Supplementary Information for

Diversity enabled sweet spots in layered architectures and speed-accuracy trade-offs in sensorimotor control

This PDF file includes:

Figs. S1 to S4
References for SI reference citations
In this supplementary document, we state the models and theories used to derive the results presented in the main text. Robust control theory is used to characterize the performance of a feedback system using its input-output relation (1, 2). The fundamental limits are given in section 1.A for the basic model, and section 1.B for the layered model, followed by a survey of relevant tools in the networked control theory in section 1.C. The neural signaling SAT is characterized by assuming a fixed space and metabolic resources to build and maintain axons. The component SATs are derived for different encoding schemes: spike-based in section 2.A and rate-based in section 2.B. Then, the system SATs are derived using the fundamental limits and the component SATs for the basic model in section 3.A and for the layered model in section 3.B. Lastly, the experimental detail and additional experimental results are presented in section 4.

Notations. We use lower case letters to denote sequences, i.e. \( x = \{x(0), x(1), x(2), \ldots \} \), and \( x(t_1 : t_2) \) to denote a truncated sequence of \( x \) from \( t_1 \) to \( t_2 \), i.e. \( x(t_1 : t_2) = \{x(t_1), x(t_1 + 1), \ldots, x(t_2)\} \). The \( \infty \)-norm of a sequence \( x \) is defined as \( \|x\|_\infty := \sup_{t} |x(t)| \), and the 1-norm of a sequence \( x \) is defined as \( \|x\|_1 = \sum_{t=0}^{\infty} |x(t)| \). The mutual information between two random variables \( x, y \) with the probability density function \( P \) is defined using \( I(x; y) = \mathbb{E} \log(P(x, y)/P(x)P(y)) \). We use \( Q \) to denote a quantizer that approximates a continuous domain with a finite set of values. The quantizer is defined by \( 2^R \) intervals partitioned by \( \{p[\ell]\} \) and their respective representative points \( \{c[\ell]\} \) such that

\[
Q(x) = \begin{cases} 
  c[1] & x \in [p[0], p[1]) \\
  c[2] & x \in [p[1], p[2]) \\
  \vdots \\
  c[2^R] & x \in [p[2^R - 1], p[2^R])
\end{cases}
\]

where \( R \) is referred to as its data rate. We denote \( \tilde{Q}_{RS} \) to be a uniform quantizer with data rate \( R \) and domain \([-\Psi, \Psi]\), which partitions its domain into \( 2^R \) intervals with equal lengths and maps the input from each interval to the middle point of that interval.

1. Fundamental limits in system performance

A. The basic model. We consider the system with delayed and quantized control:

\[
x(t + 1) = x(t) + w(t) + u(t)
\]

where \( x(t) \in \mathbb{R} \) is the system error, \( u(t) \in \mathbb{R} \) is the control action, and \( w(t) \in \mathbb{R} \) is the disturbance. We also assume zero initial condition, i.e. \( x(0) = 0 \). Next, we describe the robust control problem and its solution in a deterministic setting and a stochastic setting.

The feedback loop from sensor measurement \( x(t) \) to control action \( u(t) \) has a latency of \( T_u := T_s + T_i \) with a signaling rate \( R \). The delay \( T_u \) is composed of \( T_s \), which models the nerve signaling delay, and \( T_i \), which models other internal delays in the feedback control loop (including both sensory and motor delays). The signaling rate \( R \) is defined to be the maximum amount of information that can be transmitted by the control loop from sensors to actuators. The feedforward loop from disturbance \( w(t) \) to the control action \( u(t) \) has an advanced warning of \( T_s \), which allows the controller to get prepared for future disturbance before it hits the system dynamics. Typically, the value of \( T_s \) depends on the speed of the rider and the features on the trail. The total delay in control (i.e. delay from the moment the error dynamics are impacted by a disturbance to the moment the control acts against the disturbance) is the latency minus warning:

\[
T := T_u - T_a = T_s + T_i - T_a.
\]

Deterministic worst-case. In the worst-case setting, the disturbance \( w(t) \) is assumed to be infinity-norm bounded. The controller is characterized by the function

\[
u(t) = K(x(0 : t - T_u), w(0 : t - T_u + T_a), u(0 : t - 1)).
\]

We assume that the data rate \( R \) is minimum stabilizing, i.e. \( R > 0 \) (3). We consider minimizing the worst-case error normalized by the size of the disturbance

\[
\inf_{K} \frac{\|x\|_\infty}{\|w\|_\infty} = \inf_{K} \max_{\|w\|_\infty \leq 1} \frac{\|x\|_\infty}{\|w\|_\infty},
\]

which is equivalent to the minimum values of \( \|x\|_\infty \) when the disturbance size is normalized to \( \|w\|_\infty = 1 \). This problem admits a simple and intuitive solution. In particular, the optimal cost is given by

\[
\max(0, T) + (2^R - 1)^{-1},
\]

where \( T \) from Eq. 3 is the total delay in control.

Remark 1 As we consider marginally stable systems, the system is minimum stabilizing when \( R > 0 \). However, in terms of noise cancellation, the rate error and the total error go to \( \infty \) as \( R \to 0 \), so the actual system for sensorimotor control is unlikely to be designed to be near the minimum stabilizing rate.
Remark 2 Eq. 6 states the size of the errors and informs how the performance degrades with delay. Moreover, Eq. 6 also suggests the size of oscillations that can be induced by delays. In the experiment, if there is a large delay between the wheel and the screen cursor location, and the bump’s disturbance has a higher frequency, we expect that the delay error will get amplified. Oscillations in standing on one leg and in stick balancing can be easily tested by the readers.

Stochastic average-case. In the stochastic setting, the signal \( w(t) \) is independent and identically distributed Gaussian random variables with zero mean and unit variance. The controller is characterized by the conditional probability density function

\[
P(u(t)|x(0 : t - T_u), w(0 : t - T_u + T_a), u(0 : t - 1)).
\]

The communication in the feedback loop is done through an arbitrary discrete-time channel that satisfies the following constraint:

\[
\lim_{n \to \infty} \frac{1}{n} I\{x(0 : n - T_u), w(0 : n - T_u + T_a); u(0 : n) \} \leq R.
\]

Here, \( R \) is also assumed to be minimum stabilizing, i.e. \( R > 0 \). On special case of Eq. 7 and Eq. 8 is to have a quantizer of rate \( R \) in the feedback loop. The sensorimotor control in risk-neutral setting motivates us to consider an average error, and as such, our goal is to minimize the steady-state mean squared error normalized by error variance, i.e.

\[
\inf_{P:E} \mathbb{E}[x^2] / \mathbb{E}[w^2],
\]

which is equivalent to the following robust control problem:

\[
\inf_{P:E} \lim_{T \to \infty} \mathbb{E}\left[\frac{1}{T} \sum_{t=1}^{T} x(t)^2\right].
\]

This problem also admits a closed-form expression similar to its deterministic counterpart:

\[
\max(0, T) + (2^{2R} - 1)^{-1},
\]

where the equality is attained by an additive Gaussian channel with capacity \( R \). As a quantizer is a special case of Eq. 7 and Eq. 8, the mean squared error cannot be smaller than the left hand side of Eq. 11. Recall from the main text that the first term \( T \) in Eq. 11 can be considered as the delay error, the second term \( (2^{2R} - 1)^{-1} \) can be considered as the rate error. The delay error, the rate error, and the total error subject to the component SATs Eq. 27 is given in Fig. S2A.

B. The layered model. Next, we consider the layered system with two feedback loops

\[
x(t + 1) = x(t) + u(t) + r(t) + b(t)
\]

The disturbance is now composed of two terms: a component \( r(t) \) that is observed with advance warning \( T_a \geq 0 \) and a component \( w(t) \) that can be observed only through its impact on system response. We assume that the two disturbances are bounded by

\[
\| r \|_{\infty} \leq 1, \quad \| w \|_{\infty} \leq \epsilon.
\]

Here, \( \epsilon > 0 \) captures the relative size of \( b(t) \) in comparison to the size of \( r(t) \).

The controller is characterized by

\[
\begin{align*}
    u(t) &= u_L(t) + u_H(t) \\
    u_L(t) &= \mathcal{L}(x(0 : t - T_L - T_c), w(0 : t - T_L - T_c), u_L(0 : t - 1)) \\
    u_H(t) &= \mathcal{H}(x(0 : t - T_h), r(0 : t - T_h + T_a), u_H(0 : t - 1)).
\end{align*}
\]

The control action is generated by two nominally independent feedback loops, each having their own sensing, computation, and communication components. Both feedback loops, \( \mathcal{L}, \mathcal{H} \) act through a motor nerve pathway with data rates \( R_L, R_H \) and delays \( T_L, T_H \), respectively.

We consider minimizing the worst case infinity norm of \( x(t) \), i.e.

\[
\inf_{\mathcal{H}, \mathcal{L}, \| w \|_{\infty} \leq \epsilon, \| r \|_{\infty} \leq 1} \| x \|_{\infty}.
\]

The optimal cost is given by

\[
\left\{ T_L + T_i + \frac{1}{2R_i - 1} \right\} \epsilon + \max(0, T_h - T_a) + \frac{1}{2R_h - 1}.
\]

When \( T_a > T_h \), the value of Eq. (16) reduces to

\[
\left\{ T_L + T_i + \frac{1}{2R_i - 1} \right\} \epsilon + \frac{1}{2R_h - 1}.
\]
C. Relevant tools from networked control literature. Control under communication constraints has been extensively studied. The comprehensive surveys (3–6) cover important issues in the field of networked control. The necessary and sufficient data rate through the feedback loop in order to achieve system stability in linear stochastic control is studied in (7–9). Early study on the relation between information and estimation accuracy with causal information structures appeared in the work of (10, 11). In the context of control, the optimal controller structure, separation principles, performance bounds are studied in (10, 12–27).

2. Component SATs

In this section, we characterize the SATs for neural signaling in spike-based encoding and rate-based encoding. In a spike-based encoding scheme, information is encoded in the presence or absence of a spike in specific time intervals, analogous to digital packet-switching networks (28, 29).

A. Spike-based encoding. We model the complex size distribution of axon bundles of identical radius. We use $T_s, C_s$ to denote the delay and data rate (i.e. the amount of information in bits that can be transmitted) that can be communicated by the axon bundles, respectively. When the signaling is precise and noiseless, an axon with the achievable firing rate $\phi$ can transmit $\phi$ bits of information per unit time. For sufficiently large myelinated axons, we assume that the propagation speed $1/T_s$ is proportional to the axon radius $\rho$ (30), i.e.

$$T_s = \alpha/\rho$$  \[18\]

for some proportionality constant $\alpha$. We also model the achievable firing rate $\phi$ to be proportional to the axon radius $\rho$, i.e.

$$\phi = \beta\rho,$$  \[19\]

for some proportionality constant $\beta$. Moreover, the space and metabolic costs of a nerve are proportional to its volume (30), and given a fixed nerve length, these costs are proportional to its total cross-sectional area $s$. When the signaling is precise and noiseless, the amount of information per unit time (bits/sec) that an axon with achievable firing rate $\phi$ can transmit is simply:

$$C_s = \phi.$$  \[20\]

Combining above, we have

$$C_s = \lambda T_s.$$  \[21\]

where $\lambda = s\beta/\pi\alpha$ is proportional to the spatial and metabolic cost to build and maintain the nerves.

B. Rate-based encoding. In a rate-based encoding scheme, information is encoded in the spike rate. We can think of the rate-based encoding as a Poisson-type communication channel whose input is the spike rate $\gamma(t)$ and the output is the spike timing $M(t)$. We assume that the spike timing is a non-homogeneous Poisson point process with rate (intensity) $\gamma = \{\gamma(t) \geq 0 : t \in \mathbb{R}_+\}$, denoted by $\mathcal{P}_1(\gamma)$. The communication channel is then given by

$$M(t) = \mathcal{P}_1(\gamma).$$  \[22\]

where the spike rate is bounded by

$$\gamma(t) \leq \phi \quad t \in \mathbb{R}_+,$$  \[23\]

for some $\phi > 0$. The capacity of communication channel Eq. 22 is defined to be

$$C_r = \sup_{\gamma} \lim_{T \to \infty} \frac{1}{T} I(\gamma^T; M^T),$$  \[24\]

where the supremum is taken over all distributions of the input process $\mathcal{P}_\gamma(t)$ satisfying Eq. 23. Kabanov has shown in (31) that $C_r$ is upper-bounded by

$$C_r = \frac{(\phi + 1)^{1+\phi^{-1}}}{2} - \left(1 + \frac{1}{\phi}\right) \log(\phi + 1).$$  \[25\]

So for sufficiently large $\phi$, we have

$$C_r \to \phi/2 \quad \text{as} \quad \phi \to \infty,$$  \[26\]

which can be approximated by

$$C_r = \frac{\lambda}{2} T.$$  \[27\]
Fig. S1. The signaling rate in spike-based coding vs. rate-based coding given a fixed resource to build and maintain nerves. The solid black line shows the achievable signaling rate $C_s$ from Eq. 20 in spike-based coding; the solid blue line shows the achievable signaling rate $C_r$ from Eq. 25 in rate-based coding; and the dotted blue line shows the approximation $C_s/2$ given in Eq. 27. The rate of the rate-based encoding is less than half of that of spike-based encoding and approaches to half of the spike-based encoding rate as the achievable firing rate increases (see Eq. 26). This boost in signaling rate due to spike-based encoding may be particularly beneficial in highly constrained settings (e.g., when the achievable firing rate $\phi$ is low, or when the available resource $\lambda$ are limited).
C. Comparison of different encoding schemes. Interestingly, the SATs of spike-based encoding and rate-based encoding are qualitatively similar: given a fixed resource (space and metabolic cost to build and maintain a never), the achievable data rate is roughly proportional to delay. The amounts of information that can be transmitted in the two encoding schemes subject to a fixed resource are different and compared in Figure S1. The spike-based encoding allows more information to be transmitted than rate-based encoding and thereby more efficient. Spike based encoding is particularly beneficial in the regime of low spike rate and limited resources.

Although rate-based encoding has been believed to be the most standard encoding schemes conventionally, there is growing evidence that nerves are able to use spike-based encoding (32). An important assumption behind spike-based encoding is that axons are able to generate spikes in high timing precision. In this respect, existing literature has found that nerves are capable of spiking with highly precise timing (33–35). Moreover, a few experiments also observe that spike timing carries behaviorally relevant information (29, 34).

3. Optimizing system performance

A. System SATs in the basic model. We can combine the fundamental limits in system performance from Section 1 and the component SATs from Section 2 to obtain the system SATs. We showed the delay errors and the rate errors in control for two cases: Eq. 6 in deterministic worse-case, and Eq. 11 in stochastic average-case. We derived the component SATs in two cases: Eq. 6 in deterministic worse-case, and Eq. 11 in stochastic average-case. We can combine these results to study different scenarios and their system SATs.

Here we study the errors in the deterministic worse-case when spike-based encoding are used. Combining Eq. 6 and Eq. 21, the optimal cost is given by

\[
\max_{\|w\|_\infty \leq 1} \|x\|_\infty = \begin{cases} 
\frac{1}{\lambda} R + T_i - T_a + (2^R - 1)^{-1} & \text{if } \frac{1}{\lambda} R + T_i - T_a \geq 0 \\
(2^R - 1)^{-1} & \text{otherwise.}
\end{cases}
\]  

[28]

Next, we compute the signaling delay \(T_s\) and rate \(R\) that minimizes the worst case error in Eq. 28, i.e.

\[
R^* = \arg \min_{R > 0} \max \left(0, \frac{1}{\lambda} R + T_i - T_a\right) + (2^R - 1)^{-1}
\]  

[29]

\[
T_s^* = \frac{1}{\lambda} R^*.
\]  

[30]

Let us define

\[
C(R) := \frac{1}{\lambda} R + T_i - T_a + (2^R - 1)^{-1}.
\]  

[31]

It can be shown that

\[
\frac{dC(R)}{dR} = \frac{1}{\lambda} - \frac{2^R \log(2)}{(2^R - 1)^2}.
\]  

[32]

The derivative \(dC(R)/dR\) with domain \(R > 0\) satisfies the following three conditions.

- The function \(dC(R)/dR\) converges to negative infinity as \(R \to 0_+\).
- The function \(dC(R)/dR\) is increasing in \(R > 0\).
- If \(C(R)\) has critical points, i.e. \(dC(R)/dR = 0\), then the critical points satisfies

\[
R = \log(L(\lambda))
\]  

[33]

where \(L(\lambda)\) is defined to be

\[
L(\lambda) := \frac{1}{2} \left( (2 + \lambda \log 2) + \sqrt{4\lambda \log 2 + (\lambda \log 2)^2} \right).
\]  

[34]

The value of \(L(\lambda)\) is increasing in \(\lambda > 0\) and

\[
\lim_{\lambda \to 0_+} L(\lambda) = 1.
\]  

[35]

So for any \(\lambda > 0\), the critical points Eq. 33 take a positive value.

From the above conditions, we obtain that \(C(R)\) has a unique minimumizer (i.e. optimal signaling rate \(R\)) for any \(\lambda > 0\). This optimal rate is given by

\[
R^* = \begin{cases} 
\log(L(\lambda)) & \text{if } \frac{1}{\lambda} \log(L(\lambda)) + T_i - T_a \geq 0 \\
\lambda(T_a - T_i) & \text{otherwise.}
\end{cases}
\]  

[36]
Combining Eq. 36 with the component SAT Eq. 21 yields the optimal signaling delay

\[ T^*_s = \frac{1}{\lambda} R^*. \]  

The preceding analysis can also be applied to other scenarios when Eq. 21 is replaced by Eq. 27 and/or when Eq. 6 are replaced by 11. In these scenarios, there also exist a unique delay and data rate that minimizes the performance limits.

Beyond the component SATs in Eq. 21 and Eq. 27, the existence of a unique optimum depends on the specific forms of the component SATs. On the other hand, Diversity-Enabled Sweet Spots can exist even if the delays and rate rates that minimize the fundamental performance limits are not unique. Layering and diversity give more flexibility to optimize the delays and data rates of individual layers according to the subtasks performed at each layer. To see this, we can consider a hypothetical case that the optimal delays and rates is given by an interval (but not a point as in Fig 6B of the main text, the interval can be either connected or not connected). As long as the optimal intervals do not have an intersection for the two layers, the performance can be improved by layering and diversity.

Although there can exist some corner cases when the optimal intervals of the two layers have non-empty intersections, both the theoretical prediction and empirical observations suggest that layering and diversity help improve the system performance in the case studies in this paper: oculomotor control, balancing, and lateral control in humans.

**B. Optimal component delay and data rate in the layered model.** Next, we consider the optimal delay and data rate of the two-layer model in two scenarios: diverse cases vs. uniform cases. In the diverse case, the signaling delays and rates between layers can be heterogeneous. In the uniform case, the signaling delays and data rates are homogeneous. The two cases are assumed to have identical resource constraints to build and maintain the axons in both layers, which are quantified by the total cross-sectional area to build axons.

We use \((T_\ell, R_\ell)\) and \((T_h, R_h)\) to denote the pair of signaling delays and data rates for the lower layer and the higher layer, respectively. Both layers satisfy the component SATs

\[ R_\ell = \lambda_\ell T_\ell, \]
\[ R_h = \lambda_h T_h. \]

Recall from Section 2 that \(\lambda_\ell, \lambda_h\) is proportional to the resource used in the lower layer and the higher layer, respectively. For a fair comparison of the diverse case and the uniform case, we assume that both cases are under the same resource limitations.

So \(\lambda_\ell\) are assumed to be identical for both cases, and so does \(\lambda_h\).

The optimal cost is given by

\[ \sup_{\|x\|_\infty \leq h, \|r\|_\infty \leq 1} \|x\|_\infty = \left\{ T_\ell + T_i + \frac{1}{2R_\ell - 1} \right\} \epsilon + \max(0, T_h - T_\ell) + \frac{1}{2R_h - 1}. \]  

We assume that the advanced warning \(T_a\) is sufficiently large, i.e.

\[ T_a \gg \frac{1}{\lambda_h} \log(L(\lambda_h)), \]

where \(L(\lambda)\) is defined in Eq. 34.

In the diverse case, minimizing Eq. 40 over the pairs \((T_\ell, R_\ell)\) and \((T_h, R_h)\) subject to Eq. 38 and Eq. 39 yields

\[ R^*_\ell = \log(L(\lambda_\ell)) \]
\[ T^*_\ell = \frac{1}{\lambda_\ell} \log(L(\lambda_\ell)) \]

and

\[ T^*_h = T_a - T_i \]
\[ R^*_h = \lambda_h (T_a - T_i). \]

Here, we applied the results from Section 3.A.

In the uniform case, we additionally restrict the signaling delay to be homogeneous

\[ T_\ell = T_h. \]

As the delay is proportional to the cross-sectional area of the axons, this condition can be interpreted as having axons of uniform size. For example, in the context of the oculomotor control system, this can be interpreted as the hypothetical setting where the optic nerves and vestibular nerves have identical size (i.e. They are plotted at the same location in Fig. 2), as opposed to optic nerves being orders of smaller and numerous. Note that Eq. 46 does not imply that the total delays in the control loops are identical in the two layers since the total delays are also influenced by advanced warning and internal delays.
Although the amount of performance improvement from the uniform case to the diverse case depends on the modeling parameters, the worst case error of the diverse case is no greater than the worst case error of the uniform case. This can be seen from

\[
\inf_{T_\ell, R_\ell, T_h, R_h} C(T_\ell, R_\ell, T_h, R_h) \leq \inf_{T_\ell, R_\ell, T_h, R_h} C(T_\ell, R_\ell, T_h, R_h)
\]

where

\[
C(T_\ell, R_\ell, T_h, R_h) := \left\{ T_\ell + \frac{1}{2R_\ell - 1} \right\} \epsilon + \max(0, T_h - T_\ell) + \frac{1}{2R_h - 1}.
\]

Here, the left-hand side of Eq. 47 is the worst-case error in the diverse case, while the right-hand side of Eq. 47 is the worst-case error in the uniform case. The constraint set of the right-hand side is a subset of the constraint set of the left-hand side.

**Extensions and diversity with a layer.** In this paper, we focus on the role of diversity between layers relevant to the observed nerve size and number heterogeneity (Fig. 2). The analysis can be easily extended to the case when \(\lambda_h\) and \(\lambda_\ell\) are also optimization variables, which accounts for the varying resource use due to diverse nerve lengths. And related tools can be used to study diversity within layers. For the theoretical and experimental study of diversity within layers, we refer interested readers to our companion paper (36). The analysis methods presented in this paper and our companion paper can be combined to consider a hybrid of diversity between layers and within a layer.

4. Experiments

**A. Additional results for the multiplex experiments.** Recall from the Material and Methods in the main text that we tested driving behaviors when there are curvatures in the trail, bumps in the road, and both. The first task mainly uses the reflex layer, the second task the planning layer, and the third task both layers. The individual and combined errors are shown in Fig 4.

In addition, we measured the latencies of the two layers in these tasks as follows. We extracted the signals from 900ms in advance of the trail change or a bump to 1800ms afterward. We measured the delay in each control task from the peak in the cross-correlation between the time sequence of the disturbance (bump or trail change) and the time sequence of the control input. To be able to compare with the trail effects, we flipped the sign of bump to make the cross-correlation positive. The results for the first task, second task, and third task are shown in Fig. S4A, Fig. S4B, and Fig. S4C, respectively. The latencies in the first and second tasks are estimated to be 1000.20ms from Fig. S4A and 366.74ms from Fig. S4B, respectively.

Next, we further decomposed the errors in the third task into the error caused by the bump and the error caused by trail changes. The decomposed signal for the trail effects and bump effects are shown in Fig. S4D and Fig. S4E, respectively. The latencies in the estimated trail effects and bump effects in the third task are estimated to be approximately the same as the first task (trail only) and the second task (bump only).

**B. Test the impact of component SATs in control performance.** Recall from the Material and Methods in the main text that we compared the driving behaviors when there are additional delay, quantization, and both from the steering wheel to the trajectory. Other than the worst case framework described in the main text, we also tested the average errors in an average-case framework, illustrated below. In the average case, we generated the turning angle (alternating to left or right) of the trail from the actual trajectory and the desired one. The measured errors for the three tasks are shown in Fig. S4A, respectively. The latencies for the first task, second task, and third task are estimated to be approximately the same as the first task (trail only) and the second task (bump only).

4. References

**Fig. S2.** Theoretical and experimental system SATs in sensorimotor control (average-case). (A) Theoretical SATs in the tracking (driving) task. The delay error (blue), rate error (red), and the total error (black) in Eq. 6 are shown with varying hardware SAT $T = (R - 5)/15$. (B) Empirical SATs in the tracking (driving) task. The error under added delay (blue), the error under added quantization (red); and the error under added delayed and quantization (black) are shown. In the last case, the added delay $T$ and quantization rate $R$ satisfy $T = (R - 5)/15$. 
Fig. S3. Cross correlation between the disturbance (bump or trail change) and the control input. (A) Trail only condition; (B) Bump only condition; (C) Bump and trail together condition; (D) Filtered trail effects from the trail and bump condition; (E) Filtered bump effects from the trail and bump condition. The shadowed area indicates the standard error across subjects.
Error dynamics
\[ x(t + 1) = w(t) + x(t) + u(t) \]

Feedback controller with delay:
\[ T = T_s - T_a + T_i \]
data rate: R

Fig. S4. A block diagram of the system described in Section 1.A.


