Exp Brain Res (2011) 209:181–192
DOI 10.1007/s00221-011-2540-3

RESEARCH ARTICLE

Observation learning versus physical practice leads to different consolidation outcomes in a movement timing task

Maxime Trempe · Maxime Sabourin ·
Hassan Rohbanfard · Luc Proteau

Received: 6 August 2010 / Accepted: 25 December 2010 / Published online: 30 January 2011

Abstract Motor learning is a process that extends beyond training sessions. Specifically, physical practice triggers a series of physiological changes in the CNS that are regrouped under the term "consolidation" (Stickgold and Walker 2007). These changes can result in between-session improvement or performance stabilization (Walker 2005). In a series of three experiments, we tested whether consolidation also occurs following observation. In Experiment 1, participants observed an expert model perform a sequence of arm movements. Although we found evidence of observation learning, no significant difference was revealed between participants asked to reproduce the observed sequence either 5 min or 24 h later (no between-session improvement). In Experiment 2, two groups of participants observed an expert model perform two distinct movement sequences (A and B) either 10 min or 8 h apart; participants then physically performed both sequences after a 24-h break. Participants in the 8-h group performed Sequence B less accurately compared to participants in the 5-min group, suggesting that the memory representation of the first sequence had been stabilized and that it interfered with the learning of the second sequence. Finally, in Experiment 3, the initial observation phase was replaced by a physical practice phase. In contrast with the results of Experiment 2, participants in the 8-h group performed Sequence B significantly more accurately compared to participants in the 5-min group. Together, our results suggest that the memory representation of a skill learned through observation undergoes consolidation. However, consolidation of an observed motor skill leads to distinct behavioural outcomes in comparison with physical practice.

Keywords Motor learning · Consolidation · Observation · Off-line learning · Stabilization

Physical practice has long been regarded as the single most important determinant of motor skill acquisition. This belief was often expressed by the old adage "practice makes perfect," as advocated by early models of motor skill learning (Crossman 1959; Fitts 1964; Adams 1971; Schmidt 1975; Shea and Morgan 1979). Although the importance of physical practice for motor skill learning is undeniable, recent evidence demonstrates that important processes take place between practice sessions. Specifically, physical practice triggers a series of physiological changes in the brain, from protein synthesis to the formation of new synapses (McGaugh 2000), leading to the long-term retention of the new skill. Regrouped under the term "consolidation" (Stickgold and Walker 2007), these changes require time to occur and constitute the foundation of motor skill learning.

Current hypothesis states that brain networks activated during physical practice are reactivated during rest (Hoffman and McNaughton 2002) and/or a sleep interval (Wilson and McNaughton 1994; Ji and Wilson 2007). This reactivation is thought to favour a reorganization of the information in the CNS (see however Tononi and Cirelli 2003; Hill et al. 2008 for a different view), leading to the activation of different brain networks when the newly practiced motor skill is retested following a consolidation interval (Karni et al. 1995; Shadmehr and Holcomb 1997; Walker et al. 2005).

M. Trempe · M. Sabourin · H. Rohbanfard · L. Proteau (✉)
Département de kinésiologie, Université de Montréal,
2100 Boul. Édouard Montpetit, Montréal,
QUÉBEC H3T 1J4, Canada
e-mail: luc.proteau@umontreal.ca

Although this reorganization of the CNS does not necessarily result in a behaviourally observable change in performance (Karni et al. 1995; Shadmehr and Holcomb 1997), consolidation has often been associated with a stabilization of the participant's performance (Walker 2005; Krakauer and Shadmehr 2006). This outcome has mainly been observed in visuomanual adaptation studies in which participants adapted their reaching movements to compensate for either a rotation of the visual feedback or new forces applied to the hand (Task A). Although adaptation occurred during initial training, poor retention was observed if a second and opposed perturbation (rotation or force-field; Task B) was practiced immediately following Task A, whereas retention was hardly affected if Tasks A and B were practiced several hours apart (Brashers-Krug et al. 1996; Krakauer et al. 1999). Thus, the memory representation acquired through practice is initially kept in a labile form until it is stabilized by the consolidation processes and becomes resistant to different sources of interference.

In addition, consolidation has also been associated with a spontaneous performance increase (i.e., off-line learning) without additional training between the practice and retest sessions (Robertson et al. 2004; Press et al. 2005; Walker and Stickgold 2005; Walker et al. 2005). For example, when participants practiced a sequence of finger movements to reproduce it as quickly and accurately as possible, they were typically faster and made fewer errors when retested following a night of sleep, even if no additional training took place between the practice and retest sessions.

In the past, consolidation has been studied using mainly physical practice tasks. However, is physical practice a prerequisite for motor skill consolidation? To our knowledge, only four reports have investigated the consolidation processes following either observation (Van Der Werf et al. 2009) or motor imagery (Debarnot et al. 2009a, 2009b, 2010). In all these reports, a consolidation interval including sleep resulted in a significant increase in performance. However, the initial acquisition session also included either physical practice (Debarnot et al. 2009a, 2009b, 2010) or contractions of the muscles used to perform the task (Van Der Werf et al. 2009), therefore making it difficult to determine whether consolidation was triggered uniquely by observation/motor imagery. Therefore, whether physical practice is a prerequisite for obtaining behavioural evidence of motor skill consolidation remains an open question. In a series of three experiments, we investigated whether the CNS consolidates a motor skill learned solely through observation, without any form of physical practice. To reach our goal, we sought evidence of the two most common behavioural outcomes resulting from consolidation: off-line learning (Experiment 1) and performance stabilization (Experiment 2). Observing off-line learning and/or performance stabilization would indicate that the memory representation learned through observation has been consolidated. In a third experiment, we contrasted our results with those obtained with physical practice tasks.

Experiment 1

Experiment 1 was designed to determine whether a rest interval following an observation session can result in off-line learning. To do so, we used a timing task that consisted in producing a sequence of arm movements to knock down a series of wooden barriers in a prescribed movement time. This task was chosen because its small accuracy requirements places the focus of what has to be learned on its timing demands, and previous reports have shown that participants can learn the correct movement time simply by observing a model performing the task (Blandin et al. 1999; Blandin and Proteau 2000). In the present experiment, two groups of participants first observed a video of an expert model performing 40 near-perfect trials; participants then physically performed the task either 5 min (5-min group) or 24 h (24-h group) after observation. If the memory representation acquired through observation is sufficiently improved by consolidation to result in an observable behavioural outcome, participants of the 24-h group should demonstrate better performance during the physical practice phase compared to participants of the 5-min group; they should also perform better compared to participants of a control group performing the task without prior observation.

Methods

Participants

Thirty-nine undergraduate students (mean age = 20.5, SE = 0.4; 29 women) from the Département de kinésiologie at the Université de Montréal took part in the experiment. All subjects were naive to the purpose of the study and had no prior experience with the task. None of them reported neurological disorders, and they all had normal or corrected-to-normal vision. The study was approved by the Health Sciences Research Ethic Committee of the Université de Montréal.

Task and apparatus

The apparatus was similar to that used by Blandin et al. (1999) and is illustrated in Fig. 1. It was placed on a table (74 cm high) and was composed of a wooden base

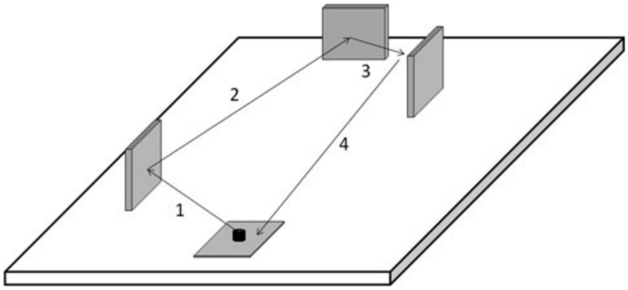

Fig. 1 Movement pattern used in all three experiments (Sequence A). Participants had to press the start microswitch and then hit the first, second and third wooden barriers with their right hand before ending their movement by pushing down on the metal plate surrounding the microswitch. Each segment had to be completed in 300 ms

(46 × 53 cm), a start/stop assembly, and three wooden barriers (11.5 × 8.25 cm). Participants sat in front of the apparatus. Aligned with their body midline was a microswitch that served as a starting position. At the beginning of each trial, the wooden barriers were placed perpendicular (vertical) to the wooden base. Participants had to press the microswitch, knock down the three barriers in a clockwise motion with their right hand, and then push down on a metal plate (11.5 × 8.25 cm) surrounding the microswitch (see Fig. 1). The movements' mean index of difficulty was 2.4 (Fitts 1954). The start/stop assembly and the barriers were connected to a computer via the I/O port of an A-D converter (National Instrument). A millisecond timer, which was activated when the participants pressed the start microswitch, recorded the time at which each barrier was knocked down and stopped when the metal plate was pushed down. This enabled us to record the total movement time, i.e., the time elapsed from the pressing of the microswitch to the pushing down of the metal plate, as well as the time needed to complete each of the four segments of the task.

During the observation phase, participants sat on a chair and watched a video displayed on a 46-inch LCD monitor (Sony KDL-46XBR4). The monitor was located two metres in front of the participants.

Procedures

In the retention phase, participants had to complete the entire movement sequence in 1200 ms. In addition, each segment of the sequence had to be completed in 300 ms. The experimental movement pattern, the total movement time and the time to complete each segment of the task were illustrated on a poster located directly in front of the apparatus (see Fig. 1), which was present during all experimental phases. Participants were instructed to use their right hand to perform the task.

Participants first observed a video of an expert model performing 40 near-perfect trials of the movements sequence (absolute error = 13.0 ms, SE = 1.4; root mean square error = 0.5, SE = 0.03; see below for details concerning the calculations). They were asked to observe the video attentively to learn the correct timing of the sequence. They were explicitly informed that they would have to physically perform the sequence after the observation session. After each trial performed by the model, the time taken to produce the entire sequence (total movement time) as well as the time taken to complete each segment of the sequence were displayed on the monitor to provide feedback to the participants. To ensure that participants remained attentive, the experimenter frequently asked the participants to comment on the performance of the model before the feedback was displayed (through questions such as "Was the previous trial performed too slowly? Too quickly?"). Importantly, participants were asked to keep their hands on their thighs during observation and to avoid moving their arms. They were clearly instructed not to try to reproduce the movements while watching the model. Thereafter, participants performed a retention test (20 trials without feedback) either 5 min (5-min group, $n = 13$) or 24 h (24-h group, $n = 13$) after the observation phase. In addition, a third group performed the retention test without prior observation (control group, $n = 13$).

Testing sessions were scheduled between 8 a.m. and 5 p.m. Participants of the 24-h group were instructed to continue with their usual activities between the sessions. They were asked to maintain their normal sleep schedule and to avoid consuming alcoholic beverages or recreational drugs. Compliance with the instructions was confirmed verbally by the participants at the beginning of the second session. Participants also completed a written questionnaire to report how many hours they slept during the night between the observation session and the retention test. In average, participants of the 24-h group slept 7.7 h (SE = 0.3).

Data reduction

To determine whether participants produced the sequence in the prescribed movement time, we measured the total movement time (i.e., the time elapsed from pressing the start microswitch to pushing down on the metal plate) of all trials performed in the retention test and computed the mean absolute error (AE)[1] and the variable error (VE)[2] for

---

[1] $AE = \sum \frac{|x_i - 1200|}{n}$ where x_i is the total movement time on trial i and n is the total number of sequences executed.

[2] $VE = \sqrt{\sum \frac{(x_i - M)^2}{n}}$ where x_i is the total movement time on trial i, M is the mean total movement time, and n is the total number of sequences executed.

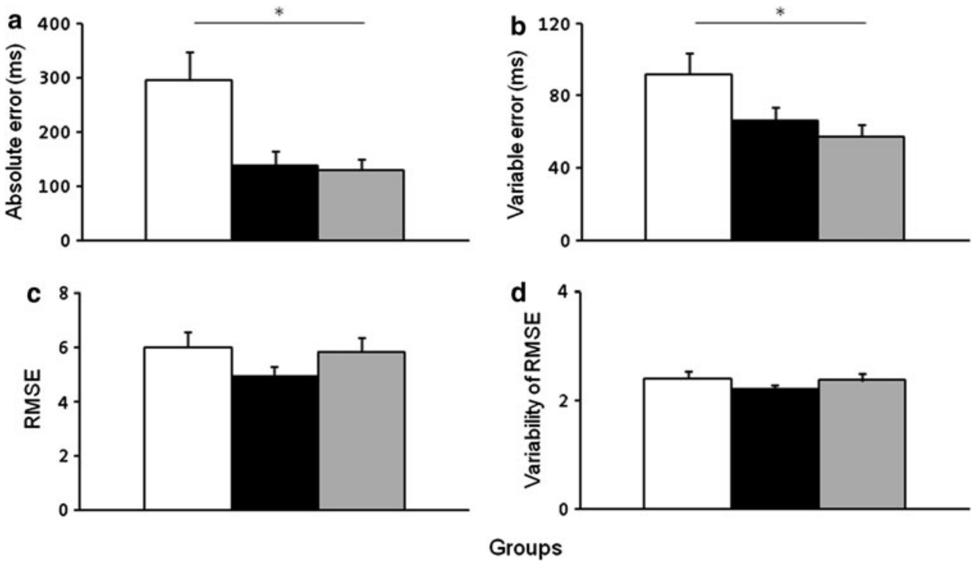

Fig. 2 Experiment 1: Retention data of the control (*white*), 5-min (*black*), and 24-h (*grey*) groups. The symbol * indicates a significant difference between the groups, and the *error bars* illustrate the standard error of the mean

each participant. These measures indicate, respectively, the accuracy and the consistency of the participants' responses. Then, to determine whether participants learned the relative timing of the sequence (i.e., produced all four segments of the sequence in equal time), the time taken to complete each segment of the sequence was expressed as a percentage of the total movement time; this value was used to compute the participants' root mean square error (RMSE; Blandin et al. 1999)[3] and its variability. Trials deviating more than two standard deviations from each participant's mean (for EA and RMSE) were removed from all analyses. Less than 4% of the trials were removed. The low spatial accuracy demands of the task resulted in participants failing to knock down one barrier on less than 1% of the trials. These trials were rejected during the experimental session and immediately re-conducted.

Statistical analyses

Unless mentioned otherwise, data were submitted to 3 Groups one-way ANOVAs. Post hoc comparisons were made using Dunnett's test to compare the performance of the 24-h group with the performances of the 5-min and control groups. To ensure no inflation of type 1 error, we assessed the normality of the distribution by calculating the z score of the skewness and kurtosis values (Tabachnick and Fidell 2007). Hartley's F_{max} test was used to assess the homogeneity of variance of the ANOVAs. When necessary, we used the average of the Greenhouse-Geisser and Huyn-Feldt correction to correct for a possible violation of the sphericity assumption (Stevens 1992). All significant effects are reported at $P < 0.05$.

Results

Total movement time

To determine whether participants learned the total movement time, we first calculated the mean absolute error and contrasted the data using a one-way ANOVA. The ANOVA revealed a significant difference between the groups, $F\ (2,\ 38) = 7.2$, $P = 0.002$, $\eta^2 = 0.29$. As illustrated in Fig. 2a, participants of the 24-h group were significantly more accurate than participants of the control group were ($P = 0.004$) but did not do better than participants of the 5-min group ($P = 0.98$). Furthermore, a second ANOVA contrasting the variable error revealed a significant difference between the groups, $F\ (2,\ 38) = 5.1$, $P = 0.01$, $\eta^2 = 0.22$. As illustrated in Fig. 2b, participants of the 24-h group were significantly less variable compared to participants of the control group ($P = 0.007$), whereas no difference between the 24-h and 5-min groups was observed ($P = 0.65$).

Relative timing

The RMSE and the variability of the RMSE were used to assess whether participants learned the relative timing of the sequence. The ANOVAs revealed no significant difference between the groups, $F(2,\ 38) = 1.3$, $P = 0.3$, $\eta^2 = 0.07$ and $F(2,\ 38) = 0.96$, $P = 0.4$, $\eta^2 = 0.05$, respectively (see Fig. 2c and d). Thus, observation alone

---

[3] $RMSE = \sqrt{\frac{\sum\left(\frac{yi}{TMT}-0.25\right)}{4}}$ where y_i is the movement time of segment i and TMT is the total movement time of the trial.

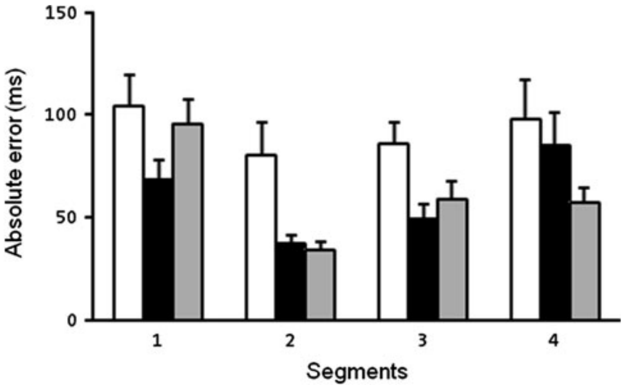

Fig. 3 Experiment 1: Absolute error of the control (*white*), 5-min (*black*) and 24-h (*grey*) groups for each segment of the sequence. The Segment × Group interaction was not significant ($P = 0.32$). The *error bars* illustrate the standard error of the mean

was not sufficient for participants to learn the relative timing of the new sequence.

Movement time of the segments

Recent reports have demonstrated that off-line learning occurs when the initial performance is relatively modest (Kuriyama et al. 2004; Trempe and Proteau 2010). In the present experiment, it is possible that consolidation led to off-line learning but only for the most difficult segment(s) of the sequence. To assess this possibility, we compared the absolute error of each segment of the sequence using a 3 Groups × 4 Segments ANOVA with repeated measurements on the second factor (see Fig. 3). Although the ANOVA revealed a significant main effect of segment, $F(3, 108) = 6.5$, $P = 0.004$, $\eta_p^2 = 0.15$, indicating that some segments were more difficult than others, the Group × Segment interaction did not reach significance, $F(6, 108) = 1.2$, $P = 0.32$, $\eta_p^2 = 0.06$. The ANOVA also revealed a significant main effect of Group, $F(2, 36) = 7.1$, $P = 0.002$, $\eta_p^2 = 028$; participants of the 24-h and 5-min groups outperformed those of the control group but did not differ significantly from one another.

Discussion

The objective of the present report was to determine whether the CNS consolidates a motor skill learned through observation. In the present experiment, we sought evidence of off-line learning by comparing the performance of participants who had either a 5-min or a 24-h break between an observation session and a physical practice session. A third group also performed a physical practice session without prior observation. Two main findings emerged from our results.

First, observation alone (i.e., without physical practice) is sufficient to learn a prescribed movement time. Specifically, participants of the 24-h group were temporally more accurate and less variable compared to participants of the control group. This better performance was noted 24 h after the initial observation phase, suggesting that the memory representation of the correct movement time was stored in long-term memory. This finding is in line with numerous reports showing that observation is beneficial to motor learning (see Ashford et al. 2006 for a meta-analysis). In contrast, observation did not allow participants to learn the relative timing of the new sequence. Specifically, both observation groups (5-min group and 24-h group) did no better than the control group at producing the four segments of the sequence in equal movement time. As previously reported, physical practice seems to be essential to learn the spatiotemporal structure of a new sequence (Blandin et al. 1999).

Second, and more importantly, our data failed to reveal any difference between the 5-min and 24-h groups. Regardless of the interval between the observation session and the physical practice session, participants were equally accurate and variable when they physically performed the sequence. This finding markedly contrasts with prior reports showing off-line learning when the task was physically practiced (Robertson et al. 2004; Press et al. 2005; Walker and Stickgold 2005; Walker et al. 2005; Trempe and Proteau 2010); it also differs from the results of a recent study reporting off-line learning following observation (Van Der Werf et al. 2009; this discrepency will be discussed in the General discussion). In the present experiment, the absence of off-line learning could not be explained by a floor effect because the participants' performance could clearly improve further. Specifically, the mean absolute error of the 24-h group was about 130 ms, whereas errors around 50 ms have been reported when participants physically practiced the task with feedback (Blandin et al. 1999; see also the results of Experiment 3 of the present paper). Furthermore, off-line learning has commonly been associated with a large effect size (with Cohen's d ranging from 1 to 2, Fischer et al. 2002; Walker et al. 2002). In our experiment, the post hoc comparisons had a power of 0.8 (considering an expected effect size of 1, Cohen 1988), which should have been sufficient to detect any significant difference. Also, off-line learning has been associated with performance increase of 20 to 30% (Fischer et al. 2002; Kuriyama et al. 2004); thus, one could have expected the 24-h group to decrease their absolute error to around 100—110 ms. This was clearly not the case as the 24-h group had a mean absolute error of 130 ms when tested the second day. In addition, all participants of the 24-h group had sufficient time (including a night of sleep) to consolidate the new sequence, and they all slept within the first 12-h interval after the video presentation, thus decreasing the possible interference from other daily activities (Van Der Werf et al. 2009).

The observation that a 24-h rest interval did not result in behavioural evidence of off-line learning does not indicate, however, that the memory representation of the movement time was not consolidated. The finding that there was no performance decay 24 h after acquisition suggests that changes did occur in the CNS to store the new memory representation in long-term memory. Thus, it is possible that following observation, consolidation results in performance stabilization.

Experiment 2

The objective of the second experiment was to test the hypothesis that a rest interval following an observation session results in a stabilization of the memory representation learned through observation. Two groups of participants observed videos of an expert model performing two distinct sequences (A and B) either 5 min or 8 h apart (Brashers-Krug et al. 1996). Retention was tested the following day. If consolidation stabilizes the memory representation learned through observation, retention should be better if the two sequences are observed 8 h apart since the memory representation of Sequence A would have become more stable and resistant to the interference of Sequence B (Brashers-Krug et al. 1996; Krakauer et al. 2005).

Participants

Twenty-six undergraduate students (mean age 21.3, *SD* = 1.6; 17 women) from the Département de kinésiologie at the Université de Montréal took part in the experiment. All subjects were naive to the purpose of the study and had no prior experience with the task. None of the subjects reported neurological disorders, and all had normal or corrected to normal vision. The study was approved by the Health Sciences Research Ethic Committee of the Université de Montréal.

Task and apparatus

Sequence A was identical to the sequence used in Experiment 1, whereas Sequence B had the same number of segments (4) but required a different timing (see Fig. 4). Specifically, Sequence B had to be completed in a total movement time of 1600 ms separated into four segments of 450, 350, 450 and 350 ms, respectively. Thus, Sequences A and B shared no timing characteristics. The movements' mean index of difficulty of Sequence B was 2.2 (Fitts 1954).

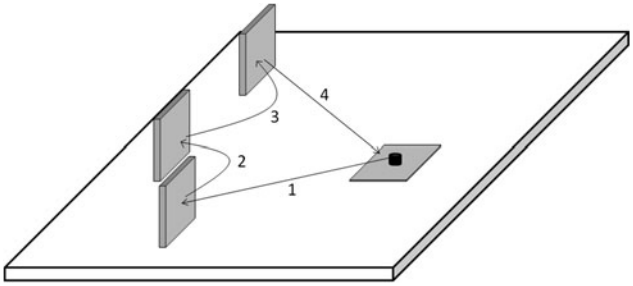

Fig. 4 Movement pattern of Sequence B (Experiments 2 and 3). Participants had to press the starting microswitch and hit the first, second and third wooden barriers with their right hand before ending their movement by pushing down on the metal plate surrounding the microswitch. The four segments had to be completed in 450, 350, 450 and 350 ms, respectively

Procedures

Testing sessions were scheduled between 8 a.m. and 5 p.m. Participants observed a video of an expert model performing 40 near-perfect trials of Sequence A (same video as in Experiment 1), then of Sequence B, either 5 min (5-min group, $n = 13$) or 8 h apart (8-h group, $n = 13$). The mean absolute error of the model performing Sequence B was 11.9 ms (SE = 1.3) and his mean RMSE was 0.44 (SE = 0.02). Participants returned to the laboratory 24 h after the observation of the first sequence and physically performed 20 trials of each sequence without feedback (starting with Sequence A). To reduce possible interference effects between Sequence A and B during the retention test, all participants performed a metronome task at the beginning of the retention test and before switching to the second sequence. The task consisted in pressing a button (1 mm) at regular intervals during a period of 2 min to match the audible beats (1 beat per 3 s) generated by a computer. The rhythm produced by the metronome was completely different from the correct rhythm of both sequences. The remaining procedures were identical to those used in Experiment 1.

Results

Sleep data

Participants slept an average of 7.1 h (*SE* = 0.29) during the night between the practice sessions and the retention test. There was no significant difference between the groups, $t(24) = 0.30$, $P = 0.77$, $d = 0.12$.

Total movement time

As in Experiment 1, we computed the mean absolute error and the variable error of the total movement time for each participant. In addition, we also computed the constant error (CE) to determine whether movement times were biased in a specific way (i.e., too fast or too slow). Data were submitted to three separate ANOVAs contrasting 2 Groups × 2 Sequences with repeated measurements on the second factor.

The ANOVA contrasting the absolute error revealed a significant Group × Sequence interaction, $F(1, 24) = 4.4$, $P = 0.046$, $\eta_p^2 = 0.16$. As illustrated in Fig. 5a, both groups were equally accurate when performing Sequence A ($P = 0.7$), whereas the 8-h group made larger errors compared to the 5-min group when performing Sequence B ($P = 0.049$). The ANOVA contrasting the CE revealed no difference between the groups, $F(1, 24) = 3.2$, $P = 0.09$, $\eta_p^2 = 0.12$, nor a Group × Sequence interaction, $F(1, 24) = 1$, $P = 0.32$, $\eta_p^2 = 0.04$ (see Fig. 5b). Finally, participants of the 8-h group were significantly less variable compared to participants of the 5-min group, regardless of the sequence, $F(1, 24) = 16.9$, $P < 0.001$, $\eta_p^2 = 0.4$ (see Fig. 5c).

Relative timing

As in Experiment 1, we computed the RMSE of relative timing and its variability to determine whether participants learned the relative timing of the movement sequences. For the RMSE, the ANOVA revealed neither a significant difference between the groups, $F(1, 24) = 0.02$, $P = 0.9$, $\eta_p^2 = 0.001$, nor a Group × Sequence interaction, $F(1, 24) = 1.3$, $P = 0.95$, $\eta_p^2 = 0.04$. Similar results were obtained when contrasting the variability of the RMSE, $F(1, 24) = 0.008$, $P = 0.93$, $\eta_p^2 < 0.001$ and $F(1, 24) = 0.5$, $P = 0.48$, $\eta_p^2 = 0.02$ for the main effect of Group and the Group × Sequence interaction, respectively.

Discussion

The objective of Experiment 2 was to determine whether we could find behavioural evidence that a memory representation learned through observation can be stabilized by consolidation. Based on previous reports (Brashers-Krug et al. 1996; Muellbacher et al. 2002; Walker et al. 2003), we hypothesised that an 8-h rest interval between the observation sessions would lead to improved learning since the memory representation of the first sequence (Sequence A) would have had sufficient time to become stable and resistant to the interference of the second sequence (Sequence B).

Interestingly, the results did not support our hypothesis. First, there was no accuracy difference between the groups when participants performed Sequence A. Thus, regardless of the interval between the observation sessions, Sequence B did not interfere with the learning of Sequence A (i.e., no retrograde interference). This result diverges from previous reports using physical practice showing interference when two tasks are practiced successively (Brashers-Krug et al. 1996; Krakauer et al. 1999).

Participants of the 8-h group performed Sequence B less accurately compared to participants of the 5-min group:

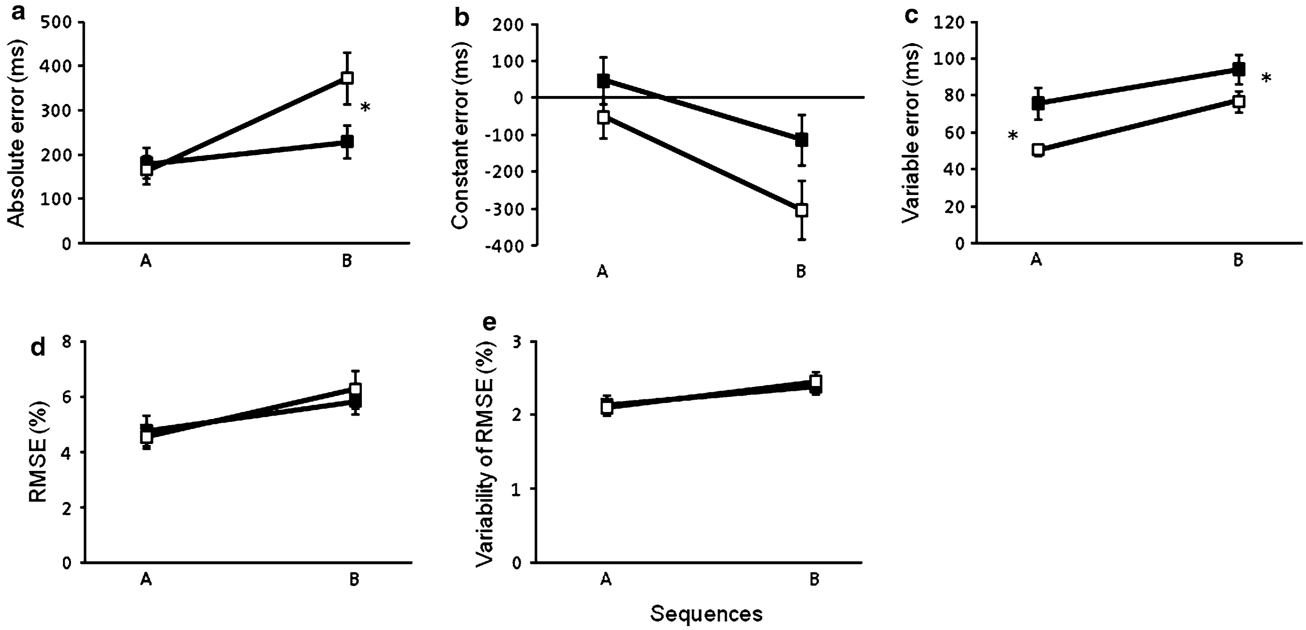

Fig. 5 Experiment 2: Retention data of the 5-min (*filled*) and 8-h (*opened*) groups. The symbol * indicates a significant difference between the groups, and the *error bars* illustrate the standard error of the mean

a longer between-session interval impaired learning of Sequence B. Noteworthy is that participants of the 8-h group performed Sequence B in around 1300 ms, i.e., in a movement time closer to the 1200 ms of Sequence A than to the required 1600 ms for Sequence B. This suggests that the memory representation of Sequence A caused anterograde interference on the learning of Sequence B. Because the two groups differed only by the length of time between the observation sessions, this anterograde interference observed only for the 8-h break group suggests that consolidation processes took place between the observation sessions and stabilized the memory representation of Sequence A. As a result, participants failed to learn the second sequence and simply reproduced both sequences the following day using the stable memory representation of Sequence A. This explanation is further supported by the finding that participants of the 8-h group performed both sequences with less variability (they kept reproducing the same stable movement time). Thus, our results argue that observation triggered consolidation processes that stabilized the memory representation of the new motor skill.

The finding that an 8-h interval resulted in anterograde interference (impaired learning of Sequence B) differs from the retrograde interference previously reported using a physical practice protocol (Brashers-Krug et al. 1996; Walker et al. 2003). However, because our sequence production task was also quite different from the adaptation and finger sequence tasks used in these reports, it is difficult to conclude that observation and physical practice trigger different consolidation processes. Experiment 3 was therefore conducted to determine whether the results of Experiment 2 were specific to observation learning or to our sequence production task.

Experiment 3

To determine whether the results of Experiment 2 were specific to our task, the observation session was replaced by a physical practice session. Participants performed 40 trials with feedback of Sequences A and B either 5 min (5-min group, $n = 12$) or 8 h (8-h group, $n = 12$) apart and were retested the following day. If the results of Experiment 2 were specific to our task, participants of the 8-h group should again demonstrate weaker retention for Sequence B. In contrast, if the results of Experiment 2 were specific to the consolidation processes taking place after observation, participants of the 8-h group should outperform those of the 5-min group in the retention test (as in previous reports using physical practice tasks Brashers-Krug et al. 1996; Walker et al. 2003). None of the participants recruited took part in the previous two experiments. All remaining procedures were identical to those described in Experiment 2.

Results

Acquisition

Acquisition data for EA and RMSE are illustrated in Fig. 6. To assess whether the 5-min and 8-h groups performed

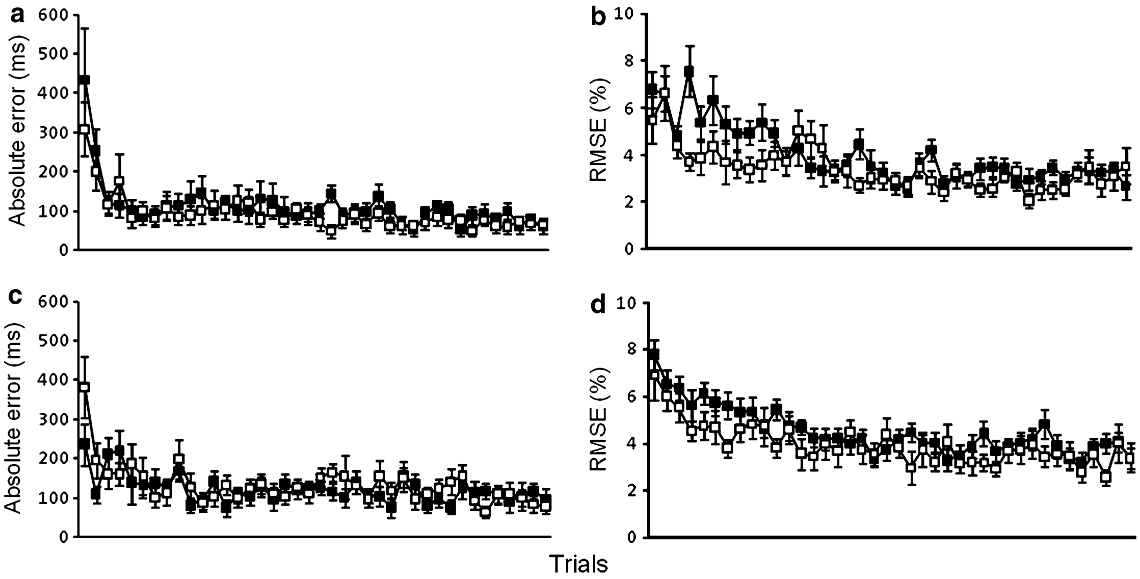

Fig. 6 Experiment 3: Acquisition data of the 5-min (*filled*) and 8-h (*opened*) groups when practicing Sequence A (**a** and **b**) and Sequence B (**c** and **d**). The *error bars* illustrate the standard error of the mean

differently during acquisition, we averaged the last 20 trials of each sequence and calculated EA, EC, EV, RMSE and the variability of RMSE. Data of all dependent variables were then submitted to separate ANOVAs contrasting 2 Groups × 2 Sequences. There was no significant difference between the groups ($P > 0.13$) with the exception that the 8-h group was significantly less variable (EV) compared to the 5-min group when practicing Sequence A ($P < 0.001$, $F(1, 22) = 6.7$, $P = 0.017$, $\eta_p^2 = 0.23$ for the Group × Sequence interaction).

Sleep data

Participants slept on average 7.9 h ($SE = 0.22$) during the night between the practice sessions and the retention test. There was no significant difference between the groups, $t(22) = 0.35$, $P = 0.73$, $d = 0.14$.

Retention

As in Experiment 2, all 20 trials performed during the retention test were used to calculate AE, EC, EV, RMSE and the variability of RMSE for each participant. Data were then submitted to separate 2 Groups × 2 Sequences ANOVAs. The ANOVA contrasting the AE revealed a significant Group × Sequence interaction, $F(1, 22) = 7.5$, $P = 0.012$, $\eta_p^2 = 0.25$. As illustrated in Fig. 7a, although both groups performed Sequence A equally accurately ($P = 0.32$), participants of the 8-h group were significantly more accurate compared to participants of the 5-min group when performing Sequence B ($P = 0.016$). This difference was not caused by a specific bias since the ANOVA contrasting the EC revealed no significant difference between the groups, $F(1, 22) = 0.01$, $P = 0.92$, $\eta_p^2 < 0.001$, nor a Group × Sequence interaction, $F(1, 22) < 0.001$, $P = 0.99$, $\eta_p^2 < 0.001$ (see Fig. 7b). The variability also did not differ between the groups, $F(1, 22) = 0.25$, $P = 0.6$, $\eta_p^2 = 0.01$ and $F(1, 22) = 0.97$, $P = 0.34$, $\eta_p^2 = 0.04$ for the main effect of Group and the Group × Sequence interaction, respectively ($P > 0.34$; see Fig. 7c).

The ANOVAs contrasting RMSE and its variability did not reveal any significant difference between the groups (see Figs. 7d and e).

Discussion

The objective of Experiment 3 was to determine whether the anterograde interference observed in Experiment 2 was specific to the consolidation processes taking place following observation or to the nature of our task. To answer this question, participants physically practiced two distinct sequences of arm movements (A and B) either 5 min or 8 h apart before being retested the following day. As in Experiment 2, participants of both groups showed equal retention of Sequence A. However, participants of the 8-h group performed Sequence B significantly more accurately compared to participants of the 5-min group. Thus, an 8-h rest interval between the physical practice sessions led to better learning. Taken together, the results of Experiment 2

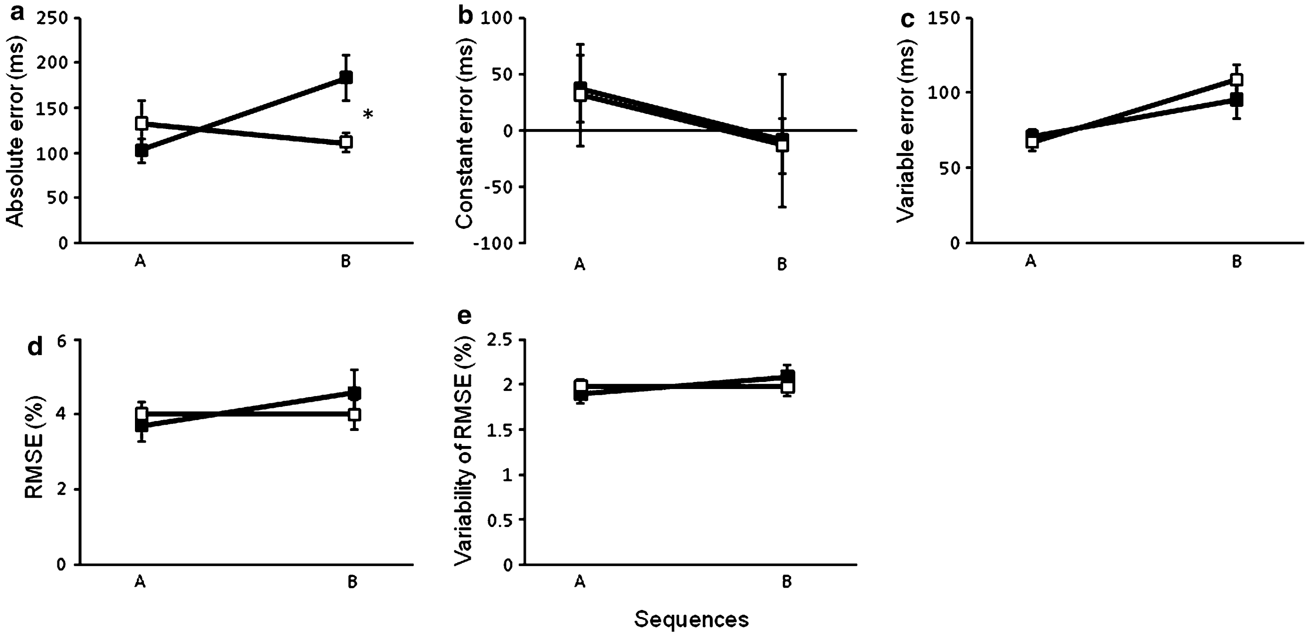

Fig. 7 Experiment 3: Retention data of the 5-min (*filled*) and 8-h (*opened*) groups. The symbol * indicates a significant difference between the groups, and the *error bars* illustrate the standard error of the mean

and 3 indicate that the consolidation processes taking place after observation lead to different behavioural outcomes than those taking place after physical practice.

General discussion

The objective of the present report was to determine whether we could find behavioural evidence that a motor skill learned through observation can be consolidated. To do so, we conducted a series of three experiments in which we sought evidence of off-line learning and stabilization, the two most common behavioural outcomes of the consolidation processes (Walker 2005; Krakauer and Shadmehr 2006). In Experiment 1, no off-line learning was observed. This was the case even when the observation and retention sessions were separated by sufficient time (including a night of sleep) and when the participants' performance could improve still further. Although it could be argued that subtle kinematic differences could have resulted from consolidation (for example, smoother and less variable bell-shape velocity profiles to perform each segment of the task), they clearly had no significant impact on what the participants were asked to do, which makes them secondary to the findings of the present study. In Experiment 1, the observers outperformed the control group in the 24-h retention test, indicating that observation resulted in a lasting representation of the task in the CNS. This proposition is supported by the results of Experiment 2, in which an 8-h break between the observation of two distinct movement sequences resulted in anterograde interference, suggesting that the memory representation of the first sequence (Sequence A) had been stabilized and kept in memory during the inter-session interval. Together, the results of Experiments 1 and 2 indicate that specific consolidation processes do occur following an observation session. To our knowledge, this is the first time that it is unequivocally shown that physical practice is not a prerequisite for motor skill consolidation.

Furthermore, the results of all three experiments demonstrate that the behavioural outcomes of the consolidation processes taking place after observation differ from those taking place after physical practice. This is particularly evident when comparing the results of Experiments 2 and 3, in which the same consolidation interval led to opposite results depending on whether the new sequence of movements was initially observed or physically practiced. In addition, the absence of off-line learning in Experiment 1 also differs from the results obtained with physical practice tasks. This new finding suggests that observation may trigger different consolidation processes than those triggered by physical practice.

Although speculative, this hypothesis could explain the contradiction between our results and those of Van Der Werf et al. (2009), who reported off-line learning following observation. In their experiment, participants were asked to press two computer keys with two fingers (corresponding to the fingers used in the sequence production task) during observation. This procedure was employed to ensure that participants did not attempt to physically practice the sequence during observation. However, by doing so, networks involved in the muscle contraction were activated during observation. Although participants were not physically practicing the sequence, this activation may have been sufficient to trigger off-line learning processes associated with physical practice. In contrast, participants in our experiment were instructed to sit still with their hands resting on their thighs during the observation phase. Thus, participants produced no muscle contraction. Similarly, this hypothesis could account for the off-line learning reported by Debarnot et al. (2009a, 2009b) following mental imagery. Because participants physically practiced the sequence before the mental imagery session, the spontaneous improvement observed during the re-test session may have reflected consolidation of the short physical practice phase or the joint effects of imagery and physical practice (see Stefan et al. 2008).

Although observation and physical practice are known to share many similarities (Jeannerod 1999), these two acquisition modalities are not identical, thus providing hints to explain the different consolidation outcomes. First, while several neuroimaging studies have shown large overlap in the brain regions activated during observation and physical practice, certain brain regions are nevertheless activated more intensely (Cross et al. 2009), or even exclusively (see Grèzes and Decety 2001 for a meta-analysis) during physical practice. In addition, observation is thought to involve the mirror neuron system (also called the observation action network [AON]), i.e., a subset of neurons that are activated whether the action is produced or observed. Located mainly within the premotor and parietal cortex, these neurons are believed to play a crucial role for our understanding of others' actions (Rizzolatti and Craighero 2004; Iacoboni and Dapretto 2006). However, the primary motor cortex (M1), a key structure for motor skill learning (Sanes and Donoghue 2000), is not known for being part of the AON and may therefore not be engaged in the long-term retention of skills learned by observation. Although M1 has been reported to be important for the short-term retention of a new internal model learned by observation (Brown et al. 2009), both rodent (Kleim et al. 2004) and human (Karni et al. 1995) experiments have reported that long-lasting plastic changes in M1 occur only when the learner's performance reaches an asymptote. Because no movement is produced during observation, no asymptote can be reached; observation may therefore not provide sufficient stimulation to drive a plastic reorganisation of M1. A second, but not exclusive, possibility is that feedback provided to participants during physical practice may be crucial for off-line learning. Specifically, Holroyd and Coles (2002) suggested that the mesencephalic system may modulate its dopaminergic signal in response to the outcome of a movement, thus making this system likely to modulate memory consolidation (Jay 2003). Again, because observers produce no movement, they obviously receive no feedback about their own performance. Consequently, the dopaminergic signal may remain silent during acquisition and fail to trigger specific consolidation processes.

As mentioned in Experiment 2, the finding that an 8-h interval between the observation sessions impaired learning of the second sequence is counterintuitive. One could argue that participants of the 8-h group were simply more tired than participants of the 5-min group were when observing Sequence B (around 4 p.m. for the 8-h group vs. 8:30 a.m. for the 5-min group) and therefore encoded the sequence less efficiently. Two lines of evidence argue against this explanation, however. First, no participant reported excessive fatigue at the time of the second observation session. Moreover, when asked by the experimenter at various points during the observation session to comment on the performance of the model (e.g., “Was the previous movement too fast? Too slow?”), participants of the 8-h group provided answers clearly indicating that they remained attentive to the video. Second, and more importantly, participants of the 8-h group in Experiment 3 also practiced Sequence B around 4 p.m. and still demonstrated better learning compared to participants of the 5-min group. Thus, if the larger error observed in Experiment 2 was caused by fatigue, Experiment 3 should have led to similar results.

Nevertheless, it remains difficult to explain why an 8-h consolidation interval led to opposite results when the sequences were observed or physically practiced. Possibly, information may have been encoded by distinct memory systems (declarative and procedural) depending on the acquisition protocol (see also Kelly et al. 2003 for a similar discussion). Declarative and procedural memories are known to recruit different neural networks and to be consolidated differently (Walker 2005; Robertson and Cohen 2006). In the present report, observation learning most likely relied on declarative memory since participants were consciously trying to identify the correct timing of each sequence. It is therefore possible that the two memory representations did not compete for the same resources when learned successively, that is, when the differences between the two sequences were most salient. As a result, no interference was observed in retention for the 5-min group in Experiment 2. This hypothesis is in line with the results of Debarnot et al. (2010), who reported no interference when two sequences were successively practiced by mental imagery. However, when the two sequences were observed 8 h apart, the stabilized memory representation of the first sequence could have been used as a reference for the acquisition of the second sequence, which would explain why we observed anterograde interference. In contrast, physical practice in Experiment 3 most likely recruited the procedural memory system since the correct rhythm developed through practice can be more easily felt than verbalized. As reported previously (Shadmehr and Holcomb 1999), two procedural skills learned successively seem to compete for the same resources and are therefore particularly subject to interference. This would explain why, in Experiment 3, a 5-min pause impaired learning of the second sequence.

This hypothesis is in line with the results of Kelly et al. (2003), suggesting that observational learning is subject to interference from a secondary task engaging the declarative memory system. Using a sequence-learning task, the authors reported that observers failed to learn the sequence when they were concurrently engaged in a tone-counting task, whereas learning occurred when the observers could devote all their attention to the primary task (sequence learning). In contrast, the secondary task did not prevent sequence learning when the primary task was performed physically. Further work is still necessary, however, to determine whether physical practice following observation (and vice versa) interferes with the consolidation processes.

In conclusion, our results indicate that observation does trigger consolidation processes that lead to a stabilization of the new motor skill and its long-term retention. Although observation and physical practice are known to share many similarities, our results indicate that they are consolidated differently.

Acknowledgments This work was supported by a Discovery grant (L.P.) and a scholarship (M. T.) provided by the Natural Sciences and Engineering Research Council of Canada.

References

Adams JA (1971) A closed-loop theory of motor learning. J Mot Behav 3:111–150

Ashford D, Bennett SJ, Davids K (2006) Observational modeling effects for movement dynamics and movement outcome measures across differing task constraints: a meta-analysis. J Mot Behav 38:185–205

Blandin Y, Proteau L (2000) On the cognitive basis of observational learning: development of mechanisms for the detection and correction of errors. Q J Exp Psychol 53A:846–867

Blandin Y, Lhuisset L, Proteau L (1999) Cognitive processes underlying observational learning of motor skills. Q J Exp Psychol 52A:957–979

Brashers-Krug T, Shadmehr R, Bizzi E (1996) Consolidation in human motor memory. Nature 382:252–254

Brown LE, Wilson ET, Gribble PL (2009) Repetitive transcranial magnetic stimulation to the primary motor cortex interferes with motor learning by observing. J Cogn Neurosci 21:1013–1022

Cohen J (1988) Statistical power analysis for the behavioral sciences. Lawrence Erlbaum Associates, New Jersey

Cross ES, Kraemer DJM, Hamilton AFC, Kelley WM, Grafton ST (2009) Sensitivity of the action observation network to physical and observational learning. Cereb Cortex 19:315–326

Crossman ERFW (1959) A theory of the acquisition of speed skill. Ergonomics 2:153–166

Debarnot U, Creveaux T, Collet C, Doyon J, Guillot A (2009a) Sleep contribution to motor memory consolidation: a motor imagery study. Sleep 32:1559–1565

Debarnot U, Creveaux T, Collet C, Gemignani A, Massarelli R, Doyon J, Guillot A (2009b) Sleep-related improvements in motor learning following mental practice. Brain Cogn 69: 398–405

Debarnot U, Maley L, De Rossi D, Guillot A (2010) Motor interference does not impair the memory consolidation of imagined movements. Brain Cogn 74:52–57

Fischer S, Hallschmid M, Elsner AL, Born J (2002) Sleep forms memory for finger skills. Proc Natl Acad Sci USA 99:11987–11991

Fitts PM (1954) Adaptation of aimed arm movements to sensorimotor discordance: evidence for direction-dependant gain control. J Exp Psychol 47:381–391

Fitts PM (1964) Perceptual-motor skills learning. In: Melton AW (ed) Categories of human learning. Academic Press, New York

Grèzes J, Decety J (2001) Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. Hum Brain Mapp 12:1–19

Hill S, Tononi G, Ghilardi AF (2008) Sleep improves the variability of motor performance. Brain Res Bull 76:605–611

Hoffman KL, McNaughton BL (2002) Coordinated reactivation of distributed memory traces in primate neocortex. Science 297:2070–2073

Holroyd CB, Coles MGH (2002) The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol Rev 109:679–709

Iacoboni M, Dapretto M (2006) The mirror neuron system and the consequences of its dysfunction. Nat Rev Neurosci 7:942–951

Jay TM (2003) Dopamine: a potential substrate for synaptic plasticity and memory mechanisms. Prog Neurobiol 69:375–390

Jeannerod M (1999) To act or not to act: perspectives on the representation of actions. Q J Exp Psychol 1999:1–29

Ji D, Wilson MA (2007) Coordinated memory replay in the visual cortex and hippocampus during sleep. Nat Neurosci 10:100–107

Karni A, Meyer G, Jezzard P, Adams MM, Turner R, Ungerleider LG (1995) Functional MRI evidence for adult motor cortex plasticity during motor skill learning. Nature 377:155–158

Kelly SW, Burton AM, Riedel B, Lynch E (2003) Sequence learning by action and observation: evidence for separate mechanisms. Br J Psychol 94:355–372

Kleim JA, Hogg TM, VanderBerg PM, Cooper NR, Bruneau R, Remple M (2004) Cortical synaptogenesis and motor map reorganization occur during late, but not early, phase of motor skill learning. J Neurosci 24:628–633

Krakauer JW, Shadmehr R (2006) Consolidation of motor memory. Trends Neurosci 29:58–64

Krakauer JW, Ghilardi MF, Ghez C (1999) Independent learning of internal models for kinematic and dynamic control of reaching. Nat Neurosci 2:1026–1031

Krakauer JW, Ghez C, Ghilardi MF (2005) Adaptation to visuomotor transformations: consolidation, interference, and forgetting. J Neurosci 25:473–478

Kuriyama K, Stickgold R, Walker MP (2004) Sleep-dependent learning and motor-skill complexity. Learn Mem 11:705–713

McGaugh JL (2000) Memory—a century of consolidation. Science 287:248–251

Muellbacher W, Ziemann U, Wissel J, Dang N, Kofler M, Facchini S, Boroojerdi B, Poewe W, Hallet M (2002) Early consolidation in human primary motor cortex. Nature 415:640–644

Press DZ, Casement MD, Pascual-Leone A, Robertson EM (2005) The time course of off-line motor sequence learning. Cogn Brain Res 25:375–378

Rizzolatti G, Craighero L (2004) The mirror-neuron system. Annu Rev Neurosci 27:169–192

Robertson EM, Cohen DA (2006) Understanding consolidation through the architecture of memories. Neuroscientist 12:261–271

Robertson EM, Pascual-Leone A, Miall RC (2004) Current concepts in procedural consolidation. Nat Rev Neurosci 5:576–582

Sanes JN, Donoghue JP (2000) Plasticity and primary motor cortex. Annu Rev Neurosci 23:393–415

Schmidt RA (1975) A schema theory of discrete motor skill learning. Psychol Rev 82

Shadmehr R, Holcomb HH (1997) Neural correlates of motor memory consolidation. Science 277:821–825

Shadmehr R, Holcomb HH (1999) Inhibitory control of competing motor memories. Exp Brain Res 126:235–251

Shea JB, Morgan RL (1979) Contextual interference effects on the acquisition, retention, and transfer of a motor skill. J Exp Psychol [Hum Learn] 5:179–187

Stefan K, Classen J, Celnik P, Cohen LG (2008) Concurrent action observation modulates practice-induced motor memory formation. Eur J Neurosci 27:730–738

Stevens J (1992) Applied multivariate statistics for the social sciences. L. Erlbaum Associates, Hillsdale, N.J

Stickgold R, Walker MP (2007) Sleep-dependent memory consolidation and reconsolidation. Sleep Med 8:331–343

Tabachnick BG, Fidell LS (2007) Using multivariate statistics. Montreal, Boston

Tononi G, Cirelli C (2003) Sleep and synaptic homeostasis: a hypothesis. Brain Res Bull 62:143–150

Trempe M, Proteau L (2010) Distinct consolidation outcomes in a visuomotor adaptation task: off-line leaning and persistent after-effect. Brain Cogn 73:135–145

Van Der Werf YD, Van Der Helm E, Schoonheim MM, Ridderikhoff A, Van Someren EJW (2009) Learning by observation requires an early sleep window. Proc Natl Acad Sci USA 106:18926–18930

Walker MP (2005) A refined model of sleep and the time course of memory formation. Behav Brain Sci 28:51–104

Walker MP, Stickgold R (2005) It's practice, with sleep, that makes perfect: implications of sleep-dependent learning and plasticity for skill performance. Clin Sports Med 24:301–317

Walker MP, Brakefield T, Morgan A, Hobson JA, Stickgold R (2002) Practice with sleep makes perfect: sleep-dependent motor skill learning. Neuron 35:205–211

Walker MP, Brakefield T, Hobson JA, Stickgold R (2003) Dissociable stages of human memory consolidation and reconsolidation. Nature 425:616–620

Walker MP, Stickgold R, Alsop D, Gaab N, Schlaug G (2005) Sleep-dependent motor memory plasticity in the human brain. Neuroscience 133:911–917

Wilson MA, McNaughton BL (1994) Reactivation of hippocampal ensemble memories during sleep. Science 265:676–679