

RESEARCH ARTICLE | *Control of Movement*

# Somatosensory working memory in human reinforcement-based motor learning

1919

 Ananda Sidarta,<sup>1</sup>  Floris T. van Vugt,<sup>1,2</sup> and David J. Ostry<sup>1,2</sup>

<sup>1</sup>Department of Psychology, McGill University, Montréal, Québec, Canada; and <sup>2</sup>Haskins Laboratories, New Haven, Connecticut

Submitted 2 July 2018; accepted in final form 24 October 2018

**Sidarta A, van Vugt FT, Ostry DJ.** Somatosensory working memory in human reinforcement-based motor learning. *J Neurophysiol* 120: 3275–3286, 2018. First published October 24, 2018; doi:10.1152/jn.00442.2018.—Recent studies using visuomotor adaptation and sequence learning tasks have assessed the involvement of working memory in the visuospatial domain. The capacity to maintain previously performed movements in working memory is perhaps even more important in reinforcement-based learning to repeat accurate movements and avoid mistakes. Using this kind of task in the present work, we tested the relationship between somatosensory working memory and motor learning. The first experiment involved separate memory and motor learning tasks. In the memory task, the participant's arm was displaced in different directions by a robotic arm, and the participant was asked to judge whether a subsequent test direction was one of the previously presented directions. In the motor learning task, participants made reaching movements to a hidden visual target and were provided with positive feedback as reinforcement when the movement ended in the target zone. It was found that participants that had better somatosensory working memory showed greater motor learning. In a second experiment, we designed a new task in which learning and working memory trials were interleaved, allowing us to study participants' memory for movements they performed as part of learning. As in the first experiment, we found that participants with better somatosensory working memory also learned more. Moreover, memory performance for successful movements was better than for movements that failed to reach the target. These results suggest that somatosensory working memory is involved in reinforcement motor learning and that this memory preferentially keeps track of reinforced movements.

**NEW & NOTEWORTHY** The present work examined somatosensory working memory in reinforcement-based motor learning. Working memory performance was reliably correlated with the extent of learning. With the use of a paradigm in which learning and memory trials were interleaved, memory was assessed for movements performed during learning. Movements that received positive feedback were better remembered than movements that did not. Thus working memory does not track all movements equally but is biased to retain movements that were rewarded.

exploration; reinforcement learning; sensory working memory; somatic

## INTRODUCTION

When learning motor skills such as swimming or dance, it is necessary to discover the limb configuration that enables successful movement. In motor tasks such as these, there is limited visual information, and the only performance measure available is success or failure. Learning under these conditions proceeds at least in part through exploration and trial and error. In this model of motor learning, sensory working memory, which enables maintenance and decision making related to prior sensory information, is presumably involved in movement selection by allowing repetition of successful movements and the avoidance of errors. However, to date little is known about the relation between sensory working memory and this kind of motor learning.

Short-term memory has been previously shown to store feedforward control of reaching transiently before being consolidated in more stable and long-term memory (Brashers-Krug et al. 1996; Krakauer et al. 1999; Tong et al. 2002). Individual differences in working memory capacity have been assessed in relation to the amount of motor learning. It was shown that estimates of visuospatial working memory capacity correlate with the rate of sequence learning and visuomotor adaptation (Bo and Seidler 2009; Bo et al. 2011). In a related study that included neuroimaging, spatial working memory was involved early in visuomotor adaptation and was associated with task-related neural activity in right dorsolateral prefrontal cortex and bilateral inferior parietal cortex (Anguera et al. 2010). The role of visuospatial working memory in visuomotor adaptation has also been demonstrated when adaptation involved either an explicit strategy or adaptation to an abruptly introduced perturbation (Christou et al. 2016). In contrast, with the use of a gradual perturbation, which minimizes explicit strategies, working memory capacity was no longer a reliable predictor of learning.

In situations in which there is only success or failure information about movement outcome (reinforcement learning), learning is partly driven by positive feedback and reward, which serve as reinforcement. Prior studies have reported the influence of reward on motor learning when other types of information are available as well. Specifically, positive feedback during training increases memory for reaching direction in a visuomotor adaptation task (Galea et al. 2015; Shmuelof et al. 2012; Therrien et al. 2016) and memory of pinch force (Abe et al. 2011). Although it has been established that reward

Address for reprint requests and other correspondence: D. J. Ostry, Dept. of Psychology, 2001 McGill College Ave., Montreal, QC, Canada, H3A 1G1 (e-mail: david.ostry@mcgill.ca).

during training plays a role in motor learning, its effects on working memory remain unclear.

Recent behavioral studies have investigated somatosensory processes involved in reinforcement-based motor learning in a task in which participants received binary feedback on their movement outcome (Bernardi et al. 2015; Therrien et al. 2018). In the present study, we assessed the relationship between somatosensory working memory and human motor learning in a similar task in which participants made movements to hidden targets and received positive feedback when the movement finished within a target zone. We hypothesized that, given the paucity of visual information, the task would be heavily reliant on somatosensory information (Bernardi et al. 2015) and accordingly that participants with better somatosensory working memory would show better learning. This hypothesis was addressed in two separate experiments. The first was an offline experiment, in which working memory and learning tasks were completed on separate days and working memory capacity was tested as a predictor for motor learning performance. The second experiment aimed to understand what movements participants held in memory during the experiment. One hypothesis is that participants held both successful and unsuccessful movements in memory because these are the movements to adopt or avoid, respectively, in the future. An alternative hypothesis is that predominantly rewarded movements are remembered because in principle repeating these movements accurately is sufficient for performing the task. These contrasting hypotheses were tested in an online experiment, in which motor learning and memory trials were presented in an interleaved fashion to assess participants' memory for their own movements. The online technique enabled us to examine in a trial-by-trial manner whether successful or unsuccessful movements were remembered more or less well. Overall, it was observed that participants who had better somatosensory memory learned more in the motor task. The online experiment also revealed that successful trials (trials with positive feedback as reinforcement) were better remembered.

## MATERIALS AND METHODS

We conducted two separate studies that measured offline and online working memory, respectively. A total of 30 right-handed participants were recruited (6 men, mean age = 22.11 yr old, SD = 2.85) for the offline working memory experiment, which consisted of two experimental sessions, one testing motor learning and the other testing working memory. Each session was completed on a separate day with the order counterbalanced across participants. For the online working memory experiment, we recruited another 30 right-handed participants (4 men, mean age = 20.9 yr old, SD = 2.45) for a single-day study. All procedures were approved by the McGill University Faculty of Medicine Institutional Review Board. The participants were healthy adults with no prior physical or neurological conditions and provided written informed consent to the study.

Both experiments used a two-degree-of-freedom robotic manipulandum (Interactive Motion Technologies) with a vertical handle attached to the end-effector. Participants were seated in front of the robot with their right shoulder abducted to  $\sim 70^\circ$  and the elbow supported by an air sled. A semisilvered mirror, which served as a display screen, was placed just below eye level and blocked the vision of the arm and the robot handle. A white start circle, 20 mm in diameter, was positioned on the display screen at  $\sim 30$  cm in front of the participant's body midline.

### Offline Working Memory Experiment

*Sensorimotor learning paradigm.* The task in this study was similar to that used in previous work (Bernardi et al. 2015; Sidarta et al. 2016). Briefly, in the left part of the workspace, the participant was shown a 1-cm-thick white target stripe or bar, within which there was a hidden rectangular target zone that also had a width of 1 cm (Fig. 1B). The center of the zone was located 15 cm from the center of the start circle. Parallel to this target stripe was a thin yellow line that indicated the distance of the hand from the stripe. A small 12-mm-diameter yellow circle attached to the yellow line corresponded to hand position. This circle was shown briefly at the beginning of each movement and disappeared as soon as the robot handle left the start position. No information about the lateral deviation of the hand was provided during movement, so the participants could not use the error information associated with lateral distance from the target as a learning signal.

The participant was first given 15 familiarization trials with instructions. In both the familiarization trials and in the actual experiment, participants were told that after a "Go" cue appeared, they had to perform straight outward reaching movements to the target stripe without making corrections. Each movement had to stop within the stripe and be completed within 500 - 700 msec. The participant was given feedback about the movement speed verbally if they were consistently too slow or too fast. However, there was no penalty if the movement did not end on time or ended outside of the stripe. Once the movement ended, the robot brought the arm back to the start position.

The experiment began with a block of 25 baseline trials without any feedback regarding movement accuracy. Participants were instructed to reach at an angle of  $45^\circ$  to the left. They then performed 4 training blocks of 50 trials each and were told to learn which movement to the unseen target was successful, that is, ended within the target zone. The goal was to reduce the deviation with respect to the hidden target. Success was determined solely by the lateral deviation at the movement end point, not the movement distance or speed. Following a successful trial, an animated explosion and the words "Nice shot!" along with a pleasant tone and a running score appeared on the screen to provide positive feedback as reinforcement. The width (1 cm) and center position of the target zone ( $45^\circ$  to the left workspace) were fixed. In this offline experiment, the participant was told to pay attention to the arm configuration when successful and to make as many successful movements as possible (no such instruction was given in the online experiment). The session ended with a final set of 25 movements without any feedback, which evaluated motor accuracy following learning.

*Offline somatosensory working memory task.* An offline somatosensory working memory task tested recognition memory. A set of memory items was presented one at a time followed by a test item (probe). The participant had to indicate whether the test item was in the memory set or not. In the present working memory task, the to-be-remembered items were passive limb displacements produced by the robot in directions in the left part of the workspace as in the sensorimotor reaching task described above. During the experiment, the view of the arm was occluded, and the screen was completely blank.

Each trial of the memory test began with the words "New Round" presented on the screen as a visual cue. The participant was instructed to remain passive as the robot displaced the right arm outward in four different directions (Fig. 1C); each had an amplitude of 15 cm and took 900 ms to complete. After a brief hold time at the destination, the manipulandum moved the arm back to the start position. There was a delay of 500 ms between consecutive movements. Once the participant had experienced the four memory items, a tone was played, which was a cue indicating the following displacement would be the test direction. The participant responded verbally after having experienced the test direction, that is, "Yes" if they felt that the test item

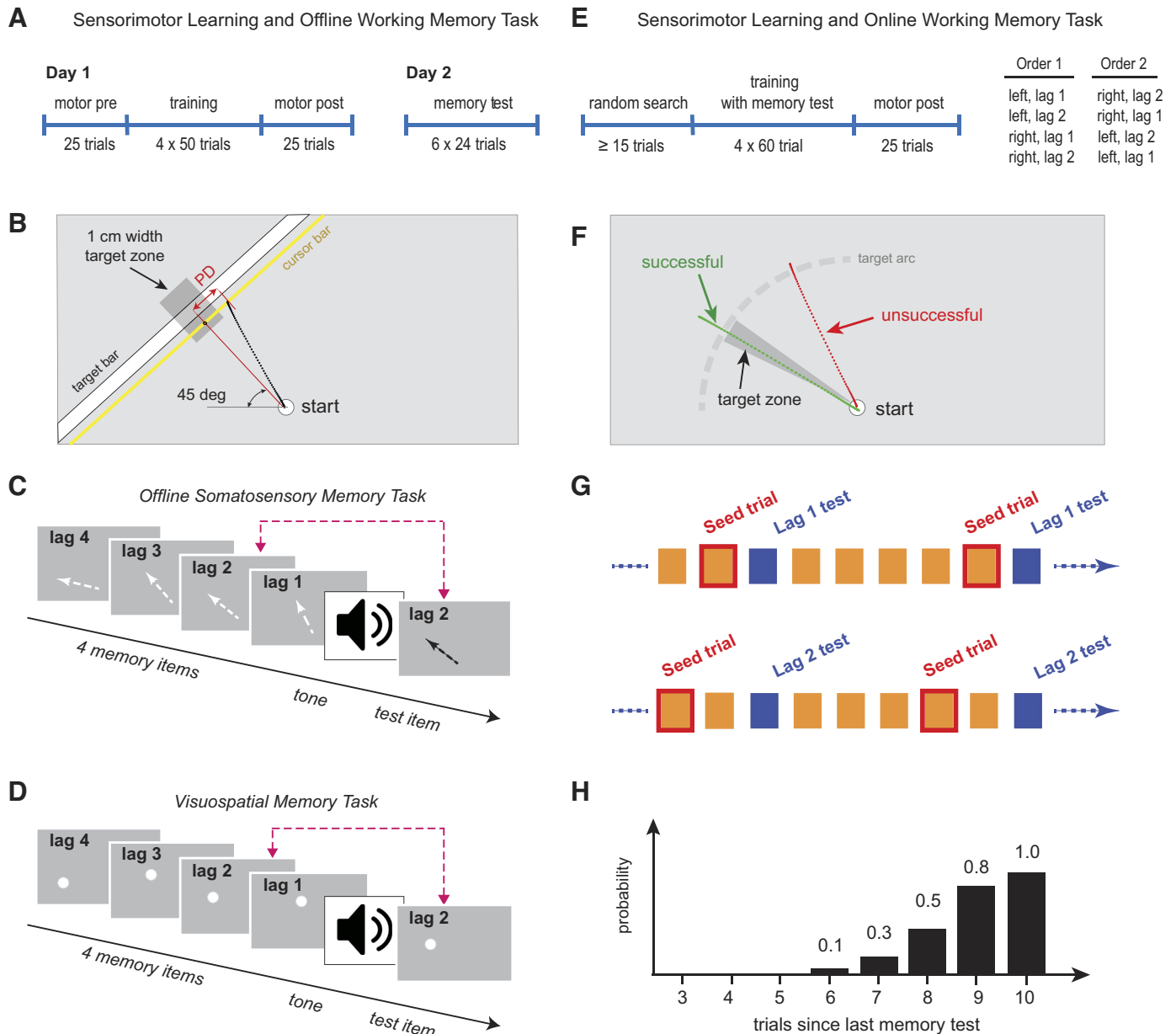


Fig. 1. Experimental setup. *A*: offline working memory testing sequence. *B*: participants learned to make movements to an unseen visual target (gray). Movement accuracy was quantified as the lateral perpendicular deviation (PD) at the movement end point. *C*: somatosensory working memory test design. *D*: visuospatial working memory test design. *E*: online working memory testing sequence. *F*: online sensorimotor learning with examples of successful (green) and unsuccessful (red) movements. A movement was successful when it fell within  $2.5^\circ$  of the target direction (shown in gray). *G*: online working memory test trials (blue) were interleaved with the sensorimotor training trials (orange). Each working memory trial was based on a previous movement (seed trial, red). *H*: working memory test trials were placed according to a predefined probability distribution, which was a function of the number of trials since the last working memory test.

was one of the four directions presented in the memory set and “No” otherwise.

In a given block of memory test trials, the test item was one of the four directions in the memory set on half of the total trials. In such trials, the test items were presented with varying lags separating the test item and the to-be-remembered item in the memory set. For example, a lag 2 memory trial means that the test direction was the same as the memory set direction presented two items ago (Fig. 1C, dashed arrow). In the remaining half, the test item was a lure; that is, it was a totally new direction. The order in which the test direction was a lure or was one of the previous memory set items was randomized across trials.

On a given memory test trial, the set of four memory items and the test item (probe) were obtained as follows. In all cases, we first

started with 6 directions equally spaced with  $10^\circ$  separation. These six directions  $d_i$  were found using the following formula:  $d_i = 10i + 100^\circ + j$ , where  $i \in \{1, 2, 3, \dots, 6\}$  and  $j \in \{-11^\circ, -10^\circ, \dots, 10^\circ, 11^\circ\}$ . Two out of these six directions were then chosen pseudorandomly to be discarded, resulting in the memory set of four directions. The directions that were discarded were not at the two extremities and not adjacent to one another. Last, the test direction was selected as follows. If this was a lure trial, one of the two removed directions became the test direction, otherwise one of the four memory directions was chosen at random to be the test direction.

Before the start of the actual task, each participant went through six familiarization trials with feedback (correct/incorrect) to ensure that they understood the task. The actual task consisted of 6 blocks of 24

trials that lasted for ~8–10 min each with a short break after each block. Overall, there were 18 trials at each lag.

**Visuospatial working memory and digit span tasks.** To determine whether the relationship between working memory and learning was specific to somatosensory memory, we invited the participants who did the above tasks to participate in a set of control conditions. The first task was a visuospatial working memory task, which assessed the ability to remember the locations of items presented visually in space. The task resembled the somatosensory working memory task with the exception that there was no displacement of the arm and the stimuli were locations of the end of movement that were shown as white circles on the screen (Fig. 1D). Each participant underwent a series of familiarization trials before the actual test. They were told to pay attention to a  $20 \times 20$  cm bounded area on the left side of the screen with the white start position in its lower right corner. Four white circles would then be presented one after the other, followed by a tone and a test circle. The participant then had to verbally indicate whether the test circle was in the memory set or not. As with the somatosensory working memory task, the actual test consisted of 6 blocks of 24 trials each, and the set of memory and test items was generated using similar procedures.

Two other cognitive tests were also employed as control tasks. Following the visuospatial working memory task, the participants were presented with forward and backward digit span tasks to tap into the verbal short-term memory (Wechsler 1999). In this task, participants were presented a sequence of digits on the screen and then had to report the sequence in forward or backward order (as specified by an instruction) using the keypad. At the beginning of the test, a message would appear on the monitor screen to tell the participants whether the task was a forward or backward task. During the experiment, a series of numbers at a pace of 1 s per digit was presented with a 1-s pause in between sequential digits. Both tasks began with a set of three-digit numbers and continued up to nine-digit numbers. Within a set, there was no single number that was repeated, and the digit sequence was random. Before the experiment, we provided the participants with three familiarization trials with instructions using two-digit numbers. Subsequently, they began the actual task, which consisted of the forward and backward digit span task (with the order counterbalanced). Task performance was quantified as the proportion of correct trials. Out of 30 original participants, 25 participated in the control conditions.

**Data analysis.** Motor performance in each trial was quantified as the perpendicular deviation (PD) at movement end point from a straight line originating at the start position and passing through the center of the target bar, which is exactly  $45^\circ$  to the left of a straight-ahead movement. If the movement ended beyond the target bar, the perpendicular lateral deviation was computed with respect to this movement end point (Fig. 1B). Movements that ended closer to the center had smaller PD scores. For each participant, the average absolute deviation (IPD) before (Pre) and after (Post) training was calculated using the 25 trials without any positive feedback, and the difference served as a measure of improvement in accuracy, with larger positive values corresponding to greater learning. Using the same set of trials, we also assessed accuracy in terms of movement bias (or the average value of signed PD) and end-point variability (or standard deviation of signed PD), which evaluated movement consistency. During training blocks, to assess whether positive feedback or its absence influenced the movement on the immediately following trial, we calculated the absolute change in movement direction following each successful and unsuccessful trial as  $\Delta m_n = \text{IPD}_{n+1} - \text{PD}_n$ , which gives the difference in PD between the trial  $n$  and  $n + 1$ , contingent on trial  $n$  being successful or not.

In the somatosensory working memory task, we quantified both the hit rate (proportion of “yes” responses when the test item was part of the memory set) and the false alarm rate (proportion of incorrect “yes” responses) for each lag, and the difference between hit and false alarm rates was obtained. Using ANOVA, we assessed differences in hit –

false alarm rates across lags. The same analyses were conducted for the visuospatial working memory test. Tests for normality and assumption of sphericity of the data set were conducted using Shapiro-Wilk test and Mauchly’s test, respectively. Relevant post hoc analyses were done with Bonferroni-Holm correction.

A composite somatosensory working memory score of each participant was computed as the average of hit – false alarm rates over all four lags. A similar approach was used to obtain an individual’s visuospatial working memory score. Performance on the forward and backward digit span tasks was measured by proportion of correct trials. Subsequently, we computed the correlation between each of the memory scores and the measures of learning together with the 95% bootstrapped confidence interval (CI).

### Online Working Memory Experiment

**Sensorimotor learning paradigm.** Whereas the first experiment was designed to test the relationship between the somatosensory working memory performance and the amount of motor learning measured separately (offline), a second experiment measured participants’ memory for their own movements on a trial-by-trial basis during the learning processes. As such, the working memory test was interleaved with the motor learning trials themselves. This experiment used the same basic setup as the offline experiment but was divided into two parts, which involved sets of movements to a hidden target at the right and the left of the workspace, respectively. Memory testing was restricted to two lags, one tested at the right and the other at the left. The assignment of movement direction and memory test lag was random across participants (see Fig. 1E). There was a 10-min break halfway through the experiment, at which time participants switched movement directions and lags.

The study began with familiarization trials in which a quarter arc with a 1-cm thickness was shown on the screen (Fig. 1F). As before, vision of the participant’s arm was blocked. Participants were instructed to move to any point on the arc after the “Go” cue appeared and to make straight movements without corrections. The yellow hand cursor position was removed once the arm moved outside of the white start circle. The required movement duration was 500–700 ms, but there was no penalty if the movement did not end on time or outside the target arc. Once the movement ended, the robot brought the arm back to the start position. Directional error was measured in terms of angular deviation (AD) from the true target direction at the maximum movement speed. The width of the target zone was  $5^\circ$ , and positive feedback was provided if the AD was within  $\pm 2.5^\circ$  (Fig. 1F).

Following the familiarization trials, the arc was removed. However, the participant was instructed to move in the direction of the arc and was told that there was a target located in the now hidden arc. The task was to search for the correct direction to the target and then to continuously move in the same direction. When the direction was correct, the trial was considered successful, and the participant was given the same positive feedback as in the offline experiment (an animated explosion, a pleasant tone, and a score). This positive reinforcement was independent of the movement length although we told the participants during the familiarization trials whether the movement was too far beyond the arc or too short. For each participant, we chose a participant-specific target direction as follows. The participant first made 15 baseline movements (without feedback). The target direction was then set to the direction of the first movement after the 15th trial that fell within the range of  $20\text{--}70^\circ$  relative to the horizontal at the right of the workspace, or  $110\text{--}160^\circ$  at the left. This provided at least 15 movements in which participants randomly explored the workspace before the first reinforced (successful) trial. It also eliminated the use of explicitly defined directions to the target. Throughout training, the width and position of the reinforced direction did not change. After the random exploration phase, participants completed 4 blocks of 60 training trials with positive feedback when



successful. This was followed by 25 further movement trials with no feedback.

**Online somatosensory working memory test.** The online working memory test was designed to assess participants' memory for their own movements during motor learning in a trial-by-trial manner (Fig. 1G). Individual reaching movements were recorded from movement start to movement end point. On a memory test trial, the robot would replay a rotated version of the previous movement (in the case of a lag 1 memory test) or the movement two trials before (in the case of lag 2). The rotated movement was 5° to the left or right of the participant's original movement, selected at random. The movement that was used for the working memory test will be referred to as the seed movement. The task in the online working memory test was to indicate the direction of the rotation relative to their seed movement direction. Participants responded "Left" or "Right" for this purpose.

The online working memory tests were presented once every five to eight trials according to a probability distribution shown on Fig. 1H. A visual cue on the display screen appeared for 1,500 ms, indicating that the upcoming movement was a memory test. After responding, participants continued the training by again making reaching movements to the occluded target. Participants were explicitly informed whether lag 1 or lag 2 memory judgements were required for a given workspace direction.

**Data analysis.** In the online experiment, movement accuracy was quantified using absolute AD (IAD) measured at the maximum movement speed. We used an arc along with the AD so that the target location could be made different for each participant while still ensuring that all participants made movements of equal distance. The total number of trials with positive feedback over the course of the four training blocks was used to quantify a reinforcement index of learning. As before, we quantified the effect of positive feedback on the present trial on movement direction on the following trial with

$\Delta m_n = |AD_{n+1} - AD_n|$ , contingent on trial  $n$  being successful or unsuccessful. Because the working memory test was interleaved in between two training trials (trial  $n$  and  $n+1$ ), we also examined whether the presence of the online working memory test had any influence on the change in movement direction ( $\Delta m$ ) immediately after the memory test.

Working memory performance at each lag was quantified using the proportion of correct responses. To assess the effects of positive feedback on memory for movements, we examined whether memory was different following successful vs. unsuccessful movements. For each participant and each lag, working memory performance contingent on successful and unsuccessful seed movements was calculated separately. A two-way repeated-measures ANOVA was used to evaluate whether the success of a seed movement affected working memory performance at different lags. Tests for normality and assumption for sphericity were conducted using Shapiro-Wilk test and Mauchly's test, respectively. The Greenhouse-Geisser corrected  $P$  value was used if the sphericity assumption was violated. The correlation between motor learning and the overall memory performance was computed together with the 95% bootstrapped CI.

## RESULTS

### Offline Working Memory Experiment

In this study, movement accuracy before and after learning was quantified as the IPDI at the movement end point, based on the 25 movements in the baseline (Pre) and motor evaluation blocks (Post). Movement bias was measured as the average value of signed PD. Overall, participants showed learning as indicated by a reliable decrease in the mean absolute lateral deviation, IPDI [ $t(29) = 4.82$ ,  $P < 0.001$ ] (Fig. 2A) and in the

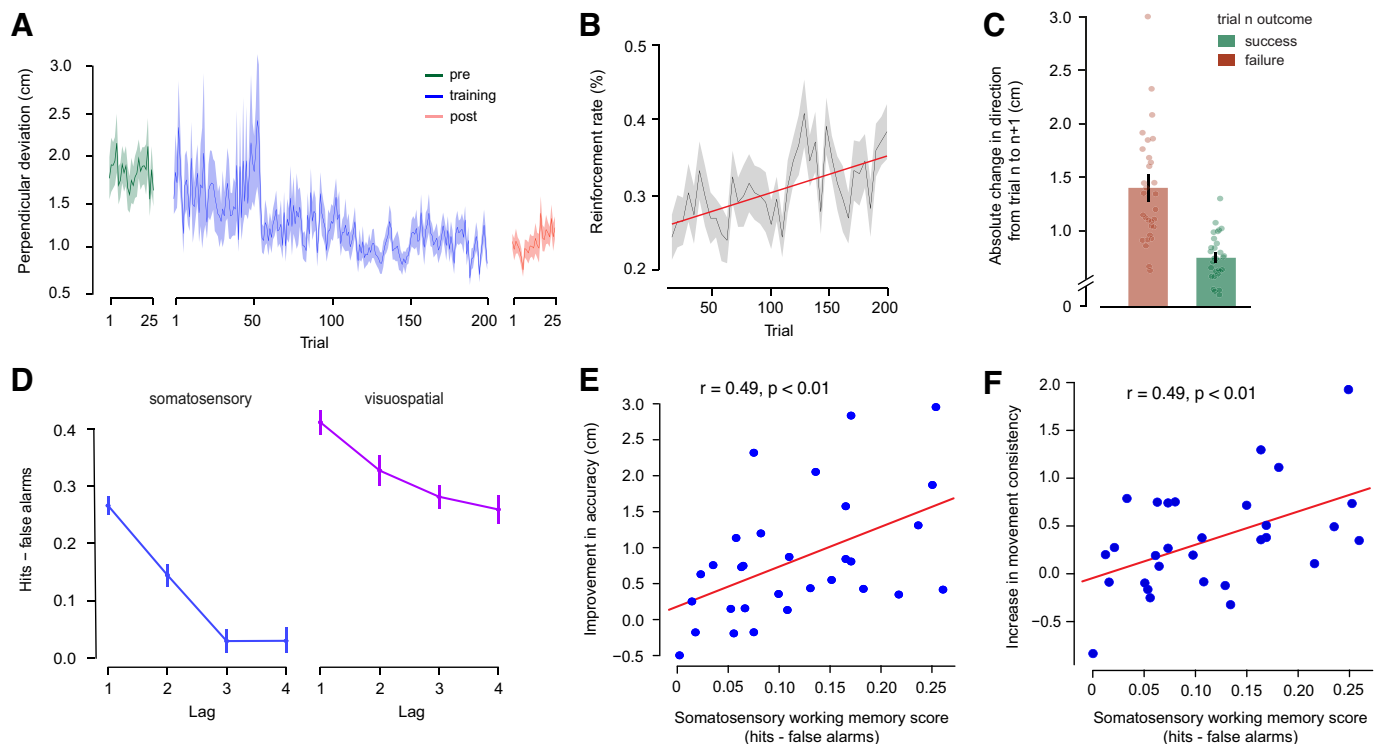


Fig. 2. Results of the offline working memory experiment. **A**: movement accuracy as measured by the perpendicular deviation at movement end point (IPDI) in 3 different task phases (pre, training, and post). **B**: reinforcement rate averaged across participants (bins of 5 trials) shows an increase in the number of successful trials over the course of training. **C**: there was a larger change in movement direction following unsuccessful trials. **D**: somatosensory and visuospatial working memory decreases for movements that were presented longer ago (increasing lag). **E**: somatosensory working memory performance was positively correlated with improvement in accuracy in the motor learning task. Participants who had better working memory learned more. **F**: somatosensory working memory performance was positively correlated with reduction in end-point variability.

magnitude of bias [Pre:  $M = 1.57$  cm,  $SD = 0.73$  cm; Post:  $M = 1.01$  cm,  $SD = 0.57$ ;  $t(29) = 5.19$ ,  $P < 0.001$ ]. With the use of the standard deviation of signed PD as a measure of movement end-point variability, it was found that there was decrease in variability from before to after learning [Pre:  $M = 1.29$  cm,  $SD = 0.47$  cm; Post:  $M = 0.96$  cm,  $SD = 0.32$  cm;  $t(29) = 3.16$ ,  $P < 0.005$ ]. Moreover, there was a correlation between the reduction in error magnitude and the reduction in variability [ $r = 0.45$ ,  $P = 0.013$ , CI (0.29, 0.67)], indicating that participants that showed greater improvement in accuracy also had a greater reduction in variability.

To quantify the reinforcement rate over time, a linear function was fit to reinforcement rate (with a bin size of 5 trials) for each participant to provide an estimate of the slope. The average slope across participants was shown to be significantly different than zero [1-sample  $t$ -test:  $t(29) = 3.18$ ,  $P < 0.005$ ], suggesting that the amount of positive reinforcement increased over training trials (Fig. 2B). Absolute change in signed PD between the current and next immediate movement ( $\Delta m$ ) was computed to assess the effect of the positive reinforcement on subsequent movements. Nonreinforced movements resulted in a greater trial-to-trial change in movement direction than reinforced movements, which presumably reflects exploration to find the correct direction when movements fail to end in the target zone [ $t(29) = -6.33$ ,  $P < 0.001$ ] (Fig. 2C).

In terms of somatosensory working memory, it was found that response accuracy decreased as a function of lag [ $F(3,87) = 54.29$ ,  $P < 0.001$ ], indicating that more recently experienced movements were remembered more accurately (Fig. 2D). Performance at the first two lags was significantly different from 0 (Bonferroni corrected,  $P < 0.01$ ), suggesting that, for this task, people could reliably maintain two previous movement directions in working memory. Analyses of working memory were also conducted for the visuospatial memory task ( $n = 25$ ), which likewise yielded differences in performance across lags [ $F(3,72) = 17.26$ ,  $P < 0.001$ ]. In general, visuospatial working memory performance was better than that for somatosensory working memory [ $F(1,24) = 106.43$ ,  $P < 0.001$ ] in a manner that varied across lags [2-way interaction:  $F(3,72) = 3.53$ ,  $P = 0.018$ ]. In somatosensory working memory, reliable differences were observed between lag 1 and lag 2 and between lag 2 and lag 3 ( $P < 0.005$ ) but not between lag 3 and lag 4 ( $P = 1.0$ ). In contrast, visuospatial memory scores between lag 1 and lag 2 were found to be different ( $P = 0.011$ ), but there was no difference in scores in the subsequent lags ( $P > 0.52$ ).

In the forward version of the digit span test, the overall proportion of correct responses was 68.6% ( $SD = 3.1\%$ ), whereas, for the more difficult backward digit span test, the proportion correct was 60.4% ( $SD = 4.5\%$ ). We estimated the degree of association between somatosensory working memory and the three other memory tasks, visuospatial working memory, forward, and backward digit span (Table 1). We found that somatosensory and visuospatial working memory scores showed a positive correlation [ $r = 0.43$ ,  $P = 0.038$ , CI (0.19, 0.790)]. In contrast, there was no reliable correlation between somatosensory working memory and either the forward digit span [ $r = 0.18$ ,  $P = 0.39$ , CI (-0.21, 0.53)] or the backward digit span test [ $r = 0.33$ ,  $P = 0.09$ , CI (0.08, 0.70)].

Somatosensory working memory performance was positively correlated with the accuracy improvement such that

Table 1. Correlation coefficient ( $r$ ) between somatosensory working memory and other measures of working memory

Task	VSWM	DSf	DSb
SWM	0.43*	0.18	0.33
VSWM		0.32	0.23
DSf			0.71†

DSb, digit span (backward); DSf, digit span (forward); SWM, somatosensory working memory; VSWM, visuospatial working memory. \* $P < 0.05$ , † $P < 0.001$ .

individuals with better memory showed greater reduction in IPDI [ $r = 0.49$ ,  $P = 0.006$ , CI (0.26, 0.81)] (Fig. 2E). Better somatosensory working memory performance was also related to lower movement variability following learning [ $r = 0.49$ ,  $P = 0.005$ , CI (0.27, 0.82)] (Fig. 2F). Visuospatial working memory had no reliable relationship with the reduction in IPDI [ $r = 0.12$ ,  $P = 0.55$ , CI (-0.28, 0.49)] but was positively correlated with the reduction in variability, such that individuals with higher visuospatial working memory performance had less variable movements [ $r = 0.65$ ,  $P < 0.005$ , CI (0.30, 0.82)]. Performance on the digit span tasks was not related to any of the learning measures ( $r < 0.20$ ,  $P > 0.10$ ).

To assess whether the relationship between the reduction in absolute error and the memory score was specific to the somatosensory modality, we conducted multiple linear regression with the reduction in error as the dependent variable and the four memory scores (somatosensory, visuospatial, and two-digit-span tasks) as predictors. It was found that somatosensory working memory was able to explain the reduction in error ( $P = 0.027$ ) but not the other predictors ( $P = 0.58$  for visuospatial,  $P = 0.25$  and  $P = 0.18$  for forward and reverse, respectively). In a second model, we used the reduction in variability as the dependent variable and found that visuospatial working memory score was a reliable predictor ( $P = 0.016$ ) but not somatosensory working memory ( $P = 0.19$ ) or the remaining two predictors ( $P = 0.89$ ,  $P = 0.67$  for forward and reverse digit span, respectively).

It has been demonstrated previously that task-relevant baseline variability in reinforcement-based learning is able to predict the amount of learning (Wu et al. 2014). To address the concern that the correlation between working memory and motor learning was driven by differences in baseline variability, we conducted the following analysis. Baseline variability was quantified using the standard deviation of the signed PD during Pre (trials without feedback). After we controlled for the baseline variability using partial correlation, the relationships remained significant between somatosensory working memory and reduction in absolute error [ $r = 0.43$ ,  $P = 0.018$ , CI (0.20, 0.73)] and in reduction end-point variability [ $r = 0.48$ ,  $P = 0.009$ , CI (0.29, 0.72)].

#### Online Working Memory Experiment

In a second experiment, working memory trials were interleaved with motor learning trials, allowing us to test memory for movements that the participants actually performed during learning.

We first obtained behavioral measures of learning in the sensorimotor task. Because this experiment involved blocks of testing in which movements were made either to the right or

the left of the workspace (with order balanced), we tested for the possibility of order effects on motor learning. The order in which participants experienced the two movement directions was not found to significantly affect the overall amount of learning in either direction as assessed using the total number of reinforced movements [ $t(29) = -0.42, P = 0.67$ ], the |ADI| measured during the last block of training trials [ $t(29) = -1.72, P = 0.1$ ], and during Post [ $t(29) = 0.68, P = 0.52$ ]. We examined differences in learning performance between movements in the left and right workspace. There was no reliable difference in terms of accuracy during the last block of learning [ $t(29) = 1.02, P = 0.31$ ] or during Post [ $t(29) = -1.19, P = 0.08$ ] between the two directions. We found a reliable difference in terms of total reinforced trials [ $t(29) = -2.87, P = 0.005$ ], indicating that movements on the left were overall less successful ( $M = 82.4, SD = 24.5$ ) than on the right ( $M = 103.3, SD = 32.7$ ). In subsequent analyses, the behavioral measures of learning for the individual participants were averaged across the two reaching directions. The mean movement distance traveled toward the hidden arc was 20.06 cm ( $SD = 5.4$  cm). Because the target arc was invisible throughout training, it is possible that differences in the extent of reaching and movement speed might have an effect on the overall accuracy. Taking together the data from all participants, we found that neither movement distance ( $r = 0.02, P = 0.92$ ) nor speed ( $r = 0.02, P = 0.07$ ) influenced the |ADI|.

Figure 3, A and B, shows the movement accuracy as defined by the |ADI| and reinforcement rate. The AD at maximum speed was significantly correlated with the AD at the movement end point ( $r = 0.82, P < 0.001$ ), as well as with the PD measured at movement end point ( $r = 0.49, P < 0.001$ ). To assess whether there was learning, a linear function was fit to the |ADI| over all training trials for each participant to provide an estimate of the learning slope. We took this approach rather than measuring differences between baseline and posttest movements because in the present experiment there was no actual target defined until its direction was set on or about *trial 16*, based on each individual participant's movement direction. We found that the average slope across participants was significantly different than 0, indicating that the error magnitude decreased over training trials [ $t(29) = -3.17, P < 0.01$ ].

Figure 3B shows that average reinforcement rate (with a bin size of 5 trials) across participants increased over the course of learning [1-sample  $t$ -test,  $t(29) = 2.17, P = 0.022$ ]. It was further found that participants that received more total reinforcement typically had less variable [ $r = -0.46, P = 0.011, CI (-0.70, -0.12)$ ] and more accurate movements (smaller |ADI|) [ $r = -0.50, P = 0.005, CI (-0.73, -0.21)$ ] during Post trials without positive feedback. In addition, participants that made more accurate movements during Post trials also produced less variable movements [ $r = 0.78, P < 0.0001, CI (0.59, 0.89)$ ]. As with the offline sensorimotor task, more positive reinforcement was associated with a smaller magnitude of change in movement direction ( $\Delta m$ ) in the next immediate trial [ $r = -0.89, P < 0.0001, CI (-0.91, -0.75)$ ], consistent with the idea that positive reinforcement reduces trial-to-trial variability.

Average proportions of correct responses for lag 1 and lag 2 test were  $M = 76.1\%$  ( $SD = 1.9\%$ ) and  $M = 71.8\%$  ( $SD = 2.2\%$ ), respectively, where 50% denotes chance level. No significant difference was observed in overall

working memory performance between lag 1 and lag 2 [ $t(29) = 1.52, P = 0.16$ ] or between the movement direction tested first and the one that was tested second [ $t(29) = 1.62, P = 0.15$ ]. Average memory performance for movements in the right and left workspace was  $M = 77.6\%$  ( $SD = 9.9\%$ ) and  $M = 70.5\%$  ( $SD = 11.7\%$ ), and the difference was reliable [ $t(29) = 2.73, P = 0.02$ ]. We investigated whether longer/shorter movements were better remembered as follows. For each subject, all memory trials were grouped according to the extent of the seed movement using median split. A similar analysis was performed to examine whether movement speed influenced memory performance. It was found that there was no reliable difference in memory performance between seed movements that were long and short [ $t(865) = 0.11, P = 0.43$ ] nor between seed movements that were fast and slow [ $t(865) = 0.41, P = 0.96$ ]. This suggests that the average memory score is insensitive to both movement extent and speed.

An overall measure of working memory performance was computed for each participant as the mean proportion of correct answers combining both lags and workspaces. This approach was adopted because there was no significant difference in online working memory performance between lag 1 and lag 2 or between the movement direction tested first and the one that was tested second ( $P > 0.10$ , respectively). Subsequently we assessed the relationship between performance during training and online working memory. The working memory score was found to be reliably associated with the total number of reinforced movements [ $r = 0.47, P = 0.009, CI (0.19, 0.82)$ ] (Fig. 3C). Participants with higher working memory scores also achieved better asymptotic performance as indicated by smaller |ADI| in the last block of training [ $r = -0.41, P = 0.039, CI (-0.64, -0.072)$ ].

By interleaving the working memory task with training movements, we were able to evaluate possible differential effects on memory of making movements that successfully ended in the target zone (reinforced movements) and those that missed the target and did not receive reinforcement. Figure 3D shows the working memory score for each lag according to whether the corresponding seed movement was successful or not. It can be seen that for both lags memory was better when tests involved successful seed movements than when seed movements were not reinforced [ $F(1,29) = 6.08, P = 0.019$ ], and this was not different across lags [ $F(1,29) = 0.153, P = 0.68$ ].

The presence of a working memory test interleaved between two consecutive training trials may affect the movement trial immediately following it. For example, it is possible that the movement direction deviates more from the target zone following a working memory test but in a certain AD. To investigate this possibility, we quantified the change in direction ( $\Delta m$ ) when working memory trial intervened between trial  $n$  and  $n + 1$ , in cases when trial  $n$  was reinforced and not reinforced. Figure 3E shows that reaching movement deviated more following a working memory test but only when the present trial  $n$  was reinforced [2-way interaction:  $F(1,29) = 33.11, P < 0.0001$ ]. Post hoc analyses showed that the effect of the memory task on  $\Delta m$  was greater following a successful trial ( $P < 0.001$ ) than an unsuccessful trial ( $P = 0.351$ ). We also found that this additional amount of change in movement direction after a work-

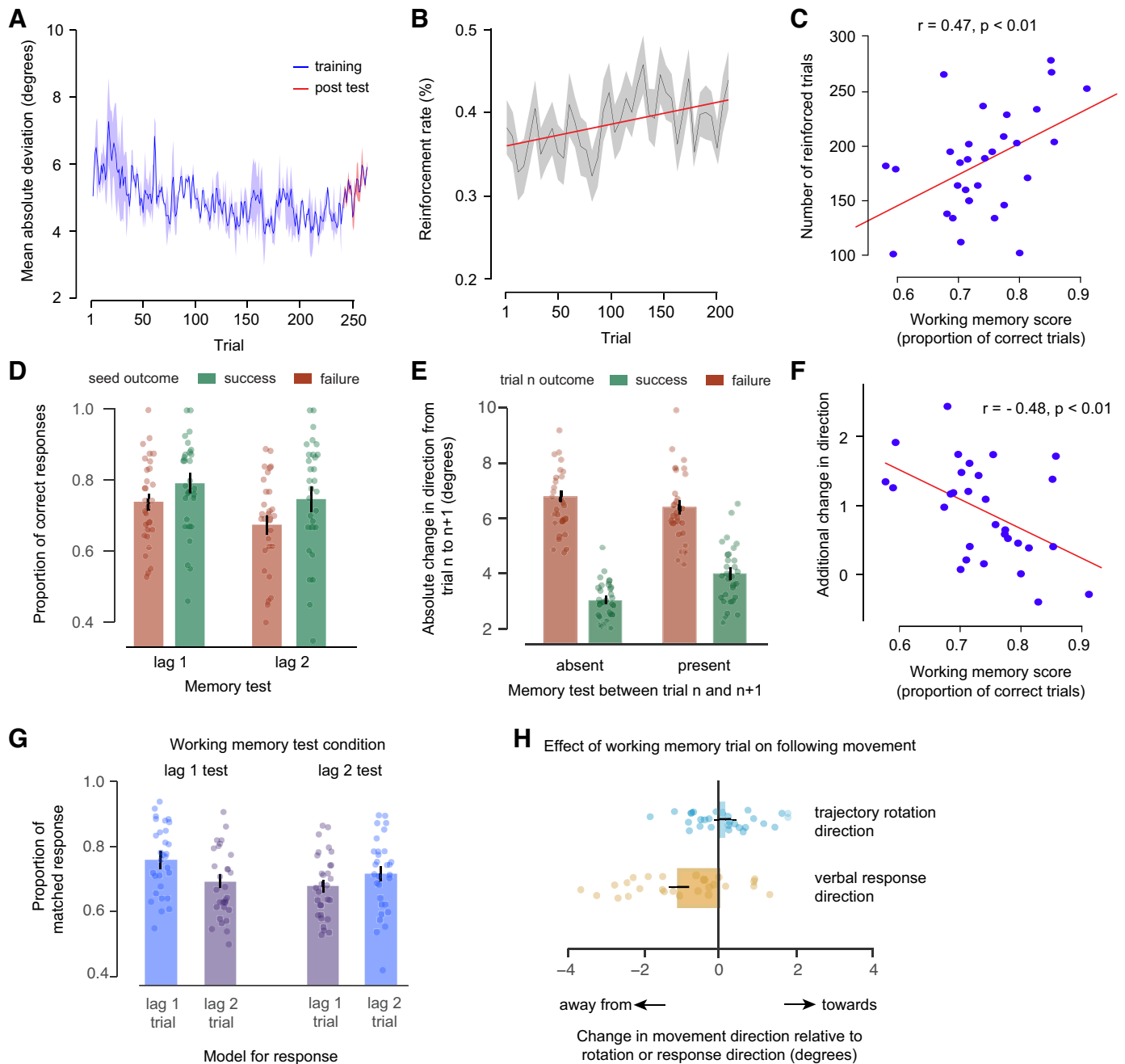


Fig. 3. Online working memory experiment. *A*: participants showed learning as indicated by a reduction in angular deviation (IAD). *B*: reinforcement rate increased over trials. *C*: participants with better working memory learned more as shown by a greater number of reinforced trials. *D*: successful movements were remembered better. *E*: presence of a working memory test resulted in an increase in directional change in the following trial but only when the preceding trial was successful. *F*: participants with better working memory show smaller changes in movement direction following memory tests, indicating that the working memory trials disrupted the learning less. *G*: participants correctly based their memory test responses on the appropriate seed (reference) movement. For each participant, model responses were computed based on lag 1 or lag 2 movements, and these were then matched to participant's actual responses (indicated by the bar height). *H*: following a working memory test, movements were biased away from the participants' judgments (verbal responses) of the memory test direction but were not affected by the actual trajectory rotations.

ing memory test,  $\Delta m_{WM} - \Delta m_{noWM}$ , was negatively correlated with the working memory score [ $r = -0.48$ ,  $P = 0.008$ , CI  $(-0.72, -0.16)$ ]; that is, participants who had higher working memory test scores were affected less (Fig. 3*F*).

In the offline experiment, the capacity of the somatosensory working memory was found to be roughly two items. In the online experiment in which participants were tested using their own movements, memory performance of each lag was reli-

ably greater than chance level, suggesting that they were able to remember two movements as well [1-sample  $t$ -test,  $t(29) = 11.91$ ,  $P < 0.001$  and  $t(29) = 10.51$ ,  $P < 0.001$ ]. Because the working memory test in the present study made use of the participant's own movements, it was not possible to control for the angular differences between two consecutive movements, one of which may serve as the seed movement for the working memory test. Accordingly, we asked whether



participants were using the actual seed movement information, in particular, when doing the lag 2 working memory tests. To assess whether this was the case, we examined participants' responses on the subset of lag 2 memory trials for which the correct answer would be different if participants in fact were using the lag 2 vs. the lag 1 movement as the reference for their judgements. For every participant, we computed the proportion of answers that matched the lag 2 reference answer and the proportion of answers that matched the lag 1 reference answer. Similar analysis was also done for the working memory lag 1 test in which the wrong reference in this context was the movement performed two trials before. It was found that participants' answers matched more often the answers that would be expected if they were basing their response on the actual seed movement than if they based it on the wrong movement (Fig. 3G) [ $F(1,29) = 16.87$ ,  $P < 0.005$ ], and the pattern was the same for both lags [ $F(1,29) = 0.99$ ,  $P = 0.34$ ]. This result suggests that participants were capable of basing their answers on the correct reference (seed) movements as instructed and not simply substituting as a basis for judgement with the wrong movements.

Finally, we assessed whether the change in direction following a working memory test was influenced by either the rotation direction used in the working memory trial (left or right) or by the direction indicated in the participants' response. Figure 3H shows the pattern of the signed directional change immediately after a working memory test. The top bar shows the change in direction relative to the direction of the memory test in which a positive value means that the movement direction is biased toward the direction of the rotation in the memory test. The bottom bar shows the change in direction relative to participant's judgement. The negative value means the movement is biased in a direction opposite to participant's verbal response. It is seen that the movement direction following the working memory trial was opposite to the participant's judgement, regardless of whether the response was correct or not [1-sample  $t$ -test,  $t(29) = 9.22$ ,  $P < 0.005$ ]. This suggests that the participant's perceptual judgment introduced a bias in planning the direction of the subsequent movement.

## DISCUSSION

The present studies demonstrated a relationship between sensory working memory and reinforcement-based motor learning. The sensorimotor learning task was based on a reinforcement learning paradigm in which participants made arm movements to unseen targets, and, when the movement ended within the target zone, participants received positive feedback as reinforcement to enable learning. In each experiment, we observed an improvement in movement accuracy over the course of training, which was also reflected in an increase in reinforcement rate. Somatosensory working memory was assessed using participants' judgements of the direction of passive displacements of the arm. In one experiment, memory tests and learning were performed separately in time. In the other, memory tests and learning trials were interleaved such that the memory tests probed the participants' memory for the movements they performed themselves in the context of learning. In both studies, we found that people with better somatosensory working memory learned more. The experiment involving interleaved memory and learning trials enabled

us to examine the contribution of positive feedback to working memory performance. It was found that successful trials, that is, trials that received positive reinforcement, were better remembered.

### *Somatosensory Working Memory Predicts Human Motor Learning*

The term somatosensory working memory is used in the present study to refer to recognition memory and decision making for arm configurations associated with reaching movements. Prior work in both humans and nonhuman primates has documented instances of working memory in the somatosensory domain. Such studies have often involved the use of tactile discrimination tasks in which, for example, one has to make judgments about the shape of an object (Kaas et al. 2007; Stoeckel et al. 2003), compare two sets of vibratory stimuli (Preuschhof et al. 2006; Romo et al. 1999), or recognize patterns by tracing lines in the absence of vision (Fiehler et al. 2008).

Other studies have documented aspects of somatosensory working memory with tasks that involve limb displacement. For example, the participant's arm was passively displaced by the experimenter to a target location, and the task was to reproduce the movement to the same location (Chapman et al. 2001; Goble et al. 2006; Jones and Henriques 2010). When a delay was introduced between the passive presentation and participant's reproduction, reaching was less accurate than immediate reaching, suggesting that short-term sensory memory decays over time. The present study also found that somatosensory working memory accuracy decreased for movements that were presented longer ago (at longer lags). In both experiments in the present data set, participants could reliably retain at least two prior movements in memory.

Individual differences in somatosensory working memory performance were found to correlate with the amount of reinforcement motor learning. This is consistent with previous work demonstrating a link between sensory working memory and visuomotor adaptation (Anguera et al. 2010; Christou et al. 2016). In another demonstration of this same relationship, when subjects perform a secondary task that depletes spatial working memory capacity, subsequent visuomotor adaptation is also impaired (Anguera et al. 2012). Likewise, in reinforcement learning, it has been shown that the use of a secondary task impairs learning (Codol et al. 2018; Holland et al. 2018). Taken together, those findings are consistent with the idea that working memory is involved in motor learning.

Is the memory involved in motor learning specific to the somatosensory domain, or is it a general memory capacity? To answer this question, we assessed whether other types of working memory might account for the individual differences in learning that were observed. To do this, a series of control tasks were used that involved visuospatial or verbal working memory. The forward and backward digit span tasks tested for the possibility that memory performance and possibly motor learning were related to verbal memory capacity. The visuospatial working memory task tested for the possibility that, although there was no explicit visual target, learning performance involved visuospatial information. Our results showed that motor learning was not related to digit span memory. In contrast, visuospatial working memory was reliably correlated

with a reduction in movement variability but not to measures of improved movement accuracy. This suggests that reinforcement-based motor learning may contain several components, such as reduction of error and reduction of variance, which proceed in parallel but may be driven by different processes and thus differentially dependent on working memory. The contribution of working memory to the observed reduction in absolute movement error was specific to the somatosensory domain. The reduction in movement variability is likely to entail more general memory processes, as it is reliably associated with both visuospatial and somatosensory working memory. Such domain-general memory capacities have been found in other studies, for example, in tasks that tap into both verbal and visuospatial working memory (Bo et al. 2011; Kane et al. 2004).

### *Working Memory and Positive Reinforcement*

In the present work, both experiments showed that movement accuracy increased over the course of learning. Positive reinforcement was shown to promote learning in terms of improvement in movement accuracy (less absolute error) and reduction in movement variability. In addition, trial-to-trial movement variability was influenced by reinforcement, with unsuccessful trials resulting in larger changes in movement direction, as was observed previously (Pekny et al. 2015; Sidarta et al. 2016).

If somatosensory working memory contributes to motor learning, are all movements equally well remembered? This issue was addressed in the present studies using a motor learning task in which, at pseudorandom intervals during training, a working memory test was delivered that tested how well participants remembered their own past movements. Specifically, we presented a participant's own movement with either a rightward or leftward deflection. This online task presumably draws on the natural learning situation in which one keeps track of prior somatosensory states. By deflecting the movement, we also probed participants' ability to make a perceptual judgment associated with their own actions by comparing test displacements with the information held in the somatosensory working memory.

Somatosensory working memory scores during motor learning were found to be higher for movements that received positive feedback (Fig. 3D). This finding, in conjunction with the observation that working memory for movements is limited to roughly two items, suggests that the nervous system deals with this limitation by prioritizing the retention of successful movements.

Memory bias toward rewarded movements may be due to factors such as attention, saliency, or the arousing effects of reinforcement that the participant received when movements were successful. This result forms part of an increasing body of evidence documenting that memory is enhanced for items or events associated with reward. Electrophysiological studies in nonhuman primates have found that reward influences neuronal discharge in areas of prefrontal cortex that are known to be implicated in working memory. Activity in a subset of dorsolateral prefrontal neurons was found to be modulated by reward of previously performed memory-guided saccades (Leon and Shadlen 1999; Tsujimoto and Sawaguchi 2004). Reward was also observed to modulate performance in a spatial memory

task such that the discharge pattern of neurons in ventrolateral prefrontal cortex was associated with both spatial cues and reward (Kennerley and Wallis 2009). The influence of reward on memory in these cases may be driven by projections of midbrain dopaminergic neurons to prefrontal cortex as shown by prior anatomical studies in nonhuman primates (Gaspar et al. 1992; Williams and Goldman-Rakic 1998). Similar influences of reward on memory are found in visual and auditory working memory in humans. Performance of visual working memory is modulated by reward, and activity in prefrontal cortex is correlated with reward value (Gong and Li, 2014; Klink et al. 2017; Krawczyk et al. 2007). On the basis of functional connectivity analyses involving auditory cortex, prefrontal cortex, and ventral striatum, pleasurable music is thought to be encoded more strongly in auditory working memory (Zatorre and Salimpoor 2013).

Previously, Pekny et al. (2015) and Holland et al. (2018) investigated trial-to-trial changes in movement direction in a reinforcement learning task as a function of the history of prior rewards. In each paper, the authors computed the difference in movement direction between a particular trial and the immediately preceding trial, as a function of the sequence of rewards going back as far as three preceding trials. It was found that the memory for a sequence of rewards influences the change in movement direction on the present trial. It is possible that the change in direction is due to memory decay with increasing distance from the last successful trial, which would be consistent with the lag effect on memory seen in the present study (Fig. 2D). An alternative possibility is that the increase in variability following a string of unsuccessful movements reflects an exploration strategy or a combination of memory decay and exploration.

It was observed that the online working memory tests presented during learning increased the variability of the next reaching movement, but people with better working memory performance were less affected. To better understand the nature of this effect, we examined whether the subsequent change in movement direction followed the direction of the rotation introduced by the robot or the direction indicated in the participant's verbal response. We found that the reaching direction did not shift toward the rotation direction presented in the memory test. Instead, it was biased in a direction opposite to participants' judgements, which may indicate an attempt to correct for the presumed direction of rotation (Fig. 3H). Because the online task required participants to make a perceptual judgment, this finding is consistent with the idea that the perceptual judgment appeared to introduce a bias in planning the next immediate movement. Such finding may be due to a top-down influence on the motor system by the prefrontal neurons (Cisek 2007).

Recent work in reinforcement-based motor learning has focused on the involvement of awareness, exploration, and explicit processes (Cashaback et al. 2017; Codol et al. 2018; Holland et al. 2018; Manley et al. 2014; McDougale et al. 2016; Pekny et al. 2015; Therrien et al. 2016), both in clinical and healthy populations, as well as in learning-related brain plasticity (Sidarta et al. 2016). Several of these studies incorporate a reward zone that shifts either gradually (Holland et al. 2018; Pekny et al. 2015) or dynamically based on the performance in previous trials (Therrien et al. 2016). In contrast, the size and position of the reinforced direction in the present study were

fixed. The advantage of using a constant target width is the consistency of the environment (and task demand) across blocks and, accordingly, the ability to interpret differences in variables such as movement accuracy or the number of reinforced trials with respect to a common reference.

It is possible that the role of working memory is different for a fixed vs. rotated reward zone. Specifically, previous work showed that spatial working memory capacity predicts explicit visuomotor adaptation (Christou et al. 2016) and that learning to aim to a shifting reward zone is dominated by explicit processes (Codol et al. 2018; Holland et al. 2018). Therefore, one might expect that a shifting reward zone involves visuospatial working memory components. However, in the present paradigm, there was an average initial bias before learning (average magnitude = 1.57 cm), and the magnitude of the bias significantly reduced following learning. To correct for the bias, participants may rely on spatial rotation, which implies that this fixed target paradigm may share some features with a shifting reward zone.

### *Potential Neural Bases of Somatosensory Working Memory in Human Motor Learning*

Although previous studies have suggested that the dorsolateral prefrontal cortex is implicated in visuospatial working memory in human motor learning (Seidler et al. 2012), the neural substrates of somatosensory working memory for limb configuration in relation to motor learning are less certain. With the use of a sensorimotor learning task similar to that in the present study, it was found that learning-related changes in resting-state function connectivity involved second somatosensory cortex, the right supramarginal gyrus, and right BA 9/46v (Sidarta et al. 2016). These areas are known to be interconnected anatomically. Specifically, studies in nonhuman primates have identified bidirectional projections linking areas PF and PFG in the inferior parietal lobe (supramarginal gyrus in humans), the parietal operculum, and ventral area 46 below the principal sulcus (Gerbella et al. 2013; Petrides and Pandya 2002; Preuss and Goldman-Rakic 1989). Indeed area 9/46v is the only region in prefrontal cortex to project to hand area structures in cortical motor areas, specifically to ventral premotor cortex and pre-smooth muscle actin (Lu et al. 1994; Luppino et al. 1993). In other primate studies, second somatosensory cortex, ventral premotor cortex, supplementary motor area, and the lateral prefrontal cortex have been shown to be involved in somatosensory memory and decision-making tasks involving vibrotactile stimuli (Romo et al. 1999, 2002, 2004). These findings suggest that lateral prefrontal cortex may be involved in online guidance of reaching movements (Goldman-Rakic 1996) by providing motor areas with sensory information stored in working memory.

One potential limitation of the present work is that baseline assessment of participants' somatosensory acuity was not performed. It is possible that the perceptual acuity may have an effect on our estimate of memory performance. Nevertheless, regardless of whether memory items were spaced far apart (10° spacing, as in the offline task) or close together (5°, as in the online task), memory performance was able to predict learning. This suggests that the memory tasks capture a type of memory that is largely invariant to spatial scale. Presumably, this in turn

makes it less likely that somatosensory acuity influences working memory estimates.

In conclusion, two experiments are presented here that provide evidence for the idea that somatosensory working memory supports reinforcement-based motor learning in humans. In the future, it would be desirable to directly modulate neural activity in areas in frontal and prefrontal cortex that contribute to working memory to assess their individual contributions to human motor learning.

### ACKNOWLEDGMENTS

We thank David St-Amand and Katrine Bergeron for assistance.

### GRANTS

This work was supported by the National Institute of Child Health and Human Development R01 HD075740, Les Fonds Québécois de la Recherche sur la Nature et les Technologies, Québec (FQRNT), and the Natural Sciences and Engineering Research Council of Canada (NSERC).

### DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

### AUTHOR CONTRIBUTIONS

A.S. performed experiments; A.S. analyzed data; A.S., F.v.V., and D.J.O. interpreted results of experiments; A.S. prepared figures; A.S., F.v.V., and D.J.O. drafted manuscript; A.S., F.v.V., and D.J.O. edited and revised manuscript; A.S., F.v.V., and D.J.O. approved final version of manuscript.

### REFERENCES

- Abe M, Schambra H, Wassermann EM, Luckenbaugh D, Schweighofer N, Cohen LG. Reward improves long-term retention of a motor memory through induction of offline memory gains. *Curr Biol* 21: 557–562, 2011. doi:10.1016/j.cub.2011.02.030.
- Anguera JA, Reuter-Lorenz PA, Willingham DT, Seidler RD. Contributions of spatial working memory to visuomotor learning. *J Cogn Neurosci* 22: 1917–1930, 2010. doi:10.1162/jocn.2009.21351.
- Anguera JA, Bernard JA, Jaeggi SM, Buschkuhl M, Benson BL, Jennett S, Humfleet J, Reuter-Lorenz PA, Jonides J, Seidler RD. The effects of working memory resource depletion and training on sensorimotor adaptation. *Behav Brain Res* 228: 107–115, 2012. doi:10.1016/j.bbr.2011.11.040.
- Bernardi NF, Darainy M, Ostry DJ. Somatosensory contribution to the initial stages of human motor learning. *J Neurosci* 35: 14316–14326, 2015. doi:10.1523/JNEUROSCI.1344-15.2015.
- Bo J, Seidler RD. Visuospatial working memory capacity predicts the organization of acquired explicit motor sequences. *J Neurophysiol* 101: 3116–3125, 2009. doi:10.1152/jn.00006.2009.
- Bo J, Jennett S, Seidler RD. Working memory capacity correlates with implicit serial reaction time task performance. *Exp Brain Res* 214: 73–81, 2011. doi:10.1007/s00221-011-2807-8.
- Brashers-Krug T, Shadmehr R, Bizzi E. Consolidation in human motor memory. *Nature* 382: 252–255, 1996. doi:10.1038/382252a0.
- Cashaback JGA, McGregor HR, Mohatarem A, Gribble PL. Dissociating error-based and reinforcement-based loss functions during sensorimotor learning. *PLOS Comput Biol* 13: e1005623, 2017. doi:10.1371/journal.pcbi.1005623.
- Chapman CD, Heath MD, Westwood DA, Roy EA. Memory for kinesthetically defined target location: evidence for manual asymmetries. *Brain Cogn* 46: 62–66, 2001. doi:10.1016/S0278-2626(01)80035-X.
- Christou AI, Miall RC, McNab F, Galea JM. Individual differences in explicit and implicit visuomotor learning and working memory capacity. *Sci Rep* 6: 36633, 2016. doi:10.1038/srep36633.
- Cisek P. Cortical mechanisms of action selection: the affordance competition hypothesis. *Philos Trans R Soc Lond B Biol Sci* 362: 1585–1599, 2007. doi:10.1098/rstb.2007.2054.



- Codol O, Holland PJ, Galea JM.** The relationship between reinforcement and explicit control during visuomotor adaptation. *Sci Rep* 8: 9121, 2018. doi:10.1038/s41598-018-27378-1.
- Fiehler K, Burke M, Engel A, Bien S, Rösler F.** Kinesthetic working memory and action control within the dorsal stream. *Cereb Cortex* 18: 243–253, 2008. doi:10.1093/cercor/bhm071.
- Galea JM, Mallia E, Rothwell J, Diedrichsen J.** The dissociable effects of punishment and reward on motor learning. *Nat Neurosci* 18: 597–602, 2015. doi:10.1038/nn.3956.
- Gaspar P, Stepniewska I, Kaas JH.** Topography and collateralization of the dopaminergic projections to motor and lateral prefrontal cortex in owl monkeys. *J Comp Neurol* 325: 1–21, 1992. doi:10.1002/cne.903250102.
- Gerbella M, Borra E, Tonelli S, Rozzi S, Luppino G.** Connectional heterogeneity of the ventral part of the macaque area 46. *Cereb Cortex* 23: 967–987, 2013. doi:10.1093/cercor/bhs096.
- Goble DJ, Lewis CA, Brown SH.** Upper limb asymmetries in the utilization of proprioceptive feedback. *Exp Brain Res* 168: 307–311, 2006. doi:10.1007/s00221-005-0280-y.
- Goldman-Rakic PS.** Regional and cellular fractionation of working memory. *Proc Natl Acad Sci USA* 93: 13473–13480, 1996. doi:10.1073/pnas.93.24.13473.
- Gong M, Li S.** Learned reward association improves visual working memory. *J Exp Psychol Hum Percept Perform* 40: 841–856, 2014. doi:10.1037/a0035131.
- Holland P, Codol O, Galea JM.** Contribution of explicit processes to reinforcement-based motor learning. *J Neurophysiol* 119: 2241–2255, 2018. doi:10.1152/jn.00901.2017.
- Jones SA, Henriques DY.** Memory for proprioceptive and multisensory targets is partially coded relative to gaze. *Neuropsychologia* 48: 3782–3792, 2010. doi:10.1016/j.neuropsychologia.2010.10.001.
- Kaas AL, van Mier H, Goebel R.** The neural correlates of human working memory for haptically explored object orientations. *Cereb Cortex* 17: 1637–1649, 2007. doi:10.1093/cercor/bhl074.
- Kane MJ, Hambrick DZ, Tuholski SW, Wilhelm O, Payne TW, Engle RW.** The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *J Exp Psychol Gen* 133: 189–217, 2004. doi:10.1037/0096-3445.133.2.189.
- Kennerley SW, Wallis JD.** Reward-dependent modulation of working memory in lateral prefrontal cortex. *J Neurosci* 29: 3259–3270, 2009. doi:10.1523/JNEUROSCI.5353-08.2009.
- Klink PC, Jeurissen D, Theeuwes J, Denys D, Roelfsema PR.** Working memory accuracy for multiple targets is driven by reward expectation and stimulus contrast with different time-courses. *Sci Rep* 7: 9082, 2017. doi:10.1038/s41598-017-08608-4.
- Krakauer JW, Ghilardi M-F, Ghez C.** Independent learning of internal models for kinematic and dynamic control of reaching. *Nat Neurosci* 2: 1026–1031, 1999. doi:10.1038/14826.
- Krawczyk DC, Gazzaley A, D’Esposito M.** Reward modulation of prefrontal and visual association cortex during an incentive working memory task. *Brain Res* 1141: 168–177, 2007. doi:10.1016/j.brainres.2007.01.052.
- Leon MI, Shadlen MN.** Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron* 24: 415–425, 1999. doi:10.1016/S0896-6273(00)80854-5.
- Lu MT, Preston JB, Strick PL.** Interconnections between the prefrontal cortex and the premotor areas in the frontal lobe. *J Comp Neurol* 341: 375–392, 1994. doi:10.1002/cne.903410308.
- Luppino G, Matelli M, Camarda R, Rizzolatti G.** Corticocortical connections of area F3 (SMA-proper) and area F6 (pre-SMA) in the macaque monkey. *J Comp Neurol* 338: 114–140, 1993. doi:10.1002/cne.903380109.
- Manley H, Dayan P, Diedrichsen J.** When money is not enough: awareness, success, and variability in motor learning. *PLoS One* 9: e86580, 2014. doi:10.1371/journal.pone.0086580.
- McDougle SD, Boggess MJ, Crossley MJ, Parvin D, Ivry RB, Taylor JA.** Credit assignment in movement-dependent reinforcement learning. *Proc Natl Acad Sci USA* 113: 6797–6802, 2016. doi:10.1073/pnas.1523669113.
- Pekny SE, Izawa J, Shadmehr R.** Reward-dependent modulation of movement variability. *J Neurosci* 35: 4015–4024, 2015. doi:10.1523/JNEUROSCI.3244-14.2015.
- Petrides M, Pandya DN.** Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *Eur J Neurosci* 16: 291–310, 2002. doi:10.1046/j.1460-9568.2001.02090.x.
- Preuschhof C, Heekeren HR, Taskin B, Schubert T, Villringer A.** Neural correlates of vibrotactile working memory in the human brain. *J Neurosci* 26: 13231–13239, 2006. doi:10.1523/JNEUROSCI.2767-06.2006.
- Preuss TM, Goldman-Rakic PS.** Connections of the ventral granular frontal cortex of macaques with perisylvian premotor and somatosensory areas: anatomical evidence for somatic representation in primate frontal association cortex. *J Comp Neurol* 282: 293–316, 1989. doi:10.1002/cne.902820210.
- Romo R, Brody CD, Hernández A, Lemus L.** Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature* 399: 470–473, 1999. doi:10.1038/20939.
- Romo R, Hernández A, Zainos A, Lemus L, Brody CD.** Neuronal correlates of decision-making in secondary somatosensory cortex. *Nat Neurosci* 5: 1217–1225, 2002. doi:10.1038/nn950.
- Romo R, Hernández A, Zainos A.** Neuronal correlates of a perceptual decision in ventral premotor cortex. *Neuron* 41: 165–173, 2004. doi:10.1016/S0896-6273(03)00817-1.
- Seidler RD, Bo J, Anguera JA.** Neurocognitive contributions to motor skill learning: the role of working memory. *J Mot Behav* 44: 445–453, 2012. doi:10.1080/00222895.2012.672348.
- Shmuelof L, Huang VS, Haith AM, Delnicki RJ, Mazzoni P, Krakauer JW.** Overcoming motor “forgetting” through reinforcement of learned actions. *J Neurosci* 32: 14617–14621, 2012. doi:10.1523/JNEUROSCI.2184-12.2012.
- Sidarta A, Vahdat S, Bernardi NF, Ostry DJ.** Somatic and reinforcement-based plasticity in the initial stages of human motor learning. *J Neurosci* 36: 11682–11692, 2016. doi:10.1523/JNEUROSCI.1767-16.2016.
- Stoeckel MC, Weder B, Binkofski F, Buccino G, Shah NJ, Seitz RJ.** A fronto-parietal circuit for tactile object discrimination: an event-related fMRI study. *Neuroimage* 19: 1103–1114, 2003. doi:10.1016/S1053-8119(03)00182-4.
- Therrien AS, Wolpert DM, Bastian AJ.** Effective reinforcement learning following cerebellar damage requires a balance between exploration and motor noise. *Brain* 139: 101–114, 2016. doi:10.1093/brain/awv329.
- Therrien AS, Wolpert DM, Bastian AJ.** Increasing motor noise impairs reinforcement learning in healthy individuals. *eNeuro* 5: 5, 2018. doi:10.1523/ENEURO.0050-18.2018.
- Tong C, Wolpert DM, Flanagan JR.** Kinematics and dynamics are not represented independently in motor working memory: evidence from an interference study. *J Neurosci* 22: 1108–1113, 2002. doi:10.1523/JNEUROSCI.22-03-01108.2002.
- Tsujimoto S, Sawaguchi T.** Neuronal representation of response-outcome in the primate prefrontal cortex. *Cereb Cortex* 14: 47–55, 2004. doi:10.1093/cercor/bhg090.
- Wechsler D. *Wechsler Adult Intelligence Scale* (3rd ed.). San Antonio, TX: Psychological Corporation, 1999.
- Williams SM, Goldman-Rakic PS.** Widespread origin of the primate mesofrontal dopamine system. *Cereb Cortex* 8: 321–345, 1998. doi:10.1093/cercor/8.4.321.
- Wu HG, Miyamoto YR, Gonzalez Castro LN, Ölveczky BP, Smith MA.** Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nat Neurosci* 17: 312–321, 2014. doi:10.1038/nn.3616.
- Zatorre RJ, Salimpoor VN.** From perception to pleasure: music and its neural substrates. *Proc Natl Acad Sci USA* 110, Suppl 2: 10430–10437, 2013. doi:10.1073/pnas.1301228110.