bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

De novo learning versus adaptation of continuous control in a manual tracking task

Christopher S. Yang[1*], Noah J. Cowan[2], and Adrian M. Haith[3]

[1]The Solomon H. Snyder Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD, USA.

[2]Department of Mechanical Engineering, Johns Hopkins University Whiting School of Engineering, Baltimore, MD, USA.

[3]Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, USA.

[*]email: christopher.yang@jhmi.edu

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Abstract

Learning to perform feedback control is critical for learning many real-world tasks that involve continuous control such as juggling or bike riding. However, most motor learning studies to date have investigated how humans learn *feedforward* but not *feedback* control, making it unclear whether people can learn new continuous feedback control policies. Using a manual tracking task, we explicitly examined whether people could learn to counter either a 90° visuomotor rotation or mirror-reversal using feedback control. We analyzed participants' performance using a frequency domain system identification approach which revealed two distinct components of learning: 1) adaptation of baseline control, which was present only under the rotation, and 2) *de novo* learning of a *continuous* feedback control policy, which was present under both rotation and mirror reversal. Our results demonstrate for the first time that people are capable of acquiring a new, continuous feedback controller via *de novo* learning.

Introduction

Many real-world motor tasks require continuous feedback control in order to perform skillfully. For example, when one is riding a bicycle, one must assess the environment and the state of the body to inform future motor commands that will minimize the chance one will topple off the bicycle. But when one learns how to ride a bicycle, does one learn a new feedback control policy to do so? Most studies investigate motor learning within the context of discrete, point-to-point movements which can be executed by predominantly using feedforward control. Therefore, while it may seem intuitive that one should be able to learn a new feedback controller, little is known about whether this is actually the case.

Humans can learn to perform new motor tasks through a variety of learning processes [1]. One of the most well-characterized mechanisms is adaptation, an error-driven learning process by which task performance is improved by experiencing and subsequently reducing sensory prediction errors [2–4]. Adaptation is known to support learning in a variety of laboratory settings including under imposed visuomotor rotations [5–7], prism goggles [8,9], split-belt treadmills [10,11], and force fields [12,13]. However, adaptation is known to be limited in the extent to which it can change behavior [6,14–16], suggesting that it cannot account for learning more complex motor skills. It has therefore been suggested that more complex skills must be learned by building a new control policy from scratch, a process termed *de novo* learning (Figure 1A) [17–19]. However, the exact mechanisms and properties of *de novo* learning remain unclear.

A simple visuomotor perturbation that is thought to require *de novo* learning is a mirror reversal of visual feedback. After having learned to compensate for mirror-reversed feedback, participants do not exhibit

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

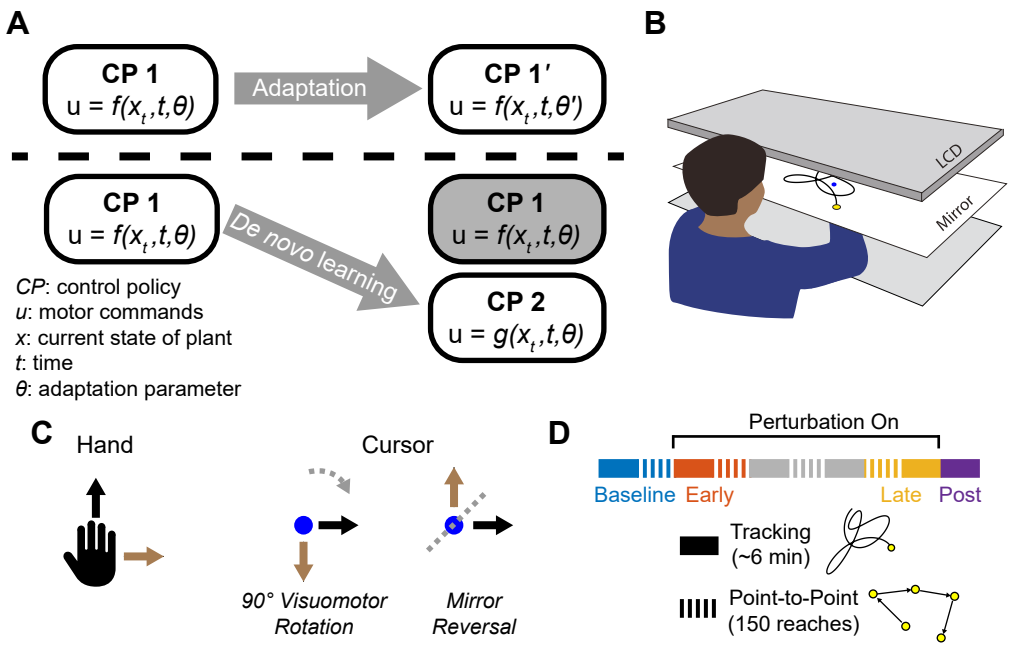

Figure 1. Conceptual overview and experimental design. **A.** We conceptualize adaptation as a parametric change to an existing control policy (changing θ to θ') and *de novo* learning as building a new control policy (g) to replace the baseline control policy (f). **B.** Participants performed planar movements with their right hand while targets were presented to them on a mirror-reversed display. In these tasks, there is a visual target (yellow) that participants are asked to track with their cursor (blue). **C.** Participants learned to control the cursor under one of two visuomotor perturbations: a 90° visuomotor rotation, or a mirror reversal. **D.** Participants were trained on their respective perturbation in both a point-to-point reaching task (1 block = 150 reaches) and a sum-of-sinusoids tracking task (1 block = ~6 mins), the latter of which allowed us to assess their feedback control capabilities. We first measured baseline performance on both tasks under veridical feedback (blue), followed by interleaved tracking and point-to-point blocks with perturbed feedback from early learning (orange) to late learning (yellow). We removed the perturbations post-learning to assess aftereffects (purple).

aftereffects [20, 21] that are characteristic of adaptation [13, 22, 23], suggesting that learning occurs through a qualitatively distinct mechanism. Furthermore, work in patient populations has suggested a dissociable neural basis for these two types of learning, with cerebellar dysfunction impairing learning in adaptation-based tasks (e.g., visuomotor rotations) but not *de novo* tasks (e.g. mirror reversal) and vice versa for basal ganglia disorders [20, 24–26].

To date, *de novo* learning has only been examined in point-to-point movements [18, 27], and the extent to which learning to counter a mirror reversal genuinely reflects learning of a new feedback control policy is unclear. Reaching to static targets under a mirror reversal could be achieved without learning any new control policy by simply re-aiming one's movements to a surrogate target opposite the mirroring axis [7]. Indeed, while people can easily learn to make accurate point-to-point movements under mirror-reversed feedback, they remain unable to generate appropriate feedback corrections during movement [18, 28, 29]. In contrast, adaptation to rotated visual feedback generalizes strongly to online feedback corrections [18, 30, 31].

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

To compensate for mirror reversal, are participants capable of learning a new feedback control policy *de novo*, or do they primarily rely on a feedforward re-aiming strategy?

Here, we used a manual tracking task to more comprehensively assess whether participants could learn to continuously control a cursor under either a visuomotor rotation (engaging adaptation) or a mirror reversal (requiring *de novo* learning). Participants were primarily exposed to the perturbation during point-to-point movements but, in periodic assessment blocks, were asked to track a target moving in a pseudorandom sum-of-sinusoids trajectory (Figure 1B-C) [32–36]. The fact that the target moved continuously and unpredictably precluded participants from engaging in time-consuming deliberate planning of the kind associated with strategic re-aiming [6, 37–39]. Participants instead had to continuously send motor output to track the target from moment to moment. We found that participants were capable of tracking the target under the mirror reversal and their feedback control performance improved with learning. Moreover, by analyzing different frequencies of hand movement, we found that these participants exhibited qualitatively different behavior from those who learned a visuomotor rotation. Ultimately, our results suggest humans are capable of acquiring new feedback control policies via *de novo* learning.

Results

Trajectory-based analyses

We recruited twenty participants for this study. Participants used their right hand to manipulate an on-screen cursor under either a 90° visuomotor rotation ($n = 10$) or a mirror reversal ($n = 10$) about an oblique 45° axis (Figure 1). These perturbations were designed such that, in both cases, motion of the hand in the x direction was mapped to cursor motion in the y direction and vice versa. Each group first practiced moving under their respective perturbation in a *point-to-point task*, reaching towards targets that appeared at random locations on the screen (Figure 1D), and we quantified their performance using initial reach-direction error. For the rotation group, this error decreased as a function of training time and plateaued near 0°, suggesting that participants successfully learned to compensate for the rotation (Figure 2A). For the mirror-reversal group, there was no clear learning curve using this simple measure, but performance was better than would be expected if participants had continued to use their baseline controller (which would manifest as reach errors spanning a range from −180° to 180°), indicating learning. Thus, both groups of participants appeared to compensate for the perturbations during point-to-point movements, consistent with previous findings.

To test the extent to which participants could learn a new feedback policy under these two perturba-

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

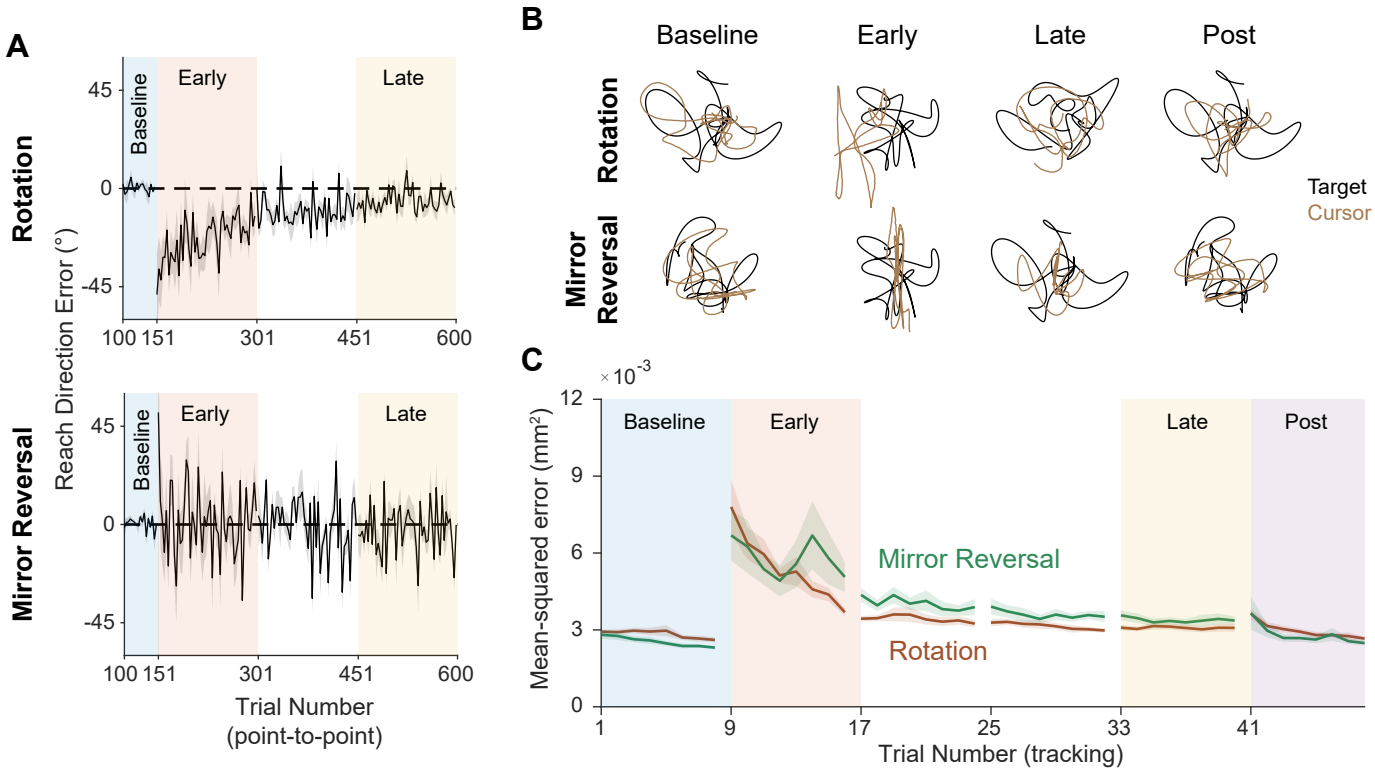

Figure 2. Performance in the point-to-point and tracking tasks. **A.** Performance in the point-to-point task, as quantified by reach-direction error, is plotted for the rotation (top) and mirror-reversal groups (bottom). The learning curves were smoothed by binning the trials into sets of 3 consecutive trials. Data from only the last 51 reaches of the baseline point-to-point block are shown (trials 100-150). Trials 1-150 occurred between the baseline and early tracking blocks, and trials 151-600 occurred between the early and late tracking blocks. **B.** Example tracking trajectories from one participant in each group. Target trajectories are shown in black while cursor trajectories are shown in brown. Each trajectory displays approximately 5 seconds of movement. **C.** Performance in the tracking task as quantified by average mean-squared positional error between the cursor and target during each trial. All error bars are SEM across participants.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

tions, we periodically assessed their performance in a tracking task with a continuously moving target. We administered this task before, during, and after blocks of practice of point-to-point movements under the perturbation. In the tracking task, participants followed a target that moved in a continuous sum-of-sinusoids trajectory at frequencies ranging between 0.1–2.15 Hz, with distinct frequencies used for x- and y-axis target movement. The resulting target motion appeared random. Furthermore, the target's trajectory was altered every block by randomizing the phases of the component sinusoids, preventing participants from learning a specific target trajectory.

As an initial assessment of learning, we measured the average mean-squared positional error (tracking error) between the target and cursor trajectories during every tracking trial. Example trajectories from single participants are presented in Figure 2B. Both groups of participants improved their tracking error over time (Figure 2C). Moreover, the tracking error was similar between both groups of participants by late learning.

We estimated how participants translated target motion into hand motion by aligning the hand and target trajectories through a linear transformation that, when applied to the target trajectory, minimized the discrepancy between the hand and target trajectories. We also included provision for possible delays between the target and hand trajectories in this analysis (estimated at approximately 380 ms; see Methods for details). At baseline, this transformation matrix should resemble the identity matrix, as the hand trajectory should be well aligned with the target trajectory. Indeed, at baseline, the transformation matrices resembled identity matrices for both groups of participants (Figure 3A). After learning, the transformation matrix should resemble the inverse of the matrix representing the applied perturbation. Figure 3A shows the estimated transformation matrices for both groups at different time points during the experiment, along with a visualization of how they affected the unit x and y vectors. These transformation matrices began to resemble the ideal transformation by late learning for both groups.

To analyze learning statistically, we focused on the off-diagonal elements of the matrices, which critically distinguish the different transformations from one another and from baseline. In late learning, both the rotation (linear mixed effects model: main effect of block ($F(2, 36) = 220.78$, $p < 0.0001$; Tukey's range test: $p < 0.0001$) and mirror-reversal groups (Tukey's range test: $p < 0.0001$) exhibited off-diagonal values that were significantly different from their baseline values (Figure 3B), and in the appropriate direction to compensate for their respective perturbations.

From these matrices, we computed specific metrics associated with each perturbation to further characterize learning. For the rotation group, we estimated the angular compensation by fitting a pure rotation matrix to the estimated transformation matrices (Figure 3C) to obtain a compensation angle θ. At baseline, we found that $\theta = 2.3 \pm 1.4°$, and this increased to $\theta = 69.6 \pm 1.7°$ by late learning. For the mirror-reversal

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

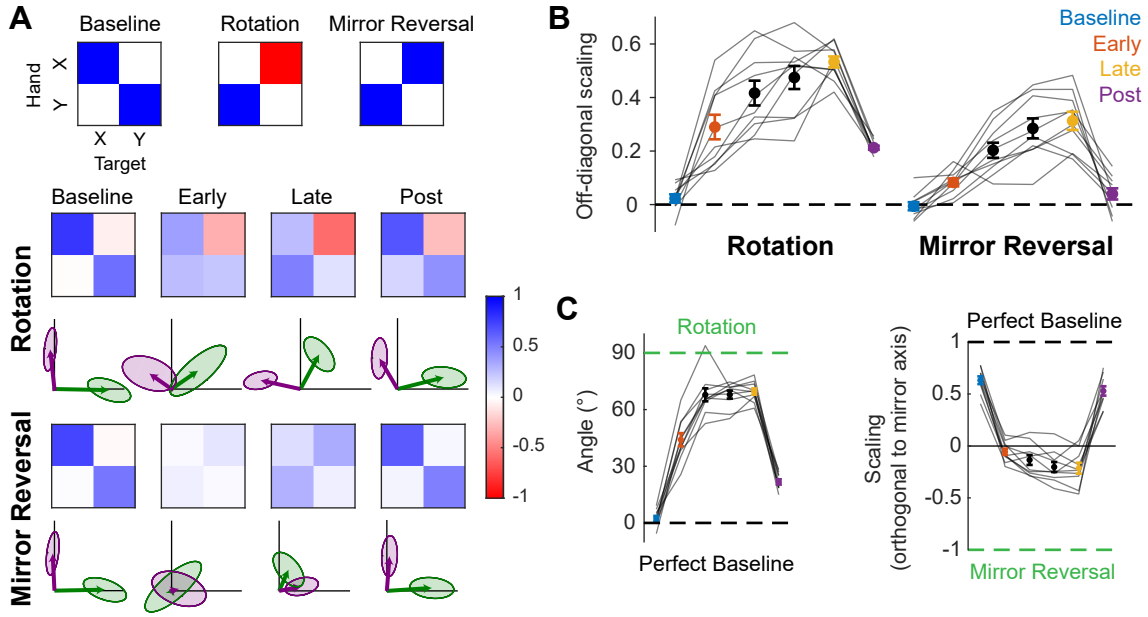

Figure 3. Trajectory-alignment analysis of tracking task. **A.** Transformation matrices relating target and hand movement. The top row illustrates the ideal transformation matrices at baseline or to successfully compensate for each perturbation. The rotation (middle row) and mirror-reversal (bottom row) groups' transformation matrices evolved during learning (averaged across participants). Each matrix can also be visualized in terms of how it transforms the unit x and y vectors (equivalent to the columns of the matrix), plotted below each matrix (x = green, y = purple).Unit x and y vectors are shown for scale. Shaded areas are 95% confidence ellipses across participants. **B.** The average of the two off-diagonal elements of the estimated transformation matrices across all blocks of the experiment (baseline through post). Black lines indicate individual participants, and circles and error bars indicate mean and SEM across participants. **C.** (Left: rotation group) Angular compensation for the rotation, estimated by approximating each transformation matrix with a pure rotation matrix. (Right: mirror-reversal group) Scaling factor orthogonal to the mirror axis. In each plot, dashed lines depict ideal performance when the perturbation is (green) or is not (black) applied. Black lines indicate individual participants and circles and error bars indicate mean and SEM across participants.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

group, we computed the scaling of the target trajectory in the direction orthogonal to the mirror axis (Figure 3C) to assess whether participants learned to flip the direction of their movements in response to target motion in this direction. This value was positive at baseline and negative by late learning, indicating that participants successfully inverted their hand trajectories relative to that of the target.

At the end of the experiment, we tested for aftereffects by repeating the tracking task one more time, but with the perturbation removed (and with participants made explicitly aware of this). We again estimated transformation matrices for this block and found that the two visuomotor perturbations had different effects on participants (linear mixed effects model: 2-way interaction between group and block, $F(2, 36) = 11.9686$). The off-diagonal elements for the rotation group were significantly different from baseline (Tukey's range test: $p < 0.0001$), indicating the presence of aftereffects. This corresponded to a compensation angle of $\theta = 21.7 \pm 1.4°$, similar to the magnitude of aftereffects reported for visuomotor rotation in point-to-point tasks [16, 40]. For the mirror-reversal group, by contrast, the off-diagonal elements of the post-learning transformation matrix were not significantly different from baseline (Tukey's range test: $p = 0.6036$; baseline range: -0.06–0.10; post-learning range: -0.06–0.14), suggesting negligible aftereffects.

These data support the idea that participants were able to compensate for the applied perturbation in the more challenging tracking task, but achieved this via qualitatively different mechanisms for the two different perturbations.

Frequency-domain analyses

The transformation matrices estimated based on trajectory alignment suggest that the mirror-reversal group was capable of learning a feedback control policy. However, this analysis has limitations. One limitation is that the mean-squared tracking error we calculate is biased towards large-amplitude deviations of the target which occurred relatively slowly (at 0.1–0.65 Hz). This could potentially have allowed the tracking task to be solved through intermittent "catch-up" movements that were strategically planned similar to an explicit re-aiming strategy for point-to-point movements, rather than through establishing a new feedback control policy qualitatively similar to the one used at baseline. The design of the tracking task made it amenable to a more more fine-grained analysis of participants' behavior using frequency-domain system identification, owing to the construction of the target trajectory as a sum of sinusoids at different frequencies.

In order to apply system identification tools to analyze behavior, we first sought to validate whether behavior was approximately linear. We did so by examining the amplitude spectra of participants' hand movement after transforming the time-domain data to the frequency domain using a discrete Fourier transform and comparing these amplitude spectra to that of the target. A linear relationship between hand and

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

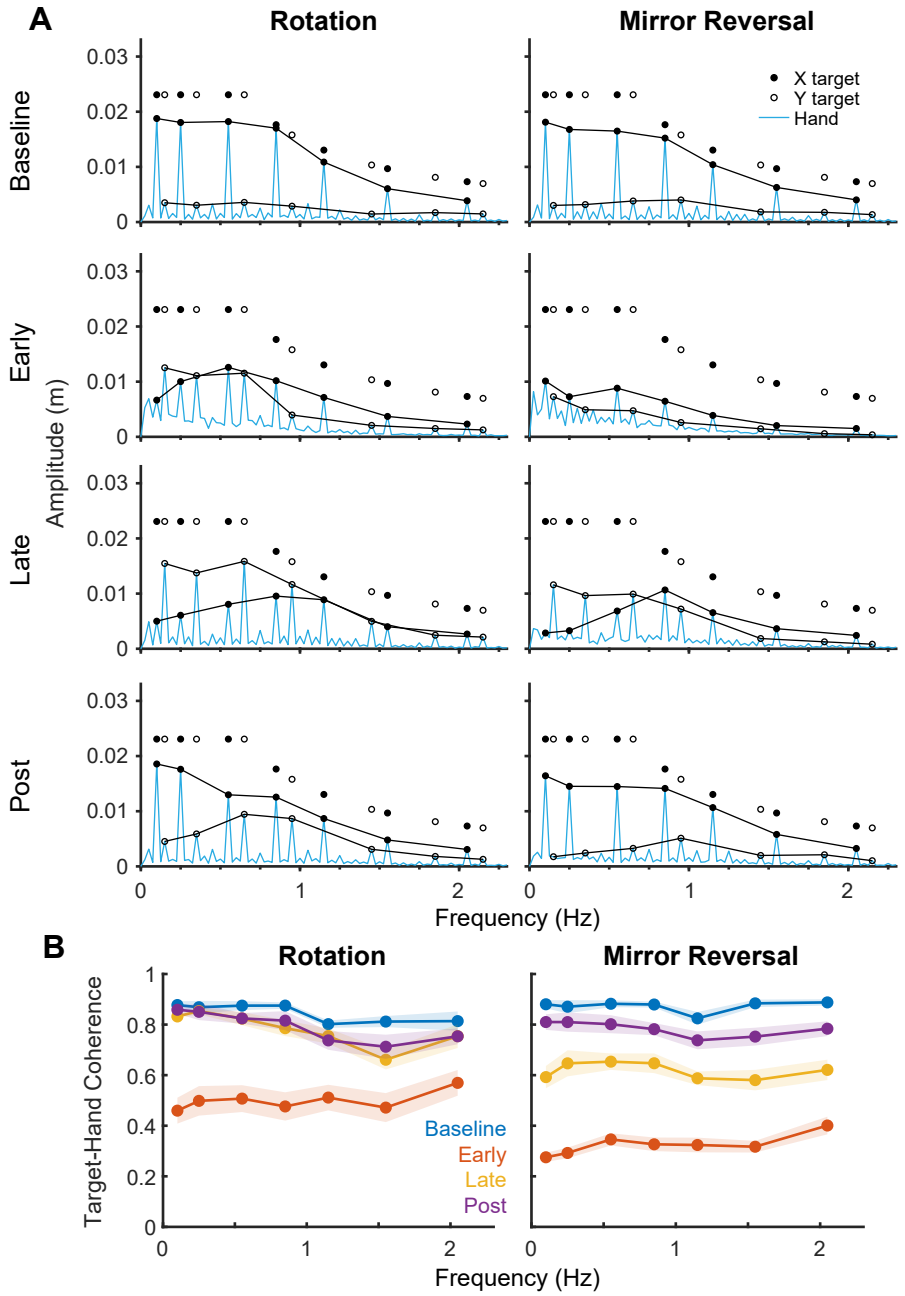

Figure 4. Frequency responses of hand movements during the tracking task. **A.** Amplitude spectra of x-hand (blue line) trajectories during different blocks of learning. In each plot, the amplitude and frequency of x- and y-target motion is depicted as the upper row of black and white dots, respectively. The peaks of the hand spectra at either the x- or y-target frequencies are connected by black lines for ease of visualization. **B.** Spectral coherence between x-target movement and both x- and y-hand movement (i.e., single-input multi-output coherence). This coherence is proportional to the linear component of the hand's response to the target. Corresponding plots for y-target movements' coherence with hand movement can be found in Figure S1B. Error bars are SEM across participants.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

target movement would imply that the hand should move at the same frequencies as the target along each axis. Lastly, target trajectories were carefully constructed to allow potential nonlinearities (which would manifest as extraneous peaks in the output spectra) to be clearly discernible (see Methods for details).

The amplitude spectra show clearly that, at baseline, both groups of participants moved almost exclusively at the frequencies of the target (Figure 4A: x-hand data, Figure S1: y-hand data). Furthermore, x-hand movements primarily occurred at x-target frequencies, not y-target frequencies. Likewise, y-hand movements primarily occurred at y-target frequencies. (Note that coupling across axes, which we observed at other stages of learning, would *not* indicate nonlinear behavior; rather it would indicate an imperfect though possibly linear sensorimotor mapping.) We further quantified how linear participants' responses were by computing the spectral coherence between target and hand movement [41]. Although coherence was low during early learning, it was much closer to the maximum value of 1 at baseline, late learning, and post-learning (Figure 4B). This is consistent with a linear relationship between target motion and hand motion. These data suggest that participants faithfully tracked the target throughout the experiment and provided validation for using linear, frequency-domain analysis to examine behavior.

The introduction of the perturbation led to a broadband increase in amplitude across all frequencies for both groups (Figure 4A, "Early"), indicating some nonlinear behavior as one might expect, particularly early on in learning. However, the peaks at the target frequencies were still clearly identifiable, enabling the use of linear systems analysis to approximate the relationship between target motion and hand motion. These nonlinearities abated with practice (Figure 4A, "Late") and remained modest after the perturbation was removed (Figure 4A, "Post"), again facilitating linear analysis.

To perfectly compensate for either the rotation or the mirror reversal, movement at x-target frequencies needed to be remapped from x-hand movements to y-hand movements, and vice versa at y-target frequencies. During early learning, participants in the rotation group did produce x-hand movements in response to y-target frequencies, but also inappropriately continued to produce x-hand movements at x-target frequencies (Figure 4A). By late learning, the amplitude of x-hand movement further increased at y-target frequencies and decreased at x-target frequencies. Behavior for the mirror-reversal group followed a similar pattern, albeit with less pronounced peaks in the amplitude spectrum during early learning. Importantly, the extent of remapping responses at different frequencies was not uniform across the frequency spectrum, but appeared frequency-dependent.

After the perturbation was removed (Post-learning), the rotation group exhibited x-hand movements at both x- and y-target frequencies, unlike baseline where movements were largely restricted to x-target frequencies (Figure 4A). This indicated that learning to compensate for a rotation led to aftereffects, corroborating our earlier analysis. In contrast, the mirror-reversal group's x-hand movements were indistinguishable

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

baseline, confirming that any aftereffects were negligible. These features of the amplitude spectra, and the differences across groups, were qualitatively the same for y-hand movements (Figure S1).

Amplitude spectra illustrate important features of learning, but do not carry information about the directionality of the response and thus do not distinguish learning of the two different perturbations; perfect compensation would lead to identical *amplitude* spectra for each perturbation. In order to distinguish these responses, we needed to determine not just the amplitude, but the direction of response along each axis, i.e. whether it was positive or negative. We used phase information to disambiguate the sign of the gain by assuming that the phase of the response at each frequency would remain similar to baseline throughout learning.

In order to understand the potential frequency-dependence of participants' compensation for the perturbation, we sought to estimate a series of transformation matrices describing participants' behavior across different frequencies. To do so, we grouped pairs of neighboring x- and y-target frequencies to build seven 2×2 signed gain matrices. These matrices represented how the transformation varied across the frequency spectrum, under the assumption that behavior was similar across neighboring pairs of frequencies. Similar to the trajectory-alignment analysis, participants should have a gain matrix close to the identity matrix at baseline while, under ideal compensation for each perturbation, each gain matrix should equal the inverse of the matrix describing the perturbation.

We visualized these estimated frequency-dependent transformation matrices through their effect on the unit x and y vectors (the columns of the gain matrices; Figure 5A). At baseline, participants in both groups responded to x- and y-target motion by moving their hands in the x- and y-axes, respectively, across all target frequencies. For the rotation group, all vectors rotated clockwise during learning, although compensation appeared to be more complete at low frequencies (darker arrows) than at high frequencies (lighter arrows). For the mirror-reversal group, compensation during late learning occurred most successfully at low frequencies, apparent as the darker vectors flipping across the mirror axis relative to baseline. At high frequencies, however, responses failed to flip across the mirror axis and remained similar to baseline.

To examine these observations statistically, we focus again on the off-diagonal elements of the gain matrices. The rotation group's gain matrices were altered in the appropriate direction to counter the perturbation at all frequencies (Figure 5B; linear mixed effects model: 2-way interaction between block and frequency, $F(12, 360) = 19.77$, $p < 0.0001$; data split by frequency for post hoc Tukey's range test: Bonferroni-adjusted $p < 0.0001$ for all frequencies). Although the mirror-reversal group's low-frequency gain matrices were also altered in the appropriate direction (Tukey's range test: Bonferroni-adjusted $p < 0.0001$ for lowest 3 frequencies), the high-frequency gain matrices were not significantly different from baseline (Tukey's range test: Bonferroni-adjusted $p > 0.7479$ for highest 4 frequencies; baseline gain range: -0.03–0.06; late-learning gain

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

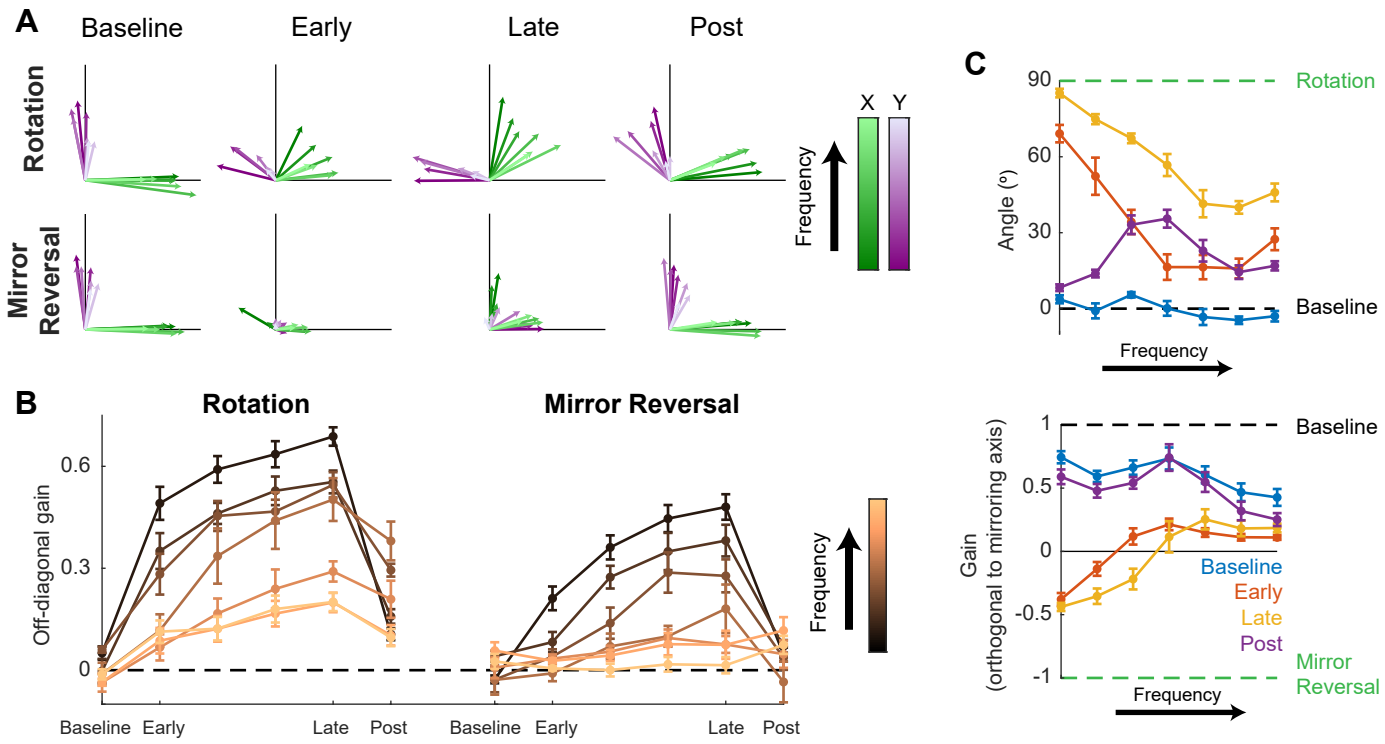

Figure 5. Gain matrix analysis of tracking task at different frequencies of hand movement. For all panels, neighboring x- and y-target frequencies were paired together in numerical order. The 7 frequency pairings were (x then y frequencies reported in each parentheses in Hz): (0.1, 0.15), (0.25, 0.35), (0.55, 0.65), (0.85, 0.95), (1.15, 1.45), (1.55, 1.85), (2.05, 2.15). **A.** Visualizations of the gain matrices (see Figure S4 for colormap illustration of the matrices). Green and purple arrows depict hand responses to x- and y-target frequencies, respectively. Darker colors represent lower frequencies and lighter colors represent higher frequencies. **B.** Average of the two off-diagonal values of the gain matrices at different blocks of learning. Darker colors represent lower frequencies and lighter colors represent higher frequencies. Error bars are SEM across participants. **C.** (Top) Angle of the pure rotation matrices fitted to the rotation group's gain matrices. (Bottom) Gain of the mirror-reversal group's movements orthogonal to the mirror axis. Green and black horizontal lines show ideal compensation when the perturbation is or is not applied, respectively. Error bars are SEM across participants.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

range: 0.01–0.18).

Fitting a rotation matrix to the rotation group's gain matrix at each frequency revealed that participants' baseline compensation angle was close to 0° at all frequencies (Figure 5C). By late learning, compensation was nearly perfect at the lowest frequency and dropped off at higher frequencies. For the mirror-reversal group, the gains of participants' low-frequency movements orthogonal to the mirror axis were positive at baseline and became negative during learning, appropriate to counter the perturbation. At high frequencies, by contrast, the gain reduced slightly during learning but never became negative. Taken together, these analyses suggest that both groups of participants were successful at compensating at low frequencies, but at high frequencies, the rotation group was only mildly successful and the mirror-reversal group was largely unsuccessful.

Post-learning, the rotation group's off-diagonal gains were significantly different from baseline for all frequencies except the lowest frequency (Figure 5B; post hoc linear mixed-effects models subset by frequency: 2-way interaction between group and block, $F(2, 36) > 6.36$ for highest 6 frequencies; Tukey's range test: Bonferroni-adjusted $p < 0.0454$ for highest 6 frequencies), indicating aftereffects. A similar trend was evident in participants' estimated compensation angles (Figure 5C). By contrast, the mirror-reversal group's matrices were not significantly different from baseline across all frequencies (Figure 5B; Tukey's range test: Bonferroni-adjusted $p > 1.50$ for all frequencies; baseline gain range: -0.03–0.06; post-learning gain range: -0.03–0.12). The gains orthogonal to the mirror axis were also similar to baseline across all frequencies, confirming the lack of aftereffects (Figure 5C).

Thus, compensation for mirror reversal was frequency-dependent and did not exhibit aftereffects. Compensation for visuomotor rotation was similarly frequency-dependent, but was augmented by an additional component of learning which was expressed similarly across frequencies and did result in aftereffects.

In a final experiment, we tested the degree to which training in the point-to-point task was necessary for participants to learn a continuous feedback controller. Participants in this experiment experienced the rotation or mirror reversal primarily using the tracking task with minimal point-to-point practice (Figure S5A). The total amount of compensation participants expressed while learning the rotation was blunted in comparison to the main experiment, but this group did gradually improve throughout the experiment (Figure S5B-D). This result is consistent with previous work comparing the generalization of rotation learning from pointing to tracking tasks and vice versa [42]. However, mirror-reversal learning was severely blunted in comparison to the main experiment and did not improve past early learning. Thus, point-to-point practice was helpful in learning the rotation but was critical for learning the mirror reversal.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Discussion

In the present study, we tested whether participants could learn to achieve continuous control over a cursor under either rotated or mirror-reversed visual feedback. Although previous work has established that participants can learn to compensate for such perturbations during point-to-point movements, this compensation could have been achievable using a simple re-aiming strategy rather than by constructing a novel, continuous feedback controller. We found that when participants performed a continuous tracking task, their behavior was close to linear (as quantified by spectral coherence), implying that they tracked the target continuously, rather than with intermittent catch-up movements. This pattern was preserved when participants learned to compensate for either a 90° visuomotor rotation or a mirror-reversal. As expected, we found that participants who learned the visuomotor rotation exhibited strong aftereffects once the perturbation was removed, amounting to an approximately 25° rotation of hand motion relative to target motion—consistent with previous findings in point-to-point tasks. In contrast, participants who learned a mirror-reversal showed no aftereffects, suggesting that they did not learn to compensate by adaptation of their existing controller, but by establishing a new feedback control policy *de novo*.

Interestingly, we found that the pattern of compensation was frequency specific (Figure 5B) with the nature of compensation at high frequencies in particular revealing distinct signatures of adaptation and *de novo* learning. At low frequencies, both groups of participants successfully compensated for their perturbations. But at high frequencies, only the rotation group were able to apply any compensation; behavior for the mirror-reversal group at high frequencies was similar to baseline behavior. There were similarities, however, in the overall time course and frequency dependence of learning under each perturbation (Figure 5B), with both groups exhibiting a steady increase in compensation over time, particularly at lower frequencies. Additionally, both groups' compensation exhibited a similar gradation as a function of frequency, decreasing as frequency increased.

We believe these results show that distinct learning processes drove two separate components of learning. One component, present only in the rotation group, was expressed uniformly at all frequencies and likely reflects a parametric adjustment of an existing baseline control policy, i.e., adaptation. This interpretation is consistent with studies demonstrating that adaptation of point-to-point movements generalizes to feedback corrections for unexpected perturbations [18, 30, 31], and is also consistent with the limited compensation of around 25° [6, 14–16], matching the extent of compensation we observed at high frequencies.

A second component of learning—which was present in both groups of participants—contributed to compensation primarily at low frequencies, exhibited a gradation as a function of frequency, and was not associated with aftereffects. We suggest this component corresponds to *de novo* learning. Based on the premise

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

that mirror-reversal learning is purely *de novo*, the mirror-reversal group's behavior in Figure 5B demonstrates the frequency-dependent characteristics of *de novo* learning and how those characteristics evolve with training. Previous studies have demonstrated that participants are unable to generate appropriate feedback corrections at low-latency after learning to make point-to-point movements under mirror-reversed visual feedback [18], paralleling our finding that the mirror-reversal group could not compensate at high-frequencies of target movement. On the other hand, the rotation group's total compensation may be conceptualized as the superposition of the two components of learning. These data thus support previous suggestions that residual learning under a visuomotor rotation that cannot be attributed to implicit adaptation may rely on the same mechanisms as those used for *de novo* learning [1].

The inability for either group to compensate at high frequencies (when tracking an unpredictable stimulus; see [33]) is likely attributable to delays in the feedback control loop; once delays become comparable to the period of oscillation, then it becomes impossible to exert precise control at that frequency. Our findings thus suggest the existence of two distinct control pathways, each capable of different forms of plasticity—one that is fast but can only be recalibrated through adaptation and one that is slower but can be reconfigured to implement arbitrary new control policies. One possibility is that these two pathways reflect feedforward control (generating motor output based purely on target motion) and feedback control (generating motor output based on the cursor location and/or distance between cursor and target). Feedback control is slower than feedforward control due to the additional delays associated with observing the effects of one's prior motor commands on the current cursor position. The observed pattern of behavior may thus be due to a fast but inflexible feedforward controller that always expresses baseline behavior (although adjustable via implicit adaptation) interacting with a slow but easily reconfigurable feedback controller that learns to compensate for perturbed visual input. At low frequencies, the target may move slowly enough that any inappropriate feedforward control to track the target is masked by corrective feedback responses. But at high frequencies, the target may move too fast for feedback control to be exerted, leaving only inappropriate feedforward responses.

An alternative explanation is that there may be multiple feedforward controllers (and/or feedback controllers) incurring different delays. A fast but inflexible baseline controller (amenable to recalibration through adaptation) might interact with a slower but more flexible controller. This organization parallels dual-process theories of learning and action selection [43–45] and raises the interesting possibility that the *de novo* learning exhibited by our participants might be, in some sense, cognitive in nature. Re-aiming to compensate for visuomotor perturbations during point-to-point movements is generally presumed to be cognitive [45, 46]. However, continuous feedback control would generally be considered to be beyond the scope of such cognitive compensation.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Cognitive compensation might typically be considered a time-consuming, deliberative process that would be too slow to account for behavior even at relatively low frequencies in our tasks. However, it is possible for action selection to occur quite rapidly but still be considered cognitive. For instance, it has been proposed that stimulus-response associations can be "cached" in working memory, enabling a cognitive response to be deployed rapidly and without deliberation [37]. However, the number of associations that can be stored in this way appears to be limited to just 2-7 [37, 47], raising doubts as to whether such a mechanism could support a continuous feedback controller. Nevertheless, recent theories have framed prefrontal cortex as a general-purpose network capable of learning to perform arbitrary computations on its inputs [48]. From this perspective, it does not seem infeasible that such a network could learn to implement an arbitrary feedback controller.

It is important to note that we use the term "*de novo* learning" to refer to any mechanism, aside from implicit adaptation, that leads to the creation of a new control policy. We suggest that *de novo* learning is initially solved explicitly before becoming cached or automatized into a more procedural form. There are, however, a number of alternative mechanisms that could be engaged to establish a new control policy. One proposal is that *de novo* learning occurs by simultaneously updating forward and inverse models by simple gradient descent [49]. Another possibility is that a new policy could be learned through reinforcement learning. In motor learning tasks, reinforcement has been demonstrated to engage a learning mechanism that is independent of implicit adaptation [50–52] potentially via basal-ganglia-dependent mechanisms [53, 54]. Such reinforcement could provide a basis for forming a new control policy. Although prior work on motor learning has focused on point-to-point movements, theoretical frameworks for reinforcement learning have been extended to continuous time and space to learn continuous control policies for robotics [55–58].

One additional interesting finding is that when participants exhibited phase lags relative to the target of 180°, they still tracked the target with non-zero gain (Figure S3). All participants exhibited this behavior, even at baseline. When the hand is phase lagged by 90°–270° at a particular frequency, the optimal gain to minimize the distance between the hand and target at that frequency is zero [33]. This peculiarity of human manual tracking behavior has been demonstrated before [59], suggesting that this was not a spurious finding. A potential explanation for this result is that participants may be relying on some strategy that is incompatible with this particular notion of optimality (e.g. reproducing the target trajectory instead of minimizing distance from the target). Alternatively, it could be that participants were unable to assess the error associated with these out-of-phase movements at high frequencies because the stimulus is masked within an inherently unpredictable stimulus; other studies have indicated that when only a ("predictable") single-sine component is presented, the nervous system makes a rational judgement to reduce the amplitude of tracking movements when there are large phase lags [33, 59], likely because the predictable stimulus makes

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

it easier for the nervous system to ascertain the cost of making extraneous movements.

Ultimately, our goal is to understand real-world skill learning. We believe that studying learning in continuous tracking tasks can help bring us closer to this goal since a critical component of many skills is the ability to continuously control an effector. Tasks like juggling and riding a bike rely heavily on the ability to perform feedback corrections at low latency in response to external events. Studies of well-practiced human behavior in continuous control tasks has a long history, such as those examining the dynamics of pilot and vehicle interactions [60]. However, most existing paradigms for studying motor *learning* have examined only point-to-point movements and, although mid-movement perturbations have been used to assess learning of feedback control [18, 29], very few studies have sought to examine how we learn to continuously manipulate an effector in response to a continuously changing goal.

Although we have described the mirror-reversal task as requiring *de novo* learning, we acknowledge that there are many types of learning which might be described as *de novo* learning that this task does not capture. For example, many skills, such as playing the cello, challenge one to learn how to *execute* new movement patterns that one has never executed before [17]. This is not the case in the tracking task which only challenges one to *select* movements one already knows how to execute. Also, in many cases, one must learn to use information from new sensory modalities for control [61, 62], such as adjusting one's action based on auditory feedback while playing the cello. Our task, by contrast, uses only very familiar visual cues. Nevertheless, we believe that learning a new controller, even with already well-practiced actions and familiar sensory feedback, is a critical part of learning many skills and represents an important aspect of any *de novo* learning. The tracking task presented here offers a simple but powerful approach for characterizing how we learn a new continuous feedback controller and, as such, provides an important new direction for studying skill learning.

Our characterization of learning made use of frequency-based system identification, a powerful tool that has been extensively used to study biological motor control such as insect flight [35, 63], electric fish refuge tracking [34, 64], human posture [36, 65], and human reaching [59]. Although measuring feedback responses to a step perturbation can be considered a form of system identification, frequency-based methods constitute an approach that is substantially more powerful [66], as well as being considerably more time-efficient in comparison to imposing intermittent target jumps on a subset of trials. However, to our knowledge, the frequency-based approach has not previously been applied to investigate motor learning. Here, we have demonstrated that a system identification approach can not only recapitulate the results of previous studies but also extend these results to identify distinct components of control. Our approach is also very general as it can be applied to assess learning of any linear visuomotor mapping (e.g., 15° rotation, bimanual perturbations). Under previous approaches, quantifying feedback responses under different types of learned

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

mappings (rotation, mirror-reversal) require different *ad hoc* analyses that cannot be directly compared [18]. In contrast, our frequency-based approach provides a general method to characterize behavior under rotations, mirror-reversals, or any linear mapping from effectors to a cursor.

The primary goal of our frequency-based analysis was to establish how participants mapped target motion into hand motion at different stages of learning. However, frequency-based system identification yields even more information than this; in principle, it provides complete knowledge of a linear control system in the sense that knowing how the system responds to sinusoidal input at different frequencies enables one to predict how the system will respond to arbitrary inputs. The full frequency-domain data, including both phase and gain information across all frequencies, can be used to formally compare different possible control system architectures [59] supporting learning. We intend to pursue this approach in future work to obtain a more algorithmic understanding of how the sensorimotor control system changes during different forms of learning.

Methods

Participants. 20 right-handed, healthy participants over 18 years of age were recruited for this study (23.55 ± 4.95 years old; 10 male, 10 female). All participants had no history of neurological disorders. All methods were approved by the Johns Hopkins School of Medicine Institutional Review Board.

Tasks

Participants made planar movements with their right arm—which was supported by a frictionless air sled on a table—to control a cursor on an LCD monitor (60 Hz). Participants viewed the cursor on a horizontal mirror which reflected the monitor (Figure 1B). Hand movement was monitored with a Flock of Birds magnetic tracker (Ascension Technology, VT, USA) at 130 Hz. The cursor was controlled under three different hand-to-cursor mappings: 1) veridical, 2) 90° visuomotor rotation, and 3) mirror reversal about a 45° oblique axis. The participants were divided evenly into two groups, one that experienced the visuomotor rotation ($n = 10$) and one that experienced the mirror reversal ($n = 10$). Both groups were exposed to the perturbed cursors while performing two different tasks: 1) the point-to-point task, and 2) the tracking task.

Point-to-point task. To start a trial, participants were required to move their cursor (circle of radius 2.5 mm) into a target (circle of radius 10 mm) that appeared in the center of the screen. After 500 ms, the target appeared 12 cm away from the starting location in a random direction. Participants were explicitly instructed to move in a straight line, as quickly and accurately as possible to the new target. Once the cursor remained stationary (speed < 0.065 m/s) in the new target for 1 sec, the target appeared in a new location

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

12 cm away, but constrained to lie within a 20 cm $\times$ 20 cm workspace. Each block used different, random target locations from other blocks. Blocks in the main experiment consisted of 150 reaches while blocks in the control experiment consisted of 15 reaches. To encourage participants to move quickly to each target, we provided feedback at the end of each trial about the peak velocity they attained during their reaches, giving positive feedback (a pleasant tone and the target turning yellow) if the peak velocity exceeded roughly 0.39 m/s and negative feedback (no tone and the target turning blue) if the peak velocity was below that threshold.

Tracking task. At the start of each trial, a motionless target (circle of radius 8 mm) appeared in the center of the screen, and the trial was initiated when the participant's cursor (circle of radius 2.5 mm) was stationary (speed < 0.065 m/s) in the target. From then, the target began to move for 46 seconds in a continuous, pseudo-random trajectory. The first 5 seconds was a ramp period where the amplitude of the cursor increased linearly from 0 to its full value, and for the remaining 41 seconds, the target moved at full amplitude. The target moved in a two-dimensional, sum-of-sinusoids trajectory; seven sinusoids of different frequencies, amplitudes and phases were summed to determine target movement in the x-axis (frequencies (Hz): 0.1, 0.25, 0.55, 0.85, 1.15, 1.55, 2.05; amplitudes (cm): 2.31, 2.31, 2.31, 1.76, 1.30, 0.97, 0.73, respectively), and the same procedure was repeated for target movement in the y-axis but with different frequencies (frequencies (Hz): 0.15, 0.35, 0.65, 0.95, 1.45, 1.85, 2.15; amplitudes (cm): 2.31, 2.31, 2.31, 1.58, 1.03, 0.81, 0.70, respectively). Different frequencies were used for the x- and y-axes so that hand movement responses at a given frequency could be attributed to either x- or y-axis target movements. All frequencies were prime multiples of a base frequency of 0.05 Hz (base period of 20 s) to ensure that the harmonics of any target frequency would not overlap with any other target frequency. Additionally, the amplitudes of the sinusoids for all but the lowest frequencies were proportional to the inverse of their frequency to ensure that each individual sinusoid had similar peak velocity. We set a ceiling amplitude for low frequencies in order to prevent target movements that were too large for participants to comfortably track. Note that due to the construction of the input signal, there were no low-order harmonic relations between any of the component sinusoids on the input, making it likely that nonlinearities in the tracking dynamics would manifest as easily discernible harmonic interactions (i.e. extraneous peaks in the output spectra). Moreover, care was taken to avoid "frequency leakage" by designing discrete Fourier transform windows that were integer multiples of the base period, improving our ability to detect such nonlinearities.

For the entire 46-second target trajectory, participants were explicitly instructed to try to keep their cursor inside the target for as long as possible during the trial. These specific instructions were provided to encourage participants to track the target, as opposed to replicating the target's trajectory. We also changed

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

the target's color to yellow anytime the cursor was inside the target to provide feedback for their success. After 46 seconds had elapsed, the trial ended. One block of the tracking task consisted of eight, 46-second trials (6 min 6 seconds of tracking per block), and the same target trajectory was used for every trial within a block. For different blocks, we randomized the phases, but not the frequencies, of target sinusoids to produce different trajectories. We produced 5 different target trajectories for participants to track in the 6 tracking blocks. The trajectory used for baseline and post-learning were the same to allow a better comparison for aftereffects. All participants tracked the same 5 target trajectories, but we randomized the order in which they experienced these trajectories in order to minimize any phase-dependent learning effects.

Experiment. We first assessed the baseline control of both groups of participants by having them perform one block of the tracking task followed by one block of the point-to-point task under veridical cursor feedback. We then applied either the VMR or MR to the cursor, and used the tracking task to measure their control capabilities during early learning. Afterwards, we alternated three times between blocks of point-to-point training and blocks of tracking. In total, each participant received 450 point-to-point reaches of training under perturbed cursor feedback. Finally, we measured aftereffects post-learning by returning to the veridical mapping and using the tracking task.

Data Analysis

Analyses were performed using MATLAB (The Mathworks, Natick, MA, USA) and RStudio (RStudio, Inc., Boston, MA, USA) [67–71]. Figures were generated using Adobe Illustrator (Adobe Inc., San Jose, CA, USA).

Trajectory-based analysis. In the point-to-point task, we assessed performance by calculating the angular error between the cursor's initial movement direction and the target direction relative to the start position. To determine the cursor's initial movement direction, we computed the direction of the cursor's instantaneous velocity vector roughly 150 ms after the time of movement initiation. Movement initiation was defined as the time when the cursor left the start circle on a given trial.

To obtain cleaner estimates of hand trajectories in the tracking task, we averaged the trajectories across all trials within each block, doing this separately for each participant. We assessed performance by measuring the average mean-squared error between the hand and target positions on every trial for each participant. We subsequently averaged this error across participants. For the trajectory-alignment analysis, we first aligned the hand and target trajectories in time by delaying all target trajectories by approximately 380 ms relative to their respective hand trajectories. We determined this delay by identifying the average delay which would

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

minimize the mean-squared error at baseline across all participants. We then fit a transformation matrix for each participant, for each block, based on minimizing the mean-squared error between the average hand and target trajectories:

$$\underset{a,b,c,d}{\operatorname{argmin}} \left\{ \begin{bmatrix} \overrightarrow{Hand}_X \\ \overrightarrow{Hand}_Y \end{bmatrix} - \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \overrightarrow{Target}_X \\ \overrightarrow{Target}_Y \end{bmatrix} \right\} \tag{1}$$

This matrix was averaged across participants to generate the transformation matrices shown in Figure 3A. The transformation matrices were visualized by plotting the matrices' column vectors.

The off-diagonal elements of each participant's transformation matrix was used to calculate the cross-axis scaling:

$$Scale_{rotation} = \frac{-b+d}{2} \qquad Scale_{mirror} = \frac{b+d}{2} \tag{2}$$

Rotation angles for the rotation group's transformation matrices were found by identifying the angle which generated pure rotation matrices that best fit the transformation matrices according to an L_2 norm. Finally, for the mirror-reversal group, the scaling orthogonal to the mirror axis was found by computing how the matrix transformed the unit vector along the orthogonal axis:

$$Scale_{orthogonal} = 0.5 \left(\begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right) \tag{3}$$

Frequency-domain analysis. To analyze trajectories in the frequency domain, we applied the discrete Fourier transform to the target and hand trajectories in every tracking trial. This produced a set of complex numbers representing the amplitude and phase of the signal at every frequency. We only analyzed the first 40 seconds of the trajectory that followed the 5-second ramp period so that our analysis period was equivalent to an integer multiple of the base period (20 s). This ensured that we would obtain clean estimates of the sinusoids at each target frequency. Amplitude spectra were generated by taking double the modulus of the Fourier-transformed hand trajectories at positive frequencies.

Spectral coherence was calculated between the target and hand trajectories. To do so, we evaluated the multiple (i.e., multi-input multi-output) coherence at every frequency of target motion, determining how target motion in one axis elicited hand movement in both axes. We chose to compute the multiple coherence as this best captured the linearity of participants' behavior; using hand movement in only one axis for the analysis would only partially capture participants' responses to target movement at a given frequency. Calculations were performed using a 1040-sample Blackman-Harris window.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

In order to estimate the frequency-dependent 2×2 transformation between target and hand movement, we assumed that a matrix, A, could capture how the sensorimotor system transformed target movement into hand movement for each participant at given frequency:

$$\begin{bmatrix} \text{Hand}_x \\ \text{Hand}_y \end{bmatrix} = A \begin{bmatrix} \text{Target}_x \\ \text{Target}_y \end{bmatrix} \qquad A = \begin{bmatrix} a_{xx} & a_{xy} \\ a_{yx} & a_{yy} \end{bmatrix} \tag{4}$$

Hand and Target are vectors of complex sinusoids corresponding to the time-domain trajectories in the x- and y-axes. Since the target only moved at a discrete set of frequencies, which were different for the x- and y- axes, we paired neighboring x- and y-frequencies to assess the target-to-hand transformation in a small bandwidth, assuming that behavior would be approximately the same across neighboring frequencies. Each element of A therefore represents the transformation of target motion along a particular axis to hand motion in a particular axis; the first subscript represents the hand-movement axis and the second subscript represents the target-movement axis. To obtain A, we first computed the ratios between the complex sinusoids of hand movement and target movement at each frequency of movement. These ratios are known as phasors, complex numbers which relate the sinusoids in terms of gain and phase [33,59]. Target movement at a given frequency, which only moves in one axis, can elicit movement in both axes. Therefore, for each x- target frequency, we computed a phasor for x-hand responses and a phasor for y-hand responses. We performed similar calculations for each y-target frequency. For a given block of learning, the two phasors for each frequency of movement were paired together to form fourteen phasor pairs. Pairs describing behavior from the same frequency were grouped together across all six blocks of learning.

Gain and phase data are inherently redundant—a positive gain with a phase of π is indistinguishable from a negative gain with a phase of 0. Conventionally, this redundancy is resolved by assuming that gain is positive. In our task, however, the sign of the gain was crucial to disambiguate the directionality of the hand responses (e.g. whether the hand moved left or right in response to upward target motion). We used phase information to disambiguate positive from negative gains, assuming that the phase of the hand response at a given frequency would be similar throughout the experiment. To do so, we introduced template phasors, with a fixed gain of 1, to capture the assumed invariant phase of the response at a given frequency. We estimated these template phasors based on behavior in the baseline block and assumed that this phase would be preserved across different axes of movement and across subsequent blocks. Therefore, we only needed to fit a scalar (potentially negative) gain a:

$$\underset{a_{xx},a_{yx}}{\text{argmin}} \left\{ \begin{bmatrix} \psi_{xx} \\ \psi_{yx} \end{bmatrix} - \begin{bmatrix} a_{xx}\psi_{\text{template1}} \\ a_{yx}\psi_{\text{template1}} \end{bmatrix} \right\} \qquad \underset{a_{xy},a_{yy}}{\text{argmin}} \left\{ \begin{bmatrix} \psi_{xy} \\ \psi_{yy} \end{bmatrix} - \begin{bmatrix} a_{xy}\psi_{\text{template2}} \\ a_{yy}\psi_{\text{template2}} \end{bmatrix} \right\} \tag{5}$$

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Here, ψ indicates an individual phasor. We therefore a set of template phasors (template 1 for the lowest frequency, template 2 for the second lowest frequency, etc.), and a set of gains a, which were assumed to vary from block to block. By pairing neighboring x- and y-frequencies, we obtained a set of seven A matrices describing the signed gain relating target movement to hand movement across frequencies ranging from low to high. Visualizations of these gain matrices, cross-axis gain, rotation angle, and gain orthogonal to the mirroring axis were calculated in the same way as in the transformation matrix analysis.

Statistics

The statistical tests for the transformation and gain matrix analyses were performed using linear mixed-effects models. The parameters of our models were the perturbation type (rotation or mirror reversal), block of learning, and element of the matrices. In both analyses, we hypothesized that there would be interactions between all three parameters. We followed the initial statistical modeling by performing post-hoc tests where we split the data by the element of the transformation/gain matrix being analyzed. Pairwise comparisons were performed using Tukey's range test. Error bars were equal to the standard error of the mean. Confidence ellipses were approximately double the standard error of the mean.

References

[1] Krakauer, J. W., Hadjiosif, A. M., Xu, J., Wong, A. L. & Haith, A. M. Motor learning. *Compr. Physiol.* **9**, 613–663 (2019).

[2] Mazzoni, P. & Krakauer, J. W. An implicit plan overrides an explicit strategy during visuomotor adaptation. *J. Neurosci.* **26**, 3642–3645 (2006).

[3] Tseng, Y.-W., Diedrichsen, J., Krakauer, J. W., Shadmehr, R. & Bastian, A. J. Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *J. Neurophysiol.* **98**, 54–62 (2007).

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

[5] Krakauer, J. W., Ghilardi, M. F. & Ghez, C. Independent learning of internal models for kinematic and dynamic control of reaching. *Nat. Neurosci.* **2**, 1026–1031 (1999).

[6] Fernández-Ruiz, J., Wong, W., Armstrong, I. T. & Flanagan, J. R. Relation between reaction time and reach errors during visuomotor adaptation. *Behav. Brain Res.* **219**, 8–14 (2011).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

[7] Morehead, J. R., Qasim, S. E., Crossley, M. J. & Ivry, R. Savings upon re-aiming in visuomotor adaptation. *J. Neurosci.* **35**, 14386–14396 (2015).

[8] Martin, T. A., Keating, J. G., Goodkin, H. P., Bastian, A. J. & Thach, W. T. Throwing while looking through prisms. I. Focal olivocerebellar lesions impair adaptation. *Brain* **119 (Pt 4)**, 1183–1198 (1996).

[9] Fernández-Ruiz, J. & Díaz, R. Prism adaptation and aftereffect: Specifying the properties of a procedural memory system. *Learn. Mem.* **6**, 47–53 (1999).

[10] Choi, J. T. & Bastian, A. J. Adaptation reveals independent control networks for human walking. *Nat. Neurosci.* **10**, 1055–1062 (2007).

[11] Finley, J. M., Long, A., Bastian, A. J. & Torres-Oviedo, G. Spatial and temporal control contribute to step length asymmetry during split-belt adaptation and hemiparetic gait. *Neurorehabil. Neural Repair* **29**, 786–795 (2015).

[12] Lackner, J. R. & Dizio, P. Rapid adaptation to Coriolis force perturbations of arm trajectory. *J. Neurophysiol.* **72**, 299–313 (1994).

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

[14] Taylor, J. A., Klemfuss, N. M. & Ivry, R. B. An explicit strategy prevails when the cerebellum fails to compute movement errors. *Cerebellum* **9**, 580–586 (2010).

[15] Taylor, J. A. & Ivry, R. B. Flexible cognitive strategies during motor learning. *PLoS Comput. Biol.* **7** (2011).

[16] Bond, K. M. & Taylor, J. A. Flexible explicit but rigid implicit learning in a visuomotor adaptation task. *J. Neurophysiol.* **113**, 3836–3849 (2015).

[17] Costa, R. M. A selectionist account of de novo action learning. *Curr. Opin. Neurobiol.* **21**, 579–586 (2011).

[18] Telgen, S., Parvin, D. & Diedrichsen, J. Mirror reversal and visual rotation are learned and consolidated via separate mechanisms: Recalibrating or learning de novo? *J. Neurosci.* **34**, 13768–13779 (2014).

[19] Sternad, D. It's not (only) the mean that matters: Variability, noise and exploration in skill learning. *Curr. Opin. Behav. Sci.* **20**, 183–195 (2018).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

[20] Gutierrez-Garralda, J. M. *et al.* The effect of Parkinson's disease and Huntington's disease on human visuomotor learning. *Eur. J. Neurosci.* **38**, 2933–2940 (2013).

[21] Lillicrap, T. P. *et al.* Adapting to inversion of the visual field: A new twist on an old problem. *Exp. Brain Res.* **228**, 327–339 (2013).

[22] Redding, G. M. & Wallace, B. Adaptive coordination and alignment of eye and hand. *J. Mot. Behav.* **25**, 75–88 (1993).

[23] Kluzik, J., Diedrichsen, J., Shadmehr, R. & Bastian, A. J. Reach adaptation: What determines whether we learn an internal model of the tool or adapt the model of our arm? *J. Neurophysiol.* **100**, 1455–1464 (2008).

[24] Schugens, M. M., Breitenstein, C., Ackermann, H. & Daum, I. Role of the striatum and the cerebellum in motor skill acquisition. *Behav. Neurol.* **11**, 149–157 (1998).

[25] Maschke, M., Gomez, C. M., Ebner, T. J. & Konczak, J. Hereditary cerebellar ataxia progressively impairs force adaptation during goal-directed arm movements. *J. Neurophysiol.* **91**, 230–238 (2004).

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

[27] Ranganathan, R., Wieser, J., Mosier, K. M., Mussa-Ivaldi, F. A. & Scheidt, R. A. Learning Redundant Motor Tasks with and without Overlapping Dimensions: Facilitation and Interference Effects. *J. Neurosci.* **34**, 8289–8299 (2014).

[28] Kasuga, S., Telgen, S., Ushiba, J., Nozaki, D. & Diedrichsen, J. Learning feedback and feedforward control in a mirror-reversed visual environment. *J. Neurophysiol.* **114**, 2187–2193 (2015).

[29] Gritsenko, V. & Kalaska, J. F. Rapid online correction is selectively suppressed during movement with a visuomotor transformation. *J. Neurophysiol.* **104**, 3084–3104 (2010).

[30] Ahmadi-Pajouh, M. A., Towhidkhah, F. & Shadmehr, R. Preparing to reach: Selecting an adaptive long-latency feedback controller. *J. Neurosci.* **32**, 9537–9545 (2012).

[31] Cluff, T. & Scott, S. H. Rapid feedback responses correlate with reach adaptation and properties of novel upper limb loads. *J. Neurosci.* **33**, 15903–15914 (2013).

[32] Miall, R. C., Weir, D. J., Wolpert, D. M. & Stein, J. F. Is the cerebellum a Smith predictor? *J. Mot. Behav.* **25**, 203–216 (1993).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

[33] Roth, E., Zhuang, K., Stamper, S. A., Fortune, E. S. & Cowan, N. J. Stimulus predictability mediates a switch in locomotor smooth pursuit performance for Eigenmannia virescens. *J. Exp. Biol.* **214**, 1170–1180 (2011).

[34] Madhav, M. S., Stamper, S. A., Fortune, E. S. & Cowan, N. J. Closed-loop stabilization of the jamming avoidance response reveals its locally unstable and globally nonlinear dynamics. *J. Exp. Biol.* **216**, 4272–4284 (2013).

[35] Sponberg, S., Dyhr, J. P., Hall, R. W. & Daniel, T. L. Luminance-dependent visual processing enables moth flight in low light. *Science* **348**, 1245–1248 (2015).

[36] Kiemel, T., Oie, K. S. & Jeka, J. J. Slow dynamics of postural sway are in the feedback loop. *J. Neurophysiol.* **95**, 1410–1418 (2006).

[37] McDougle, S. D. & Taylor, J. A. Dissociable cognitive strategies for sensorimotor learning. *Nat. Commun.* **10**, 40 (2019).

[38] Leow, L.-A., Gunn, R., Marinovic, W. & Carroll, T. J. Estimating the implicit component of visuomotor rotation learning by constraining movement preparation time. *J. Neurophysiol.* **118**, 666–676 (2017).

[39] Haith, A. M., Huberdeau, D. M. & Krakauer, J. W. The influence of movement preparation time on the expression of visuomotor learning and savings. *J. Neurosci.* **35**, 5109–5117 (2015).

[40] Morehead, J. R., Taylor, J. A., Parvin, D. E. & Ivry, R. B. Characteristics of implicit sensorimotor adaptation revealed by task-irrelevant clamped feedback. *J. Cogn. Neurosci.* **29**, 1061–1074 (2017).

[41] Roddey, J. C., Girish, B. & Miller, J. P. Assessing the performance of neural encoding models in the presence of noise. *J. Comput. Neurosci.* **8**, 95–112 (2000 Mar-Apr).

[42] Abeele, S. & Bock, O. Transfer of sensorimotor adaptation between different movement categories. *Exp. Brain Res.* **148**, 128–132 (2003).

[43] Hardwick, R. M., Forrence, A. D., Krakauer, J. W. & Haith, A. M. Time-dependent competition between goal-directed and habitual response preparation. *Nat Hum Behav* (2019).

[44] Day, B. L. & Lyon, I. N. Voluntary modification of automatic arm movements evoked by motion of a visual target. *Exp Brain Res* **130**, 159–168 (2000).

[45] Huberdeau, D. M., Krakauer, J. W. & Haith, A. M. Dual-process decomposition in human sensorimotor adaptation. *Curr. Opin. Neurobiol.* **33**, 71–77 (2015).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

[46] McDougle, S. D., Ivry, R. B. & Taylor, J. A. Taking Aim at the Cognitive Side of Learning in Sensorimotor Adaptation Tasks. *Trends Cogn. Sci. (Regul. Ed.)* **20**, 535–544 (2016).

[47] Collins, A. G. E. & Frank, M. J. How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. *Eur. J. Neurosci.* **35**, 1024–1035 (2012).

[48] Wang, J. X. *et al.* Prefrontal cortex as a meta-reinforcement learning system. *Nat. Neurosci.* **21**, 860–868 (2018).

[49] Pierella, C., Casadio, M., Solla, S. A. & Mussa-Ivaldi, F. A. The dynamics of motor learning through the formation of internal models. *bioRxiv* 652727 (2019).

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

[51] Cashaback, J. G. A., McGregor, H. R., Mohatarem, A. & Gribble, P. L. Dissociating error-based and reinforcement-based loss functions during sensorimotor learning. *PLoS Comput. Biol.* **13**, e1005623 (2017).

[52] Holland, P., Codol, O. & Galea, J. M. Contribution of explicit processes to reinforcement-based motor learning. *J. Neurophysiol.* **119**, 2241–2255 (2018).

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

[54] Hikosaka, O., Nakamura, K., Sakai, K. & Nakahara, H. Central mechanisms of motor skill learning. *Curr. Opin. Neurobiol.* **12**, 217–222 (2002).

[55] Doya, K. Reinforcement learning in continuous time and space. *Neural Comput* **12**, 219–245 (2000).

[56] Theodorou, E., Buchli, J. & Schaal, S. Reinforcement learning of motor skills in high dimensions: A path integral approach. In *2010 IEEE International Conference on Robotics and Automation*, 2397–2403 (2010).

[57] Smart, W. D. & Kaelbling, L. P. Practical reinforcement learning in continuous spaces. In *Proceedings of the Seventeenth International Conference on Machine Learning*, ICML '00, 903–910 (Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2000).

[58] Todorov, E. Efficient computation of optimal actions. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 11478–11483 (2009).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

[59] Zimmet, A. M., Bastian, A. J. & Cowan, N. J. Cerebellar patients have intact feedback control that can be leveraged to improve reaching. *bioRxiv* 827113 (2019).

[60] McRuer, D. & Jex, H. A Review of Quasi-Linear Pilot Models. *IEEE Trans. Hum. Factors Electron.* **HFE-8**, 231–249 (1967).

[61] van Vugt, F. T. & Ostry, D. J. The structure and acquisition of sensorimotor maps. *J. Cogn. Neurosci.* **30**, 290–306 (2018).

[62] Bach-y-Rita, P. & W. Kercel, S. Sensory substitution and the human–machine interface. *Trends in Cognitive Sciences* **7**, 541–546 (2003).

[63] Roth, E., Hall, R. W., Daniel, T. L. & Sponberg, S. Integration of parallel mechanosensory and visual pathways resolved through sensory conflict. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 12832–12837 (2016).

[64] Cowan, N. J. & Fortune, E. S. The critical role of locomotion mechanics in decoding sensory systems. *J. Neurosci.* **27**, 1123–1128 (2007).

[65] Oie, K. S., Kiemel, T. & Jeka, J. J. Multisensory fusion: Simultaneous re-weighting of vision and touch for the control of human posture. *Brain Res. Cogn. Brain Res.* **14**, 164–176 (2002).

[66] Schoukens, J., Pintelon, R. & Rolain, Y. Time domain identification, frequency domain identification. Equivalencies! Differences? In *Proceedings of the 2004 American Control Conference*, vol. 1, 661–666 vol.1 (2004).

[67] R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2016).

[68] Lenth, R. V. *Estimability: Tools for Assessing Estimability of Linear Predictions* (2015).

[69] Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & R Core Team. *Nlme: Linear and Nonlinear Mixed Effects Models* (2016).

[70] Lenth, R. V. Least-squares means: The R package lsmeans. *J. Stat. Softw.* **69**, 1–33 (2016).

[71] Fox, J. & Weisberg, S. *An R Companion to Applied Regression* (Sage, Thousand Oaks CA, 2011), second edn.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.15.906545. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Acknowledgements We thank A. Zimmet for technical advice and J. Krakauer, A. Bastian, and C. Fetsch for helpful discussions. This material is based upon work supported by the National Science Foundation under Grant No. 1825489. C.S.Y was supported by NIH grants (5T32NS091018-17, 5T32NS091018-18) and the Link Foundation Modeling, Training & Simulation Fellowship.

Author Contributions All authors conceived the project and designed the experiments. C.S.Y. performed all data collection and analysis while N.J.C. and A.M.H. supervised. C.S.Y wrote the manuscript and N.J.C. and A.M.H. edited the manuscript.

Competing Interests The authors declare no competing interests.