Human Random Generation as a Locally-Bound Process

Lucas Castillo Pablo León-Villagrá Nick Chater
Adam Sanborn

16th July 2021

The concept of randomness is central to human probabilistic inference, and so research into how people perceive and produce randomness has a long tradition in psychology and cognitive science (e.g. Baddeley, 1966; Wagenaar, 1970). However, when producing random sequences, people systematically deviate from randomness, which is problematic for any model of cognition that assumes that people can generate independent, identically-distributed samples.

One prominent explanation of people's performance in these tasks is that they generate sequences based on learned schemas, such as counting up or down, and switch between schemas when a monitoring process perceives randomness to be declining (see Cooper, 2016). Instead, we propose that people are tapping into a general cognitive ability to produce samples for inference – a sophisticated mental algorithm that does not sample independently, but locally (Chater et al., 2020).

In the types of experiments commonly used in human randomness research, such as generating random numbers or letters of the alphabet, the schema and sampling account produce similar predictions. To distinguish between the two accounts, we introduced a novel experimental paradigm that directly manipulated the dimensionality and order of the domain. In our task, participants first learned a set of syllables arranged in either one- or two-dimensional displays (with varying item order) and then generated random sequences of the items they had learned at a constant pace (Figure 1).

To contrast schema and sampling approaches, we constructed models that computed the likelihood of each sequence based on either the frequency of syllables in English and the frequency of syllables in participants' learning stage (schema model), or based on the spatial structure of the display (local sampling model). We found that the local sampling model predicted participants' sequences best. Analysing the relative contributions of each models' parameters, we found that participants' sequences were best characterized by a sampler that avoided item repetition (No Repeat) and kept its spatial trajectory (Momentum; see Figure 3 for an illustration).

Our model results were consistent with global descriptors of the sequences that were used in previous literature (see Figure 1 for an example): participants repeated items less often, made transitions to adjacent items more often, and made fewer abrupt direction changes than expected from a random sequence, although some of these patterns varied according to domain dimensionality. In addition, longer delays between items made repetitions and direction changes more likely and adjacent items less likely, consistent with more items being sampled in between utterances (Figure 2).

Together, our results suggest that human random sequences result from of a local sampling mechanism where successive samples share a trajectory. These patterns could be produced by a sampling algorithm with momentum which partially reuses the momentum of previous iterations (Horowitz, 1991), or one where multiple samples across a common trajectory are proposed (for instance, Nishimura & Dunson, 2020).

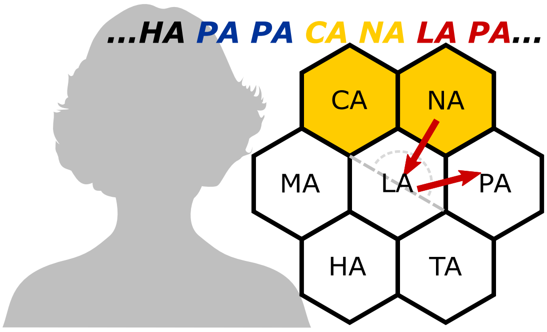

Figure 1: Participants generated a sequence of random syllables without seeing the display. We computed the proportion of repetitions, adjacent items and abrupt turns (Examples of these in blue, yellow and red, respectively).

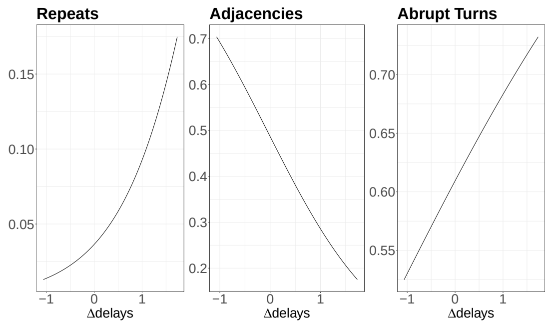

Figure 2: Estimated probability of the next transition being a repetition, an adjacent item, or an abrupt turn, according to the temporal delay between items, relative to the sequence mean (Δdelays).

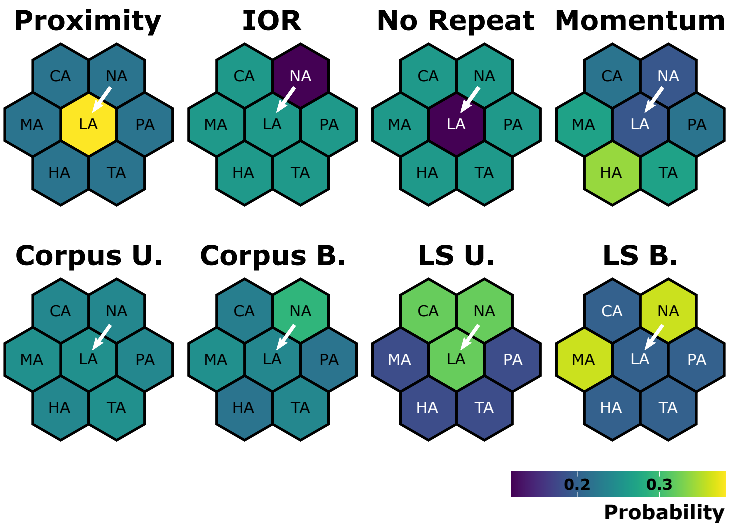

Figure 3: Example model predictions after a *na, la* transition. The constituent parts of the local sampling model (top) calculate the likelihood of the next item based on the topology of the display, whereas the schema model (bottom) uses the frequency of syllables in English and in the learning sequence.

References

Baddeley, A. D. (1966). The Capacity for Generating Information by Randomization. *Quarterly Journal of Experimental Psychology*, *18*(2), 119–129. https://doi.org/10.1080/14640746608400019

Chater, N., Zhu, J.-Q., Spicer, J., Sundh, J., León-Villagrá, P., & Sanborn, A. (2020). Probabilistic Biases Meet the Bayesian Brain. *Current Directions in Psychological Science*, *29*(5), 506–512. https://doi.org/10.1177/0963721420954801

Cooper, R. P. (2016). Executive Functions and the Generation of "Random" Sequential Responses: A Computational Account. *Journal of Mathematical Psychology*, *73*, 153–168. https://doi.org/10.1016/j.jmp.2016.06.002

Horowitz, A. M. (1991). A Generalized Guided Monte Carlo Algorithm. *Physics Letters B*, *268*(2), 247–252. https://doi.org/10.1016/0370-2693(91)90812-5

Nishimura, A., & Dunson, D. (2020). Recycling Intermediate Steps to Improve Hamiltonian Monte Carlo. *Bayesian Analysis*, *15*(4), 1087–1108.

Wagenaar, W. (1970). Appreciation of Conditional Probabilities in Binary Sequences. *Acta Psychologica*, *34*, 348–356.