Title

When Does Motor Skill Learning Occur?

Author

Gupta, Mohan Wunand

Publication Date

2024

UNIVERSITY OF CALIFORNIA SAN DIEGO

When Does Motor Skill Learning Occur?

A Dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Experimental Psychology

by

Mohan W. Gupta

Committee in charge:

Professor Timothy C. Rickard, Chair
Professor Taylor Berg-Kirkpatrick
Professor Judith Fan
Professor Hal Pashler

2024

The Dissertation of Mohan W. Gupta is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Completing this dissertation and the continuation of my academic journey would not have been possible without the support of numerous people. Firstly, without the persistent mentorship and support from my advisor Dr. Timothy Rickard, I would not be the scientist that I have grown into today. His insights, patience, and dedication to my intellectual growth have been invaluable, and I am profoundly grateful for his mentorship. The members of my dissertation committee, Drs. Berg-Kirkpatrick, Fan, and Pashler have provided me with insightful comments and invaluable intellectual conversations. In particular, I extend my heartfelt gratitude to Dr. Judith Fan for not only her academic guidance but also for connecting me with my future post-doctoral advisor, an opportunity for which I am immensely grateful.

I am profoundly grateful to my fellow PhD compatriots, particularly Kyros and Leo, my roommates for several years, whose camaraderie has been a constant source of strength. Their support, along with that of Thomas who allowed me to explore my musical craft over the years has been invaluable throughout this journey.

I am grateful beyond words to my friends in Los Angeles, Buck, Jordan, Abbas, Abiha, Arlen, Jackson, Fabio, Frank, Kim, Nick, Gabbi, Cedric, Giga Chad Paul, Sam, Shannon, and countless others. Many of my reprieves during my PhD have come from the many hikes, camping trips, dinosaur digs, or celebrations throughout the nights, you have truly made my time in southern California very special. Cozy house for now, and forever.

Last but certainly not least, I extend my deepest gratitude to my family. My mom, Joell, whose investment in my education through homeschooling and the commutes to a better school system have undoubtedly given me the educational base so that I can thrive today. To my siblings, Rama and Forest, your unwavering support and guidance through my life has helped my life a brighter place. Lastly, to my late papa who passed away during my first year. I wish you could see me finish this journey and fulfill that immigrant dream. Your kindness, patience, and love will always persist through me.

Chapter 1, in full, is a reprint of the material as it appears in npj Science of Learning. Gupta, Mohan W.; Rickard, Timothy C., Spring Nature 2022. The dissertation author was the primary investigator and author of this paper.

Chapter 2, in full, is a reprint of the material as it appears in Scientific Reports, Gupta, Mohan W.; Rickard, Timothy C., Springer Nature, 2024. The dissertation author was the primary researcher and author of this paper.

Chapter 3, in part is currently in submission for publication of the material. Gupta, Mohan W.; Rickard, Timothy C. The dissertation author was the primary researcher and author of this material.

VITA

2017	Bachelor of Science in Psychology, Michigan State University
2017	Bachelor of Science in Neuroscience, Michigan State University
2021	Master of Arts in Experimental Psychology, University of California, San Diego
2024	Doctor of Philosophy in Experimental Psychology, University of California San Diego

ABSTRACT OF THE DISSERTATION

When Does Motor Skill Learning Occur?

by

Mohan W. Gupta

Doctor of Philosophy in Experimental Psychology
University of California San Diego, 2023
Professor Timothy C. Rickard, Chair

Motor skill learning, or learning a sequence of movements, is a fundamental process that occurs throughout one's life, from learning how to pick up your coffee cup to learning to play Für Elise. The primary perspective is that motor skill learning occurs during rest periods, offline, rather than concurrently with practice, online. However, those interpretations have conflated learning with performance. In this dissertation, I will propose an account that assumes learning occurs online, concurrently with practice, and that reactive inhibition – the transient worsening of performance during continuous trials – presents the illusion that learning occurs offline. Chapter 1 provides evidence that the online account plus reactive inhibition is sufficient to explain both learning and the performance improvement following a rest period. However, this experiment was unable to rule out the offline learning account. Chapter 2 advances a computational modeling framework that compares both the online and offline accounts. Additionally, we consider a third hybrid possibility that learning occurs both online and offline. Overall, a hybrid model was able to best fit the data suggesting that both online and offline learning occur. Further, I demonstrated the necessity of assuming reactive inhibition in computational models, regardless of the type of learning model. Finally, these models suggested that there may be learning differences resulting from training schedules. Chapter 3 examined if learning rates were dependent on training schedules. Contrary to prior work, I found no evidence that training schedule influenced learning rates in a motor skill task where few errors are made. Taken together, my dissertation work provides a novel computational framework for testing the temporality of motor skill learning, compelling evidence that reactive inhibition is a necessary component to consider and challenges the predominant view that motor skill learning only occurs offline.

INTRODUCTION

Motor skill learning, the process by which individuals acquire and refine the ability to perform sequences of movements, constitutes a cornerstone of human experience, manifesting from the simplest daily actions to the mastery of intricate musical compositions. The prevailing perspective posits that motor skill learning unfolds offline, during periods of rest. However, recent inquiries have cast doubt upon that perspective, prompting a critical reevaluation of the temporal dynamics underlying motor skill acquisition.

In this dissertation, I endeavor to challenge the predominant view that motor skill learning occurs exclusively offline by advancing an alternate framework that assumes that learning occurs online, concurrently with practice. Central to this framework is the concept of reactive inhibition, a phenomenon wherein performance temporarily deteriorates over the course of prolonged task engagement. Through a comprehensive exploration of empirical evidence and computational modeling, this dissertation aims to delineate when motor skill learning occurs, while elucidating the underlying mechanisms driving motor skill performance.

Chapter 1 presents empirical evidence supporting the viability of the proposed online learning framework, demonstrating its capacity to account for the observed performance enhancements following rest periods. However, this work failed to rule out an offline learning account, prompting the subsequent development of computational models in Chapter 2. I advanced a computational framework that compared three types of learning models: online, offline, and hybrid. This chapter examines the viability of each learning account while elucidating the role of reactive inhibition in shaping skill acquisition trajectories.

Building upon the insights gleaned from computational modeling, Chapter 3 examines the influence of training schedules on learning rates. Contrary to prevailing assumptions, learning rates were agnostic to training schedules, raising fundamental questions a pure offline account must contend with in the future.

In summation, this dissertation endeavors to furnish compelling evidence in support that online learning indeed occurs, challenging the entrenched notion of exclusive offline learning mechanisms in motor skill acquisition. By unraveling the complex interplay between online learning, offline consolidation, and reactive inhibition, this study not only advances our understanding of motor skill acquisition but also underscores the necessity for a paradigm shift in conceptualizing the temporal dynamics of skill learning processes.

Chapter 1 Dissipation of Reactive Inhibition is Sufficient to Explain Post-rest Improvements in Motor Sequence Learning.

Introduction

A fundamental question about motor learning is whether it occurs online (concurrently with performance) or offline (during break periods). Pertinent to that question is the repeated observation that after a rest period – whether involving sleep, five minutes, or even 10 seconds – there are reaction time (RT) improvements in motor sequence tasks[1-8]. Numerous authors have concluded that those improvements are due to a form of offline memory consolidation that enhances learning and results in superior behavioral performance, rather than merely stabilizing memory as in the case of declarative learning[1-3,5,7-8]. However, the consolidation account is unable to explain several phenomena across different rest time scales. First, in sleep studies that adequately controlled for various factors like circadian rhythms and reactive inhibition (RI; the slowing of RT as one continuously performs a motor task[9]), no improvements in RT are observed after rest[4,6,10-13]. Second, there are improvements in RT over some rest intervals but not others[2-3]. For example, a rest period of five minutes, like the one used in the current study, shows an RT improvement. However, if that rest is increased to four hours, there is no improvement[2]. Third, RT improvements after a rest period greater than five minutes only occur in "massed" training conditions when the on-task trial time is 30 s, and not in "spaced" conditions when it is 10 s, despite the total amount of on-task time being equated[4,10]. The consolidation account fails to explain these phenomena.

On the other hand, the accrual and dissipation of RI may be sufficient to explain those phenomena, without the need to infer offline facilitating consolidation. All else held constant, the longer a motor task continues, the greater the RI build-up, resulting in progressively slower RTs. During breaks between trials, RI dissipates. The longer the break, the greater the expected dissipation[4,9-10,12]. Thus, long on-task trial times and short breaks between trials should yield the largest build-up of RI over trials, and consequently, the largest post-rest RT improvement due to RI dissipation. Conversely, short on-task trial time and a long break period will yield the smallest build-up of RI over trials and the smallest post-rest RT improvement.

Those RI effects have failed to be considered in several recent studies, nor in many past studies in which offline facilitating consolidation has been inferred. This raises the possibility that there is no facilitating consolidation. Rather, it may be that learning occurs concurrently with performance and that dissipation of reactive inhibition creates illusory learning during breaks. The main goal of this work was to test the sufficiency of that alternative account. We investigated whether the amount of post-rest RT improvement that is observed over variations in trial time (10 s vs. 30 s) and break time between trials (10 s vs. 30 s) can be explained by an RI account, without invoking a facilitating consolidation process.

The facilitating consolidation account has two possible interpretations. The first and simpler interpretation is that consolidative processes only occur during the post-training rest period[2]. Thus, as long as the total amount of both training time (the amount of online learning) and the rest periods (amount of consolidation) are equated over groups, then both the amount of consolidation and the associated RT improvement after rest will be the same over groups. The second interpretation has arisen from recent evidence that consolidation may also occur during short breaks between trials and that all learning occurs during those breaks[7-8]. In this interpretation, groups with more frequent and longer breaks will have undergone more consolidation by the end of training. Further, if we assume that there is a finite amount of facilitating consolidation that can occur over the time course of the experiment, consolidation during breaks may reduce the amount of additional consolidation that occurs during the post-training rest period. This assumption has not been made previously in the literature. This version of the consolidation account and the RI account make similar predictions for the amount of post-rest improvements: the smallest post-rest improvement should occur in the 10 s-on, 30 s-break group (for which there are 1,050 s of cumulative break time during practice; 35 breaks at 30 s per break) and largest post-rest improvement in the 30 s-on, 10 s-break group (for which there are 110 s of cumulative break time; 11 breaks at 10 s per break).

Results

Reactive Inhibition

To confirm the presence of RI in the 30 s on-task trial groups, we divided each 30 s trial into three consecutive 10 s bins[4,10]. We then compared mean RTs between the first and third bins. A paired-samples t-test, averaged over all practice trials, yielded evidence of RI in both the 30 s break group, $t(36) = -3.42$, $p = 0.0016$, $d = -0.56$, and the 10 s break group, $t(40) = -3.26$, $p = 0.0023$, $d = -0.51$. For the 10 s on-task trial groups, we split the 10 s trials into three 3.33 s bins. A paired-samples t-test, averaged over all practice trials, yielded evidence of RI in both the 30 s break group, $t(38) = -8.2$, $p < .0001$, $d = -1.32$ and the 10 s break group, $t(41) = -6.3$, $p < .0001$, $d=-0.98$.

Post-Rest Improvement

With the presence of RI confirmed, we investigated how the amount of on-task time and break time affected the post-rest improvement. As in Brawn et al. (2011), we compared the RT means of the last two training trials (11 and 12) with post-rest trials (13 and 14). A 2x2 mixed-factors Analysis of Variance (ANOVA) revealed a significant effect of break time on the post-rest improvement, $F(1, 155) = 14.49$, $p < 0.001$, $\eta^2 = .08$ (Figure 1.1b), as well as a significant effect of on-task trial time, $F(1, 155) = 5.79$, $p < 0.01$, $\eta^2 = .04$ (figure 1.1b). There was no significant interaction between the two factors, $F(1, 155) = 1.217$, $p = 0.272$, $\eta^2 = .007$. The same results were found when analyzing the number of correctly completed sequences between the last two training trials and the post-rest trials. The ANOVA revealed a significant effect of break time, F(1, 155) =

9.23, $p = 0.0027$, $\eta^2 = .04$, as well as a significant effect of on-task trial time, $F(1, 155) = 85.38$, $p < 0.001$, $\eta^2 = .34$, with no significant interaction, $F(1, 155) = 1.722$, $p = 0.19$, $\eta^2 = .007$.

To investigate if the last two training trials (11 and 12) and post-rest trials (13 and 14) had significantly different RTs within each group, we ran four paired-samples t-tests. The 30 s on-task trial, 30 s break group showed a significant decrease in RT, $t(36) = -3.97$, $p = .0003$, $d = -0.65$, as did the 30 s on-task trial, 10 s break group, $t(40) = -9.01$, $p < .0001$, $d = -1.41$. In contrast, the 10 s on-task trial, 30 s break group showed no evidence of RT decrease, $t(38) = -0.7$, p-value = .49, $d = -0.11$ whereas the 10 s on-task trial, 10 s break group did, $t(41) = -7.9$, $p < .0001$, $d = -1.22$. The null post-rest result for the 10 s on-task trial, 30 s break group suggests that a 30 s break was sufficient to resolve most if not all of the RI build-up that occurred on each trial.

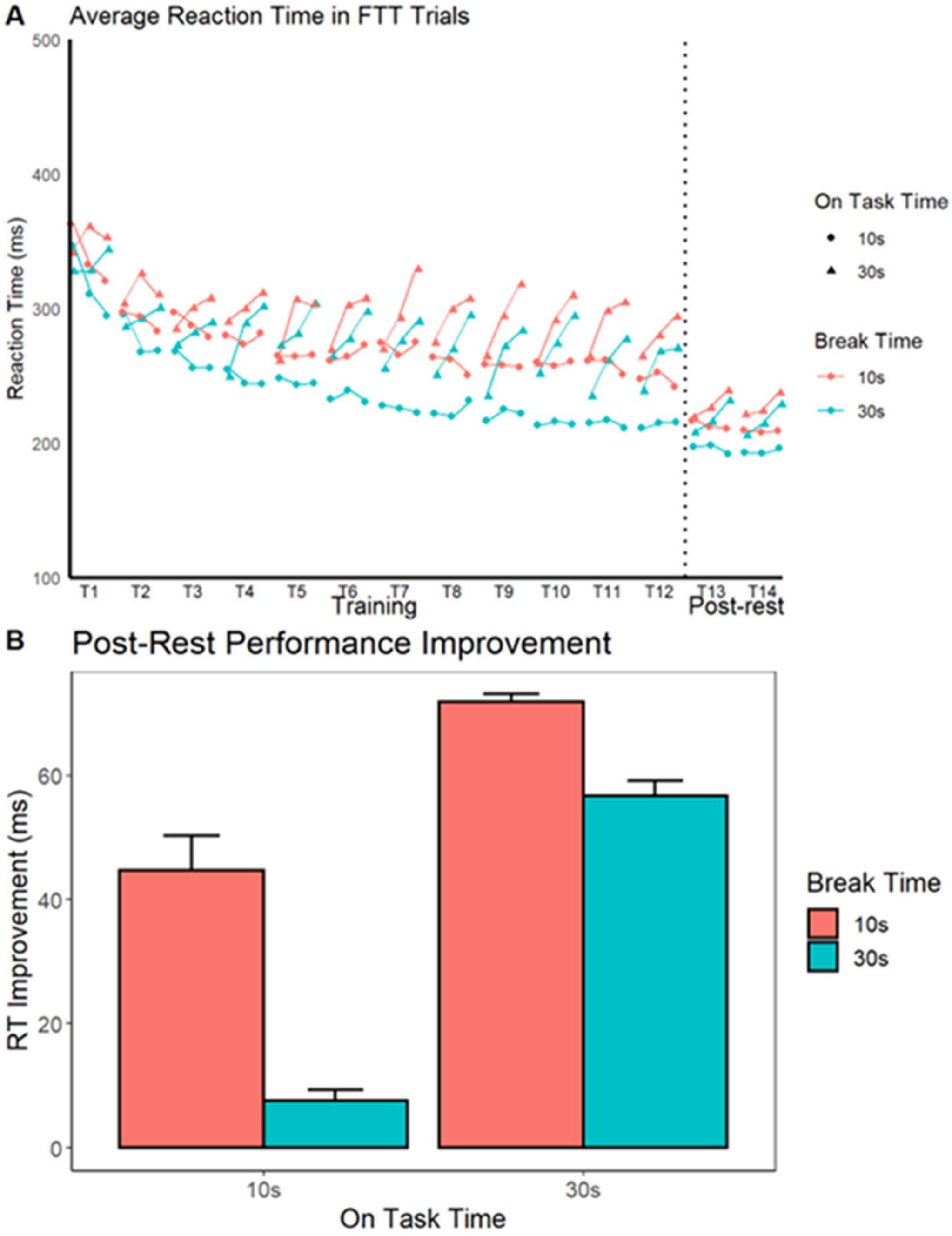


Figure 1.1 A) Each point represents 10 s of on-task trial time. In the 30 s on-task trial time condition, the triangles connected by lines are not separated by breaks. In the 10 s on-task trial time condition, each circle is separated by a break, even if the line connects them. B The y-axis shows the amount of RT improvement after the rest. The x-axis indicates the amount of on-task trial time, whereas the color indicates the amount of break time. Holding the amount of on-task trial time constant (x-axis), break time has a strong effect on the amount of RT improvement. When the break time (color) is held constant, on-task trial time also has a strong effect on RT improvement. Error bars in standard error.

Discussion

We investigated how the post-rest RT improvement is moderated by on-task trial time and break time. We found that both factors significantly affect the post-rest improvement. The longer the trial and the shorter the breaks, the greater the post-rest improvement, and vice-versa. Those results, along with the clear build-up of RI within trials, are fully consistent with our RI account. This account assumes that learning is concurrent with performance, RI builds during continuous performance, and that it dissipates gradually during breaks.

The simpler interpretation of the consolidation hypothesis predicted that only the amount of total training and the length of the rest interval will affect the post-rest improvement. Because both of those factors were held constant across the four groups, that interpretation predicted null effects, which were not observed. An alternative interpretation of the consolidation hypothesis predicts that consolidation occurs during the much shorter breaks between trials[8], and that the greater the consolidation during the breaks will result in less consolidation during the rest period (although this latter prediction has not been previously hypothesized in the literature). That prediction is also consistent with the observed post-rest improvement over groups.

Our findings provide the first systematic evidence that an RI-based model assuming online learning and no offline facilitating consolidation can explain motor sequence learning and performance in the context of short breaks and rest periods. Although the revised consolidation account with an additional assumption can also explain those post-rest improvement effects, the RI account has two advantages. First, RI and its dissipation during breaks is clearly a necessary factor in understanding motor performance, whereas facilitating consolidation does not appear to be necessary. Second, our finding that the post-rest RT improvement was negligible and non-significant in the 10 s on-task trial, 30 s break group is not surprising in light of the RI model, given that RI has long been understood as a short-lasting phenomenon[4,9-10,12]. In contrast, there is no precedent in the literature suggesting that facilitating consolidation can be exhausted by a series of 30 s breaks between trials, such that no additional facilitating effect occurred during a subsequent 300 s rest period. Finally, in our 10 s on-task trial and 10 s break group that is analogous to groups used in two recent studies[7-8], we found evidence of RI, both within trial and across the rest period. This raises the possibility that the offline facilitating consolidation that the authors of those studies inferred in fact reflects solely the dissipation of RI.

This study was not designed to estimate the accrual and dissipation rates of the RI, but we can gain some insight based on the non-significant post-rest RT improvement for the 10 s on-task, 30 s break group, whereas there was a statistically significant post-rest improvement for the other three groups. Within the RI theoretical framework, the null effect in the 10 s on-task trial, 30 s break group indicates that 30 s is sufficient to fully resolve the RI that builds over 10 s trial. Conversely, we know that 10 s of break between 10 s on-task trials is insufficient. Hence, for the case of 10 s on-task trial time at least, RI resolves at a rate that is somewhere between one and three times smaller than the rate at which it accrues. This suggests that the rate of dissipation is longer than the rate of accrual. More research is needed to understand what the exact relative rate is and whether it is a constant over different on-task time periods.

In conclusion, the RI account is sufficient to explain the post-rest improvements after a 300 s rest. This finding reinforces the claim that the accrual and dissipation of RI is a critical factor for understanding motor learning and performance over short time scales (for related conclusions in the case of implicit sequence learning, see Török et al., 2017), whereas facilitating consolidation may not be.

Methods

Participants

All participants were right-handed. Thirty-seven participants were in the 30 s on, 30 s break group (age = 20.16, F = 67.6%). Thirty-nine participants were in the 10 s on, 30 s break group (age = 20.31, F = 79.5%). Forty-two participants were in the 30 s on, 10 s break group (age = 21.07, F = 79.5%). Forty-four participants were in the 10 s on, 10s break group (age = 20.54, F = 75%). One participant was removed from the 30 s on, 10 break group and two participants were removed from the 10 s on, and 30 s break group due to corrupted data. Participants provided informed consent via button press. All procedures were approved by the institutional review board of the University of California, San Diego.

Experimental design and procedure

Participants performed a standard finger-tapping-task where they repeated the sequence, 4-1-3-2-4, as quickly and accurately as possible with their non-dominant left hand[14]. A 2x2, between-participant design was used, with factors of Trial Length (10 or 30 seconds) and Break Period between trials (10 s or 30 s). After the 360 seconds of on-task training, there was a 300 second rest where participants performed a distraction task of double-digit addition. Afterwards, they performed 60 s of test trials with breaks in between in the same conditions that they trained on (Figure 1.2).

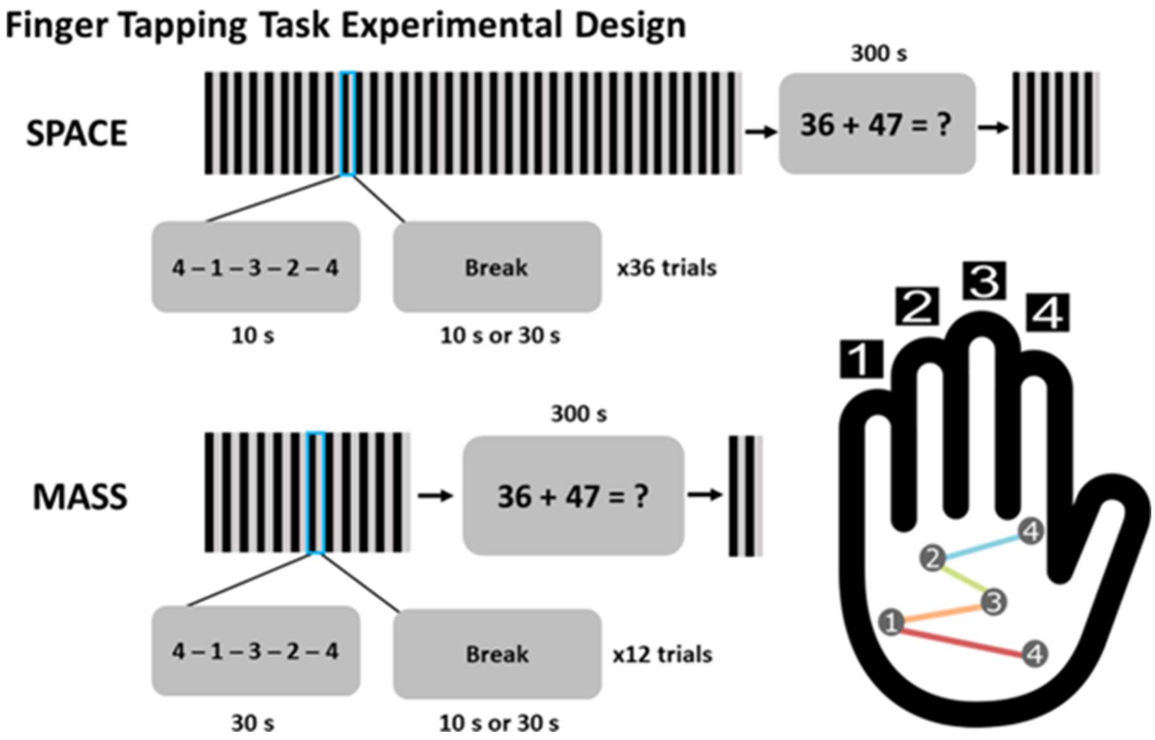


Figure 1.2 Participants learned a motor sequence over one session. They were instructed to repeatedly type a sequence, 41324, with their non-dominant left hand as fast and as accurately as possible. Keypress 4 was performed with the index finger, keypress 3 with the middle finger, keypress 2 with the ring finger, and keypress 1 with the pinky finger. Participants trained for a total of 360 s in either 10 s or 30 s trials. In between practice trials were either breaks of 10 s or 30 s. After training, participants performed 300 s of double digit addition. They were then tested on the practiced sequence for 60 s with the same trial and break lengths during training.

Statistical analysis

The first completed sequence of each trial was considered a warm-up trial and was removed prior to data analysis. RT was defined as the time between temporally adjacent keypresses, where the first keypress RT for a trial was the time since the last keypress of the preceding trial. Keypresses were logged as ‘KEYUP’ events in JavaScript. This event registers the keypress once the key has been released. The post-rest RT improvement was defined as the difference between mean RT of the last two training trials (11 and 12) and the mean RT of the post-rest trials (13 and 14).

Data availability

All data and code (stimuli and analyses are available online (https://osf.io/khaqv/). Further information and requests for resources should be directed to and will be fulfilled by the corresponding author, TCR (trickard@ucsd.edu).

Competing Interest Statement: We have no competing interests to report.

Author Contributions: MWG conducted the experiments and analyzed the data. Both MWG and TCR designed the experiment and wrote the article.

References

1. Walker, M. P., Brakefield, T., Morgan, A., Hobson, J. A., & Stickgold, R. Practice with sleep makes perfect: Sleep-dependent motor skill learning. Neuron, 35(1), 205–211 (2002). https://doi.org/10.1016/S0896-6273(02)00746-8

2. Hotermans, C., Peigneux, P., De Noordhout, A. M., Moonen, G., & Maquet, P. Early boost and slow consolidation in motor skill learning. Learning and Memory, 13(5), 580–583 (2006). https://doi.org/10.1101/lm.239406

3. Albouy, G., Ruby, P., Phillips, C., Luxen, A., Peigneux, P., & Maquet, P. Implicit oculomotor sequence learning in humans: Time course of offline processing. Brain Research, 1090(1), 163–171 (2006). https://doi.org/10.1016/j.brainres.2006.03.076

4. Rickard, T. C., Cai, D. J., Rieth, C. A., Jones, J. & Ard, M. C. Sleep does not enhance motor sequence learning. J. Exp. Psychol. Learn Mem. Cogn. 34, 834–842 (2008).

5. Debarnot, U., Clerget, E., & Olivier, E. Role of the primary motor cortex in the early boost in performance following mental imagery training. PLoS ONE, 6(10) (2011). https://doi.org/10.1371/journal.pone.0026717

6. Pan, S. C., & Rickard, T. C. Sleep and motor learning: Is there room for consolidation? Psychological Bulletin, 141(4), 812–834 (2015). https://doi.org/10.1037/bul0000009

7. Bönstrup, M., Iturrate, I., Hebart, M. N., Censor, N. & Cohen, L. G. Mechanisms of offline motor learning at a microscale of seconds in large-scale crowdsourced data. Npj Sci. Learn. 5, 1–10 (2020).

8. Buch, E. R., Claudino, L., Quentin, R., Bönstrup, M., & Cohen, L. G. Consolidation of human skill linked to waking hippocampo-neocortical replay. Cell Reports, 35(10) (2021). https://doi.org/10.1016/j.celrep.2021.109193

9. Hull, C. L. (1943). Principles of behavior. New York: Appleton-Century- Crofts.

10. Brawn, T. P., Fenn, K. M., Nusbaum, H. C. & Margoliash, D. Consolidating the effects of waking and sleep on motor-sequence learning. J. Neurosci. 30, 13977–13982 (2010).

11. Nettersheim, A., Hallschmid, M., Born, J., & Diekelmann, S. The Role of Sleep in Motor Sequence Consolidation: Stabilization Rather Than Enhancement. Journal of Neuroscience, 35(17), 6696-6702 (2015). https://doi.org/10.1523/JNEUROSCI.1236-14.2015

12. Török, B., Janacsek, K., Nagy, D. G., Orbán, G. & Nemeth, D. Measuring and filtering reactive inhibition is essential for assessing serial decision making and learning. J. Exp. Psychol. Gen. 146, 529–542 (2017).

13. Rickard, T. C., Pan, S. C., & Gupta, M. W. Severe publication bias contributes to illusory sleep consolidation in the motor sequence learning literature. J. Exp. Psychol. Learn Mem. Cogn, (2022). https://doi.org/10.1037/xlm0001090

14. Nissen, M. J. & Bullemer, P. Attentional requirements of learning: evidence from performance measures. Cogn. Psychol. 19, 1–32 (1987).

Acknowledgements

Chapter 1, in full, is a reprint of the material as it appears in npj Science of Learning. Gupta, Mohan W.; Rickard, Timothy C., Spring Nature 2022. The dissertation author was the primary investigator and author of this paper.

Chapter 2 Comparison of Online, Offline, and Hybrid Hypotheses of Motor Sequence Learning Using a Quantitative Model that Incorporate Reactive Inhibition

Introduction

Multiple researchers have advanced the hypothesis that motor learning occurs *offline,* both during sleep[1-3] and in more recent work during brief waking breaks[4-7]. Conversely, motor sequence learning has been posited to take place concurrently with task performance (i.e., *online* learning)[9,11]. Discrimination among those and related hypotheses should have fundamental implications for the properties of the underlying neural system.

The hypothesized motor learning during waking breaks is believed to involve facilitating *micro-consolidation*, as opposed to the stabilizing consolidation (i.e., protection from forgetting and memory integration) that is thought to occur for declarative learning. The conclusion favoring that hypothesis is based on the findings that (1) there is often negative or no response time improvement over motor sequence repetitions within a performance trial, (2) performance at the beginning of a trial is often better than that at the end of the preceding trial, and (3) neural evidence of hippocampal replay appears to occur during rest periods.

The micro-consolidation hypothesis was challenged, however, by Gupta and Rickard (2022). They advanced evidence for a diametrically opposed learning model that assumes (1) that all learning occurs online, concurrently with task performance, and (2) that reactive inhibition (RI) accrues over sequences within a trial and dissipates over time during breaks. Although the mechanism underlying RI is not well established, the empirical effect it describes has been replicated over studies spanning more than 80 years[10-13]. Nevertheless, the phenomenon has not played a central role in recent studies of facilitating micro-consolidation. Gupta and Rickard (2022) explored the sufficiency of their hypothesis using a standard finger tapping task where participants repeatedly typed a five-key sequence with their non-dominant hand[14]. There were four groups, crossing 10 or 30 s per performance trial with a 10 or 30 s break between trials, while equating total time on task. After the training phase, there was a 5-minute rest period, followed by additional trials. They found strong evidence during training for both accrual and dissipation of RI in all groups. Further, they observed that the largest RT gain (i.e., decrease in RT) in correct sequence completion time occurred in the 30 s on-task, 10 s break group and the smallest in the 10 s on-task, 30 s break group. These results are in-line with the hypothesis that RI builds-up substantially across training trials in the former group (because the 10 s breaks were insufficient to fully resolve the build-up of RI during each 30 s trial), but largely resolved during breaks in the latter group. Those results can account for the behavioral evidence underlying the micro-consolidation hypothesis without the need to posit offline learning.

Although Gupta and Rickard concluded that the online learning plus RI hypothesis may be sufficient to account for the ordinal pattern of post-rest RT gains in their data, they did not offer a quantitative model of the complex patterns that were observed across sequences, trials, and the rest period. Indeed, no such model has been advanced to date for an explicit motor sequence learning task (for a model of RI effects in the implicit motor sequence task, see Torok et al., 2017). If achievable with a modest number of psychologically plausible free parameters, a quantitative learning models should advance research in this area by (1) assessing the sufficiency of different models at a finer temporal grain size than in past work, (2) providing new insights into the properties of RI, and (3) setting a new reference for future theory development. Here we advance three types of learning models: online, offline, and hybrids of the two.

The motor sequence task used here is closely related to that used by Gupta and Rickard and by many prior authors. As one exception, a trial in the current study was defined as a fixed number of correctly completed sequences, rather than a fixed amount of training time. This change was made because in the time constrained version of the task, participants complete a variable number of correct sequences during each trial, complicating model fitting to averaged data (Figure

2.1b). In the two *spaced practice* (S) groups each trial ended after completion of 5 correct sequences and in the two *massed practice* (M) groups, each trial ended after the completion of 15 correct sequences. We indicate break time between trials, in seconds, as the number after either M or S; hence, the group labels are S30, S10, M30, and M10. The total number of correct sequences (henceforth, *sequences*) was held constant across groups by implementing three trials in the spaced groups for every one trial in the massed groups, for a total of 180 training sequences and 30 test sequences in each group. Training and test phases were separated by a five-minute distractor task of double digit addition.

Quantitative Models

We first develop the general quantitative framework in the context of the Online model and then develop the Offline and Hybrid learning models. All models that include RI have an identical quantitative implementation. All models feature a variant of a *skill* function for mapping learning onto correct sequence completion time (henceforth, response time, or *RT*). Prior work on skill learning indicates that, for tasks that do not exhibit major strategy shifts with practice[15], RT gain over trials is a smooth, negatively accelerating function of repetition. We adopted that assumption here. Because the exact mathematical function that governs achieved skill is unknown for motor sequence learning, we adopted a flexible function that combines power and exponential terms (see Supp 3; also advanced as a general practice function for cognitive skills[16]). Depending on the model, achieved skill occurs as a function of sequence practice (online learning), micro-consolidation during breaks (offline learning), or both (hybrid learning).

In all models we assume that no forgetting occurs during either the short breaks between trials or the 5-minute rest period prior to the test. We further assume that the five minute post-training rest is sufficient to resolve all residual RI built-up across training trials. In the hypothetical case of no RI, the observed RT would directly reflect the achieved skill. However, in the presence of RI, achieved skill is a latent variable on all trials except on the first training trial and the first test trial, as elaborated below.

Online Skill Model

The Online learning model assumes that learning starts immediately on the first sequence of a trial and runs to completion by the last sequence, such that there is no learning during breaks. In this model, RT for the underlying skill is solely a function of cumulative sequences over training and test phases (S),

$$RT_{\text{skill-Online}} = a + b*e^{(-c\ *S)}*S^{-k}. \qquad \textbf{\textit{(Equation 1)}}$$

As motivated later, we removed the first sequence of the first trial (and of later trials) as a warm-up sequence and then started the cumulative sequence variable S at a value of one on the second sequence of the first trial; i.e., modeled sequence number one was actually experimental sequence two. Because there was a sequence learning event prior to modeled sequence one, Equation 1 is parameterized appropriately. Parameter *a* is the asymptotically achievable skill, *b* is the improvement that can occur with practice, *c* is the exponential rate parameter, and *k* is the power rate parameter (for a summary of the estimated parameters and known variables across all models, see Table 1). Together, the two rate parameters determine the shape of the skill curve and how quickly it approaches asymptote. In the absence of RI, Equation 1 would describe the *observed* RTs. That skill curve is illustrated by the blue solid line in Figure 2.1a.

Table 2.1 Parameter and variable descriptions in the model equation. All parameters are constrained to have values > 0; j < Stot/Btot.

		Description
Parameters	a	asymptotic skill
	b	magnitude of skill improvement that can occur with practice
	c	Power rate
	k	Exponential rate
	y	rate of the RT slowing effect over sequences within each training and test trial due to the build-up of within-trial RI
	z	rate of the RT slowing effect over training and test trials due to build-up of residual RI
	g	additional increment in offline learning during the post-training rest period in version 2 of the offline model
	j	index of offline learning in the HybridJ model
	S	modeled sequence number, cumulative across training and test phases
	ST	modeled sequence number within trial (1-14 for massed groups; 1-4 for spaced groups)
	T	training phase trial number (1 to N)
Variables	T_{test}	test phase trial number, starting a 1 on the first test trial
	X	dummy variable that takes a value of 0 for training trials and 1 for test trials
	B_{tot}	total number of breaks including the rest period
	S_{tot}	total number of modeled sequences
	S’	describes the allocation of learning units between online and offline components in the HybridJ model

Modeling RI. RI was implemented in the same way for all models. Two effects on the observed RT were implemented equivalently across all models that incorporate RI. The first is the increase in RT over sequences within each trial due to the build-up of RI (i.e., the *within-trial* RI effect). We assumed the simplest case of linear RT increase over sequences within-trial, with the same rate parameter (y) for all trials across both training and test phases. Although this linear effect may not hold when there are a very large number of sequences per trial, prior work[9,11] suggests that it is a reasonable approximation for the current experiment. We assume that RI operates with the same force and magnitude across all trials.

The second effect is the build-up of residual RI from trial to trial due to its incomplete dissipation during trials (parameter z). Consider the M10 group, in which there are 15 experimental sequences within each trial and 10 s breaks between trials. Based on prior work, that group should exhibit both substantial within-trial RI build-up and incomplete resolution of RI during the short break, yielding a residual RI effect at the beginning of the next trial. That residual RI is expected to accrue over trials, and for simplicity we assume that the accrual rate is constant across trials. Thus, the difference between the observed RT and RT_{skill} on the first modeled sequence of each trial increases as a linear function of trial number during both training and test phases. That residual build-up should be largest for the M10 group and smallest for the S30 group, as implied by the results of Gupta and Rickard (2022). Given a linear relation between RI and its effect on RT, the predicted RT change within and across training trials due to RI is,

$$RT_{RI\text{-}training} = (ST-1)*y + (T-1)*z, \qquad \textbf{\textit{(Equation 2)}}$$

where (ST-1) reflects the fact that within-trial RI is defined to be zero on the first sequence of each trial, and (T-1) reflects the fact that residual RI is by definition zero on trial one. The effect of within-trial RI on RT is illustrated in Figure 2.1a, where the effect of the residual RI build-up is illustrated by the widening gap over trials between the dashed and solid curves.

Based on our earlier work[9], we assume that a 5-minute rest period is sufficient to completely resolve residual RI for all four groups. Hence, the observed RT for the first test phase sequence is expected – like that on the first training trial – to be a pure measure of RT_{skill} (see Figure 2.1a). The build-up of both within-trial and residual RI during the test, and the consequential RT effects, are assumed to occur at the same rate as during training. Thus, for the test phase, the effect of RI on RT is:

$RT_{RI\text{-}test} = (ST - 1)*y + (T_{test} - 1)*z$, **(Equation 3)**

where T_{test} is the test trial number, which starts at a value of 1.

Hence, the effect of RI across both training and test phases is given by the mixture equation:

$RT_{RI} = (1\text{-}X)*\{(ST\text{-}1)*y + (T\text{-}1)*z\} + X*\{(ST\text{-}1)*y + (T_{test}\text{-}1)*z\}$, **(Equation 4)**

where X takes a value of zero during training and one during the test.

Full Online Model Equation. The simultaneous least squares nonlinear model fitting across both training and test phases was accomplished separately for each model and group by combining the appropriate skill equation for the model with the common RT_{RI} equation. For the Online model:

$RT_{overall\text{-}Online} = RT_{skill\text{-}Online} + RT_{RI}$ **(Equation 5)**

Offline Model

This model assumes that learning occurs exclusively during breaks and that the amount of learning is equivalent across all of the equal duration breaks. The proposed micro-consolidation account[5-8] is consistent with both assumptions, given that hippocampal replay occurs at the same rate across all breaks. In this model, S in the online skill equation is replaced by T-1, given that learning occurs during breaks between trials,

$RT_{skill\text{-}Offline} = a + b*e^{\{-c*(T-1)\}} * (T\text{-}1)^{-k}$, ***(Equation 6a)***

where (T-1) places the first offline consolidation event appropriately after the first trial.

Two versions of the Offline model were under primary consideration. First, because the reference Offline model in the literature assumes negligible effects of RI on performance, we considered a version (V1) with no RI. The overall equation for observed RT in this case is just the offline skill equation (Equation 6a):

$RT_{\text{observed-Offline-V1}} = RT_{\text{skill-Offline}}$ ***(Equation 6b)***

In version 2 we added RI to version 1:

$RT_{\text{observed-Offline-V2}} = RT_{\text{skill-Offline}} + RT_{\text{RI}}$. (***Equation 6c***)

We also considered variants of those two offline models in which more micro-consolidation occurs during the 5 minute rest period than during the training phase breaks. Given that X (see Table 2) takes a value of zero during training and of one during the test, and representing the RT effect of extra consolidation during the rest period with a offset new parameter (*g*), the skill equation for that variant is,

$RT_{\text{skill-Offline-offset}} = a + b*e^{\{-c*(T-1)\}} * (T-1)^{-k} - g*X$ ***(Equation 6d)***

When that variant was fitted in the absence of assumed RI, in no group was the fit better by the Bayesian Information Criterion (BIC) than that for offline model version 2. When that variant was fitted with RI included, the residual sum of squares (RSS) was slightly reduced relative to offline model version 2, but version 2 again provided better BIC fits. We thus do not consider those two variants further in this paper.

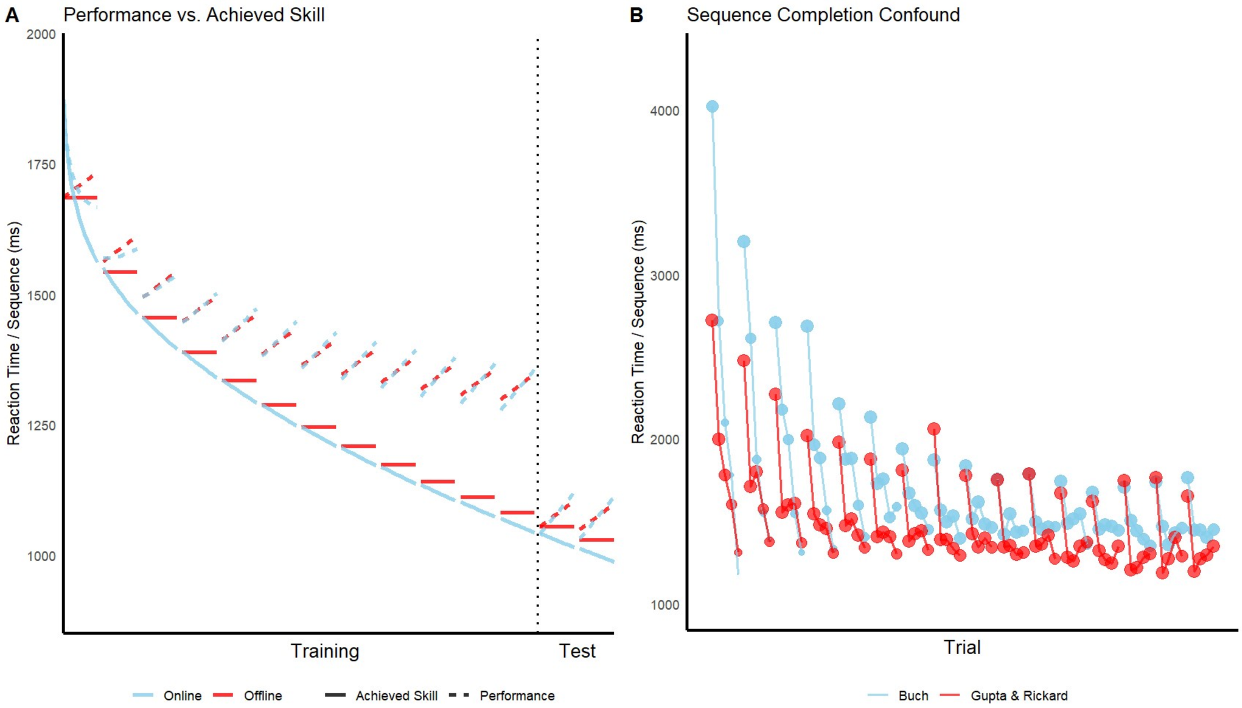


Figure 2.1 (a) Model predictions of the Online and Offline models assuming RI with respect to both underlying achieved skill (solid lines; equations 1 and 6b) and observed performance as moderated by accrual and dissipation of RI (dashed lines; equations 5 and 6c). b) Data from Buch et al. (2019) and Gupta and Rickard (2022) are plotted for the first 16 trials. Both datasets consist of 10 s trials and 10 s breaks. Those data are plotted as a function of how quickly a sequence is completed, artificially truncated at five sequences per trial. Many participants completed more than five sequences. Because the constraint on all of these trials is time rather than number of correct sequences, participants complete a variable number of sequences, as evidenced by the size of the circles: the larger, the more participants that have completed that sequence number. Data of this type would unnecessarily complicate model fitting to averaged data, particularly over the first 11 trials in spaced conditions, which much of the prior work on the micro-consolidation hypothesis has been focused.

Hybrid Models

A third class of models assumes that learning can occur both online and offline. This assumption is plausible because multiple systems are likely to underlie motor sequence learning[25-26,28]. Three variants are considered: *HybridJ, HybridE, and HybridP*.

HybridJ. The goal of the HybridJ model is to estimate the relative proportion of learning that is due to offline and online components. Conceptually, each executed sequence across both training and test phases is treated as yielding one unit of learning. If both offline and online learning occur, then some of those learning units occur concurrently with sequence execution and some occur during the break periods. As is implicitly the case for the Online and Offline models, the number of learning units per trial or break is held constant over practice and test phases, such that the non-linear RT improvement (skill) is solely a property of the mapping from learning to performance. It may be more accurate to assume that the learning rate itself decreases across practice sequences and trials. However, we cannot differentiate between those two possibilities in the current work. For convenience, we implement the non-linear effect fully within the RT skill equation and assume that the underlying learning increments are constant over all sequences and trials.

A single new parameter (j) estimates the number of sequence learning units that occur during each break (offline learning). Defining B_{tot} as that total number of breaks in each group (including the 5-minute rest period), then $B_{tot}*j$ is the total number of (offline) learning units that occur during all breaks. In the massed groups, for example, there are 11 breaks during training, a break between training and test phase, and a break between the first and second test trials, so that $B_{tot} = 13$. The number of remaining sequence-level learning units available for online learning is then $S_{tot} – B_{tot}*j$, where S_{tot} is the total number of modeled sequences across training and the test phases.

The cumulative effective number of learning units accrued across sequences and trials is then,

$S' = S*(S_{tot} – B_{tot}*j)/S_{tot} + (T-1)*j$, **(Equation 7a)**

where $(T-1)*j$ implements the cumulative number of sequence learning units that occur across breaks (offline) and $S*(S_{tot} – B_{tot}*j)/S_{tot}$ implements the remaining fractional learning units that accrue across sequences. The total number of sequence units is conserved, so that on the last test sequence, the value of S’ converges on the value of S. In this HybridJ model, S’ replaces S in the skill equation, yielding,

$RT_{\text{skill-HybridJ}} = a + b{*}e^{(-c\ {*}S')}{*}S'^{-k}$, **(Equation 7b)**

and the observed RT equation is,

$RT_{\text{overall-HybridJ}} = RT_{\text{skill-HybridJ}} + RT_{\text{RI}}$. **(Equation 7c)**

If all learning is either online or offline, then we expect the HybridJ model to yield a higher (less favorable) BIC value than either for the Online or Offline model, due to the extra free parameter, *j,* which in those cases would not yield improved fits. Conversely, if both types of learning play an important role, then we expect a best fit of this hybrid model with the estimated value of *j* somewhere between zero (all online learning) and its maximum value of Stot/Btot (all offline learning). HybridJ is also appropriate for a variant of the Online model not considered earlier, in which all learning is triggered by the act of performance but in which learning runs to completion over time that includes break periods.

HybridE and HybridP. The HybridE and HybridP models assume that there is both offline and online learning, but they posit that those two terms of the skill equation map exclusively to either online or offline learning. Hence, these two models and the HybridJ model address independent rather than competitive hypotheses. In the HybridE model, the exponential RT improvement [$e^{(-c*T-1)}$] occurs offline and power RT improvement [S^{k}] occurs online.

$RT_{\text{skill-HybridE}} = a + b{*}e^{\{-c*(T-1)\}}{*}S^{-k}$, **(Equation 8a)**

yielding,

$RT_{\text{overall-HybridE}} = RT_{\text{skill-HybridE}} + RT_{\text{RI}}$. **(Equation 8b)**

The HybridP model assume the reverse,

$RT_{\text{skill-HybridP}} = a + b{*}e^{(-c*S)}{*}(T-1)^{-k}$. **(Equation 8c)**

yielding,

$RT_{\text{overall-HybridP}} = RT_{\text{skill-HybridP}} + RT_{\text{RI}}$. **(Equation 8d)**

Results

Errors

The error rate was calculated for each participant as the number of incorrect key presses prior to each (correct) sequence within each trial. Averaging over sequences, trials, and participants in the training phase, the error rate was .17, .20, .39, and .20 key presses in the M30, M10, S30, and S10 groups, respectively. Given five key presses for a correct sequence, the proportional key press error rate is ranged from .033 to .073 across groups. A mixed factors Factorial Analysis of Variance (ANOVA) on the error rate revealed no effect of either trial type (massed vs. spaced), F(1, 164) = .71, p = .4, d = .0043 or break time (10 s vs. 30 s), F(1, 164) = .3, p = .59, d = .0018, and no interaction F(1, 164) = .71, p = .4, d = .0043. The mean error rates across sequences within-trial, averaged over training trials and participants, are depicted in Figure 2.2. The Error rate prior to the first sequence was relatively high for all groups, suggesting a "warm-up effect" on performance at the beginning of each trial. For the 5 sequence groups, the number of errors made prior to sequences 2 through 5 is roughly constant with no significant Pearson correlation between error rate and sequence number, r(330) = .032, p = .56. However, for the 15 sequence groups, there is a gradual increase in error rate from sequence 2 onward, r(1188) = .12, p < .0001. This suggests that in the 15 sequence groups only, within-trial RI manifested not just in correct sequence RTs but also to some extent in the error rate.

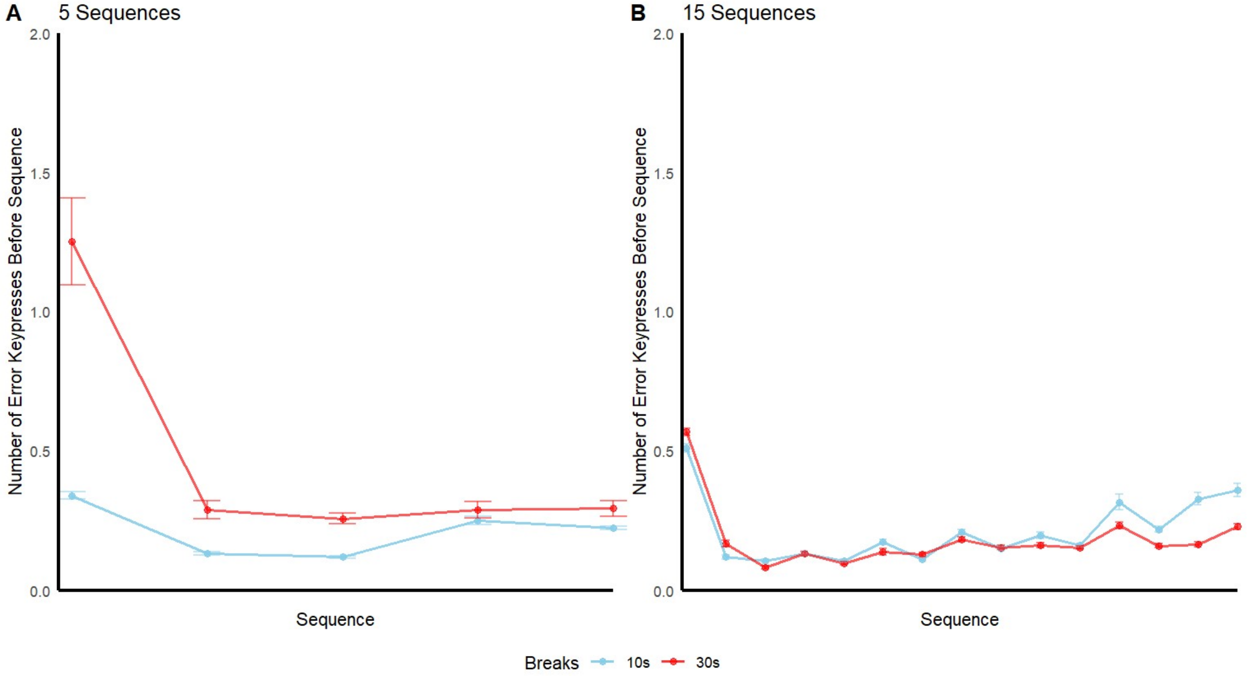


Figure 2.2 The x-axis corresponds to the number of the sequences completed within the trial. The y-axis is the average number of error keypresses before a completed sequence. For example, the first data point refers to the average number of error keypresses before the first sequence. Error bars are standard error.

Correct Sequence RTs and Model Fits

As shown in Supplement 1, the first sequence of each training and test trial had RTs that were far longer than that for other sequences, mirroring the higher error rate prior to those sequences. Those outlier sequences were removed prior to model fitting. Individual participant RTs can be seen in Supp 8. The mean sequence RTs over participants are shown in Figure 2.3 for all groups, along with fits of the Online model, Offline model version 1 that assumes no RI, and the HybridJ model. Version 1 of the Offline model is unable to capture either the pronounced RI effects (RT increase) over sequences within-trial or the prominent residual RI effect across trials, particularly for the massed groups. However, the Online and HybridJ models were also able to capture those and other major patterns in the data. Note the curvilinear RT prediction over sequences for the first few trials that is most prominent for the Online model and present for the HybridJ model but is absent for the Offline model version 1 (as well as version 2). That curved form reflects the combined effects of the linear RI effects over sequences and the non-linear decrease on RT skill curve. For all of the model fits overlaid, see Supplement 5.

Also evident in Figure 2.3 are at least two deviations of the data from the predictions, even for the best fitting HybridJ model (see discussion below), that seem unlikely to reflect chance. First, RTs on the first test trial are systematically larger than the model predictions. This suggests that there is a longer warm-up period after a 5-minutes rest than after 10 or 30 s breaks between trials during training (there was no warmup trial prior to the test). Second, for the M10 group, the predicted RTs on the last test trial are longer than the observed RTs for all sequences.

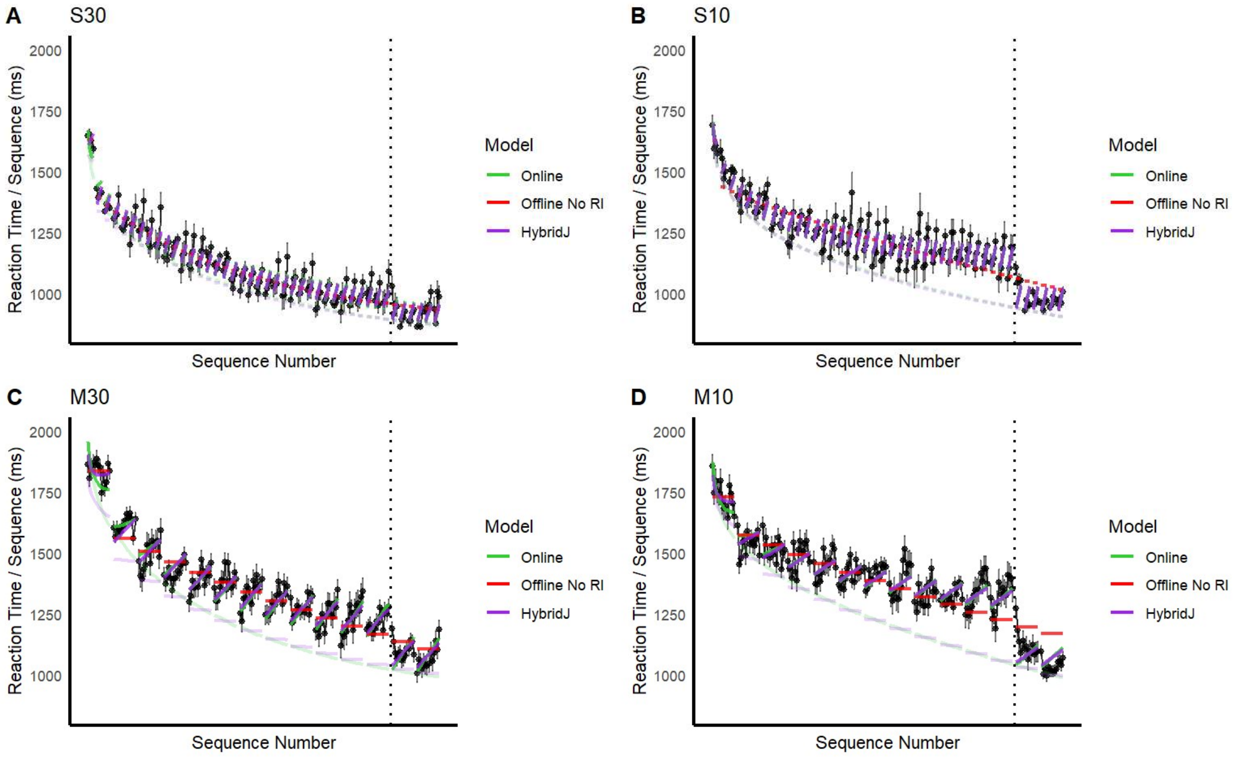


Figure 2.3 Each black dot is the RT of one completed sequence. Error bars are the standard error. The darker lines depict the overall model fits. The faint lines underneath is the estimated achieved skill.

Model fits to correct sequence RTs were assessed using the Bayesian Information Criterion (BIC) applied to least-squares parameter estimation, BIC = n*ln(RSS/n) + h*ln(n),

where lower values correspond to better fits. This criterion combines evidence of fit quality (the residual sums of squares; RSS) with a penalty for model variants with more free parameters (h). Results of each group and model, along with the mean BIC scores over groups for each model, are shown Table 2.

Version 1 of the Offline model, which assumes no RT and best represents the micro-consolidation account in the literature, yielded the worst fits across all models for all four experimental groups. In the rankings described next, we will ignore that model. Version 2 of the Offline model that assumed RI provided the best fit to the S30 group, the worst fit to the S10 group, an intermediate fit to the M30 group, and the second worst fit to the M10 group. The Online model yielded better fits overall than the Offline model. It provided the best fit for the S10 group, intermediate fits across models for the S30 and M10 groups, and the second worst fit for the M30 group.

Those results are qualified, however, by those for the hybrid models. HybridJ provides the best fits overall, including the best fits for the M10 and M30 groups. In those fits parameter *j* provides an estimate of the proportion of the observed RT improvement that is due to the Offline model (version 2) as opposed to the Online model. The fitted values of *j* were 4.097, zero, 13.42, and 11.81 for the S30, S10, M30, and M10 groups, respectively (Supplemental Figure 2.4). Hence, the HybridJ fits suggest that 100%, 0%, 89%, and 78% of the learning, respectively for those groups, was due to offline learning. Those results suggest a larger contribution of offline learning for three of the four groups, and fully Online learning for the S10 group. We take those results as preliminary, given the relatively low correlation coefficients of the j parameter in our parameter recovery analysis (Supplemental Figure 2.7). Fine grained estimation of the relative influence of online and offline components awaiting further research.

Given that the HybridJ results suggest both online and offline learning, the HybridE and HybridP models can address the complementary question of whether the power and exponential components of the skill function might selectively map onto those two types of learning. The substantially better BIC fits for HybridE compared to HybridP, combined with the second-best fits overall for HybridE, suggests tentatively that online learning may manifest mostly as power RT gains, whereas offline learning may manifest mostly as exponential RT gains. We will return to that finding in the Discussion.

Table 2.2 For each model the BIC was calculated. Models with better fits, penalizing for the number of parameters have lower BIC values. Numbers in green indicate the best fit for that group, whereas red indicates worse fits for that group.

Group	Offline No RI	Offline	Online	HybridJ	HybridE	HybridP
S30	1315.558	1229.712	1241.242	1232.569	1240.359	1254.786
S10	1477.377	1294.249	1279.425	1284.630	1279.534	1293.744
M30	1545.205	1475.403	1501.229	1468.726	1484.094	1472.465
M10	1611.154	1597.050	1569.606	1569.044	1569.079	1595.570
Mean	1487.323	1399.103	1397.876	1388.742	1393.266	1404.141

Reactive Inhibition and the Skill Function Across Groups

For simplicity in characterizing the patterns for RI, we will focus here on the Online model. The following patterns held across all models that included RT with minor variations in the estimated parameter values (Supplemental Figure 2.4). First, in the context of both massed and spaced training, the rate of within-trial RI build-up, as estimated by the parameter *y*, was minimally related to break time. This result suggests that the within-trial RI effect is relatively independent of the rate of accrual of residual RI over trials, which was greater in the massed groups. In contrast, the *y* estimates were significantly smaller for the two massed practice (M) groups (11.68 and 8.11) than for the two spaced practice (S) groups (28.52 and 32.25). Hence, the rate of per sequence RT increase was substantially larger than in the spaced groups, even though total slowing over sequences and accrual of residual RI was greater in the massed groups. That result was not expected. However, in light of the increasing error rate over sequences for the massed but not the spaced groups (Figure 2.2), it may reflect a speed-accuracy trade-off for the massed groups. In that speculative account, participants in massed groups achieved a smaller rate of RT increase across sequences within-trial at the expense of a progressively increasing error rate over those sequences. The accrual of residual RI across trials (parameter *z*), differed across groups in an ordinal pattern that is consistent with expectations based on our prior work, having estimated values of 1.032 ms, 4.286 ms, 10.88 ms, and 19.12 ms per trial in the S30, S10, M30, and M10 groups, respectively.

The results for achieved skill are best understood by plotting the curves for the four groups on the same graph (Figure 2.4). In this comparison we used the HybridJ model fits as a reference, but similar patterns were observed for the other models (Supplemental Figure 2.5). The curves are highly similar for the two massed groups, and similar, to a somewhat lesser degree, for the two spaced groups. In contrast, there is a clear gap between the massed and spaced groups throughout training. We will consider the theoretical implications of these results in the Discussion.

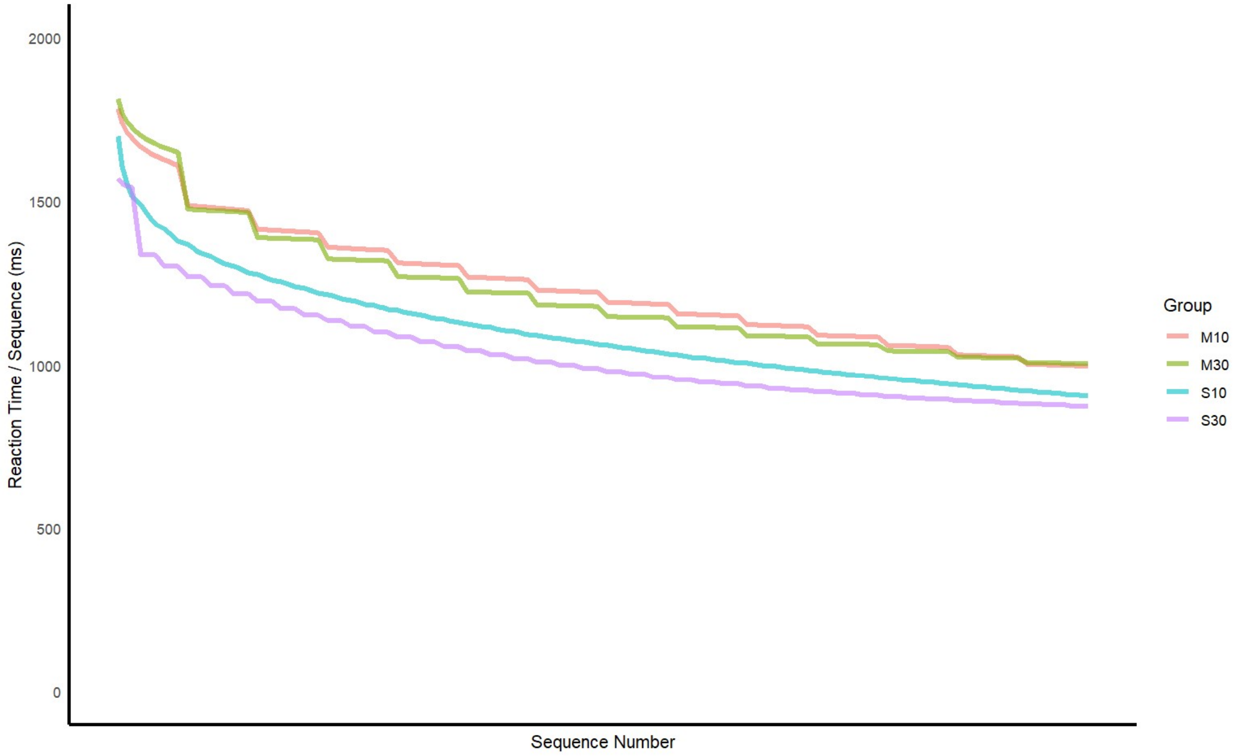


Figure 2.4 Achieved skill estimates from the HybridJ model for training trials for each group.

Discussion

We explored three hypotheses for learning over the time course of explicit motor sequence practice. The models differ with respect to whether learning occurs online, offline, or both. Each model incorporates a flexible function for mapping learning to RT that includes both exponential and power rate parameters. Each involves the same quantitative treatment of both within-trial and residual RI (with the exception of the Offline version 1 model), with assumed constant magnitude within-trial and residual RI effects across training and test phases.

Comparison of the reference Offline models that did and did not account for RI confirmed that inclusion including RI yields much better fits. Across all models that did include RI, fit quality by the BIC values differed. The HybridJ model fitted best, suggesting both online and offline learning occur. Comparison of the HybridE and HybridP model fits favored HybridE, suggesting that online learning yields primarily power function RT gains, whereas offline learning yields primarily exponential RT gains.

Reactive Inhibition and Skill Learning

Although we expected the current finding that the residual RI over trials would be greatest in the M10 group and least in the S30 group, we were agnostic about the relative magnitude of that effect in the M30 vs S10 groups because the relative rate of RI accrual during performance vs. RI resolution during a break was not known. We observed greater residual RI build-up in the M30 group than the S10 group (see CIs for the z parameter in Supplemental Figure 2.4), indicating that the difference in accrual of RI for the 15 vs. 5 sequences per trial exceeds the difference in the resolution of RI for a 10 s break vs. a 30 s break.

The clear difference in model fits between version 1 and version 2 of the Offline model strongly indicate the necessity of RI. Some authors have assumed or stated that RI is minimal on early training trials and becomes progressively more pronounced over trials once asymptotic performance has approached[5-8,11]. However, in the current models the linear effect of RI on RT within-trial was constrained to be the same across all training and test trials, yielding good overall fits. More generally, our results make it clear that modeling of RI effects should be central to any future work on the nature of motor sequence learning.

Given the relatively small step of RT improvement from the end of training to the beginning of the test for the S30 group, the estimated RT_{skill} curve for that group (Figure 2.3) may be a close approximation of true underlying achieved skill, as anticipated by experimental design. Given that the achieved skill curves in that Figure are highly similar for the two spaced groups, we can tentatively conclude that the curve for the S10 group also approximates achieved skill on the first sequence of every trial. The longer RT_{skill} values in the massed groups in Figure 2.4 is consistent with either of two interpretations, between which we cannot distinguish here. First, it may be that a 5-minute break was not sufficient to fully resolve the accumulated RI for the massed groups, and that the residual RI that remained at the beginning of the test in those groups was absorbed into the achieved skill curve in the least squares fits. By that account, actual achieved skill may be the same in all four groups, but the "achieved skill" estimate for the massed practice groups is contaminated by persistent residual RI, yielding longer RTs. Alternatively, the curves for the massed groups may reflect a lower level of achieved skill; that is, it may be that the higher level of RI in the massed groups adversely affected not only the observed performance but also the amount of underlying skill learning. The question of whether RI affects both performance and learning has clear relevance for both learning theory and optimization of skill training, and warrants further investigation. A few early studies addressed that question for other types of motor skill tasks, but no strong consensus was reached[18-20].

Fast and Slow Learning

Our model includes two learning rate components, a power component that yields rapid RT gains during early practice and slower gains later and an exponential component that yields constant proportion RTs gains from sequence to sequence or trial to trial (i.e., relatively smaller RT gains early in practice; Supplemental Figure 2.6). As noted earlier, those RT non-linearities are presumed in the current models to reflect solely the mapping from learning to RT. However, they may instead reflect differences between a *fast* learning process that yields early RT gains and a *slow* learning process that can yield more evenly distributed RT gains. The plausibility of separate fast and slow learning processes in motor sequence learning stems from the discovery of fast and slow processes in motor adaptation studies[25-26]. In those studies, the fast process has been linked to declarative (i.e., hippocampally mediated) learning and the slow process to nondeclarative (e.g., basal ganglia mediated) learning. In the HybridE model, we forced the fast power RT gains to occur exclusively online and the slower exponential gain to occur exclusively offline. In the HybridP model, we forced the reverse. Overall, the BIC results clearly favored the HybridE model between those two, and the overall fit of HybridE lagged only to the complementary HybridJ model. Those results suggest that in the motor sequence task, the fast declarative learning occurs online and the slower nondeclarative learning occurs offline. That conclusion, though speculative, appears to be inconsistent with a micro-consolidation perspective that assumes hippocampal to neocortical replay[7-8] – which presumably involves the same mechanism as does traditional declarative memory consolidation – occurs offline.

Sleep Research Bearing on Online vs. Offline Learning

There is a long-held belief that facilitating motor consolidation occurs during sleep[11-12,21-23]. However, when experimental design and analysis confounds – including RI, circadian rhythms, and averaging over online learning in the data analysis – are controlled for or mitigated, the post-sleep performance gain in motor sequence learning is virtually eliminated[11-12,21]. Most recently, meta-analytic evidence for substantial publication bias in that literature has been reported[23]. When the effects of both publication bias and confounding factors are simultaneously adjusted for, the data suggest that some degree of forgetting rather than performance improvement occurs during the hours-long offline periods of both wakefulness and sleep. The current results leave open the possibility of facilitating offline motor consolidation over brief waking periods but not over sleep periods.

Limitations

The current study is subject to several limitations. First, it does not explain the slower than expected RTs across the first several sequences on the first test trial. We assume that is due to an extended "warm-up" effect. We also do not explain the mechanistic basis of the pronounced first sequence "warm-up" effect that was observed on all trials (Supplemental Figure 2.1). Our approach to analysis in which those initial events of a trial are ignored is not unique in the literature, however, and it does not appear to compromise our main conclusions. Second, our inference that the lower rate of RI build-up over sequences within-trial in the massed groups may be a result of a speed-accuracy trade-off is speculative. Third, our use of an achieved skill function that includes two rate parameters is not strongly motivated a priori, although it is consistent with recent claims of separate learning rates for the declarative and nondeclarative components of motor adaptation learning[24-27]. As a practical matter, we used a function with two rate parameters because it yielded better data fits than did either the exponential or power function alone (Supplemental Figure 2.3). Fourth, our assumption that the estimated achieved skill curve reflects the true latent skill level, while plausible for the spaced groups, is less certain for the massed groups, as discussed above. Fifth, although the HybridJ model had superior fits, in our model recovery analysis (Supplemental Figure 2.7), the correlation between the recovered and known parameters was relatively low for the j parameter and others. Hence, one should approach the estimated values of j (reflecting both online and offline learning) with caution. Finally, we attempted to fit the non-linear models at the participant level, but were unable to achieve consistent convergence, and we further suspect that fits at that level are complicated by local minimum solutions. To assure optimal fits to the averaged data, we performed the non-linear fits independently in two statistical programming languages (R and SAS) – confirming that they both converged on the same RSS values – and in all fits we used an extensive starting parameter grid search prior to commencement of gradient descent.

Directions for Future Work

Conclusive differentiation among the candidate models and estimation of the relative influence of online and offline learning will likely require experiments that are designed to specifically differentiate between parameter values and a shift to state-space modeling, along with datasets with decreased error variance. This will allow for greater ease of fitting a 7-parameter model, compared to our experiment which had little group differentiation.

State-space modeling would further elucidate how changes in learning and performance unfold over time. In the current context, there are at least three processes that may occur exclusively over time: dissipation of RI during breaks, saturation of offline learning during breaks, and saturation of online learning over time in the noted alternative online account in which learning is triggered immediately by performance but runs to completion during the break. In all three cases, initial modeling might assume a single-parameter exponential time function, with progression to more complex functions as justified by the BIC measure. The state-space approach will provide stronger constraints, and perhaps deeper biological insights, in subsequent work. Strategies for achieving more systematic data in future work include reducing the variability in participant-level sequence RTs (e.g., by increasing the number of key presses in each sequence), substantially increasing the participant sample size, and a closer examination of initial learning where predictions between the models are most varied (Supplemental Figure 2.9) will be needed.

Conclusion

Three classes of models based on when motor learning occurs were tested: online, offline, and hybrid. We showed the necessity of including RI as a central component in any such model. The results favor a hybrid model in which both online and offline learning occur. The quantitative modeling framework described here provides researchers with a new systematic and integrated approach to investigate mechanisms that underlie motor sequence learning and performance.

Methods

Participants

All 148 participants were right-handed. 42 participants in 15 correct sequences per trial, 30 s breaks (age = 20.38, F = 90.5%). 42 participants were in the 5 correct sequences per trial, 30 s break group (age = 20.83, F = 83.3%). 43 participants were in the 15 correct sequence per trial on, 10 s break group (age = 21.28, F = 79.1%). 41 participants were in the 5 correct sequences per trial on, 10 s break group (age = 20.02, F = 82.9%). Participants provided informed consent via button press. All procedures were approved by the institutional review board of the University of California, San Diego and all methods were performed in accordance with the relevant guidelines and regulations.

Experimental design and procedure

Participants performed a standard finger-tapping-task where they repeated the sequence, 4-1-3-2-4 (see Figure 2.5), as quickly and accurately as possible with their non-dominant left hand[14]. A 2x2, between-participant design was used, with factors of Number of Sequences (5 or 15 sequences) and Break Period between trials (10 s or 30 s). All participants completed the same total number of sequences. After the 180 completed sequences of on-task training, there was a 300 second rest where participants performed a distraction task of double-digit addition. Afterwards, they performed 30 sequences of test trials with breaks in between in the same conditions that they trained on. With this design, we were able to replicate our previous results that used a time based constraint (Supplemental Figure 2.2; Gupta & Rickard, 2022).

Figure 2.5. Finger tapping task. Participants learned a motor sequence task during a single session. They were instructed to repeatedly type a sequence, 41324, with their non-dominant left hand as fast and as accurately as possible. Keypress 4 was performed with the index finger, keypress 3 with the middle finger, keypress 2 with the ring finger, and keypress 1 with the pinky finger. Participants trained for a total of 180 sequences with either 5 or 15 sequences per trial. In between practice trials were either breaks of 10 s or 30 s. After training, participants performed 300 s of double-digit addition. They were then tested on the practiced sequence for another 30 sequences with the same trial and break lengths during training.

Statistical analysis

The first completed sequence of each trial was considered a warm-up sequence and was removed prior to the RT data analysis[9] (see fig S2). Thus, participants in the 15 sequence group had more data analyzed because they had fewer trials.

Keypress latency was defined as the time between temporally adjacent keypresses. We first log-transformed individual keypresses for correct sequences. The log-transformation reduced noise without changing the overall data pattern. Next, the log latency was averaged over the 5 keypresses of each sequence. Those averaged log latencies were then anti-logged and multiplied by 5 to get the RT for each sequence. The graphed data points are those sequence RTs average over participants.

Model Fitting

The power component of Equation 6a, $(T-1)^{-k}$, would yield a divide by zero error for the pure case of (T-1), because on trial one, (T-1) would equal 0. To resolve that issue both here and in later described models that have T in the skill function, in the model fitting program we substituted (T - 0.9999999999) for (T-1). This minor adjustment on the other model fits made no difference when using 1 or 0.999999999. Thus, for each equation where one was subtracted either from the trial number or cumulative sequence number, we used 0.999999999.

Data availability

All data and code (stimuli and analyses are available online (https://osf.io/j3uc8/). Further information and requests for all data and code should be directed to and will be fulfilled by the corresponding author, TCR (trickard@ucsd.edu).

Competing Interest Statement: We have no competing interests to report.

Acknowledgements: We graciously thank Dr. Ethan Buch for sharing their data with us.

Author Contributions: MWG conducted the experiments and analyzed the data. Both MWG and TCR created the model, designed the experiment, and wrote the article.

References

1. Walker, M. P., Brakefield, T., Morgan, A., Hobson, J. A., & Stickgold, R. Practice with sleep makes perfect: Sleep-dependent motor skill learning. Neuron, 35(1), 205–211 (2002). https://doi.org/10.1016/S0896-6273(02)00746-8

2. Hotermans, C., Peigneux, P., De Noordhout, A. M., Moonen, G., & Maquet, P. Early boost and slow consolidation in motor skill learning. Learning and Memory, 13(5), 580–583 (2006). https://doi.org/10.1101/lm.239406

3. Albouy, G., Ruby, P., Phillips, C., Luxen, A., Peigneux, P., & Maquet, P. Implicit oculomotor sequence learning in humans: Time course of offline processing. Brain Research, 1090(1), 163–171 (2006). https://doi.org/10.1016/j.brainres.2006.03.076

4. Robertson, E. M. (2019). Skill Memory: Mind the Ever-Decreasing Gap for Offline Processing. In *Current Biology* (Vol. 29, Issue 8, pp. R287–R289). Cell Press. https://doi.org/10.1016/j.cub.2019.03.007

5. Bönstrup, M., Iturrate, I., Thompson, R., Cruciani, G., Censor, N., & Cohen, L. G. (2019). A Rapid Form of Offline Consolidation in Skill Learning. *Current Biology*, *29*(8), 1346-1351.e4. https://doi.org/10.1016/j.cub.2019.02.049

6. Bönstrup, M., Iturrate, I., Hebart, M. N., Censor, N. & Cohen, L. G. Mechanisms of offline motor learning at a microscale of seconds in large-scale crowdsourced data. Npj Sci. Learn. 5, 1–10 (2020).

7. Jacobacci, F., Armony, J.L., Yeffal, A., Lerner, G., Amaro Jr, E., Jovicich, J., Doyon, J. and Della-Maggiore, V., 2020. Rapid hippocampal plasticity supports motor sequence learning. *Proceedings of the National Academy of Sciences*, *117*(38), pp.23898-23903.
8. Buch, E. R., Claudino, L., Quentin, R., Bönstrup, M., & Cohen, L. G. Consolidation of human skill linked to waking hippocampo-neocortical replay. Cell Reports, 35(10) (2021). https://doi.org/10.1016/j.celrep.2021.109193

9. Gupta, M.W., Rickard, T.C. Dissipation of reactive inhibition is sufficient to explain post-rest improvements in motor sequence learning. *npj Sci. Learn.* **7**, 25 (2022). https://doi.org/10.1038/s41539-022-00140-z

10. Hull, C. L. (1943). Principles of behavior. New York: Appleton-Century- Crofts.

11. Rickard, T. C., Cai, D. J., Rieth, C. A., Jones, J. & Ard, M. C. Sleep does not enhance motor sequence learning. J. Exp. Psychol. Learn Mem. Cogn. 34, 834–842 (2008).

12. Brawn, T. P., Fenn, K. M., Nusbaum, H. C. & Margoliash, D. Consolidating the effects of waking and sleep on motor-sequence learning. J. Neurosci. 30, 13977–13982 (2010).

13. Török, B., Janacsek, K., Nagy, D. G., Orbán, G. & Nemeth, D. Measuring and filtering reactive inhibition is essential for assessing serial decision making and learning. J. Exp. Psychol. Gen. 146, 529–542 (2017).

14. Nissen, M. J. & Bullemer, P. Attentional requirements of learning: evidence from performance measures. Cogn. Psychol. 19, 1–32 (1987).

15. Rickard, T. C. (2004). Strategy Execution in Cognitive Skill Learning: An Item-Level Test of Candidate Models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(1), 65–82. https://doi.org/10.1037/0278-7393.30.1.65

16. Heathcote, A., Brown, S., & Mewhort, D. J. K. (2000). The power law repealed: The case for an exponential law of practice. Psychonomic Bulletin and Review, 7, 185–207.

17. Spiess A.N., Neumeyer N. An evaluation of R^2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach. BMC Pharmacol. 2010 Jun 7;10:6. doi: 10.1186/1471-2210-10-6. PMID: 20529254; PMCID: PMC2892436.

18. Kimble, G. A. (1949). An experimental test of a two-factor theory of inhibition. Journal of Experimental Psychology, 39(1), 15–23. https://doi.org/10.1037/h0058281

19. Carron, A. V. (1969). Physical fatigue and motor learning. Research Quarterly of the American Association for Health, Physical Education and Recreation, 40(4), 682–686. https://doi.org/10.1080/10671188.1969.10614902

20. Carron, A. V. (1969). Performance and learning in a discrete motor task under massed vs. distributed practice. Research Quarterly of the American Association for Health, Physical Education and Recreation, 40(3), 481–489. https://doi.org/10.1080/10671188.1969.10614866

21. Nettersheim, A., Hallschmid, M., Born, J., & Diekelmann, S. (2015). The Role of Sleep in Motor Sequence Consolidation: Stabilization Rather Than Enhancement. *Journal of Neuroscience*, *35*(17), 6696–6702. https://doi.org/10.1523/JNEUROSCI.1236-14.2015

22. Pan, S. C., & Rickard, T. C. (2015). Sleep and motor learning: Is there room for consolidation? *Psychological Bulletin*, *141*(4), 812–834. https://doi.org/10.1037/bul0000009

23. Rickard, T. C., Pan, S. C., & Gupta, M. W. (2022). Severe publication bias contributes to illusory sleep consolidation in the motor sequence learning literature. *Journal of Experimental Psychology: Learning Memory and Cognition*, (2008). https://doi.org/10.1037/xlm0001090

24. Brown, R. M., & Robertson, E. M. (2007). Inducing motor skill improvements with a declarative task. *Nature Neuroscience*, *10*(2), 148–149. https://doi.org/10.1038/nn1836

25. Taylor, J. A., Krakauer, J. W., & Ivry, R. B. (2014). *Explicit and Implicit Contributions to Learning in a Sensorimotor Adaptation Task*. *34*(8), 3023–3032. https://doi.org/10.1523/JNEUROSCI.3619-13.2014

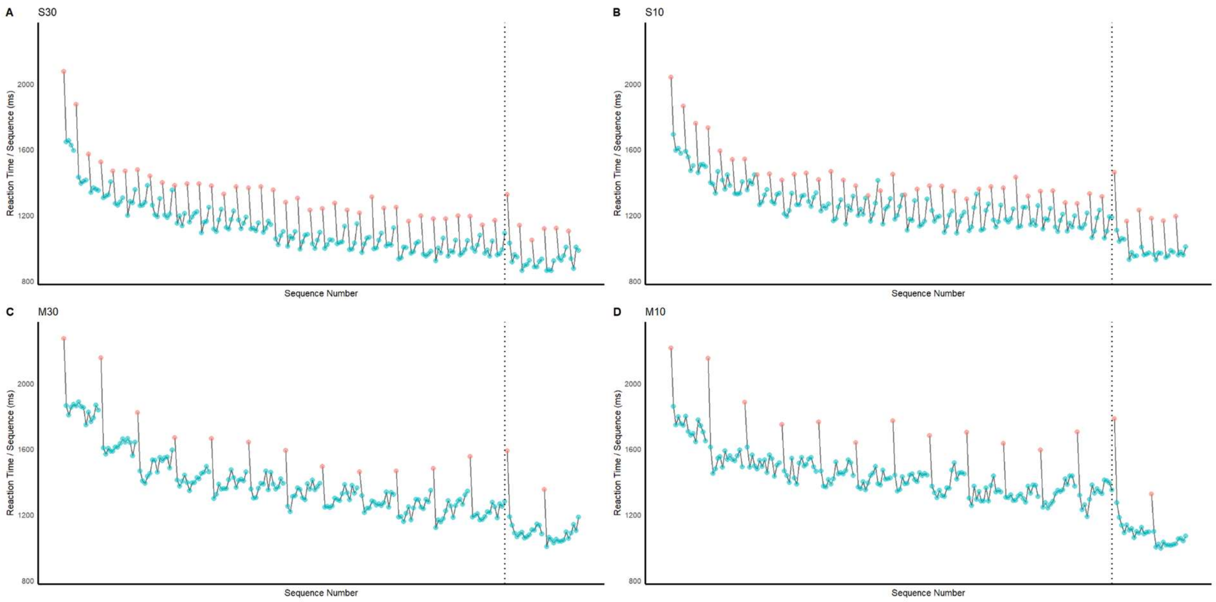
26. McDougle, S. D., Bond, K. M., & Taylor, J. A. (2015). Explicit and implicit processes constitute the fast and slow processes of sensorimotor learning. *Journal of Neuroscience*, *35*(26), 9568–9579. https://doi.org/10.1523/JNEUROSCI.5061-14.2015

27. Keisler, A., & Shadmehr, R. (2010). A shared resource between declarative memory and motor memory. *Journal of Neuroscience*, *30*(44), 14817–14823. https://doi.org/10.1523/JNEUROSCI.4160-10.2010

28. Krakauer, J. W., Hadjiosif, A. M., Xu, J., Wong, A. L., & Haith, A. M. (2019). Motor learning. Compr Physiol, 9(2), 613-663.

Supplementary

For modeling and statistical tests, we removed the first sequence of each trial because that sequence had systematically longer RTs (Supplemental Figure 2.1), as was done in Gupta and Rickard. Averaged across all trials, there is a significant difference for each group between the first and second sequence, $t(41) = 10.2$, $p < .0001$, $d = 1.57$; $t(40) = 12.4$, $p < .0001$, $d = 1.94$; $t(41) = 9.69$, $p < .0001$, $d = 1.49$; $t(42) = 10.1$, $p < .0001$, $d = 1.53$ for the S30, S10, M30 and M10 groups respectively. It is possible that since our data was collected online, this phenomenon is reflective of participants not paying attention at the beginning of each trial. However, the data from Buch et al., (2020), which was an in-person experiment, exhibit the same pattern or first sequence slowdown when plotting by sequence (Supplemental Figure 2.1). This was not noted in their original publication.

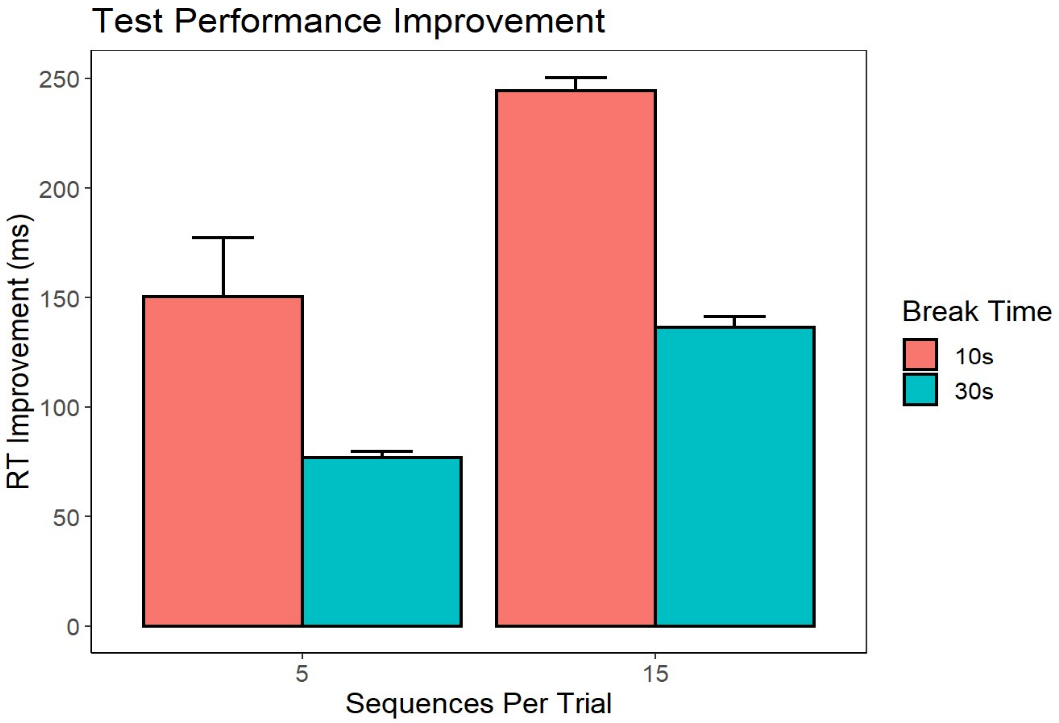


Supplemental Figure 2.1 RTs for each sequence across trials, including the first sequence, marked in red. In every trial across groups, aside for the first couple, the first sequence is markedly slower than the rest. This systematicity warranted the removal of it because it is clear there is a third variable problem influencing the RTs of that sequence.

Statistical Analysis of RI and Post-rest Gains

The current method of controlling the number of sequences completed instead of controlling for the total practice time yielded analogous results to our previous study[2], and in line with previous findings from Brawn et al., (2011) and Rickard et al., (2008). To confirm the presence of RI in the 15 sequence groups, we compared the second sequence to the 15th sequence averaged across all training trials. A one-tailed paired-samples t-test, averaged over all practice trials, yielded evidence of RI in the 30 s break group, t(41) = -3.83, p = .0002, d = -.42, but not the 10 s break group, t(42) = .33, p = .43, d = -.019. We believe this may be due to a consistently longer warmup time for this group. To confirm the presence of RI in the five sequence trial groups, we compared the second sequence to the last sequence. A one-tailed paired-samples t-test, averaged over all practice trials, yielded evidence of RI in the 30 s break group, t(41) = -2.94, p =.0027, d = -.32, and the 10 s break group t(40) = -6.35, p = <.0001, d = -.7.

We determined whether the current results replicate our previous findings regarding how the number of sequences completed and the break time affected the post-rest gain (Supplemental Figure 2.2). As in Gupta and Rickard (2022), we compared the RT means (sequence 2 onward) over the last two training trials (11 and 12) with the post-rest trials (13 and 14). A 2x2 mixed-factors Analysis of Variance (ANOVA) revealed a significant effect of break time on the post-rest gain, F(1, 164) = 9.08, p = .003, $\eta^2 = .05$, as well as a significant effect of the number of sequences during the trial on the post-rest gain, F(1, 164) = 6.66, p = .012, $\eta^2 = .04$ (Supplemental Figure 2.2). There was no significant interaction between the two factors, F(1, 164) = .321, p = .57, $\eta^2 = .002$. These results replicated the same pattern observed in Gupta and Rickard.



Supplemental Figure 2.2 The bar chart shows the total amount of gain from the last 30 sequences of training compared to the 30 test sequences after rest. The y-axis is the average amount of RT gain. The x-axis indicates the number of sequences performed during a trial, whereas the color indicates the amount of break time. These results replicate our previous findings in Gupta and Rickard (2022) that both break time and the number of sequences performed affect the amount of gain.

Parameter Necessity: Offline Model with RI

We tested for improved fits with each added parameter by the BIC criterion, using the offline model version 2 as the reference. We started with the simplest case exponential RT gain parameter (*c*) only, with no RI (SM1), then added the within-trial RI parameter (*y*; SM2), then the cumulative RI parameter (*z*; SM3), and finally power RT gain parameter (*k*; SM4). As shown in Supp 3, for each group each added parameter reduced BIC.

Supplemental Table 2.1 This table reflects the BIC values at each stage of the model development. Lower values indicate a better fit of the data for that model relative to other models.

Group	SM1	SM2	SM3	SM4
S30	1363.196	1303.667	1282.190	1241.242
S10	1456.864	1416.645	1301.861	1279.505
M30	1711.199	1632.545	1531.512	1514.275
M10	1768.305	1748.233	1612.573	1583.348

To obtain robust confidence intervals for the parameter estimates separately for each group, we performed bootstrapping, creating 999 sample datasets (Table 2). The pattern of parameter estimates provides insight regarding similarities and differences across groups.

Supplemental Table 2.2 Median parameter values are reported with 95% confidence intervals in brackets estimated from bootstrapping 999 samples. The number next to the letter is the amount of break time between trials. r^2 is reported for model fits. Note that it is inappropriate to use r^2 to compare non-linear least square models[4].

Model Type	Group	a	b	c	k	y	z	j	r^2
Online	S30	734.70 [.0051494, 816.80]	940.01 [825.89, 1673.8]	.007042 [.0011389, .011652]	.15030 [.096666, .17675]	28.52 [23.76,34.09]	1.0328 [34.09, 1.742]		.78
	S10	.2852 [0,770.1]	1664 [913.1, 1749]	.001519 [.001163, .007562]	.08734 [.07120, .1386]	32.25 [26.64, 37.75]	4.286 [3.577, 5.011]		.69
	M30	891.7 [803.3, 942.3]	1073 [985.2,1207]	.009453 [.006431,.01219]	.1020 [.07267, .1288]	11.68 [10.18, 13.04]	10.88 [8.635, 13.12]		.78
	M10	.02784 [0, 783.9]	1832 [1066, 1928]	.001782 [.001438,.005921]	.06570 [.05187, .09530]	8.111 [6.322, 9.968]	19.12 [16.83, 21.65]		.61
Offline No RI	S30	863.3 [798.6, 904.4]	556.9 [520.9, 606.2]	.04904 [.03694,.06071]	.01991 [.01558, .02406]				.66
	S10	0 [0,811.0]	1430 [653.9, 1464]	.008461 [.007261, .02525]	.00830 [.00523, .014]				.18

Supplemental Table 2.3 Median parameter values are reported with 95% confidence intervals in brackets estimated from bootstrapping 999 samples. The number next to the letter is the amount of break time between trials. r^2 is reported for model fits. Note that it is inappropriate to use r^2 to compare non-linear least square models[4]. (Continued)

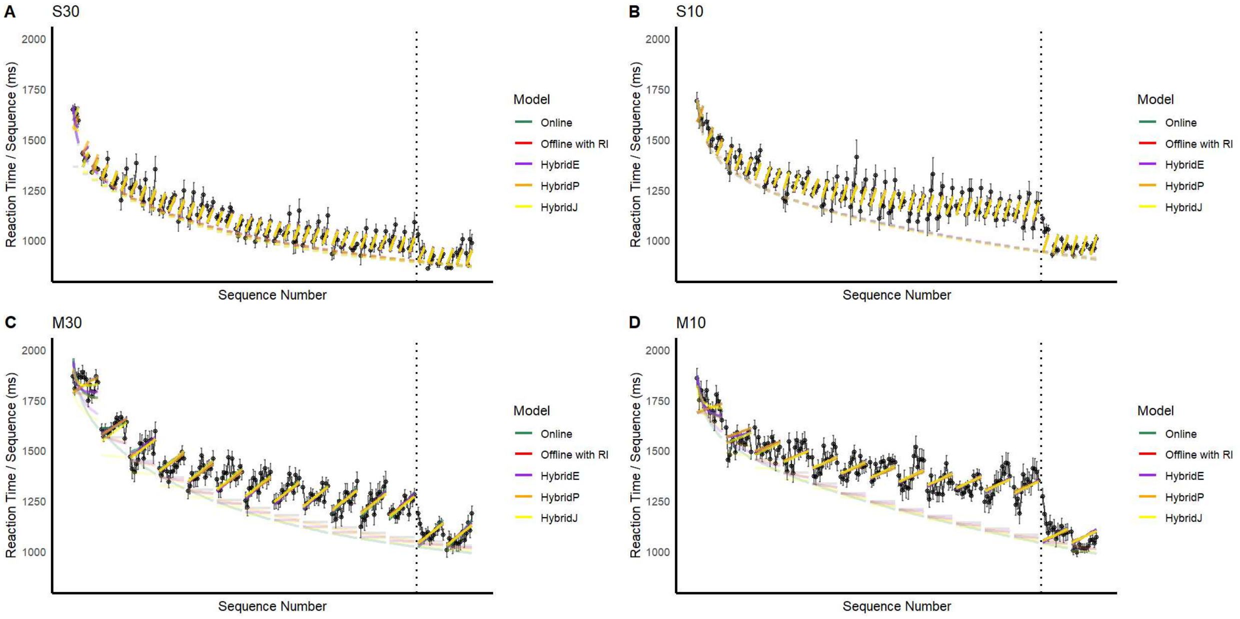
	M30	80.26 [0, 758.2]	1516 [862.2, 1617]	.02884 [.02547, .06526]	.009336 [.007567, .01416]				.63
	M10	93.21[0,952.6]	1519 [692.0, 1632]	.02630 [02183,.07940]	.005009 [.002930, .009278]				.094
Offline	S30	825.7 [794.0, 854.7]	571.2 [543.2, 596.8]	.05841 [.04845, .06876]	.01435 [.01161, .01722]	23.62 [19.04, 28.50]	1.390 [.6978, 1.979]		.80
	S10	54.47 [0, 800.8]	1546 [796.1, 1620]	.005286 [.003678, .03276]	.1087 [.08642, .1707]	27.23 [21.87, 32.95]	4.431 [3.716, 5.207]		.66
	M30	431.8 [0, 900.7]	1375 [960.6, 1805]	.01588 [.003483, .09875]	.2288 [.1686, .2950]	6.853 [5.456, 8.227]	10.45 [8.480, 12.46]		.81
	M10	3.227 [0,906.4]	1685 [863.4, 1735]	.02034 [.01365, .1216]	.1084 [.06478, .1567]	3.849 [1.966, 5.589]	19.96 [17.37, 22.80]		.55

Supplemental Table 2.4 Median parameter values are reported with 95% confidence intervals in brackets estimated from bootstrapping 999 samples. The number next to the letter is the amount of break time between trials. r^2 is reported for model fits. Note that it is inappropriate to use r^2 to compare non-linear least square models[4]. (Continued)

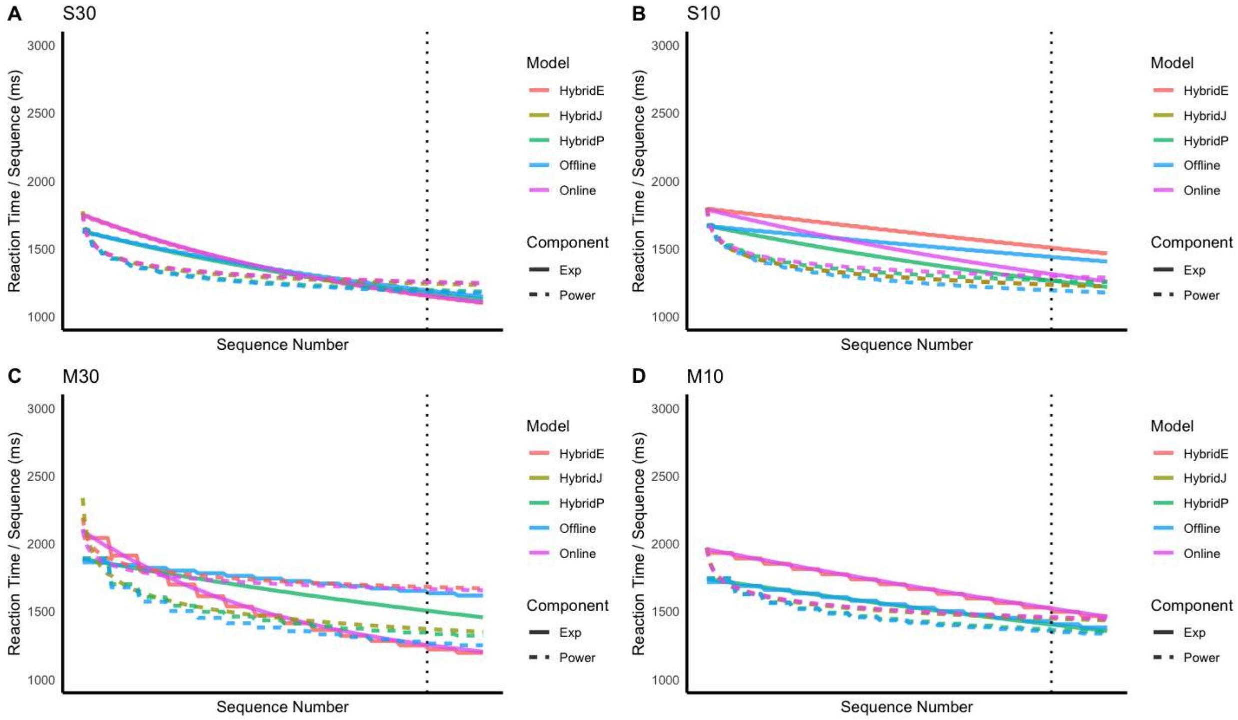
HybridE	S30	745.3 [.1324,821.0]	944.1 [847.9, 1665]	.02913 [.004932, .04626]	.1520 [.09923, .1778]	26.57 [21.67, 32.06]	1.066 [.3250, 1.697]		.78
	S10	81.67 [0, 776.6]	1643 [924.3, 1753]	.006282 [.004606, .03105]	.08860 [.07074, .1367]	30.56 [24.80, 36.44]	4.290 [3.590,5.084]		.69
	M30	922.1 [861.7, 966.7]	1170 [1100, 1252]	.1366 [.1043, .1711]	.1122 [.08739, .1362]	8.592 [7.140, 10.05]	11.23 [9.167, 13.34]		.80
	M10	.01814 [0, 778.4]	1871 [1143, 1963]	.02452 [.02, .07781]	.06658 [.05195, .09747]	6.038 [4.136, 7.731]	19.16 [16.69, 21.70]		.61
HybridP	S30	716.3 [0,821.7]	844.1 [730.6, 1562]	.005595 [.0007127, .01164]	.1873 [.1165 , .2291]	25.47 [19.99, 30.74]	1.088 [.3808, 1.844]		.77
	S10	610.9 [0, 835.9]	973.8 [752.2, 1606]	.003841 [.001005, .01017]	.1352 [.08664, .1821]	28.99 [23.24, 35.22]	4.531 [3.807, 5.374]		.66
	M30	689.6 [0, 892.56]	1104 [915.5, 1797]	.002558 [.0002829, .00696]	.261 [.171, .3115]	8.250 [6.592, 9.991]	10.55 [8.591, 12.44]		.81

Supplemental Table 2.5 Median parameter values are reported with 95% confidence intervals in brackets estimated from bootstrapping 999 samples. The number next to the letter is the amount of break time between trials. r^2 is reported for model fits. Note that it is inappropriate to use r^2 to compare non-linear least square models[4]. (Continued)

	M10	40.69 [0, 856.9]	1647 [852.7, 1716]	.001494 [.0009690, .007239]	.1111 [.07706, .1626]	5.636 [3.592, 7.778]	19.82 [17.22, 22.68]		.55
HybridJ	S30	820.6 [783.5, 850.5]	598.4 [569.0, 627.8]	.01347 [.01110, .01600]	.02508 [.02045, .02982]	23.88 [19.11, 28.60]	1.297 [.6970, 1.935]	4.097 [4.097, 4.097]	.80
	S10	598.9 [0, 805.4]	1116 [786.3, 1702]	.003989 [.001314, .009086]	.1144 [.06449, .1409]	31.91 [26.13, 37.65]	4.414 [3.706, 5.177]	.1377 [0, 3.128]	.69
	M30	860.3 [713.9, 931.8]	904.4 [800.3, 1083]	.007084 [.004475, .009520]	.06546 [.04074, .08926]	7.997 [6.487, 9.570]	9.990 [7.725, 11.87]	13.42 [10.28, 14.88]	.82
	M10	.1019 [0, 785.2]	1647 [928.2, 1821]	.001850 [.001535, .005613]	.04025 [.02181, .07056]	5.226 [3.239, 7.593]	18.82 [16.48, 21.36]	11.81 [4.623, 14.84]	.62



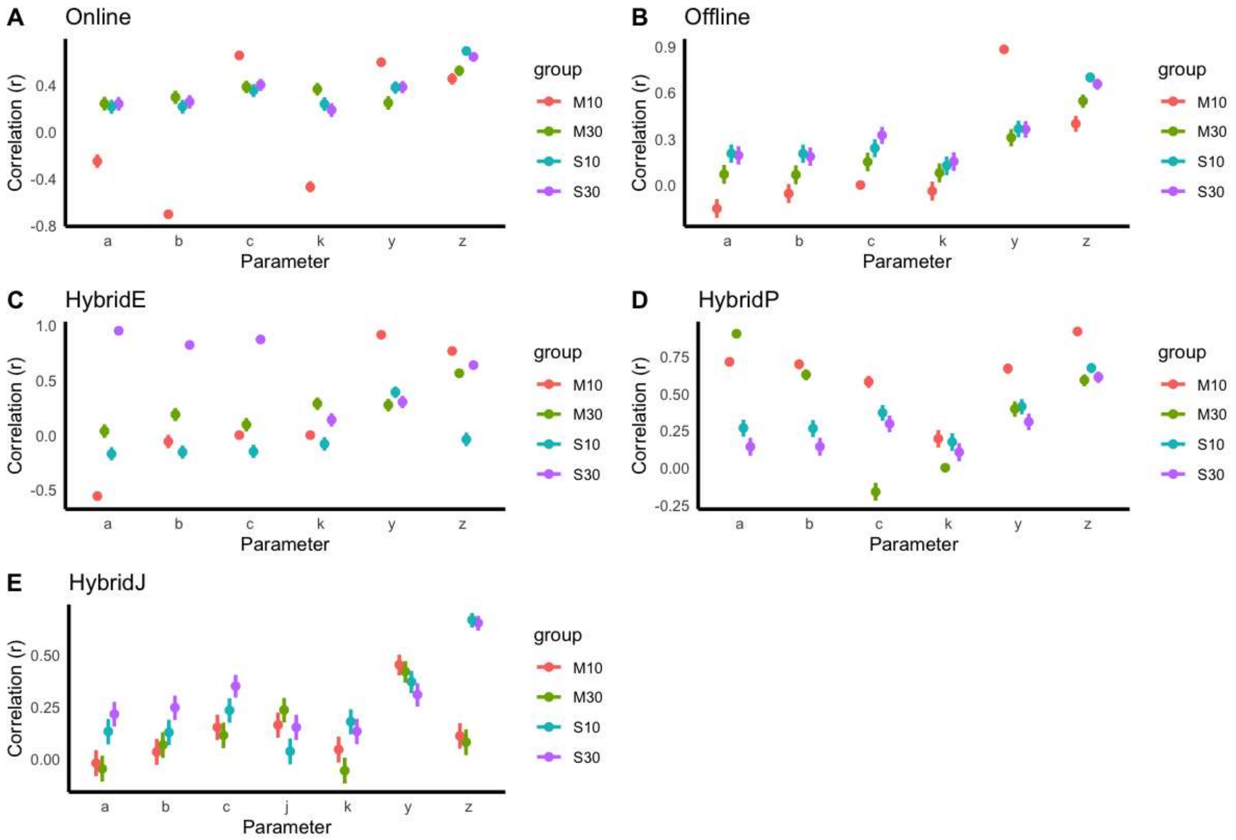
Supplemental Figure 2.3 Each black dot is the RT of one completed sequence. Error bars are the standard error. This graph shows all model fits. The faint lines underneath are the estimated achieved skill.



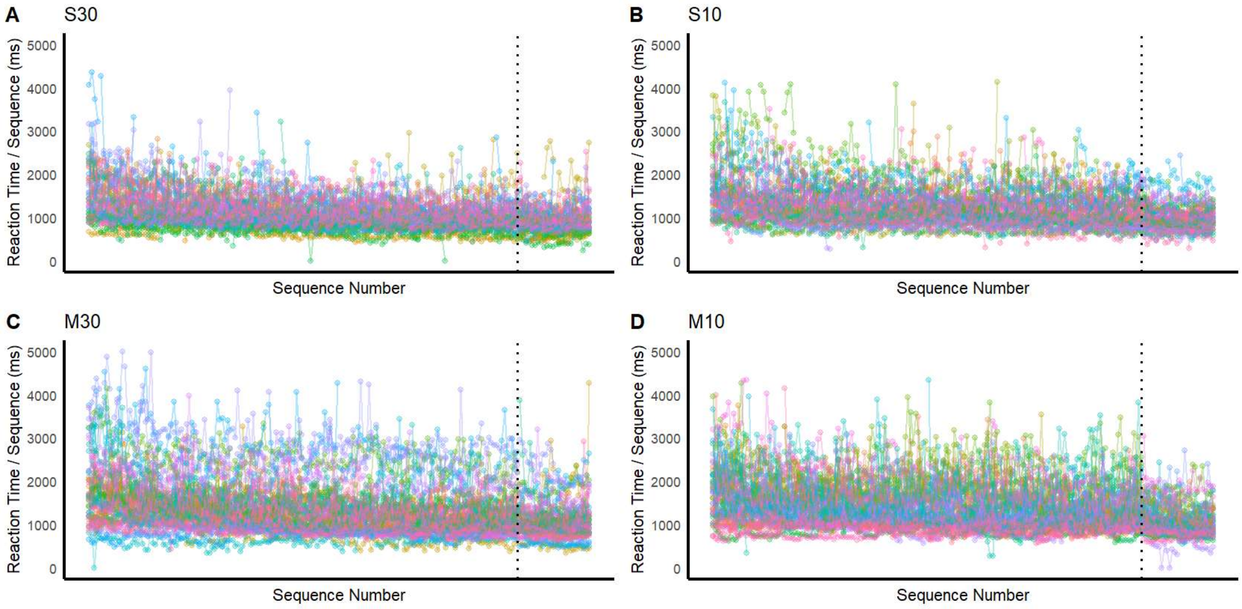
Supplemental Figure 2.4 For each model, we plotted the exponential and power components plus the *a* and *b* parameters. For each model there appears to be a fast component that is typically the power learning rate and a slow component that is typically the exponential learning rate.

Parameter Recovery

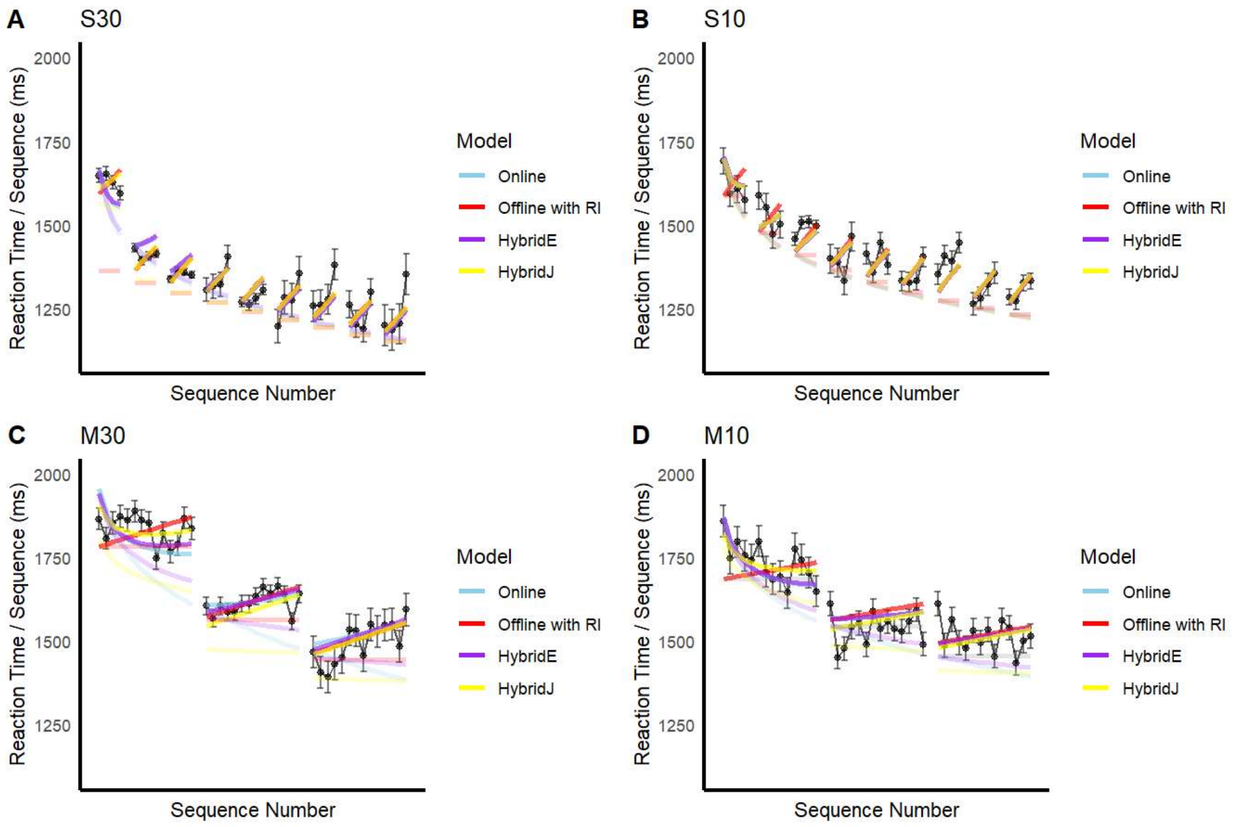
To conduct parameter recovery, we used the bootstrapped parameter values that were used to calculate the confidence intervals of the models. Since we know the true parameter values, theoretically we should be able to fit the model again and recover the known parameter values. From those known parameter values, we simulated data for each model, for each group. We then injected uniform noise into the simulated dataset that was the plus or minus of the standard deviation of the entire RT of all sequences independently for each observed dataset. We ran this simulation 999 times to obtain estimates of the known parameter values. Finally, we correlated the recovered parameter values with the known real values.



Supplemental Figure 2.5 Each model underwent parameter recovery in which the recovered values were correlated with the known values. The lines indicate the confidence intervals for each correlation. For some parameters, in each recovery the same value was obtained (e.g. k = .1 for each recovered fit). In these cases we could not obtain sensible correlations. Thus, we filled those values 0s.



Supplemental Figure 2.6 Individual participant sequence RTs for each group are plotted. Each participant has an individual color.



Supplemental Figure 2.7 Each black dot is the RT of one completed sequence. Error bars are the standard error. This graph shows model fits over the first 3 (massed) or 9 (spaced) trials for online, offline with RI, HybridE, and HybridJ. The faint lines underneath are the estimated achieved skill. Much of the model differences occurred within these trials.

References

1. Gupta, M.W., Rickard, T.C. Dissipation of reactive inhibition is sufficient to explain post-rest improvements in motor sequence learning. *npj Sci. Learn.* **7**, 25 (2022). https://doi.org/10.1038/s41539-022-00140-z

2. Brawn, T. P., Fenn, K. M., Nusbaum, H. C. & Margoliash, D. Consolidating the effects of waking and sleep on motor-sequence learning. J. Neurosci. 30, 13977–13982 (2010).

3. Rickard, T. C., Cai, D. J., Rieth, C. A., Jones, J. & Ard, M. C. Sleep does not enhance motor sequence learning. J. Exp. Psychol. Learn Mem. Cogn. 34, 834–842 (2008).

4. Spiess A.N., Neumeyer N. An evaluation of R^2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach. BMC Pharmacol. 2010 Jun 7;10:6. doi: 10.1186/1471-2210-10-6. PMID: 20529254; PMCID: PMC2892436.

Acknowledgements

Chapter 2, in full, is a reprint of the material as it appears in Scientific Reports, Gupta, Mohan W.; Rickard, Timothy C., Springer Nature, 2024. The dissertation author was the primary researcher and author of this paper.

Chapter 3 Motor Sequence Learning is Independent of Spacing, Micro-consolidation, and Reactive Inhibition

Introduction

There are many reasons why one may choose to take a break during motor skill training (e.g., playing the piano); among them are muscle fatigue, cognitive fatigue, faltering motivation, and promotion of learning. In the cognitive neuroscience literature, two non-exclusive hypotheses have been advanced regarding the effect of brief breaks on learning. The micro-consolidation hypothesis suggests that learning occurs primarily or exclusively during breaks (offline), driven by hippocampal to neocortical replay of learned sequences on the time scale of seconds[6-8,11]. In apparent support of that hypothesis, motor sequence performance immediately after a break is often better than at the end of the preceding training trial[6]. However, Gupta and Rickard (2022) showed that response time slowing due to accrual of reactive inhibition (RI) during performance, along with dissipation of RI during breaks, can account for post-break performance improvements without invoking facilitating micro-consolidation (for a candidate neurological mechanistic account of RI, see Bächinger et al., 2019). Hence, the hypothesis that motor learning occurs online (i.e., exclusively during performance) rather than during breaks remains viable. Most recently, Gupta and Rickard (in press) developed a computational model that incorporates both learning and RI effects. They advanced preliminary evidence that both online and offline learning may occur.

Although the classic effect of RI is a transient worsening of performance, several studies have addressed the possibility that RI may also negatively affect motor learning. Across alphabet printing, peg board learning, the stabilometer task, and the Tsai-Partington numbers tasks, every possible result has been found; greater learning with spaced training[3,14], no learning difference between massed and spaced training[9-10], and greater learning with massed training[2]. In a meta-analysis, Lee and Genovese (1988) concluded that massed training negatively affected both performance and learning. However, given the variety of findings, methodological limitations, and small sample sizes in some cases, we view that literature as inconclusive. Further, the relation between RI and learning has not been explored for the case of motor sequence learning that is explored here and that dominates the contemporary literature on motor learning and memory consolidation.

In the current study we investigated both whether spaced training promotes more learning than does massed training and whether RI affects not only performance but also learning. We addressed these questions by minimizing RI and maximizing the opportunity for offline micro-consolidation in a *spaced* training group (by using short duration performance trials and long breaks) and by maximizing RI and minimizing the opportunity for micro-consolidation time in a *massed* training group (by using long trials and short breaks). The spaced group in Experiment 1 performed five correct sequences during each trial with 30 s breaks between trials. The massed group performed 25 correct sequences during each trial with 10 s breaks. After a 15 minute rest period during which the transient effect of RI on performance should be completely resolved, there was a test involving the spaced task for both groups (five correct sequences per trial and 30 s breaks). The total number of training and test sequences was the same for the two groups. Because both groups performed the same task on the test, learning and RI effects across test trials should be equated. Hence, any performance difference between groups on the test should exclusively reflect differences in learning during training.

Based on prior results[12-13], the effect of RI on performance in the spaced group should be limited primarily to sequences within individual trials, with minimal RI accrual across trials. For the massed group, however, there should be substantial accrual of unresolved RI across trials. Hence, if RI impairs learning in a dose-response manner, we should see better final test performance in the spaced group. In addition, because there is a much longer total break time across training in the spaced group (1,020 s, vs 70 s in the massed group), there is more opportunity for offline micro-consolidation in the spaced group, again suggesting more learning and better final test performance. Hence, if either RI impairs learning or greater break time facilitates it (or both), we should observe better final test performance in the spaced group. However, if RI is exclusively a performance phenomenon, with no effect on learning, and if the increased break time in the spaced groups does not promote more offline micro-consolidation, then we should observe equivalent test performance in the two groups.

Experiment 1

Methods

Participants

Eighty-eight right-handed participants were recruited, 45 in the massed group (age = 21.13, F = 71.1%) and 43 in the spaced group (age = 19.98, F = 81.4%). The experiment was conducted online. Participants provided informed consent via button press. All procedures were approved by the institutional review board of the University of California, San Diego.

Experimental design and procedure

Participants performed a classic finger-tapping-task where they repeated the sequence, 4-1-3-2-4, as quickly and accurately as possible with their non-dominant left hand[16] (See Figure 3.1). Participants performed one correct warm-up sequence before starting the main task. A between-participant design was used, where the massed group completed 25 correct sequences per training trial with 10 s breaks between trials and the spaced group completed 5 correct sequences per training trial with 30 s breaks. After 175 completed sequences during the training phase in both groups, there was a 15 minute rest wherein participants in both groups performed a distraction task of double-digit addition. Afterwards, both groups performed 10 trials with 5 correct sequences per trial and 30 s breaks.

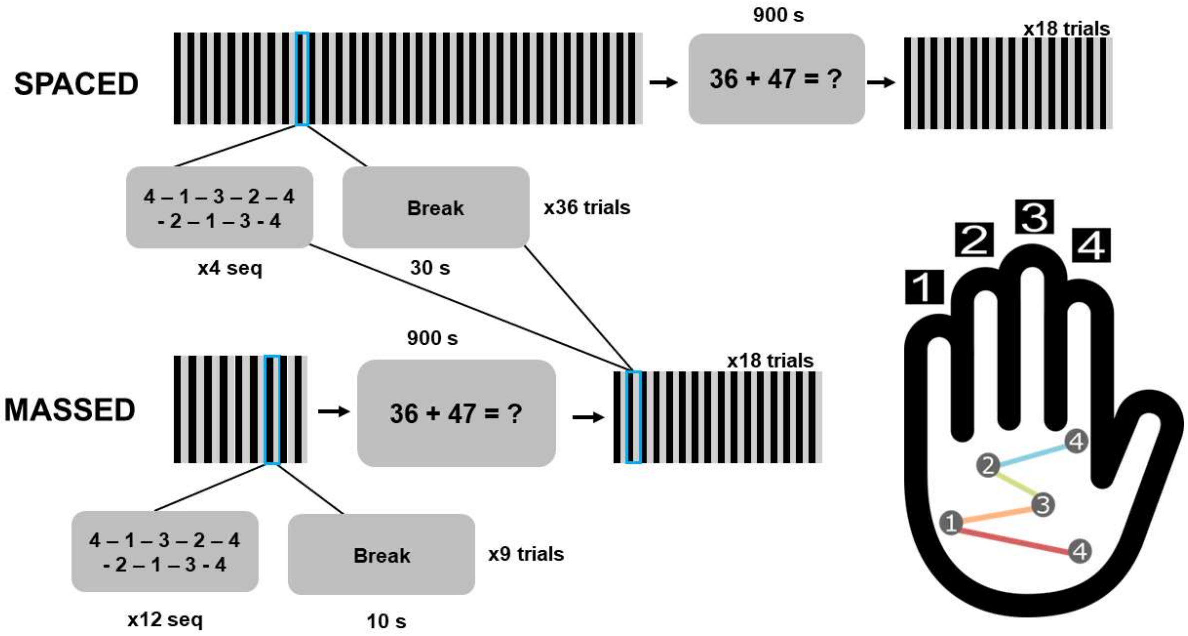


Figure 3.1 Participants learned to type a motor sequence over one session with their non-dominant left hand. They were instructed to repeatedly type a sequence, 41324, as fast and as accurately as possible. Keypress 4 was performed with the index finger, keypress 3 with the middle finger, keypress 2 with the ring finger, and keypress 1 with the pinky finger. Participants trained for a total of 175 sequences with either 5 or 25 sequences per trial. In between training trials were either breaks of 30 s or 10 s. After training, participants performed 900 s of double digit addition during the rest period. After, they were tested on the trained sequence for another 50 sequences with 5 sequences per trial and 30 s breaks.

Statistical Analysis

The dependent measure was time in seconds to complete a correct sequence. Keypress latency within sequence was measured as the time (in seconds) between temporally adjacent keypresses. To reduce noise in the data, we log-transformed the keypresses latencies. Mean of the logged keypresses was then calculated for each sequence and participant. We then anti-logged those means and multiplied by 5 to obtain a measure of sequence RT in seconds. The first completed sequence was removed from each trial prior to further analysis due to the consistently longer RTs on those sequences, indicative of warm-up[12-13].

To test our hypotheses we used both frequentist statistics and Bayes factors. All t-tests were two-tailed. The r (prior) value was set to 0.707 for all Bayes factor tests based on the recommended default Cauchy prior value[19]. We used Raftery's guidelines to interpret the Bayes factor, where 1–3 is weak, 3–20 is positive, 20–150 is strong, and > 150 is very strong evidence for the null hypothesis[17]. The Bayes factor for the null is labeled BF_{01}, whereas the Bayes factor test for the alternative hypothesis is BF_{10}.

Results

Errors

The error rate was calculated for each participant as the number of incorrect key presses prior to each correct sequence within each trial. Averaging over sequences, trials, and participants in the training phase, the error rate was 0.348 and 0.383 key presses in the spaced and massed groups, respectively. A two-sample t-test on the error rate revealed a non-significant effect between groups, $t(86) = 0.505$, $p = .61$, $d = 0.076$, $BF_{01} = 4.0$. On the test trials the error rate was 0.389 and 0.155 key presses in the spaced and massed groups, respectively. A two-sample t-test revealed a significant effect between groups, $t(86) = –2.37$, $p = .019$, $d = -0.36$, $BF_{10} = 2.53$.

Trial-level RI

To confirm the accrual of RI over sequences within trial in the massed group, for each participant we calculated the RT on the 25th sequence and subtracted the RT on the second sequence (the first sequence of each trial having been removed as warm-up). We then averaged those difference scores over trials, excluding the first trial on which there was no RT slowing over sequences. A paired-samples t-test yielded evidence of RI, $t(44) = -2.51$, $p = 0.016$, $d = -0.37$, $BF_{10} = 2.63$. In the spaced group, we compared the second sequence to the fifth sequence averaged over all training trials except for the first five trials. A paired-samples t-test again yielded evidence of

RI, $t(42) = -4.75$, $p < .0001$, $d = -0.73$, $BF_{10} = 898$. Note that the smaller p-value and larger effect size for the spaced group presumably reflects the larger number of trials in that group and hence the larger amount of data averaging for each participant in that group.

Training RTs

To test for performance differences at the end of training between the massed and spaced groups, a two-sample t-test on the mean RT for the last 25 sequences for the massed vs. spaced groups was performed, $t(86) = 1.93$, $p = 0.057$, $d = 0.29$, $BF_{10} = 1.12$ (Figure 3.2a). As a post-hoc approach to gaining more power for that test, we conducted a mixed-factors ANOVA with trials and groups as factors. For the spaced group, we combined successive sets of three trials into one to match the number of trials in the massed group. This analysis was conducted on all training trials. There was a significant effect of trial, $F(7, 560) = 19.07$, $p < .0001$, $np^2 = 0.19$, $BF_{10} = 4.55e^{15}$ but no significant effect of group, $F(1, 86) = 1.78$, $p = .19$, $np^2 = 0.02$, $BF_{01} = 1.46$. There was, however, a significant interaction between trial and group, $F(6, 560) = 20.26$, $p < .0001$, $np^2 = 0.18$, $BF_{10} = 5.14e^{18}$. Hence, there was evidence of greater accrual of residual RI across trials in the massed group.

Post-rest Improvement

Confirming the interaction apparent in Figure 3.2, a two-sample t-test on the participant-level RT difference scores (mean RT on the last 25 training sequences minus the first 25 test sequences) for the massed vs. spaced groups was significant, $t(86) = 4.41$, $p < .0001$, $d = 0.66$. The Bayes Factor test positively favored the alternative, $BF_{10} = 646$.

Final Test RTs

A two-sample t-test on the mean RTs of all of the test sequences for the massed vs. spaced groups was not significant, $t(86) = -0.74$, $p = 0.46$, $d = -0.11$ (Figure 3.2b). The Bayes Factor test positively favored the null, $BF_{01} = 3.52$. The results are comparable when only performing the test on the average of the first 25 test sequences.

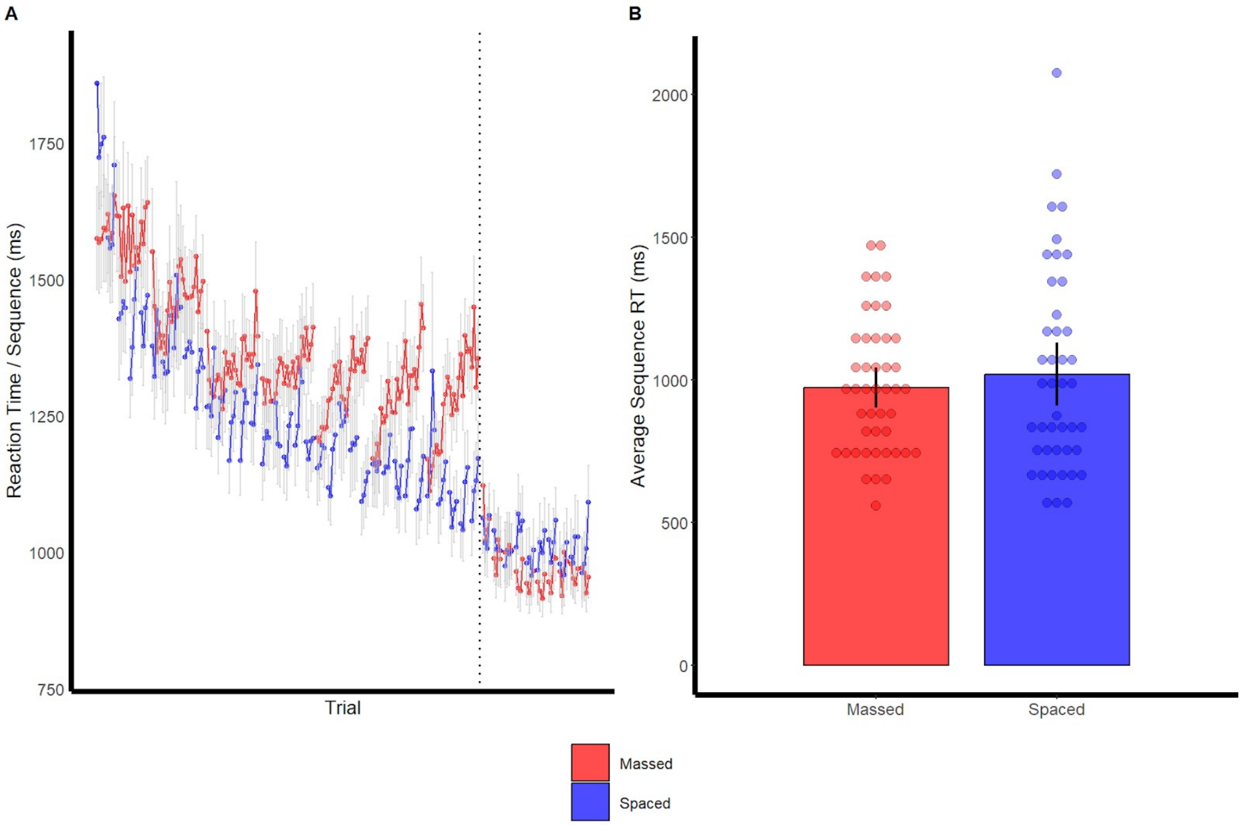


Figure 3.2 a) Each dot represents the average RT of one correctly completed sequence. The dots connected by lines indicate they are part of the same trial. The spaced group had more first sequence warm-up trials removed due to there being more trials in that group. Thus, fewer sequences are shown for that group. The gray error lines are standard errors. b) The bar plot shows the final test mean RTs averaged over all test trials with the dots representing an individual participant average. Error bars are 95% confidence intervals.

Experiment 1 Discussion

There was a significant interaction between group and trial number in the training phase, indicating greater build-up of residual RI over trials in the massed group. After a 15 minute rest period, however, the groups performed equivalently on the same spaced task. Learning during training thus appears to have been equivalent in the two groups.

Although Experiment 1 was relatively highly powered compared to other studies in the literature using the same or similar tasks, it is still possible that the power was not sufficient to detect a difference on the test. Further, the five-item motor sequence is simple and relatively easy to learn. In the second experiment we more than doubled the number of participants per group to increase power and used a nine-item sequence in a bid to further increase both trial-level and across-trial RI differences between the two groups.

Experiment 2

Methods

Participants

We recruited 196 right-handed participants, 98 in the massed group (age = 20.5, F = 73.4%) and 98 in the spaced group (age = 21.1, F = 76.5%). Participants provided informed consent via button press. All procedures were approved by the institutional review board of the University of California, San Diego.

Experimental design and procedure

Participants performed the same task as in Experiment 1, however, the sequence length was increased to nine items, 4-1-3-2-4-2-3-1-4 (we reversed the first four items of the original sequence and then appended it to the end). Participants in the massed group completed 12 correct sequences per training trial with 10 s breaks between trials. The spaced group completed 4 correct sequences per training trial with 30 s breaks. After 108 completed sequences during training, there was a 15 minute rest wherein participants both groups performed a distraction task of double-digit addition. Before the test trials, participants completed a single warmup sequence and then had a break. Afterwards, both groups performed 15 test trials with 4 sequences per trial and 30 s breaks. All other aspects of the methods were identical to those of Experiment 1.

Results

Errors

The error rate was calculated for each participant as the number of incorrect key presses prior to each (correct) sequence within each trial. Averaging over sequences, trials, and participants in the training phase, the error rate was 1.37, and 1.89, key presses in the spaced and massed groups, respectively. In a two-sample t-test on the error rate, there was a significant effect between groups, $t(194) = 2.18$, $p = .03$, $d = 0.22$, $BF_{10} = 1.42$. On the test, the error rate was 1.3 and 0.88 key presses in the spaced and massed groups, respectively. A two-sample t-test on the error rate revealed a non-significant effect between groups, $t(194) = -1.38$, $p = .17$, $d = -0.14$, $BF_{01} = 2.65$.b

Trial-level RI

A paired-samples t-tests equivalent to those of Experiment 1 yielded evidence of trial-level RI in both the massed group, $t(97) = -7.52$, $p < .0001$, $d = -0.75$, $BF_{10} = 3e^{8}$, and the space group, $t(97) = -10.33$, $p < .0001$, $d = -1.04$, $BF_{10} = 7e^{14}$.

Training RTs

To confirm the performance differences at the end of training between the massed and spaced groups, we performed a two-sample t-test on the mean RT for the last 24 sequences, yielding $t(194) = 6.15$, $p < .0001$, $d = 0.62$, $BF_{10} = 2.27e^{6}$ (Figure 3.3a). We also conducted a mixed-factors ANOVA using all of the training trials, as in Experiment 1. For the spaced group, trials were recoded to match the massed group such that trials one, two, three, four and five became trial one, etc. There was a significant effect of trial, $F(9, 1649) = 136.81$, $p < .0001$, $np^2 = 0.43$, $BF_{10} = 7.12e^{151}$, a significant effect of group, $F(1, 193) = 24.38$, $p < .0001$, $np^2 = 0.11$, $BF_{10} = 8.59$, and a significant effect of the interaction term between trial and group, $F(9, 167) = 58.09$, $p < .0001$, $np^2 = 0.22$, $BF_{10} = 1.67e^{79}$. Hence there was compelling evidence of greater accrual of residual RI across trials in the massed group.

Post-rest Improvement

Confirming the interaction apparent in Figure 3.3, a two-sample t-test on the participant-level mean RT difference scores (mean RT on the last 24 training sequences minus first 24 test sequences) for the massed vs. spaced groups was significant, $t(194) = 8.99$, $p < .0001$, $d = 0.91$. The Bayes Factor test positively favored the alternative, $BF_{10} = 2.14e^{13}$.

Final Test RTs

A two-sample t-test on the group means of all RT test sequences (fig. 1b) did not approach significance, $t(194) = 1.25$, $p = .21$, $d = .13$. The Bayes Factor test again favored the null, $BF_{01} = 3.1$ (Figure 3.3b). The results are comparable when only performing the test on the average of the first 24 test sequences.

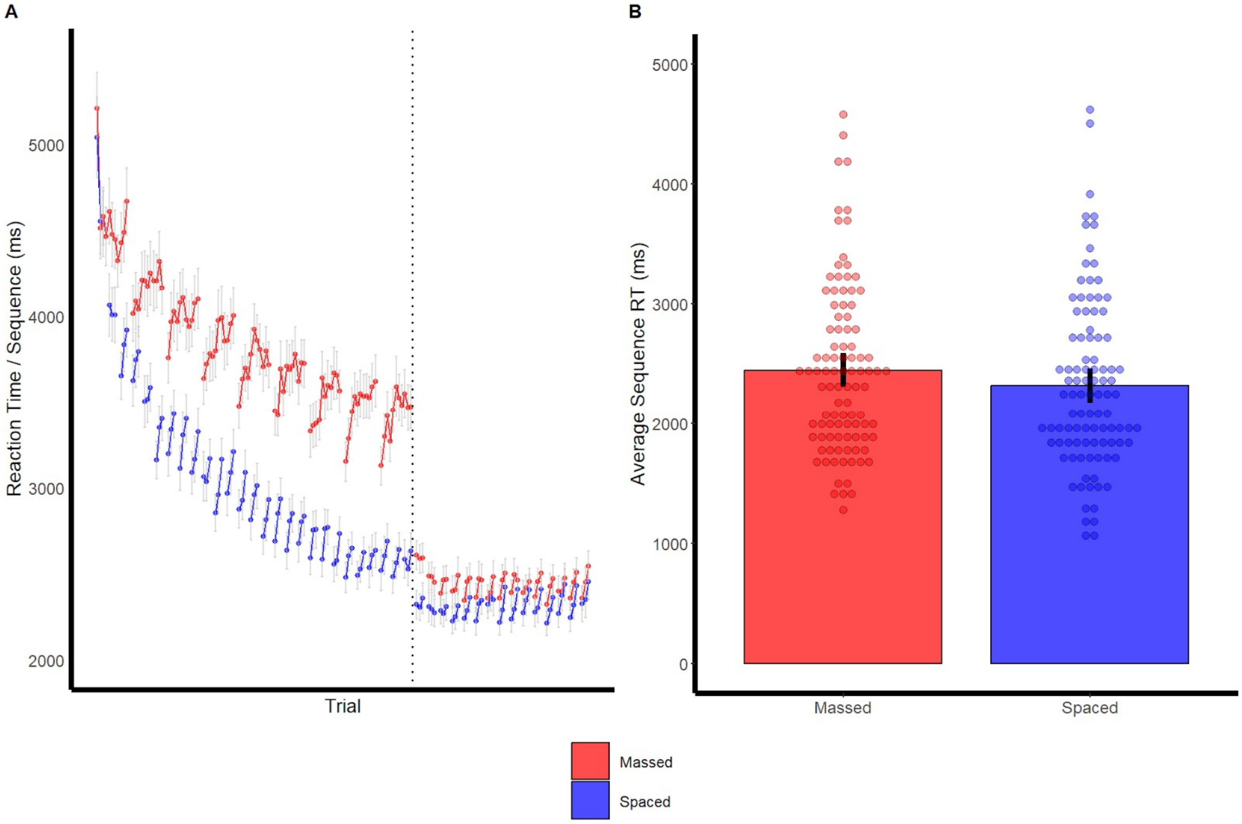


Figure 3.3 a) Each dot represents the average RT of one correctly completed sequence. The dots connected by lines indicate they are part of the same trial. The spaced group had more first sequence warm-up trials removed due to there being more trials in that group. Thus, fewer sequences are shown for that group. The gray error lines are standard error b) The bar plot shows the final test mean RTs averaged over all test trials with the dots representing an individual participant average. Error bars are 95% confidence intervals.

General Discussion

We investigated whether the severity of RI or the time available for offline micro-consolidation moderates motor sequence learning. In both experiments during training, a larger amount of RI was induced in the massed group by limiting the number of breaks and the break duration, and by requiring participants to complete a relatively large number of correct sequences per trial. After a 15 minute rest to dissipate RI, both groups performed a test on the same task (the spaced task from the training phase). In both experiments, test performance in the two groups was statistically equivalent, suggesting no effect of RI on learning and no effect of the large group differences in break duration on the amount of offline micro-consolidation. In Experiment 2 there was a non-significant trend toward better final test performance in the spaced group (although the reverse was true for Experiment 1), and thus we cannot rule-out a small difference in learning for the two groups in that experiment. It is possible that the 15 minute rest period was insufficient to completely dissipate RI in the massed group of Experient 2, perhaps due to the 9 key presses required per sequence. That possibility seems unlikely, however, given that on average the training session lasted around seven minutes, not including breaks, with the 15 minute rest period more than doubling that duration. In any case, any differences in achieved learning between the two groups in Experiment 2 is at best minimal. The absence of a spacing effect in the current experiments was surprising in light of the abundant evidence that learning during spaced training is far superior to that during massed training in not only declarative memory tasks, but also for at least some types of cognitive skill learning[18].

The conclusions in the preceding paragraph are drawn from the correct sequence RT results. There were also weak statistical differences in error rate for the two groups. In Experiment 1, the error rate was statistically equivalent during training, whereas the spaced group had a statistically higher error rate on the test. In Experiment 2, there were statistically higher error rates during training for the massed group, but no statistical differences in error rate on the test. Both of those significant effects are weak as measured by Bayes Factors, and they are inconsistent across experiments. Previously, Gupta & Rickard (in press) speculated that there may be a speed-accuracy tradeoff where the quicker spaced group has more errors than the slower massed group. However, those results do not appear to fully replicate in this study. Hence, across several experiments there is no consistent pattern of error differences between groups. Regardless, the diverging RT pattern across training trials for the two groups in each experiment appears to be explainable only by pronounced group differences in the build-up of cumulative RI.

The simplest account of our results is that the same amount of training phase learning occurred in both groups of the two experiments and that the performance differences between spaced and massed groups during training solely reflects differences in the build-up of RI. Both the online and offline (i.e., micro-consolidation) learning hypotheses (or a hybrid of them) are potentially consistent with that account. In the online learning account advanced by Gupta and Rickard (2022; in press), learning occurs immediately during each executed sequence (and not during breaks) and is independent of the schedule of trials and brief breaks. Because the same number of sequences were performed in both the massed and spaced groups, the online account straightforwardly predicts equivalent learning by the end of training and hence equivalent performance on the test.

Our results appear to place some boundary conditions, however, on the offline account. First, given the results, the amount of offline learning that occurred during the training breaks must have been distributed differently for the massed and spaced groups. In Experiment 2, for example, more offline learning must have occurred during each of the eight 10 s breaks (plus the rest period) of the massed group than during each of twenty-seven 30 s breaks (plus the rest period) of the spaced group. Second, either (a) a 10 s break in the massed groups was sufficient to fully exhaust all possible offline learning potential after performing both 25 sequences of a five-item sequence (Experiment 1) and 12 sequences of a nine-item sequence trial (Experiment 2) or (b) if a 10 s break is insufficient to exhaust that learning potential, then the unrealized learning potential must have accumulated over training trials without loss, such that it came to full fruition during the 15 minute rest period. The first and simpler possibility is consistent with the finding by Buch et al. (2021) that hippocampal activity associated with micro-consolidation may run to completion within about 2.5 s after the end of a performance trial.

Finally, our results raise the possibility that massed training and the associated RI is not detrimental to naturalistic motor sequence learning – such as training on a musical instrument or sport – even when worsening of performance across training trials is palpable. However, there may well be factors that limit the generalizability of that conclusion. In the current tasks, for example, the error rate was not systematically higher in the massed groups. Longer trial durations, particularly if accompanied by increased error rate, may hamper learning. That possibility hints at an optimization scenario wherein the practical advantage of fewer, relatively long training trials is balanced against the possibility of impaired learning. That pattern would be consistent with the concept of desirable difficulty[5] and is consistent with recent findings for a motor skill game[1].

Data availability

All data and code (stimuli and analyses are available online (https://osf.io/ukwf9/). Further information and requests for resources should be directed to and will be fulfilled by the corresponding author, TCR (trickard@ucsd.edu).

References

1. Al-Fawakhiri, N., Kayani, S., & McDougle, S. D. (2023). Evidence of an optimal error rate for motor skill learning. bioRxiv, 2023-07.

2. Archer, E. J. Postrest performance in motor learning as a function of prerest degree of distribution of practice. *Journal of Experimental Psychology*, 47(1), 47–51 (1954).

3. Ammons, R. B. Acquisition of motor skill: I. Quantitative analysis and theoretical formulation. *Psychol. Rev.*, 54, 263-281 (1947).

4. Bächinger, M., Lehner, R., Thomas, F., Hanimann, S., Balsters, J., & Wenderoth, N. Human motor fatigability as evoked by repetitive movements results from a gradual breakdown of surround inhibition. *ELife*, 8, 1–30 (2019).

5. Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. Psychology and the real world: Essays illustrating fundamental contributions to society, 2(59-68).

6. Bönstrup, M., Iturrate, I., Thompson, R., Cruciani, G., Censor, N., & Cohen, L. G. A Rapid Form of Offline Consolidation in Skill Learning. *Current Biology*, *29*(8), 1346-1351 (2019).

7. Bönstrup, M., Iturrate, I., Hebart, M. N., Censor, N. & Cohen, L. G. Mechanisms of offline motor learning at a microscale of seconds in large-scale crowdsourced data. NOJ Sci. Learn. 5, 1–10 (2020).

8. Buch, E. R., Claudino, L., Quentin, R., Bönstrup, M., & Cohen, L. G. Consolidation of human skill linked to waking hippocampo-neocortical replay. *Cell Reports*, 35(10) (2021).

9. Carron, A. V. Physical fatigue and motor learning. *Research Quarterly of the American Association for Health, Physical Education and Recreation*, 40(4), 682–686 (1969a).

10. Carron, A. V. Performance and learning in a discrete motor task under massed vs. distributed practice. *Research Quarterly of the American Association for Health, Physical Education and Recreation*, 40(3), 481–489 (1969b).

11. Jacobacci, F., Armony, J.L., Yeffal, A., Lerner, G., Amaro Jr, E., Jovicich, J., Doyon, J. and Della-Maggiore, V.. Rapid hippocampal plasticity supports motor sequence learning. *Proceedings of the National Academy of Sciences*, *117*(38), pp.23898-23903 (2020).

12. Gupta, M.W., Rickard, T.C. Dissipation of reactive inhibition is sufficient to explain post-rest improvements in motor sequence learning. *npj Sci. Learn.* **7**, 25 (2022).

13. Gupta, M.W. and Rickard, T.C. A Quantitative Model that Incorporates Reactive Inhibition but No Micro-Consolidation Explains Motor Sequence Practice Effects. (*Scientific Reports, In Press*).

14. Kimble, G. A. An experimental test of a two-factor theory of inhibition. Journal of Experimental Psychology, 39(1), 15 (1949).

15. Lee, T. D., & Genovese, E. D. Distribution of practice in motor skill acquisition: Learning and performance effects reconsidered. *Research Quarterly for exercise and Sport*, 59(4), 277-287 (1988).

16. Nissen, M. J. & Bullemer, P. Attentional requirements of learning: evidence from performance measures. Cogn. Psychol. 19, 1–32 (1987).

17. Raftery, A. E. . Bayesian model selection in social research. In P. V. Marsden (Ed.), Sociological methodology, 111–196 (1995).

18. Rickard, T. C., Lau, J. S. H., & Pashler, H. (2008). Spacing and the transition from calculation to retrieval. Psychonomic Bulletin & Review, 15(3), 656-661.

19. Wagenmakers, EJ., Love, J., Marsman, M. et al. Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review* 25, 58–76 (2018).

Acknowledgements

Chapter 3, in part is currently in submission for publication of the material. Gupta, Mohan W.; Rickard, Timothy C. The dissertation author was the primary researcher and author of this material.

CONCLUSION

In conclusion, my dissertation endeavored to challenge the predominant view of the temporal dynamics of motor skill learning. I proposed an alternative framework wherein learning occurs concurrently with practice, rather than exclusively during rest periods. By integrating empirical experimentation and computational modeling, my dissertation has shed light on the intricate interplay between online and offline learning processes, while elucidating the role of reactive inhibition in shaping motor skill acquisition trajectories.

My research findings have provided compelling evidence in support of the proposed online learning framework, demonstrating its capacity to account for performance enhancements observed following rest periods. Moreover, through developing a computational framework, I have delineated the contributions of online and offline learning, culminating in the identification of a hybrid model that best fits the complex dynamics of motor skill learning. Furthermore, my examination of the influence of training schedules on learning rates yielded surprising results. Contrary to expectations, we found no evidence to support schedule-dependent learning effects in tasks characterized by minimal error rates, highlighting the need for further exploration in this area.

By challenging the predominant view that motor skill learning exclusively occurs offline, my dissertation has contributed to a deeper understanding of the temporal dynamics underlying motor skill acquisition processes. It is imperative that future research continues to explore the possibility that both online and offline learning occur, while considering the role of reactive inhibition. Ultimately, the insights gained from my dissertation not only advance theoretical understanding but also hold practical implications for skill acquisition and motor learning interventions. By embracing a more nuanced perspective that acknowledges the concurrent nature of learning and practice, we can develop more effective strategies for enhancing skill acquisition and performance across a range of domains.