

Title: Individual differences in implicit motor learning: Task specificity in sensorimotor adaptation and sequence learning

Running title: Individual differences in implicit motor learning

Authors: Stark-Inbar Alit¹, Raza Meher¹, Taylor A. Jordan², Ivry B. Richard^{1,2}

¹Department of Psychology and ²Helen Wills Neuroscience Institute, University of California, Berkeley, CA 94720, USA, ³Department of Psychology, Princeton University, Princeton, New Jersey 08544.

Corresponding author: Stark-Inbar Alit, alitstark@gmail.com

Keywords: individual differences, implicit learning, adaptation, sequence learning, reliability

Abstract:

In standard taxonomies, motor skills are typically treated as representative of implicit or procedural memory. We examined two emblematic tasks of implicit motor learning, sensorimotor adaptation and sequence learning, asking whether individual differences in learning are correlated between these tasks, as well as how individual differences within each task are related to different performance variables. As a prerequisite, it was essential to establish the reliability of learning measures for each task. Participants were tested twice on a visuomotor adaptation task and on one of two sequence learning tasks, the serial reaction time task or the alternating reaction time task. Learning was evident in all tasks at the group level, and reliable at the individual level in visuomotor adaptation and the alternating reaction time task, but not in the serial reaction time task. Performance variability was predictive of learning in both domains, yet the relationship was in the opposite direction for adaptation and sequence learning. For the former, faster learning was associated with lower variability, consistent with models of sensorimotor adaptation in which learning rates are sensitive to noise. For the latter, greater learning was associated with higher variability and slower reaction times, factors that may facilitate the spread of activation required to form predictive, sequential associations. Interestingly, learning measures on the different tasks were not correlated. Together, these results argue against a shared process for implicit learning in sensorimotor adaptation and sequence learning, and provide insight into factors that account for individual differences in learning within each task domain.

New & Noteworthy:

We investigated individual differences in the ability to implicitly learn motor skills. As a prerequisite, we assessed whether individual differences were reliable across test sessions. The results show that two commonly used tasks of implicit learning, visuomotor adaptation and the alternating serial reaction time task, exhibited good test-retest reliability in measures of learning and performance. However, the learning measures did not correlate between the two tasks, arguing against a shared process for implicit motor learning.

Introduction:

Studies of motor learning and performance tend to focus on group level measures, characterizing the behavior of the “average” person (Kanai and Rees 2011; King et al. 2012). Inter-individual differences are typically ignored, treated as one source of “noise” that can be averaged out across the group or factored out by random group assignment. However, everyday life experience makes evident that people vary considerably in their abilities to learn new skills or modify the way in which they perform a learned skill. While the study of individual differences in skilled performance has, historically, been of considerable interest in the field of human factors (Seashore and Tiffin 1930; Keele and Hawkins 1982), recent work in motor learning has tended to ignore this topic. Determining the underlying sources of individual differences and the extent to which they generalize across tasks is critical to developing a comprehensive account of motor learning.

In this paper, we take an individual differences approach to examine implicit motor learning. We focus on two task domains that have been hypothesized to be emblematic of procedural motor learning, sensorimotor adaptation and sequence learning. In adaptation studies, the environment is perturbed in some novel way and the participant learns to adjust their behavior to minimize the consequences of that perturbation. For example, in visuomotor rotation tasks, an angular displacement is introduced between movements of the hand and a cursor, and successful performance requires moving the hand in the opposite direction to counter the perturbation. Sequence learning has been studied with a range of tasks in which a series of stimulus-response pairs form a sequence that is either deterministic (e.g., Nissen and Bullemer, 1987) or probabilistic (e.g., Howard and Howard, 1997), with learning defined as the difference in RT between predictable and non-predictable events.

As suggested by their taxonomic linkage with procedural learning, adaptation and sequence learning can occur implicitly. For example, the perturbation in adaptation studies might be small, and/or introduced in a gradual manner, conditions in which the participants are unaware of the perturbation (Kagerer et al. 1997; Michel et al. 2007; Criscimagna-Hemminger et al. 2010; Taylor and Ivry 2012; Schlerf et al. 2013). Similarly, various manipulations have been devised to reduce or eliminate awareness during sequence learning, including the use of long (Pascual-Leone et al. 1993; Sanchez and Reber 2012) or probabilistic sequences (Howard and Howard 1997; Nemeth et al. 2010), or by using a dual-task procedure (Hazeltine et al., 1997;

Gheysen et al., 2009; Sanchez et al., 2010). While the magnitude of learning is reduced under such conditions (Rüsseler and Rösler 2000; Spencer et al. 2006), robust learning is still observed (Reber and Squire 1994).

Here, we examine three issues regarding individual differences across these two implicit motor learning domains. First, how do individual differences in motor learning generalize across tasks? The earliest studies of individual differences tended to focus on stable, or trait-based variability, asking, for example, whether the abilities of an individual across a wide range of tasks were related to single construct such as intelligence (e.g. Spearman's "g factor"; Spearman, 1904). Motivated by models of cognition, studies of human performance have identified individual differences in the operation of core components that generalize across task domains (Henry, 1958). For example, Keele et al. (Keele et al., 1985) showed that individual differences in temporal precision account for performance differences across motor and perceptual timing tasks. In a related vein, sensorimotor adaptation and sequence learning might entail overlapping processes since learning in both tasks entail predictions of spatial events (Keele et al. 1985; Spencer and Ivry 2009; Sanchez et al. 2010; Seidler et al. 2012), and both have been associated with cerebellar function (Seidler 2006; Galea et al. 2011; Ferrucci et al. 2013). On the other hand, there are good reasons to expect that learning in these two task domains might entail very different processes. Computationally, adaptation is assumed to depend on mechanisms designed to reduce sensory prediction errors (Shadmehr et al. 2010), whereas sequence learning is hypothesized to depend on associative processes (Kinder et al. 2008; Spencer and Ivry 2009). Even in terms of the cerebellum, adaptation and sequence learning may be associated with different cortico-cerebellar loops (Kelly and Strick 2003; Krienen and Buckner 2009) and cerebellar regions (Rabe et al. 2009; Burciu et al. 2014) (Hazeltine et al. 1997; Ramnani et al. 2000; Seidler et al. 2002; Bernard and Seidler 2013).

Second, we ask whether individual differences in implicit motor learning are stable. At the heart of studies of individual differences is an interest in understanding sources of variability. Variability is observed at all levels, be it the expression of proteins, the firing of neurons, or movements of the body (Faisal et al., 2008). There is also measurement noise, reflecting the fact that our measurement tools are imprecise. An important distinction can be made between stable and unstable sources of variability in performance. The latter emerges from the specific conditions at the time of measurement (e.g. random noise, fatigue, attentiveness), factors that

tend to average out with repetitions. The former, on the other hand, refers to stable characteristics of the system, and measurements of these sources of variability should be relatively consistent over test sessions. The reliability of our measurements (i.e. consistency over repeated tests) limits the ability to infer relationships between variables, as the size of the correlation between two variables is limited by the reliability of either (or both) of the variables (Gullicksen, 2013).

Surprisingly, we are not aware of any studies that have examined the reliability of measures of implicit motor learning. Studies that tested people over multiple days have focused on consolidation (off-line learning), using the same task and stimulus parameter to look at retention of learning from one session to the next (Spencer et al. 2006; Zarahn et al. 2008; Doyon et al. 2009; Krakauer 2009; Nemeth et al. 2010; Meier and Cock 2014). Not only do these studies rarely discuss the reliability of individual differences across days, but also the consolidation design limits the ability to draw inferences about reliability (i.e. individual differences in off-line consolidation will introduce variation in the estimates of learning across sessions). Given this concern, we designed our study such that each participant was tested twice on each task, employing variants that should minimize any contribution (savings or interference) from the initial test session on the second test session. We employed conditions in which awareness was absent or minimized, ensuring that learning was entirely implicit. In this manner, we can ask if individual differences on each task are stable or unstable across repeated measurements, a prerequisite for examining factors that account for individual differences in performance.

Third, we used an individual difference approach to examine how variability contributes to, and constrains, skill acquisition. We often think of variability as an unwanted feature that interferes with learning and performance (Faisal et al. 2008; Wu et al. 2014). This idea is formally captured in models such as that offered by a Kalman filter within the framework of Bayesian estimation. For example, as sensory noise increases (e.g., when the position of a cursor, indicating hand position is artificially blurred), the rate of learning rate decreases (Burge et al. 2008; Wei and Körding 2010). Theories of optimal control posit that actions are planned to minimize the impact of variability on performance (Körding and Wolpert 2004; Körding et al. 2007). In this framework, part of skill acquisition entails a reduction in variability. Alternatively, movement variability can also be seen as a signature of exploration, allowing the system to gather knowledge concerning new or unstable environments (Sutton and Barto 1998; Ishii et al., 2002). In this view, an increase in variability might come about from a propensity to explore. It

is also possible that the causation is reversed, such that high variability increases the likelihood that the system will explore new associations. We sought to identify factors that might account for individual differences in learning on these two tasks, focusing on measures of performance variability that were obtained outside the learning context or early during learning, asking how differences in variability might impede or enhance learning (Wu et al., 2014).

Methods:

Participants

A total of 123 neurologically healthy young adults were recruited through the Research Participant Pool at the University of California, Berkeley (63% females; age 21.2 ± 2.4). All were right-handed as assessed by the Edinburgh Handedness Inventory (0.83 ± 0.19 , where -1 is strictly left-handed and 1 is strictly right-handed). Participants received course credit for completing the study. The protocol was approved by the UC Berkeley institutional review board.

Experimental tasks

There were three experimental tasks, one used as a measure of sensorimotor adaptation and the other two to measure sequence learning. Participants were tested twice on each task, in two runs. Runs were either completed on separate sessions (at least 2 days apart), or during the same session (SRT task only; following a 10-15 min break).

I. Visuomotor adaptation (VMA) task:

Participants made center-out reaching movements to visually displayed targets, sliding a digitizing stylus with their right hand across a digitizing tablet (33 cm x 20cm; Wacom Technology Corporation, WA, USA). Movement trajectories were recorded by sampling the stylus tip at 100 Hz with customized software, written in Python (www.python.org). The visual stimuli were displayed on a 15-in. monitor, oriented in the normal upright position. The tablet was oriented horizontally and positioned below a wooden box that occluded vision of the arm. To provide feedback, a small white cursor (4 mm diameter) was displayed on the screen to indicate the position of the stylus. The system was calibrated such that there was a 1:1

correspondence between movement distance of the stylus and the displacement distance of the feedback cursor on the screen.

To begin each trial, a circle (5 mm diameter) appeared at the center of the monitor, indicating the start position. Participants moved the stylus to center the feedback cursor within the start position. After holding this position for 500 ms, four circles appeared. The radial distance from the start circle to each of the four circles was 8 cm. The circles were separated by 30°, forming a virtual arc on the upper part of the monitor (locations 45°, 75°, 105°, and 135°, where 90° is directly above the starting position). Three of the circles were colored blue (5 mm diameter) and one, the target, was colored white and was slightly larger (8 mm).

Participants were instructed to reach to the target “as fast and as accurate” as possible. The start circle disappeared when the movement amplitude exceeded 1 cm. Online visual feedback, by means of the feedback cursor, was presented while the participant reached outwards toward the target. The trial ended when the participant intersected the target. At that moment, the target turned green and auditory feedback was provided based on movement speed. If the movement time was less than 500 ms, a pleasant high-pitch “ding” sound was played. If the movement time exceeded this criterion, an unpleasant low-pitch “buzz” sound was played. Note, reaction time was not emphasized; the auditory feedback was based solely on movement time. The cursor endpoint position, defined at the point where the reach amplitude crossed 8 cm, was displayed for 500 ms.

At the end of the feedback interval, the cursor and four circles disappeared and the start circle reappeared. The participant was visually guided back to the starting location using radial feedback, a white circle with a diameter that corresponded to the distance of the hand from the starting position (i.e., providing no information on hand angle). The participant was instructed to move so as to reduce the diameter of this circle. When the hand was within 1 cm of the start position, the circle changed to a cursor, and the participant moved the cursor into the start location. When this position had been maintained for 500 ms, the next target appeared.

Participants were briefly introduced to the task with a short training block of 24 trials, six reaches to each of the four possible target locations. The experimental session was composed of 640 trials, divided into six blocks (**Fig. 1a**). The first block (Baseline1) consisted of 40 trials in which the mapping between hand position and cursor position was veridical. In the second block (240 trials, Rotate1) a pseudorandom walk visuomotor perturbation was imposed by rotating the

position of the feedback cursor with respect to the true position of the hand (see below). Following the end of the rotation block, the participant completed a third block consisting of 40 trials with veridical feedback (Wash1). The participant was then given a short break, before repeating each of the three blocks a second time (Baseline2, Rotate2, Wash2). The task transitioned smoothly between blocks without indicating to the participants the change between blocks. The order of target locations was pseudorandom, such that each target location was used twice every eight trials, avoiding successive repetitions of each location. The 640-trial run took approximately 35 minutes to complete.

During the rotation blocks, the perturbation of the visual feedback followed a global sinusoidal pattern with local jitter (modified from Cheng and Sabes, 2006), increasing or decreasing by 1-2° on each trial. The maximum rotation was 12° in either the clockwise or counterclockwise direction, with the cycle spanning 80 trials (**Fig. 1a**). In terms of actual distance on the screen, a displacement of 1° corresponded to 0.14 cm. We chose to use small trial-to-trial changes in the size of the rotation along with a maximum rotation size of only 12° (corresponding to 1.67 cm displacement) to minimize awareness of the perturbation (Cheng and Sabes 2007; Schlerf et al. 2013).

Participants completed the VMA task twice, with each run performed on different days. Two different versions of the task were created by varying the local jitter of the perturbation and the order of the target locations. A different version was used for each of the two runs.

II. Serial Reaction Time (SRT) task:

We employed a version of the serial reaction time (SRT) task that is designed to minimize the contribution of explicit learning (e.g. Robertson, 2007; Spencer and Ivry, 2009). Four horizontal lines (3.5 cm, with 1.5 cm edge-to-edge spacing) were continuously displayed on the monitor throughout the experiment. At the start of each trial, an "X" symbol (1.8 x 2 cm) appeared above one of the lines. Participants were instructed to press the spatially compatible key as fast as possible. Responses were made with one of four fingers of the right hand, with the fingers positioned on the V, B, N, and M keys on a standard computer keyboard. The target stimulus disappeared 100 ms after the participant pressed a key (either the correct key or a wrong one) or, after an interval of 2000 ms if no key press was detected. Once the target disappeared, a 100 ms inter-trial interval was initiated, followed by the presentation of the next stimulus.

The experimental run consisted of 1260 trials, organized into 15 blocks of 84 trials each. There were two types of blocks, Random or Sequence. On Random blocks, the stimuli were selected randomly with the constraints that the stimulus did not appear at the same location on successive trials (e.g., 2-2 or 4-4), did not form 3-element trills (e.g., 1-3-1 or 2-4-2) or 3-element runs (e.g. 1-2-3 or 4-3-2), and that each stimulus appeared 3 times every 12 trials. On Sequence blocks, the stimuli followed a repeating 12-element deterministic sequence (repeating 7 times/block). Two sequences were created (SeqA: 241342314213; SeqB: 312413421324). The second sequence was created by shifting each element of the first by one position, with position 4 becoming position 1. Based on past work, we expected that a 12-element sequence would minimize awareness (Pascual-Leone et al. 1993; Spencer and Ivry 2009). We opted to not use a dual-task procedure as a way to minimize awareness because participants vary in how they allocate attention between the sequence learning and secondary tasks, making it difficult to associate individual differences to sequence learning per se. Blocks 1, 7, 13, and 15 were Random blocks, and the other 11 blocks were Sequence blocks (**Fig. 1b**). Overall, it took approximately 15 minutes to complete the 1260 trials, including a short break provided after Block 8. Participants completed two 1260-trial runs, with different sequences (SeqA or SeqB) used in each run.

III. Alternating Serial Reaction Time (ASRT) task:

As described in the Results section below, the SRT task proved to be problematic for an individual differences study given that within-subject measures of learning showed low reliability. As such, we switched to a different assay of implicit sequence learning, the alternating serial reaction time (ASRT) task (Howard and Howard 1997; Remillard 2008; Janacsek et al. 2012). Unlike the deterministic SRT task, the ASRT task involves a probabilistic second-order grammar in which odd numbered elements follow a fixed sequence and even numbered elements are selected at random.

The ASRT task used similar stimulus-response mapping as the SRT task (Howard and Howard 1997; Nemeth et al. 2010; Janacsek et al. 2012). Four empty circles (1.5 cm diameter, with 2.5 cm edge-to-edge spacing) were continuously displayed on the computer monitor throughout the experiment. The stimulus on each trial was a dog's head that appeared over one of the circles (2 cm diameter; Nemeth et al., 2010). Participants had to press the spatially

compatible response key as fast as possible. Responses were made with the four fingers of the right hand, with each finger mapped to one key on a keyboard (keys V, B, N, and M). The trial ended after the correct key was pressed; thus, if the wrong key was pressed, the stimulus remained presented until its associated key was pressed. The target stimulus disappeared immediately after the correct key was pressed, replaced by an empty circle. This initiated a 120 ms inter-trial interval before the onset of the next stimulus.

The stimuli followed an 8-element sequence of the structure, 1r2r3r4r, where the four odd numbered elements (1, 2, 3 and 4) follow a fixed sequence (1-2-3-4 in this example), and the even numbered elements (the r's) were selected at random (**Fig. 1c**). For each participant, the four fixed elements were randomly mapped to the four stimulus locations, and the only constraint on the random elements was that they were chosen with an equal probability (allowing successive elements, trills and runs). The experimental session consisted of 45 blocks of 85 trials each (3825 trials in total). The first five stimuli in each block were selected at random and are not included in the analysis. The 8-element sequence was then repeated 10 times for each block. At the end of each block, the participant received feedback on the screen, indicating mean reaction time and overall accuracy. The task took approximately 60 minutes to complete.

Procedure

We present the results in terms of four groups, based on the tasks that the participants completed. Group 1 (n=38) was tested on the VMA task twice, with a break of 2-5 days between the sessions. These participants were also tested on a reaching version of the SRT task (Spencer and Ivry 2009). However, we did not observe consistent sequence learning with this task, either at the group or individual level, and, thus, do not report these data here. This led us to switch to the keyboard variants of the SRT tasks (as described above). Group 2 (n=23) was only tested on the keyboard SRT task, with the task repeated twice within a single session. Groups 3 and 4 completed both a sensorimotor adaptation task and a sequence-learning task. Group 3 (n=37) was tested on the VMA and SRT tasks, with 17 starting with the VMA task and 20 starting with the SRT task. The two tasks were repeated in a separate session with the same order used for each participant. Group 4 (n=25) was tested on both the VMA and ASRT tasks, with all participants first tested on the VMA task. This two-hour session (given the length of time required for each task) was repeated a second time, again with a 2-5 day break between sessions.

For all groups, the first and second runs with a given task used a different stimulus pattern. For the VMA task, this meant that the order of targets and perturbation function differed in the two runs; for the SRT task, different sequences were used in the two runs. The order of the two versions of each task was counterbalanced across participants and days. As noted above, the same grammar was used for all participants on the ASRT task, although the mapping of the grammar elements to stimuli/responses was randomly set for each participant, with the constraint that different mappings were used in the two runs.

At the end of the second run, a questionnaire was administered to determine participants' awareness of the task manipulations, with separate sets of items used to probe awareness of the VMA and sequencing tasks. The first question asked if the participant had noticed any changes over the course of the specified task. For the second question, participants were told there were two groups and they should report, using a 7-point Likert scale, which group they thought they belonged to. For the VMA task, the choices were between a group for whom the feedback matched the position of the hand and a group for whom the feedback and hand did not match. For the sequencing tasks, the choices were between a group for whom the stimuli were selected at random and a group for whom the stimuli followed a pattern. The third question tested the participants' knowledge of the manipulations in the last run (regardless of their response to the previous question). For the VMA task, the participants were forced to judge if the feedback cursor had been shifted, relative to hand position, to the left, right, or both. For the sequence learning tasks, the participants were told the length of the sequence (12 for SRT; 8 for ASRT) and asked to generate, as best they could, the repeating sequence.

We recognize that there are limitations with the use of post-experimental questionnaires as probes of probe awareness (Perruchet and Amorim 1992; Howard and Howard 1997). Based on previous research, we were confident that awareness would be low in the VMA task given the subtle changes in the size and direction of the perturbation from trial-to-trial, as well as limited size of the maximum perturbation (Cheng and Sabes 2007; Schlerf et al. 2013). Similarly, awareness has been shown to be very low or non-existent in the ASRT task due to the insertion of the random elements (Howard and Howard 1997; Janacsek and Nemeth 2012). The SRT task is more problematic in terms of awareness, with performance measures sometimes providing results that are at odds with post-experimental subjective reports (e.g. Perruchet and Amorim

1992; Shanks et al. 2005). Given this, we also examined performance measures as possible indicators of awareness in the analysis of these data.

Data analysis

For each task, we calculated measures of performance, variability, and learning. These data were calculated separately for each run, allowing tests of reliability. When appropriate, we averaged the data over the two runs to make comparisons between tasks. All data was analyzed using Matlab (Mathworks, Natwick, MA).

Correlational analyses were used to evaluate reliability (comparing the same measure from two different runs) and to compare different performance measures (comparing two measures, each averaged over the two runs). Pearson correlation coefficients represent the strength of the linear dependence between the variables of interest. We also employed orthogonal (Deming) regression to evaluate the relationship between variables of interest (Cornbleet and Gochman 1979; Linnet 1993, 1998). Regular least squares regression assumes that the x-axis variable is noise free and that all of the variance is in the y-axis variable. In contrast, the Deming method is symmetric with respect to the two variables, making no assumptions regarding their dependence or independence, and does not assume there is no variance for the variable plotted on the x-axis. This approach is appropriate for the current study since we make no assumptions about dependence (e.g., second run performance is not influenced by first run performance). Thus, Deming regression provides the best estimate of the slope of the relationship between the x and y variables, based on estimates of noise in both variables.

I. Visuomotor adaptation (VMA) task:

Using the output from the stylus, movement trajectories were reconstructed off-line. To minimize the contribution of on-line corrections, the initial heading angle was used as the primary dependent variable. This was calculated as the angle between a line connecting the center of the start circle and the center of the target, and a line connecting the center of the start circle and hand position midway to the target (4 cm from start position). We also calculated reaction time and movement time. The former was measured as the time from target onset until the time at which the hand crossed a 1 cm ring around the center of the starting circle. The latter

was measured as the time between the reaction time and time at which the cursor intersected the target circle.

Trials were discarded when the heading angle was more than ± 3 SD from the mean heading angle or the movement time was more than 3 SD longer than the mean movement time (calculated per participant and run). Two percent of the trials were removed based on these criteria ($2.0 \pm 0.68\%$, ranging from 0.2 to 4.4% over participants). We did not exclude data based on the RT data since the instructions did not emphasize reaction time. All 100 participants tested on VMA were included in the analysis.

Learning in visuomotor rotation tasks is manifest by changes in heading angle that counter the direction of the perturbation. As a global measure of learning, we calculated the cross-correlation between the sequence of heading angles and the sequence of perturbations across all rotation trials (Rotate1 and Rotate2). We repeated this calculation by shifting the heading angle time series up to a quarter of a cycle to identify the lag at which the correlation was largest (based on goodness-of-fit).

Our main analysis of the learning data focused on the participants' learning-rate, measured by the trial-by-trial response to the evolving rotation. We employed a standard state-space model of adaptation (Thoroughman and Shadmehr 2000; Cheng and Sabes 2007), where the internal model (x) of the motor system (in this case heading angle) is updated at each trial ($n+1$) according to a learning rule:

$$1) \quad e_n = x_n + r_n$$

$$2) \quad x_{n+1} = x_n - B * e_n$$

The error (e_n) experienced at trial n is given by the difference between the hand heading angle (x) and the rotation (r). The hand heading angle for the next trial (x_{n+1}) is updated based on the weighted (B) error on the current trial. Effectively, B is a learning rate, reflecting the percentage of error that is corrected from trial-to-trial. As such, large values of B indicate faster learning. This model was fit to estimate each participant's learning rate (B) by minimizing the root-mean-square (RMS) error between the actual time series of heading angles and the model output (equation. 1) using the *fminsearch* function in Matlab. We obtained a single measure of learning for each session by fitting the data from both the Rot1 and Rot2 blocks.

A main question of interest in the current experiment is to assess individual differences in sensorimotor adaptation, and to ask if these individual differences are related to movement

variability. To measure performance variability that was not contaminated by learning, we focused on the data from the Baseline1 block. Spatial variability (sSD), defined as the standard deviation of the heading angles across the Baseline1 block, served as our measure of performance variability for the VMA task. Movement time (MT) was also taken from the Baseline1 block.

II. Serial Reaction Time (SRT) task:

The median reaction time (RT) was calculated for each block. We used the median, rather than mean RT to minimize the effect of outliers, although the results are similar with either measure. Only correct trials were used in the analysis. Learning in the SRT task is defined as the increase in RT on random blocks following training with sequence blocks.

It is important to note that the learning measure for the SRT task is not based on a trial-by-trial measure, but rather involves a comparison across blocks in which the stimuli are predictable or selected at random. Our initial plan was to focus on the contrast of learning at the end of the task (as done in previous SRT studies, e.g. Spencer and Ivry, 2009), comparing the final two random blocks (Blocks 13 and 15) and the final two sequence blocks (blocks 12, 14). After finding that the learning measure defined on the final random blocks showed poor reliability (see below), we also defined a second learning measure (initial learning) based on the random block included in the middle of the task, or what we will refer to as the mid-task probe. Here we compared the median RT on this random block (block 7) and the two adjacent blocks (blocks 6, 8).

In our analysis of individual differences, we wanted to see if variation in SRT learning was related to any performance variables. To this end, we used measures of reaction time (RT) from blocks 2-4. At this point in training, we expected learning to be modest. The means of these RT data were used to estimate basic motor speed and temporal variability (tSD). For the latter, we took the standard deviation of the RT data and normalized this by dividing by the mean RT for these blocks.

We did not have data for both runs for four of the 60 participants from the SRT task (one participant was accidentally tested twice on the same sequence; the program did not run properly on the second session for three participants). In addition, two participants were excluded from the final analysis because their RTs increased by >200 ms on the final random blocks compared

to the neighboring sequence blocks (and correspondingly their learning scores were >3 standard deviations above the group mean learning score). We interpret this as a performance-based indicator that these participants had developed explicit knowledge of the sequence during at least one of the testing sessions. Finally, from the questionnaire data, one participant was fully aware of the sequence, recalling the entire sequence, and we thus excluded this participant from the analysis. In total, the data from 53 participants were included in the final analysis.

As described above in the *Procedure* section, the third questionnaire question required participants to generate a 12-element pattern, attempting to produce the sequence employed in the second run. From these data, we calculated a match-index (*MI*) based on the generation of three or more successive elements that corresponded to the actual sequence (Reber and Squire 1998; Wilkinson and Shanks 2004; Bo et al. 2011; Martini et al. 2013). This composite score *MI* served as a proxy of awareness, with scores ranging from 0 (no match) to 1 (full recall). We calculated the probability of recalling triplets, quartets, and quintets from the original sequence (for example, SeqA includes triplets of 241, 413, 134, etc., quartets of 2413, 4134, 1342, etc., and quintets of 24134, 41342, etc.), comparing this to the probability of randomly recalling matches (avoiding repetitions). For the latter, the probabilities to report, by chance, a triplet, a quart, and a quint given 12-element sequences are 0.33, 0.11, and 0.037 respectively. We gave larger weight to longer sequences, using weights equal to the length of mini-sequences. Thus, by chance the *MI* score would be 0.316 ($3 \times 0.33 + 4 \times 0.11 + 5 \times 0.037$). We did not look for matches longer than 5-elements since these matches have minimal impact on the expected value by chance. The *MI* scores were only calculated for the second sequence since we assume this information, if available, would be most active at the time of the questionnaire. This analysis was limited to 52 participants since questionnaire data was missing from one participant.

III. Alternating Serial Reaction Time (ASRT) task:

RTs from correct trials were grouped over triplets of 3 sequential trials. Two types of triplets were created from the data: High-frequency triplets, in which the first and third elements are sequential elements of the basic sequence, and low-frequency triplets, in which the first and third elements do not follow the basic sequence. In the 1r2r3r4r example above (r indicating random elements), triplets of 1x2, 2x3, 3x4 and 4x1 (x indicating any possible element, either random or sequence) would occur more often, either from including two consecutive stimuli of

the repeating sequence (e.g. 132 created from 1r2) or from two consecutive random elements that, by chance, follow the basic sequence (e.g. the same 132 for a triplet of r3r). By contrast, triplets of 1x3 or 4x2 (e.g., 123) would occur less frequently because they could never be obtained from two consecutive sequence elements. Rather, they would only occur when an r2r triplet results in a specific random selection (e.g., where the initial and final r's are chosen to be 1 and 3). Overall, there were 68% high-frequency triplets and 32% low frequency ones in the experiments (as in Nemeth et al., 2013).

Following the criterion used in previous studies (Howard and Howard 1997; Nemeth et al. 2010; Janacsek et al. 2012), and similar to the restrictions imposed in the SRT task, two kinds of low-frequency triplets were excluded from our analyses: repetitions and trills. The elimination of these special triplets ensured that differences between high- and low-frequency triplets were not confounded by response patterns that might produce fast RTs.

Sequence learning in the ASRT is calculated by comparing RT for high-frequency and low-frequency triplets (Nemeth et al., 2010; Janacsek et al., 2012). The difference between the two types of triplets was calculated separately for each block by subtracting the median RT of the high-frequency triplets from that of the low-frequency triplets (Nemeth et al., 2010). Note that, unlike the SRT task, learning in the ASRT is based on a within-block, trial-by-trial measure. To test for reliability between sessions, we calculated an average learning value over the entire session (from blocks 4-45).

To relate learning to performance measures, the overall mean and standard deviation of the RT data were calculated from blocks 2-3, again choosing an early time point where learning should be minimal. We did not include the block 1 data, given the assumption that participants are becoming familiar with the task during this block. As with the SRT data, we created a normalized measure of temporal variability (tSD) by dividing the standard deviation of RT during blocks 2-3 by the mean RT of these blocks.

Four of the 25 participants in Group 4 were excluded from the final analysis. The mean RTs for two participants were more than 3 standard deviations from the group mean, and the temporal variability of the other two participants was more than 3 standard deviations from the group mean. In total, the data from 21 participants were included in the final analysis.

Results:

Individual Differences in Visuomotor Adaptation

To look at implicit sensorimotor adaptation, we used a pseudorandom perturbation of visual feedback, cycling, every 80 trials, between a 12° clockwise rotation and 12° counterclockwise rotation. Overall, there was a group bias as participants tended to reach in a slightly CCW direction of the targets in the absence of any perturbation (baseline block), a pattern that was evident throughout the task. The participants adapted to the perturbation, modifying their initial heading angle in a direction opposite the perturbation (**Fig 2a** shows the group average over both runs). In terms of our global measures of learning (over all participants and runs), the mean correlation between the perturbation and initial heading was 0.66 ± 0.13 , and the best fit was obtained when the heading angle function lagged the perturbation function by 4.9 ± 1.5 trials. Note that despite the similar low frequency of the perturbation and behavior, the high frequency local jitters reduced the overall correlation between the graphs. The mean learning rate, calculated by the state-space model, was 0.13 ± 0.04 .

Three different groups were tested on the VMA task, differing in whether they performed this task alone or in combination with either the SRT or ASRT task. For all three groups, the VMA task was repeated in two separate runs (on separate days), with different orders of target locations and perturbations for the two runs. We first assessed if there were any differences between groups on our three primary dependent variables: movement time, spatial variability (both calculated from Baseline1 trials before the visuomotor perturbation was introduced), and learning rate (estimated from the Rotate1 and Rotate2 trials using the state-space model). For this analysis, we conducted mixed ANOVAs with the between-subject factor Group and within-subject factor Run. In all of the ANOVAs, there was no effect of Group, nor a Group X Run interaction ($p > 0.1$). In a secondary analysis limited to Group 3, where some participants completed the VMA task before the SRT one and vice versa for others, we confirmed that there was no effect of order ($p > 0.15$), indicating that performance of the VMA task was unaffected by whether the participants had first completed the SRT task. Over the three groups, there was a significant effect of Run on movement time (Run 1 = 451.9 ± 55.1 ms; Run 2 = 431.6 ± 50.5 ms; $p < 0.001$) and learning rate (Run 1 = 0.133 ± 0.037 ; Run 2 = 0.120 ± 0.040 ; $p < 0.001$) indicating that participants reached faster and learned slower during the second run. There was no change in spatial variability across runs (Run 1 = $4.6 \pm 1.1^\circ$; Run 2 = $4.6 \pm 1.8^\circ$; $p = 0.46$).

Our main interest in this study was to explore individual differences in implicit motor learning. Such questions are only sensible if our individual measures are reliable over time. Importantly, even though there was a significant change in the group means between runs, measures can be reliable if the rank order of the samples remains more or less constant. Given that there was no effect of Group, we can look at this question in a large sample by combining the 100 participants into one group, and comparing performance on the VMA task between the two runs (with minimum separation of two days). As shown in **Figure 2b**, reliability was high for movement time ($r=0.77$; 95% CI=[0.68, 0.84]), spatial variability ($r=0.52$; 95% CI=[0.36, 0.65]), and learning rate ($r=0.78$; 95% CI=[0.69, 0.85]), (all p 's $<<0.001$). These data demonstrate that there is some stable factor or set of factors that underlie individual differences in terms of how the participants reached, as well as in terms of their ability to respond to a visuomotor perturbation. Note that while movement time was constrained experimentally, there remained considerable variability in how fast the participants moved and this remained consistent over runs.

We next looked at the relationship between these three variables, averaging the scores for the two runs for each individual (**Fig. 2c**). There was a trend for movement time and learning rate to be positively correlated (**Fig. 2c** middle panel; $r=0.17$, $p=0.09$, 95% CI=[-0.03, 0.36]), with faster learning associated with slower movements. However, movement time and spatial variability were not correlated (**Fig. 2c** left panel; $r=-0.094$, $p=0.35$, 95% CI=[-0.28, 0.10]), indicating that at least in terms of initial heading, there was no speed-accuracy trade-off. Importantly, there was a significant negative relationship between spatial variability and learning rate (**Fig. 2c** right panel; $r=-0.41$, $p<<0.001$, 95% CI=[-0.55, -0.22]): Participants who exhibited higher reaching variability during baseline blocks adapted to the visuomotor rotation at a slower rate.

To examine the influence of awareness on learning, we looked at the questionnaire data. Since we didn't want to bias participants to expect a perturbation, the questionnaire was administered only at the end of the second run. The data indicated that participants had no or minimal awareness of the visuomotor perturbation. Only a third (38%) of the participants reported that they "experienced any change in apparatus during the test session." When forced to explain the change, none described a mismatch of the visual feedback; rather they offered a large variety of explanations (e.g., "I became better with practice" or "the distance between the

starting point and end point is becoming longer”). When asked if they belonged to a no-perturbation (match) or perturbation (no-match) group with a 7-point Likert scale, the mean rating was 2.5 ± 1.6 , falling between the “2: Somewhat confident in match group” and “3: Minimal confidence but expect in match group” categories (**Fig. 5c**). When forced to indicate the direction of a “mismatch”, the majority (57.7%) selected the answer that there was a unidirectional mismatch, while the actual perturbation included both clockwise and counter clockwise rotations.

To summarize, we observed impressive reliability in measures of performance and learning on a visuomotor rotation task. Adaptation to the pseudorandom perturbation was implicit for all participants and faster for those with lower baseline variability.

Individual Differences in Sequence Learning:

A) The Serial Reaction Time task (SRT)

To study implicit sequence learning, we used a 12-element version of the SRT task in which participants responded to the location of visual stimuli by pressing spatially compatible keys on a computer keyboard. Accuracy was relatively high ($93\% \pm 7\%$ across participants over all blocks). Over all participants and the two runs, there was a significant increase in RT on the final two random blocks, indicative of sequence learning (**Fig. 3a**, left panel: 36.1 ± 23.6 ms; t-test vs. 0; $t(52) = 11.1$, $p < 0.001$). There was also an increase in RT for the mid-task probe, although to a smaller extent (11.2 ± 22.5 ms; $t(52) = 3.6$, $p < 0.005$). A similar pattern was evident in the accuracy data (**Fig. 3a**, right panel), with mean increases in error rates of $2.2 \pm 3.3\%$ and $1.8 \pm 4.9\%$ in the late and mid-task probes, respectively ($t(52) = 4.8$, $p < 0.001$ and $t(52) = 2.7$, $p < 0.01$, respectively). We note that these learning costs were evident, despite the fact that RT showed only a modest decrease over the course of the sequence blocks.

Two groups were tested on the SRT task. One group (Group 2) was tested in a single session, completing two runs with different sequences. The other group (Group 3) was tested over two separate sessions, with each session also including VMA testing. As with the VMA analysis, we first assessed if there were any differences between groups on our three primary dependent variables: Two performance variables, RT and temporal variability, calculated from the early stages of the task (blocks 2-4) where we assume learning is relatively modest; and learning, calculated from the final-task random probes. These data were analyzed using a set of

mixed-design ANOVAs with the between-subject factor Group and the within-subject factor Run. In all of the ANOVAs, there was no effect of Group or Group x Run interaction ($p > 0.34$). In an analysis limited to Group 3, there was also no effect of order ($p > 0.15$), indicating that SRT performance was unaffected by whether the participants had first completed the VMA task.

There was a significant Run effect on RT. Participants responded faster in the early baseline blocks of Run2 compared to Run1 (Run 1 = 373 ± 65 ms; Run 2 = 343 ± 59 ms; $p < 0.001$). The reduction in reaction time came at a modest, non-significant cost in accuracy (Run 1 = $96.5 \pm 4.1\%$; Run 2 = $93.5 \pm 8.1\%$; $p = 0.11$) and an increase in variability, operationalized as the standard deviation of these baseline RTs (divided by the mean RT), although this effect was not significant (Run 1 = 0.041 ± 0.029 ; Run 2 = 0.052 ± 0.031 ; $p = 0.05$). Learning was comparable for the two runs (Run 1 = 10.1 ± 8.3 ; Run 2 = 8.9 ± 8.6 ; $p = 0.43$). In terms of sequence specific learning, participants showed faster reaction times for SeqA than for SeqB (SeqA: 44.8 ± 33.8 ms; SeqB: 27.3 ± 28.4 ms, $p < 0.005$).

In the absence of Group effects, we combined the data from the 53 participants to assess reliability on the SRT task. Baseline RTs were reliable across runs (**Fig. 3b** left panel; $r = 0.63$, $p < 0.001$, 95% CI = $[0.43, 0.77]$). In contrast, our measure of variability, the standard deviation of these baseline RTs, was not reliable (**Fig. 3b** middle panel; $r = -0.12$, $p = 0.38$, 95% CI = $[-0.38, 0.15]$). Even more striking, the measures of sequence learning were not reliable when taken at the end of the runs (**Fig. 3b** right panel; $r = 0.07$, $p = 0.63$, 95% CI = $[-0.21, 0.33]$). Thus, while we observed consistent learning on this task in the group analysis, the degree of learning in Run1 was not predictive of the degree of learning in Run2 on an individual basis. We also looked at reliability by correlating learning scores between the sequences, correlating Sequence A with learning scores on Sequence B. Here, too, we failed to observe consistency in terms of individual differences ($r = 0.14$, $p = 0.3$, 95% CI = $[-0.14, 0.39]$).

There are a few reasons why reliability might be poor for our measure of sequence learning. First, whereas learning in the VMA task is based on trial-to-trial adjustments, learning in the SRT task is based on a comparison across blocks. The latter may be more susceptible to contamination from other factors (e.g., fluctuations across blocks in motivation or fatigue). We return to this issue below when examining performance on the Alternating-SRT task where learning is based on trial-to-trial measures.

Second, a subset of participants may have developed some degree of awareness in one session or the other. The cost observed on random blocks is much larger when people are aware of the sequence, in large part because participants can anticipate the stimulus (as reflected in RTs under 200 ms and/or in learning scores above 200 ms). As such, we would expect a poor correlation if some participants were aware of the sequence in one run and unaware in the other. As noted in the Methods, we excluded participants with high levels of sequence awareness by excluding the data from the two participants who had learning scores greater than 3 standard deviations of the mean score, and the one participant who recalled the full sequence.

Nonetheless, there may be some graded level of contamination from awareness. Indeed, the questionnaire data indicated that participants might have had some awareness of the repeating sequence, with slightly more than a third (40%) of the participants agreeing with the statement that they had “experienced any change in the experiment during the test session.” When asked to mark their confidence in being assigned to a random group or a sequence group on a 7-levels Likert-scale (**Fig. 5c**), the mean rating was 4.17 ± 1.85 , not significantly different from “4: Not sure which group” ($p=0.51$). We also asked participants to recall or make their best guess of the repeating sequence on Run2. The amount of recall was measured by a match-index, ranging from 0 (no match to actual sequence) to 1 (perfect match), with random guessing expected to produce a value of 0.136. The mean MI value over all participants was 0.21 ± 0.15 , significantly above chance (two-tailed t-test: $p<0.001$), suggesting participants had partial awareness of some segments of the sequence.

The absence of reliable individual differences in SRT learning, coupled with the awareness concern, led us to examine an alternative measure of learning. In particular, we reasoned that awareness (as well as factors such as fatigue) would be less problematic during the initial stages of the experimental run. As such, we performed a post-hoc analysis of learning using the mid-task random probes (see **Fig. 1b**). As expected, learning was lower at this initial probe compared to the late probe ($t(52)=7.74$, $p<<0.001$, between the two probes ; **Fig. 3d**). Interestingly, these learning scores showed modest reliability across runs, despite having a more restricted range than that observed for the late learning scores (**Fig. 3d**, left panel; learning: $r=0.27$, $p<0.05$, 95% CI=[0.00, 0.50]). We also observed reliability between the initial and late stages of learning ($r=0.49$, $p<0.005$, 95% CI=[0.25,0.67]), suggesting that despite the possible

contamination of explicit factors on the late learning measure, some aspect of implicit learning was maintained over the course of the experimental task.

Last, we looked at the relationship between our dependent variables, averaging the scores for the two runs for each individual (**Figs. 3b, 3d**). There was no correlation between RT and temporal variability (**Fig. 3b**, left; $r=-0.15$, $p=0.3$, 95% CI=[-0.4, 0.13]). Participants who responded faster during the baseline blocks exhibited greater learning, as indicated by significant negative relationships between baseline reaction time and learning, during both late (**Fig. 3c**, middle; $r=-0.45$, $p<0.01$, 95% CI=[-0.64, -0.20]) and initial (**Fig. 3d**, middle; $r=-0.58$, $p<<0.001$, 95% CI=[-0.73, -0.36]) stages of the task. There was no correlation between learning and baseline variability for the late learning probe (**Fig. 3c**, left; $r=0.04$, $p=0.79$, 95% CI=[-0.24, 0.30]), whereas there was a weak correlation with the initial learning (**Fig. 3d**, left; $r=0.31$, $p<0.05$, 95% CI=[0.05, 0.54]).

In summary, we observed robust learning on the SRT task. However, when we focused on the traditional measure used in SRT studies, namely RT differences between the final sequence and random blocks, we failed to observe reliable individual differences: Learning in one run was not predictive of learning in a second run. This lack of reliability on the learning scores stands in contrast to the reliable individual differences in mean RT. We suspect that two factors contribute to the lack of reliability. First, SRT learning is based on comparing RTs across blocks and various factors can produce considerable variation in RT across blocks. Second, the late probes may be contaminated by awareness. Consistent with these hypotheses, we did observe a modest degree of reliability when the analysis was limited to the initial probe. Moreover, the degree of learning at this earlier probe was larger for participants who responded slower and exhibited larger variability in the early stages of the task.

B) The Alternating Serial Reaction Time Task (ASRT)

Given the concerns with the SRT task, we opted to use the alternating serial reaction time (ASRT) task as a different way to assay sequence learning. One major advantage of the ASRT is that learning is calculated continuously from each triplet of trials (see Methods). Moreover, as reported in previous studies, learning in the ASRT task appears to be completely implicit (e.g. (Howard and Howard 1997; Nemeth et al. 2010; Janacek et al. 2012)). Only a single group was

tested on the ASRT task, with each participant performing the task on two separate sessions. Accuracy was reasonably high in this task ($90\pm3\%$ across participants).

We observed robust improvements in performance on the ASRT task (**Fig. 4a**). Overall, RTs became faster over the successive blocks and there was a substantial overall reduction in RT on Run2 compared to Run1. In terms of sequence learning, we focus on the comparison of RTs to high- and low-frequency triplets. Sequence learning was evident in both runs. In separate repeated measures ANOVAs, the effect of Triplet type (high vs. low) was significant for both Run1 ($p<<0.001$) and Run2 ($p<<0.001$). The effect of Block was only significant for the Run1 data (Run1: $p<<0.001$; Run2: $p=0.28$). However, the Triplet Type x Block interaction was significant on both runs (Run1: $p<0.001$; Run2: $p<0.05$), reflective of the fact that the difference between the high and low frequency triplets increased over the test session, indicative of learning.

As can be seen in **Figure 4**, the participants were much faster in the second run. This performance difference was evident in blocks 2 and 3, which serve as our proxy for baseline performance where sequence learning is assumed to be low. Not only were RTs faster in Run2 compared to Run1 (Run1= 379.7 ± 40.0 ms, Run2= 345.3 ± 26.4 ms, $p<<0.001$), but variability was also lower in these blocks (SD/Mean: Run1= 0.192 ± 0.034 , Run2= 0.174 ± 0.034 ms, $p<0.001$). Mean sequence learning across the run (see Methods) was also lower in Run2 (Run1= 14.4 ± 7.3 , Run2= 10.1 ± 5.5 , $p<0.01$).

We next turned to our assessment of reliability (**Fig. 4b**). In terms of baseline measures of performance, individual differences in mean RT and temporal variability were reliable across runs (RT: $r=0.64$, $p<0.005$, 95% CI=[0.29, 0.84]; tSD: $r=0.51$, $p<0.02$, 95% CI=[0.096, 0.77]). Learning on the ASRT was also reliable (RT: $r=0.46$, $p<0.04$, 95% CI=[0.04, 0.74]). As can be seen in the right panel of **Fig. 4b**, most of the data points fall below the unity line, reflective of the fact that learning was lower on Run2 compared to Run1. Nonetheless, participants who showed higher levels of learning on Run1 tended to show higher levels of learning on Run2. This correlation is especially impressive given the very modest level of learning (and thus restricted range) on the ASRT.

Given that individual differences in probabilistic sequence learning are reliable, we can ask if these differences are related to baseline performance measures (**Fig. 4c**). First, we observed a trend for a positive correlation between RTs and temporal variability, a pattern that is

in the opposite direction to that seen for the SRT task (**Fig. 3c**, left panel; RT: $r=0.41$, $p=0.06$, 95% CI=[-0.03, 0.72]). Also in contrast to the SRT task, we observed a trend for a positive correlation between RT and learning with the ASRT task: Better learning was associated with slower RTs (**Fig. 4c**, middle panel; RT: $r=0.42$, $p=0.05$, 95% CI=[-0.01, 0.72]). There was also a positive correlation between sequence learning and temporal variability: Better learning was associated with higher variability (**Fig. 4c**, right panel; RT: $r=0.56$, $p<0.01$, 95% CI=[0.17, 0.80]), a result consistent with that observed in the initial probe data for the SRT task. We note that in the ASRT task, high temporal variability could reflect a difference in RT between high and low frequency triplets, a difference that is accentuated in good learners. However, the positive correlation is also observed if the analysis is performed separately for the high and low triplets (high: $r=0.52$, $p<0.05$; low: $r=0.57$, $p<0.01$).

Analysis of questionnaire data (**Fig. 5c**) indicated that participants were not aware to the presence of a repeating probabilistic sequence. Only a third (33%) of the participants reported that they “experienced any change in the experiment during the test session.” From the Likert scale, participants were more likely to judge that they had been in the random (no sequence) group (2.62 ± 1.32).

To summarize, we observed impressive reliability in measures of performance and learning on probabilistic sequence learning. Learning the probabilistic sequence was implicit for all participants and greater for those with slower baseline RTs and higher baseline variability.

Individual differences across tasks:

The participants in Groups 3 and 4 completed two implicit motor learning tasks, the VMA and the SRT tasks (Group 3) or the VMA and ASRT tasks (Group 4). We compared performance across tasks, asking if there are consistent individual differences in implicit learning in two distinct task domains. Within each task, we averaged performance across the two runs. Note that given the lack of reliable learning in the late probe for the SRT task, the data from this task should be viewed cautiously.

Table 1 presents the correlation matrix and **Figure 5** depicts correlations between the different measures for each task. Overall, there was little correlation between the rate of adaptation and the magnitude of sequence learning, the key measures of implicit learning. There was essentially no correlation between VMA adaptation and SRT learning (**Fig 5a**),

measured either at the end of training or at the midpoint of the task (where learning is likely implicit and the measure was reliable). Regarding ASRT, there was a positive correlation, albeit non-significant, between VMA adaptation and ASRT learning over all participants (**Fig 5b**; $r=0.38$, $p=0.17$, 95% CI=[-0.14, 0.65]). However, this correlation is largely driven by one participant who showed faster adaptation and greater sequence learning. We consider the effects of this outlier in two ways. First, when we simply removed this data point, the correlation coefficient not only became much weaker but the direction actually reversed ($r=-0.17$, $p=0.48$, 95% CI=[-0.57, 0.3]). Second, we used a more conservative winzoring procedure, replacing the values of the outlier with the SRT/VMA values from the closest neighbors. This procedure also indicated that there was no correlation between the two learning scores ($r=0.03$, $p=0.87$, 95% CI=[-0.4, 0.46]). In sum, both trimming methods indicate that there is no relationship between learning on the ASRT and VMA tasks. Thus, the correlational data fail to support the hypothesis that there is some common process that contributes to implicit motor learning across task domains.

The correlation matrix does highlight one other important feature of the results. Learning was correlated with motor variability for both the adaptation and sequence learning tasks. However, the direction of the correlation was opposite. For adaptation, the correlation was negative, suggesting that people with high spatial variability adapt at a slower rate. For sequence learning, the correlation was positive, suggesting that people with high temporal variability learn sequential associations at a faster rate. This effect was only significant for the ASRT task; nonetheless, a similar pattern was observed in the late phase of the SRT task.

Discussion:

Motor skills occupy a prominent place as representative of implicit, or procedural memory in standard memory taxonomies (Squire and Zola 1996; Sun et al. 2007). We used an individual differences approach to examine the relationship between two classes of tasks commonly used to study implicit motor learning, sensorimotor adaptation and sequence learning. Participants were tested in two separate runs, allowing us to first examine if measures of learning and performance were reliable – a prerequisite for evaluating individual differences. We then

examined factors that might account for individual differences within each task, as well as the relationship in learning between the two task domains.

Reliability of implicit learning

Surprisingly, we were unable to find any previous work examining the reliability of individual differences in implicit motor learning. Studies that tested people over multiple days have generally focused on consolidation, and thus repeat the exact same task over successive days to look at variables such as forgetting or off-line learning (e.g. Doyon et al., 2009; Krakauer, 2009; Nemeth et al., 2010; Meier and Cock, 2014). These consolidation factors limit our ability to draw inferences about reliability since there may be individual differences in consolidation and other memory processes. Savings, the phenomenon in which people show faster relearning of a previously forgotten memory, has been observed in visuomotor adaptation tasks, but these studies have not reported individual differences and have used conditions in which participants are likely aware of the perturbation (Zarahn et al. 2008; Morehead et al. 2015).

In the current study, we observed reliable individual differences in core performance measures. Focusing just on the data from the baseline phase of the tasks, movement time was highly correlated across runs in the VMA task, and RT was reliable in both the SRT and ASRT tasks. We do not view these differences as indicative of some sort of individual limit in performance. Rather, they may be best viewed as a “signature” of a consistent individual style, perhaps reflecting an idiosyncratic criterion for how people choose to trade-off speed and accuracy in the context of each task. During the second exposure to the tasks, participants reached faster in the VMA task and had faster reaction times in the sequence learning tasks, suggesting that familiarity with the tasks led to similar improvements (broadly speaking) across participants.

Our main interest concerned the reliability of measures of implicit learning. The rate of trial-by-trial adaptation exhibited reasonable reliability across runs on the VMA task with trial-by-trial adjustments that, on average, corrected for about 10% of the error. Similarly, learning, measured by the difference in RT to high and low predictable stimuli, was reliable in the ASRT task with a mean overall difference of 12 ms between the high- and low-frequency elements. Post-experiment questionnaires confirmed that learning in the VMA and ASRT tasks was

implicit. We employed a very gradual perturbation in the VMA task and never allowed the perturbation to exceed 12 deg. For the ASRT task, various lines of evidence suggest that the presence of random elements for 50% of the trials completely disrupts awareness of the repeating elements (e.g. Howard and Howard, 1997; Nemeth et al., 2010). The modest learning scores are also consistent with the hypothesis that learning was implicit.

Whereas we found consistent group-level learning on the SRT task, the standard probes of learning were not reliable across runs: On both runs there was a marked increase in RT on the final random blocks relative to surrounding sequence blocks, but the magnitude of this increase was not consistent between the two runs at the individual level. We believe there are at least two factors for the lack of reliability on the SRT task. First, learning on the form of the SRT task used here is based on a measure that requires averaging data across all of the trials within a block, and then comparing means across blocks. Variables that might introduce block-by-block variation or low frequency changes in performance can have a large impact on such measures. For example, the motivation level or idiosyncratic speed-accuracy criterion might fluctuate from one block to the next. The VMA and ASRT tasks are not impacted by such factors given that their learning measures are continuously based on trial-to-trial performance.

Second, the lack of reliable SRT learning could be related to awareness. There is an extensive literature concerning the role of awareness in the SRT task, as well as the methodological challenges for assessing or eliminating/minimizing awareness (e.g. by adding a concurrently secondary task: Hazeltine et al., 1997; Grafton et al., 1998; Gheysen et al., 2009; Sanchez et al., 2010; or by using long sequences: Pascual-Leone et al., 1993; Spencer and Ivry, 2009; Sanchez and Reber, 2012). Empirically, a number of the participants exhibited a large increase in RT on the random blocks in the late learning probe (e.g., 30% had an increase >50 ms), a difference that, coupled with fast RTs on the sequence blocks, is suggestive of awareness. Moreover, while recognizing limits in drawing comparisons between the tasks on the questionnaire data, it is noteworthy that participants' confidence that there was a repeating pattern was higher in the SRT task compared to the ASRT task and higher than participants' confidence that there was a perturbation in the VMA task. Awareness would contaminate our measure of implicit learning (Willingham et al. 1989; Reber and Squire 1998) and impact assessments of reliability. Consistent with this hypothesis, we observed a modest correlation of

the learning scores at the mid-task probe, a time point at which we expect awareness would be lower.

Note the contamination from awareness could either increase or decrease reliability. If a subset of the participants were aware (and thus, had larger learning scores) on both Run1 and Run2, the correlation would be inflated. On the other hand, awareness could decrease reliability if different subgroups of participants become aware in Run1 and Run2. The current results suggest that the latter is more likely to have occurred here. We are not claiming that explicit learning is unreliable. Had we used a completely explicit sequence-learning task, reliability may have been high. The problem, though, is when a measure of learning is a composite of multiple processes, with the weight given to those processes varying across measurements.

In summary, the current results indicate implicit measures of learning were reliable at the individual level in response to a pseudorandom visuomotor rotation and in becoming sensitized to the predictability of sequential elements. The lack of reliability in the SRT task presents a challenge for studies that have used this task to study individual differences (Unsworth and Engle, 2005; Norman et al., 2007; Noohi et al., 2014) and the presence of awareness also compromises the utility of this task for studying implicit learning (Hazeltine et al. 1997; Grafton et al. 1998; Keele et al. 2003). For example, despite concurrently performing a secondary task, participants could have developed different levels of awareness to the repeating sequence (which would be difficult to track by directly asking participants at the end of the task).

Is there a common “implicit motor learning” component?

Studies of individual differences in motor control and learning have a long history. One prominent claim has been that differences are task specific, a hypothesis supported by studies using large test batteries (e.g. Parker and Fleishman, 1960) or studies of expertise (Bachman, 1961; Chase and Simon, 1973; Mann et al, 2003). The alternative perspective, advocated by Franklin Henry’s “specificity hypothesis” (see Henry, 1958) is that some component of individual differences arises from a set of core operations that are shared across tasks. Building on this idea, Keele and colleagues (Keele et al. 1985, 1987; Ivry and Keele 1989) provided evidence that individual differences in motor control might reflect variation in how consistent people are in controlling certain parameters of movement such as timing, force, or rate. For example, people who were consistent in timing repetitive movements also exhibited consistent

845 timing when producing isometric force pulses or in judging the duration of time events. These
846 ideas, coupled with evidence from neuropsychological studies (e.g., Ivry and Keele, 1989) led to
847 models in which different neural systems were associated with particular component operations.

848 Extending the logic of that work, we asked here if “implicit motor learning” might
849 constitute a core component of skill acquisition (see Conway et al., 2010). To this end, we
850 looked at between-task correlations. Given the reliability in the VMA and ASRT tasks, we
851 combined the results from the two runs in making these comparisons. We also included the SRT
852 task, although these data should be viewed cautiously given the lack of reliability in the measure
853 of learning. The results showed no correlation between the learning measures on the different
854 tasks, arguing against the notion of a common implicit learning process shared across motor
855 learning tasks, at least for visuomotor adaptation and sequence learning.

856 The lack of correlation may be surprising given that learning in both tasks has been
857 associated with cerebellar function. Patients with cerebellar pathology, either from focal insult or
858 degeneration, have pronounced impairments in learning when tested on sensorimotor adaptation
859 (Martin et al. 1996; Shin and Ivry 2003; Smith and Shadmehr 2005; Tseng et al. 2007; Rabe et
860 al. 2009; Taylor et al. 2010; Werner et al. 2010; Donchin et al. 2012; Schlerf et al. 2013) and
861 sequence learning tasks (Pascual-Leone et al., 1993; Molinari et al., 1997; Gómez-Beldarrain et
862 al., 1998; Shin and Ivry, 2003; but see Spencer & Ivry, 2009). Neuroimaging studies in humans
863 (Della-Maggiore and McIntosh 2005; Danckert et al. 2008; Bernard and Seidler 2013) also point
864 to an essential role for the cerebellum, either in generating the error signals to guide learning
865 (Diedrichsen et al. 2005; Schlerf et al. 2013) and/or generating the predictions of expected
866 sensory events (Ramnani et al. 2000). Based on the implicit nature of these tasks and the
867 association of the cerebellum with both sensorimotor adaptation and sequence learning, one
868 might have expected to observe a correlation in learning between the learning tasks.

869 However, there are a number of reasons that may explain why the measures of learning
870 were not correlated. First, based on computational considerations, adaptation and sequence
871 learning require very different learning mechanisms and operations. Adaptation is error driven,
872 entailing the modification of an internal model based on the difference between predicted and
873 observed sensory feedback (Taylor and Ivry 2012). Error-based learning is likely to make
874 minimal contribution to sequence learning; here, learning entails the formation of associations
875 between successive stimulus-response pairs, supporting the establishment of predictions from

one S-R element to the next. This process, at least when arising implicitly, may be more of a Hebbian-like process (Lu et al. 1998; Nixon and Passingham 2000; Spencer and Ivry 2009). As such, adaptation involves changes in terms of refining execution whereas sequence learning is more about priming for action selection.

Second, the linkage of both tasks through their association with the cerebellum is, at best, superficial. The cerebellum is a large structure and although it has been associated with both adaptation and sequence learning, one need not assume that these two task domains engage similar regions and/or cerebellar operations. Indeed, the evidence suggests that adaptation and sequence learning tasks may engage distinct cerebellar regions and cortico-cerebellar loops (Kelly and Strick 2003; Krienen and Buckner 2009; Strick et al. 2009; Buckner et al. 2011). Neuropsychological and neuroimaging evidence point to the involvement of motor regions of the cerebellum in adaptation tasks, including cerebellar regions with relatively clear somatotopy. These regions show strong functional connectivity with motor cortex (Bernard and Seidler 2013; Burciu et al. 2014), as was even found during sensorimotor adaptation tasks (Rabe et al., 2009; Burciu et al., 2014). Sequence learning, on the other hand, has been linked with neocerebellar regions that are functionally connected with premotor and prefrontal cortex (Nixon and Passingham 2000; Ramnani et al. 2000; Seidler et al. 2002; Bernard and Seidler 2013; although it remains unclear if this holds for both explicit and implicit conditions, see Hazeltine et al., 1997). Moreover, it remains unclear if the cerebellar contribution to sequence learning is directly related to learning per se, or is indirect, related to the formation and/or maintenance of the stimulus-response associations (Nixon and Passingham 2000; Spencer and Ivry 2009).

It will be interesting in future work to exploit individual differences in comparing cerebellar contributions to adaptation and sequence learning. For example, we might expect that functional connectivity between cerebellar regions and primary motor cortex will be predictive of learning rates in sensorimotor adaptation (Bernard et al. 2012; although see Rabe et al., 2009), whereas functional connectivity between cerebellar regions and premotor/prefrontal cortex will be predictive of sequence learning.

We recognize that our assays of learning for visuomotor adaptation and sequence learning involve very different measures (spatial accuracy vs. reaction time). There are cases where performance measured in different units and on different scales do correlate (e.g., mathematical skill, verbal fluency, and spatial visualization, among others, composing a

construct of general intelligence, *g*) (Spearman 1904). Thus the use of different measures does not inherently impose a bias against observing a relationship between the two domains of implicit learning. Nonetheless, it would be interesting to compare tasks of implicit learning in which more similar measures were employed, for example, by using adaptation to a temporal delay.

Within-task behavioral predictors of implicit learning

The preceding discussion indicates that there is little commonality between processes involved in implicit adaptation and implicit sequence learning. A different question concerns predictors of each of these forms of learning. To address this question, we focused on baseline measures of performance, using data obtained prior to the onset of the perturbation in the VMA task or during the early stages of the sequence tasks when there was little evidence of learning. We looked at two features of baseline performance: speed and variability.

As a measure of performance speed, we used MT in the VMA task and RT in the sequence-learning task, the temporal measures that were emphasized in the instructions used for each task. There was no relationship between MT and the rate of adaptation. In contrast, RT was correlated with sequence learning. However, the direction of the correlation was in opposite directions for the two sequence learning tasks: For SRT, the correlation was negative, with faster responders exhibiting the largest amount of learning; for ASRT, the correlation was positive with the slower responders exhibiting the largest amount of learning.

The ASRT task is consistent with other work showing that the benefits of associative retrieval processes become greater as RT increases (Chenery et al. 1994). A similar idea could account for the positive correlation between RT and learning in the ASRT task. However, this hypothesis would also predict a similar pattern in the SRT task, whereas, here, the correlation was reversed. It may be that the reversal is related to the concern discussed previously that some participants may have developed a degree of awareness in the SRT task. Awareness would not only lead to the largest cost on the random blocks (high magnitude of learning), but should also lead to fast RTs (i.e. expecting and even predicating the next stimulus-response pair). We recognize that our interpretations of these RT relationships are speculative and require further study.

Next, we considered how performance variability was related to individual differences in learning. For variability, we used a measure of spatial variability for the VMA task, using the standard deviation of the heading angle during the baseline block (no perturbation). For the sequence learning tasks, we used the standard deviation of the RTs in the early blocks. Here we observed reliable correlations between measures of variability and learning on both the VMA and ASRT tasks. The rate of adaptation was negatively correlated with variability, such that faster learning was associated with more consistent reaches. In contrast, the magnitude of sequence learning was positively correlated with variability, such that the amount of learning was greatest for participants who were most variable in reaction time. A similar pattern was also observed in the SRT task. We recognize that there are differences in our measures of variability for the two task domains; in particular, one is a spatial measure and the other a temporal measure. Nonetheless, we chose these variability measures since they are closest to the dependent variable related to learning on each task.

The negative relationship between variability and learning rate on the VMA task is consistent with models in which motor output or sensory noise modulates the weight given to an error signal (Baddeley et al. 2003; Körding and Wolpert 2004; Burge et al. 2008; van Beers 2009; Schlerf et al. 2013). In conditions of high noise, these models would predict that the motor system would reduce its learning rate because confidence (or certainty) in the sensory signal is low; the converse is true in conditions of low noise. We would expect this relationship to be especially pronounced in the current study given our use of a small perturbation, one in which learning was entirely implicit. Indeed, the size of the perturbation change from trial-to-trial never exceeded 2°, falling within one standard deviation of the expected distribution of endpoints for reaches in the absence of any perturbation.

An error-based account does not seem appropriate when considering sequence learning, since the fundamental learning mechanism here entails an associative process, one that learns to predict successive elements in a chain of events. As such, the absence of a negative correlation between our performance measure of variability, the standard deviation of RT, and sequence learning, is not surprising. However, the reversal to a positive correlation is intriguing. Wu and colleagues (Wu et al. 2014) observed a similar relationship, albeit in measures of spatial variability, in a task in which participants had to learn a complex trajectory or force-field. They hypothesized that greater variability enabled increased exploration of the workspace, and thus

968 allowed participants to discover the appropriate solution. An extension of the exploration idea to
969 account for the positive correlation in the sequence learning tasks in the current study is not
970 straightforward: It is not obvious how greater temporal variability promotes, or reflects
971 exploration. Perhaps temporal variability is indicative of a variable retrieval process and this
972 variability helps with the formation of sequential associations. Related to this idea, temporal
973 variability was correlated with mean RT in the ASRT task (even after normalizing the temporal
974 variability by mean RT) and, as described above, longer (and more variable RTs) may have
975 allowed greater spreading activation from one element to the next. Our sequence learning data
976 suggest that motor variability should not be considered merely the inevitable consequence of
977 signal-dependent neural noise in the motor system that should be overcome, but can, instead, be
978 viewed as a key ingredient of learning, centrally driven and actively regulated, which the motor
979 system leverages during learning (Wu & Smith, 2014).

982 **Acknowledgments:**

983 The authors would like to thank Princess Lim for help with data collection. This work was
984 supported by NIH grants NS092079 and NS074917 to R.B.I, and NS084948 to J.A.T.

985

References

- Bachman JC.** Specificity vs. Generality in Learning and Performing Two Large Muscle Motor Tasks. *Research Quarterly* 32:3-11, 1961.
- Baddeley RJ, Ingram HA, Miall RC.** System identification applied to a visuomotor task: near-optimal human performance in a noisy changing task. *J Neurosci Off J Soc Neurosci* 23: 3066–3075, 2003.
- van Beers RJ.** Motor learning is optimally tuned to the properties of motor noise. *Neuron* 63: 406–417, 2009.
- Bernard JA, Seidler RD.** Cerebellar contributions to visuomotor adaptation and motor sequence learning: an ALE meta-analysis. *Front Hum Neurosci* 7: 27, 2013.
- Bernard JA, Seidler RD, Hassevoort KM, Benson BL, Welsh RC, Wiggins JL, Jaeggi SM, Buschkuehl M, Monk CS, Jonides J, Peltier SJ.** Resting state cortico-cerebellar functional connectivity networks: a comparison of anatomical and self-organizing map approaches. *Front Neuroanat* 6: 31, 2012.
- Bo J, Jennett S, Seidler RD.** Working memory capacity correlates with implicit serial reaction time task performance. *Exp Brain Res* 214: 73–81, 2011.
- Buckner RL, Krienen FM, Castellanos A, Diaz JC, Yeo BTT.** The organization of the human cerebellum estimated by intrinsic functional connectivity. *J Neurophysiol* 106: 2322–2345, 2011.
- Burciu RG, Reinold J, Rabe K, Wondzinski E, Siebler M, Müller O, Theysohn N, Gerwig M, Donchin O, Timmann D.** Structural correlates of motor adaptation deficits in patients with acute focal lesions of the cerebellum. *Exp Brain Res* 232: 2847–2857, 2014.
- Burge J, Ernst MO, Banks MS.** The statistical determinants of adaptation rate in human reaching. *J Vis* 8: 20.1-19, 2008.
- Chenery HJ, Ingram JC, Murdoch BE.** The effect of repeated prime-target presentation in manipulating attention-induced priming in persons with dementia of the Alzheimer's type. *Brain Cogn* 25: 108–127, 1994.
- Cheng S, Sabes PN.** Calibration of visually guided reaching is driven by error-corrective learning and internal dynamics. *J Neurophysiol* 97: 3057–3069, 2007.
- Conway CM, Bauernschmidt A, Huang SS, Pisoni DB.** Implicit statistical learning in language processing: word predictability is the key. *Cognition* 114: 356–371, 2010.
- Cornbleet PJ, Gochman N.** Incorrect least-squares regression coefficients in method-comparison analysis. *Clin Chem* 25: 432–438, 1979.
- Criscimagna-Hemminger SE, Bastian AJ, Shadmehr R.** Size of error affects cerebellar contributions to motor learning. *J Neurophysiol* 103: 2275–2284, 2010.

- 1021 **Danckert J, Ferber S, Goodale MA.** Direct effects of prismatic lenses on visuomotor control:
1022 an event-related functional MRI study. *Eur J Neurosci* 28: 1696–1704, 2008.
- 1023 **Della-Maggiore V, McIntosh AR.** Time course of changes in brain activity and functional
1024 connectivity associated with long-term adaptation to a rotational transformation. *J Neurophysiol*
1025 93: 2254–2262, 2005.
- 1026 **Diedrichsen J, Hashambhoy Y, Rane T, Shadmehr R.** Neural correlates of reach errors. *J*
1027 *Neurosci Off J Soc Neurosci* 25: 9919–9931, 2005.
- 1028 **Donchin O, Rabe K, Diedrichsen J, Lally N, Schoch B, Gizewski ER, Timmann D.**
1029 Cerebellar regions involved in adaptation to force field and visuomotor perturbation. *J*
1030 *Neurophysiol* 107: 134–147, 2012.
- 1031 **Doyon J, Korman M, Morin A, Dostie V, Hadj Tahar A, Benali H, Karni A, Ungerleider**
1032 **LG, Carrier J.** Contribution of night and day sleep vs. simple passage of time to the
1033 consolidation of motor sequence and visuomotor adaptation learning. *Exp Brain Res* 195: 15–26,
1034 2009.
- 1035 **Faisal AA, Selen LPJ, Wolpert DM.** Noise in the nervous system. *Nat Rev Neurosci* 9: 292–
1036 303, 2008.
- 1037 **Ferrucci R, Brunoni AR, Parazzini M, Vergari M, Rossi E, Fumagalli M, Mameli F, Rosa**
1038 **M, Giannicola G, Zago S, Priori A.** Modulating human procedural learning by cerebellar
1039 transcranial direct current stimulation. *Cerebellum Lond Engl* 12: 485–492, 2013.
- 1040 **Galea JM, Vazquez A, Pasricha N, de Xivry J-JO, Celnik P.** Dissociating the roles of the
1041 cerebellum and motor cortex during adaptive learning: the motor cortex retains what the
1042 cerebellum learns. *Cereb Cortex N Y N 1991* 21: 1761–1770, 2011.
- 1043 **Gheysen F, Gevers W, De Schutter E, Van Waelvelde H, Fias W.** Disentangling perceptual
1044 from motor implicit sequence learning with a serial color-matching task. *Exp Brain Res* 197:
1045 163–174, 2009.
- 1046 **Gómez-Beldarrain M, García-Moncó JC, Rubio B, Pascual-Leone A.** Effect of focal
1047 cerebellar lesions on procedural learning in the serial reaction time task. *Exp Brain Res* 120: 25–
1048 30, 1998.
- 1049 **Grafton ST, Hazeltine E, Ivry RB.** Abstract and effector-specific representations of motor
1050 sequences identified with PET. *J Neurosci Off J Soc Neurosci* 18: 9420–9428, 1998.
- 1051 **Hazeltine E, Grafton ST, Ivry R.** Attention and stimulus characteristics determine the locus of
1052 motor-sequence encoding. A PET study. *Brain J Neurol* 120 (Pt 1): 123–140, 1997.
- 1053 **Howard JH, Howard DV.** Age differences in implicit learning of higher order dependencies in
1054 serial patterns. *Psychol Aging* 12: 634–656, 1997.
- 1055 **Ivry RB, Keele SW.** Timing functions of the cerebellum. *J Cogn Neurosci* 1: 136–152, 1989.

- 1056 **Janacsek K, Fiser J, Nemeth D.** The best time to acquire new skills: age-related differences in
1057 implicit sequence learning across the human lifespan. *Dev Sci* 15: 496–505, 2012.
- 1058 **Janacsek K, Nemeth D.** Predicting the future: from implicit learning to consolidation. *Int J*
1059 *Psychophysiol Off J Int Organ Psychophysiol* 83: 213–221, 2012.
- 1060 **Kagerer FA, Contreras-Vidal JL, Stelmach GE.** Adaptation to gradual as compared with
1061 sudden visuo-motor distortions. *Exp Brain Res* 115: 557–561, 1997.
- 1062 **Kanai R, Rees G.** The structural basis of inter-individual differences in human behaviour and
1063 cognition. *Nat Rev Neurosci* 12: 231–242, 2011.
- 1064 **Keele SW, Hawkins HL.** Explorations of individual differences relevant to high level skill. *J*
1065 *Mot Behav* 14: 3–23, 1982.
- 1066 **Keele SW, Ivry R, Mayr U, Hazeltine E, Heuer H.** The cognitive and neural architecture of
1067 sequence representation. *Psychol Rev* 110: 316–339, 2003.
- 1068 **Keele SW, Ivry RI, Pokorny RA.** Force control and its relation to timing. *J Mot Behav* 19: 96–
1069 114, 1987.
- 1070 **Keele SW, Pokorny RA, Corcos DM, Ivry R.** Do perception and motor production share
1071 common timing mechanisms: a correctional analysis. *Acta Psychol (Amst)* 60: 173–191, 1985.
- 1072 **Kelly RM, Strick PL.** Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman
1073 primate. *J Neurosci Off J Soc Neurosci* 23: 8432–8444, 2003.
- 1074 **Kinder A, Rolfs M, Kliegl R.** Sequence learning at optimal stimulus-response mapping:
1075 evidence from a serial reaction time task. *Q J Exp Psychol* 2006 61: 203–209, 2008.
- 1076 **King AC, Ranganathan R, Newell KM.** Individual differences in the exploration of a
1077 redundant space-time motor task. *Neurosci Lett* 529: 144–149, 2012.
- 1078 **Kording KP, Tenenbaum JB, Shadmehr R.** The dynamics of memory as a consequence of
1079 optimal adaptation to a changing body. *Nat Neurosci* 10: 779–786, 2007.
- 1080 **Körding KP, Wolpert DM.** Bayesian integration in sensorimotor learning. *Nature* 427: 244–
1081 247, 2004.
- 1082 **Krakauer JW.** Motor learning and consolidation: the case of visuomotor rotation. *Adv Exp Med*
1083 *Biol* 629: 405–421, 2009.
- 1084 **Krienen FM, Buckner RL.** Segregated fronto-cerebellar circuits revealed by intrinsic functional
1085 connectivity. *Cereb Cortex N Y N 1991* 19: 2485–2497, 2009.
- 1086 **Linnet K.** Evaluation of regression procedures for methods comparison studies. *Clin Chem* 39:
1087 424–432, 1993.

- 1088 **Linnet K.** Performance of Deming regression analysis in case of misspecified analytical error
1089 ratio in method comparison studies. *Clin Chem* 44: 1024–1031, 1998.
- 1090 **Lu X, Hikosaka O, Miyachi S.** Role of monkey cerebellar nuclei in skill for sequential
1091 movement. *J Neurophysiol* 79: 2245–2254, 1998.
- 1092 **Martin TA, Keating JG, Goodkin HP, Bastian AJ, Thach WT.** Throwing while looking
1093 through prisms. I. Focal olivocerebellar lesions impair adaptation. *Brain J Neurol* 119 (Pt 4):
1094 1183–1198, 1996.
- 1095 **Martini M, Furtner MR, Sachse P.** Working memory and its relation to deterministic sequence
1096 learning. *PloS One* 8: e56166, 2013.
- 1097 **Meier B, Cock J.** Offline consolidation in implicit sequence learning. *Cortex J Devoted Study*
1098 *Nerv Syst Behav* 57: 156–166, 2014.
- 1099 **Michel C, Pisella L, Prablanc C, Rode G, Rossetti Y.** Enhancing visuomotor adaptation by
1100 reducing error signals: single-step (aware) versus multiple-step (unaware) exposure to wedge
1101 prisms. *J Cogn Neurosci* 19: 341–350, 2007.
- 1102 **Molinari M, Leggio MG, Solida A, Ciorra R, Misciagna S, Silveri MC, Petrosini L.**
1103 Cerebellum and procedural learning: evidence from focal cerebellar lesions. *Brain J Neurol* 120 (
1104 Pt 10): 1753–1762, 1997.
- 1105 **Morehead JR, Qasim SE, Crossley MJ, Ivry R.** Savings upon Re-Aiming in Visuomotor
1106 Adaptation. *J Neurosci Off J Soc Neurosci* 35: 14386–14396, 2015.
- 1107 **Nemeth D, Janacsek K, Fiser J.** Age-dependent and coordinated shift in performance between
1108 implicit and explicit skill learning. *Front Comput Neurosci* 7: 147, 2013.
- 1109 **Nemeth D, Janacsek K, Londe Z, Ullman MT, Howard DV, Howard JH.** Sleep has no
1110 critical role in implicit motor sequence learning in young and old adults. *Exp Brain Res* 201:
1111 351–358, 2010.
- 1112 **Nixon PD, Passingham RE.** The cerebellum and cognition: cerebellar lesions impair sequence
1113 learning but not conditional visuomotor learning in monkeys. *Neuropsychologia* 38: 1054–1072,
1114 2000.
- 1115 **Noohi F, Boyden NB, Kwak Y, Humfleet J, Burke DT, Müller MLTM, Bohnen NI, Seidler**
1116 **RD.** Association of COMT val158met and DRD2 G>T genetic polymorphisms with individual
1117 differences in motor learning and performance in female young adults. *J Neurophysiol* 111: 628–
1118 640, 2014.
- 1119 **Pascual-Leone A, Grafman J, Clark K, Stewart M, Massaquoi S, Lou JS, Hallett M.**
1120 Procedural learning in Parkinson’s disease and cerebellar degeneration. *Ann Neurol* 34: 594–602,
1121 1993.

- 1122 **Perruchet P, Amorim MA.** Conscious knowledge and changes in performance in sequence
1123 learning: evidence against dissociation. *J Exp Psychol Learn Mem Cogn* 18: 785–800, 1992.
- 1124 **Rabe K, Livne O, Gizewski ER, Aurich V, Beck A, Timmann D, Donchin O.** Adaptation to
1125 visuomotor rotation and force field perturbation is correlated to different brain areas in patients
1126 with cerebellar degeneration. *J Neurophysiol* 101: 1961–1971, 2009.
- 1127 **Ramnani N, Toni I, Josephs O, Ashburner J, Passingham RE.** Learning- and expectation-
1128 related changes in the human brain during motor learning. *J Neurophysiol* 84: 3026–3035, 2000.
- 1129 **Reber PJ, Squire LR.** Parallel brain systems for learning with and without awareness. *Learn*
1130 *Mem Cold Spring Harb N* 1: 217–229, 1994.
- 1131 **Reber PJ, Squire LR.** Encapsulation of implicit and explicit memory in sequence learning. *J*
1132 *Cogn Neurosci* 10: 248–263, 1998.
- 1133 **Remillard G.** Implicit learning of second-, third-, and fourth-order adjacent and nonadjacent
1134 sequential dependencies. *Q J Exp Psychol* 2006 61: 400–424, 2008.
- 1135 **Robertson EM.** The Serial Reaction Time Task: Implicit Motor Skill Learning? *J Neurosci* 27:
1136 10073–10075, 2007.
- 1137 **Rüsseler J, Rösler F.** Implicit and explicit learning of event sequences: evidence for distinct
1138 coding of perceptual and motor representations. *Acta Psychol (Amst)* 104: 45–67, 2000.
- 1139 **Sanchez DJ, Gobel EW, Reber PJ.** Performing the unexplainable: implicit task performance
1140 reveals individually reliable sequence learning without explicit knowledge. *Psychon Bull Rev* 17:
1141 790–796, 2010.
- 1142 **Sanchez DJ, Reber PJ.** Operating characteristics of the implicit learning system supporting
1143 serial interception sequence learning. *J Exp Psychol Hum Percept Perform* 38: 439–452, 2012.
- 1144 **Schlerf JE, Xu J, Klemfuss NM, Griffiths TL, Ivry RB.** Individuals with cerebellar
1145 degeneration show similar adaptation deficits with large and small visuomotor errors. *J*
1146 *Neurophysiol* 109: 1164–1173, 2013.
- 1147 **Seashore CE, Tiffin J.** AN OBJECTIVE METHOD OF EVALUATING MUSICAL
1148 PERFORMANCE. *Science* 72: 480–482, 1930.
- 1149 **Seidler RD.** Differential effects of age on sequence learning and sensorimotor adaptation. *Brain*
1150 *Res Bull* 70: 337–346, 2006.
- 1151 **Seidler RD, Bo J, Anguera JA.** Neurocognitive contributions to motor skill learning: the role of
1152 working memory. *J Mot Behav* 44: 445–453, 2012.
- 1153 **Seidler RD, Purushotham A, Kim S-G, Uğurbil K, Willingham D, Ashe J.** Cerebellum
1154 activation associated with performance change but not motor learning. *Science* 296: 2043–2046,
1155 2002.

- 1156 **Shadmehr R, Smith MA, Krakauer JW.** Error correction, sensory prediction, and adaptation
1157 in motor control. *Annu Rev Neurosci* 33: 89–108, 2010.
- 1158 **Shanks DR, Rowland LA, Ranger MS.** Attentional load and implicit sequence learning.
1159 *Psychol Res* 69: 369–382, 2005.
- 1160 **Shin JC, Ivry RB.** Spatial and temporal sequence learning in patients with Parkinson’s disease
1161 or cerebellar lesions. *J Cogn Neurosci* 15: 1232–1243, 2003.
- 1162 **Smith MA, Shadmehr R.** Intact ability to learn internal models of arm dynamics in
1163 Huntington’s disease but not cerebellar degeneration. *J Neurophysiol* 93: 2809–2821, 2005.
- 1164 **Spearman C.** “General Intelligence,” Objectively Determined and Measured. *Am J Psychol* 15:
1165 201–292, 1904.
- 1166 **Spencer RMC, Ivry RB.** Sequence learning is preserved in individuals with cerebellar
1167 degeneration when the movements are directly cued. *J Cogn Neurosci* 21: 1302–1310, 2009.
- 1168 **Spencer RMC, Sunm M, Ivry RB.** Sleep-dependent consolidation of contextual learning. *Curr*
1169 *Biol CB* 16: 1001–1005, 2006.
- 1170 **Squire LR, Zola SM.** Structure and function of declarative and nondeclarative memory systems.
1171 *Proc Natl Acad Sci U S A* 93: 13515–13522, 1996.
- 1172 **Strick PL, Dum RP, Fiez JA.** Cerebellum and nonmotor function. *Annu Rev Neurosci* 32: 413–
1173 434, 2009.
- 1174 **Sun R, Zhang X, Slusarz P, Mathews R.** The interaction of implicit learning, explicit
1175 hypothesis testing learning and implicit-to-explicit knowledge extraction. *Neural Netw Off J Int*
1176 *Neural Netw Soc* 20: 34–47, 2007.
- 1177 **Taylor JA, Ivry RB.** The role of strategies in motor learning. *Ann N Y Acad Sci* 1251: 1–12,
1178 2012.
- 1179 **Taylor JA, Klemfuss NM, Ivry RB.** An explicit strategy prevails when the cerebellum fails to
1180 compute movement errors. *Cerebellum Lond Engl* 9: 580–586, 2010.
- 1181 **Thoroughman KA, Shadmehr R.** Learning of action through adaptive combination of motor
1182 primitives. *Nature* 407: 742–747, 2000.
- 1183 **Tseng Y-W, Diedrichsen J, Krakauer JW, Shadmehr R, Bastian AJ.** Sensory prediction
1184 errors drive cerebellum-dependent adaptation of reaching. *J Neurophysiol* 98: 54–62, 2007.
- 1185 **Unsworth N, Engle RW.** Individual differences in working memory capacity and learning:
1186 evidence from the serial reaction time task. *Mem Cognit* 33: 213–220, 2005.
- 1187 **Wei K, Körding K.** Uncertainty of feedback and state estimation determines the speed of motor
1188 adaptation. *Front Comput Neurosci* 4: 11, 2010.

1189 **Werner S, Bock O, Gizewski ER, Schoch B, Timmann D.** Visuomotor adaptive improvement
1190 and aftereffects are impaired differentially following cerebellar lesions in SCA and PICA
1191 territory. *Exp Brain Res* 201: 429–439, 2010.

1192 **Wilkinson L, Shanks DR.** Intentional control and implicit sequence learning. *J Exp Psychol*
1193 *Learn Mem Cogn* 30: 354–369, 2004.

1194 **Willingham DB, Nissen MJ, Bullemer P.** On the development of procedural knowledge. *J Exp*
1195 *Psychol Learn Mem Cogn* 15: 1047–1060, 1989.

1196 **Wu HG, Miyamoto YR, Gonzalez Castro LN, Ölveczky BP, Smith MA.** Temporal structure
1197 of motor variability is dynamically regulated and predicts motor learning ability. *Nat Neurosci*
1198 17: 312–321, 2014.

1199 **Zarahn E, Weston GD, Liang J, Mazzoni P, Krakauer JW.** Explaining savings for
1200 visuomotor adaptation: linear time-invariant state-space models are not sufficient. *J*
1201 *Neurophysiol* 100: 2537–2548, 2008.

1202

1203

Figure Captions

Figure 1. Experimental tasks. **A. Visuomotor adaptation (VMA) task.** The perturbation followed a pseudorandom walk of a global sinusoidal pattern. **B. Serial Reaction Time (SRT) task.** On Random blocks (R), the stimulus positions were selected randomly and on Sequence blocks (S), the stimulus positions followed a repeating 12-element sequence. **C. Alternating Serial Reaction Time (ASRT) task.** Odd numbered elements follow a fixed sequence and even numbered elements are selected at random (r). This creates high and low frequency triplets (see text).

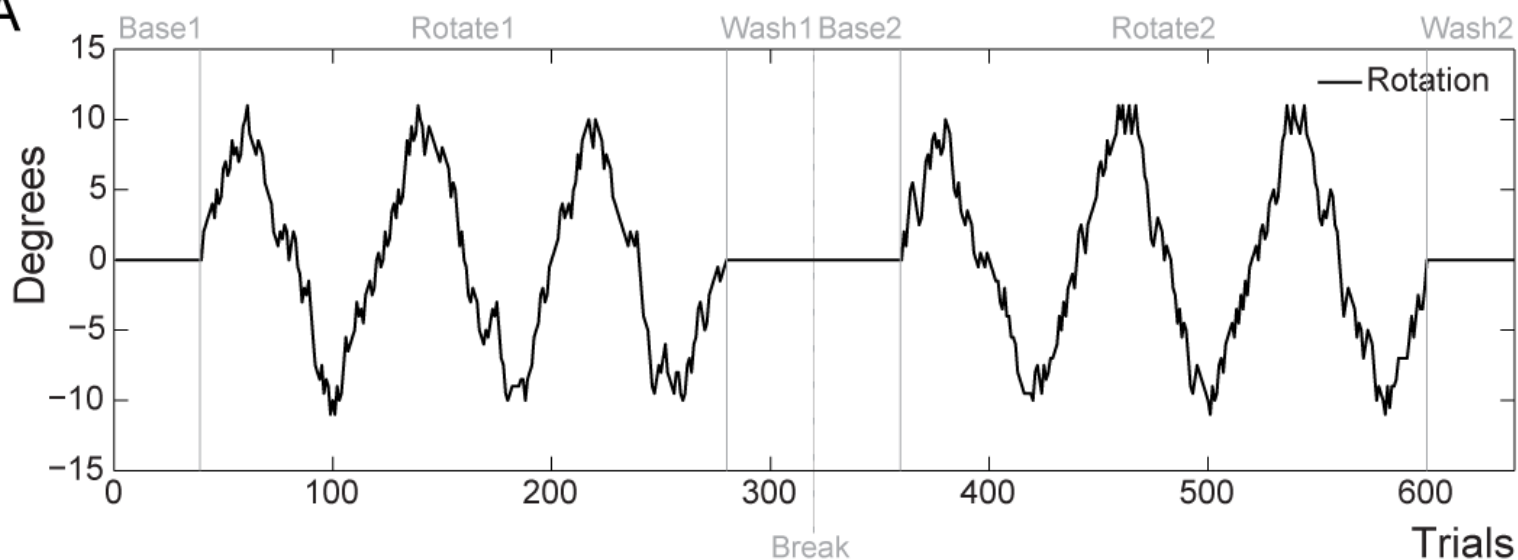
Figure 2. VMA results (n=100). **A.** Group average response (red, shaded region is group SEM) to a gradual perturbation (black). For visualization purposes, the response function has been flipped, although the actual changes in movement heading were in the opposite direction of the perturbation; **B.** Reliability of each movement time, spatial variability, and learning rate between Run1 and Run2 scores (circles correspond to individual participants). Baseline movement time and spatial variability measures are taken from the baseline block. Learning rate is estimated from the model fit of the data from the perturbation blocks. **C.** Correlations between different measures of performance. In all figures, r and p values represent the strength and significance of the Pearson correlation coefficients of the linear dependence between the variables of interest. Orthogonal (Deming) regression lines evaluate the relationship between variables of interest, without making assumptions concerning their dependence or independence.

Figure 3. SRT results (n=53). **A.** Group average of median RT (left) and accuracy (right), for Run1 (blue) and Run2 (red). Shaded areas represent group SEM. **B.** Reliability of reaction time, temporality variability, and learning between the Run1 and Run2 scores. Baseline metrics of RT and the standard deviation of RT are taken from blocks 2-4 (early in training). Learning is calculated from the last four blocks (random blocks 13 and 15 minus sequence blocks 12 and 14). **C.** Correlation between different measures of performance. **D.** Reliability of learning (left) and correlation of learning and baseline metrics of learning at the mid-way probe (random block 7 minus sequence blocks 6 and 8).

Figure 4. ASRT results (n=25). **A.** Group average of median RT (left) for Run1 (blue) and Run2 (red) (left panel), divided into low and high frequency triplets (middle), and as difference scores (right). **B.** Reliability reaction time, temporal variability, and learning between Run1 and Run2 scores. Baseline metrics of RT and the standard deviation of RT are based on data from blocks 2-3 (early in training); learning is averaged over blocks 4-45. **C.** Correlations between different measures of performance.

Figure 5. Between tasks correlations of learning scores. **A.** Correlation between learning measures of visuomotor adaptation SRT using the final probe of SRT learning (left) and mid-way SRT learning (right). **B.** Correlation between learning measures on visuomotor adaptation and ASRT. Note that the positive correlation is largely influenced by the participant who had the fastest rate of adaptation and exhibited the largest amount of sequence learning. When the correlation is re-calculated without this individual, there is no correlation between the learning measures for the two tasks ($r=-0.17$, $p=0.48$). **C.** Histogram of responses on Likert-scale to survey question probing awareness of the perturbation (VMA) or sequence (SRT and ASRT). Low values correspond to low awareness; high values correspond to high awareness. Although all of the scores are towards lower end, there is a rightward shift of the distribution for the SRT task, indicative of higher awareness of the presence of a sequence.

A



B



Sequence A: 241342314213

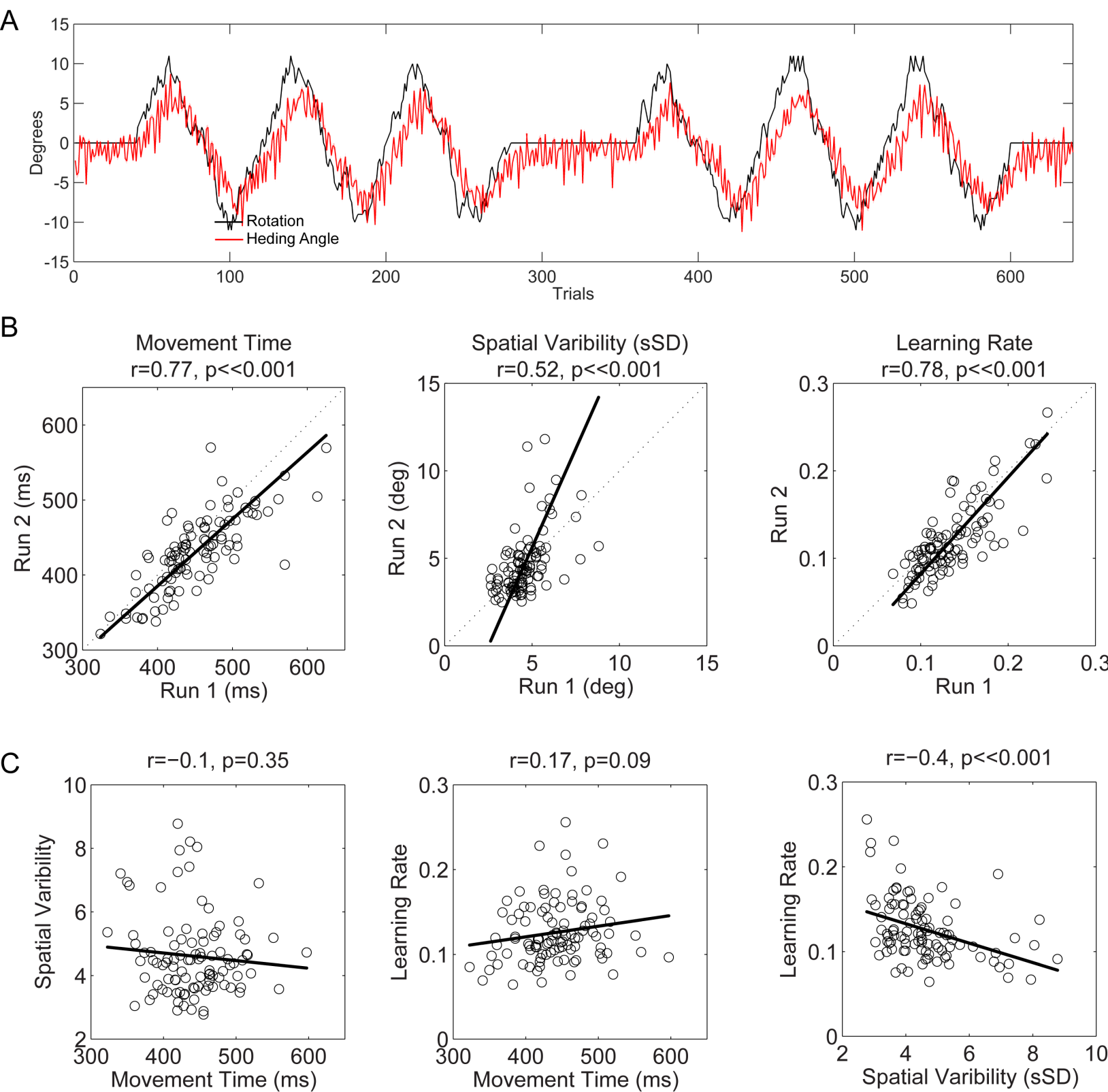
Sequence B: 312413421324

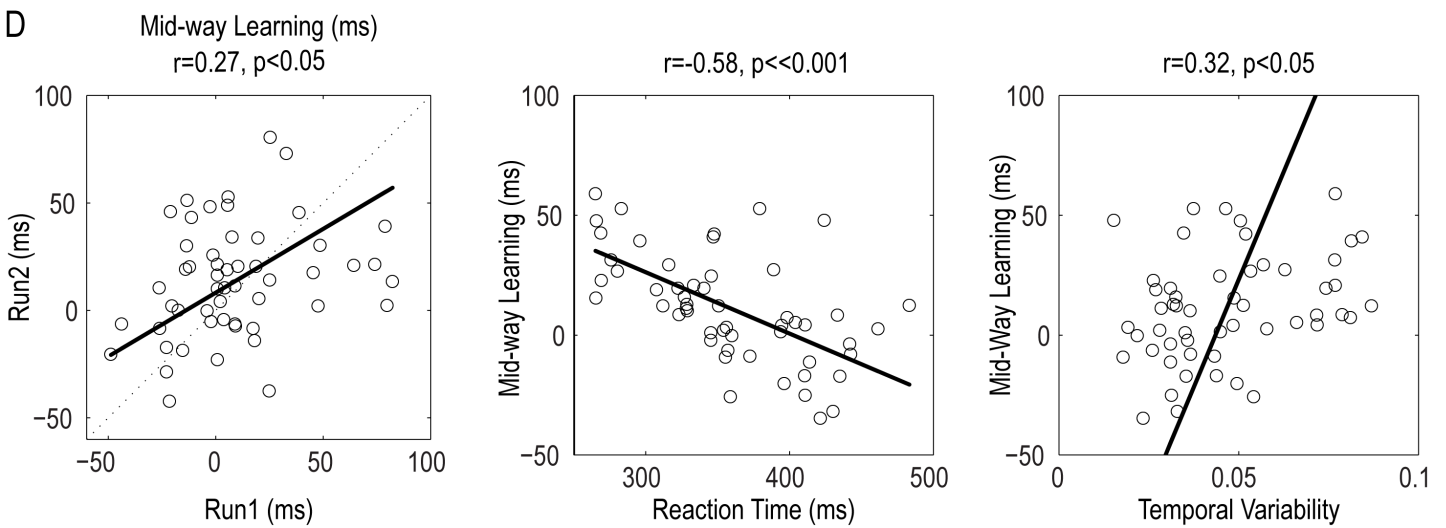
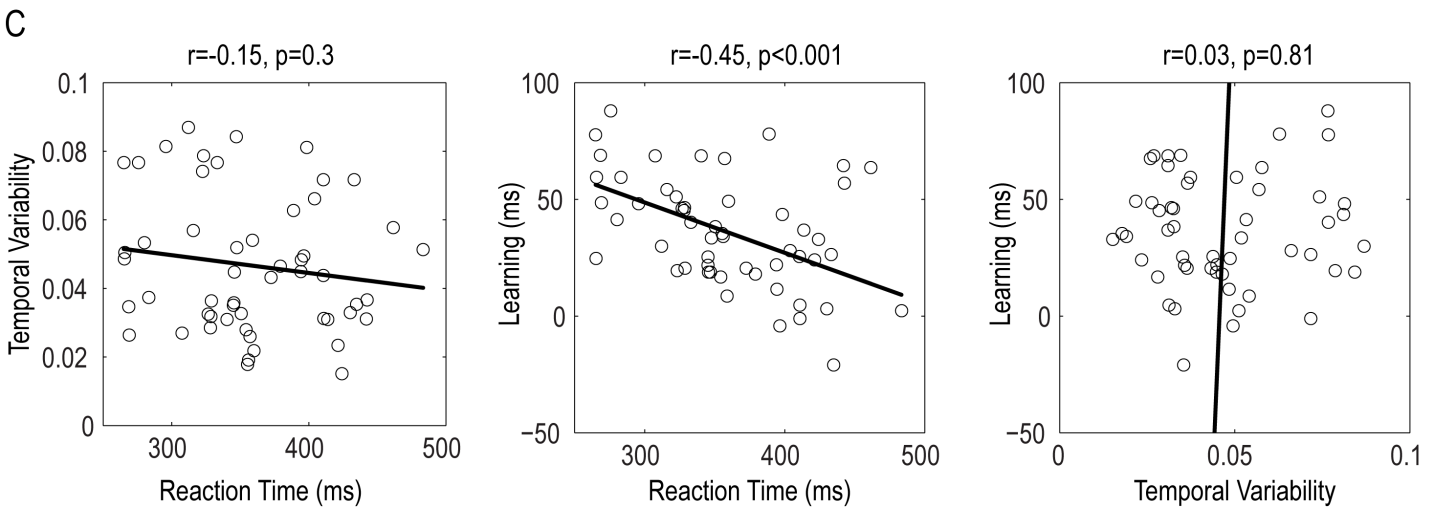
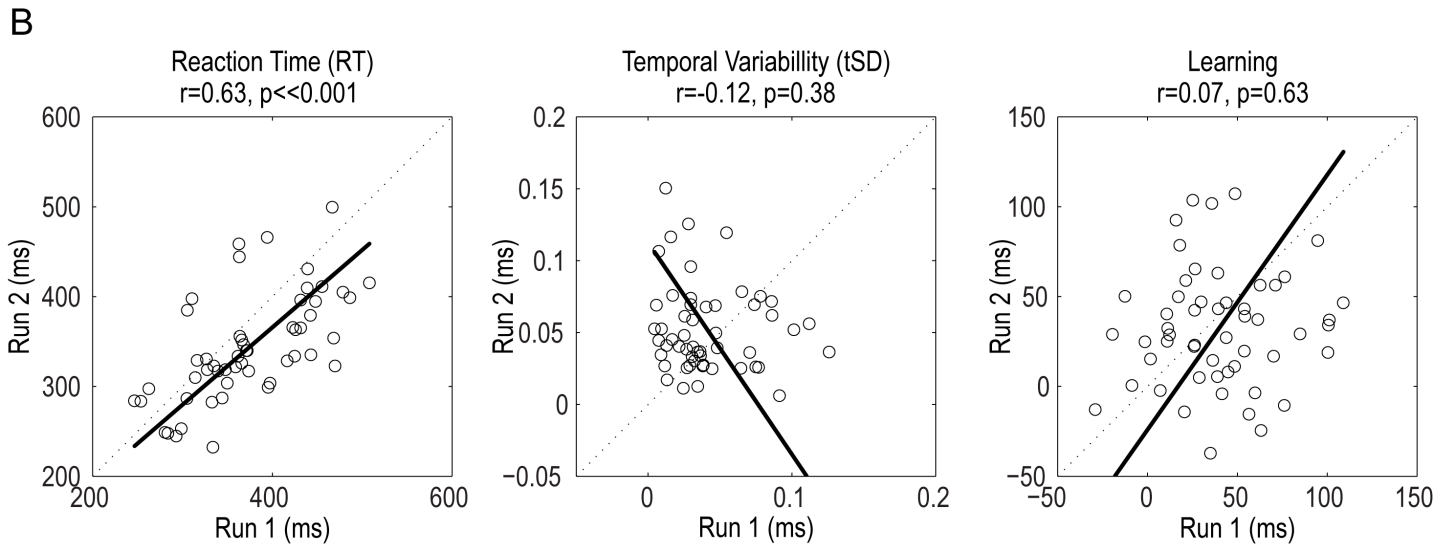
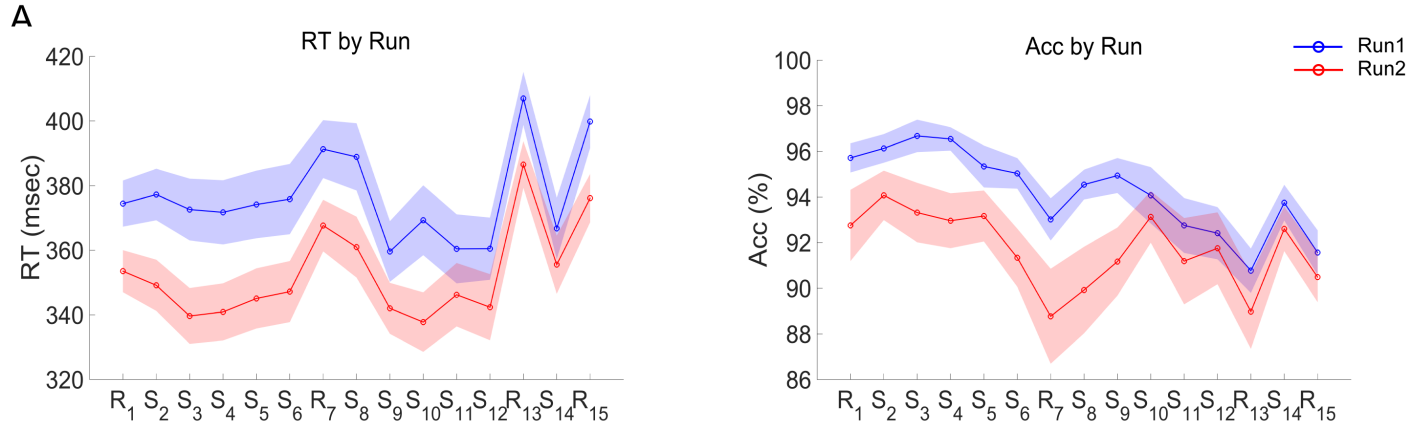
C

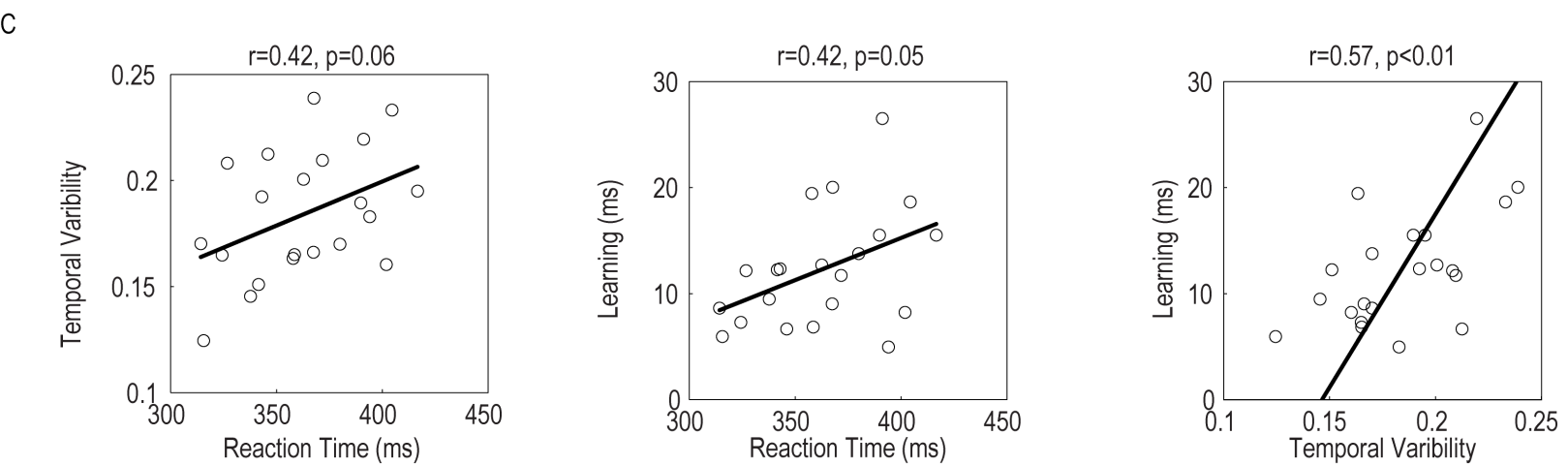
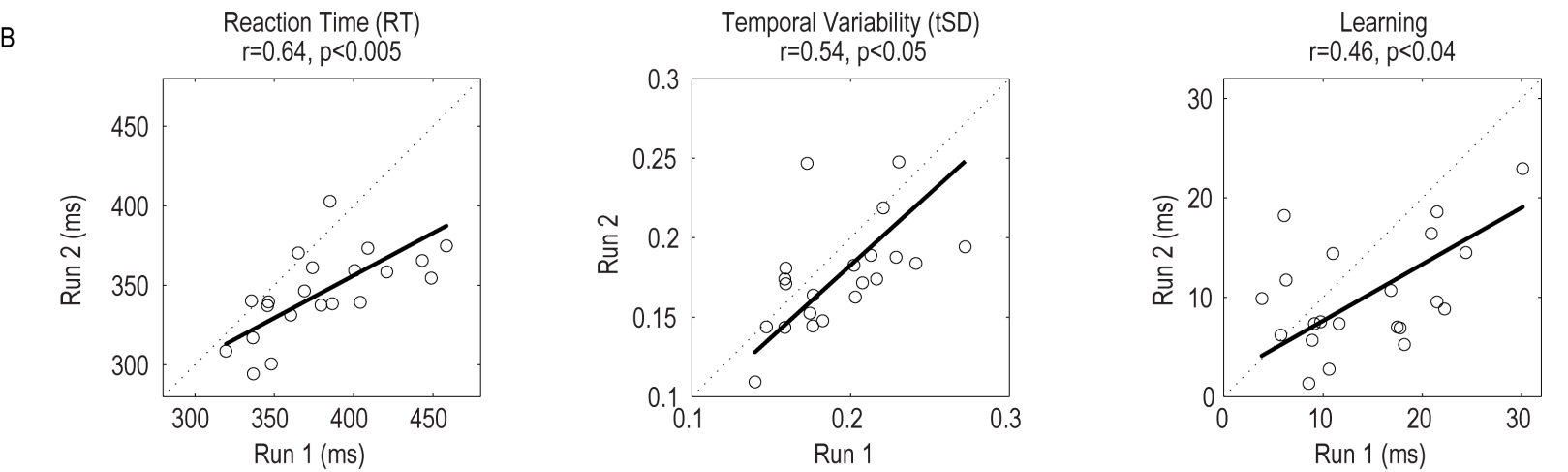
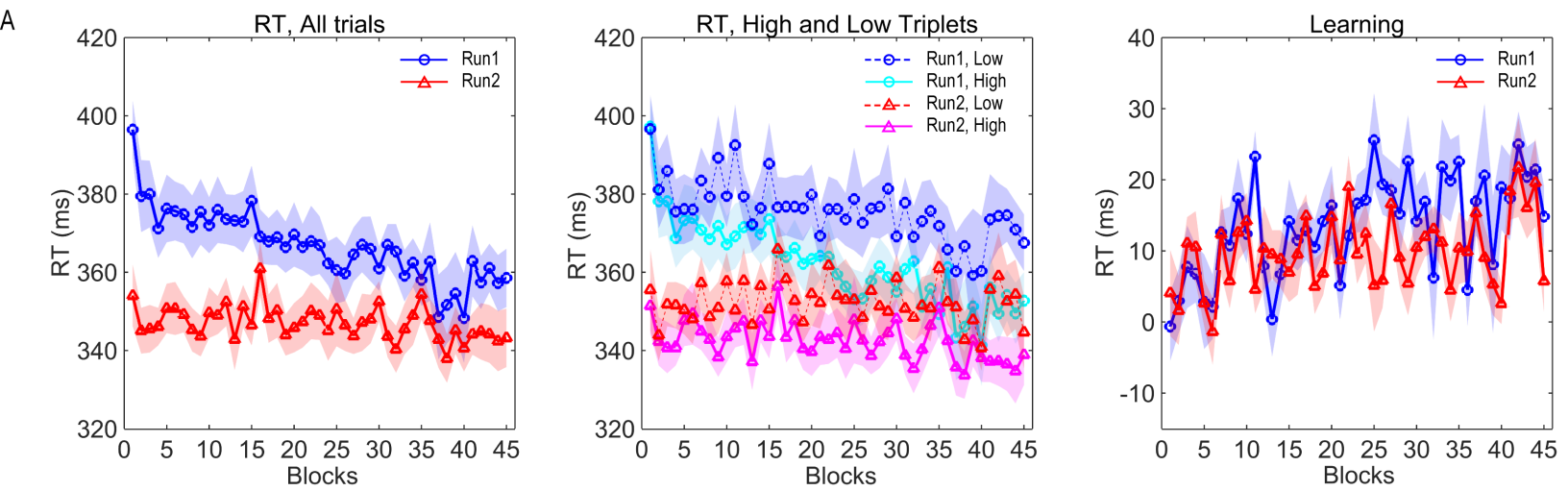
Alternating sequence syntax: 1r2r3r4r...

Triplet	High-Frequency Triplets	Low-Frequency Triplets
3-r-4	e.g. 3-2-4	N/A
r-2-r	e.g. 3-2-4	e.g. 3-2-1

1 2 2 4 **3 2 4** 2 1 **3 2 1** **3 2 4** 2 1 **3 2 1** 3 4 4 2 1 **3 2 4** 3 3 4 2 1 1 2 4 **3 2 4** 2 1 4 2
 4 3 1 4 2 1 4 2 4 **3 2 4** 2 1 **3 2 4** **3 2 4** 2 1 **3 2 4** 3 1 4 2 1 **3 2 1** 3 1 4 2 1 3 2 4 3 1 ...

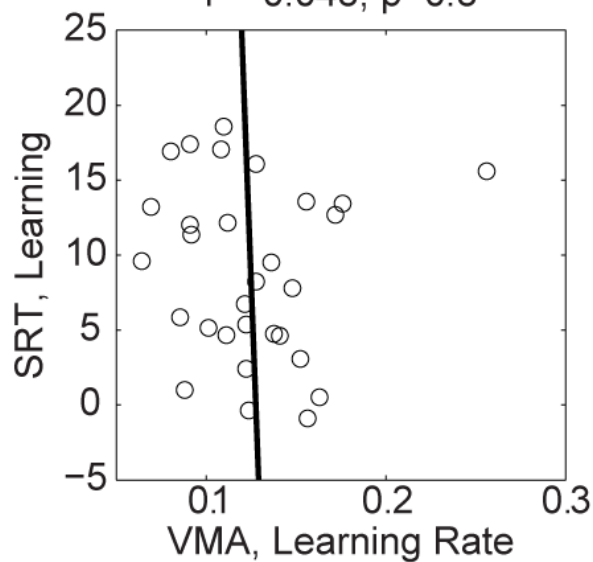




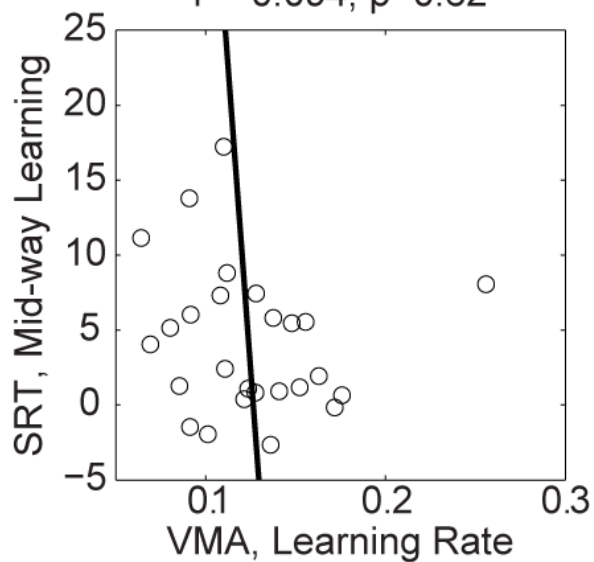


A

VMA-SRT (n=30)
 $r=-0.048$, $p=0.8$

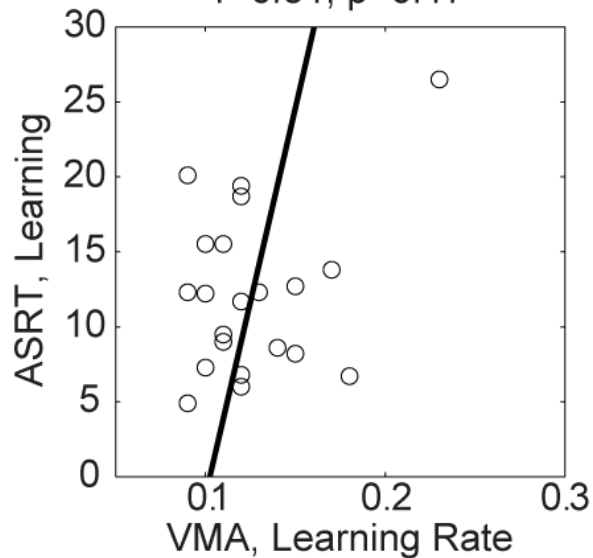


VMA-SRT (n=30)
 $r=-0.094$, $p=0.62$



B

VMA-SRT (n=21)
 $r=0.31$, $p=0.17$



C

Likert Scale

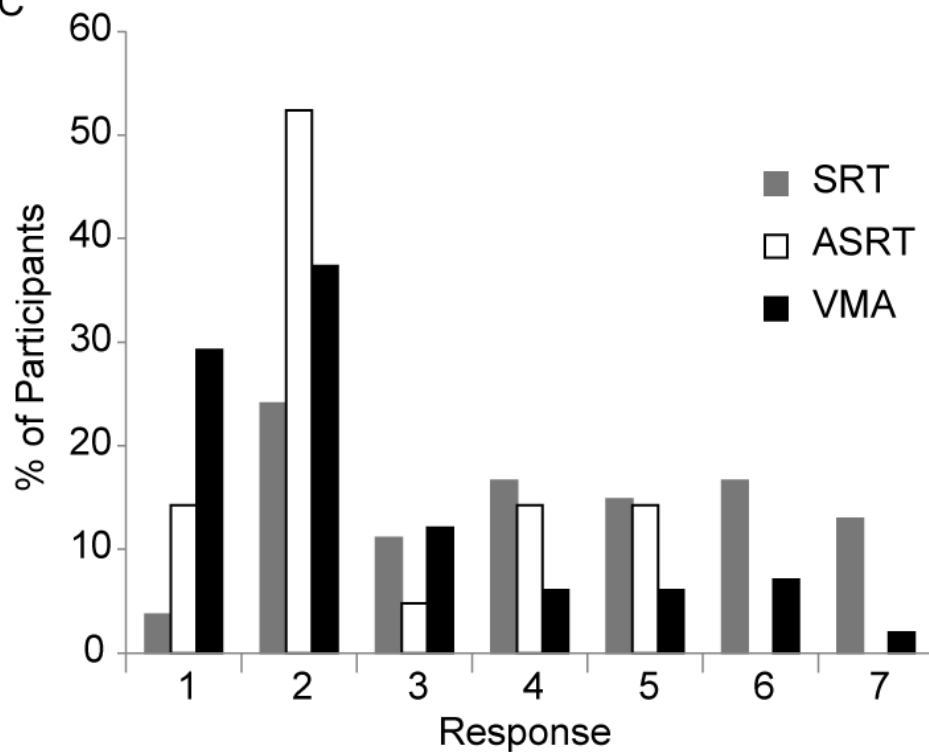


Table 1**Correlation matrices of key performance and learning variables**

Subsets of participants who completed both visuomotor adaptation and sequence learning (*
 $p < 0.05$, ** $p < 0.005$, **** $p < 0.001$)

Table 1A. VMA and SRT

		SRT			VMA		
		tSD	Learning	I. Learning	MT	sSD	LR
SRT	RT	0.02	-0.31*	-0.58***	0.08	0.37*	0.08
	tSD		-0.06	0.28	-0.05	-0.02	0.04
	Learning			0.54**	0	-0.04	-0.06
	Initial Learning				-0.27	-0.22	-0.07
VMA	MT					-0.08	0.21
	sSD						-0.4*

Table 1B. VMA and ASRT

		ASRT		VMA		
		tSD	Learning	MT	sSD	LR
ASRT	RT	0.44*	0.42	0.30	-0.17	0.14
	tSD		0.57**	-0.02	-0.06	0.11
	Learning			-0.04	-0.19	0.31
VMA	MT				-0.20	0.04
	sSD					-0.45*