A Multiple Trace Theory of Statistical Learning

Dock H. Duncan, Jan Theeuwes, Sander A. Los

Abstract

The attentional effects of statistical learning and intertrial priming have long been treated as separate selection history effects. This is despite their many similarities, improving performance in the same direction. There is thus motivation to either formally dissociate these two effects, or unify them under a single theoretical framework. We suggest that multiple trace theory is the framework to unify these two cognitive effects. We used a Kalman filter approach to model reaction times while participants performed the additional singleton task with biased distractor presentations - a paradigm known to engender both statistical learning and intertrial effects. Initial results suggest this by-trial modelling approach aptly captures learning effects and their effect on reaction times. Subsequent steps will now compare unified versus divided models of intertrial priming and statistical learning to provide evidence for their dissociation or union.

Keywords: Attention, Statistical Learning, Multiple Trace Theory, Instance Theory, Intertrial Priming, Distractor Suppression

Intro

Maljkovic and Nakayama in a pair of highly influential papers (1994, 1996) identified that spatial and color features that repeated in subsequent displays of a visual search task greatly improved the performance of participants. This effect was dubbed intertrial priming. Several years later, it was identified that statistical tendencies of a visual search task could be learned by participants and automatically incorporated into their search behavior, improving performance when these regularities are adhered to and impairing performance when they were violated. This automatic encoding of task regularities was labeled statistical learning.

Already in the early works on statistical learning, painstaking steps were taken to ensure that the observed effect was not one and the same as intertrial priming. This was primarily done by demonstrating that the effect persisted some time after task regularities were removed. In the current research we attempt to demonstrate that such persistence can in fact be accounted for in a single parsimonious system based on the multiple trace accumulation framework (Logan 1988; Los, Kruijne, and Meeter 2014; Salet et al. 2022). We propose that the accumulation of memory traces with an exponential decay function can account for both short term intertrial effects as well as long term and persistent statistical learning effects using a single priority accumulation function.

Aggregate reaction time results

Methods. 80 participants took part in this experiment using the online platform Prolific. Data cleaning and preprocessing was done in line with previous experiments (Duncan, van Moorselaar, and Theeuwes 2024). Participants performed a version of the additional singleton task with an imbalanced spatial distribution of distractors, a paradigm known to result in statistically learned suppression (Wang and Theeuwes 2018). Participants completed seven blocks of the experimental task. The first three blocks contained the high-probability distractor location which held a distractor on 50% of distractor present trials (HP locations counterbalanced across participants). Participants next moved on to two transfer blocks, where the regularity was removed, and distractors now appeared at each location with equal frequency. Following two of these transfer blocks, the regularity was reintroduced for a final two reacquisition blocks.

Aggregate Behavioral Results. Grand averaged reaction times showed that participants reliably learned to suppress the high-probability distractor location, responding significantly faster when a distractor was presented there as opposed to another low-probability location (Figure 1).

Learning, Extinction and Reacquisition. When separating trials into learning, transfer and reacquisition blocks, we observed that learning was only present during the learning and reacquisition blocks, and disappeared during the extinction blocks (Figure 2). This indicates that participants were able to flexibly adjust to changes in the statistical regularities underlying the task.

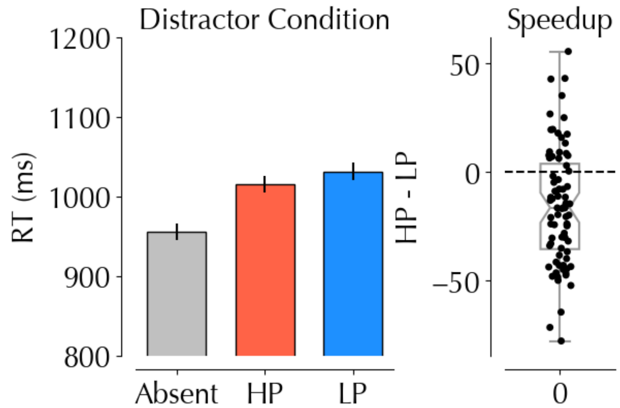

Figure 1: Grand average reaction times collapsed over all blocks in the experiment and separated between trials with no distractor, HP distractors and LP distractors. Shown on the right is the speedup effect calculated by subtracting each participant's average reaction time on LP trials form their average reaction time on HP trials. A negative score thus represents that that the participant was faster when the distractor appeared at the HP location relative to any other LP location. Error bars represent within subject 95% confidence intervals (Cousineau 2005; Morey 2008).

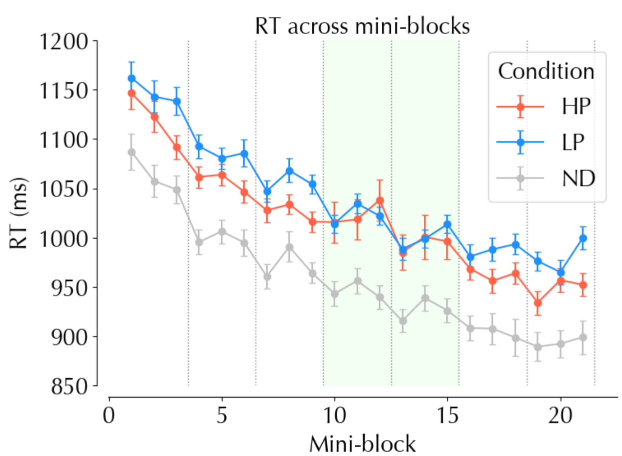

Figure 2: The progression of learning over time illustrated by plotting average reaction times on distractor absent trials, HP distractor trials and LP distractor trials separated among mini blocks of 48 successive trials (1/3 of the 144 trials in an experimental block). Error bars are within subject standard error. Shaded green area represents the period of time in which participants were in the transfer phase, where no distractor regularity was present.

Modelling reaction times using a Multiple Trace Framework

Preliminary results have shown that a Kalman filter approach can capture trial-by-trial fluctuations in reaction times by incorporating previous trial information using long-term memory traces. Further work will now compare model attributes to elucidate whether statistical learning and intertrial effects are more parsimoniously accounted for as a unified system, or as separate systems working in conjunction

Initial Results

We used a Kalman Filter approach to model reaction time data on a trial-by-trial basis while incorporating previous trial information into an iterative internal model of the task statistics. Our model varied three parameters - attentional enhancement per trial, suppression per trial, attention trace decay, attentional spread and baseline reaction time. Initial results provide variable fits to the data per participant. As a next step, we will test whether model fits can predict reacquisition when fit to data from the learning and transfer phases separately. Next, we will observe whether model fits can be improved by providing two separate mechanisms to account for short- and long-term regularities, thereby representing a dissociation between statistical learning and intertrial priming. Lastly, we will compare a ramp versus hill function to model learning, thereby comparing whether behavior better matches a model based on gradual implicit learning, or sudden explicit realization (Musfeld, Souza, and Oberauer 2023).

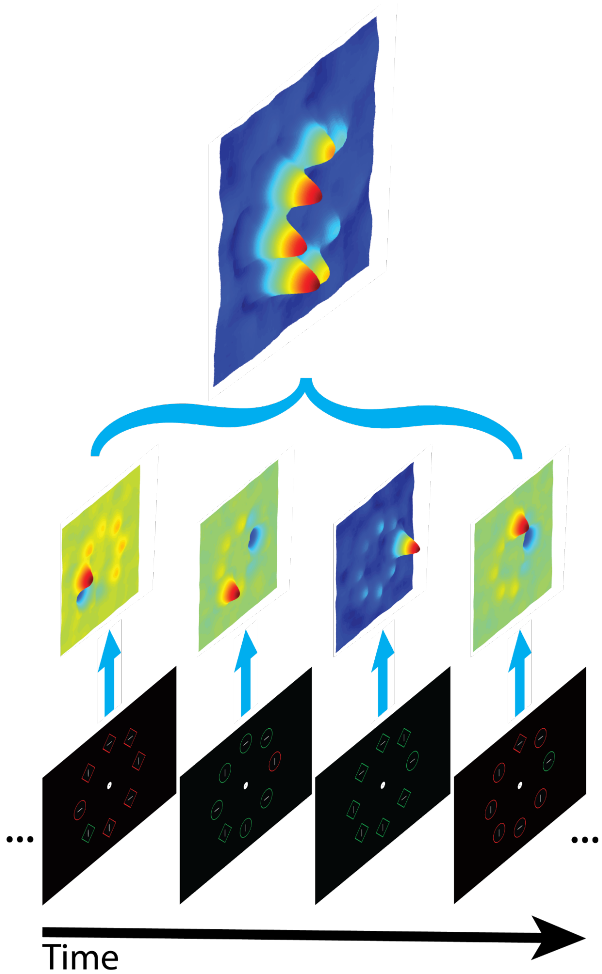

FIgure 3: The conceptual trial-by-trial Kalman filter behavior as individual trials generate attention traces representing either positive enhancement at target locations, or negative suppression at distractor locations. The accumulation of traces are integrated into an overall priority map represented at the top. Trace influence is modeled as a power unction wherein recent trial features influence immediate trial behavior more than distant trial features. The accumulation of trial history will represent the statistics underlying the task, which in this visual example would be suppression of the rightmost location.

References

Cousineau, Denis. 2005. “Confidence Intervals in Within-Subject Designs: A Simpler Solution to Loftus and Masson’s Method.” *Tutorials in Quantitative Methods for Psychology* 1(1):42–45.

Duncan, Dock H., Dirk van Moorselaar, and Jan Theeuwes. 2024. “Visual Statistical Learning Requires Attention.” *Psychonomic Bulletin & Review*. doi: 10.3758/s13423-024-02605-1.

Logan, Gordon D. 1988. “Toward an Instance Theory of Automatization.” *Psychological Review* 95(4):492.

Los, Sander A., Wouter Kruijne, and Martijn Meeter. 2014. “Outlines of a Multiple Trace Theory of Temporal Preparation.” *Frontiers in Psychology* 5:1058.

Maljkovic, Vera, and Ken Nakayama. 1994. “Priming of Pop-out: I. Role of Features.” *Memory & Cognition* 22(6):657–72.

Maljkovic, Vera, and KEN Nakayama. 1996. “Priming of Pop-out: II. The Role of Position.” *Perception & Psychophysics* 58(7):977–91.

Morey, Richard D. 2008. “Confidence Intervals from Normalized Data: A Correction to Cousineau (2005).” *Tutorials in Quantitative Methods for Psychology* 4(2):61–64.

Musfeld, Philipp, Alessandra S. Souza, and Klaus Oberauer. 2023. “Repetition Learning Is Neither a Continuous nor an Implicit Process.” *Proceedings of the National Academy of Sciences* 120(16):e2218042120.

Salet, Josh M., Wouter Kruijne, Hedderik van Rijn, Sander A. Los, and Martijn Meeter. 2022. “FMTP: A Unifying Computational Framework of Temporal Preparation across Time Scales.” *Psychological Review* 129(5):911.

Wang, Benchi, and Jan Theeuwes. 2018. “Statistical Regularities Modulate Attentional Capture.” *Journal of Experimental Psychology: Human Perception and Performance* 44(1):13.