Limits of prediction:
False memory for predictable words is unaffected by testing delay

Katja Haeuser (khaeuser@coli.uni-saarland.de)
Department of Psychology, Saarland University

Prediction has long been known to affect online language comprehension. More recently, it has also been suggested to have detrimental effects on long-term memory by engendering false memories. For example, after reading the sentence “At the hospital, the nurse stitches up the pants”, participants show elevated levels of false recognition memory for the predictable but not presented word “injury” (henceforth, “lure”) in memory retrieval tests. But are these false-memory effects truly driven by prediction? A recent view challenges this assumption [1], arguing instead that false remembering is constructed during retrieval when participants rely on semantically associative strategies to “jog” their memory of the previously encoded sentences.

I sought to adjudicate between prediction and retrieval accounts of false memory. Native German speakers read predictable and unpredictable sentences in a self-paced reading (SPR) task and completed two recognition memory tests (within subjects): An immediate test (T1; N = 68; Mean Age = 35; M/F/NB = 24/41/3) and a delayed test one week later (T2; N = 22; Mean Age = 37; M/F/NB = 8/12/2; the study was run online, only a subset of the participants from T1 returned for T2). If false remembering is related to prediction during reading, false-alarm rates for lures should be lower at T2, as prediction effects should fade quickly, akin to spreading-activation effects [2]. If false remembering emerges at retrieval, false-memory rates for predictable lures should stay the same or even increase over time [3]. Results reported take into account data from all available participants for each task and time point (Fig 1A+B); as paired data showed the same pattern of results statistically (Fig 1C+D). During SPR, participants read unpredictable items more slowly, suggesting greater integration difficulty (Fig. 1A+C). In the recognition task, even though the magnitude of the false-memory *effect* (false alarms to lures vs. new words) decreased from T1 to T2 (in line the with the prediction account), absolute rates of false remembering for lures increased minimally from T1 to T2 in the full-data sample (Fig. 1B) or stayed the same in the paired sample (Fig. 1D) – both inconsistent with the prediction account. For correct recognition memory, word condition interacted with time: At T1, memory was better for unpredictable words, whereas at T2, memory was better for predictable words.

In sum, false remembering was comparable at immediate and delayed testing, consistent with a retrieval-based account of false memory, which argues that false memory is constructed during retrieval by means of semantic association. Notably, correct recognition memory was differentially influenced by prediction error and schema congruency over time: Prediction error only improved memory at shorter, and not longer, time intervals [4].

References

[1] Haeuser, K. I., & Kray, J. (2026). Do you remember? Effects of lexical predictability on false and true recognition memory. Journal of Experimental Psychology: Learning, Memory and Cognition.

[2] Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. Psychological Review, 82(6), 407–428.

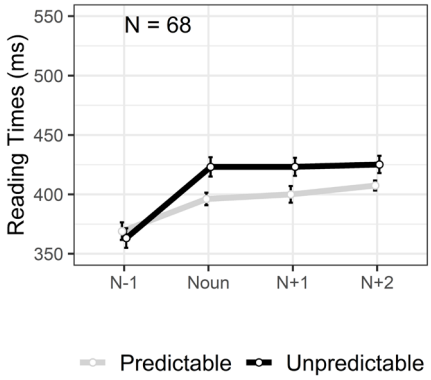
[3] Reyna, V. F., Corbin, J. C., Weldon, R. B., & Brainerd, C. J. (2016). How fuzzy-trace theory predicts true and false memories for words, sentences, and narratives. Journal of Applied Research in Memory and Cognition, *5*(1), 1–9.

[4] Greve, A., Cooper, E., Tibon, R., & Henson, R. N. (2019). Knowledge is power: Prior knowledge aids memory for both congruent and incongruent events, but in different ways. Journal of Experimental Psychology: General, 148(2), 325–341.

Figure 1. A: Self-paced reading times for words in the critical region (… the $_{N-1}$ injury/pants $_{Noun}$ of $_{N+1}$ the $_{N+2}$ little girl). **B:** Recognition memory performance (proportion of items judged as old) as a function of time point and condition. At each time point, participants were presented with predictable and unpredictable (old) words, new words and lures, and asked to indicate if presented words were “old” or “new”. For lure and new items, the graph shows false alarm rates. For old items, the graph shows hit rates. Panels **C** and **D** show reading and memory rates for the subset of data from fully completed observations (N = 22). The study was run online; only a subset of the participants from T1 returned for T2.

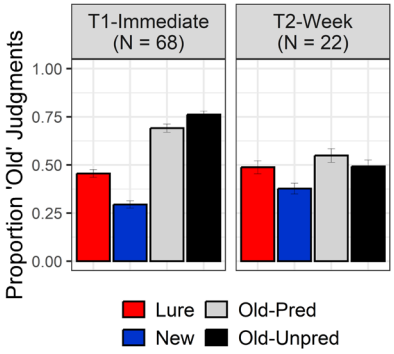
A

Model syntax for SPR: lmer(log(RT) ~ condition + frequ + length + position + random_effects, data=data)

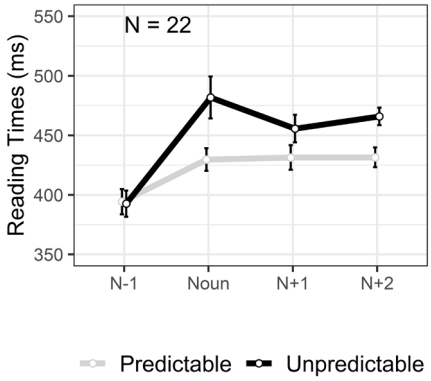


B

Model syntax for recognition memory: Judged_as_old ~ condition * time + frequ + length + random_effects, family = binomial



C



D