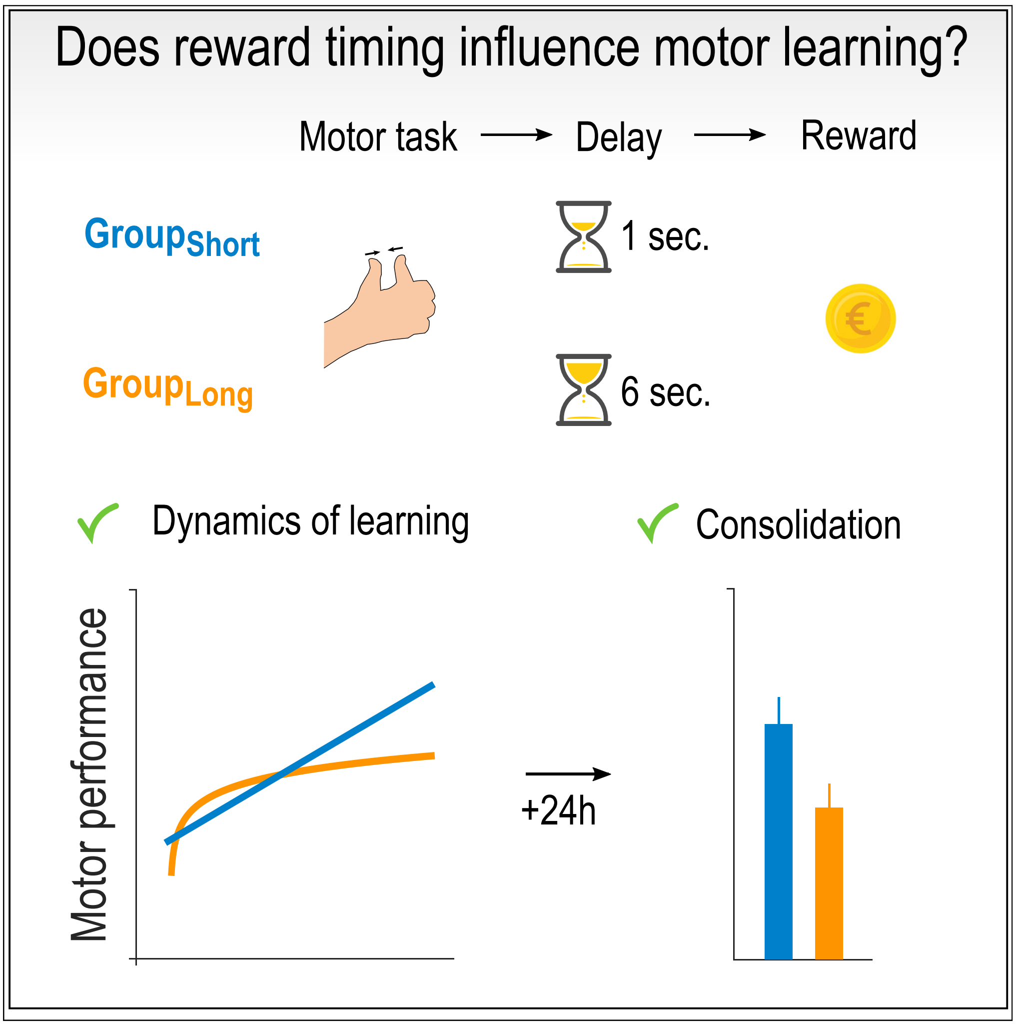
iScience

CellPress
OPEN ACCESS

Article

Reward timing matters in motor learning



Pierre Vassiliadis,
Aegryan Lete,
Julie Duque,
Gerard Derosiere

pierre.vassiliadis@uclouvain.be

Highlights

Reward timing impacts the dynamics of motor learning

Training with short reward delay generates continuous improvements in performance

Long reward delay favors fast initial learning and reduces endpoint performance

Training with long reward delay disrupts memory consolidation in learners

Vassiliadis et al., iScience *25*, 104290
May 20, 2022
https://doi.org/10.1016/j.isci.2022.104290

iScience

CellPress OPEN ACCESS

Article

Reward timing matters in motor learning

Pierre Vassiliadis,[1,2,3,*] Aegryan Lete,[1] Julie Duque,[1] and Gerard Derosiere[1]

SUMMARY

Reward timing, that is, the delay after which reward is delivered following an action is known to strongly influence reinforcement learning. Here, we asked if reward timing could also modulate how people learn and consolidate new motor skills. In 60 healthy participants, we found that delaying reward delivery by a few seconds influenced motor learning. Indeed, training with a short reward delay (1 s) induced continuous improvements in performance, whereas a long reward delay (6 s) led to initially high learning rates that were followed by an early plateau in the learning curve and a lower performance at the end of training. Participants who learned the skill with a long reward delay also exhibited reduced overnight memory consolidation. Overall, our data show that reward timing affects the dynamics and consolidation of motor learning, a finding that could be exploited in future rehabilitation programs.

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., 2016, 2020; 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 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, whereas 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). In addition, 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 et al., 2012; Foerde and Shohamy, 2011; Gabay et al., 2018; Weismüller et al., 2018), whereas 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-centered or hippocampus-centered 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., 2008, 2013). 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, impeding one to test this hypothesis directly. Here, we tested this idea by evaluating the performance of sixty healthy participants in

[1]Institute of Neuroscience, Université Catholique de Louvain, 53, Avenue Mounier, 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

[3]Lead contact

*Correspondence: pierre.vassiliadis@uclouvain.be

https://doi.org/10.1016/j.isci.2022.104290

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 continuous improvement in performance across training, whereas 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, whereas 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.

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 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 with a monetary reward. 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).

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), whereas 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 intertrial interval (ITI; 6 s in $Group_{Short}$ and 1 s in $Group_{Long}$). Before, immediately and 24 h 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 pretraining 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.

Training with long reward delays modifies the dynamics of motor learning

As a first step, we evaluated performance on the task by computing the average success rate per $Training_{Block}$ (T1 to T6, Figure S1). To compare the learning process between the groups, we performed a Linear Mixed Model (LMM), with $TRAINING_{BLOCK}$ and $GROUP_{TYPE}$ (and their interaction) modeled as categorical fixed factors. Overall, participants of both groups significantly improved their success rates over training (main effect of $TRAINING_{BLOCK}$: $F_{(5,\ 290)} = 4.30$; $p < 0.001$; Figure 2A). Most importantly, the improvement in success rate over the blocks depended on the $Group_{Type}$, as revealed by a significant $TRAINING_{BLOCK} \times GROUP_{TYPE}$ interaction ($F_{(5,\ 290)} = 2.69$; $p = 0.021$; Figure 2A). Interestingly, between-groups post hoc comparisons further revealed that endpoint performance (*i.e.*, success rate at T6) was significantly lower in $Group_{Long}$ than in $Group_{Short}$ ($p = 0.045$; Figure 2B). Note though that this significant result would not survive multiple comparisons corrections, and therefore needs to be taken with caution. Conversely, success rates at all other $Training_{Blocks}$ were comparable between the two groups (all $p > 0.22$; Figure 2A). This result suggests that reward timing influenced the dynamics of learning leading to a poorer endpoint performance in $Group_{Long}$.

To confirm these results, we ran another LMM on the single-trial Error data (Figure 3C, Table S1) with the predictors $TRAINING_{TRIAL}$ (continuous) and $GROUP_{TYPE}$ (categorical). Focusing on the Error allowed us to evaluate the effect of reward timing on motor learning without having to bin the data in any way. This analysis confirmed that learning was influenced by the timing at which rewards were provided (Figure 2D; $TRAINING_{TRIAL} \times GROUP_{TYPE}$ interaction: $F_{(1,\ 12,114)} = 9.00$; $p = 0.0027$). This interaction reflected the fact that the slope of learning (*i.e.*, a proxy of the learning rate) was steeper in $Group_{Short}$ than in $Group_{Long}$ (Figure 2D). Importantly, comparison of the intercepts in both groups did not show any significant difference ($p = 0.60$), suggesting that the learning effect could not be explained by differences in

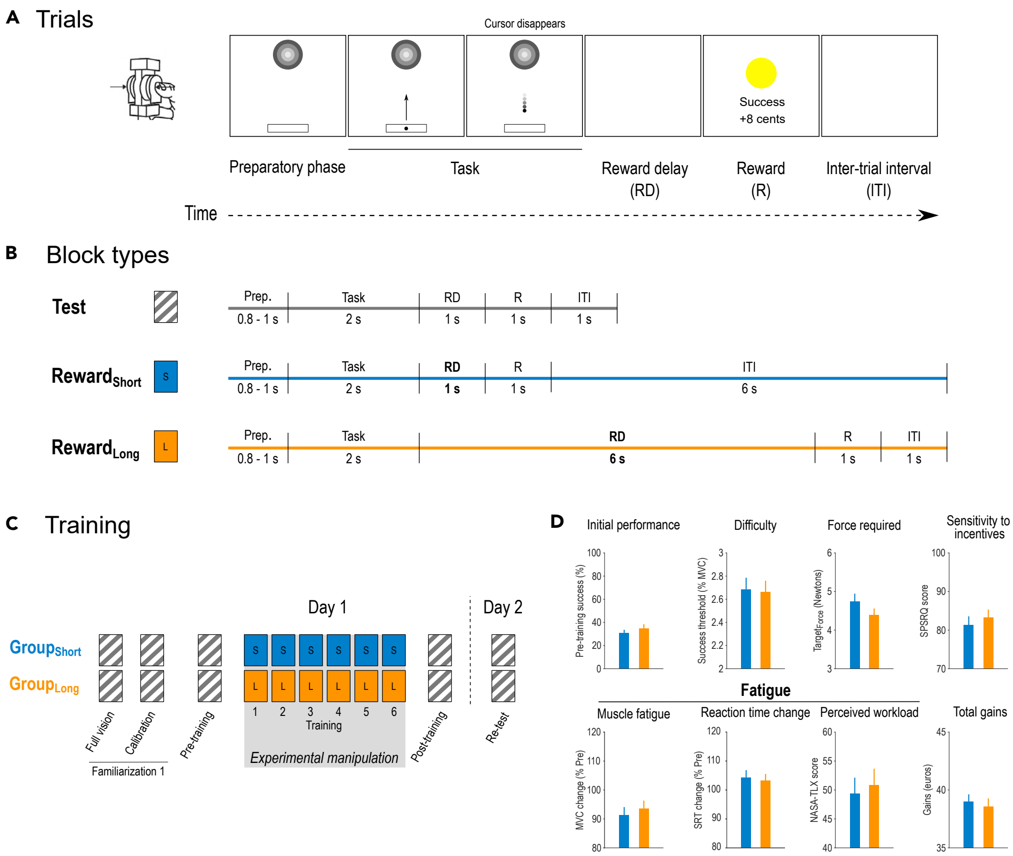


Figure 1. Motor skill learning task

(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 toward the target (as displayed here). Then, after a delay, a reward (R) appeared consisting of reinforcement feedback and a monetary reward (a successful trial is shown here). Trials ended with an intertrial 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). $\text{Reward}_{\text{Short}}$ and $\text{Reward}_{\text{Long}}$ 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 $\text{Reward}_{\text{Short}}$ and $\text{Reward}_{\text{Long}}$ 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, whereas the second one provided only partial vision and served to calibrate the difficulty of the task on an individual basis (See STAR Methods). Then, pretraining and post-training Test blocks assessments were separated by six blocks of training in the condition corresponding to each individual group ($\text{Reward}_{\text{Short}}$ for $\text{Group}_{\text{Short}}$, $\text{Reward}_{\text{Long}}$ for $\text{Group}_{\text{Long}}$). Day 2 consisted in a short re-familiarization (5 trials with full vision, not represented) followed by a retest assessment (1 Test block).

(D) Control analyses. $\text{Group}_{\text{Short}}$ and $\text{Group}_{\text{Long}}$ 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). Data are represented as mean $\pm$ SE.

initial performance. Put together, these two analyses show that training with long reward delays impairs the acquisition of a new motor skill.

An important aspect of our experimental design is that we increased the duration of the ITI in $\text{Group}_{\text{Short}}$ relative to $\text{Group}_{\text{Long}}$ (6 and 1 s, respectively; Figure 1B) to match the total duration of the trials in both groups despite differences in reward timing. To evaluate how such manipulation may have impacted 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; $\text{Group}_{\text{Short-PastStudy}}$, n = 30; from Vassiliadis et al., 2021). We

Table 1. Group features, initial performance and fatigue in the three experimental groups (mean ± SE)

	$Group_{Short}$ (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

The two last columns provide the results of independent samples t-tests.

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 $Group_{Long}$. As described previously, we ran a first LMM on the Success data with the factors $TRAINING_{BLOCK}$ and $GROUP_{TYPE}$. Consistent with our hypothesis, we found a significant $TRAINING_{BLOCK}$ × $GROUP_{TYPE}$ interaction ($F_{(10,\ 435)} = 2.84$; $p = 0.0020$) and post hoc tests showed (1) no significant difference between $Group_{Short\text{-}PastStudy}$ and $Group_{Short}$, at any $Training_{Block}$ (all $p > 0.22$) and (2) a marginally significant difference in endpoint performance when comparing $Group_{Short\text{-}PastStudy}$ and $Group_{Long}$ (*i.e.*, $p = 0.048$ and 0.052 at T5 and T6, respectively; Figures S2A and S2B). To confirm these effects on non-binned, single-trial data, we ran the same LMM on the Error variable (*i.e.*, same analysis as in Figure 2D) but with the addition of the data from $Group_{Short\text{-}PastStudy}$ (Figure S2C). Again, there was a $TRAINING_{TRIAL}$ × $GROUP_{TYPE}$ interaction ($F_{(2,\ 18,213)} = 14.99$; $p < 0.001$), that was driven by differences in the slopes of the learning curves between the groups (Figure S2D). As expected, post hoc tests showed that the slopes were steeper in $Group_{Short\text{-}PastStudy}$ than in $Group_{Long}$ ($p < 0.001$). However, slopes were also steeper in $Group_{Short\text{-}PastStudy}$ than in $Group_{Short}$ ($p = 0.018$), suggesting that longer ITIs may also have some detrimental effect on the learning rates. Notably, no differences were found when comparing the intercepts (all $p > 0.59$). The apparent discrepancy between the LMM results obtained for Success vs. for Error data possibly arises from the fact that the former analysis was based on block-averaged performance, while the latter focused on the learning rates estimated based on single-trial data (reflected by the coefficient associated with $TRAINING_{TRIAL}$ in the LMM). Notably, the difference in learning rate between $Group_{Short}$ and $Group_{Short\text{-}PastStudy}$ must be taken with caution as in addition to presenting a longer ITI duration, $Group_{Short}$ also presented a slightly longer reward delay relative to $Group_{Short\text{-}PastStudy}$ (1 s vs. 0.5 s, respectively), which might also have been detrimental for learning rates. Still, if anything, this analysis suggests that long ITIs had rather a negative impact on learning. Yet, participants of $Group_{Short}$ exhibited better learning rates than participants of $Group_{Long}$, indicating that the positive effect of the shorter reward delays overcame the negative impact of the longer ITI in this group. Overall, this analysis suggests that both reward delay and ITI duration influence motor skill learning but that reward delay plays a more prominent role in shaping learning. A direct corollary to this is that we may have underestimated the negative impact of long reward delays on learning when comparing $Group_{Long}$ with $Group_{Short}$, given that this negative effect was partially counteracted by the longer ITI duration in the latter group.

To evaluate total learning, we computed success rates at post-training, which was performed in a Test block setting in both groups. Importantly, whereas comparing performance at T6 informed us about the effect of our particular training features on learning within each training condition, Post-training performance provides information about total learning in the task, in identical Test block conditions. Overall, success rates at post-training increased by 22.8 ± 4.69% in $Group_{Short}$, and 14.5 ± 6.42% in $Group_{Long}$ with respect to pretraining. Interestingly, success rates at post-training were significantly different from 0 in $Group_{Short}$ despite Bonferroni correction of the significance threshold (cutoff for significance: $p = 0.025$; $t_{(29)} = 4.87$, $p < 0.001$), indicative of a significant improvement in performance with respect to Pre-training. In contrast, success rates at Post-training were not significantly different from 0 in $Group_{Long}$ after

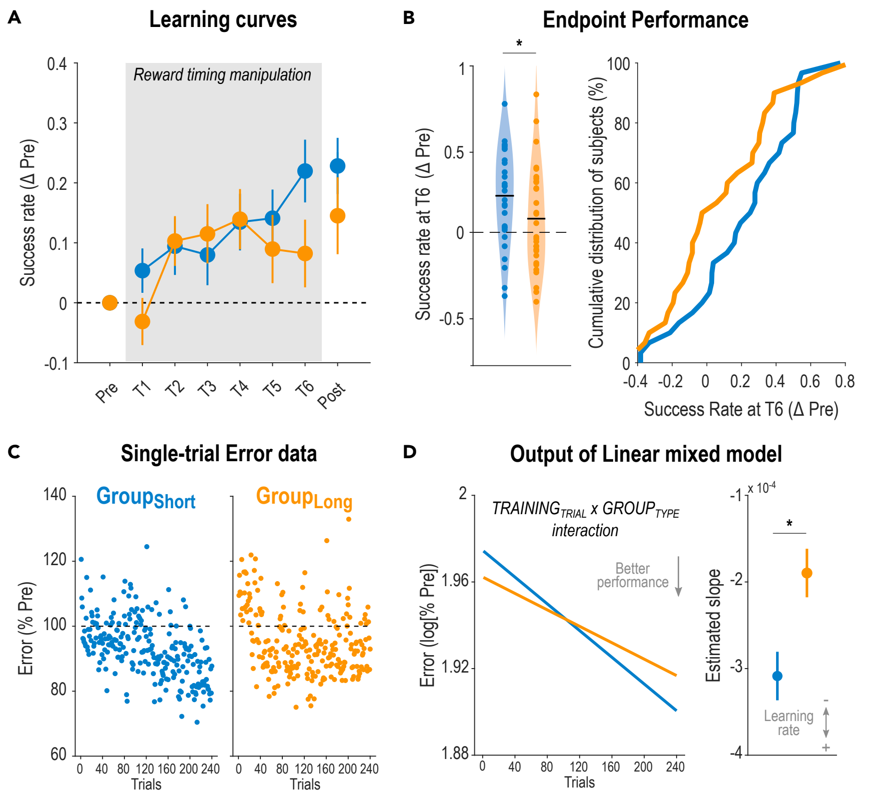


Figure 2. Effect of reward timing on motor skill learning

(A) Learning curves Proportion of successful trials (expressed as a difference with the individual pretraining success rate) is represented across practice for the two experimental groups (blue: $Group_{Short}$, n = 30, orange: $Group_{Long}$, n = 30). The gray shaded area highlights the blocks concerned by the reward timing manipulation. The remaining blocks were test blocks.

(B) Endpoint performance. Violin plot showing success rates at the end of the training period (*i.e.*, at T6) for each participant (left panel) and the corresponding cumulative distributions of the data (right panel).

(C) Single-trial error data. Normalized error data obtained during training are averaged across groups and plotted for each single trial. Note that lower Errors were associated with better performance.

(D) Output of LMM on the error data. Output of LMM run on the log-transformed error data is plotted for each group (left panel). The error data was log-transformed before respect key assumptions of LMMs (see STAR Methods section). The significant $TRAINING_{TRIAL}$ × $GROUP_{TYPE}$ interaction shows that the slope of learning was steeper in $Group_{Short}$ than in $Group_{Long}$ (right panel). Estimated intercepts were not different between groups (p = 0.60). Notably, more negative slopes reflect larger learning rates. *: significant difference (p < 0.05; F-tests on LMM coefficients). Data are represented as mean ± SE.

Bonferroni correction of the significance threshold ($t_{(29)} = 2.26$, p = 0.031). However, a *t*-test on these data did not show any significant difference between the $Group_{TYPES}$ ($t_{(58)} = 1.05$; p = 0.30). Hence, reward timing only induced a subtle change in total learning that did not reach significance when comparing directly the groups.

Results of the first analysis showed that training with long reward delays was generally associated with lower learning rates (Figure 2D), leading to a reduced endpoint performance (Figure 2B). Inspection of the raw data (Figures 2A and 2C) also suggested that the learning dynamics could be different between the groups. To evaluate this, we ran three additional analyses. First, we asked each group of participants whether the learning curves were best modeled as a linear or non-linear logarithmic function. Interestingly, we found that the data from $Group_{Short}$ were better approximated by a linear function (linear fit: Adjusted R^2 = 0.25; logarithmic fit: Adjusted R^2 = 0.21; Figure 3A), whereas the $Group_{Long}$ learning curve was better modeled with a logarithmic fit (linear fit: Adjusted R^2 = 0.063; logarithmic fit: Adjusted R^2 = 0.18; Figure 3B). This suggests that training with short reward delays was associated with generally stable learning rates, whereas training with long reward delays was related to fast learning rates early on during practice that

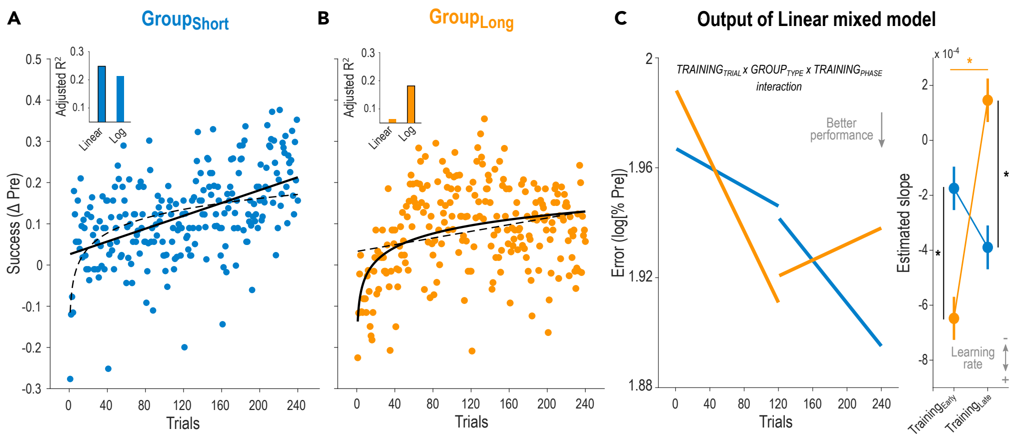


Figure 3. Effect of reward timing on the dynamics of learning

(A and B) Linear and nonlinear fits on Success learning data. The group-averaged single-trial Success data in $\text{Group}_{\text{Short}}$ (A, n = 30) and $\text{Group}_{\text{Long}}$ (B, n = 30) were fitted with either a linear or a nonlinear logarithmic function. Importantly, the best fit (i.e., represented by the solid trace) was linear for $\text{Group}_{\text{Short}}$ and logarithmic for $\text{Group}_{\text{Long}}$, suggesting that the dynamics of learning were different in both groups.

(C) Output of the LMM on the error data including the factor $\text{TRAINING}_{\text{PHASE}}$. Output of LMM run on the log-transformed Error data is plotted for each group (left panel). The significant $\text{TRAINING}_{\text{TRIAL}}$ x $\text{GROUP}_{\text{TYPE}}$ × $\text{TRAINING}_{\text{PHASE}}$ interaction shows that, in the early phase of practice, the learning rates – reflected by the slope of learning – were steeper in $\text{Group}_{\text{Long}}$ than in $\text{Group}_{\text{Short}}$ ($p < 0.001$). This was the opposite in the late phase of practice ($p < 0.001$). Notably, there was also a significant reduction of the learning rates from $\text{Training}_{\text{Early}}$ to $\text{Training}_{\text{Late}}$ in $\text{Group}_{\text{Long}}$ ($p < 0.001$; orange star), while there was a tendency for an increase in learning rates in $\text{Group}_{\text{Short}}$ ($p = 0.056$). Note that lower Errors were associated with better performance and that more negative slopes reflect larger learning rates. *: significant difference ($p < 0.05$; F-tests on LMM coefficients). Data are represented as mean ± SE.

was quickly followed by a plateau in performance. Indeed, simple linear regressions on the success data showed that 76.7% (23/30) of participants of $\text{Group}_{\text{Long}}$ exhibited higher learning rates in the early than in the late phase of training, whereas this percentage was 46.7% (14/30) in $\text{Group}_{\text{Short}}$ (Fisher exact test on the proportions: $p = 0.033$; Figure S3). To further evaluate how learning rates varied across early and late phases of practice, we ran the same LMM on the Error data as described above (Figure 2D) with the addition of the factor $\text{TRAINING}_{\text{PHASE}}$ which was modeled as a categorical fixed effect with two modalities ($\text{Training}_{\text{Early}}$ vs. $\text{Training}_{\text{Late}}$ for the first and last 120 trials of training, respectively; Table S2). Interestingly, we found a triple $\text{TRAINING}_{\text{TRIAL}}$ x $\text{GROUP}_{\text{TYPE}}$ × $\text{TRAINING}_{\text{PHASE}}$ interaction ($F_{(112,110)} = 40.62$; $p < 0.001$), demonstrating that learning rates (reflected by the coefficients associated with the factor $\text{TRAINING}_{\text{TRIAL}}$) varied not only depending on the group but also based on the phase of practice. As illustrated on Figure 3C, this interaction was due to the fact that at $\text{Training}_{\text{Early}}$, the estimated learning rate was significantly higher in $\text{Group}_{\text{Long}}$ than in $\text{Group}_{\text{Short}}$, whereas it was the opposite at $\text{Training}_{\text{Late}}$ (both $p < 0.001$). Moreover, learning rates were significantly higher at $\text{Training}_{\text{Early}}$ than at $\text{Training}_{\text{Late}}$ in $\text{Group}_{\text{Long}}$ ($p < 0.001$). In $\text{Group}_{\text{Short}}$, there was a trend for the opposite effect (*i.e.*, higher learning rates at $\text{Training}_{\text{Late}}$ than at $\text{Training}_{\text{Early}}$; $p = 0.056$). Again, intercepts at $\text{Training}_{\text{Early}}$ were not significantly different ($p = 0.36$), indicating comparable initial levels of performance in the two groups. Hence, this analysis confirms that reward timing impacts learning dynamics. More specifically, training with short reward delay appears to induce continuous gains in performance during training, whereas long reward timings favor nonlinear dynamics with larger initial learning rates that then drop significantly, indicative of a plateau in learning.

Training with long reward delay impairs overnight skill consolidation in learners

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). To evaluate consolidation, we ran an LMM on the normalized success rates obtained at post-training of Day 1 and at retest of Day 2 (*i.e.*, both performed in a test block setting) with $\text{TEST}_{\text{BLOCK}}$ and $\text{GROUP}_{\text{TYPE}}$ as fixed effects. This analysis did not reveal any main effect of $\text{TEST}_{\text{BLOCK}}$ ($F_{(1,\ 58)} = 0.75$; $p = 0.39$) and $\text{GROUP}_{\text{TYPE}}$ ($F_{(1,\ 86.14)} = 1.18$; $p = 0.28$) nor any $\text{TEST}_{\text{BLOCK}}$ × $\text{GROUP}_{\text{TYPE}}$ interaction ($F_{(1,\ 58)} = 0.48$; $p = 0.49$). The same results were obtained when running the LMM on the single-trial Error data. However, a potential caveat of theses analyses is that they included

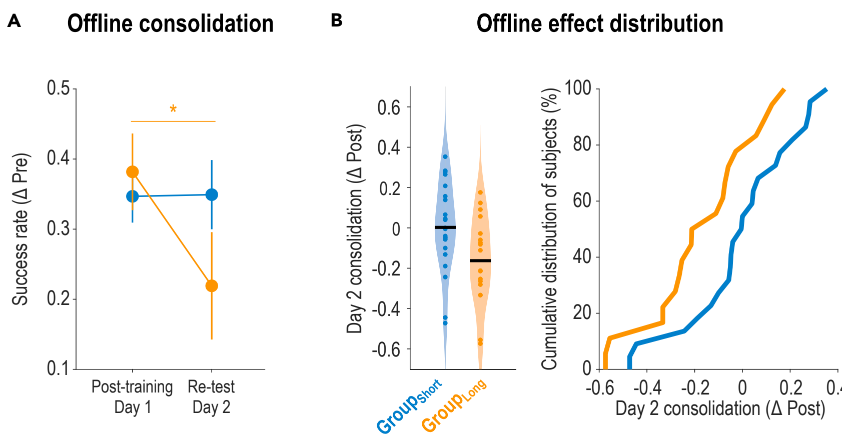


Figure 4. Effect of reward timing on overnight consolidation of the motor skill

(A) Offline consolidation of the motor skill. Proportion of successful trials (expressed as a difference with the individual pretraining success rate) at post-training (Day 1) and retest (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 $TEST_{PHASE} \times GROUP_{TYPE}$ interaction showed that while $Group_{Short}$ participants performed comparably well in post-training and retest, demonstrating offline consolidation of the skill, this process was altered in $Group_{Long}$.

(B) Offline effect distribution. Violin plot showing the offline effect (Success rate in Day 2 – Success rate in Post-training) for each participant (left panel) and the corresponding cumulative distributions of the data (right panel). *: significant difference ($p < 0.05$; F-test on LMM coefficients). Data are represented as mean ± SE.

participants who did not learn the task on Day 1 and even exhibited a deterioration of performance with practice. In these participants, a retest performance similar to the pretraining level would be considered as evidence for an offline gain in performance, when it would actually only reflect a return to the baseline level of performance. In a second step, we therefore focused on the learners – that is, participants who exhibited an improvement of performance with practice on Day 1 (n = 22 and 18 in $Group_{Short}$ and $Group_{Long}$, respectively). This allowed us to compare offline consolidation in participants who actually responded to the training and who also happened to be very close in terms of Post-training success rates (Figure 4A), a crucial aspect in order to interpret any overnight change in performance. Interestingly, this analysis revealed a $TEST_{PHASE} \times GROUP_{TYPE}$ interaction ($F_{(1, 38)} = 5.77$; $p = 0.021$). In fact, as mentioned previously, performance was strongly similar between learners of the two groups at post-training on Day 1 ($p = 0.65$) but diverged between the groups on Day 2. Indeed, success rates were significantly reduced on Day 2 relative to Day 1 in $Group_{Long}$ ($p = 0.0021$) but remained stable from one day to another in $Group_{Short}$ ($p = 0.96$, Figure 4B). The difference in performance on Day 2 between the groups was only at the trend level ($p = 0.096$). Notably, this interaction was replicated when running the LMM on the single-trial Error data ($F_{(12,615.9)} = 7.25$; $p = 0.0071$). This indicates that delaying rewards on Day 1 impaired consolidation of the motor skill on Day 2 in learners. 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.

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 continuous gains in performance, whereas 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, whereas 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.

An important finding of our study is the overall impairment of learning when training with long compared to short reward delays, which was reflected by a reduction of global learning rates as well as endpoint performance during training. 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 observed in the subjects of Group$_{Long}$. Another complementary interpretation is that reward delays affected the precision of dopaminergic reward prediction errors in the striatum (Fiorillo et al., 2008; Kobayashi and Schultz, 2008). In this case, the temporal uncertainty caused by increased reward delays would alter the association between the movement and the corresponding outcome because of imprecise learning signals in the reward system (Fiorillo et al., 2008). Overall, the present data indicate that the temporal contingency between movements and rewards is a decisive aspect of reward-based motor learning.

Despite clear effects of reward delay during the training phase, we did not find any between-group difference at post-training (*i.e.*, performed in a test block setting, with short reward delay and ITI). There are several ways to interpret this finding. First, it is possible that reward timing has dissociable effects on motor performance and learning (Schmidt and Bjork, 1992; Soderstrom and Bjork, 2015). As such, the introduction of reward delays during training may generally alter motor performance, but not the learning of the skill, as evaluated in the post-training test block. A second interpretation is that the reward timing manipulation affected the learning process but was not sufficient to evoke lasting behavioral differences. This would be in line with previous work on associative learning showing that reward delays modulate brain signatures of reward processing in healthy subjects but not behavioral learning in the test phase (Foerde and Shohamy, 2011). Yet, the same researchers also found robust learning effects when testing populations of patients that presented specific dysfunctions of the striatum or the hippocampus (Foerde et al., 2012, 2013; Foerde and Shohamy, 2011). A possibility is therefore that our reward delay manipulation was not sufficient to modulate behavioral learning in young healthy individuals (potentially because of other compensatory learning mechanisms) but may still prove efficient when testing populations of patients exhibiting specific lesions of the networks involved in reward processing.

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 behavior and is thought to reflect implicit learning (McDougle et al., 2015; Trewartha et al., 2014). In contrast, the fast process entails both a high learning rate and a quick forgetting of the new behavior and supports the explicit learning of new motor behaviors (McDougle et al., 2015; Trewartha et al., 2014). The nature of our task did not allow us to evaluate the relationship between reward timing and the relative contribution of implicit and explicit learning. Still, people who trained with a short reward delay exhibited learning dynamics that presented a low initial learning rate and a clear overnight consolidation – reminiscent of the slow process, whereas those who trained with a long reward delay exhibited a high initial learning rate and an overnight forgetting of the motor memory – evocative of the fast process. Based on these results, one may suggest that short reward delays preferentially facilitate the slow (putatively more implicit) process, whereas long reward delays may favor the fast (potentially more explicit) learning process, accentuating their respective contribution to subjects' improvements. Interestingly, the striatum and hippocampus, which are involved in processing rewards offered after short and long delays, respectively (Foerde et al., 2012, 2013; Foerde and Shohamy, 2011), exhibit a pattern of activation during motor learning that is consistent with this interpretation. As such, the striatum displays slow, continuous changes in activity over the course of motor learning whereas the hippocampus usually exhibits a fast increase in activity in the early phase of learning that wanes later on (Albouy et al., 2008, 2012, 2013; Doyon et al., 2018; Rieckmann et al., 2010; Schendan et al., 2003). Notably though, this parallel between our behavioral results and previous neurophysiological findings in motor learning needs to be taken with caution as the aforementioned studies mainly used motor sequence learning tasks that may engage partially different brain mechanisms than our motor skill learning task (Krakauer et al., 2019). Altogether, these elements suggest that the different learning dynamics observed in individuals training with short and long reward delays could result from the preferential engagement of distinct brain networks that exhibit different activation patterns during motor learning.

The impairment of motor consolidation observed in subjects who trained with a long reward delay also suggests that reward timing does not only affect the acquisition of the skill, but also the offline processing of the acquired motor memory. The reduction of overnight consolidation in learners of $\mathrm{Group}_{\mathrm{Long}}$ may appear discordant with previous work showing improved episodic memory consolidation after training with long reward delays (Foerde and Shohamy, 2011). Notably though, the beneficial effect of long reward delays on episodic memory previously reported was not observed in Parkinson's disease patients nor in their age-matched controls (Foerde et al., 2012). Our results may also seem to differ from those of former motor learning studies showing consolidation improvements in hippocampal-related skills (Albouy et al., 2008, 2015). However, an important difference with respect to these studies is the nature of our task. As such, the hippocampus is known to be involved to various degrees in motor learning depending on the type of skill that is practiced (McDougle et al., 2022), contributing more to learning in settings requiring to build a spatial representation of the task (Albouy et al., 2015) or to learn a perceptual component (Rose et al., 2011). Although the hippocampus is potentially involved in skill learning tasks involving the flexible selection of force parameters (i.e., as in the current study, McDougle et al., 2022), its engagement may have been limited as learning did not involve a strong spatial or perceptual component. Another 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., 2010, 2019), reducing their beneficial effect on offline consolidation mechanisms (Ambrose et al., 2016; Sterpenich et al., 2021).

Beyond reward timing, another feature that could have altered both the learning dynamics and consolidation 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). First, the comparison of $\mathrm{Group}_{\mathrm{Short}}$ and $\mathrm{Group}_{\mathrm{Short\text{-}Replication}}$ suggests that lengthening the post-reward delay had a rather negative impact on the learning dynamics, inducing a reduction in learning rates (Figure S2). Despite this detrimental impact, participants of $\mathrm{Group}_{\mathrm{Short}}$ (ITI = 6 s) still exhibited better learning rates than participants of $\mathrm{Group}_{\mathrm{Long}}$ (ITI = 1 s), suggesting that the positive effect of the shorter reward delays overcame the negative impact of the longer ITI in $\mathrm{Group}_{\mathrm{Short}}$. Overall, this analysis suggests that both reward delay and ITI duration influence motor learning but that reward delay plays a more prominent role in shaping learning. Second, the presence of resting periods of a few seconds during learning was recently shown to induce a rapid form of consolidation during motor sequence learning (Bönstrup et al., 2019, 2020; Buch et al., 2021; Jacobacci et al., 2020). We cannot rule out that the longer ITI experienced by $\mathrm{Group}_{\mathrm{Short}}$ could have facilitated this form of consolidation. Notably though, this rapid form of consolidation was not correlated with overnight consolidation, suggesting different mechanisms for between-trials and between-days consolidation (Bönstrup et al., 2019). Hence, we believe it is unlikely that the longer ITIs in $\mathrm{Group}_{\mathrm{Short}}$ drove the effect of reward timing on overnight consolidation.

In conclusion, our data indicate that the timing at which reward is delivered during motor training alters the dynamics of learning and the consolidation of the new motor memory. Research is now required to gain further knowledge as to the brain networks involved in these time-dependent effects of reward on motor learning. Such knowledge would prove useful for the design of future reward-based rehabilitation programs, in which reward timing may be individualized depending on the brain networks and learning processes affected in specific populations of patients. For instance, short reward delays may be preferred during rehabilitation when brain lesions affect the medial temporal lobe (Foerde et al., 2013), whereas long reward delays may prove more efficient when patients suffer from dysfunction of the striatal network (Foerde et al., 2012; Foerde and Shohamy, 2011; Gabay et al., 2018; Weismüller et al., 2018). In addition, our study suggests that short reward delays and short ITIs should be generally preferred in motor rehabilitation when the motor deficit is not associated with any lesion of the reward circuitry, as occurs after spinal cord injury or lesions of the peripheral nervous system.

Limitations of the study

Even if initial performance was not significantly different between the groups in any analysis, the fact that it was slightly lower in $\mathrm{Group}_{\mathrm{Long}}$ may have caused an overestimation of early learning rates in this group. In this case, higher early learning rates in $\mathrm{Group}_{\mathrm{Long}}$ (relative to $\mathrm{Group}_{\mathrm{Short}}$) would reflect a quick recovery from an initial perturbation caused by the introduction of long reward delays at T1. The present data do not allow us to rule out this interpretation completely. Notably, although a decrement in initial performance in $\mathrm{Group}_{\mathrm{Long}}$ may have contributed to bias our estimation of early learning rates, it cannot explain the between-group differences observed when considering the late phase of training, strongly suggestive of an effect of reward timing on learning dynamics.

Relatedly, the nature of our research question required us to employ different timings in the Training and in the Test blocks. As such, in Reward$_{Short}$ blocks, reward delay (1 s) was identical to the Test blocks but the ITI (6 s) was different. Conversely, in Reward$_{Long}$ blocks, ITI duration was identical to the Test blocks (1 s) but the reward delay was different (6 s). Therefore, strictly speaking, the overall similarity between the training and the Test blocks was identical in both groups. However, our results suggest that changes in reward delay have a stronger impact on motor performance than changes in ITIs, implying that performance in the Test blocks may be more affected in Group$_{Long}$ due to the difference in the reward delay experienced during training versus during the Test block. One may hypothesize that this could have subsequently altered performance on Day 2, which was reduced in learners of Group$_{Long}$. Even if we cannot definitely refute or confirm this hypothesis, we believe that it is unlikely. First, if this was true, Group$_{Long}$ should be more disturbed than Group$_{Short}$ when transitioning from the end of the training phase to the post-training Test block. Importantly though, we observed the opposite pattern of results, with a tendency to improve performance from training to post-training. Second, analysis of consolidation showed a reduction of performance on Re-test (on Day 2) compared to post-training in learners of Group$_{Long}$, with both assessments being test blocks. Any disturbance of Group$_{Long}$ subjects because of the difference between the reward delay experienced during training and test blocks should have similarly affected both post-training and re-test blocks. Overall, characterizing the impact of dynamic changes in reward delay on motor performance represents an interesting avenue for future research.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Motor skill learning task
  - Sensory and reinforcement feedbacks
  - Reward timing manipulation
  - Motor skill learning protocol
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Motor skill learning
  - Motor skill consolidation
  - Group features, initial performance and fatigue

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.104290.

ACKNOWLEDGMENTS

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.

AUTHOR CONTRIBUTIONS

Conceptualization, P.V., A.L., J.D., and G.D.; Methodology, P.V., A.L., J.D., and G.D.; Formal Analysis, P.V.; Investigation, P.V. and A.L.; Data Curation, P.V. and A.L.; Writing – Original Draft, P.V.; Writing – Review & Editing, P.V., A.L., J.D., and G.D.; Visualization, P.V.; Funding Acquisition, P.V., J.D., and G.D.; Supervision, J.D. and G.D.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 21, 2021
Revised: February 25, 2022
Accepted: April 20, 2022
Published: May 20, 2022

REFERENCES

Abe, M., Schambra, H., Wassermann, E.M., Luckenbaugh, D., Schweighofer, N., and 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., Fogel, S., King, B.R., Laventure, S., Benali, H., Karni, A., Carrier, J., Robertson, E.M., and Doyon, J. (2015). Maintaining vs. enhancing motor sequence memories: respective roles of striatal and hippocampal systems. Neuroimage *108*, 423–434. https://doi.org/10.1016/j.neuroimage.2014.12.049.

Albouy, G., King, B.R., Maquet, P., and 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., et al. (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.

Albouy, G., Sterpenich, V., Vandewalle, G., Darsaud, A., Gais, S., Rauchs, G., Desseilles, M., Boly, M., Dang-Vu, T., Balteau, E., et al. (2012). Neural correlates of performance variability during motor sequence acquisition. Neuroimage *60*, 324–331. https://doi.org/10.1016/j.neuroimage.2011.12.049.

Ambrose, R.E., Pfeiffer, B.E., and Foster, D.J. (2016). Reverse replay of hippocampal place cells is uniquely modulated by changing reward. Neuron *91*, 1124–1136. https://doi.org/10.1016/j.neuron.2016.07.047.

Bernardi, N.F., Darainy, M., and 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.

Bönstrup, M., Iturrate, I., Hebart, M.N., Censor, N., and Cohen, L.G. (2020). Mechanisms of offline motor learning at a microscale of seconds in large-scale crowdsourced data. Npj Sci. Learn. *5*, 1–10. https://doi.org/10.1038/s41539-020-0066-9.

Bönstrup, M., Iturrate, I., Thompson, R., Cruciani, G., Censor, N., and Cohen, L.G. (2019). A rapid form of offline consolidation in skill learning. Curr. Biol. *29*, 1346–1351.e4. https://doi.org/10.1016/j.cub.2019.02.049.

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

Buch, E.R., Claudino, L., Quentin, R., Bönstrup, M., and Cohen, L.G. (2021). Consolidation of human skill linked to waking hippocampo-neocortical replay. Cell Rep. *35*. https://doi.org/10.1016/j.celrep.2021.109193.

Chen, X., Holland, P., and 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., and 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., and Perrey, S. (2014). Similar scaling of contralateral and ipsilateral cortical responses during graded unimanual force generation. Neuroimage *85*, 471–477. https://doi.org/10.1016/j.neuroimage.2013.02.006.

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

Dhawale, A.K., Smith, M.A., and Ölveczky, B.P. (2017). The role of variability in motor learning. Annu. Rev. Neurosci. *40*, 479–498. https://doi.org/10.1146/annurev-neuro-072116-031548.

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

Doyon, J., Gabitov, E., Vahdat, S., Lungu, O., and Boutin, A. (2018). Current issues related to motor sequence learning in humans. Curr. Opin. Behav. Sci. *20*, 89–97. https://doi.org/10.1016/j.cobeha.2017.11.012.

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

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

Foerde, K., Braun, E.K., and Shohamy, D. (2012). A Trade-Off between feedback-based learning and episodic memory for feedback events: evidence from Parkinson's disease. Neurodegener. Dis. *11*, 93–101. https://doi.org/10.1159/000342000.

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

Foerde, K., and 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., and 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., and 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., and 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., and 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.

Jacobacci, F., Armony, J.L., Yeffal, A., Lerner, G., Amaro, E., Jovicich, J., Doyone, J., and Della-Maggiore, V. (2020). Rapid hippocampal plasticity supports motor sequence learning. Proc. Natl. Acad. Sci. U S A *117*, 23898–23903. https://doi.org/10.1073/pnas.2009576117.

Klein-Flügge, M.C., Hunt, L.T., Bach, D.R., Dolan, R.J., and 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., and 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., and Haith, A.M. (2019). Mot. Learn. *9*, 613–663. https://doi.org/10.1002/cphy.c170043.

Lardi, C., Billieux, J., D'Acremont, M., and Van der Linden, M. (2008). A French adaptation of a short version of the sensitivity to punishment and sensitivity to reward questionnaire (SPSRQ). Pers. Indiv. Differ. *45*, 722–725. https://doi.org/10.1016/j.paid.2008.07.019.

Mawase, F., Uehara, S., Bastian, A.J., and Celnik, P. (2017). Motor learning enhances use-dependent plasticity. J. Neurosci. *37*, 2673–2685. https://doi.org/10.1523/JNEUROSCI.3303-16.2017.

McDougle, S.D., Bond, K.M., and Taylor, J.A. (2015). Explicit and implicit processes constitute the fast and slow processes of sensorimotor learning. J. Neurosci. *35*, 9568–9579. https://doi.org/10.1523/JNEUROSCI.5061-14.2015.

sequence learning. Neuron *37*, 1013–1025. https://doi.org/10.1016/S0896-6273(03)00123-5.

McDougle, S.D., Wilterson, S.A., Turk-Browne, N.B., and Taylor, J.A. (2022). Revisiting the role of the medial temporal lobe in motor learning. J. Cogn. Neurosci. *34*, 532–549. https://doi.org/10.1162/jocn_a_01809.

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

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

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

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

Quattrocchi, G., Greenwood, R., Rothwell, J.C., Galea, J.M., and Bestmann, S. (2017). Reward and punishment enhance motor adaptation in stroke. J. Neurol. Neurosurg. Psychiatry *88*, 730–736. jnnp-2016-314728. https://doi.org/10.1136/jnnp-2016-314728.

Rieckmann, A., Fischer, H., and Bäckman, L. (2010). Activation in striatum and medial temporal lobe during sequence learning in younger and older adults: relations to performance. Neuroimage *50*, 1303–1312. https://doi.org/10.1016/j.neuroimage.2010.01.015.

Rose, M., Haider, H., Salari, N., and Büchel, C. (2011). Functional dissociation of Hippocampal mechanism during implicit learning based on the domain of associations. J. Neurosci. *31*, 13739–13745. https://doi.org/10.1523/JNEUROSCI.3020-11.2011.

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

Schmidt, R.A., and Bjork, R.A. (1992). New conceptualizations of practice: common principles in three paradigms suggest new concepts for training. Psychol. Sci. *3*, 207–217. https://doi.org/10.1111/j.1467-9280.1992.tb00029.x.

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

Shadmehr, R., Reppert, T.R., Summerside, E.M., Yoon, T., and Ahmed, A.A. (2019). Movement vigor as a Re fl ection of subjective economic utility. Trends Neurosci. *xx*, 1–14. https://doi.org/10.1016/j.tins.2019.02.003.

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

Sidarta, X.A., van Vugt, X.F.T., and Ostry, D.J. (2018). Somatosensory working memory in human reinforcement-based motor learning. J Neurophysiol. *120*, 3275–3286. https://doi.org/10.1152/jn.00442.2018.

Smith, M.A., Ghazizadeh, A., and Shadmehr, R. (2006). Interacting adaptive processes with different timescales underlie short-term motor learning. PLoS Biol. *4*, e179. https://doi.org/10.1371/journal.pbio.0040179.

Soderstrom, N.C., and Bjork, R.A. (2015). Learning versus performance: an integrative review. Perspect. Psychol. Sci. *10*, 176–199. https://doi.org/10.1177/1745691615569000.

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

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

Therrien, A.S., Statton, M.A., and Bastian, A.J. (2020). Reinforcement signaling can be used to reduce elements of cerebellar reaching ataxia. Cerebellum *20*, 62–73. https://doi.org/10.1007/s12311-020-01183-x.

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

Torrubia, R., Ávila, C., Moltó, J., and Caseras, X. (2001). The sensitivity to punishment and sensitivity to reward questionnaire (SPSRQ) as a measure of gray's anxiety and impulsivity dimensions. Pers. Indiv. Differ. *31*, 837–862. https://doi.org/10.1016/S0191-8869(00)00183-5.

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

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

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

Vassiliadis, P., Derosiere, G., and 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., and 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.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Motor learning data ('All_Var_table.mat')	This paper	https://osf.io/4kqpe/
Subjects characteristics ('Subjects_characteristics_Timing.xlsx')	This paper	https://osf.io/4kqpe/
Software and algorithms		
Matlab vR2007 7.5 and R2018a	Mathworks	www.mathworks.com/products/matlab.html
Statistica 10	StatSoft Inc.	https://www.statistica.com/en/
Psychophysics Toolbox	Psychtoolox.org	http://psychtoolbox.org/

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

- Motor learning data (All_Var_table.mat') and de-identified subjects characteristics ('Subjects_characteristics_Timing.xlsx') are freely available via an open-access data sharing repository (https://osf.io/4kqpe/).
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

A total of sixty right-handed healthy volunteers participated in the present study (46 women, 23.7 $\pm$ 0.3 years old; mean $\pm$ SE). Data from a previous group of thirty participants was also re-analyzed (20 women, 23.9 $\pm$ 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 compensated for their participation. Finally, all participants were asked to fill out a French adaptation of the Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ; (Lardi et al., 2008; Torrubia et al., 2001)) and a NASA Task Load Index questionnaire (NASA-TLX, (Hart and Staveland, 1988)).

METHOD DETAILS

Motor skill learning task

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

Reward timing manipulation

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

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

Motor skill learning protocol

Subjects were tested on two consecutive days (Day 1 and Day 2; Figure 1C). On Day 1, we first measured the individual MVC to calculate the $Target_{Force}$. Notably, MVCs and simple 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 h 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 we verified the homogeneity of the variances systematically and non-parametric tests were used when variances were non-homogeneous. Linear mixed models (LMM) were fitted using the *fitlme* function in Matlab, with the restricted maximum likelihood fitting method. As random effects, we added intercepts for participants. Normality of residuals, skewness and homoscedasticity of the data were systematically tested and logarithmic transformations were applied when necessary. Significance of fixed effects was tested by conducting ANOVAs on the models' coefficients (with Satterthwaite approximation of the degrees of freedom) with the function *anova* and post-hoc comparisons were conducted using the *coefTest* function (F-test on the corresponding coefficients). 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 LMM with the categorical fixed effects $\text{GROUP}_{\text{TYPE}}$ ($\text{Group}_{\text{Short}}$ and $\text{Group}_{\text{Long}}$, n = 30 each) and $\text{TRAINING}_{\text{BLOCK}}$ (T1 to T6). In order to confirm these results using single-trial data, we used the Error allowing us to obtain a continuous variable at each trial. Notably, for each participant, Errors measured during training were expressed in percentage of the average Pre-training level. In this case, we ran a LMM with the categorical fixed effect $\text{GROUP}_{\text{TYPE}}$ ($\text{Group}_{\text{Short}}$ and $\text{Group}_{\text{Long}}$) and the continuous fixed effect $\text{TRAINING}_{\text{TRIAL}}$ (trial 1 to 240). When the analysis revealed a significant interaction, we then compared the coefficient associated to $\text{TRAINING}_{\text{TRIAL}}$ to evaluate potential between-group differences in learning rates. Then, to characterize the effect of the ITI's duration on motor learning, we replicated these analyses with the inclusion of the $\text{Group}_{\text{Short-PastStudy}}$.

As a second step, we aimed at evaluating the effect of reward timing on the dynamics of the learning process. To do so, we ran the same LMM as described above with the addition of the fixed effect $\text{TRAINING}_{\text{PHASE}}$ which was modeled as a categorical fixed effect with two modalities ($\text{Training}_{\text{Early}}$ or $\text{Training}_{\text{Late}}$ for the first and last 120 trials or training, respectively). We were especially interested in a potential triple $\text{TRAINING}_{\text{TRIAL}}$ x $\text{GROUP}_{\text{TYPE}}$ x $\text{TRAINING}_{\text{PHASE}}$ interaction which would indicate that learning rates varied not only depending on the group but also depending on the phase of practice.

As a supplementary analysis to support our differences of learning dynamics between the groups, we also ran regression analysis for each subject on binned Success rates (presented in Figure S3). Specifically, 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 $\text{Group}_{\text{Long}}$ (n = 30 each). 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 $Group_{Short}$ and $Group_{Long}$ at Post-training, using an independent sample t-test. 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 (threshold for significance 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. We first performed this analysis on the whole cohort (n = 30 per group). However, a potential caveat of theses analyses is that they included participants who did not learn the task on Day 1 and even exhibited a deterioration of performance with practice on Day 1. In these participants, a Re-test performance (*i.e.*, on Day 2) similar to the Pre-training level would be considered as evidence for an offline stabilization or even gain in performance, when it would actually only reflect a return to the baseline level of performance. In a second step, we therefore focused only on participants who demonstrated skill learning on Day 1 ($Success_{Post\text{-}training} - Success_{Pre\text{-}training} > 0$). This allowed us to compare offline consolidation in participants who responded to the training on Day 1 and who also happened to have very close Post-training success rates (Figure 4A), a crucial aspect in order to interpret any overnight change in performance. 40 participants were considered in this analysis (22 and 18 in $Group_{Short}$ and $Group_{Long}$, respectively). Pre-training normalized Success rates (averaged per block) and Error (single-trial) data were analyzed by means of LMMs with $GROUP_{TYPE}$ ($Group_{Short}$ and $Group_{Long}$) and $TEST_{BLOCK}$ (Post-training and Day 2) as categorical fixed effects.

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.