j.neunet.2012.12.005

966 Poulton EC. 1974. Tracking Skill and Manual Control. Academic Press.

967 Pruszynski JA, Johansson RS, Flanagan JR. 2016. A rapid tactile-motor reflex automatically
968 guides reaching toward handheld objects. *Curr Biol* **26**:788–792.
969 doi:10.1016/j.cub.2016.01.027

970 Rincon-Gonzalez L, Buneo CA, Tillery H. 2011. The Proprioceptive Map of the Arm Is
971 Systematic and Stable, but Idiosyncratic. *PLoS One* **6**:25214.
972 doi:10.1371/journal.pone.0025214

973 Roth E, Sponberg S, Cowan NJ. 2014. A comparative approach to closed-loop computation.
974 *Curr Opin Neurobiol* **25**:54–62. doi:10.1016/j.conb.2013.11.005

975 Roth E, Zhuang K, Stamper SA, Fortune ES, Cowan NJ. 2011. Stimulus predictability mediates
976 a switch in locomotor smooth pursuit performance for Eigenmannia virescens. *J Exp Biol*
977 **214**:1170–1180. doi:10.1242/jeb.048124

978 Saunders JA, Knill DC. 2003. Humans use continuous visual feedback from the hand to control
979 fast reaching movements. *Exp Brain Res* **152**:341–352. doi:10.1007/s00221-003-1525-2

980 Smith M, Brandt J, Shadmehr R. 2000. Motor disorder in Huntington's disease begins as a
981 dysfunction in error feedback control. *Nature* **403**:544–549. doi:10.1038/35000576

982 Smith MA, Shadmehr R. 2005. Intact Ability to Learn Internal Models of Arm Dynamics in
983 Huntington's Disease But Not Cerebellar Degeneration. *J Neurophysiol* **93**:2809–2821.
984 doi:10.1152/jn.00943.2004.

Taylor MM, Creelman CD. 1967. PEST: Efficient Estimates on Probability Functions. *J Acoust Soc Am* **41**:782–787. doi:10.1121/1.1910407

Trouillas P, Takayanagi T, Hallett M, Currier RD, Subramony SH, Wessel K, Bryer A, Diener HC, Massaquoi S, Gomez CM, Coutinho P, Hamida M Ben, Campanella G, Filla A, Schut L, Timann D, Honnorat J, Nighoghossian N, Manyam B. 1997. International Cooperative Ataxia Rating Scale for pharmacological assessment of the cerebellar syndrome. *J Neurol Sci* **145**:205–211. doi:10.1016/S0022-510X(96)00231-6

Tyryshkin K, Coderre AM, Glasgow JI, Herter TM, Bagg SD, Dukelow SP, Scott SH. 2014. A robotic object hitting task to quantify sensorimotor impairments in participants with stroke. *J Neuroeng Rehabil* **11**. doi:10.1186/1743-0003-11-47

Vilis T, Hore J. 1980. Central neural mechanisms contributing to cerebellar tremor produced by limb perturbations. *J Neurophysiol* **43**:279–291.

Weeks HM, Therrien AS, Bastian AJ. 2017. Proprioceptive Localization Deficits in People with Cerebellar Damage. *Cerebellum* **16**:427–437. doi:10.1007/s12311-016-0819-4

Wolpert DM, Miall RC, Kawato M. 1998. Internal models in the cerebellum. *Trends Cogn Sci* **2**:338–347. doi:10.1016/S1364-6613(98)01221-2

Yon D, Gilbert SJ, de Lange FP, Press C. 2018. Action sharpens sensory representations of expected outcomes. *Nat Commun*. doi:10.1038/s41467-018-06752-7

Zimmet AM, Cowan NJ, Bastian AJ. 2018. Patients with Cerebellar Ataxia Do Not Benefit from Limb Weights. *The Cerebellum*. doi:10.1007/s12311-018-0962-1

Zimmet, Amanda M.; Cao, Di; Bastian, Amy J.; Cowan, Noah J., 2020. "Data associated with the publication: Cerebellar patients have intact feedback control that can be leveraged to improve reaching", *Johns Hopkins University Data Archive*, https://doi.org/10.7281/T1/BCARLC, V1.

CODE/DATA REPOSITORY

The dataset and code used in this paper is publicly available via the Johns Hopkins University Data Archive at https://doi.org/10.7281/T1/BCARLC

FIGURES

Figure 1 – Experimental Setup

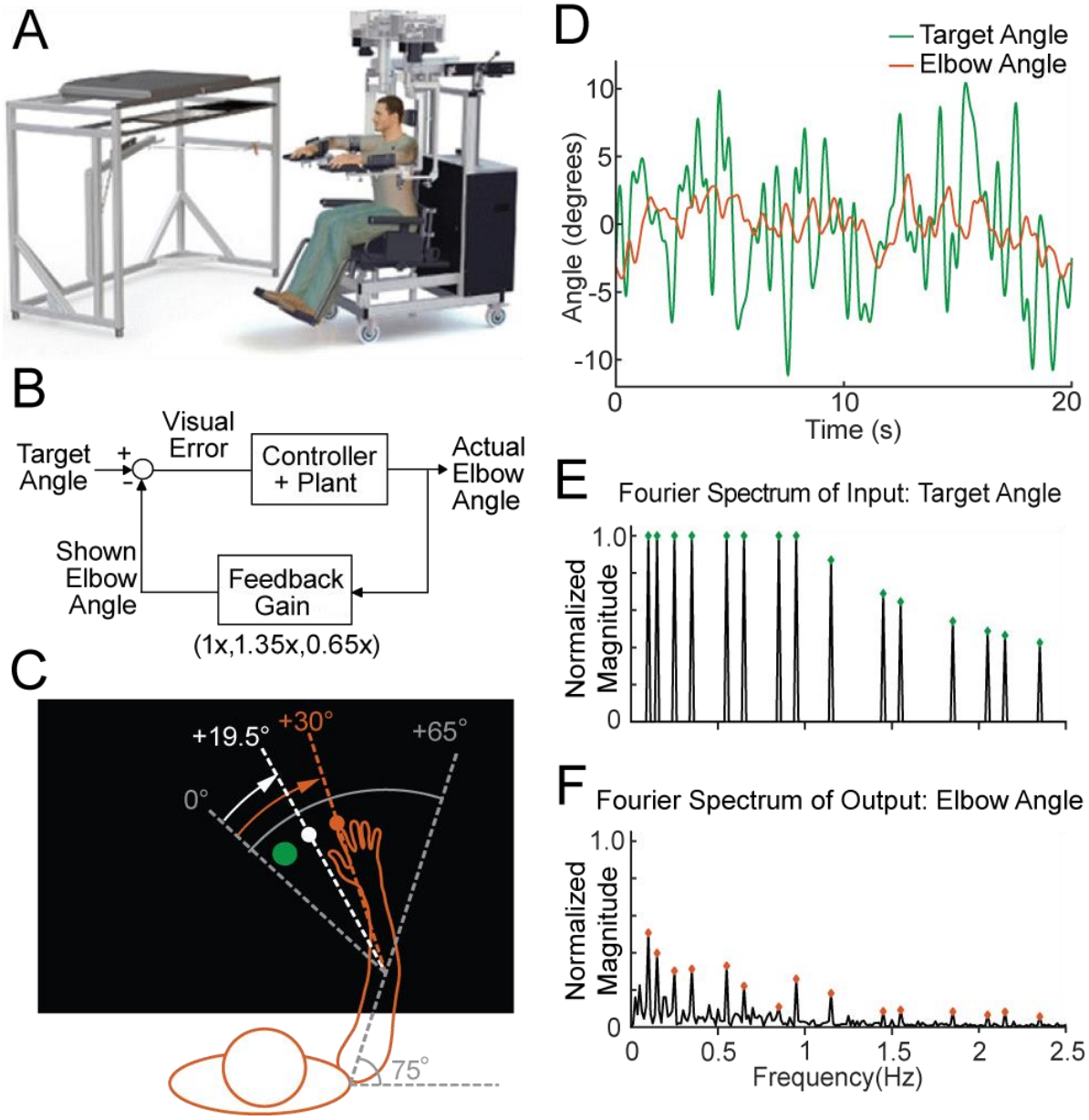


Figure 1 *The experimental paradigm*. **A** The Kinarm Exoskeleton robot system (BKIN Technologies Ltd.) where all tasks are executed. Image is taken from (Tyryshkin et al., 2014). **B** This block diagram shows our experimental paradigm for experiment 1. The actual elbow angle (orange in C) is multiplied by a feedback gain to compute the shown elbow angle (white dot in C). **C** A schematic view of the Kinarm screen in the 0.65 Feedback Gain condition. The subject's shoulder is locked at a 75° angle. The green dot is the target dot. The green dot oscillates according to the sum-of-sines pattern along an arc with a radius equivalent to the forearm + hand + finger length. The sum-of-sines pattern is centered about the 0° angle. This 0° centerline is at a 65° angle from the subject's straight arm position. The subject could not see their fingertip (orange dot) or arm position. The angle of the white dot (visible to the subject) is equal to 0.65 times the actual fingertip angle in orange (0.65 · 30° = 19.5°). **D** Sample data (20 s) for a single cerebellar patient showing the angles traversed by the target dot during the sum-of-sines task (green) and by the fingertip (orange, averaged over 5 trials). **E** The normalized Fourier spectrum of the input sum-of-sines signal (green trace in D). **F** The spectrum of the response by a single cerebellar patient over five trials (orange trace in D).

Figure 2 – Interpreting Phasor Plots

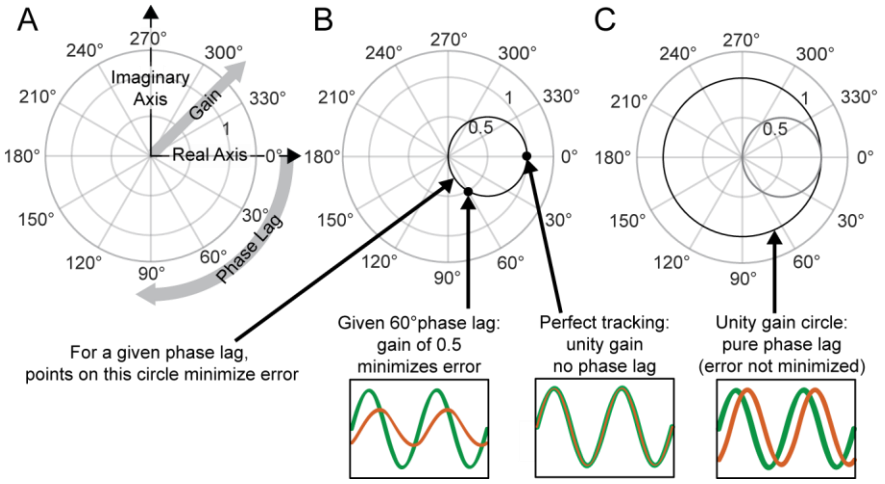


Figure 2 *Interpreting Phasor Plots*. **A-C** Brief explanation of phasor plots, which are explained in depth in Methods in "Phasor Plot." **A** Illustration of the relationship between the real and imaginary components of the complex numbers and the gain and phase lag. **B** The bold black circle illustrates the gains that minimize error given a particular phase lag (veridical feedback condition only). **C** Points on the bold black circle illustrate a different control strategy than what was asked of our subjects. Points on this circle replicate the input signal without minimizing error between the input and output signals (veridical feedback condition only) due to the phase lag.

Figure 3 – Motion Rescaling under Different Feedback Gains

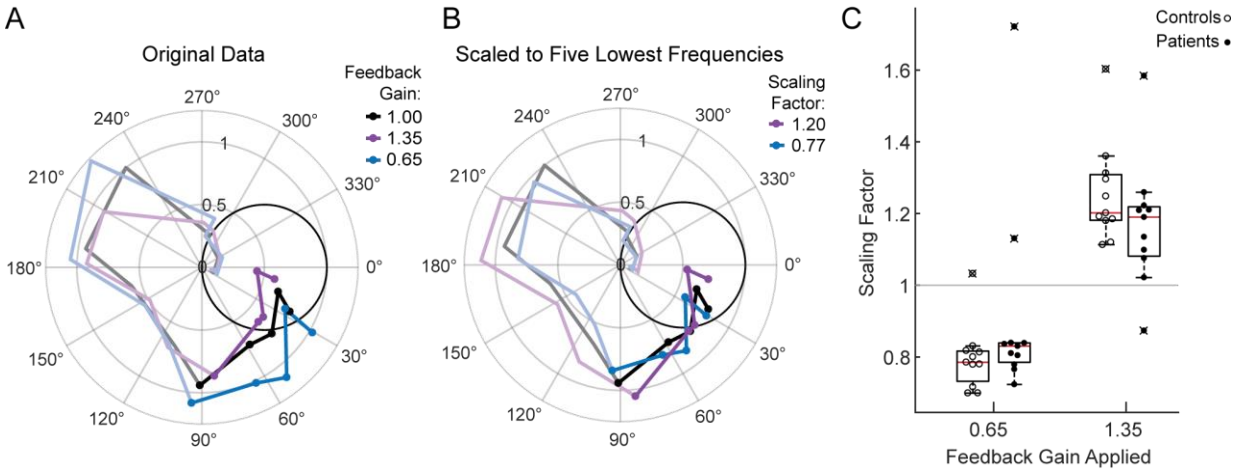


Figure 3 *Motion Rescaling under Different Feedback Gains*. **A** Sample phasor plot data from a single cerebellar patient under all three feedback gain conditions. The dots represent the frequency response at the lowest five frequencies. The solid line traces through the responses at all tested frequencies. The dimmed lines show the frequency responses at the other frequencies tested beyond the five lowest frequencies. **B** Scaled phasor plot data from the same single cerebellar patient. The gain value from D has been multiplied by the scaling factor shown in the legend here to create an overlaid phasor plot. The higher the Scaling Factor, the less effort (in terms of movement magnitude) the subject expended in that condition. **C** Patients and controls respond similarly to the applied Feedback Gain, indicating that they incorporate visual feedback similarly. The higher the Scaling Factor, the less effort (in terms of movement magnitude) the subject expended in that condition. Outliers are marked with an x. The outlier subject with the highest Scaling Factor in the 0.65 condition was from the same subject who had the lowest Scaling Factor in the 1.35 condition.

Figure 4 – Modeling

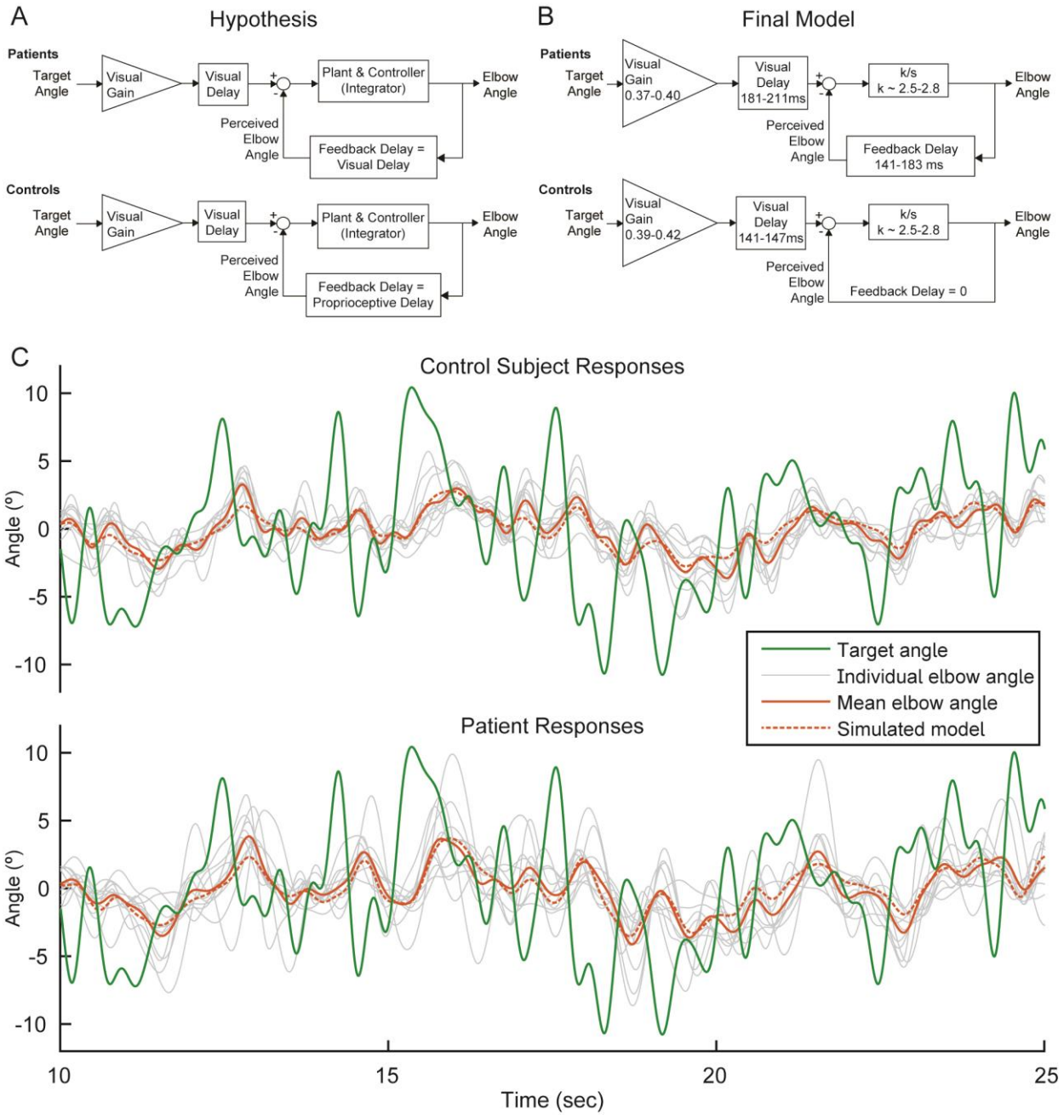


Figure 4 *Modeling.* **A** Our modeling framework was based on the McRuer Crossover model (McRuer and Krendel, 1974). This simple model structure lumps the controller and plant as a scaled integrator, and assigns different delays to the measurement of the target from that of self-movement feedback. We hypothesized that the patient and control models would be equivalent except for the magnitude of the feedback delay. The Visual Gain is necessary because, given the complexity and speed of the signal, participants were unable to match the full magnitude of the signal. **B** After our model selection procedure, the Final Model structure was distilled from the top models to capture the general features shared among them. It is similar to our hypothetical model, but with subtle differences: the feedback delay for the controls was determined to be zero instead of equivalent to the proprioceptive delay, the feedback delay for patients was shorter than or equal to the visual delay, and the visual delay for patients was slightly longer than healthy controls. **C** Time domain visualization of subject and model responses for a typical 15 second time window of our experiments. The same visual target trajectory (green) was played to all patients and controls, and was used in simulation. Individual time-domain elbow-angle responses (light gray) are shown for the 11 control subjects (top plot) and 11 patients (bottom plot). The average subject responses (orange, solid) match with reasonable accuracy

1061 the simulated model responses (orange, dashed). The simulated model structure used here is the modeled named “Best 4
1062 (Lowest Err)” in Figure 4S1B.

Figure 4 Supplemental 1 – Model Fitting and Model Selection

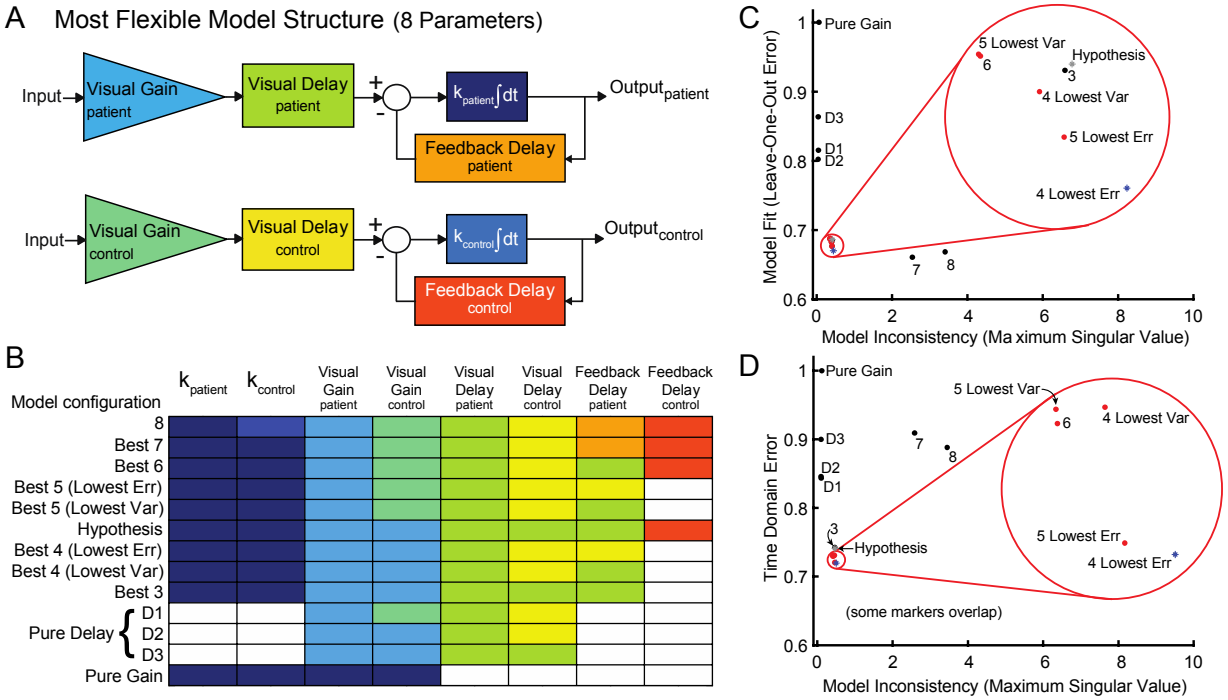


Figure 4S1 *Model Fitting and Model Selection*. We used sum-of-sine data to fit the parameters of a modified McRuer Crossover model (McRuer and Krendel, 1974). **A** In the most flexible setting, there are a total of eight free parameters, allowing for patients and controls to have completely independent model parameters. **B** In the model fitting and model selection process, we tested all possible variations of the model structure, for a total of 2667 model configurations. For example, some model structures yoke parameters together or eliminate parameters, and some yoke variables together across patient and control models. Each row illustrates an individual model configuration. Variables with the same color within that row are yoked together for that model configuration and white indicates that the specified variable was removed from the model. The table includes the following: the best model structures for all possible numbers of free parameters (Best 1-8); the hypothesized model structure (Hypothesis, grey asterisk shown in C & D); two special pure delay (D1-2) model structures. Best 1 is a pure gain model structure. Best 2 is a pure delay (D3) model structure. The model names are shown in the left column. Lowest Err means lowest Leave-One-Out error. Lowest Var means lowest model inconsistency. **C** Results of model selection procedure in frequency domain for the models in B. The models in red dots and blue asterisk, highlighted in the red circle, provide the best trade-off between minimizing model inconsistency (i.e., minimizing parameter variability and over-fitting) and minimizing error (improving model fit). The top four models are labeled 4 Lowest Err, 4 Lowest Var, 5 Lowest Err and 5 Lowest Var. These top models had nearly equivalent performance according to the model selection criteria. **D** Time domain validation. As in Figure 4S1C, we plotted error versus model inconsistency, but here we used the time-domain error (see Methods). Critically, the time-domain data were not used directly for model fitting, and yet the top models in left bottom corner of the frequency domain are also the top models in the time domain, validating that our frequency-domain model selection process selected the models that were also best in the best time-domain.

Figure 4 Supplemental 2 –Frequency Responses

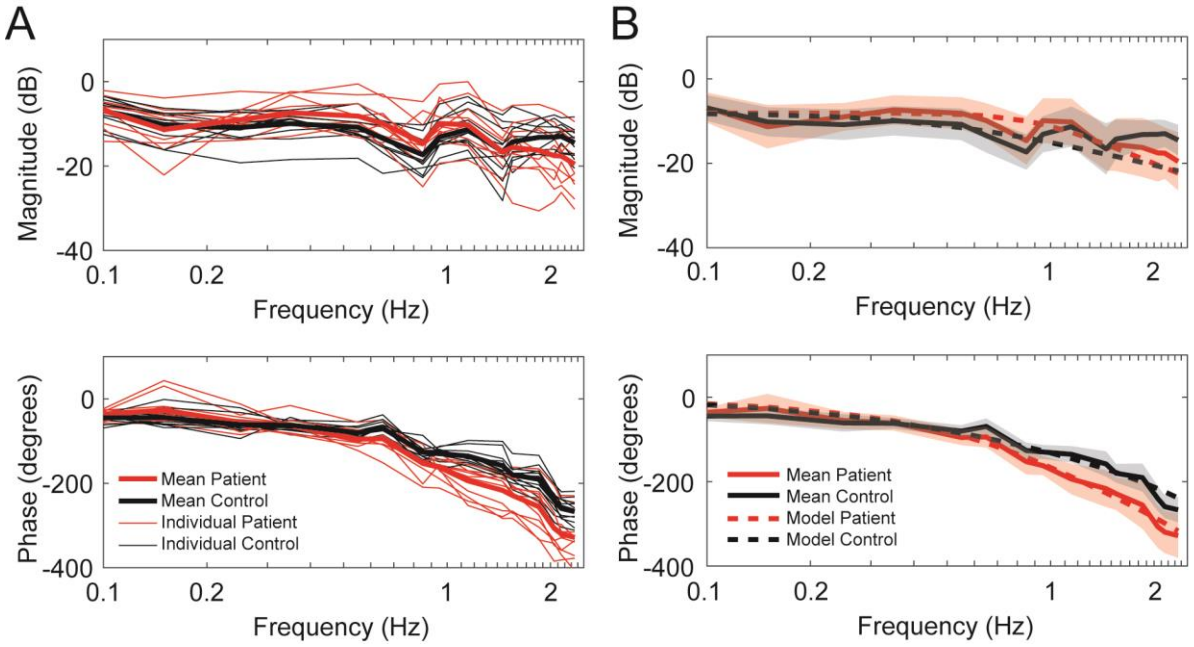
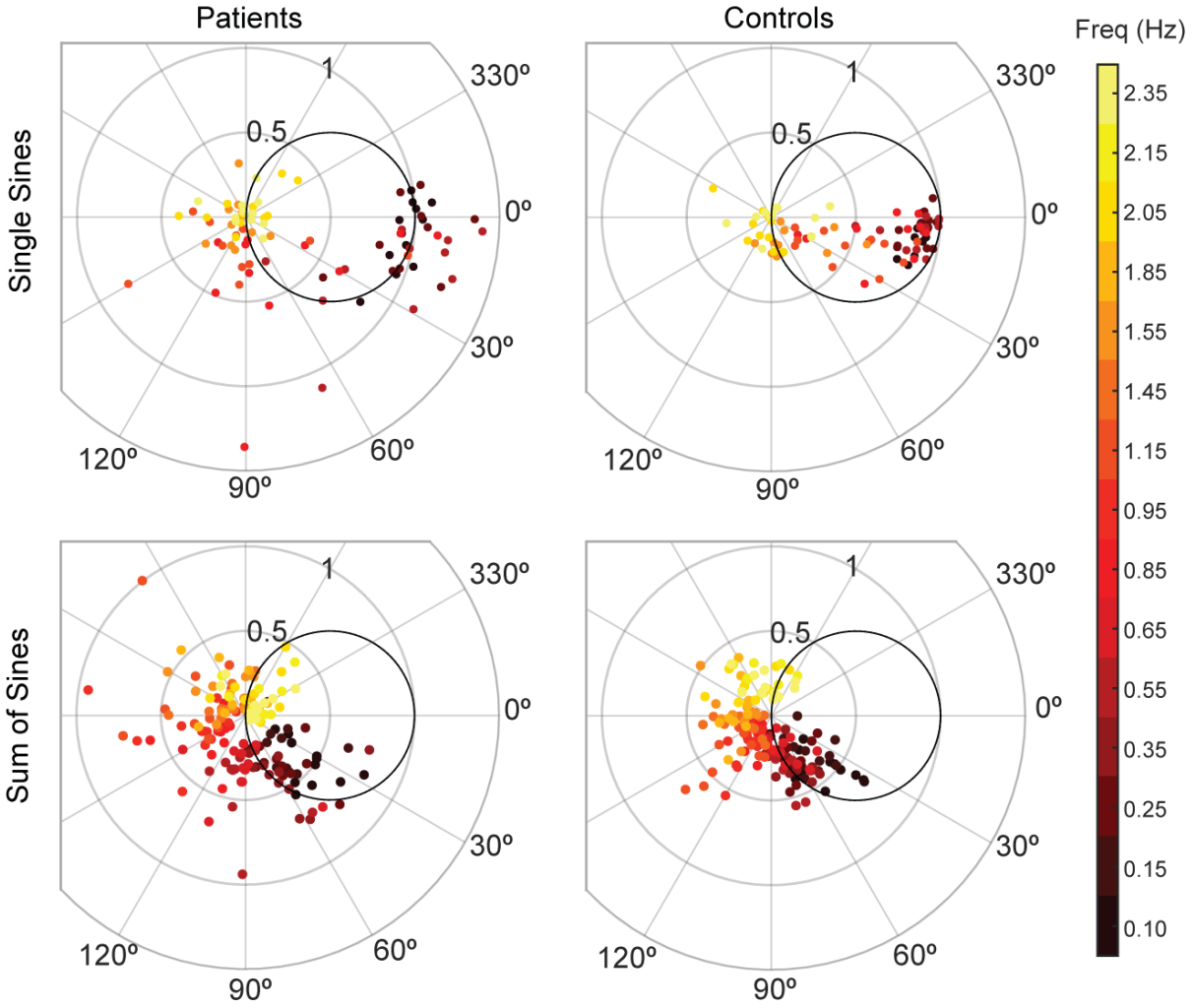


Figure 4S2 *Frequency Responses*. **A** Individual frequency responses for the 11 control subjects (black) and 11 patients (red). It also shows the average responses for controls (black bolded) and patients (red bolded). **B** Mean responses with standard deviation for controls (black bold line with grey shaded region) and patients (red bold line with red shaded region). The simulated model frequency responses (dashed lines) match the actual mean responses well. The frequency responses of the simulated models capture qualitative features, such as the crossing of the response curves between patients and controls both in the magnitude and phase. The simulated model here is Best 4 (Lowest Err).

1094 Figure 5 –Single-Sine and Sum-of-Sines Phasor Plots



1095
1096 Figure 5 *Patients and controls have categorically different responses to single-sine and sum-of-sine stimuli.* Each data point is
1097 the response at a single frequency (magnitude and phase) of a single subject. The patient data in both the single and sum-of-
1098 sines conditions are more variable than that of the control subjects. In the single-sines condition, control subjects are able to
1099 consistently remain inside the effort/error circle at low frequencies, balancing the tradeoff between effort and error, while
1100 patients are not. This result is consistent with previous studies that show patients have impaired prediction.

1101

Figure 6 – Comparison of Phase Lags

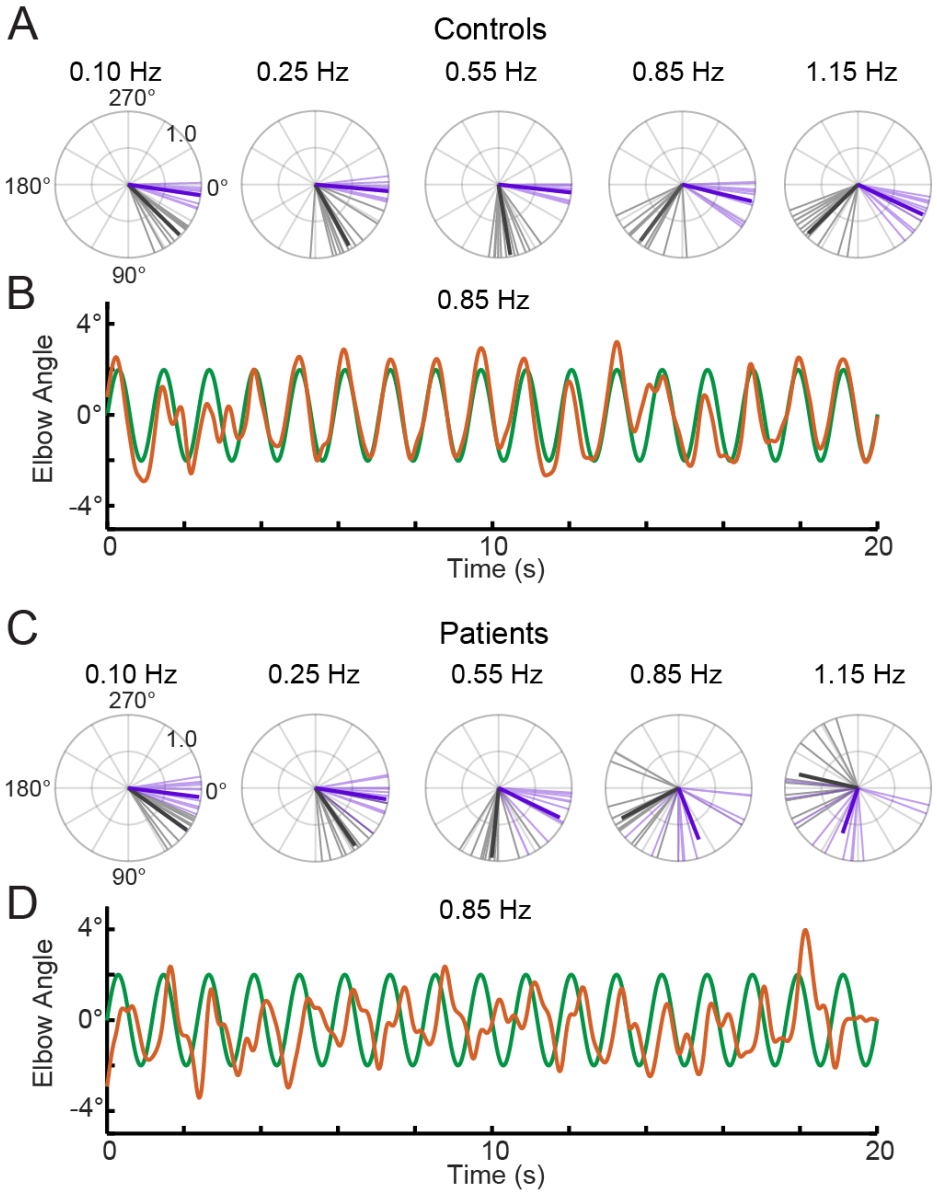


Figure 6 *Phase lags show different patterns between groups.* Polar plot representations showing phase lag for the 5 lowest frequencies common to the single-sine and sum-of-sines conditions. **A.** Control group phase lags. Individual subjects are represented as lighter colored unit vectors and the group mean vector is plotted in bold. Purple indicates the single-sine condition and black represents the sum-of-sines condition. Note that only phase lag is represented on the polar plot (gain is not represented). The Control group is able to track the single-sine stimuli with little or no phase lag, but shows phase lags that increase with frequency in the sum-of-sines condition. **B.** Example response from a control subject tracking a 0.85 Hz single-sine stimuli. The target sine wave is represented in green and the subject performance in orange. Note that this control could predict the single-sine and track it with little or no phase lag. **C.** Cerebellar group phase lags plotted as in A. The Cerebellar patients were able to track the two slowest single-sine stimuli with little phase lag (purple, 0.10 and 0.25 Hz) but then shows increasing phase lags as a function of frequency. Note that 0.10 and 0.25 Hz are extremely slow frequencies that could be followed using only visual feedback control. In the sum-of-sines condition, the cerebellar group shows phase lags that increase with frequency. **D.** Example response from a cerebellar subject tracking a 0.85 Hz single-sine stimuli. In contrast to the control subject, this cerebellar patient shows phase lags relative to the target.

1118

Figure 7 – Acceleration-Dependent Feedback Reduces Dysmetria

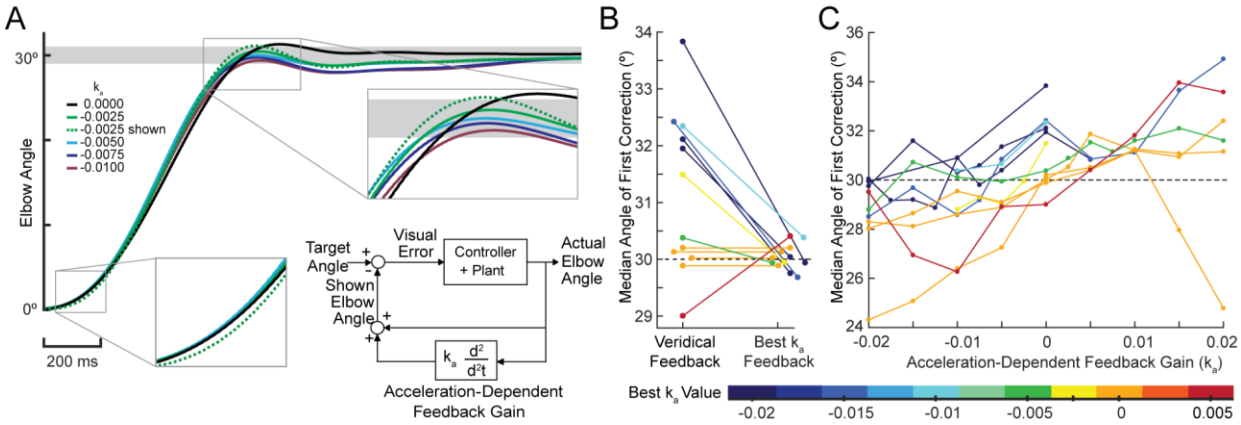


Figure 7 *Acceleration-dependent feedback reduces dysmetria*. **A** The trajectories from a single cerebellar patient, moving to the grey target zone. Overshooting or undershooting the grey region is categorized as hypermetric or hypometric, respectively. The solid lines are average trajectories (across trials for a single subject) at a given Acceleration-Dependent Feedback Gain value. A decrease in the angle of first correction (i.e., the peak of the first peak of the elbow angle) is apparent as the Acceleration-Dependent Feedback Gain decreases. The best Acceleration-Dependent Feedback Gain for this subject was determined to be -0.0025, and the resulting movement trajectory is shown as a solid green line. The dashed green line illustrates the trajectory that is displayed to the subject on the screen for the -0.0025 condition. This example also shows that oscillation around the target remains with Acceleration-Dependent Feedback Gain. The block diagram in the lower right shows how the Acceleration-Dependent Feedback Gain was applied to the visual representation of the fingertip position. **B** The reduction in dysmetria exhibited by patients when the best Acceleration-Dependent Feedback Gain was applied. Each line represents a single cerebellar patient. Lines are color coded to indicate the best Acceleration-Dependent Feedback Gain for that subject. **C** Increasing Acceleration-Dependent Feedback Gain causes increased hypermetria. Again, each line represents a single cerebellar patient. At the highest gain values, the visual feedback diverged enough from the fingertip position so that some subjects paused mid-reach, presumably due to the conflicting visual feedback. This yielded the observed drop in median angle of first correction seen in some subjects in the higher gain conditions. This figure shows only the cerebellar patients. Control subject data is shown in Figure 8.

Figure 8 – Acceleration-Dependent Feedback: Control Subjects

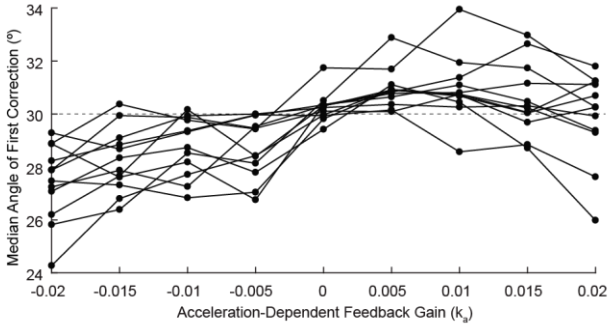
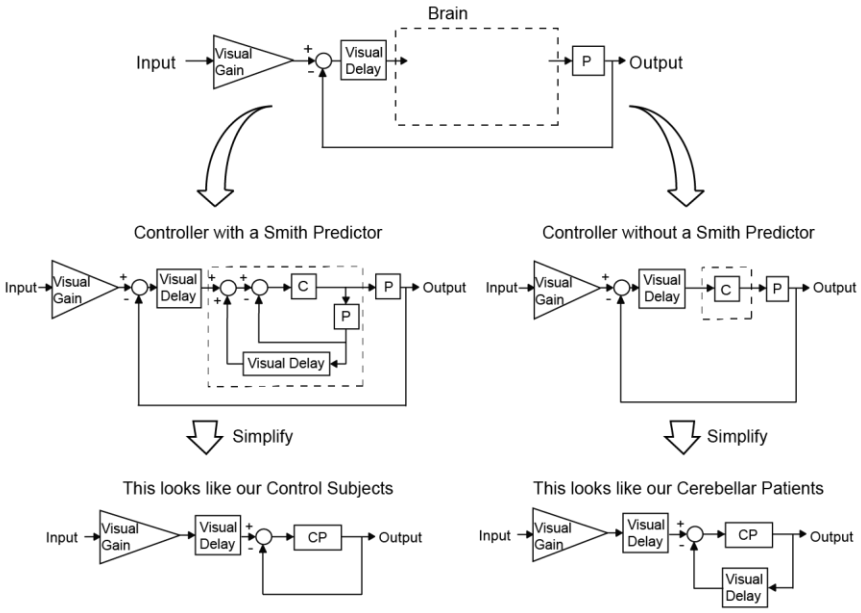


Figure 8 *Acceleration-dependent feedback also affects single reaches in control subjects.* Control subjects' median angle of first correction when completing thirty-degree flexion movements. The expected upward trend in median angle of first correction is visible across subjects. At the largest Acceleration-Dependent Feedback Gains, there is a slight drop off in median angle of first correction. At high gain values, the discrepancy between the cursor and the fingertip became more apparent, causing some subjects to pause mid-reach, resulting in a lower median angle of first correction for these high gain conditions.

1144 Figure 9– Smith Predictor



1145

1146 Figure 9 *Our modeling results agree with the theory of the cerebellum as a Smith Predictor*. A basic model for the control system
1147 is shown in the top panel. In our experiment, the plant, P, represents the entire mechanical arm system, including the robot
1148 exoskeleton arm. A single visual delay block delays the input from both the reference signal and the feedback information. A
1149 visual gain scales the amount of the input the subject will attempt to replicate. The brain, illustrated by a dashed box, can be
1150 modeled as containing just a controller, C, as shown on the right side. Or, the brain can be modeled as a controller with an
1151 accompanying Smith Predictor, as shown on the left side. When these two alternate structures are simplified, we are left with
1152 the exact model structures yielded by our previous modeling results.

TABLES

Subject No.	Patient Age	Sex	Pathology	ICARS	Upper Limb ICARS	Experiments
P1	44	M	SCA8	62	17	1,2,3
P2	52	M	ADCAIII	28	8	3
P3	53	M	Sporadic	59	19	1,2,3
P4	55	F	SCA8	41	17	3
P5	55	M	OPCA	46	14	1,2
P6	60	M	MSA-C	63	14	1,2,3
P7	62	F	Sporadic	36	16	1,2
P8	62	M	SCA6/8	62	19	1,2,3
P9	63	M	SCA6	41	10	3
P10	64	F	SCA6	58	13	1,2
P11	64	M	ADCAIII	11	3	1,2,3
P12	65	M	Idiopathic	34	10	3
P13	66	F	SCA6	41	11	1,2
P14	67	F	Left Cerebellar Stroke	27	10	1,2,3
P15	69	F	ADCAIII	52	13	3
P16	72	F	SCA6	49	18	3
P17	78	M	Sporadic (Vermis Degen.)	39	6	1,2

Table 1 The cerebellar patient population that was tested. Not all subjects completed all experiments. Some subjects completed the experiments over two visits. Upper Limb ICARS contains the sum of the upper-limb kinetic function elements of the test (ICARS subsections 10-14, out of 20).

Model	$k_{patient}$	$k_{control}$	Visual $Gain_{patient}$	Visual $Gain_{control}$	Visual $Delay_{patient}$ (ms)	Visual $Delay_{control}$ (ms)	Feedback $Delay_{patient}$ (ms)	Feedback $Delay_{control}$ (ms)
Best 5 (Lowest Err)	2.8 (0.09)	2.8 (0.09)	0.38 (0.01)	0.41 (0.02)	211 (0.008)	143 (0.003)	143 (0.003)	0
Best 5 (Lowest Var)	2.6 (0.08)	2.6 (0.08)	0.37 (0.01)	0.42 (0.02)	183 (0.005)	147 (0.003)	183 (0.005)	0
Best 4 (Lowest Err)	2.7 (0.10)	2.7 (0.10)	0.39 (0.01)	0.39 (0.01)	210 (0.007)	141 (0.003)	141 (0.003)	0
Best 4 (Lowest Var)	2.5 (0.09)	2.5 (0.09)	0.40 (0.01)	0.40 (0.01)	181 (0.005)	145 (0.003)	181 (0.005)	0

Table 2 Model parameters with standard deviation for the top models (Best 5 and Best 4). Variables with the same color within a row are yoked together for that model configuration. Top models are similar; the feedback delay was smaller than or equal to the visual delay for the patients, and zero for controls, and the patient visual delay was longer than that of the healthy controls.

Feedback Gain	k	Visual Gain	Visual Delay (s)	Feedback $Delay_{patient}$ (s)	Feedback $Delay_{control}$ (s)
1.0		<1	Visual Delay	Visual Delay	Proprioceptive Delay
1.35	Decrease	Increase	Unchanged	Unchanged	Unchanged
0.65	Increase	Decrease	Unchanged	Unchanged	Unchanged

Table 3 This is our hypothesis of how we expect the model parameter values indicated in the bottom two rows (1.35 and 0.65 Feedback Gain) will change (with respect to the value obtained in the veridical Feedback Gain condition, Feedback Gain=1.0) with changing Feedback Gains. We expect Visual Delay of approximately 140 to 250 ms, and Proprioceptive Delay approximately 110 to 150 ms (Cameron et al., 2014). We also expect Visual Gain to be less than one in the veridical feedback condition because we do not expect the subjects to attempt to replicate the full magnitude of the signal, given that it is a challenging, fast, and unpredictable task.

Feedback Gain	$k_{patient}$	$k_{control}$	Visual $Gain_{patient}$	Visual $Gain_{control}$	Visual $Delay_{patient}$ (ms)	Visual $Delay_{control}$ (ms)	Feedback $Delay_{patient}$ (ms)	Feedback $Delay_{control}$ (ms)
1	2.7	2.7	0.39	0.39	210	141	141	0
1.35	2.2	2.2	0.44	0.44	209	140	140	0
0.65	3.6	3.6	0.33	0.33	210	144	144	0

Table 4 The model (Best 4 Lowest Err) parameters changed with different applied Feedback Gains in a manner consistent with our prediction.