The influence of blocked versus interleaved training regimes and sleep on multi-task learning

Mina Habibi[a], Mehdi Senoussi[a,b], Pieter Verbeke[a,c], & Senne Braem[a]
[a]Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, 9000 Ghent, Belgium
[b]CLLE Lab, CNRS UMR 5263, Université de Toulouse, Toulouse, France
[c]AI Lab, Howest University of Applied Sciences, Kortrijk, Belgium

Abstract

Recent studies suggest that humans benefit from blocked training in continual learning by promoting more factorized task representations. We investigated whether interleaved training may support human continual learning more when task separation is also aided by consolidation during sleep. Participants learned three tasks under blocked versus interleaved regimes across two experiments: Experiment 1 (contextual cues presented before stimuli; semantically non-informative labels) and Experiment 2 (stimuli presented before contextual cues; semantically informative labels). Testing occurred both immediately and after 24 hours. People trained under the blocked training regime showed higher accuracy during learning, but this advantage did not persist in the test phase of Experiment 1. In Experiment 2, blocked training resulted in higher accuracy across learning and testing. Critically, no sleep related benefits were found in either training regime for both experiments. RNNs fit to human data, however, revealed increased task separation from Day 1 to Day 2 in Experiment 2 across both training regimes. Our findings suggest that humans can benefit from both training regimes, and that the order and way in which context and stimulus are presented—rather than the regime alone—may play an important role in continual learning.

Keywords: Blocked Regime, Interleaved Regime, Continual Learning, Task Representation

Introduction

Continual learning presents a significant challenge for artificial systems, while humans typically navigate this process more efficiently. A key aspect of this challenge is balancing stability and plasticity: too much plasticity leads to interference or "catastrophic forgetting," while excessive stability impairs the ability to learn new information (Abraham & Robins, 2005). One approach to mitigating interference is reducing the representational overlap between tasks (Flesch et al., 2018, 2022; Musslick & Cohen, 2021). Recent studies suggest that blocked training supports this process in humans by promoting more orthogonal task representations, thereby reducing interference (Flesch et al., 2018). On the other hand, other studies have shown that interleaved training enhances learning, retention, and transfer, especially over delays or following sleep, potentially via improved consolidation and pattern separation (Carvalho & Goldstone, 2015; Kim et al., 2024; Diekelmann & Born, 2010; Magill & Hall, 1990).

In this study, we investigated whether sleep-based consolidation enhances the benefits of interleaved training for human continual learning. Additionally, we explored the often-overlooked role of semantic labeling and presentation order — the sequence in which context and stimulus appear — and how it influences learning. While previous studies predominantly employed context-first designs (where the task cue precedes the stimulus), this structure may naturally support task separation by providing a temporal "gating" signal for task representations (Masse et al., 2018; Verbeke, & Verguts, 2022). However, in real-world contexts, stimuli also appear before their associated context (e.g., seeing ingredients before deciding what to cook), suggesting that stimulus-first scenarios may engage different encoding processes, such as mapping the same stimulus to multiple potential action plans.

To test this, we contrasted the effects of interleaved versus blocked training and investigated the effect of sleep in two experiments. Experiment 1 followed the conventional context-first structure, while Experiment 2 reversed the order, presenting the stimulus before the context, using semantically meaningful task cues (see Methods). This design allowed us to disentangle the effects of presentation order and training regime on learning and retention in continual learning.

Methods

Participants:

Participants included 89 participants in Experiment 1 (blocked: 46; interleaved: 43) and 103 in Experiment 2 (blocked: 50; interleaved: 53).

Stimuli and Design:

In both experiments, participants learned to categorize insects based on three binary features—legs (thick/thin), antennae (thick/thin), and mandible (shovel/pincer)—across three tasks. Participants were randomly assigned to a blocked (tasks presented in separate blocks) or interleaved (tasks randomly mixed within blocks) training regime, followed by two interleaved test phases without feedback (Figure 1A): one immediately after training (Day 1) and one after 24 hours (Day 2). Each participant completed 288 training and 288 test trials. In Experiment 1 (Figure 1B), the task context was indicated by the color of a planet (pink, blue, or brown), presented before the insect image. Participants were asked to decide whether each insect could survive on a given planet. In Experiment 2 (Figure 1C), the insect was presented before the contextual cue, which consisted of semantic response labels (e.g., "Cold vs. Warm," "Urban vs. Rural," "High vs. Low Altitude"). The planet image remained constant and participants were asked to determine which environment best suited each insect.

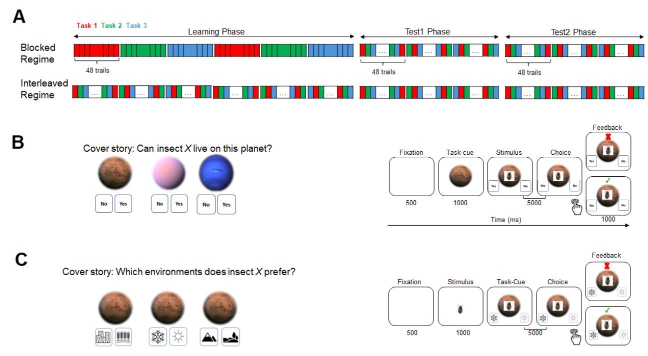

Figure 1: A) Experimental design and timeline for Exp1 and Exp2: B) Exp1 context-frist and C) Exp2 stimulus-first trial structure.

Data Analysis and Modeling:

We averaged accuracies across three-block segments, comparing late training (Blocks 4–6) to immediate testing (Test Day 1: Blocks 7–9) to assess learning, and Test Day 1 vs. Test Day 2 (Blocks 10–12) for consolidation. To explore task representations, we trained Simple Recurrent Neural Networks (RNNs) on Human data from Test Days 1 and 2. The network had stimulus/context inputs, a hidden layer (48 units), and a binary output. We measured Euclidean distances between hidden-layer activations across tasks (A vs. B, B vs. C, A vs. C) to assess representational separation. Larger distances indicate reduced overlap. Networks were trained per participant using categorical cross-entropy loss, Adam optimization, and early stopping.

Results:

Experiment 1 (Context-First): Blocked training led to higher accuracy (M = 0.80 vs. M = 0.67), $t(87) = 3.89$, $p < .001$. However, this advantage disappeared at Test Day 1, where accuracy was similar across groups (Blocked: M = 0.75, Interleaved: M = 0.73), $p = .62$. A significant interaction between training regime and time, $F(1,87) = 23.74$, $p < .001$, showed a drop in blocked performance ($p = .008$) and an improvement in interleaved ($p < .001$) (Figure 2A). Accuracy remained stable across Test Days 1 and 2, suggesting no sleep-related improvements. RNN modeling revealed no effects of training regime or time on representational distance, indicating no increased task separation after sleep (Figure 2B–C).

Experiment 2 (Stimulus-First): Blocked training resulted in higher accuracy across both learning and Test Day 1, $F(1,101) = 14.05$, $p < .001$. After sleep, accuracy declined ($p = .015$), but the blocked group maintained a higher accuracy compared to the interleaved group ($p = .009$), with no interaction between factors ($p = .26$) (Figure 3A). RNN modeling revealed increased representational distance over time ($p = .015$), but no differences between training regimes. These

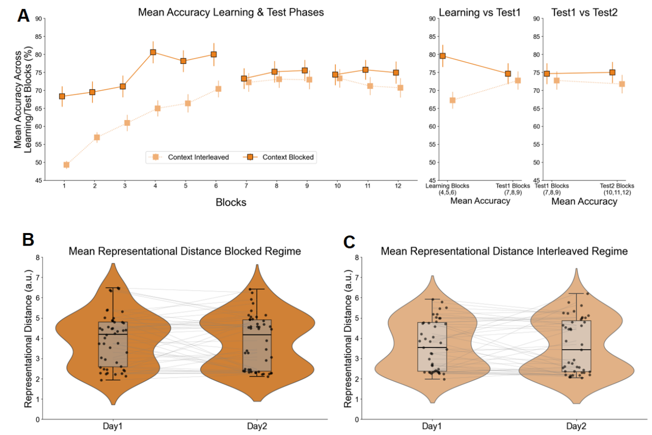

Figure 2: Context-first results.

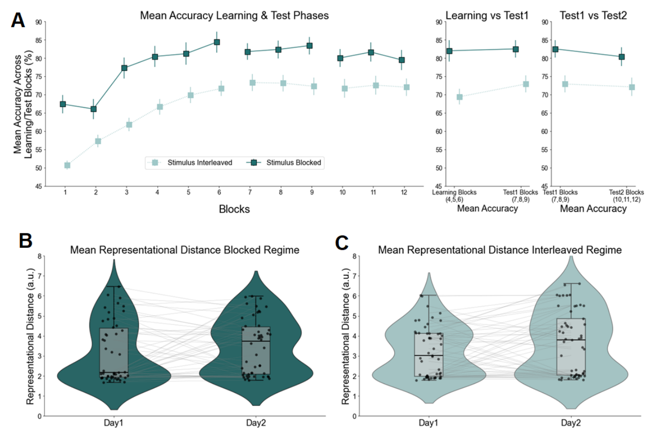

Figure 3: Stimulus-first results.

findings suggest that sleep increased task separation, irrespective of training regime (Figure 3B–C).

Discussion and conclusion

Across both experiments, our results suggest that blocked training alone does not always guarantee improved continual learning, and sleep consolidation does not always improve task separation. In Experiment 1, the blocked regime led to an advantage during training that disappeared during testing. In Experiment 2, a sustained benefit was observed, but this effect may be attributable to the order of stimulus-context presentation or semantic labeling, rather than an inherent advantage of a blocked training regime. Sleep was expected to enhance performance and task separation following interleaved training. However, there was no improvement in task separation or accuracy post-sleep. Our results show that humans can benefit from both training regimes for continual learning, challenging the assumption that blocked training alone provides a beneficial effect on continual learning and suggesting a more complex relationship between learning structure, sleep, and performance.

References

Abraham, W. C., & Robins, A. (2005). Memory retention–the synaptic stability versus plasticity dilemma. *Trends in neurosciences*, *28*(2), 73-78.

Carvalho, P. F., & Goldstone, R. L. (2015). The benefits of interleaved and blocked study: Different tasks benefit from different schedules of study. *Psychonomic bulletin & review*, *22*, 281-288.

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature reviews neuroscience*, *11*(2), 114-126.

Flesch, T., Balaguer, J., Dekker, R., Nili, H., & Summerfield, C. (2018). Comparing continual tasklearning in minds and machines. *Proceedings of the National Academy of Sciences*, *115*(44), E10313-E10322.

Flesch, T., Juechems, K., Dumbalska, T., Saxe, A., & Summerfield, C. (2022). Orthogonal representations for robust context-dependent task performance in brains and neural networks. *Neuron*, *110*(7), 1258-1270.

Kim, T., Kim, H., Philip, B. A., & Wright, D. L. (2024). M1 recruitment during interleaved practice is important for encoding, not just consolidation, of skill memory. *npj Science of Learning*, *9*(1), 77.

Magill, R. A., & Hall, K. G. (1990). A review of the contextual interference effect in motor skill acquisition. *Human movement science*, *9*(3-5), 241-289.

Masse, N. Y., Grant, G. D., & Freedman, D. J. (2018). Alleviating catastrophic forgetting using context-dependent gating and synaptic stabilization. *Proceedings of the National Academy of Sciences*, *115*(44), E10467-E10475.

Musslick, S., & Cohen, J. D. (2021). Rationalizing constraints on the capacity for cognitive control. *Trends in cognitive sciences*, *25*(9), 757-775.

Verbeke, P., & Verguts, T. (2022). Using top-down modulation to optimally balance shared versus separated task representations. *Neural networks*, *146*, 256-271.