

# Procedural Learning with Graded Entropy

## Abstract

Your abstract text here. The abstract should be one paragraph. Following the abstract should be keywords.

**Keywords:** procedural learning, implicit learning, statistical learning

## Introductino

How many times have you typed “form” instead of “from”? Or ended a word with “-tino” instead of “-tion”, perhaps even in an important, embarrassingly glaring context? The typos we make are not random; they reflect the statistical patterns in the language we use. Experienced typists learn that some keystroke sequences are highly probable, causing their fingers to automatically perform ones like “for” mistakenly, or perform them so quickly they slip up, like in the case of “-tino”. Yet typists cannot consciously report these probabilities or articulate their finger movements at the speed of execution—this knowledge is expressed in the automatic motor performance of typing, not in conscious awareness. This implicit form of learning, where statistical regularities are unconsciously acquired and expressed as behavior rather than declared knowledge, is called procedural learning (Squire (2004), Nissen & Bullemer (1987)).

Procedural learning is not limited to typing. It underlies skill acquisition across many domains, from proprioceptive skills like riding a bike to perceptual-cognitive skills like reading mirrored text (Kassubek et al. (2001)). More recently, procedural learning has been proposed to support increasingly diverse and sophisticated abilities, notably grammar acquisition in language (Ullman & Pierpont (2005), Lammertink et al. (2017), Lum et al. (2014), Kidd & Arciuli (2016)). These proposals invite closer scrutiny of its scalability: can the procedural learning system handle the large-scale, complex statistical structures that characterize domains like grammar? Consider a suggestive historical case: while QWERTY keyboards (26 letter keys) enabled fluent touch typing through procedural learning, Chinese typewriters (trays of 2,000+ characters) never achieved comparable automaticity and were eventually abandoned (for QWERTY keyboards, among others!) (Mullaney (2017)). This raises a fundamental question: does procedural learning fail beyond a certain level of statistical complexity and state space size? As a first step toward answering these questions, we examine whether procedural learning tracks fine-grained statistical patterns—both the predictability of individual transitions and the uncertainty of different states—and how it scales across state space sizes from 4 to 8 positions.

The primary method for measuring procedural learning has been the Serial Reaction Time Task (SRTT; Nissen & Bullemer (1987)). In the SRTT, participants respond to stimuli appearing in one of four positions on screen using key presses. Unbeknownst to them, the sequence of positions follows a deterministic pattern. Over time, reaction times (RT) decrease for patterned trials relative to random trials, demonstrating procedural learning of the key press sequence corresponding

to the positional sequence’s deterministic pattern. While the SRTT provided early evidence for procedural learning, it suffers from a critical confound: participants often develop explicit awareness of the sequence, making it difficult to isolate implicit procedural learning effects (Song et al. (2008), Lustig (2022)). This has contributed to low test-retest reliability of the task (West et al. (2018), Oliveira et al. (2023)).

The Alternating Serial Reaction Time Task (ASRT) addressed these issues by introducing a probabilistic design. In ASRT, deterministic (d) and random (r) trials alternate (e.g., d-r-d-r), resulting in positional sequences with two levels of frequency (high vs. low) (Howard Jr. & Howard (1997)). This probabilistic design minimizes explicit awareness and dramatically improves reliability (Farkas et al. (2024), Oliveira et al. (2023)). However, a key limitation of ASRT is that its bimodal frequency distribution represents an overtly simplistic statistical structure, lacking the graded frequency distributions found in real-world domains like language. Finally, across both SRTT and ASRT, experiment designs have mostly tested sequences with four possible positions (e.g., Janacek et al. (2012), Hedenius et al. (2011)), despite real-world domains where procedural learning is implicated often involving far larger state spaces. Whether procedural learning can track graded statistical structures, and whether it scales to larger state spaces, remains untested.

Here, we introduce a new paradigm that retains ASRT’s probabilistic design but introduces a graded statistical structure with a smooth continuum of transition probabilities. We construct transition matrices to generate positional sequences where each position has a different entropy, ranging from deterministic (low entropy) to uniformly distributed (high entropy). We also vary the number of positions (4, 5, 6, 7, 8) to test whether procedural learning scales with state space size. This allows us to ask: (1) Are learners sensitive to fine-grained differences in transition probabilities (surprisal) and state-level uncertainty (entropy), beyond simple high vs. low frequency contrasts? (2) Do learning trajectories change as the number of positions increases? Across 219 participants, we find evidence that procedural learning is sensitive to graded statistical structure, and that this sensitivity emerges early and strengthens over time. These results suggest that procedural learning is more flexible than prior paradigms can reveal, with implications for understanding how procedural learning operates in naturalistic environments.

## Hypotheses

We make the following predictions: First, if participants are learning the statistical structure and not merely practicing motor responses, trial-level RT and accuracy will be predicted by surprisal—the negative log probability of the observed transition—such that higher-surprisal (less predictable) transitions elicit slower and less accurate responses (surprisal effect). Second, previous positional entropy—the entropy of the transition distribution from the previous position—will modulate performance independently, with higher-entropy (more uncertain) previous positions leading to slower and

Table 1: Key mappings by number of positions

$n$ positions	keys (left to right)
4	{D, F, J, K}
5	{S, D, F, J, K}
6	{S, D, F, J, K, L}
7	{A, S, D, F, J, K, L}
8	{A, S, D, F, J, K, L, ;}

less accurate responses (entropy effect). We further predict that surprisal effects will be stronger following low-entropy positions, where predictions are more confident and prediction errors more costly (surprisal  $\times$  entropy interaction). Third, these effects will be attenuated as the number of positions increases, reflecting greater difficulty in learning more complex statistical structures (state space size effect).

## Method

### Participants

We recruited 251 participants (133 female, 118 male) to complete a web-based experiment using the online crowd-sourcing platform Prolific, screening for participants with an approval rating above 95% from previous studies and limiting participation to be on desktop or laptop computers. Participants ranged in age from 18 to 75 years ( $\mu=35.6$ ,  $\sigma=11.9$ ), representing 32 nationalities and 20 primary languages. Participants were compensated at \$12.00 per hour, with median completion time varying by condition (~20-40 minutes). All participants provided informed consent prior to the experiment. The study was approved by the Institutional Review Board at the Mass General Hospital Institute of Health Professions.

### Experiment Design

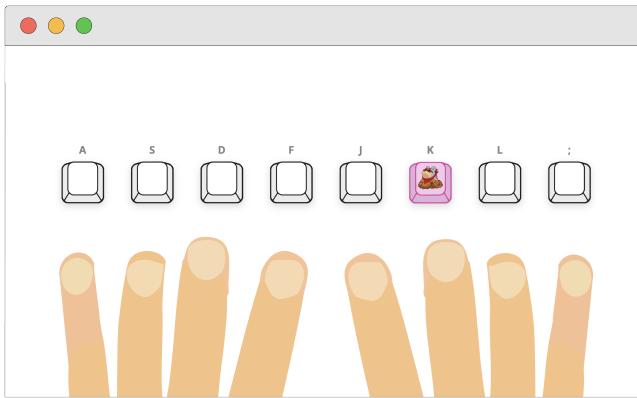


Figure 1: An example trial screen in the 8-position condition of the experiment

Like other serial reaction time task designs, our experiment tasks participants to respond to a visual stimulus that appears in one of several evenly-spaced positions on screen as quickly and accurately as possible by pressing a corresponding key on the keyboard. In our implementation, this visual stimulus is a mole garbed in a bright red bib and matching sunglasses

(Figure 1). Participants were evenly divided into 5 groups of 50 per condition, with the conditions differing in the number of positions (4, 5, 6, 7, 8) the mole can appear in. ((**tab:key-mapping?**)) shows the position-to-key mapping for each condition.

A participant is first given a text-based tutorial accompanied by an animated demonstration on which key to press in response to each position, then instructed to work through a set of practice trials where each position is visited twice in random order. After the practice section, the participant moves on to complete 20 blocks of trials, each separated with a self-paced break. Each block consists of 10 trials per position. This design choice results in longer blocks for conditions with more positions, but ensures each position is visited an equal number of times on average across conditions. In both practice and main trials, participants are given feedback on correctness via a pop-up short message (a checkmark vs. “Try again!” + an error tone), and allowed to retry each failed trial until they respond correctly. Following standard SRTT protocol, there is also a 120ms response-stimulus interval (RSI) between trials (Nissen & Bullemer (1987), Janacsek et al. (2012), Howard et al. (2004)). We find during piloting that allowing retries, combined with the 120ms RSI, prevents participants from making compensatory errors, where they unintentionally press the wrong key in the next trial in an attempt to correct their current trial. We record response time, keyboard response and correctness for each trial. After each block, participants are given adaptive feedback based on their accuracy and speed. At the end of the experiment, participants complete a brief questionnaire probing their explicit awareness of any patterns in the mole’s appearances (Table).

In addition to our dapper mole, a number of design choices were made to facilitate participant understanding and engagement. During the practice phase, each position on screen is labelled with its corresponding key character (e.g. “A”, “S”, “D”) to help familiarize participants with the keyboard mappings. These key labels disappear in the main trials to prevent explicitly encoding of the sequence of positions via their character labels, but reappear during retry of failed trials. To make the correspondence between the positions on screen and keyboard keys as automatic and intuitive as possible, we chose to use conventional QWERTY-based finger layouts for the key mappings across conditions ((**tab:key-mapping?**)). We found during piloting that while this decision helped reduce cognitive load from learning novel finger placements, the larger gap between keys “F” and “J” on the keyboard introduced a discrepancy between the evenly-spaced on-screen positions and the unevenly spaced keyboard keys. This made participants more prone to making errors in the positions towards the middle of the screen compared to those on the outer edges, especially when number of positions increases [TODO: maybe run stats on pilot to find this effect]; and was not mitigated by shifting either hand’s placement to close the extra gap, because the positions were still corresponding half to one hand and half to the other. After further piloting testing different visual aids, we settled on adding two hands visually to the bottom of the screen, each finger aligned with its target position (Figure). We received verbal feedback that this visual aid helped participants better map the spatial layout of the on-screen positions to their corresponding keys.

Lastly, the visual design of the “positions” on screen mimicked blank 3D keyboard keys that light up in pink when the mole appears on top, and look visibly “pressed down” when the participant presses the corresponding keyboard key. This design choice links the on-screen positions diegetically to the physical keyboard keys to maximally reduce mental distance between the two. This is important to enforce as much learning via implicit motor memory as possible, rather than via visual-spatial memory of the on-screen positions.

## Transition Matrices

Our experiment deviates from other serial reaction time tasks in how the sequence of positions and its probabilistic structure is constructed. In each run of the experiment, we use a first-order Markov process to generate the sequence of positions the mole appears in. This process is defined by a transition matrix, where each entry in the matrix defines the probability of transitioning from one position to another. We designed five different transition matrices of sizes 4x4, 5x5, 6x6, 7x7, and 8x8 respectively corresponding to the five conditions in our experiment. These “original” matrices were constructed to have graded and increasing levels of entropy across rows, such that position 1 has lowest entropy—having the most predictable next-position-transitions, followed by the second, with the last position, position  $n$ , having highest entropy. For example, in the 4x4 matrix, position 1 has a high probability ( $\sim 0.97$ ) of transitioning to position 2, a low probability ( $\sim 0.03$ ) of transitioning to position 1, and never transitions to position 3 or 4; while position 4 has a more uniform distribution of transition probabilities to all other positions (to 1:  $\sim (0.39)$ , to 2:  $\sim (0.03)$ , to 3:  $\sim (0.24)$ , to 4:  $\sim (0.35)$ ). This gives rise to a gradient of entropy values across the positions, which results in a richer, more varied distributed statistical structure underlying the sequence of positions the mole appears in, allowing us to investigate how the procedural learning system differentially picks up varying levels of predictability in a given statistical structure.

Several design criteria/constraints were imposed on the construction of these original transition matrices. First, these matrices had to be doubly stochastic to ensure uniform visitation to each position over time, preventing confounding effects from position frequency on learning. Second, self-transitions (e.g. position 1 to position 1) should have near-zero probabilities to reduce data loss, as data from repeated trials have been historically thrown out from analysis in serial reaction time tasks to prevent confounding motor effects from repeating the same keystroke [CITE]. Third, the graded entropy levels across the positions should translate to graded probabilities across bigrams as well, as having more fine-grained, distributed levels of transition probabilities is our experiment’s key distinction from other SRTT tasks, such as the ASRT paradigm which only has 2 levels of trigram probability. Fourth, the increase in entropy across positions (i.e. rows) should be significant between one another and as linear as possible to ensure smooth a smooth gradient with meaningfully distinct levels in predictability across positions.

Strictly satisfying all these constraints at once turned out to be not only non-trivial, but mathematically impossible for our range of  $n_{\text{positions}}$  (i.e. matrix sizes), so we wrote an optimization algorithm that explores the candidate space for

each matrix size, treating these constraints as soft penalties to find a global minimum of the total constraint violation, balancing trade-offs among the competing constraints.

The five original matrices we constructed using this algorithm had these properties: 1) near fully doubly stochastic—all row and column sums sum to 1 ( $\pm 0.0001$ ); 2) approach a uniform stationary distribution after 7 timesteps where the stationary visitation for each state (position) are within 0.001 of each other; 3) had a graded entropy structure across states—each matrix had a positional entropy range of 0.061-0.181 bits for position 1 ( $\mu=0.107$ ,  $\sigma=0.048$ ), and 1.695-2.835 bits for position  $n$  ( $\mu=2.319$ ,  $\sigma=0.451$ ). Within a matrix each position increases in entropy from the previous by 0.295-0.699 bits ( $\mu=0.442$ ,  $\sigma=0.085$ ); and 4) avoided self loops as much as possible—all of the matrices had  $< 0.08$  probabilities in the left diagonal up until the last two positions. Within each matrix size condition, all participants experience the same underlying entropy gradient across positions, but the mapping between screen positions and entropy levels is randomized through a double permutation procedure that shuffles the original matrix (of that size) using the same random permutation. This preserves the statistical properties of the original matrix—transition probabilities and entropy values—but decouples entropy levels from specific screen positions. The position sequence for each participant is subsequently generated using their individual shuffled transition matrix, so the sequence each participant experiences is different both within and between matrix size conditions.

## Statistical Analysis

### Variables and Exclusion Criteria

The main independent variables are surprisal and previous positional entropy. Surprisal is defined as the negative log probability of the actual next-position given the current position [TODO: math], derived from the transition matrix used to generate the sequence for that participant. Previous positional entropy is defined as the Shannon entropy of the transition probabilities from the previous position, also derived from the transition matrix [TODO: math]. Both surprisal and previous positional entropy are mean-centered within each matrix size condition (`surprisal_c` and `prev_entropy_c`, respectively) to facilitate interpretation of main effects and interactions in subsequent models. We also create a binary variable indicating whether the current position is a repetition of the previous position (i.e. self-transition). This variable is included as a covariate in subsequent models to control for potential motor effects from repeating the same keystroke.

The main dependent variables are response time (RT) and accuracy (proportion of correct responses) on non-practice/main trials. Incorrect trials, trials with RT  $> 1000\text{ms}$ , and trials with RTs greater than 3 median absolute deviations above the participant’s median RT (outliers) are excluded. On the participant level, participants with overall accuracy below 85%, median RT  $> 1000\text{ms}$ , or with outliers making up more than 30% of total trials are excluded from analysis. We also discard data from the first trial of each block to account for blocks being implicitly treated as “restarts” and not following an existing probability distribution by participants. RTs are log-transformed to reduce skewness.

## Modeling Approach

We use linear mixed effects regression (LMER) to model log-transformed RTs and generalized linear mixed effects regression (GLMER) with a binomial link function to model accuracy. Models are fit using the `lmerTest` package in R. *P*-values were obtained using the Satterthwaite approximation for degrees of freedom, as implemented in the `lmerTest` package (Kuznetsova et al. (2017)).

We first determine the optimal random and fixed effects structure using the final block of trials in each condition, fitting candidate models to each of the five matrix size conditions separately and selecting the structure with the lowest summed AIC across all conditions. Models are initially fit with the maximal random effects structure supported by the design (Barr et al. (2013)): random intercepts for participant and position, and random slopes by participant for surprisal, previous positional entropy, their interaction, and whether the current position is a repetition of the previous position. Where maximal models fail to converge, we use bounded linear mixed models (`blmer/bglmer`) as a first remedy. If singularity persists, we iteratively simplify the random effects structure, removing random slopes one at a time and retaining the most complex structure that converges without singularity. Among these converging models, we select the simplest structure without increasing AIC. For fixed effects, starting with the full model including surprisal, previous positional entropy, their interaction, and repetition, we use AIC-based stepwise backward selection. An effect is retained if removing it increases AIC.

Using this selected structure, we then fit separate models to each individual block within each matrix size condition to examine learning trajectories. Finally, we extract the effect size estimates from each model and test whether they are moderated by matrix size and block using linear regression without random effects, as the data are aggregated coefficients.

## Results

As a preliminary sanity check, we confirmed that RTs decreased across blocks ( $\beta = -0.01$ ,  $t = -11.72$ ,  $p < .001$ , 95% CI [-0.007, -0.005]), establishing that performance improvement occurred.

### Surprisal and Previous Position Entropy Effects

We first determined the optimal effects structure using final-block data, comparing candidate models by summing AIC values across all five matrix sizes. This comparison favored the model: `log_rt ~ surprisal_c * prev_entropy_c + is_repetition + (surprisal_c * prev_entropy_c || subject_id) + (1 | position)`

We then fit this model structure each block in each matrix size condition separately, further pruning the random effects structure until all models converge. This pruning favored a simpler random effects structure, resulting in the following final model: `log_rt ~ surprisal_c * prev_entropy_c + is_repetition + (prev_entropy_c | subject_id) + (1 | position)`

In all blocks across all matrix sizes, this model revealed significant effects for all predictors (Table 2). This structure converged without singularity across all 100 models (20 blocks  $\times$  5 conditions) and is therefore used for all subsequent

Table 2: Fixed effects from all per-block models (100 total: 20 blocks  $\times$  5 matrix sizes)

Predictor	Mean $\beta$	Mean $t$	$p$
Surprisal	0.024	3.28	
Previous Entropy	0.040	2.14	
Surprisal $\times$ Entropy	-0.008	-1.01	
Repetition	-0.180	-9.37	***

analyses.

### State Space Size Effect

To examine learning trajectories, we fit this selected model structure separately to each block within each matrix size condition, yielding 100 models total (20 blocks  $\times$  5 conditions). This approach allows us to track how sensitivity to surprisal and entropy emerges and changes over the course of learning, and whether these trajectories differ across state space sizes.

## Discussion

## References

- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Farkas, B. C., Krajcsi, A., Janacek, K., & Nemeth, D. (2024). The complexity of measuring reliability in learning tasks: An illustration using the Alternating Serial Reaction Time Task. *Behavior Research Methods*, 56(1), 301–317. <https://doi.org/10.3758/s13428-022-02038-5>
- Hedenius, M., Persson, J., Tremblay, A., Adi-Japha, E., Verissimo, J., Dye, C. D., Alm, P., Jennische, M., Bruce Tomblin, J., & Ullman, M. T. (2011). Grammar predicts procedural learning and consolidation deficits in children with Specific Language Impairment. *Research in Developmental Disabilities*, 32(6), 2362–2375. <https://doi.org/10.1016/j.ridd.2011.07.026>
- Howard, D. V., Howard, J. H., Japikse, K., DiYanni, C., Thompson, A., & Somberg, R. (2004). Implicit Sequence Learning: Effects of Level of Structure, Adult Age, and Extended Practice. *Psychology and Aging*, 19(1), 79–92. <https://doi.org/10.1037/0882-7974.19.1.79>
- Howard Jr., J. H., & Howard, D. V. (1997). Age differences in implicit learning of higher order dependencies in serial patterns. *Psychology and Aging*, 12(4), 634–656. <https://doi.org/10.1037/0882-7974.12.4.634>
- Janacek, K., Fiser, J., & Nemeth, D. (2012). The best time to acquire new skills: Age-related differences in implicit sequence learning across the human lifespan. *Developmental Science*, 15(4), 496–505. <https://doi.org/10.1111/j.1467-7687.2012.01150.x>
- Kassubek, J., Schmidtke, K., Kimmig, H., Lücking, C. H., & Greenlee, M. W. (2001). Changes in cortical activation during mirror reading before and after training: An fMRI study of procedural learning. *Cognitive Brain Research*, 10(3), 207–217. <https://doi.org/https://doi.org/10.1016/S0>

- Kidd, E., & Arciuli, J. (2016). Individual Differences in Statistical Learning Predict Children's Comprehension of Syntax. *Child Development*, 87(1), 184–193. <https://doi.org/10.1111/cdev.12461>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Lammertink, I., Boersma, P., Wijnen, F., & Rispens, J. (2017). Statistical Learning in Specific Language Impairment: A Meta-Analysis. *Journal of Speech, Language, and Hearing Research*, 60(12), 3474–3486. [https://doi.org/10.1044/2017\\_JSLHR-L-16-0439](https://doi.org/10.1044/2017_JSLHR-L-16-0439)
- Lum, J. A. G., Conti-Ramsden, G., Morgan, A. T., & Ullman, M. T. (2014). Procedural learning deficits in specific language impairment (SLI): A meta-analysis of serial reaction time task performance. *Cortex*, 51, 1–10. <https://doi.org/10.1016/j.cortex.2013.10.011>
- Lustig, C. (2022). *The transition from implicit to explicit knowledge representations in implicit sequence learning* [Text.thesis.doctoral, Universität zu Köln]. <http://www.uni-koeln.de/>
- Mullaney, T. S. (2017). *The Chinese Typewriter: A History*. MIT Press.
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19(1), 1–32. [https://doi.org/10.1016/0010-0285\(87\)90002-8](https://doi.org/10.1016/0010-0285(87)90002-8)
- Oliveira, C. M., Hayiou-Thomas, M. E., & Henderson, L. M. (2023). The reliability of the serial reaction time task: Meta-analysis of test–retest correlations. *Royal Society Open Science*, 10(7), 221542. <https://doi.org/10.1098/rsos.221542>
- Song, S., Howard, J. H., & Howard, D. V. (2008). Perceptual sequence learning in a serial reaction time task. *Experimental Brain Research*, 189(2), 145–158. <https://doi.org/10.1007/s00221-008-1411-z>
- Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory*, 82(3), 171–177. <https://doi.org/10.1016/j.nlm.2004.06.005>
- Ullman, M. T., & Pierpont, E. I. (2005). Specific Language Impairment is not Specific to Language: The Procedural Deficit Hypothesis. *Cortex*, 41(3), 399–433. [https://doi.org/10.1016/S0010-9452\(08\)70276-4](https://doi.org/10.1016/S0010-9452(08)70276-4)
- West, G., Vadillo, M. A., Shanks, D. R., & Hulme, C. (2018). The procedural learning deficit hypothesis of language learning disorders: We see some problems. *Developmental Science*, 21(2), e12552. <https://doi.org/10.1111/desc.12552>