A resource-rational model of self-paced learning

Si Ma[1], Qiong Zhang[1,2]

[1] Department of Psychology, [2] Department of Computer Science, Rutgers University

si.ma@rutgers.edu, qiong.z@rutgers.edu

Abstract

Self-paced learning, where individuals freely decide how much time to spend encoding each item, is often associated with better long-term memory retention and is further regulated by time pressure. Multiple competing verbal theories exist, yet a computationally explicit explanation is lacking. We hypothesize that the relationship between self-paced learning and memory emerges from people's optimizing task performance under limited cognitive resources during encoding. Without a pre-determined presentation schedule, self-paced participants can better allocate and recover cognitive resources according to both endogenous and exogenous factors, leading to better learning. These hypotheses were validated via a metacognitive model, where an object-level process uses allocated resources to strengthen memory encoding, and a meta-level process controls how to optimize total recall by adaptively allocating and recovering resources. Our model predictions align with empirical findings and reconcile previous verbal accounts of self-paced learning.

Introduction

Self-paced learning, which allows individuals to decide how much time to spend encoding each item, is associated with better memory retention (De Jonge et al., 2015; Markant et al., 2014; Tullis & Benjamin, 2011). According to the discrepancy-reduction theory, individuals possess an internal criterion for when an item is considered learned, and self-pacing allows selectively allocating more study time to items furthest from their learning criterion (De Jonge et al., 2015; Dunlosky & Thiede, 1998; Tullis & Benjamin, 2011). While studies predominantly support this theory as people allocate more study time to difficult items, exceptions exist where easy items are favored under increased time pressure (Son & Metcalfe, 2000). To account for this, Metcalfe (2002) proposed the Region of Proximal Learning theory, suggesting that people should prioritize studying information that is just beyond their current state of knowledge. Yet, a third account attributes self-paced advantage to one's ability to coordinate stimulus presentation with moment-to-moment attentional state (Markant et al., 2014).

To reconcile these distinct accounts of self-paced learning, following the approach of resource-rational analysis (Lieder & Griffiths, 2020), we build a computational model that considers the behavioral patterns of self-paced learning jointly shaped by the computational goal of the memory task and the architectural and resource constraints during memory encoding. Our model posits that when there is an additional control to decide study time, people can adaptively coordinate the allocation and recovery of limited cognitive resources during memory encoding to maximize overall memory performance. The architectural and resource constraints have been independently characterized in previous studies (Ma et al., 2024, 2026; Popov & Reder, 2020; Reder et al., 2007), which are fixed in the current work, under which an optimal policy is derived to understand adaptive behavior with and without self-pacing. In the remainder of this paper, we demonstrate that our model explains the empirical results of self-paced learning, as well as unifying seemingly distinct prior verbal theories.

Methods

Our proposed model captures the encoding processes of a list-learning paradigm, where participants first study a list of items and then recall items from the list in any order (Murdock Jr, 1962). The model consists of an object-level component and a meta-level component. The object-level component incorporates information about the cognitive constraints during memory encoding. Specifically, cognitive resources are at their maximum (W_{max}) at the beginning of a study list. When encoding an item, a proportion of the currently available resources is allocated for semantic processing (τ), and a proportion for episodic encoding (δ), and the depleted resources require time to recover at a rate of r. The extent of episodic encoding determines recall probability (θ_{epi}, σ_{epi}). We treat the object-level component as a fixed property, and inherit the same model formulation and parameters describing cognitive constraints (τ, W_{max}, r, σ_{epi}) from a previous study (Ma et al., 2024). The meta-level component monitors the state of memory encoding (i.e., the available resources, item position, item difficulty, cumulative time spent) and adaptively determines resource allocation (δ) and recovery time (t) for each item to maximize performance. We used the Proximal Policy Optimization (PPO) algorithm (Schulman et al., 2017) to find the optimal behavior of the model.

We simulated a self-paced group and a fixed group, each comprising 25 simulated participants who com-

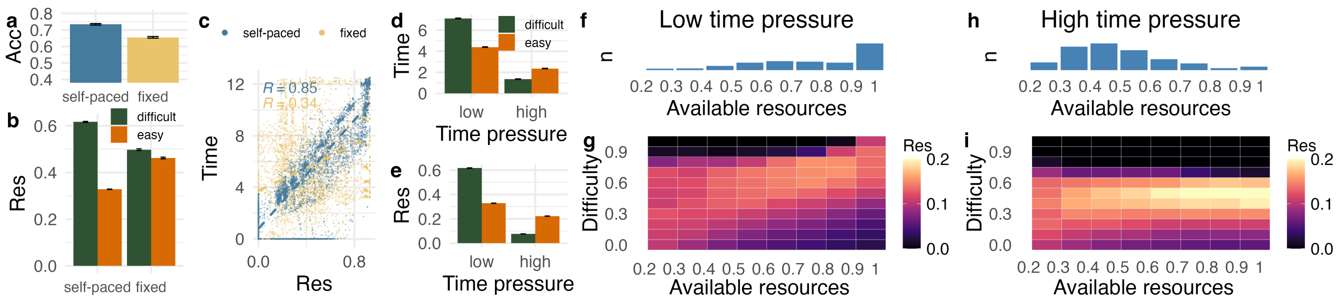

Fig 1: A resource-rational account of self-paced advantage and study time allocation. Compared with the fixed group, the self-paced group achieves higher overall recall (a), better differentiates item difficulty during resource allocation (b), and demonstrates a stronger correlation between allocated resources and allocated time (c). Under increased time pressure, the self-paced group shifts to allocating more time (d) and resources (e) to easy items. The moment-to-moment resource allocation policy reveals that the model is frequently in states with high available resources under low time pressure (f), where it is optimal to allocate resources to difficult items (g), and in states with low available resources under high time pressure (h), where it is optimal to prioritize easier items (i).

pleted 20 lists each. Each list contained 11 items with item difficulty ranging from 0 to 1. The self-paced agent optimized both the allocation of resources (δ) and the allocation of time (t), while the fixed agent – which viewed the same items under the same study times matched to their paired self-paced agent – optimized the allocation of resources (δ) only. To capture individual variability in the subjective perception of item difficulty, item difficulties were shuffled for the fixed condition. The total time limit is used to simulate low versus high time pressure.

Results

We show that our model can account for memory performance and study-time allocation patterns observed in the literature. These patterns were derived solely from optimal adaptation rather than being fit to the empirical data that the model seeks to explain.

The model predicted higher recall performance in the self-paced group compared to the fixed group (Fig 1a), aligning with empirical findings. Although both groups can freely decide their resource allocation, the self-paced group more effectively differentiates difficult versus easy items, allocating more resources to the difficult items that are further from the learning criterion (Fig 1b), showing a significant interaction between condition and item difficulty in a linear mixed-effects model ($b = 0.25$, $SE = 0.01$, $t(9,996) = 24.39$, $p < .001$). In this regard, our model behavior is consistent with the discrepancy-reduction theory (Dunlosky & Thiede, 1998). Despite their similarity, our model provides additional insight into why the self-paced group can afford to allocate more resources to difficult items. With the additional control over allocated time, the self-paced group can better coordinate their resource recovery time following resource allocation to each encoded item (Fig 1c). Consistent with this explanation, the positive relationship between allocated resources and allocated time was significantly stronger in the self-paced condition ($b = -3.88, SE = 0.12, t(22,000) = -32.00, p < .001$). Our model results also connect to the account of coordinating one's preparatory states (Markant et al., 2014), by specifying exactly what a good state is, i.e., a state with well-recovered cognitive resources.

Study time allocation depends on time pressure. Under low time pressure, the model focuses resources on items that are further from the learning criterion (discrepancy-reduction theory; Dunlosky and Thiede, 1998). However, under high time pressure, it is more adaptive to prioritize easy items, which gives the most learning gains (the region of proximal learning theory; Metcalfe, 2002). Reconciling prior theories, our model successfully captures this behavioral shift when the learning environment changes. It devoted more study time (Fig 1d) and resources (Fig 1e) to difficult items under low time pressure, but prioritized easier items under high time pressure.

Patterns of time allocation are explained by moment-to-moment resource allocation. We examined how resource allocation to each item depends on the item's difficulty and current resource availability (Fig 1g, Fig 1i). Under low time pressure, more resources are allocated to difficult items as the available resources increase (Fig 1g). Since the model is more frequently in states with a high amount of available resources (Fig 1f), this leads to greater overall resource allocation to difficult items. In contrast, under high time pressure, the model is more frequently in states with low available resources (Fig 1h), where the model allocates more resources to moderately difficult items (Fig 1i). This supports our hypothesis that self-paced participants adaptively allocate and recover limited cognitive resources during memory encoding.

Disclosure

Generative AI tools are used to assist with grammar correction, typo detection, and language polishing.

References

De Jonge, M., Tabbers, H. K., Pecher, D., Jang, Y., & Zeelenberg, R. (2015). The efficacy of self-paced study in multitrial learning. *Journal of experimental psychology: Learning, memory, and cognition*, *41*(3), 851.

Dunlosky, J., & Thiede, K. W. (1998). What makes people study more? an evaluation of factors that affect self-paced study. *Acta psychologica*, *98*(1), 37–56.

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and brain sciences*, *43*, e1.

Ma, S., Popov, V., & Zhang, Q. (2024). A neural index reflecting the amount of cognitive resources available during memory encoding: A model-based approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Ma, S., Popov, V., & Zhang, Q. (2026). Cognitive constraints and reward environments jointly shape memory formation. *Psychological Review*.

Markant, D., DuBrow, S., Davachi, L., & Gureckis, T. M. (2014). Deconstructing the effect of self-directed study on episodic memory. *Memory & cognition*, *42*(8), 1211–1224.

Metcalfe, J. (2002). Is study time allocated selectively to a region of proximal learning? *Journal of Experimental Psychology: General*, *131*(3), 349.

Murdock Jr, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*(5), 482.

Popov, V., & Reder, L. M. (2020). Frequency effects on memory: A resource-limited theory. *Psychological Review*, *127*(1), 1–46.

Reder, L. M., Paynter, C., Diana, R. A., Ngiam, J., & Dickison, D. (2007). Experience is a double-edged sword: A computational model of the encoding/retrieval trade-off with familiarity. *Psychology of Learning and Motivation*, *48*, 271–312.

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.

Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(1), 204.

Tullis, J. G., & Benjamin, A. S. (2011). On the effectiveness of self-paced learning. *Journal of memory and language*, *64*(2), 109–118.