An “Aha!” moment precedes the strategic response to a visuomotor rotation

Max Townsend[1], Matthew Warburton[1], Carlo Campagnoli[1], Mark Mon-Williams[1,2,3], Faisal Mushtaq[1,3,4*], J. Ryan Morehead[1]

[1] School of Psychology, University of Leeds, Leeds, LS2 9JT, UK
[2] Bradford Institute for Health Research, Bradford Hospitals National Health Service Trust, Bradford, BD9 6RJ, UK
[3] Leeds NIHR Biomedical Research Centre, Leeds, LS7 4SA, UK
[4] Centre for Immersive Technologies, University of Leeds, Leeds, LS2 9JT, UK
* Corresponding author: f.mushtaq@leeds.ac.uk

Abstract

Strategic behaviour in sensorimotor adaptation tasks is typically modelled either as an error minimisation process or as a process of learning through trial-and-error. The former predicts a gradual reduction in error, until some asymptote, while the latter predicts behavioural exploration to discover an efficacious solution. An alternative explanation is that a sufficiently rich understanding of the task culminates in an “Aha!” moment, which then allows the generation of a new solution. This predicts some period of perseveration in baseline behaviour, followed by an abrupt single-trial shift to a new strategic solution. To avoid obfuscation caused by the motor system, we investigate these hypotheses in a strategy-only aiming game, where participants aim and fire a cannon at targets. Most participants exhibited a period of baseline perseveration followed by a single-trial shift to good performance. We then applied the same analyses to reaching data from a visuomotor rotation task that inhibited implicit adaptation through delaying feedback presentation (Brudner et al., 2016). Similarly, we found that participants typically perseverated in reaching towards the target before suddenly switching, in a single trial, to good performance. These findings suggest that traditional descriptions of strategic behaviour are insufficient and must be updated if we want to understand how humans respond to sensorimotor perturbations.

Introduction

Correcting for movement errors is integral to the success of an embodied agent. Humans can strategically modify their behaviour to expedite this error-correcting process (Bond & Taylor, 2015; McDougle et al., 2015, 2016; Morehead et al., 2011; Taylor et al., 2010, 2014; Taylor & Ivry, 2011, 2012). While recent progress has been made in enumerating core strategic phenomena, and in formulating frameworks from which to investigate these phenomena (Tsay et al., 2024), our understanding of strategic error correction remains shallow.

Traditionally, strategic behaviour in visuomotor rotation (VMR) tasks has been characterised as a gradual error-reduction process. This view predicts a gradual, monotonic increase in learning that tends towards some asymptote, typically modelled by a state-space model (SSM; Coltman et al., 2019; McDougle et al., 2015; Taylor & Ivry, 2011). While consistent with group-averaged learning curves, it is unclear if the individual timecourse of learning evolves in the same manner. Indeed, this gradual monotonic shape can emerge when averaging over variably parameterised individual functions which differ in kind from the group curve (Bahrick et al., 1957; Gallistel et al., 2004). As such, we should also consider that strategic behaviour may instead reflect trial-and-error learning, where actions are selected based on their expected value or inferred efficacy. Here, the agent explores action space (or hypothesis space) in an attempt to discover an efficacious solution. In this view, the evolution of behaviour is either the result of a reinforcement learning (RL) process, or of updating a set of hypotheses about the context, from which the efficacy of actions is inferred (Collins & Koechlin, 2012; Darshan et al., 2014; Donoso et al., 2014; Griffiths et al., 2010, 2019; Piantadosi et al., 2016; Taylor & Ivry, 2011, 2012; Tsay et al., 2024). Learning in this manner typically generates clear signals of behavioural exploration before settling on a strategy. These behavioural patterns are inconsistent with, and dissociable from, gradual error reduction.

Alternatively, it may be the case that a strategic solution to a VMR task is selected upon experiencing an “Aha!” moment (Taylor & Ivry, 2012; Tsay et al., 2024). This “Aha!” moment might reflect a gain of insight—a qualitatively richer understanding of the task and its demands—that proceeds from reorganising one’s internal representation of the problem (Carpenter, 2020; Jones, 2003; Knoblich et al., 1999, 2001; Kounios & Beeman, 2009; Sprugnoli et al., 2017). Indeed, we see this kind of sudden learning in many areas of cognitive problem solving, such as the compound-word remote associations task (Jung-Beeman et al., 2004), anagrams (Novick & Sherman, 2003), rebus puzzles (Salvi et al.,

2016), the nine-dot problem and it's variations (MacGregor et al., 2001), among others (e.g. Sawyer, 2012; Vartanian & Goel, 2007).

The strategic problem presented during the baseline (veridical) phase of a VMR task may be simply formulated as the requirement to aim at the target before firing. A change in the action-feedback mapping renders this parsimonious and effective representation uninformative as to how to act successfully. By reorganising this representation to embody more semantic and relational information, an appropriate solution can be brought psychologically closer (Trope & Liberman, 2010). For example, when the action-feedback mapping is modified, a representation of the task that encodes a relationship between the aimed direction and the feedback direction would more naturally lead to an efficacious aiming solution. Learning would manifest, behaviourally, as a period of perseveration of baseline behaviour, followed by a sudden, substantial shift in behaviour following the "Aha!" moment.

To date, it has remained difficult to elucidate the true nature of strategic behaviour in VMR tasks as extant task designs tend to leave strategic behaviour inextricably linked to implicit adaptation and other contaminants of the motor system. This is either because one measure is used to infer the output of two or more processes, where it is difficult to correctly assign credit to these underlying processes (e.g. Hegele & Heuer, 2010), or because these processes interact in a way which we do not currently understand. To make matters worse, different methods of probing strategic behaviour generate distinct patterns of results (Langsdorf et al., 2021; Maresch, Mudrik, et al., 2021; Maresch, Werner, et al., 2021). This is likely because each method changes the nature of the interaction between the strategic system and the motor system. A different approach is to first understand each subprocess in relative isolation, such that we can use the findings as priors when studying the whole process. We can already remove implicit learning from the strategic context by presenting consistent clamped feedback which is irrelevant to performance (Morehead et al., 2017). We aim to complement this by isolating strategic learning from the motor-learning context, while enforcing the same strategic demands.

Here, we investigated aiming behaviour in a series of cannon-shooting tasks where the participant had to overcome a rotation of the visual feedback (Figure 1a). We found that aiming behaviour evolved in a manner that was consistent with an "Aha!" moment, but not with trial-and-error learning or gradual error minimisation. This pattern of baseline perseveration, followed by a delayed single-trial shift to good performance, is clear both visually and through formal model comparison. We quantitatively confirmed that our task presented similar strategic demands as in a standard VMR task and then carried out the same modelling analyses on a previous dataset from a delayed-feedback VMR task

(Brudner et al., 2016). We found the same pattern of results, where almost all participants were better described by an “Aha!” Moment than by trial-and-error learning or gradual error minimisation.

Methods

Participants and ethics

A total of 612 humans (252 males, 355 females, 5 other genders, age = 34.7 ± 11.8 years) from the United Kingdom and the United States were recruited via the Prolific Academic recruitment platform. Ethical approval was obtained from the University of Leeds school of Psychology Research Ethics Committee (ref PSYC-81).

Experimental Apparatus

Participants aimed a cannon and shot targets using keyboard keys. All participants completed the task remotely on their own desktop or laptop computers, with no further constraints on the hardware they used. The experimental task was constructed using the Unity game engine and default libraries, with custom scripts written in C# and data saved to a remote database.

Aiming Task

Participants were instructed to rotate the cannon by pressing the ‘A’ and ‘D’ keys, and to fire it with the ‘Spacebar’ key (Figure 1a). They were told to hit as many targets as possible and were not encouraged to act quickly. Tapping the rotation keys rotated the cannon by 1° increments and holding down these keys rotated the cannon at the speed of 125°sec^{-1}. An aiming reticule indicated the aimed location, whose radial distance was held constant (equal the radial distance of the target). Upon firing, the cannonball flight lasted 300 ms, and would terminate at the same radial distance as the target. The target had a radius of 3°, and the cannonball had a radius of 0.5°. Upon being hit, the target exploded and was accompanied by a green ‘Hit!’ text lasting 500 ms; if missed, a ‘buzzing’ sound accompanied red ‘Miss!’ text lasting 500 ms. The Cannon Game task would only play in full-screen mode. At any point, if the participant left full-screen mode, the task was paused and the participant was presented with a pause screen asking them to enter full-screen when they were ready to resume.

Participants were told ‘the cannonball may not always go exactly where you are aiming’. The direction of the shot was subject to noise of similar magnitude to motor noise in reaching tasks (Figure 1b), such that the cannonball endpoint feedback was

$$EndPoint° = aim° + Rotation° + \epsilon°, \qquad \epsilon° \sim N(0,3)° \tag{1}$$

where the integer aim is the angle aimed, ϵ is gaussian noise, and $Rotation$ is the feedback rotation applied to the endpoint feedback. At the start of each trial, the cannon direction was sampled from a uniform random distribution spanning 60° either side of, but at least 5° away from, the angle equal to $-Rotation°$ (i.e., the ideal aim angle).

Experiment 1

The first experiment sought to explore how participants updated their aiming behaviour in a simple task which requires the participant to aim at targets, much like in a visuomotor rotation (VMR) task.

In *experiment 1a* (n = 98), the target was presented in the same position on every trial. The experiment comprised six blocks, where each block included 30 rotation trials with a rotation applied to the direction of feedback. The rotation phase was sandwiched between 15 baseline and 15 washout trials, each with veridical feedback (Figure 1a). Within each block, the rotation phase presented a single feedback rotation for all trials, such that one group of participants each experienced every sign/magnitude combination of $\{\pm 15, \pm 30, \pm 45\}°$, and the other group experienced all of $\{\pm 45, \pm 60, \pm 90\}°$ by the end of the six experimental blocks. Five participants were excluded for producing too many outlying data points (see Data Analysis).

In *experiment 1b* (n = 152), the target was presented in one of eight locations, evenly spaced in 45° increments from the target. The order of target location was pseudorandomised, where each location occurred once in each eight-trial cycle. There was a single block of 40 cycles of rotated feedback, sandwiched between five baseline and five washout cycles. During the rotation phase, each participant experienced one rotation, sampled from $\{\pm 15, \pm 30, \pm 45, \pm 60, \pm 90\}°$. Twenty participants were excluded for producing too many outlying data points or having extreme, persistent variability in aiming (see Data Analysis).

Experiment 2

Experiment 2 implemented aiming requirements which were much more similar to those in VMR tasks. To this end, it included an autocorrection module (i.e., a spatially-generalising SSM adapted from McDougle et al. (2017)) which learns a bias that is added to the feedback direction. This bias gradually minimises the difference between the aimed location and the endpoint feedback (Figure 2a & b). The resultant formula for the endpoint feedback direction is

$$EndPoint° = aim° + Rotation° + \overrightarrow{x_t}(aim)° + \epsilon°, \qquad \epsilon° \sim N(0,3)° \tag{2}$$

where $\overrightarrow{x_t}(aim)°$ is the output of the autocorrection module for integer angle aim on trial *t*. The state vector $\overrightarrow{x_t}$ is updated after each trial *t* according to

$$\overrightarrow{x_{t+1}} = A\overrightarrow{x_t} - B\exp[(\vec{x} - \mu_t)^2/2\sigma^2] \times e_t \quad (3)$$

where $\overrightarrow{x_{t+1}}$ is the state vector on trial *t+1*, A is the retention factor, B is the learning rate, e_t is the error $(EndPoint_t - TargetAngle_t)$, $\vec{x}$ is the ordered vector of all integer angles around the workspace (-179,-178…,180), and μ_t is the centre of gaussian generalisation on trial *t.* The parameter μ_t is set to the aim angle on trial *t*, such that the generalisation of learning is centred on the aimed angle (Figure 2a & b). Parameter σ, which defines the width of generalisation, was set to 22° (selected to be consistent with previous generalisation widths (Krakauer et al., 2000; Mattar & Ostry, 2007; McDougle et al., 2017; Taylor & Ivry, 2013)).

For *Experiment* 2a (n = 61), the parameters A and B were set to 0.99 and 0.0275 respectively (consistent with the values found in McDougle et al. (2015)). The target was presented in a single location and the task comprised six 60-trial blocks, each including a 30-trial rotation phase sandwiched between a 15-trial baseline phase and 15-trial washout phase. One group experienced one of $\{\pm15, \pm30, \pm45\}°$ in each block, and the other group experienced one of $\{\pm45, \pm60, \pm90\}°$, akin to experiment 1a. Three participants were excluded for producing too many outlying data points or having extreme, persistent variability in aiming (see Data Analysis).

For *Experiment 2b* (n = 301) we wanted to approximate actual aiming requirements experienced by participants of a VMR study. To do this, we fit the learning rate (A) and retention rate (B) parameters of the autocorrection module to data from Bond & Taylor's (2015) eight-target reaching task. Parameter fitting was carried out by finding the maximum likelihood over all participants within each rotation size separately, using the *optimize.minimize* function (with the 'L-BFGS-B' method) of the *scipy* (scientific python) package for Python (parameter fits in Table 1). As in experiment 1b, the target was presented in one of eight locations, evenly spaced in 45° increments around the workspace. The order of target location was pseudorandomised, where each location occurred once in each eight-trial cycle. There was a single block of 40 cycles of rotated feedback, sandwiched between five baseline and five washout cycles. During the rotation phase, each participant experienced one rotation, sampled from $\{\pm15, \pm30, \pm45, \pm60, \pm90\}°$. Fifty-three participants were excluded for producing too many outlying data points, showing extreme, persistent variability in aiming, or for aiming at the target throughout the task (see Data Analysis).

Table 1. Learning rate and retention rate parameter values for the autocorrection module used in experiment 2b. These are the result of fitting to data from Bond & Taylor (2015).

Rotation (°)	Autocorrection module learning rate	Autocorrection module retention rate
15	0.0385	0.9985
30	0.0152	0.9956
45	0.0152	0.9951
60	0.0118	0.9875
90	0.0044	0.9918

Data Analysis

Analyses were completed with custom scripts written in Python 3.7. We used an α-value of 0.05, such that any p-value less than α indicates a statistically significant difference. Unless stated otherwise, all t-tests were two-tailed and Bonferroni corrections were applied for multiple comparisons. Our primary dependent variable was the signed difference between the aim angle and the target angle, with the resulting angle multiplied by -1 for positively signed rotations so that a positive angle was one that counteracted the applied perturbation.

Eighty-one participants were excluded (all previously listed *n* values were calculated before participant exclusion) because they met at least one of the following criteria: four participants were excluded because they didn't learn the task (their mean aim was within two degrees of zero); 19 participants were excluded because their median absolute deviation in task error was above 15 (generally where a participant never converged on an appropriate solution to counteract the applied rotation); 58 participants were excluded because they had greater than 20% of trials classified as outliers. An outlier satisfied any of the following criteria: the aiming time was less than 100ms, or the aim error, aiming time, or reaction time on a given trial (other than the first five trials of a rotation period) was greater than 3 * Inter-Quartile Range (IQR) away from that participant's median for that epoch. For experiments 1a and 2a, each non-overlapping epoch comprised 15 trials; for experiments 1b and 2b, each non-overlapping epoch comprised 5 cycles. The use of the IQR across an epoch meant that periods of high variability were preserved, for example highly explorative aiming, whereas outlier values in otherwise stable epochs were removed, for example lapses once an appropriate solution had been found that were not informative to how the solution was discovered. After participant removal, 7.6% of the remaining trials were excluded as outliers.

In experiments 1 & 2, we used t-tests to check that baseline aiming was not significantly different from zero, and to check whether aiming was significantly different from zero during mid learning and late learning. We also tested whether or not the washout aiming (excluding the first trial/cycle) significantly differed from zero. In experiments 1a & 1b, t-tests were used

to assess whether aiming was significantly different from the ideal aim angle during mid and late learning. In experiments 2a & 2b, we instead tested if the 'final cannonball angle' (aim + offset learned by autocorrection module) was significantly different from the ideal cannonball angle during late learning and mid learning. We also applied a t-test on the implicit autocorrection vs zero during washout. In experiments 1a & 2a, mid learning was defined as the middle five rotation trials, and late learning was defined as the final five rotation trials. In experiments 1b & 2b, mid learning was defined as the middle five rotation cycles, and late learning was defined as the final five rotation cycles. In experiments 1a and 2a, where participants encountered a given rotation magnitude twice with opposite signs, data was averaged over both blocks.

For the direct comparison to VMR data, we compared data from experiment 2b to data from Bond & Taylor (2015). A 2 (experiment type) x 5 (rotation size) x 3 (learning phase) ANOVA was carried out, where experiment type was denoted whether the experiment was ours or from Bond & Taylor, and learning phase was either early learning (first 10 rotation cycles), mid learning (middle 10 rotation cycles), or late learning (final 10 rotation cycles). We also applied t-tests to compare the following properties across experiments: the cycle of peak aiming, the magnitude of peak aiming, and the decrease in aiming from peak to final rotation cycle.

Model-based analyses

As part of our investigation into the individual timecourse of aiming, we compare three models of noisy aiming: a single-process SSM to model monotonic, gradual learning; a step-function to model learning that involves a singular "Aha!" moment; and a trial-and-error learning model.

The state of the gradual-learning model, x_t, determines the aim on trial t. We assume that the process of executing an intended plan is inherently noisy, such that the aim on trial t is defined

$$aim_{t,Gradual} = x_t + \epsilon \tag{4}$$

where x_t is the internal state of the SSM, and the execution noise ϵ follows an unbiased normal distribution

$$\epsilon = \mathcal{N}\left(\mu_c, \sigma^2_{c,Gradual}\right) \tag{5}$$

defined by $\mu_c = 0$, and variance of execution variability $\sigma^2_{c,Gradual}$. The state is updated according to

$$x_{t+1} = Ax_t - Be_t \tag{6}$$

where A is the retention rate, B is the learning rate, and e_t is the error on trial t, as defined in equation 3. The SSM has three free parameters: A, B, and $\sigma^2_{c,Gradual}$.

According to the "Aha!" model, the aim on trial t is defined as

$$aim_{t,Insight} = \begin{cases} 0 + \epsilon, & t < changepoint \\ h + \epsilon, & t \geq changepoint \end{cases} \tag{7}$$

where h is the step-size parameter determining the change in aim, $changepoint$ is the trial at which the step change occurs, and ϵ is unbiased gaussian execution noise as defined in equation 5 but with $\sigma^2_{c,Insight}$ as the variance parameter. This single-step-learning model has three free parameters: h, $changepoint$, and $\sigma^2_{c,Insight}$.

For the trial-and-error learning model, we reframe the model from Cashaback et al. (2019) to specify that exploration is carried out in aiming space rather than motor space. We assume that, on each trial, there is a distribution of candidate aims from which the participant selects their executed aim. Therefore, the aim on trial t is defined

$$aim_{t,Reinforce} = \mathcal{N}(\overline{aim}_t, \sigma_t^2) \tag{8}$$

Where $\overline{aim}_t$ is the mean of the current candidate aims (which we assume here to be normally distributed), and σ_t^2 is the variance of aiming variability. The mean of the current candidate aims and the variance of aiming variability are updated on each trial according to

$$\overline{aim}_{t+1}, \sigma^2_{t+1} = \begin{cases} \overline{aim}_t + \alpha\left(aim_{t,Reinforce} - \overline{aim}_t\right), \sigma^2_{c,Reinforce} & r = 1 \\ \overline{aim}_t, \sigma^2_{c,Reinforce} + \sigma_e^2 & r = 0 \end{cases} \tag{9}$$

such that the variance in aiming is equal to execution variance ($\sigma^2_{c,Reinforce}$) after a successful trial, but equal to the sum of both execution variance and exploratory variance (σ_e^2) after an unsuccessful trial. The mean execution variability used in the other models is zero ($\mu_c = 0$), so incorporating the variance of execution variability ($\sigma^2_{c,Reinforce}$) is equivalent to adding execution variability (ϵ) to the model prediction. Therefore, the resulting prediction is directly comparable to the predictions of the two other models. After a successful trial, the mean intended aim for the next trial ($\overline{aim}_{t+1}$) is updated by a proportion (α) of the difference between the observed aim (aim_t) and the current mean intended aim. The RL model has three free parameters: α, $\sigma^2_{c,Reinforce}$, and σ_e^2.

All models were fit to the rotation phase of the first block of each participant in Experiment 1a separately. Subsequent blocks were excluded as behaviour there might be contaminated by learning from previous blocks. Parameters were optimised through maximum likelihood estimation (MLE). For the gradual-learning model, the negative log-likelihood was minimised using the *optimize.minimize* function. However, the fitness function of the single-step-

learning model is non-differentiable, as there is a vertical tangent at each real-valued setting of the $changepoint$ parameter. As a simple solution, we ran the *optimize.minimize* function on the σ_c^2 and h parameters at every appropriate real-valued setting of $changepoint$, and then selected the best set of parameters. The behaviour of the Cashaback model at time t depends on the outcome of a random variable at time $t-1$. To account for this stochasticity, we produced, for each function evaluation, 1000 simulated runs using the selected parameter values, and then minimised the mean negative log-likelihood over all these runs. This typically resulted in millions of simulated runs for each participant. The objective functions for each of the three models were optimised using the gradient-free 'Powell' method as it handled the stochasticity of the Cashaback model much better than the 'BFGS' algorithm (it also produced marginally better fits for the gradual-learning model).

Results

Experiment 1

We set out to characterise aiming behaviour in a shooting task, where participants must aim and fire a cannon to hit targets (Figure 1a). This experiment was designed to explore aiming in its purest form; isolated from any implicit motor-learning system (such as adaptation). In condition *a* of this experiment, participants were exposed to either the ±15°, ±30°, and ±45° rotations, or the ±45°, ±60°, and ±90° rotations, with only one target location. Participants quickly learnt to fully compensate for all rotation sizes by aiming in the opposite direction (Figure 1b, *left*).

During the baseline phase, prior to all rotations, there was no bias in aiming (t-tests against zero; Bonferroni-corrected *p*s > 0.334). By the final five rotation trials of the block for all rotation sizes, participants had learned to respond to the feedback rotation by aiming in the opposite direction (Figure 1c, left; t-tests against zero; *p*s < 0.001). For all rotation sizes, aiming during these last five trials was not significantly different from the ideal aim angle (t-tests against ideal aim angle; *p*s > 0.118). For all rotations (*p*s > 0.389), except 90° (95% CI, 77.2°, 85.1°, $t_{(44)}$, $p = 0.016$), this ideal aim angle was already achieved by the middle five trials (trials 13 to 17). After the first trial of the washout phase, there was no residual bias in aiming (t-test against zero; *p*s > 0.064).

In condition *b,* five new groups of participants were each exposed to one of ±15°, ±30°, ±45°, ±60°, or ±90° rotations. With eight target locations, this condition comprised a single block presenting one rotation size. For all rotation sizes, participants learnt to fully compensate for the rotation.

For all groups, there was no bias in baseline aiming (*p*s > 0.086), and by the final five cycles of the rotation phase, aims had significantly changed (Figure 1c; *p*s < 0.001). During this late-rotation phase, aiming was smaller than the ideal aim angle in the 30° group (95% CI, 28.42°, 29.75°, $t_{(27)}$, $p = 0.045$), but no other group (*p*s > 0.072). By the middle five cycles (cycles 18 to 22), aims, all groups, except the 30° group (95% CI, 28.63°, 29.79°, $t_{(27)}$, $p = 0.049$), achieved aims which were not different from ideal (*p*s > 0.067). After the first washout cycle, there remained a small bias in aiming towards the previously ideal aim angle in the 15° (95% CI, 0.15°, 0.95°, $t_{(26)}$, $p = 0.044$) and 30° (95% CI, 0.16°, 0.51°, $t_{(27)}$, $p = 0.003$) groups, but no other (*p*s > 0.105).

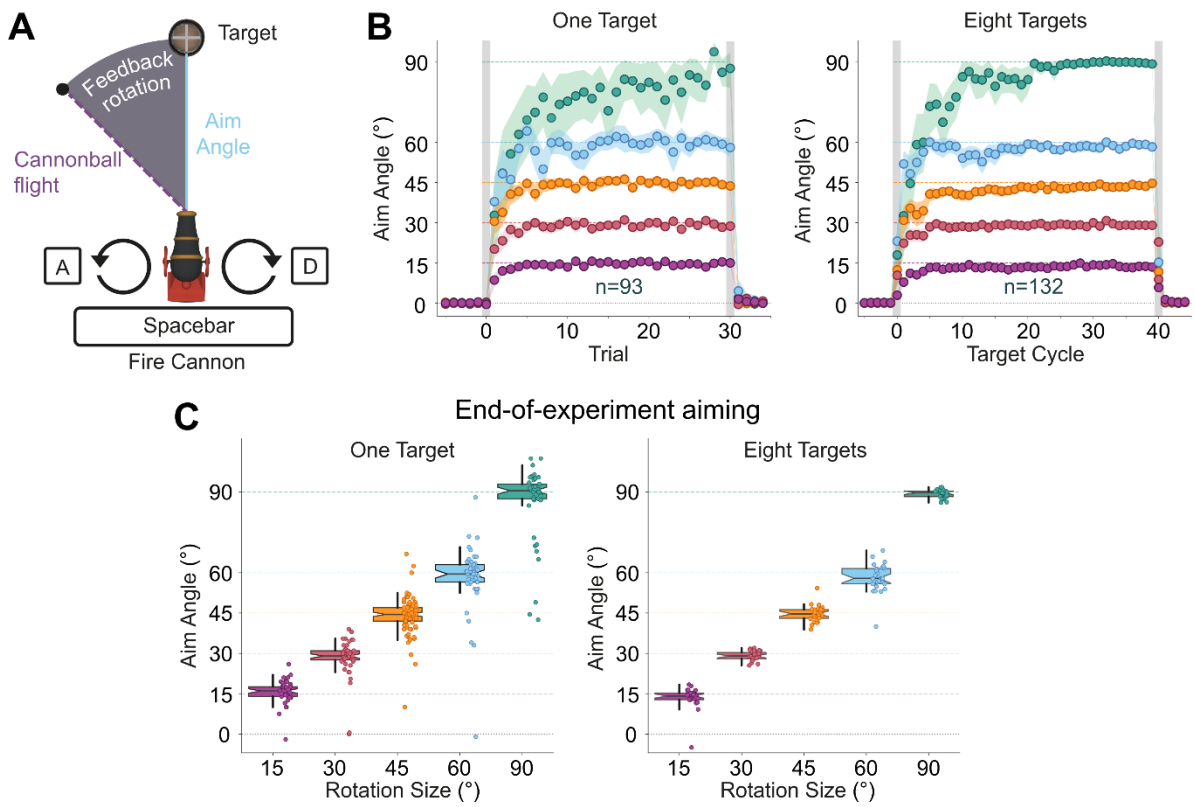

Figure 1. **(A)** The participant must aim the cannon to shoot a target. The cannon is rotated using the 'A' and 'D' keys and fired using the 'Spacebar' key. **(B)** Aiming behaviour over the course of experiments 1a, with one target location (left), and 1b, with eight target locations (right). **(C)** Aim angles during the final trial or cycle, for experiments 1a (left) and 1b (right), respectively. Dashed lines indicate ideal aim angles. Error bars/bands = 95% CI.

Experiment 2

The second experiment sought to bring the task requirements closer to those in a VMR task. This was achieved by adding in an autocorrection module to the cannon (Figure 2a & b). The autocorrection module automatically and covertly corrected for task error in a monotonic fashion (akin to the typical SSM formulation of implicit adaptation; see Methods section). Participants learned to fully compensate for all feedback rotations. The initial rapid increase in aiming was followed by a slow decrease in the amount reaimed, such that the cannonball

would persist in going towards the target region as the autocorrection module learns to compensate more (Figure 2c & d).

In condition *a* of experiment 2, there was one target position, and participants were exposed to either the ±15°, ±30°, and ±45°, or the ±45°, ±60°, and ±90° rotations. During the baseline phase, there was no bias in aiming for any rotation size (*p*s > 0.713). By the middle five trials of the block, aiming under all rotation sizes was greater than zero (*p*s < 0.001), but also smaller than what the ideal aim would be if there was no autocorrection (*p*s < 0.001). When combined with the output of the autocorrection module, the resulting cannonball angle during the middle five trials was different from the ideal angle in the 45° rotation (95% CI, 45.75°, 48.75°, $t_{(57)}$, $p = 0.020$), but not in any other rotation (*p*s > 0.320). By the final five trials of the rotation phase, the final cannonball angle was not different from ideal in any rotation size (*p*s > 0.507). At the end of the rotation phase, the aim angle had decreased from the midpoint of the block for all rotations (*p*s < 0.044) except the 90° rotation (95% CI, -13.54°, 2.41°, $t_{(28)}$, $p = 0.820$). After the first trial of the washout period, aiming in all rotations was biased in the direction opposite the previously ideal angle (*p*s < 0.001). This effect was because of the persisting effects of the autocorrection module, whose learned bias now offset the cannonball angle in the opposing direction (*p*s < 0.001).

In condition *b,* there were eight target positions, and five new groups of participants were each exposed to one of ±15°, ±30°, ±45°, ±60°, or ±90° rotations in a task comprising a single block. During the baseline phase, there was no aiming bias for any group (*p*s > 0.166). In the middle five cycles of the task, participants (combined with autocorrection) did fully compensate for the rotation in all groups (t-test vs ideal; *p*s > 0.056). There was much variability in the cycle index at which participants exhibited their greatest cycle-averaged aim (Figure 2f, rightmost panel); the mean (±standard deviation) number of cycles until peak aim were 3.9 (±5.2), 5.4 (±5.6), 6.6 (±7.8), 8.9 (±8.6), and 10.5 (±10.4) for the 15°, 30°, 45°, 60°, and 90°, respectively, such that people tended to reach their peak aims later in the experiment under larger rotations. During their cycle of peak aim, participants aimed 12.98° (±3.33°), 30.53° (±3.76°), 43.75° (±5.73°), 63.54° (±11.22°), and 96.55° (±9.85°), in their respective groups. By the final rotation cycle, aiming reduced to 2.68° (±2.35°), 18.80° (±2.35°), 27.57° (±6.02°), 48.31° (±7.02°), and 81.57° (±6.29°), resulting in a decrease from peak to final cycle of 10.30° (±3.86°), 11.73° (±4.03°), 16.18° (±6.52°), 15.23° (±13.29°), and 14.98° (±11.41°), respectively.

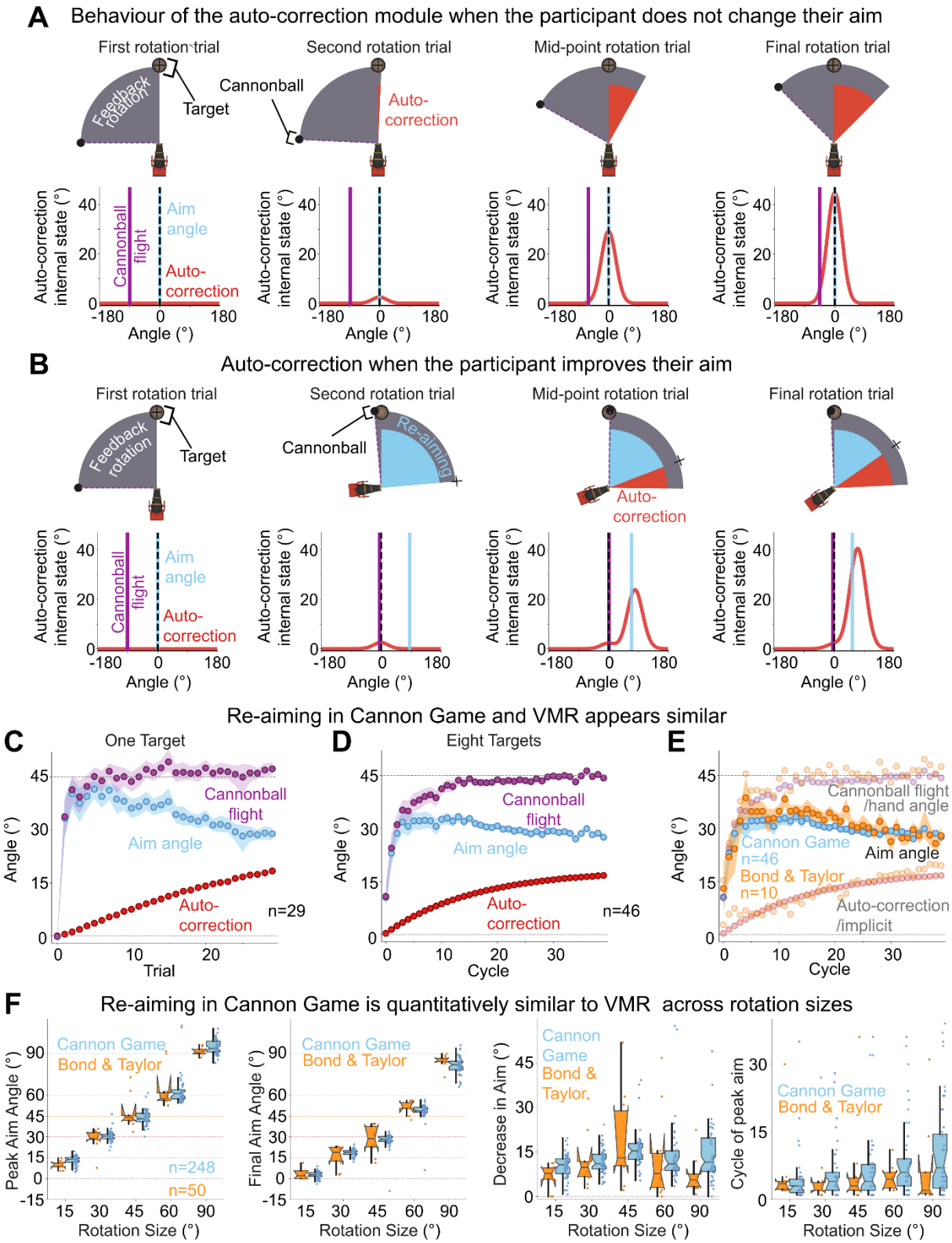

Figure 2. In experiments 2a and 2b, the cannon gradually learned to auto-correct for the feedback rotation. **(A)** & **(B)** The direction of cannonball flight (purple) is equal to auto-correction (red) + re-aiming (blue) - feedback rotation (grey). Timecourse of learning for the auto-correction module when a participant doesn't adjust their aim, and timecourse of learning when a participate learns to re-aim, respectively. Top row: The contribution of each learning system (human re-aiming and cannon auto-correction), and the task outcome, on a given trial. Bottom row: internal state of the auto-correction module (red) generalises around the workspace. Performance in the 45° condition,

under a one-target **(C)** or eight-target **(D)** experiments. **(E)** Performance during the rotation phase in experiment 2b (blue) compared to data from Bond & Taylor (2017; orange). Auto-correction model (SSM) parameters were fit to implicit learning in Bond & Taylor; see methods section. Auto-correction/implicit learning and cannonball flight/hand angle are included for transparency, but de-emphasised as the aiming comparison is primary. **(F)** Comparison of summary statistics for each rotation size. In order, from left to right: each participant's greatest cycle-averaged aiming value; each participant's aiming value on the final rotation cycle; each participant's final aim subtracted from their peak aim; the cycle number at which each participant's peak aim was observed. Error bars/bands = 95% CI.

Direct comparison with VMR task

To explore whether aiming behaviour is consistent with that observed in VMR tasks, we now directly compare key properties of learning between experiment 2b and the eight-target experiment of Bond & Taylor (Figure 2e & f). We find remarkable similarity between the two sets of aiming data, indicating that insights discovered in Cannon Game may also apply to the strategic decision-making processes employed in a VMR task.

We first conducted a 2 (experimental paradigm; Explicit vs Reaching) x 5 (rotation size) x 3 (learning phase) ANOVA on aim angles and found no main effect of experimental paradigm [$F_{(1, 298)} = 2.17$, $p = 0.141$] and no significant interaction effect of learning phase x paradigm [$F_{(2, 298)} = 0.41$, $p = 0.661$], rotation size x paradigm [$F_{(4, 298)} = 2.33$, $p = 0.055$], or the three-way interaction [$F_{(8, 298)} = 0.95$, $p = 0.475$]. Between the two experiments, there was no difference in the cycle index of peak aim for the for any rotation size (*p*s > 0.075; Figure 2f, rightmost panel). Within this cycle of peak aiming, participants in Cannon Game overshot more than Bond & Taylor (Figure 2f, leftmost panel) for the 90 ° rotation (96.55° versus 91.62°; 95% CI, 1.51°, 8.35°, $t_{(50.56)}$, $p = 0.028$), but there was no difference for any other rotation (*p*s > 0.077). In the final rotation cycle, there was no difference in aiming between experiments for any rotation (*p*s > 0.554; Figure 2f, second panel). There was a greater decrease in aiming (the difference between peak aim and final-cycle aim; Figure 2f; third panel) in Cannon Game (14.98°) than Bond & Taylor (7.17°) for the 90° rotation (95% CI, 2.47°, 13.14°, $t_{(24.45)}$, $p = 0.029$) , but no other (*p*s > 0.155).

The timecourse of individual aiming

Learning in all experiments thus far appears to evolve in a gradual, monotonic fashion until the ideal aim angle is reached. However, this gradual, monotonic group curve could result from averaging over variably parameterised individual functions which may not be gradual and monotonic themselves (Bahrick et al., 1957; Gallistel et al., 2004). Through model-based analysis of data from experiment 1a, we investigate learning at the individual level.

We use three models in this analysis: to model gradual, error-based learning, we use a single-process SSM; to model "Aha!" moment learning, we use a simple step-function model; and to model trial-and-error learning, we reframe the model formulated in Cashaback et al. (2019; see Methods section for more details). In Figure 3, we present idealised simulations,

without execution noise, but with a realistic amount of variability in model parameters. While the SSM and “Aha!” models would yield a similar group curve on a traditional trial-series plot, we can see clear qualitative differences when using a heatmap to preserve individual heterogeneity (Figure 3a vs b). The Cashaback model is not capable of learning fast enough to generate the same group curve, and results in aiming behaviour that remains highly stochastic (centred on the target) for many “participants” throughout (Figure 3c). For comparison, we present a heatmap of the actual aiming data from experiment 1a in Figure 3d & e.

We fit all three three-parameter models to the first block of each individual’s aiming data in Experiment 1a (see methods for details). Parameter fits for all models can be seen in Figure 4. Of the 93 participants, 82 were better fit by the “Aha!” model, 11 by the SSM, and none by the Cashaback model (Figure 3f). Combined with visual inspection of the heatmaps, this superior fit of insight learning makes a clear case for a (delayed) single large shift in aiming being a better description of most people’s behaviour, as opposed to gradual error-based learning or trial-and-error learning.

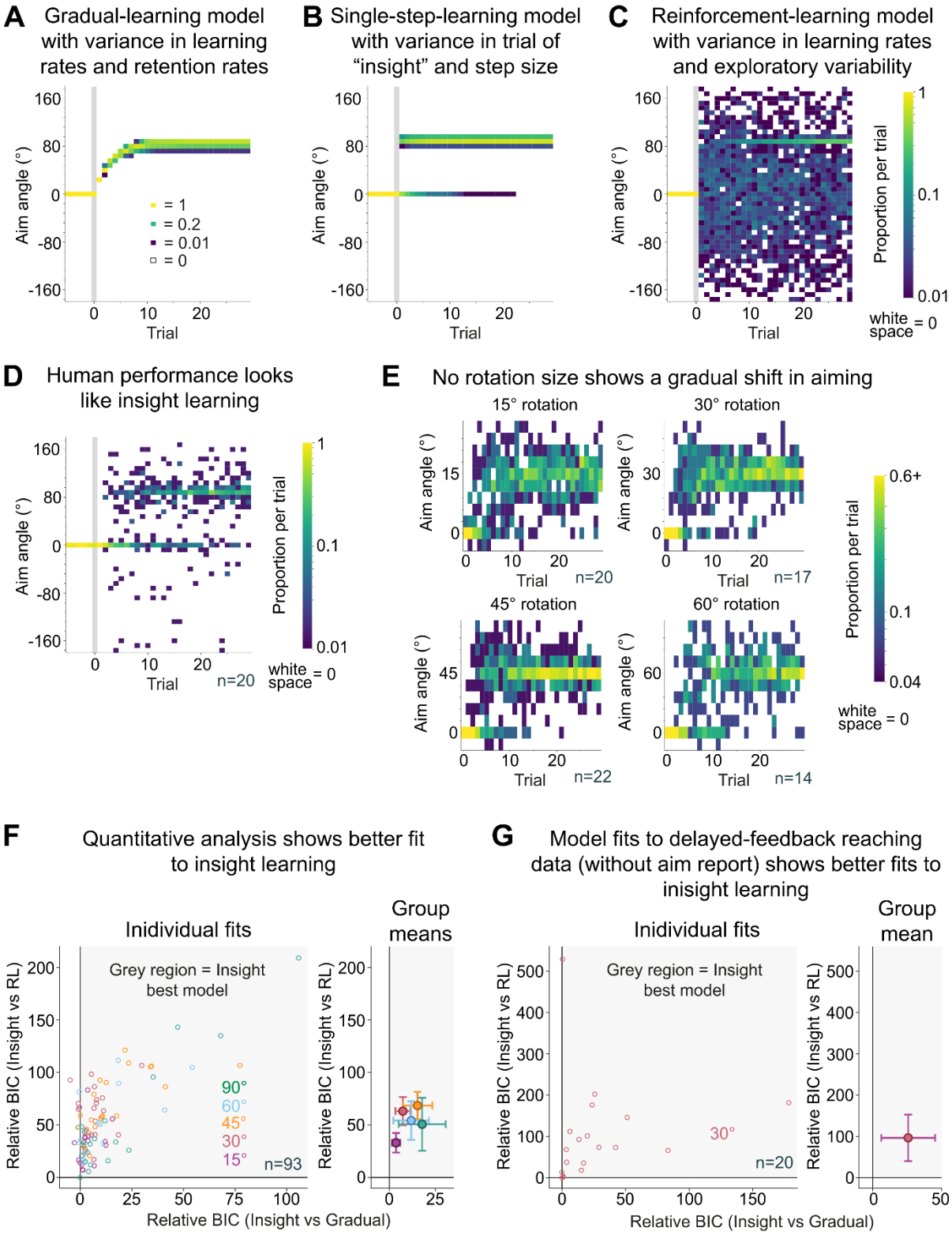

Figure 3. Proportion of simulated aims within each eight-degree arc when aiming increases gradually and monotonically **(A)** versus when aiming changes abruptly as in insight learning **(B),** and when learning occurs through reinforcement (**C**; model from Cashaback et al. (2019)). One hundred participants, subjected to a 90° rotation, were simulated for each model, and execution noise was set to zero. **(D)** The actual aiming data from the first block of experiment 1a (only if the first block was a 90° rotation). Note log scale on colourmap. **(E)** Binned aiming data for all other rotation sizes; only data from the participant's first block was used. Note adjusted colourmap to emphasise medium vs low population bins. **(F)** Relative BIC scores for the insight vs gradual learning

(x-axis) and insight vs RL (y-axis) models. More positive scores indicate a better fit by the insight model, such that the grey region means that the insight model performed better than both the other models. **(G)** The same modelling analysis carried out on reaching data from experiment 2 of Brudner et al. (2016), where both binary and cursor feedback were delayed 5000ms. Note that there were no aim reports used in the Brudner et al task, and there were 160 trials with four target locations. Error bars = 95% CI.

Individual aiming in delayed-feedback reaching tasks

To investigate the extent to which the “Aha!” moment is also present in reaching tasks, we carried out the same modelling analyses as above on data from experiment 2 of Brudner et al. (2016). Importantly, participants in this task did not report their aims, nor were they explicitly prompted to think about re-aiming as a solution. Instead, to diminish the contribution of implicit processes, both cursor and binary feedback were delayed by 5000ms. As can be seen in Figure 3g, the insight model greatly outperforms both other models, such that it better fits 14 out of the 16 participants who learned the task. For completeness, the four non-learners who didn’t change their hand heading angle over the block (ps > 0.1) are included in the analysis. Parameter fits for all models can be seen in Figure 4.

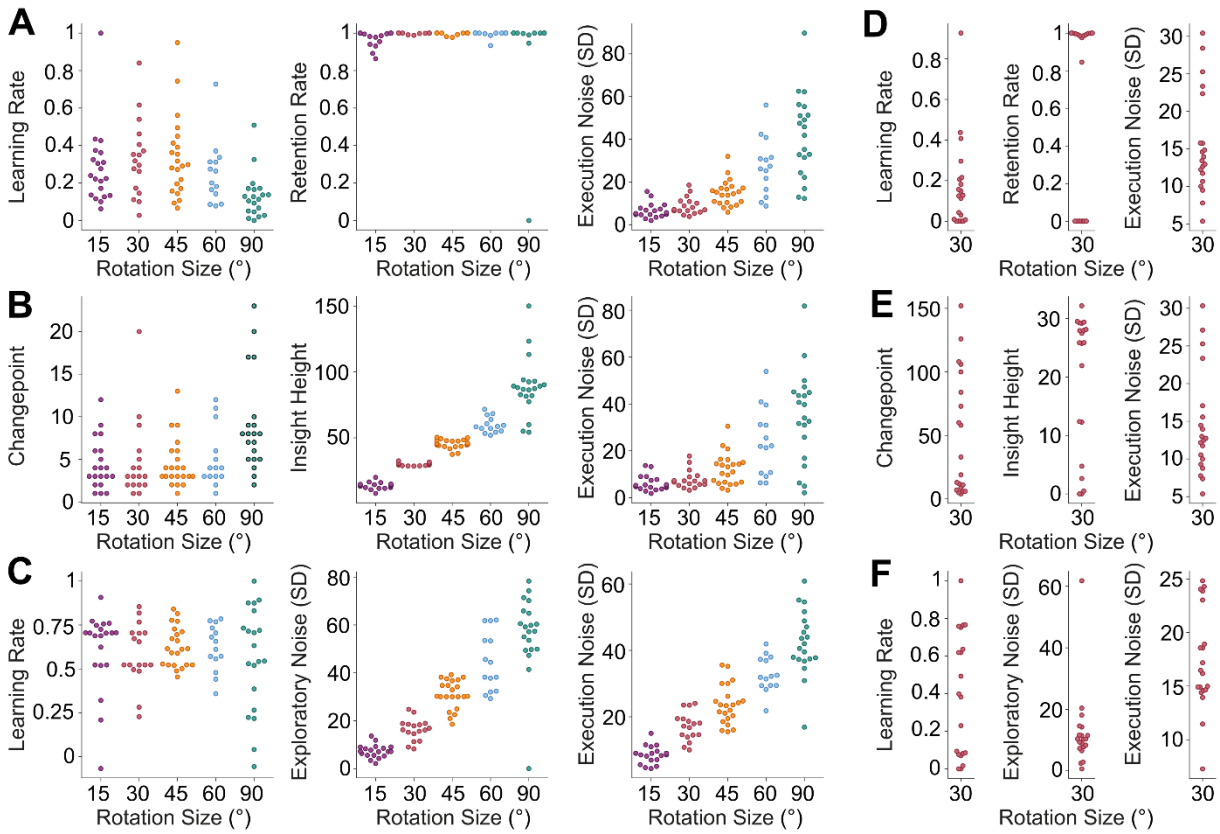

Figure 4. Parameter fits to Cannon Game (**A-C**) and Brudner et al. (2016) (**D-F**) used in model comparison. State-Space Model parameter fits in panels **(A)** and **(D)**, Insight-Learning Model parameter fits in panels **(B)** and **(E)**, and Reinforcement-Learning Model (Cashaback et al., 2019) parameter fits in panels **(C)** and **(F)**.

Discussion

Strategic behaviour in VMR tasks is commonly characterised as either a gradual error-minimisation process, or as a process of trial-and-error learning. We sought to isolate strategic behaviour, and to investigate whether either of these characterisations hold true at the individual level. We found that individual learning functions do not gradually increase towards the ideal aim angle, nor do they exhibit the exploratory behaviour predicted by trial-and-error learning. Instead, it appears that individuals perseverate in aiming at the target before experiencing an “Aha!” moment and suddenly shifting their aim near to the ideal angle. The gradual, group-level learning curve is an artefact of averaging over the heterogeneous duration of perseveration which precedes the “Aha!” moment. We re-examined data from a delayed-feedback VMR task, which had no aim reports or prompts (Brudner et al., 2016), and uncovered the same individual pattern of perseveration in baseline reaching followed by an “Aha!” moment and a sudden shift to good performance.

Our aiming task asks participants to aim and fire a cannon at on-screen targets. This allowed us to measure aiming without the motor contaminants present in VMR reaching tasks. We explored aiming behaviour in both a single-target and eight-target context. We also investigated the effects of adding an autocorrection module which learned to counteract the feedback rotation in a manner akin to the implicit adaptation observed in humans. In all cases, behaviour was consistent with aiming behaviour in VMR tasks. To quantitatively characterise this similarity, we directly compared behaviour in an eight-target, autocorrecting-cannon task against VMR aiming behaviour in Bond & Taylor (2015). We found that aiming was remarkably similar across both paradigms, for all rotation sizes. Thus, this approach may be a useful tool for investigating the strategic behaviour observed in VMR tasks.

Strategic behaviour without motor contamination

To avoid the contaminating factors introduced by the motor system, we investigated strategic behaviour in a cannon-shooting game removed from the motor context. In experiments 1a and 1b, with one target and eight targets, respectively, we established that behaviour was consistent with that in VMR literature. We replicated the gradual, monotonic improvement in task performance followed by asymptote (Figure 1b). One fundamental difference in behaviour, however, was the lack of any aftereffects during washout. This is expected, as aftereffects are produced by the persisting state of implicit adaptation, a part of the learning system which we have intentionally removed. To reconcile this difference, and to bring the aiming requirements of our task closer to those in a VMR task, we added an autocorrection module into the cannon. This autocorrection module covertly and gradually learned to

reduce the discrepancy between the aimed angle and the feedback angle in a similar manner to implicit adaptation acting on sensory prediction error (Figure 2a & b; see Methods for details). In experiment 2a, with one target location, we observed aiming behaviour akin to that in VMR tasks that ask the participant to report their aims: we found a quick, monotonic increase in aiming, followed by a slow decline in reaiming as the autocorrection module slowly learns to counteract more of the feedback rotation (Figure 2c). The degree of similarity, however, remains unclear. We are also left unsure as to whether or not the same underlying processes are generating the similar behaviour in both cases. We attempt to satisfy both concerns in the remainder of this discussion.

We attempt to characterise the degree of similarity between behaviour in our task and in traditional VMR tasks through a direct comparison to VMR data where performance is broken down into aim reports and implicit adaptation (Bond & Taylor, 2015). Aiming behaviour in experiment 2b (Figure 2d), which was designed to be directly comparable to Bond & Taylor (e.g., Figure 2e), did not differ from Bond & Taylor during the early, mid, or late rotation phases under any rotation size. This similarity in aiming behaviour, despite replacing endogenous implicit adaptation with exogenous autocorrection, is consistent with the idea that strategic behaviour will correct for errors left by the sensorimotor system (Taylor & Ivry, 2011). For a more intentional comparison, we defined three essential properties which characterise an individual's learning curve, but which are agnostic to the underlying algorithm which generated that behaviour. This set of properties included: the index of the cycle that had the greatest mean aim for each participant; the mean aim value during the peak cycle for a given participant; and the subsequent change in aim from that peak to the final cycle. These properties in our experiment were generally similar to Bond & Taylor (2015) across all rotation sizes (Figure 2f). The only difference was seen in the 90-degree rotation, where participants in the Cannon Game had a larger peak aim, and subsequent decrease in aim, than in Bond & Taylor. Overall, behaviour appears to be very similar across paradigms, prompting a deeper investigation into individual learning.

The individual timecourse of aiming and the "Aha!" moment

At the group level, aiming increases gradually and monotonically towards the ideal aim location. The individual, however, tends to perseverate in aiming at the target until suddenly shifting their aim, in a single trial, near to the ideal aim location (Figure 3d & e). These disparate patterns of behaviour are reconciled by the fact that taking the mean over variably-delayed step-functions yields a smooth curve (Bahrick et al., 1957; Gallistel et al., 2004). By comparing observed behaviour (Figure 3d & e) to gradual error-based learning and single-step learning "Aha!" simulations (Figure 3a & b, respectively), we see that human aiming generally exhibits the hallmarks of an "Aha!" moment much more so than of gradual learning

or trial-and-error learning. That is, instead of learning to gradually aim closer to the ideal aim angle as predicted by gradual learning (Figure 3a), and instead of exploratory behaviour predicted by trial-and-error learning (Figure 3b), we see that many participants perseverate in aiming at the target until some variable number of trials passes and then the participant abruptly shifts to aiming near the ideal angle. This aligns with recent work identifying qualitative aiming strategies (Tsay et al., 2024).

Aiming behaviour may instead be generated by an algorithm which learns through trial-and-error. To represent this hypothesis, we adapt the exploration-modulating model formulated by Cashaback et al. (2019). This model assumes that aiming on each trial is a normally distributed random variable, where the mean and variance depend on the outcome of the previous trial; after a failed trial, variance is increased to promote exploration, but after a successful trial, the mean is moved closer to the successful aim location and variance is decreased to increase exploitation (see Methods for algorithm). As can be seen in Figure 3c, this model can yield highly exploratory behaviour, centred on the target, until participants ‘randomly’ hit a successful trial. At which point, the mean is shifted closer to the aimed location, and the variance of aiming is determined purely by execution noise (which is set to zero for the simulations in Figure 3a, b & c). Much like the “Aha!” model, this produces a gradually increasing portion of participants that suddenly lock their aims to be near the target. Crucially, the preceding unbiased exploration around the target is only present with trial-and-error learning. Exploration—biased or otherwise—is clearly absent from the aiming behaviour of most participants (Figure 3d & e). Furthermore, given the size of the target and cannonball, it is not possible for unbiased exploration to learn fast enough on average to reliably reproduce the canonical group curve.

With visual inspection indicating only the “Aha!” model, we employ quantitative analyses on the model fits to explore further. For the model comparison, we fit the same three models used to generate the heatmap simulations—a three-parameter step-function model, a three-parameter SSM, and the three-parameter RL model—to the rotation phase of the first block of each participant (such that there could be no transfer of learning from a previous block). The insight model yielded a superior fit over both the gradual-learning model and the RL model for most participants under all rotation sizes (Figure 3f). Aiming behaviour of 82 of the 93 participants of experiment 1a are best described by the insight model, and the other 11 best described by the SSM. Participants in the fifteen and thirty-degree rotations generated data which was less distinctly fit by the “Aha!” model, relative to the SSM, than in the larger rotations (Figure 3f, right). This may be because the noise added to the cannonball flight is a larger proportion of the rotation, such that it is harder to estimate the feedback rotation and so more implicit and gradual learning processes are recruited. Indeed, this is consistent with

previous literature indicating that explicit reaiming contributes less to performance in reaching tasks with a smaller feedback rotation (Morehead et al., 2015; Neville & Cressman, 2018; Werner et al., 2015).

The “Aha!” moment in the motor context

It is possible that our task yields the same behavioural patterns as in standard VMR tasks, but through distinct strategic processes. Indeed, we know from recent literature that asking participants to explicitly report their aim does yield different aiming behaviour to that inferred from exclusion trials, and does alter the overall pattern of learning (Langsdorf et al., 2021; Maresch, Mudrik, et al., 2021; Maresch, Werner, et al., 2021). For example, our task prompts the participant to aim on every trial, whereas many VMR tasks do not. This prompting may act as a prior on solution space, such that it modifies the likelihood of adopting each considered strategy, or of using stimulus-response caching instead of a strategy.

To investigate this concern, we carried out the same model-fitting analysis on pre-existing reaching data from a task without aim reports, and without any prompt that an aiming solution might exist or that one should consider aiming (Brudner et al., 2016). Instead, the task presented feedback after a delay period of 5 seconds, which is known to inhibit implicit adaptation and emphasise strategic contributions (Schween & Hegele, 2017). Of the 16 participants who learned the task, 14 were best fit by the “Aha!” model (Figure 3g). This indicates that the VMR paradigm, in general, might elicit a perseveration in aiming followed by an “Aha!” moment with a sudden substantial change in behaviour. It is, however, possible that delaying feedback prompts the recruitment of strategic processing that would otherwise not be recruited when no such delay occurs. However, we see no reason, *a priori*, to think that strategic processing would be qualitatively—rather than quantitatively—affected by this delay.

By incorporating our insights into strategic behaviour back into the motor context, we can attack the problem of implicit behaviour in new ways. For example, we can utilise our finding that the cycle at which the participants exhibit their greatest mean aim value varies greatly, ranging from the second to the final rotation cycle (Figure 2f, rightmost plot). Combining this with the finding that implicit learning is generalised around the aimed location (McDougle et al., 2017) indicates that a lot of the heterogeneity observed in measures of implicit adaptation may instead reflect heterogeneity in strategic behaviour. For example, consider a participant in a VMR task who perseverates in aiming at the target for 20 cycles. This participant will accumulate implicit learning locally around that region of aiming space. When they experience an “aha!” moment and switch to the roughly ideal region of aiming space, a

much smaller degree of implicit learning will have been accumulated for the new region, so their implicit learning curve will appear partially “reset”. This pattern of perseveration and then a delayed shift to good performance also has implications for many other forms of learning in the motor domain. Indeed, performance in tasks which were long thought to evoke purely implicit learning (e.g., use-dependent motor learning or motor sequence learning) may largely be a function of strategic processes (Tsay et al., 2024).

Understanding the “Aha!” moment

What is happening, computationally, underneath this apparent “Aha!” moment? This behavioural pattern may reflect the reorganisation of the participant’s internal representation of the problem, and this reorganisation may then yield a richer understanding of the problem and solution space (Knoblich et al., 1999, 2001; Sprugnoli et al., 2017). It could be that the task is initially represented as a simple problem of aiming at the target, which is a parsimonious and efficacious representation during baseline. Upon enforcing the feedback rotation, this representation becomes impotent in generating efficacious solutions. Constructing a more sophisticated representation that encodes a relationship between the aim and the cannonball direction could bring the representation closer to an aiming solution for nonveridical feedback. Importantly, the participant does not know, *a priori*, that the feedback has been rotated; they just see that the cannonball follows a different path than expected. At this level of representation—that the cannonball isn’t behaving how it used to, such that previously successful actions are no longer successful—the solution of rotating one’s aim to compensate for the new error may not be immediately obvious. But this solution may become much closer if the representation of the problem is imbued with more semantic and relational information (Trope & Liberman, 2010). Framing the behaviour of the cannonball in terms relative to the aimed location, and semantically understanding the contingency therein, brings the problem closer to the idea of changing the aimed location to compensate for this altered action-feedback mapping.

Alternatively, the ”Aha!” moment may reflect the accumulation of evidence for the presence of a new context having breached some evidence threshold. In this view, upon inferring the requirement of a different solution, most participants are immediately able to implement an efficacious solution and adjust to good performance in a single trial. This abrupt shift to good performance appears counter to findings that humans tend to explore solution space upon inferring that a different solution is required (Cohen et al., 2007; Collins & Koechlin, 2012). Nevertheless, our findings indicate that an efficacious solution can be produced without any additional observations being generated, such that a novel solution must have already been covertly formulated (Bowden & Jung-Beeman, 2003; Carpenter, 2020; Smith & Kounios, 1996), was derived through simulation (Allen et al., 2020), memory replay (Foster, 2017), or

a generalisable solution is retrieved from long-term memory. Our ‘apparent’ but delayed one-shot learning may be facilitated by a simpler convex objective function than is typical in those tasks used in more cognitive domains, but this simplicity is likely shared across related motor tasks (albeit with greater stochasticity and uncertainty in some cases).

In any case, the presence of an “Aha!” moment and the initial process of solution generation still leaves space for subsequent learning processes to refine the new solution or to search for alternative solutions.

Conclusion

Together, these findings indicate that strategic behaviour, both in the Cannon Game and in VMR reaching tasks, is not a gradual error-minimisation process or a process of learning through trial-and-error. Instead, there is perseveration of baseline behaviour, followed by an “Aha!” moment which marks a dramatic single-trial shift in behaviour. This abrupt, discontinuous nature of strategic behaviour needs to be considered when modelling behaviour in VMR tasks. This is important not only for understanding how strategies are acquired and evolve over time, but also in our attempt to understand implicit adaptation and motor behaviour more generally. By employing this class of model to capture behavioural variance generated by the strategic system, we can more accurately explain the variance generated by non-strategic processes.

Future work should try to better understand the process that leads to this sudden shift in behaviour. Is there a fundamental change in the representation of the task, or does the preceding perseveration reflect the time taken to infer the presence of a new context? Answering this would also expedite the motor rehabilitation process by helping patients to learn strategic behaviours more quickly and effectively. If plan-based generalisation of implicit adaptation occurs in more naturalistic contexts, this may also enhance adaptation of the sensorimotor system during rehabilitation: we can generate the desired strategic solution for a patient early to avoid the partial “reset” in adaptation that occurs upon the delayed adoption of a new strategy.

Acknowledgements

We thank Jordan Taylor for providing the raw data used to compare Gannon Game with VMR data (Bond & Taylor, 2015; Brudner et al., 2016).

Conflict of interest

The authors declare no conflicts of interest.

Financial support

M.T. was supported by an ESRC White Rose Doctoral Training Partnership Scholarship. F.M. was supported by the UK Research and Innovation Biotechnology and Biological Sciences Research Council (BB/X008428/1) and the National Institute for Health and Care Research (NIHR) Leeds Biomedical Research Centre (NIHR203331).

Data availability

Data and code will be made available upon acceptance on the Open Science Framework.

Ethics statement

Ethical approval was obtained from the University of Leeds school of Psychology Research Ethics Committee (ref PSYC-81). All participants gave informed consent via a Qualtrics web form prior to starting the experiments.

Author contributions

MT: Conceptualization, Methodology, Software, Formal analysis, Investigation, Visualization, Writing - Original Draft, Writing - Review & Editing. **MW:** Validation, Writing - Review & Editing. **CC:** Supervision, Writing - Review & Editing. **MMW:** Supervision, Writing - Review & Editing, Funding acquisition. **FM:** Supervision, Writing - Review & Editing, Funding acquisition. **JRM:** Conceptualization, Methodology, Supervision, Writing - Review & Editing, Funding acquisition.

References

Allen, K. R., Smith, K. A., & Tenenbaum, J. B. (2020). Rapid trial-and-error learning with simulation supports flexible tool use and physical reasoning. *Proceedings of the National Academy of Sciences*, *117*(47), 29302–29310. https://doi.org/10.1073/pnas.1912341117

Bahrick, H. P., Fitts, P. M., & Briggs, G. E. (1957). Learning curves—Facts or artifacts? *Psychological Bulletin*, *54*(3), 256–268. https://doi.org/10.1037/h0040313

Bond, K. M., & Taylor, J. A. (2015). Flexible explicit but rigid implicit learning in a visuomotor adaptation task. *Journal of Neurophysiology*, *113*(10), 3836–3849. https://doi.org/10.1152/jn.00009.2015

Bowden, E. M., & Jung-Beeman, M. (2003). Aha! Insight experience correlates with solution activation in the right hemisphere. *Psychonomic Bulletin & Review*, *10*(3), 730–737. https://doi.org/10.3758/BF03196539

Brudner, S. N., Kethidi, N., Graeupner, D., Ivry, R. B., & Taylor, J. A. (2016). Delayed feedback during sensorimotor learning selectively disrupts adaptation but not strategy use. *Journal of Neurophysiology*, *115*(3), 1499–1511. https://doi.org/10.1152/jn.00066.2015

Carpenter, W. (2020). The Aha! Moment: The Science Behind Creative Insights. In S. Manuel Brito (Ed.), *Toward Super-Creativity—Improving Creativity in Humans, Machines, and Human—Machine Collaborations*. IntechOpen. https://doi.org/10.5772/intechopen.84973

Cashaback, J. G. A., Lao, C. K., Palidis, D. J., Coltman, S. K., McGregor, H. R., & Gribble, P. L. (2019). The gradient of the reinforcement landscape influences sensorimotor learning. *PLOS Computational Biology*, *15*(3), e1006839. https://doi.org/10.1371/journal.pcbi.1006839

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 933–942. https://doi.org/10.1098/rstb.2007.2098

Collins, A., & Koechlin, E. (2012). Reasoning, Learning, and Creativity: Frontal Lobe Function and Human Decision-Making. *PLoS Biology*, *10*(3), e1001293. https://doi.org/10.1371/journal.pbio.1001293

Coltman, S. K., Cashaback, J. G. A., & Gribble, P. L. (2019). Both fast and slow learning processes contribute to savings following sensorimotor adaptation. *Journal of Neurophysiology*, *121*(4), 1575–1583. https://doi.org/10.1152/jn.00794.2018

Darshan, R., Leblois, A., & Hansel, D. (2014). Interference and Shaping in Sensorimotor Adaptations with Rewards. *PLoS Computational Biology*, *10*(1), e1003377. https://doi.org/10.1371/journal.pcbi.1003377

Donoso, M., Collins, A. G. E., & Koechlin, E. (2014). Foundations of human reasoning in the prefrontal cortex. *Science*, *344*(6191), 1481–1486. https://doi.org/10.1126/science.1252254

Foster, D. J. (2017). Replay Comes of Age. *Annual Review of Neuroscience*, *40*(1), 581–602. https://doi.org/10.1146/annurev-neuro-072116-031538

Gallistel, C. R., Fairhurst, S., & Balsam, P. (2004). The learning curve: Implications of a quantitative analysis. *Proceedings of the National Academy of Sciences*, *101*(36), 13124–13131. https://doi.org/10.1073/pnas.0404965101

Griffiths, T. L., Callaway, F., Chang, M. B., Grant, E., Krueger, P. M., & Lieder, F. (2019). Doing more with less: Meta-reasoning and meta-learning in humans and machines. *Current Opinion in Behavioral Sciences*, *29*, 24–30. https://doi.org/10.1016/j.cobeha.2019.01.005

Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in Cognitive Sciences*, *14*(8), 357–364. https://doi.org/10.1016/j.tics.2010.05.004

Hegele, M., & Heuer, H. (2010). Implicit and explicit components of dual adaptation to visuomotor rotations. *Consciousness and Cognition*, *19*(4), 906–917. https://doi.org/10.1016/j.concog.2010.05.005

Jones, G. (2003). Testing two cognitive theories of insight. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(5), 1017–1027. https://doi.org/10.1037/0278-7393.29.5.1017

Jung-Beeman, M., Bowden, E. M., Haberman, J., Frymiare, J. L., Arambel-Liu, S., Greenblatt, R., Reber, P. J., & Kounios, J. (2004). Neural Activity When People Solve Verbal Problems with Insight. *PLoS Biology*, *2*(4), e97. https://doi.org/10.1371/journal.pbio.0020097

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint Relaxation and Chunk Decomposition in Insight Problem Solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(6), 1534–1555. https://doi.org/10.1037/0278-7393.25.6.1534

Knoblich, G., Ohlsson, S., & Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory & Cognition*, *29*(7), 1000–1009. https://doi.org/10.3758/BF03195762

Kounios, J., & Beeman, M. (2009). The *Aha!* Moment: The Cognitive Neuroscience of Insight. *Current Directions in Psychological Science*, *18*(4), 210–216. https://doi.org/10.1111/j.1467-8721.2009.01638.x

Krakauer, J. W., Pine, Z. M., Ghilardi, M.-F., & Ghez, C. (2000). Learning of Visuomotor Transformations for Vectorial Planning of Reaching Trajectories. *The Journal of Neuroscience*, *20*(23), 8916–8924. https://doi.org/10.1523/JNEUROSCI.20-23-08916.2000

Langsdorf, L., Maresch, J., Hegele, M., McDougle, S. D., & Schween, R. (2021). Prolonged response time helps eliminate residual errors in visuomotor adaptation. *Psychonomic Bulletin & Review*, *28*(3), 834–844. https://doi.org/10.3758/s13423-020-01865-x

MacGregor, J. N., Ormerod, T. C., & Chronicle, E. P. (2001). Information processing and insight: A process model of performance on the nine-dot and related problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 176–201. https://doi.org/10.1037/0278-7393.27.1.176

Maresch, J., Mudrik, L., & Donchin, O. (2021). Measures of explicit and implicit in motor learning: What we know and what we don't. *Neuroscience & Biobehavioral Reviews*, *128*, 558–568. https://doi.org/10.1016/j.neubiorev.2021.06.037

Maresch, J., Werner, S., & Donchin, O. (2021). Methods matter: Your measures of explicit and implicit processes in visuomotor adaptation affect your results. *European Journal of Neuroscience*, *53*(2), 504–518. https://doi.org/10.1111/ejn.14945

Mattar, A. A. G., & Ostry, D. J. (2007). Modifiability of Generalization in Dynamics Learning. *Journal of Neurophysiology*, *98*(6), 3321–3329. https://doi.org/10.1152/jn.00576.2007

McDougle, S. D., Bond, K. M., & Taylor, J. A. (2015). Explicit and Implicit Processes Constitute the Fast and Slow Processes of Sensorimotor Learning. *Journal of Neuroscience*, *35*(26), 9568–9579. https://doi.org/10.1523/JNEUROSCI.5061-14.2015

McDougle, S. D., Bond, K. M., & Taylor, J. A. (2017). Implications of plan-based generalization in sensorimotor adaptation. *Journal of Neurophysiology*, *118*(1), 383–393. https://doi.org/10.1152/jn.00974.2016

McDougle, S. D., Ivry, R. B., & Taylor, J. A. (2016). Taking Aim at the Cognitive Side of Learning in Sensorimotor Adaptation Tasks. *Trends in Cognitive Sciences*, *20*(7), 535–544. https://doi.org/10.1016/j.tics.2016.05.002

Morehead, J. R., Butcher, P. A., & Taylor, J. A. (2011). Does Fast Learning Depend on Declarative Mechanisms? *The Journal of Neuroscience*, *31*(14), 5184–5185. https://doi.org/10.1523/JNEUROSCI.0040-11.2011

Morehead, J. R., Qasim, S. E., Crossley, M. J., & Ivry, R. (2015). Savings upon Re-Aiming in Visuomotor Adaptation. *Journal of Neuroscience*, *35*(42), 14386–14396. https://doi.org/10.1523/JNEUROSCI.1046-15.2015

Morehead, J. R., Taylor, J. A., Parvin, D. E., & Ivry, R. B. (2017). Characteristics of Implicit Sensorimotor Adaptation Revealed by Task-irrelevant Clamped Feedback. *Journal of Cognitive Neuroscience*, *29*(6), 1061–1074. https://doi.org/10.1162/jocn_a_01108

Neville, K.-M., & Cressman, E. K. (2018). The influence of awareness on explicit and implicit contributions to visuomotor adaptation over time. *Experimental Brain Research*, *236*(7), 2047–2059. https://doi.org/10.1007/s00221-018-5282-7

Novick, L. R., & Sherman, S. J. (2003). On the Nature of Insight Solutions: Evidence from Skill Differences in Anagram Solution. *The Quarterly Journal of Experimental Psychology Section A*, *56*(2), 351–382. https://doi.org/10.1080/02724980244000288

Piantadosi, S. T., Tenenbaum, J. B., & Goodman, N. D. (2016). The logical primitives of thought: Empirical foundations for compositional cognitive models. *Psychological Review*, *123*(4), 392–424. https://doi.org/10.1037/a0039980

Salvi, C., Bricolo, E., Kounios, J., Bowden, E., & Beeman, M. (2016). Insight solutions are correct more often than analytic solutions. *Thinking & Reasoning*, *22*(4), 443–460. https://doi.org/10.1080/13546783.2016.1141798

Sawyer, R. K. (2012). *Explaining creativity: The science of human innovation, 2nd ed.* (pp. xi, 555). Oxford University Press.

Schween, R., & Hegele, M. (2017). Feedback delay attenuates implicit but facilitates explicit adjustments to a visuomotor rotation. *Neurobiology of Learning and Memory*, *140*, 124–133. https://doi.org/10.1016/j.nlm.2017.02.015

Smith, R. W., & Kounios, J. (1996). Sudden insight: All-or-none processing revealed by speed–accuracy decomposition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 1443. https://doi.org/10.1037/0278-7393.22.6.1443

Sprugnoli, G., Rossi, S., Emmendorfer, A., Rossi, A., Liew, S.-L., Tatti, E., di Lorenzo, G., Pascual-Leone, A., & Santarnecchi, E. (2017). Neural correlates of Eureka moment. *Intelligence*, *62*, 99–118. https://doi.org/10.1016/j.intell.2017.03.004

Taylor, J. A., & Ivry, R. B. (2011). Flexible Cognitive Strategies during Motor Learning. *PLoS Computational Biology*, *7*(3), e1001096. https://doi.org/10.1371/journal.pcbi.1001096

Taylor, J. A., & Ivry, R. B. (2012). The role of strategies in motor learning: The role of strategies in motor learning. *Annals of the New York Academy of Sciences*, *1251*(1), 1–12. https://doi.org/10.1111/j.1749-6632.2011.06430.x

Taylor, J. A., & Ivry, R. B. (2013). Context-dependent generalization. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00171

Taylor, J. A., Klemfuss, N. M., & Ivry, R. B. (2010). An Explicit Strategy Prevails When the Cerebellum Fails to Compute Movement Errors. *The Cerebellum*, *9*(4), 580–586. https://doi.org/10.1007/s12311-010-0201-x

Taylor, J. A., Krakauer, J. W., & Ivry, R. B. (2014). Explicit and Implicit Contributions to Learning in a Sensorimotor Adaptation Task. *Journal of Neuroscience*, *34*(8), 3023–3032. https://doi.org/10.1523/JNEUROSCI.3619-13.2014

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*(2), 440–463. https://doi.org/10.1037/a0018963

Tsay, J. S., Kim, H. E., McDougle, S. D., Taylor, J. A., Haith, A., Avraham, G., Krakauer, J. W., Collins, A. G., & Ivry, R. B. (2024). Fundamental processes in sensorimotor learning: Reasoning, refinement, and retrieval. *eLife*, *13*, e91839. https://doi.org/10.7554/eLife.91839

Vartanian, O., & Goel, V. (2007). Neural correlates of creative cognition. In C. Martindale, P. Locher, V. M. Petrov, & A. Berleant (Eds.), *Evolutionary and neurocognitive approaches to aesthetics, creativity and the arts* (pp. 195–207). Routledge.

Werner, S., Van Aken, B. C., Hulst, T., Frens, M. A., Van Der Geest, J. N., Strüder, H. K., & Donchin, O. (2015). Awareness of Sensorimotor Adaptation to Visual Rotations of Different Size. *PLOS ONE*, *10*(4), e0123321. https://doi.org/10.1371/journal.pone.0123321