Reward timing matters in motor learning

Pierre Vassiliadis[1,2]*, Aegryan Lete[1], Julie Duque[1], Gerard Derosiere[1]

[1] Institute of Neuroscience, Université Catholique de Louvain, 1200, Brussels, Belgium.

[2] Defitech Chair for Clinical Neuroengineering, Center for Neuroprosthetics (CNP) and Brain Mind Institute (BMI), Swiss Federal Institute of Technology (EPFL), 1202, Geneva, Switzerland

*** CORRESPONDENCE TO:**

Pierre Vassiliadis

Institute of Neuroscience

Université catholique de Louvain

53, Avenue Mounier

COSY- B1.53.04

B-1200 BRUSSELS

Belgium

Email: pierre.vassiliadis@uclouvain.be

Web: http://www.coactionslab.com

Abstract

Reward can improve motor learning and the consolidation of motor memories. Identifying the features of reward feedback that are critical for motor learning is a necessary step for successful integration into rehabilitation programs. One central feature of reward feedback that may affect motor learning is its timing – that is, the delay after which reward is delivered following movement execution. In fact, research on associative learning has shown that short and long reward delays (*e.g.*, 1 and 6 s following action execution) activate preferentially the striatum and the hippocampus, respectively, which both contribute with varying degrees to motor learning. Given the distinct functional role of these two areas, we hypothesized that reward timing could modulate how people learn and consolidate a new motor skill. In sixty healthy participants, we found that delaying reward delivery by a few seconds influenced motor learning dynamics. Indeed, training with a short reward delay (*i.e.,* 1 s) induced slow, yet continuous gains in performance, while a long reward delay (*i.e.,* 6 s) led to initially high learning rates that were followed by an early plateau in the learning curve and a lower endpoint performance. Moreover, participants who successfully learned the skill with a short reward delay displayed overnight consolidation, while those who trained with a long reward delay exhibited an impairment in the consolidation of the motor memory. Overall, our data show that reward timing affects motor learning, potentially by modulating the engagement of different learning processes, a finding that could be exploited in future rehabilitation programs.

Keywords:

Motor learning, motor skill learning, memory consolidation, reward, motivation, reinforcement learning, timing.

1. Introduction

When delivered following well-executed movements, reward can boost motor learning (Chen et al., 2017; Dhawale et al., 2017; Galea et al., 2015; Vassiliadis et al., 2021) and the consolidation of motor memories (Abe et al., 2011). This observation has raised hope for rehabilitation, where reward is regarded as a promising means to magnify the positive effects of practice on motor control (Quattrocchi et al., 2017; Therrien et al., 2020, 2016; Vassiliadis et al., 2019; Vassiliadis and Derosiere, 2020). Yet, this branch of research is only burgeoning, and a current challenge in the field is to identify the features of reward feedback that may be critical for motor learning.

Recent studies have started to tackle this issue, showing that the magnitude (Vassiliadis et al., 2021), the valence (Galea et al., 2015) and the stochasticity (Dayan et al., 2014) of reward feedback bear all a decisive impact on motor learning. Another key feature of reward feedback that may directly affect motor learning is its timing – that is, the delay after which reward is delivered following movement execution. As such, previous studies have shown that reward prediction error signals, which are key for reward-based learning, are not only modulated by the value of the reward but also depend on the timing at which it is delivered (Fiorillo et al., 2008; Klein-Flügge et al., 2011; Kobayashi and Schultz, 2008). Moreover, converging lines of evidence from neuroimaging and electroencephalographic studies indicate that different brain structures – known to be involved with varying degrees in motor learning – exhibit activity changes in response to reward feedback depending on its timing. Indeed, in associative learning tasks, short reward delays (*e.g.*, provided 1 s following action execution) activate a fronto-striatal network, while long reward delays (*e.g.*, 6 s following execution) evoke changes in the activity of the hippocampus primarily (Foerde and

Shohamy, 2011; Peterburs et al., 2016). Further, Parkinson's disease and ADHD patients, both known to exhibit striatal dysfunction (Mehler-Wex et al., 2006), are impaired in learning action-outcome associations based on short reward delays (Foerde and Shohamy, 2011; Gabay et al., 2018; Weismüller et al., 2018), while amnesic patients with damage to the hippocampus are unable to learn associations with long reward delays (Foerde et al., 2013). Altogether, these findings indicate that the processing of reward preferentially engages striatum- or hippocampus-centred networks depending on the timing at which it is delivered.

The striatum and the hippocampus show varying contributions during motor learning and consolidation (Doyon and Benali, 2005; Fernández-Seara et al., 2009; Krakauer et al., 2019; Schendan et al., 2003), which are thought to underlie the operation of distinct learning processes (Albouy et al., 2013, 2008). Hence, it is sensible to assume that reward may boost different motor learning processes – potentially relying on the striatum or the hippocampus – depending on the timing at which it is delivered. Notably, previous studies on reward-based motor learning have only exploited short reward delays, preventing one to test this hypothesis directly. Here, we tested this idea by evaluating the performance of sixty healthy participants in a skill learning task (Vassiliadis et al., 2021), where reward was delivered either at a short or at a long delay following movement execution. We found that delaying reward delivery by a few seconds influenced the dynamics of learning. Indeed, training with a short reward delay induced slow, yet continuous gains in performance, while a long reward delay led to initially high learning rates that were followed by an early plateau in the learning curve and a lower endpoint performance. Moreover, participants who successfully learned the skill with a short reward delay displayed overnight consolidation, while those who trained with a long reward delay exhibited an impairment in the consolidation of

the motor memory. Altogether, the present results provide evidence that reward timing can strongly influence motor learning, a finding that could be exploited in future rehabilitation protocols.

2. Results

Sixty healthy participants practiced a pinch-grip force task over two consecutive days. Participants were required to hold a pinch grip transducer in their right hand and to squeeze it as quickly as possible in order to move a cursor displayed on a computer screen in front of them, from an initial position to a fixed target (**Figure 1A**; (Vassiliadis et al., 2021)). The force required to reach the target ($Target_{Force}$) corresponded to 10 % of the individual maximum voluntary contraction (MVC). In most of the trials (90 %), participants practiced the task with very limited sensory feedback: the cursor disappeared when the generated force reached half of the $Target_{Force}$ (see STAR Methods for more details on the task). To learn the task, subjects were provided with six Training blocks (T1 to T6; 40 trials each; *i.e.*, total of 240 training trials; **Figure 1B**) in which they received reinforcement feedback (*i.e.*, indicating Success or Failure) associated to a monetary reward. Success on the task was determined based the Error, defined as the absolute force difference between the $Target_{Force}$ and the exerted force (Abe et al., 2011; Steel et al., 2016).

In different groups of participants, we varied the delay between the end of the movement period and the delivery of the reward during the Training blocks. As such, $Group_{Short}$ subjects trained with a short reward delay (*i.e.,* 1 s) while participants of the $Group_{Long}$ performed the task with a long reward delay (*i.e.,* 6 s). The total duration of the trials was kept constant by modulating the inter-trial interval (ITI; 6 s in $Group_{Short}$ and 1 s in $Group_{Long}$). Before, immediately and 24 hours after training, all participants performed Test blocks with no reward, a short reward delay (1 s) and a short ITI (1 s). Notably, the groups were comparable for a variety of features including Pre-training success rates, difficulty of the task, force required, sensitivity to reward and

punishment, fatigue and final monetary gains (**Figure 1C**, **Table 1**). Altogether, this design allowed us to investigate the specific effect of reward timing on motor learning and consolidation.

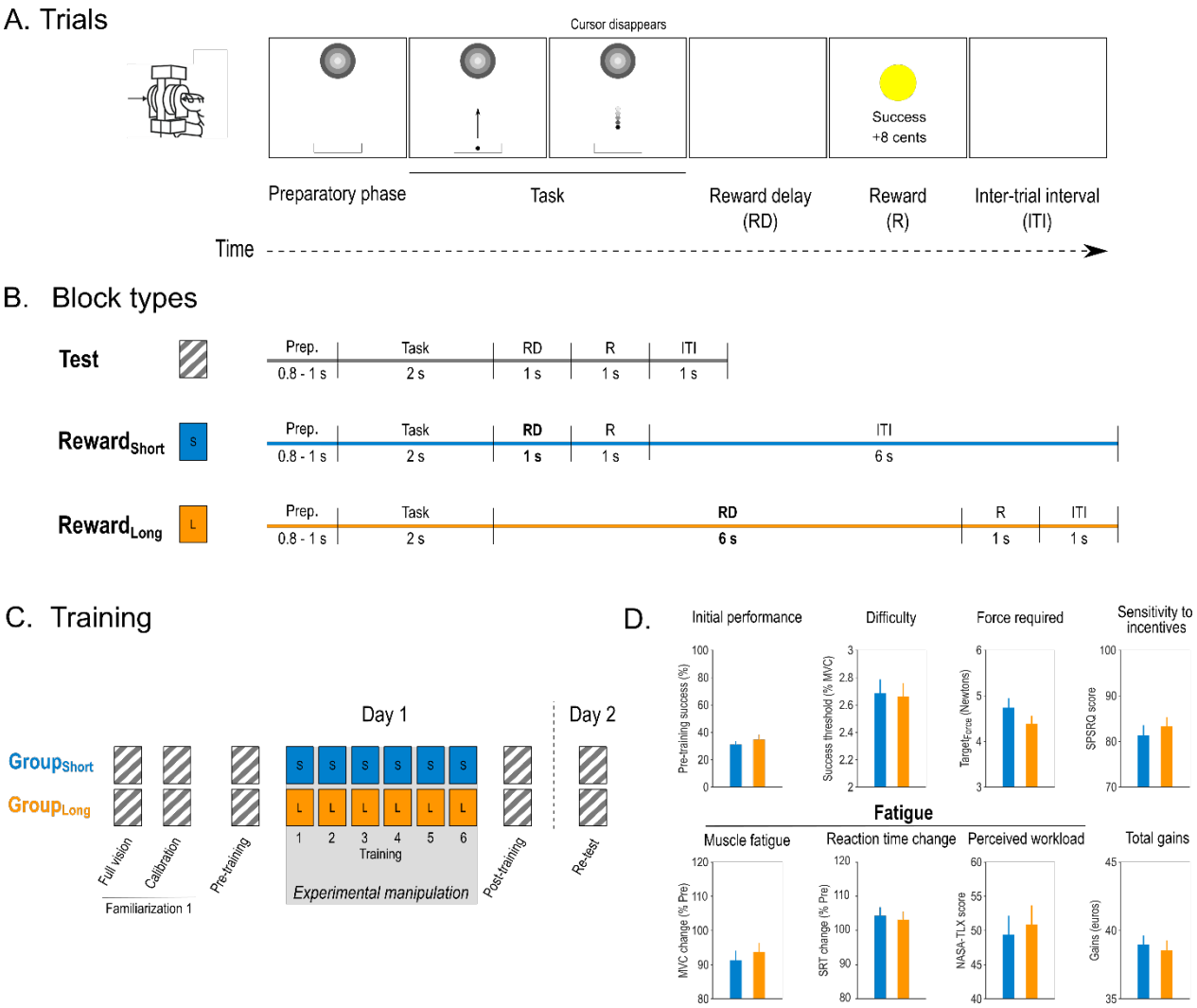


Figure 1. Motor skill learning task. A. Time course of a trial in the motor skill learning task. Each trial started with the appearance of a sidebar and a target. After a variable preparatory period (0.8-1s), a cursor appeared in the sidebar, playing the role of a "Go" signal. At this moment, participants were required to pinch the force transducer to bring the cursor into the target as quickly as possible and maintain it there until the end of the task (2 s). Notably, on most trials, the cursor disappeared halfway towards the target (as displayed here). Then, after a delay, a reward (R) appeared consisting in a reinforcement feedback and a monetary reward (a successful trial is shown here). Trials ended with an inter-trial interval (ITI). **B. Durations in the different block types.** Reward delays (RD) and ITIs were manipulated. Test blocks included a short reward delay (1 s), a short ITI (1 s) and no monetary reward (*i.e.,* only reinforcement). Reward$_{Short}$ and Reward$_{Long}$ blocks included monetary rewards and were performed with a short (1 s) and long (6 s) reward delay, respectively. The total duration of the trials was kept constant between Reward$_{Short}$

and $Reward_{Long}$ by varying the ITI. **C. Training procedure.** On Day 1, all participants performed two familiarization blocks in a Test blocks condition. The first one involved full vision of the cursor while the second one provided only partial vision and served to calibrate the difficulty of the task on an individual basis (See STAR Methods). Then, Pre- and Post-training Test blocks assessments were separated by 6 blocks of training in the condition corresponding to each individual group ($Reward_{Short}$ for $Group_{Short}$, $Reward_{Long}$ for $Group_{Long}$). Day 2 consisted in a short re-familiarization (5 trials with full vision, not represented) followed by a Re-test assessment (1 Test block). **D. Control analyses.** $Group_{Short}$ and $Group_{Long}$ were comparable for a variety of factors including initial performance, task difficulty, required force to reach the target, sensitivity to reward and punishment (as assessed by the SPSRQ questionnaire), muscular and cognitive fatigue and final monetary gains (see also Table 1).

Training with long reward delays induces an early plateau in learning and a reduction of endpoint performance

As a first step, we evaluated performance on the task by computing the average success rate per $Training_{Block}$ (T1 to T6). To compare the learning process between the groups, we performed a two-way ANOVA, with $TRAINING_{BLOCK}$ and $GROUP_{TYPE}$ as within- and between-subject factors, respectively. Overall, participants of both groups significantly improved their success rates over training (main effect of $TRAINING_{BLOCK}$: Greenhouse-Geisser [GG]-corrected $F_{(3.56, 206.77)} = 6.01$; $p < 0.001$, partial $\eta^2 = 0.094$; **Figure 2A**). Most importantly, the improvement in success rate over the blocks depended on the $Group_{Type}$, as revealed by a significant $TRAINING_{BLOCK}$ x $GROUP_{TYPE}$ interaction (GG-corrected $F_{(3.56, 206.77)} = 2.69$; $p = 0.038$, partial $\eta^2 = 0.044$; **Figure 2A**). As such, in $Group_{Short}$, performance only started to improve significantly at T4 relative to T1 (T4 vs. T1: $p = 0.044$; T5 vs T1: $p = 0.030$; T6 vs T1: $p < 0.001$) and kept improving in late $Training_{Blocks}$ (T6 vs T3: $p < 0.001$; T6 vs T4: $p = 0.034$; T6 vs T5: $p = 0.051$). In contrast, $Group_{Long}$ improved early on (*i.e.*, T2 vs T1: $p < 0.001$) but then stopped improving quickly during training, with the success rate at T2 being not different from the success rate at any of the following $Training_{Blocks}$ (all $p > 0.37$). Interestingly, post-hoc tests further revealed that endpoint performance (*i.e.*, success rate at T6)

was significantly lower in $Group_{Long}$ than in $Group_{Short}$ ($p = 0.047$; **Figure 2B**). Conversely, success rates at all other $Training_{Blocks}$ were comparable between the two groups (all $p > 0.22$; **Figure 2A**). Altogether, these findings suggest that reward timing can influence the dynamics of motor learning. More specifically, it appears that training with long reward delays is associated with an early plateau in learning and a related reduction of endpoint performance.

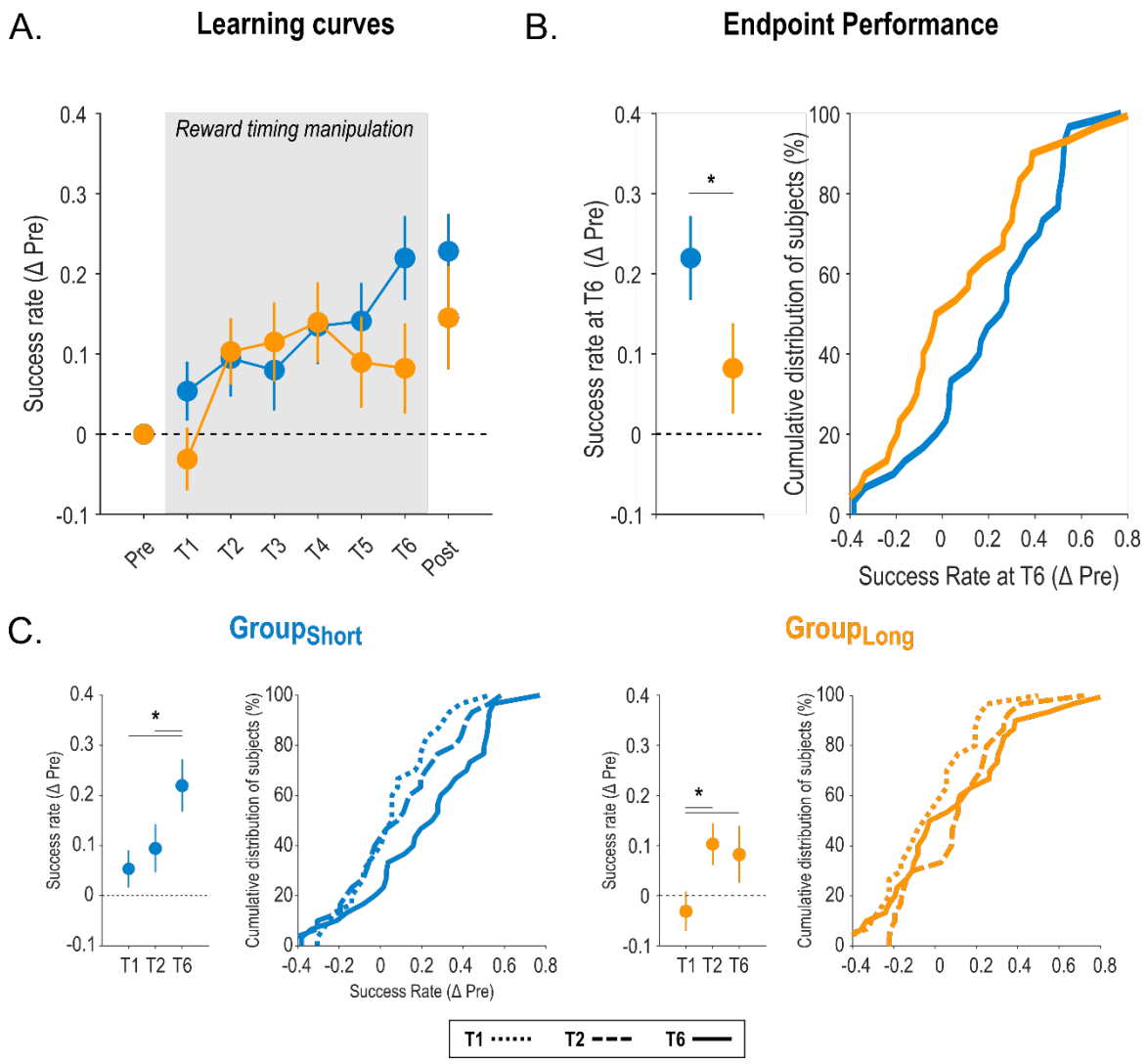


Figure 2. Effect of reward timing on motor skill learning. A. Learning curves. Proportion of successful trials (expressed as a difference with the individual Pre-training success rate) is represented across practice for the two experimental groups (blue: $Group_{Short}$, orange: $Group_{Long}$). The grey shaded area highlights the

blocks concerned by the reward timing manipulation. The remaining blocks were Test blocks. **B. Endpoint performance**. Average success rates in the end of the training period (i.e., measured at T6) in the two groups (left panel) and the corresponding cumulative distributions of the data (right panel). **C. Learning dynamics.** Average success rates observed at T1 (dotted line), T2 (dashed line) and T6 (solid line) and the corresponding cumulative distributions for $Group_{Short}$ (left) and $Group_{Long}$ (right panel). Performance improved very early in $Group_{Long}$ (i.e., between T1 and T2) and then plateaus while it improved later in $Group_{Short}$. Note that performance at T1 was not different between the groups (all $p > 0.21$). *: significant difference ($p<0.05$). Data are represented as mean ± SE.

An important aspect of our experimental design is that we increased the duration of the ITI in $Group_{Short}$ relative to $Group_{Long}$ (6 s and 1 s, respectively; **Figure 1B**), in order to match the total duration of the trials in both groups despite differences in reward timing. To verify that such manipulation did not impact learning in our task, we added in the analysis another group of participants, who trained with a short reward delay (0.5 s) and an intermediate ITI (3 s; $Group_{Short\text{-}PastStudy}$, n = 30; from (Vassiliadis et al., 2021)). We reasoned that, if differences in learning dynamics were truly driven by differences in reward timing but not by differences in ITI duration, learning in $Group_{Short\text{-}PastStudy}$ should be similar than in $Group_{Short}$, and therefore different than in $Group_{Long}$. Consistently, post-hoc tests run on the $TRAINING_{BLOCK}$ x $GROUP_{TYPE}$ interaction (GG-corrected $F_{(3.78, 328.92)} = 2.84$; $p = 0.0020$, partial $\eta^2 = 0.061$) showed (1) no difference at any $Training_{Blocks}$ between $Group_{Short\text{-}PastStudy}$ and $Group_{Short}$ (all $p > 0.23$) and (2) a trend for a difference in endpoint performance when comparing $Group_{Short\text{-}PastStudy}$ and $Group_{Long}$ (*i.e.*, $p = 0.049$ and 0.053 at T5 and T6, respectively), similarly as when comparing $Group_{Short}$ and $Group_{Long}$ (**Supplementary Figure 1**). Hence, these results confirm that a change of a few seconds in ITI duration does not impact motor learning, as previously reported (McDougle et al., 2015), implying that the differences in learning dynamics observed in groups exposed to short and long reward delays did not result from differences in ITI duration but rather from actual differences in reward timing.

199 To evaluate total learning we computed success rates at Post-training, which was performed
200 in a Test block setting in both groups. Overall, success rates increased by 22.8 ± 4.69 % in
201 $Group_{Short}$, and 14.5 ± 6.42 % in $Group_{Long}$ with respect to Pre-training. Interestingly, success rates
202 at Post-training were significantly different from 0 in $Group_{Short}$ despite Bonferroni correction
203 (cutoff for significance: $p = 0.025$; $t_{(29)} = 4.87$, $p < 0.001$), indicative of a significant improvement
204 in performance with respect to Pre-training. In contrast, success rates at Post-training were not
205 significantly different from 0 in $Group_{Long}$ after Bonferroni correction ($t_{(29)} = 2.26$, $p = 0.031$).
206 However, a t-test on these data did not show any significant difference between the $Group_{TYPES}$
207 ($t_{(58)} = 1.05$; $p = 0.30$). Hence, reward timing only induced a subtle change in total learning that did
208 not reach significance when comparing directly the groups.

209 Results of the first analysis suggest that reward timing influenced the dynamics of the
210 learning process with long reward delays being associated with a fast improvement of performance
211 quickly followed by a plateau. To confirm this, we aimed at directly quantifying the learning rates
212 for early and late phases of training ($Training_{Early}$ and $Training_{Late}$, corresponding to T1-T3 and T4-
213 T6, respectively). To do so, for each subject and each training phase, we split the data into 12 bins
214 of 10 trials, computed the average success rate for each bin and then performed linear regressions
215 over the bins for each $Training_{Phase}$ (*i.e.*, 12 bins per $Training_{Phase}$; **Figure 3A**). The slopes of the
216 linear fits (*i.e.*, expressed in ΔSuccess rate per bin of trials) were exploited as a proxy of the learning
217 rate at a given $Training_{Phase}$ (**Figure 3B, C**). A two-way ANOVA did not show any main effect of
218 the factor $TRAINING_{PHASE}$ ($F_{(1, 58)} = 2.80$; $p = 0.10$, partial $\eta^2 = 0.046$), nor any effect of
219 $GROUP_{TYPE}$ on the learning rates ($F_{(1, 58)} = 0.14$; $p = 0.71$, partial $\eta^2 = 0.0023$). Interestingly though,
220 we found a strong $TRAINING_{PHASE}$ x $GROUP_{TYPE}$ interaction on these data ($F_{(1, 58)} = 9.13$; $p =$
221 0.0037, partial $\eta^2 = 0.14$; **Figure 3A, B, C**). This interaction was driven by the fact that learning

rates were comparable across training phases in $Group_{Short}$ ($p = 0.34$), while they were significantly lower at $Training_{Late}$ than at $Training_{Early}$ in $Group_{Long}$ ($p = 0.0016$; **Figure 3B, D, E**). Consistently, learning rates were higher in $Group_{Short}$ than in $Group_{Long}$ at $Training_{Late}$ ($p = 0.017$). This tended to be the opposite at $Training_{Early}$ ($p = 0.061$). Notably, intercepts of the linear fits were not significantly different between the groups, neither at $Training_{Early}$ ($t_{(58)} = 1.07$, $p = 0.29$), nor at $Training_{Late}$ ($t_{(58)} = -1.33$, $p = 0.19$), indicating that between-group differences in learning rates could not be accounted for by differences in initial performance. Hence, this analysis confirms that training with short reward delay induces slow, yet continuous gains in performance, while long reward timings engender large initial learning rates that then drop significantly, indicative of a plateau in learning. Overall, the data show that reward timing impacts learning dynamics during a motor training.

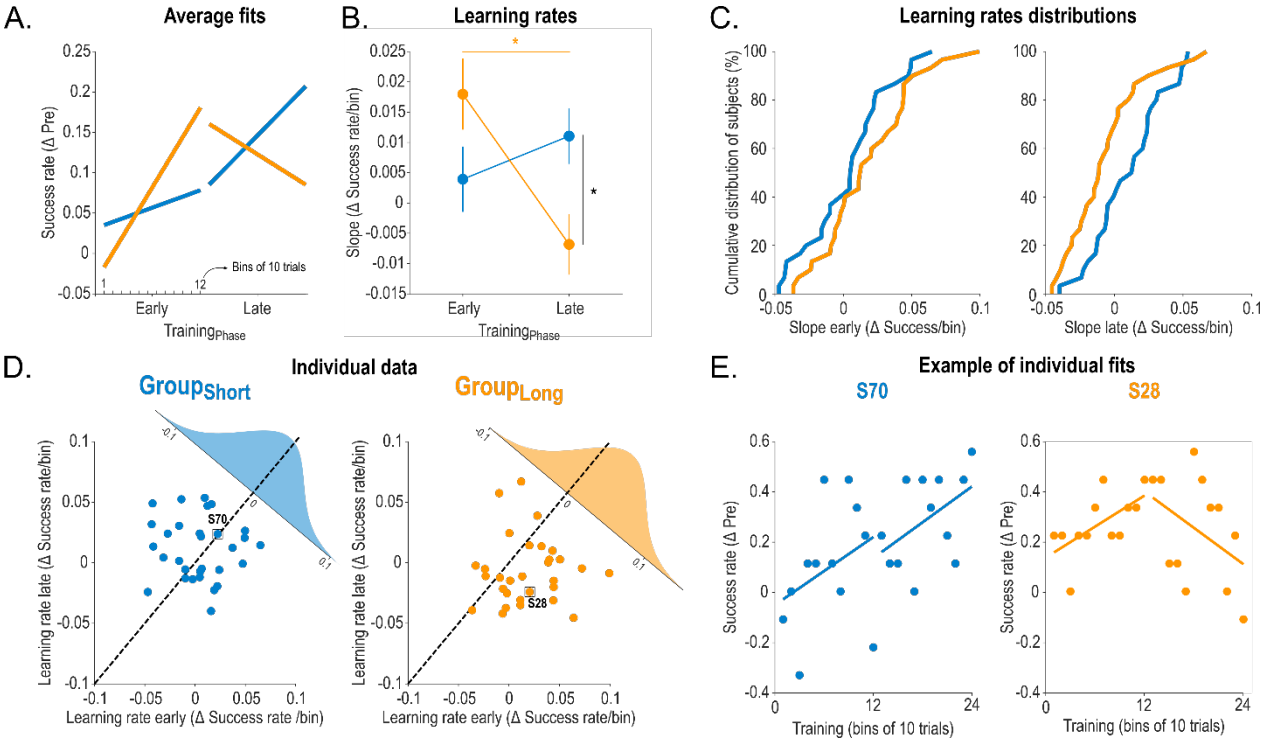


Figure 3. Effect of reward timing on learning rates. **A. Average fits**. Average fits for $Group_{Short}$ (blue) and $Group_{Long}$ (orange) obtained through linear regression on each subject's binned success data (expressed as a difference from Pre-Training) at early and late phases of training. **B. Learning rates.** The slope of the

individual fits – expressed as a delta of success rate per bin on the y-axis – was exploited as a proxy of the learning rate for early and late phases of training. Note the significant $TRAINING_{BLOCK}$ x $GROUP_{TYPE}$ interaction which was driven by the fact that while learning rates were stable in $Group_{Short}$, they were initially high in $Group_{Long}$ and dropped significantly at $Training_{Late}$ (orange star). Moreover, learning rates at $Training_{Late}$ were significantly lower in $Group_{Long}$ than in $Group_{Short}$ (black star). **C. Learning rates distributions.** Cumulative distributions of the group data for early (left panel) and late (right panel) learning phases. **D. Individual learning rates.** Scatter plot representing each subject's learning rate for early (x-axis) vs late (y-axis) $Training_{Phases}$. The group distributions of the change in learning rates ($Training_{Early}$ – $Training_{Late}$) are also represented (upper right corner of each plot). Note the shift in $Group_{Long}$ below the identity line reflecting the higher learning rates at $Training_{Early}$ than at $Training_{Late.}$ **E. Example of individual fits.** Binned success rates and the corresponding linear fits are represented for an exemplar subject of $Group_{Short}$ (left) and $Group_{Long}$ (right). *: significant difference ($p<0.05$). Data are represented as mean ± SE.

Training with long reward delay impairs overnight skill consolidation

As a last step, we investigated the impact of the reward timing experienced during training on Day 1 on overnight consolidation of the skill (*i.e.*, on Day 2). Because this analysis was focused on the consolidation of a learned motor skill, we only included participants who demonstrated learning on Day 1 (n = 22 and 18 in $Group_{Short}$ and $Group_{Long}$, respectively). To evaluate consolidation, we performed a two-way ANOVA on the normalized success rates obtained at Post-training of Day 1 and at Re-test of Day 2 (*i.e.*, both performed in a Test block setting) with $TEST_{BLOCK}$ and $GROUP_{TYPE}$ as within and between-subjects factors, respectively. The ANOVA showed a significant effect of $TEST_{BLOCK}$ ($F_{(1, 38)} = 5.44$; $p = 0.025$, partial $\eta^2 = 0.13$): success rates were on average smaller in the Re-test block on Day 2 than in the Post-training block on Day 1. However, this effect was mostly driven by the $Group_{Long}$, in which performance deteriorated overnight. As such, the ANOVA also revealed a $TEST_{PHASE}$ x $GROUP_{TYPE}$ interaction ($F_{(1, 38)} = 5.77$; $p = 0.021$, partial $\eta^2 = 0.13$). While performance was comparable between learners of the two groups at Post-training on Day 1 ($p = 0.65$), success rates were significantly reduced on Day 2 relative to Day 1 in $Group_{Long}$ ($p = 0.0028$), but remained stable from one day to another in

$Group_{Short}$ (p = 0.95). This indicates that delaying rewards on Day 1 impaired consolidation of the motor skill on Day 2. Overall, our results support the view that short or long reward delays support qualitatively different motor learning processes during training, leading to different consolidation of the skill.

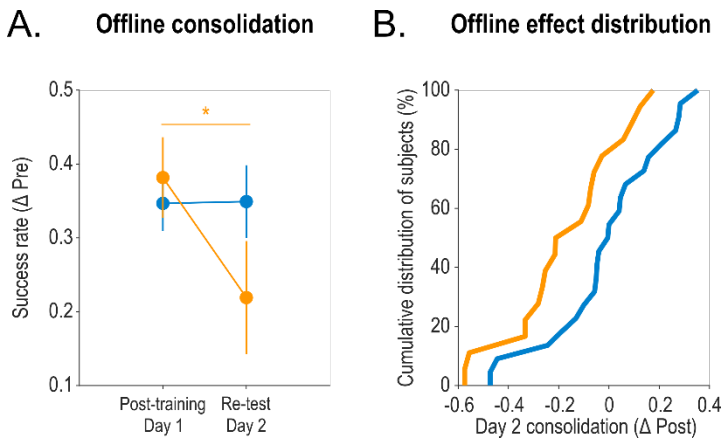


Figure 4. Effect of reward timing on overnight consolidation of the motor skill. **A. Offline consolidation of the motor skill.** Proportion of successful trials (expressed as a difference with the individual Pre-training success rate) at Post-training (Day 1) and Re-test (Day 2). Both assessments were Test blocks. This analysis only considered participants who demonstrated skill learning on Day 1 (n = 22 and 18 in $Group_{Short}$ and $Group_{Long}$, respectively). Notably, a significant $TEST_{PHASE}$ x $GROUP_{TYPE}$ interaction showed that while $Group_{Short}$ participants performed comparably well in Post-training and Re-test, demonstrating offline consolidation of the skill, this process was altered in $Group_{Long}$. **B. Offline effect distribution.** Cumulative distributions of the offline effect (Success rate in Day 2 – Success rate in Post-training) for each $Group_{Type}$

3. Discussion

Previous studies have shown that reward timing can influence the response of brain structures involved in reward processing during associative learning (Fiorillo et al., 2008; Foerde et al., 2013; Foerde and Shohamy, 2011; Klein-Flügge et al., 2011; Kobayashi and Schultz, 2008). Inspired by these neurophysiological findings, we asked whether reward timing can also influence how people learn and consolidate a new motor skill. We found that delaying reward delivery by a few seconds influences motor learning dynamics: training with a short reward delay induced slow, yet continuous gains in performance, while a long reward delay allowed high initial learning rates that were followed by an early plateau in the learning curve and a lower endpoint performance. Moreover, among participants who successfully learned the skill, those who trained with a short reward delay displayed overnight consolidation, while those who learned the task with a long reward delay exhibited an impairment in the consolidation of the motor memory. Overall, our findings show that reward timing can influence how the brain learns and consolidates new motor skills.

The differences in learning dynamics observed in subjects trained with short and long reward delays may indicate that reward boosted processes presenting different temporal dynamics. As such, a prevalent view in the field is that motor learning entails the operation of distinct processes, with either slow (*i.e.,* developing over a few trials) or fast (*i.e.,* developing over tens/hundreds of trials) temporal dynamics (Smith et al., 2006). The slow process is characterized by both a low learning rate and a sluggish forgetting of the acquired behaviour and is thought to reflect implicit learning (McDougle et al., 2015; Trewartha et al., 2014). In contrast, the fast process entails both

302 a high learning rate and a quick forgetting of the new behavior and supports the explicit learning
303 of new motor behaviors (McDougle et al., 2015; Trewartha et al., 2014). The nature of our task did
304 not allow us to evaluate the relationship between reward timing and the relative contribution of
305 implicit and explicit learning. Still, people who trained with a short reward delay exhibited learning
306 dynamics that presented a low learning rate and a clear overnight consolidation – reminiscent of
307 the slow process, while those who trained with a long reward delay exhibited a high initial learning
308 rate and an overnight forgetting of the motor memory – evocative of the fast process. Based on
309 these results, one may suggest that short reward delays preferentially facilitate the slow (implicit)
310 process, while long reward delays may favour the fast (explicit) learning process, accentuating their
311 respective contribution to subjects' improvements, an interesting issue for future investigation.

312

313 Interestingly, neurophysiological data also support the latter interpretation. Indeed, the
314 striatum – which is preferentially recruited by short reward delays (Foerde and Shohamy, 2011) –
315 exhibits slow changes in activity over the course of motor learning (Albouy et al., 2013, 2008) and,
316 relatedly, was shown to be central for slow, implicit learning (Albouy et al., 2013; Hayes et al.,
317 2015; Siegert et al., 2006). Conversely, the hippocampus – which is mostly recruited by long
318 reward delays – usually exhibits a fast increase in activity in the early phase of learning that wanes
319 later on (Albouy et al., 2013, 2008). In this line, the hippocampus was evidenced to be central for
320 fast, explicit learning (Döhring et al., 2017). Altogether, these elements suggest that the different
321 learning dynamics observed in individuals training with short and long reward delays could result
322 from the fact that these delays recruit the striatum and the hippocampus, respectively, shaping thus
323 their activity and boosting preferentially the slow and fast processes in which those structures are
324 mostly involved. Further neurophysiological work is required to test this hypothesis directly.

The impairment of motor consolidation observed in subjects who trained with a long reward delay also suggests that the strength of learning was altered in these participants. As such, efficient reward-based motor learning relies on the mapping between somatosensory sensations (*e.g.*, elicited by the generated force in the present task) and the associated reward (Bernardi et al., 2015; Sidarta et al., 2016; Vassiliadis et al., 2021), and somatosensory working memory is known to decay quickly following movement execution, after only a few seconds (Harris et al., 2001; Sidarta et al., 2018). Hence, it is possible that delaying reward delivery blunted the reinforcement of somatosensory working memory (Sidarta et al., 2018), explaining the limited learning and consolidation observed in the subjects of $\text{Group}_{\text{Long}}$. Another, potentially complementary interpretation is that rewards delivered after a long delay are temporally discounted and perceived as subjectively less valuable relative to when the delay is short (Shadmehr et al., 2019, 2010). If this is true, the reduction in subjective value in $\text{Group}_{\text{Long}}$ may have contributed to reducing the beneficial effect of reward on offline memory consolidation (Sterpenich et al., 2021). Altogether, these elements indicate that the temporal contingency between a movement and the ensuing reward is a decisive aspect of reward-based motor learning.

Beyond reward timing, another feature that could have altered motor learning in the present study is the post-reward delay – *i.e.*, the delay between reward delivery and the execution of the subsequent movement (referred to as ITI in the *Results* section, above). However, our data suggest that post-reward delay did not affect motor learning in the present study. Indeed, subjects of $\text{Group}_{\text{Short}}$ and of $\text{Group}_{\text{Short-PastStudy}}$, which trained with a short reward delay but with different ITIs, displayed very similar learning (Supplementary Figure 1). This finding is in line with the idea that

348 motor learning remains largely unaffected by the ITI duration (even for ITIs of 30 s; (McDougle
349 et al., 2015)).

350

351 In conclusion, our data indicate that the timing at which reward is delivered during motor
352 training alters the dynamics of learning and the consolidation of the new motor memory. Research
353 is now required to gain further knowledge as to the brain networks involved in these time-
354 dependent effects of reward on motor learning as well as to characterize the relative contribution
355 of implicit and explicit learning processes to these effects. Such knowledge would prove useful for
356 the design of future reward-based rehabilitation programs, in which reward timing may be
357 individualized depending on the brain networks and learning processes affected in specific
358 populations of patients (Foerde et al., 2013; Foerde and Shohamy, 2011).

4. STAR METHODS

RESOURCE AVAILABILITY

Lead contact

Further information and requests should be directed to the lead contact, Pierre Vassiliadis (contact: pierre.vassiliadis@uclouvain.be).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Data used in this study are available upon request (contact: pierre.vassiliadis@uclouvain.be) and will be freely available at the time of publication via an open-access data sharing repository (https://osf.io).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

A total of sixty right-handed healthy volunteers participated in the present study (46 women, 23.7 ± 0.3 years old; mean ± SE). Data from a previous group of thirty participants was also re-analyzed (20 women, 23.9 ± 0.43 years old; (Vassiliadis et al., 2021)). Handedness was determined via a shortened version of the Edinburgh Handedness inventory (Oldfield, 1971). None of the participants suffered from any neurological or psychiatric disorder, nor were they taking any centrally-acting medication. All participants gave their written informed consent in accordance with the Ethics Committee of the Université Catholique de Louvain (approval number: 2018/22MAI/219) and the principles of the Declaration of Helsinki. Subjects were financially

380 compensated for their participation. Finally, all participants were asked to fill out a French
381 adaptation of the Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ;
382 (Lardi et al., 2008; Torrubia et al., 2001)) and a NASA Task Load Index questionnaire (NASA-
383 TLX, (Hart and Staveland, 1988)).

384

385 **METHOD DETAILS**

386 **Motor skill learning task**

387 Participants were seated approximately 60 cm in front of a cathode-ray tube screen (refresh
388 rate: 100 Hz) with their right forearm positioned at a right angle on the table. The task was
389 developed on Matlab 7.5 (the Mathworks, Natick, Massachusetts, USA) exploiting the
390 Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) and consisted in a previously
391 described force modulation task (Vassiliadis et al., 2021). Briefly, the task required participants to
392 squeeze a force transducer (Arsalis, Belgium) between the index and the thumb to control a cursor
393 displayed on the screen. Increasing the force exerted resulted in the cursor moving vertically and
394 upward. Each trial started with a preparatory period in which a sidebar appeared at the bottom of
395 the screen and a target at the top (**Figure 1A**). After a variable time interval (0.8 to 1 s), a cursor
396 popped up in the sidebar, indicating the start of the movement period. Participants had to pinch the
397 transducer to move the cursor as quickly as possible from the sidebar to the target and maintain it
398 there for the rest of the movement period, which lasted 2 s. The level of force required to reach the
399 target ($Target_{Force}$) was individualized for each participant and set at 10 % of maximum voluntary
400 contraction (MVC). Notably, squeezing the transducer before the appearance of the cursor was
401 considered as an anticipation and therefore led to the interruption of the trial. Anticipation trials

were discarded from further analyses. At the end of each trial, a binary reinforcement feedback was presented to the subject (yellow or blue circle for success or failure, respectively).

Sensory and reinforcement feedbacks

We provided only limited visual feedback to the participants in order to increase the impact of the reinforcement feedback on learning (Mawase et al., 2017). As such, on 90 % of the trials, the cursor disappeared shortly after the start of the movement period: it became invisible as soon as the generated force became larger than half of the $Target_{Force}$ (*i.e.,* 5 % of MVC). Conversely, the remaining trials (10 % of the trials) provided a continuous vision of the cursor (full vision trials). Full vision trials were not considered in the analyses.

As mentioned above, each trial ended with the presentation of a binary reinforcement feedback, indicating success or failure. Success on the task was determined based on the Error, defined as the absolute force difference between the $Target_{Force}$ and the exerted force (Abe et al., 2011; Steel et al., 2016). The Error was first computed for each frame refresh from 0.15 s to the end of the trial (*i.e.,* providing 185 data points at 100 Hz), then averaged across the data points for each trial (Steel et al., 2016), and expressed in percentage of MVC. This indicator of performance allowed us to classify a trial as successful or not based on an individualized success threshold (see below). When the Error on a given trial was below the threshold, the trial was classified as successful, and when it was above the threshold, the trial was considered as failed. Hence, task success depended on the ability to approximate the $Target_{Force}$ as quickly and as accurately as possible.

424 **Reward timing manipulation**

425

426 The protocol involved Training and Test blocks (see *Experimental protocol*, below). During
427 Training blocks, reinforcement feedbacks were associated with a reward of 8 cents on successful
428 trials, and failed trials led to 0 cent. Importantly, in two block types, we manipulated the timing at
429 which the reinforcement feedback, and therefore the associated reward, was delivered after the
430 movement period (**Figure 1A**). Indeed, the reward was displayed after either a short or a long delay
431 – that is, 1 or 6 s following the movement period in $\text{Reward}_{\text{Short}}$ and $\text{Reward}_{\text{Long}}$ blocks, respectively
432 (see (Foerde et al., 2013; Foerde and Shohamy, 2011) for the use of similar delays in decision-
433 making tasks). In order to keep the total duration of the trial constant in these two block types,
434 inter-trial intervals (ITI, which followed reward occurrence) were set to 6 and 1 s in the $\text{Reward}_{\text{Short}}$
435 and the $\text{Reward}_{\text{Long}}$ blocks, respectively. Finally, we re-analyzed data from a previous study
436 (Vassiliadis et al., 2021), in which the training blocks involved a short reward delay timing (0.5 s)
437 and an intermediate ITI (3 s; $\text{Reward}_{\text{Short-PastStudy}}$ blocks). The latter analysis allowed us to test for
438 the reproducibility of the effects of training obtained in the $\text{Reward}_{\text{Short}}$ block.

439 In the Test blocks, reinforcement feedback occurred 1 s after the movement period,
440 involved an ITI of 1 s, and was not associated with any reward.

441

442 **Motor skill learning protocol**

443 Subjects were tested on two consecutive days (Day 1 and Day 2; **Figure 1C**). On Day 1,
444 we first measured the individual MVC to calculate the $\text{Target}_{\text{Force}}$. Notably, MVCs and simple
445 reaction times (SRT) were measured before and after the training blocks to assess potential fatigue

related to the training (see **QUANTIFICATION AND STATISTICAL ANALYSIS**). Participants then performed 2 blocks of Familiarization, in a Test block setting. The first Familiarization block comprised 20 full vision trials. Subsequently, all blocks were composed of a mixture of partial vision trials (90 % of total trials) and full vision trials (10 % of total trials), as described above. The second Familiarization block involved 40 trials and allowed us to determine baseline performance to calibrate the difficulty of the task for the rest of the experiment (Calibration block; please see (Vassiliadis et al., 2021) for details on the Calibration procedure).

Following Familiarization, participants performed 320 trials divided in 8 blocks. All subjects started and ended the session with the realization of a Test block of 40 trials, allowing us to evaluate initial performance and total learning (*i.e.*, Pre- and Post-training blocks, respectively). In between, 6 Training blocks (T1 to T6) of 40 trials were performed by the participants (see **Figure 1B**). During the Training blocks, individuals were split into 2 separate groups depending on the type of training blocks they performed. As such, $\text{Group}_{\text{Short}}$ and $\text{Group}_{\text{Long}}$ trained with $\text{Reward}_{\text{Short}}$ and $\text{Reward}_{\text{Long}}$ blocks, respectively. The group trained with $\text{Reward}_{\text{Short-PastStudy}}$ blocks was referred to as $\text{Group}_{\text{Short-PastStudy}}$. Comparing performance between the groups during the training period allowed us to test the effect of reward timing on the learning dynamics.

Day 2 was realized 24 hours later. Subjects performed the task again with the same $\text{Target}_{\text{Force}}$ and success threshold. This assessment was composed of 5 full vision trials followed by a Test block of 40 trials (Re-test) and allowed us to assess the effect of reward timing on skill consolidation.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were carried out with Matlab 2018a (the Mathworks, Natick, Massachusetts, USA) and Statistica 10 (StatSoft Inc., Tulsa, Oklahoma, USA). In the case of independent samples t-tests and Analysis of Variance (ANOVA), we verified the homogeneity of the variances systematically and non-parametric tests were used when variances were non-homogeneous. All data were systematically tested for the sphericity assumption using Maunchley's tests. The Greenhouse-Geisser (GG) correction was used for sphericity when necessary. Post-hoc comparisons were conducted using the Fisher's LSD procedure. The significance level was set at $p \leq 0.05$, except in the case of correction for multiple comparisons (see below).

Motor skill learning

As a first step, we tested the impact of reward timing on motor performance during each block of Test and Training block. We quantified for each subject the percentage of successful trials (*i.e.*, the success rate) for each block and then normalized the data according to individuals' initial performance by subtracting the success rate values measured at Pre-training from the values obtained in every block. To evaluate the impact of reward timing on success rates across training, we performed a two-way ANOVA with $GROUP_{TYPE}$ ($Group_{Short}$ and $Group_{Long}$) and $TRAINING_{BLOCK}$ (T1 to T6) as between- and within-subjects factors, respectively. Then, to evaluate the effect of the ITI's duration on motor learning, we replicated this analysis with the inclusion of the $Group_{Short\text{-}PastStudy}$.

As a second step, we aimed at evaluating how reward timing influenced learning rates during the task. To do so, we split the data into 24 non-overlapping bins of 10 trials, computed the

success rate for each bin and normalized the data according to individuals' initial performance, as done in the first analysis. The bins were then separated into two equal parts (*i.e.*, of 12 bins each) depending on whether they belonged to the early or to the late phase of training ($\text{Training}_{\text{Early}}$ and $\text{Training}_{\text{Late}}$ phases, corresponding to T1-T3 and T4-T6, respectively). Finally, we performed linear regressions on these data and extracted the slope of the fits for the $\text{Training}_{\text{Early}}$ and the $\text{Training}_{\text{Late}}$ phases of the $\text{Group}_{\text{Short}}$ and the $\text{Group}_{\text{Long}}$. The slope values – exploited here as a proxy of the learning rate – were compared using a two-way ANOVA with $\text{GROUP}_{\text{TYPE}}$ ($\text{Group}_{\text{Short}}$ and $\text{Group}_{\text{Long}}$) and $\text{TRAINING}_{\text{PHASE}}$ ($\text{Training}_{\text{Early}}$ and $\text{Training}_{\text{Late}}$) as between- and within-subjects factors, respectively.

Finally, we tested for any effect of reward timing on total learning, by comparing the success rates of the three groups at Post-training, using a one-way ANOVA with the factor $\text{GROUP}_{\text{TYPE}}$ ($\text{Group}_{\text{Short}}$, $\text{Group}_{\text{Long}}$ and $\text{Group}_{\text{Short-PastStudy}}$). Further, in order to test the statistical significance of total learning within each group, we conducted two single sample t-tests on Post-training success rate, against a constant value of 0 (Bonferroni-corrected at $p \leq 0.025$).

Motor skill consolidation

A secondary goal of the study was to evaluate the effect of reward timing on skill consolidation. For this analysis, we thus considered participants who demonstrated skill learning on Day 1 ($\text{Success}_{\text{Post-training}} - \text{Success}_{\text{Pre-training}} > 0$). This step not only allowed us to exclude non-learners but also to compare consolidation in groups that exhibited a similar learning on Day 1 (*i.e.*, a similar success rate at Post-training). 40 participants were considered in this analysis (22 and 18 in $\text{Group}_{\text{Short}}$ and $\text{Group}_{\text{Long}}$, respectively). We compared the success rate data through a two-way

ANOVA with $GROUP_{TYPE}$ ($Group_{Short}$ and $Group_{Long}$) and $TEST_{BLOCK}$ (Post-training and Day 2) as between- and within-subjects factors, respectively.

Group features, initial performance and fatigue

As a control, we verified that the $Group_{Short}$ and the $Group_{Long}$ were comparable in terms of age, success threshold, $Target_{Force}$, sensitivity to reward and to punishment (*i.e.,* as assessed by the SPSRQ questionnaire), initial performance (*i.e.,* at Pre-training) and received monetary gains. As displayed in **Table 1**, independent sample two-tailed t-tests performed on these data did not reveal any significant differences between the groups (see also **Figure 1C**).

We also assessed if potential motor and cognitive fatigue generated by Day 1 training was different between the groups (Derosière et al., 2014; Derosiere and Perrey, 2012). To do so, we expressed MVCs, and SRTs obtained after training (MVC_{POST} and SRT_{POST}) in percentage of the values measured initially (MVC_{PRE} and SRT_{PRE}). We also assessed the perceived workload after training through the NASA-TLX questionnaire. Notably, these data did not differ between the groups (**Table 1**), suggesting that motor and cognitive fatigue were not responsible for the effect of reward timing on motor learning.

	$Group_{Short}$ (n = 30)	**$Group_{Long}$ (n = 30)**	**t-value**	**p-value**
Age (in years)	22.8 ± 0.58	23.0 ± 0.52	-0.26	0.80
Gender (number of females)	22	24	/	/

Success Threshold (% MVC)	2.7 ± 0.01	2.7 ± 0.01	0.18	0.86
$Target_{Force}$ (Newtons)	4.74 ± 0.21	4.39 ± 0.17	1.32	0.19
Sensitivity to reward and punishment (score)	82.0 ± 2.32	83.3 ± 2.04	-0.43	0.67
Pre-training success rate	31.0 ± 2.61	34.9 ± 3.62	-0.86	0.39
Monetary Gains (euros)	39.0 ± 0.64	38.5 ± 0.73	0.45	0.66
Muscle fatigue (MVC_{POST} in % of MVC_{PRE})	91.3 ± 2.83	93.7 ± 2.67	-0.62	0.54
Simple Reaction Time change (SRT_{POST} in % of SRT_{PRE})	104.35 ± 2.51	103.21 ± 2.32	0.33	0.74
Perceived workload (NASA-TLX score)	49.4 ± 2.74	50.89 ± 2.79	-0.39	0.70

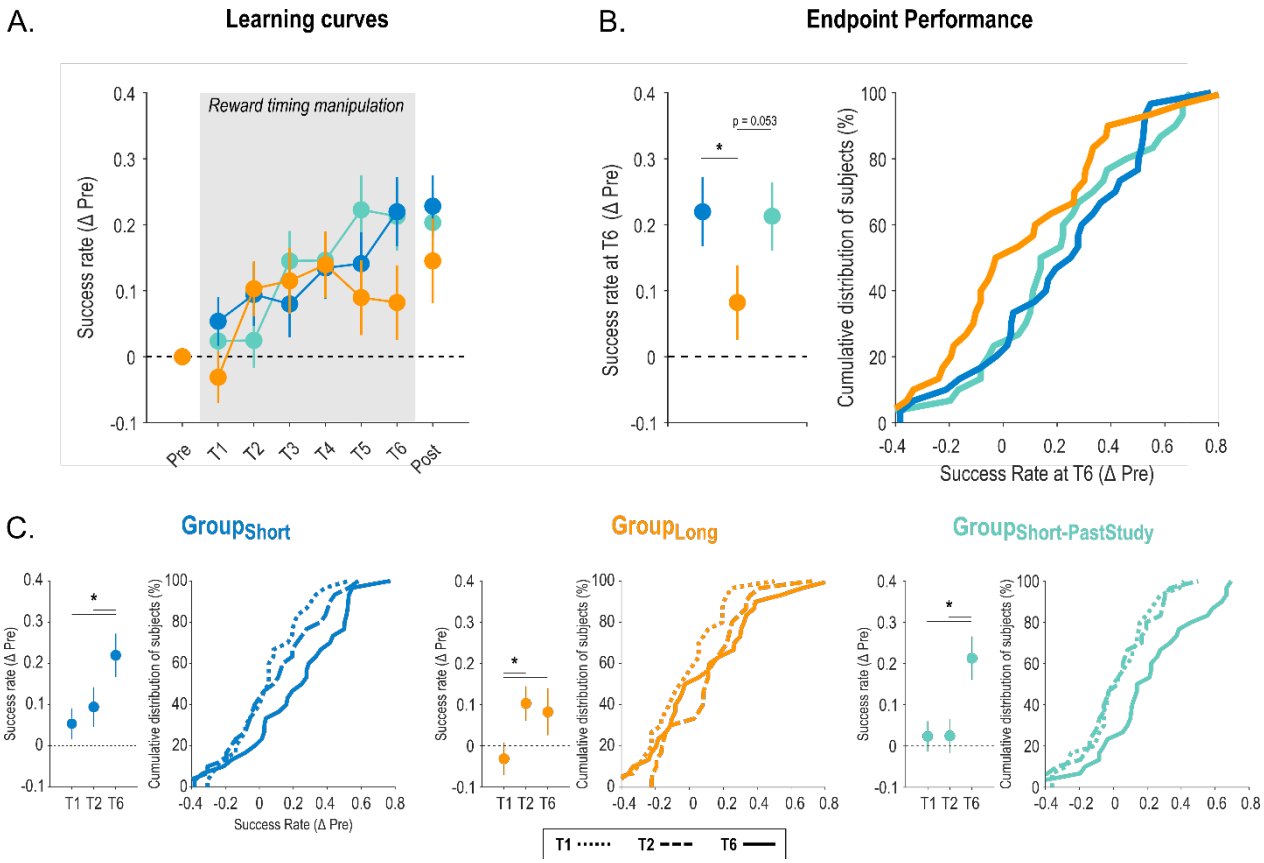
Table 1. Group features, initial performance and fatigue in the three experimental groups (mean ± SE). The 2 last columns provide the results of independent samples t-tests.

Acknowledgements

P.V. was a PhD student supported by the Fund for Research training in Industry and Agriculture (FRIA/FNRS; FC29690), and grants by the Platform for Education and Talent (Gustave Boël - Sofina Fellowships) and Wallonie-Bruxelles International. J.D. was supported by grants from the Belgian FNRS and the Fondation Médicale Reine Elisabeth (FMRE). G.D. was supported by the Belgian FNRS and by ERC StG 2018 801872 Rhythm and Brains.

Competing interests

The authors declare no conflict of interest.



Supplementary Figure 1. Effect of reward timing on motor skill learning with the inclusion of $Group_{PastStudy}$. **A. Learning curves**. Proportion of successful trials (expressed as a difference with the individual Pre-training success rate) is represented across practice for the three experimental groups (blue: $Group_{Short}$, orange: $Group_{Long}$, turquoise: $Group_{Short\text{-}PastStudy}$). The grey shaded area highlights the blocks concerned by the reward timing manipulation. The remaining blocks were Test blocks. Note the similarity in the learning curves of $Group_{Short}$ and $Group_{Short\text{-}PastStudy}$ despite the difference in ITI. **B. Endpoint performance**. Average success rates in the end of the training period (i.e., measured at T6) in the three groups (left panel) and the corresponding cumulative distributions of the data (right panel). **C. Learning dynamics.** Average success rates observed at T1 (dotted line), T2 (dashed line) and T6 (solid line) and the corresponding cumulative distributions for $Group_{Short}$ (left), $Group_{Long}$ (middle) and $Group_{Short\text{-}PastStudy}$ (right panel). Performance improved very early in $Group_{Long}$ (i.e., between T1 and T2) and then plateaus while it improved later in $Group_{Short}$ and $Group_{Short\text{-}PastStudy}$. Note that performance at T1 was not different between the groups (all $p > 0.21$). *: significant difference ($p<0.05$). Data are represented as mean ± SE.

References

Abe, M., Schambra, H., Wassermann, E.M., Luckenbaugh, D., Schweighofer, N., Cohen, L.G., 2011. Reward improves long-term retention of a motor memory through induction of offline memory gains. Curr. Biol. 21, 557–562. https://doi.org/10.1016/j.cub.2011.02.030

Albouy, G., King, B.R., Maquet, P., Doyon, J., 2013. Hippocampus and striatum: Dynamics and interaction during acquisition and sleep-related motor sequence memory consolidation. Hippocampus 23, 985–1004. https://doi.org/10.1002/hipo.22183

Albouy, G., Sterpenich, V., Balteau, E., Vandewalle, G., Desseilles, M., Dang-Vu, T., Darsaud, A., Ruby, P., Luppi, P.H., Degueldre, C., Peigneux, P., Luxen, A., Maquet, P., 2008. Both the Hippocampus and Striatum Are Involved in Consolidation of Motor Sequence Memory. Neuron 58, 261–272. https://doi.org/10.1016/j.neuron.2008.02.008

Bernardi, N.F., Darainy, M., Ostry, D.J., 2015. Somatosensory contribution to the initial stages of human motor learning. J. Neurosci. 35, 14316–14326. https://doi.org/10.1523/JNEUROSCI.1344-15.2015

Brainard, D.H., 1997. The Psychophysics Toolbox. Spat. Vis. 10, 433–436. https://doi.org/10.1163/156856897X00357

Chen, X., Holland, P., Galea, J.M., 2017. The effects of reward and punishment on motor skill learning. Curr. Opin. Behav. Sci. 20, 83–88. https://doi.org/10.1016/j.cobeha.2017.11.011

Dayan, E., Averbeck, B.B., Richmond, B.J., Cohen, L.G., 2014. Stochastic reinforcement benefits skill acquisition. Learn. Mem. 21, 140–142. https://doi.org/10.1101/lm.032417.113

Derosière, G., Alexandre, F., Bourdillon, N., Mandrick, K., Ward, T.E., Perrey, S., 2014. Similar scaling of contralateral and ipsilateral cortical responses during graded unimanual force

588 generation. Neuroimage 85, 471–477. https://doi.org/10.1016/j.neuroimage.2013.02.006

589 Derosiere, G., Perrey, S., 2012. Relationship between submaximal handgrip muscle force and
590 NIRS-measured motor cortical activation. Adv. Exp. Med. Biol. 737, 269–274.
591 https://doi.org/10.1007/978-1-4614-1566-4_40

592 Dhawale, A.K., Smith, M.A., Ölveczky, B.P., 2017. The Role of Variability in Motor Learning.
593 Annu. Rev. Neurosci. 40, 479–498. https://doi.org/10.1146/annurev-neuro-072116-031548

594 Döhring, J., Stoldt, A., Witt, K., Schönfeld, R., Deuschl, G., Born, J., Bartsch, T., 2017. Motor
595 skill learning and offline-changes in TGA patients with acute hippocampal CA1 lesions.
596 Cortex. 89, 156–168. https://doi.org/10.1016/j.cortex.2016.10.009

597 Doyon, J., Benali, H., 2005. Reorganization and plasticity in the adult brain during learning of
598 motor skills. Curr. Opin. Neurobiol. 15, 161–167.
599 https://doi.org/10.1016/j.conb.2005.03.004

600 Fernández-Seara, M.A., Aznárez-Sanado, M., Mengual, E., Loayza, F.R., Pastor, M.A., 2009.
601 Continuous performance of a novel motor sequence leads to highly correlated striatal and
602 hippocampal perfusion increases. Neuroimage 47, 1797–1808.
603 https://doi.org/10.1016/j.neuroimage.2009.05.061

604 Fiorillo, C.D., Newsome, W.T., Schultz, W., 2008. The temporal precision of reward prediction
605 in dopamine neurons. Nat. Neurosci. 11, 966–973. https://doi.org/10.1038/nn.2159

606 Foerde, K., Race, E., Verfaellie, M., Shohamy, D., 2013. A role for the medial temporal lobe in
607 feedback-driven learning: Evidence from amnesia. J. Neurosci. 33, 5698–5704.
608 https://doi.org/10.1523/JNEUROSCI.5217-12.2013

609 Foerde, K., Shohamy, D., 2011. Feedback timing modulates brain systems for learning in

humans. J. Neurosci. 31, 13157–13167. https://doi.org/10.1523/JNEUROSCI.2701-11.2011

Gabay, Y., Shahbari-Khateb, E., Mendelsohn, A., 2018. Feedback Timing Modulates Probabilistic Learning in Adults with ADHD. Sci. Rep. 8, 1–11. https://doi.org/10.1038/s41598-018-33551-3

Galea, J.M., Mallia, E., Rothwell, J., Diedrichsen, J., 2015. The dissociable effects of punishment and reward on motor learning. Nat Neurosci 18, 597–602. https://doi.org/10.1038/nn.3956

Harris, J.A., Harris, I.M., Diamond, M.E., 2001. The topography of tactile working memory. J. Neurosci. 21, 8262–8269. https://doi.org/10.1523/jneurosci.21-20-08262.2001

Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research, in: Power Technology and Engineering. pp. 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

Hayes, H.A., Hunsaker, N., Dibble, L.E., 2015. Implicit motor sequence learning in individuals with Parkinson disease: A meta-analysis. J. Parkinsons. Dis. 5, 549–560. https://doi.org/10.3233/JPD-140441

Klein-Flügge, M.C., Hunt, L.T., Bach, D.R., Dolan, R.J., Behrens, T.E.J., 2011. Dissociable Reward and Timing Signals in Human Midbrain and Ventral Striatum. Neuron 72, 654–664. https://doi.org/10.1016/j.neuron.2011.08.024

Kobayashi, S., Schultz, W., 2008. Influence of reward delays on responses of dopamine neurons. J. Neurosci. 28, 7837–7846. https://doi.org/10.1523/JNEUROSCI.1600-08.2008

Krakauer, J.W., Hadjiosif, A.M., Xu, J., Wong, A.L., Haith, A.M., 2019. Motor Learning 9, 613–663. https://doi.org/10.1002/cphy.c170043

Lardi, C., Billieux, J., D'Acremont, M., Linden, M. Van der, 2008. A French adaptation of a

632 short version of the Sensitivity to Punishment and Sensitivity to Reward Questionnaire
633 (SPSRQ). Pers. Individ. Dif. 45, 722–725. https://doi.org/10.1016/j.paid.2008.07.019

634 Mawase, F., Uehara, S., Bastian, A.J., Celnik, P., 2017. Motor Learning Enhances Use-
635 Dependent Plasticity. J. Neurosci. 37, 2673–2685.
636 https://doi.org/10.1523/JNEUROSCI.3303-16.2017

637 McDougle, S.D., Bond, K.M., Taylor, J.A., 2015. Explicit and Implicit Processes Constitute the
638 Fast and Slow Processes of Sensorimotor Learning. J. Neurosci. 35, 9568–9579.
639 https://doi.org/10.1523/JNEUROSCI.5061-14.2015

640 Mehler-Wex, C., Riederer, P., Gerlach, M., 2006. Dopaminergic dysbalance in distinct basal
641 ganglia neurocircuits: Implications for the pathophysiology of parkinson's disease,
642 schizophrenia and attention deficit hyperactivity disorder. Neurotox. Res. 10, 167–179.
643 https://doi.org/10.1007/BF03033354

644 Oldfield, R.C., 1971. The assessment and analysis of handedness: The Edinburgh inventory.
645 Neuropsychologia 9, 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

646 Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: transforming numbers
647 into movies. Spat. Vis. 10, 437–442. https://doi.org/10.1163/156856897X00366

648 Peterburs, J., Kobza, S., Bellebaum, C., 2016. Feedback delay gradually affects amplitude and
649 valence specificity of the feedback-related negativity (FRN). Psychophysiology 53, 209–
650 215. https://doi.org/10.1111/psyp.12560

651 Quattrocchi, G., Greenwood, R., Rothwell, J.C., Galea, J.M., Bestmann, S., 2017. Reward and
652 punishment enhance motor adaptation in stroke. J. Neurol. Neurosurg. Psychiatry jnnp-
653 2016-314728. https://doi.org/10.1136/jnnp-2016-314728

654 Schendan, H.E., Searl, M.M., Melrose, R.J., Stern, C.E., 2003. An fMRI study of the role of the
655 medial temporal lobe in implicit and explicit sequence learning. Neuron 37, 1013–1025.
656 https://doi.org/10.1016/S0896-6273(03)00123-5

657 Shadmehr, R., De Xivry, J.J.O., Xu-Wilson, M., Shih, T.Y., 2010. Temporal discounting of
658 reward and the cost of time in motor control. J. Neurosci. 30, 10507–10516.
659 https://doi.org/10.1523/JNEUROSCI.1343-10.2010

660 Shadmehr, R., Reppert, T.R., Summerside, E.M., Yoon, T., Ahmed, A.A., 2019. Movement
661 Vigor as a Re fl ection of Subjective Economic Utility. Trends Neurosci. xx, 1–14.
662 https://doi.org/10.1016/j.tins.2019.02.003

663 Sidarta, A., Vahdat, S., Bernardi, N.F., Ostry, D.J., 2016. Somatic and reinforcement-based
664 plasticity in the initial stages of human motor learning. J. Neurosci. 36, 11682–11692.
665 https://doi.org/10.1523/JNEUROSCI.1767-16.2016

666 Sidarta, X.A., Vugt, X.F.T. Van, Ostry, D.J., 2018. Somatosensory working memory in human
667 reinforcement-based motor learning 3275–3286. https://doi.org/10.1152/jn.00442.2018

668 Siegert, R.J., Taylor, K.D., Weatherall, M., Abernethy, D.A., 2006. Is implicit sequence learning
669 impaired in Parkinson's disease? A meta-analysis. Neuropsychology 20, 490–495.
670 https://doi.org/10.1037/0894-4105.20.4.490

671 Smith, M.A., Ghazizadeh, A., Shadmehr, R., 2006. Interacting Adaptive Processes with Different
672 Timescales Underlie Short-Term Motor Learning. PLoS Biol. 4, e179.
673 https://doi.org/10.1371/journal.pbio.0040179

674 Steel, A., Silson, E.H., Stagg, C.J., Baker, C.I., 2016. The impact of reward and punishment on
675 skill learning depends on task demands. Sci. Rep. 6, 1–9. https://doi.org/10.1038/srep36056

676 Sterpenich, V., van Schie, M.K.M., Catsiyannis, M., Ramyead, A., Perrig, S., Yang, H.D., Van
677 De Ville, D., Schwartz, S., 2021. Reward biases spontaneous neural reactivation during
678 sleep. Nat. Commun. 12. https://doi.org/10.1038/s41467-021-24357-5

679 Therrien, A.S., Statton, M.A., Bastian, A.J., 2020. Reinforcement Signaling Can Be Used to
680 Reduce Elements of Cerebellar Reaching Ataxia. Cerebellum.
681 https://doi.org/10.1007/s12311-020-01183-x

682 Therrien, A.S., Wolpert, D.M., Bastian, A.J., 2016. Effective Reinforcement learning following
683 cerebellar damage requires a balance between exploration and motor noise. Brain 139, 101–
684 114. https://doi.org/10.1093/brain/awv329

685 Torrubia, R., Ávila, C., Moltó, J., Caseras, X., 2001. The Sensitivity to Punishment and
686 Sensitivity to Reward Questionnaire (SPSRQ) as a measure of Gray's anxiety and
687 impulsivity dimensions. Pers. Individ. Dif. 31, 837–862. https://doi.org/10.1016/S0191-
688 8869(00)00183-5

689 Trewartha, K.M., Garcia, A., Wolpert, D.M., Flanagan, J.R., 2014. Fast but fleeting: Adaptive
690 motor learning processes associated with aging and cognitive decline. J. Neurosci. 34,
691 13411–13421. https://doi.org/10.1523/JNEUROSCI.1489-14.2014

692 Vassiliadis, P., Derosiere, G., 2020. Selecting and executing actions for rewards. J. Neurosci. 40,
693 6474–6476. https://doi.org/10.1523/JNEUROSCI.1250-20.2020

694 Vassiliadis, P., Derosiere, G., Dubuc, C., Lete, A., Crevecoeur, F., Hummel, F.C., Duque, J.,
695 2021. Reward boosts reinforcement-based motor learning. iScience 24, 102821.
696 https://doi.org/10.1016/j.isci.2021.102821

697 Vassiliadis, P., Derosiere, G., Duque, J., 2019. Beyond Motor Noise : Considering Other Causes

of Impaired Reinforcement Learning in Cerebellar Patients. Eneuro 6, 1–4. https://doi.org/10.1523/ENEURO.0458-18.2019

Weismüller, B., Ghio, M., Logmin, K., Hartmann, C., Schnitzler, A., Pollok, B., Südmeyer, M., Bellebaum, C., 2018. Effects of feedback delay on learning from positive and negative feedback in patients with Parkinson's disease off medication. Neuropsychologia 117, 46–54. https://doi.org/10.1016/j.neuropsychologia.2018.05.010