



Retrieval Practice and Memory Precision: A Metacognitive Approach

Brendan A. Schuetze, Luke G. Eglington, & Sean H.K. Kang
Dartmouth College, Hanover, NH USA

Motivation

- Recent memory research has begun to analyze the effects of interventions on continuous measures of memory (e.g., Donkin et al., 2014; Sutterer & Awh, 2016).
- Previous research on retrieval practice (RP) has predominantly featured stimuli with discrete right-or-wrong answers.
- A more continuous measure potentially offers greater sensitivity in assessing the effects of RP.
- We used Koriat and Goldsmith’s (1996) work on memory precision as a basis for understanding how retrieval practice (RP) affected the grain size at which participant’s reported their memories for the color of line drawn objects.

Question

How does retrieval practice affect participant-reported grain sizes, which act as metacognitive indicators of memory precision?

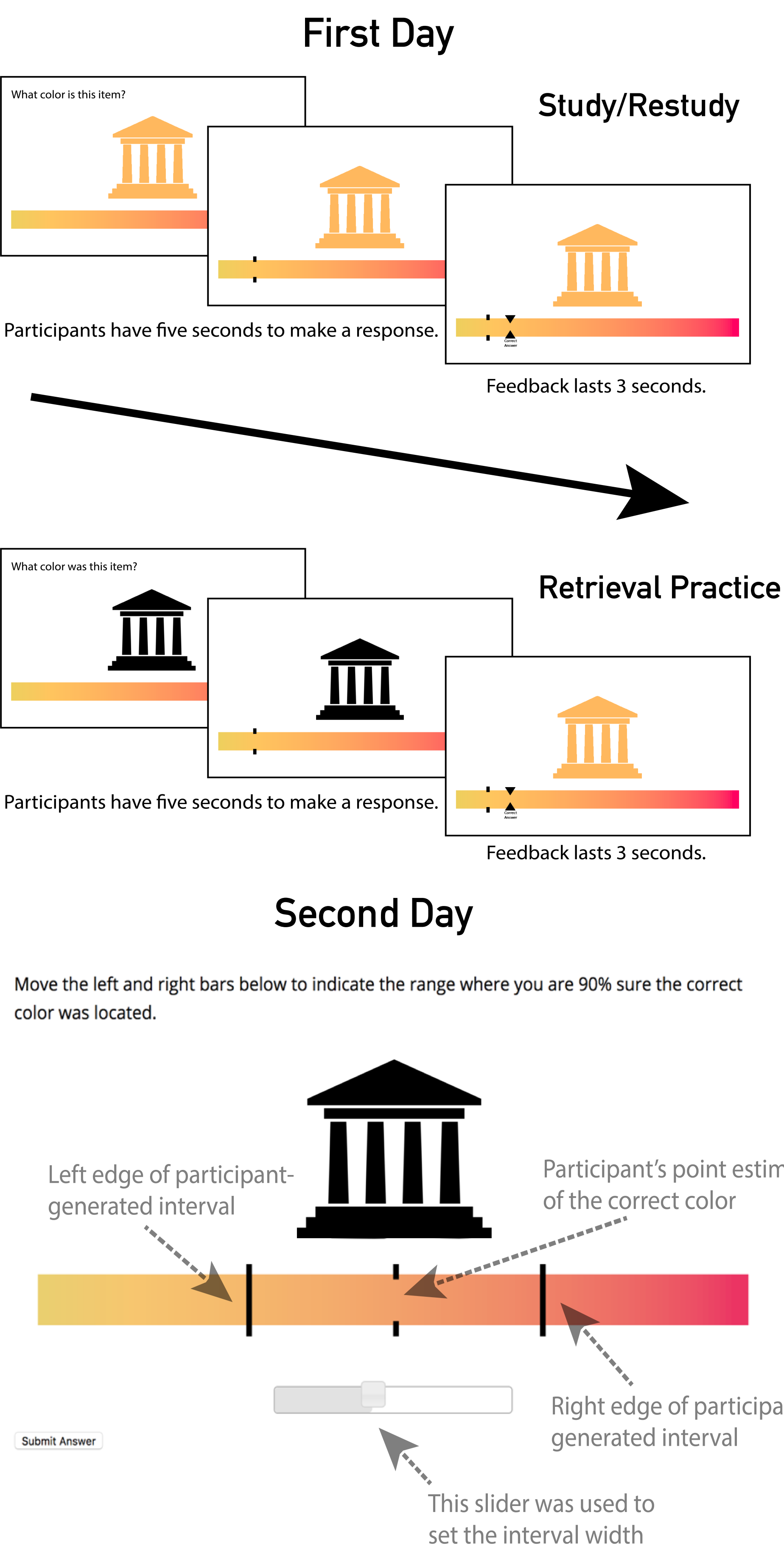
Hypotheses

- Retrieval Practice will lead to a reduction in error (i.e., better performance).
- Retrieval Practice will result in smaller interval widths, indicating higher metacognitive evaluations of memory fidelity.

Methods

- MTurkers (n = 67) learned a set of 80 line-drawn items filled with random colors taken from a subset of the LAB color space.
- Items were manipulated via the number of presentations (1 vs. 3) and also the manner in which they were studied (restudy vs. retrieval practice).

Procedure on Each Trial



Results

Mean Abs. Error

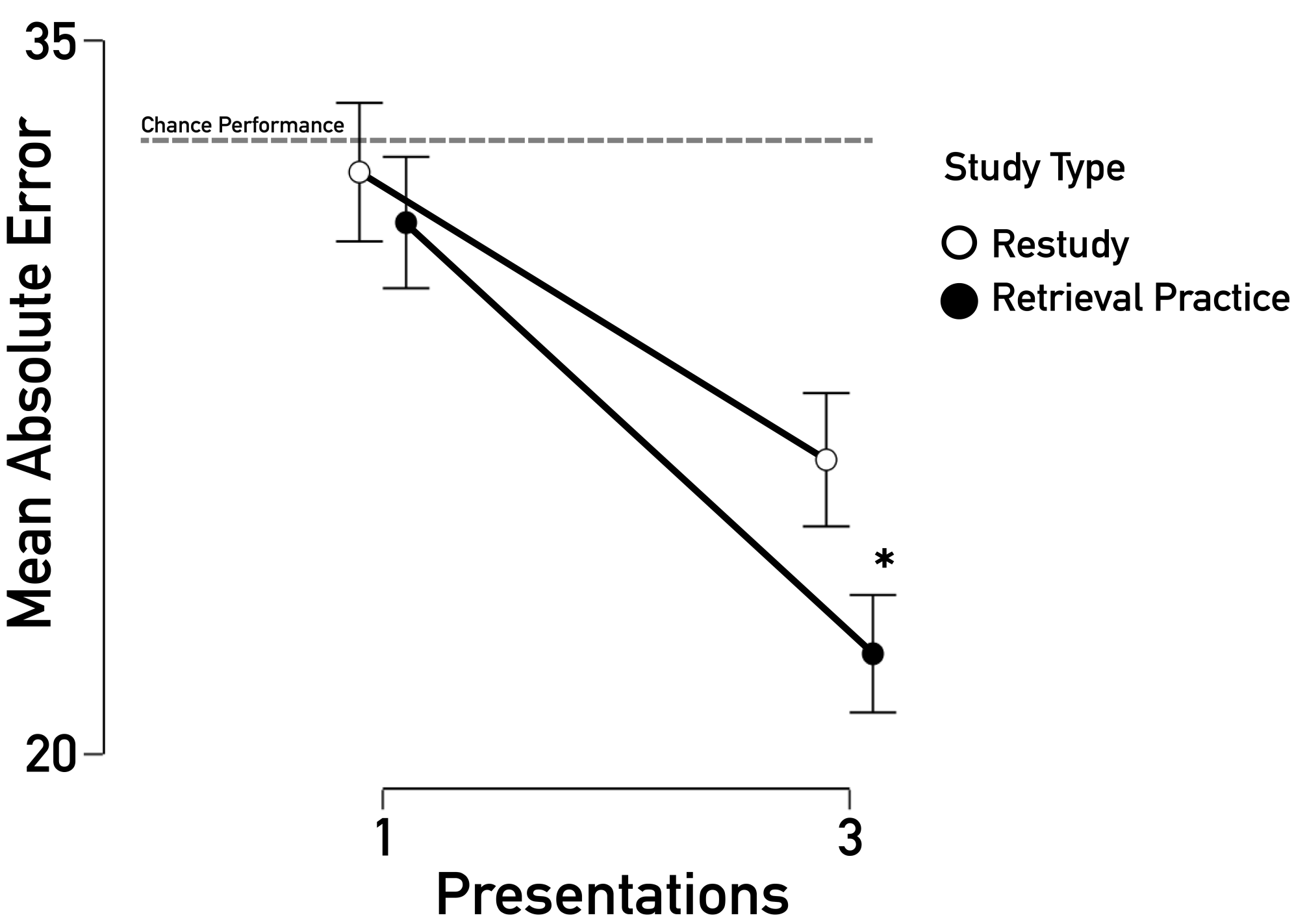


Figure 1. A linear mixed model accounting for between-subjects variability and bar positions found that increased number of presentations reduced mean absolute error ($p < .0001$; $d = 0.23$). There was a significant interaction between number of presentations and study condition ($p = .012$). Max error was equal to 125 units. 95% confidence intervals shown.

All Intervals, Median Width

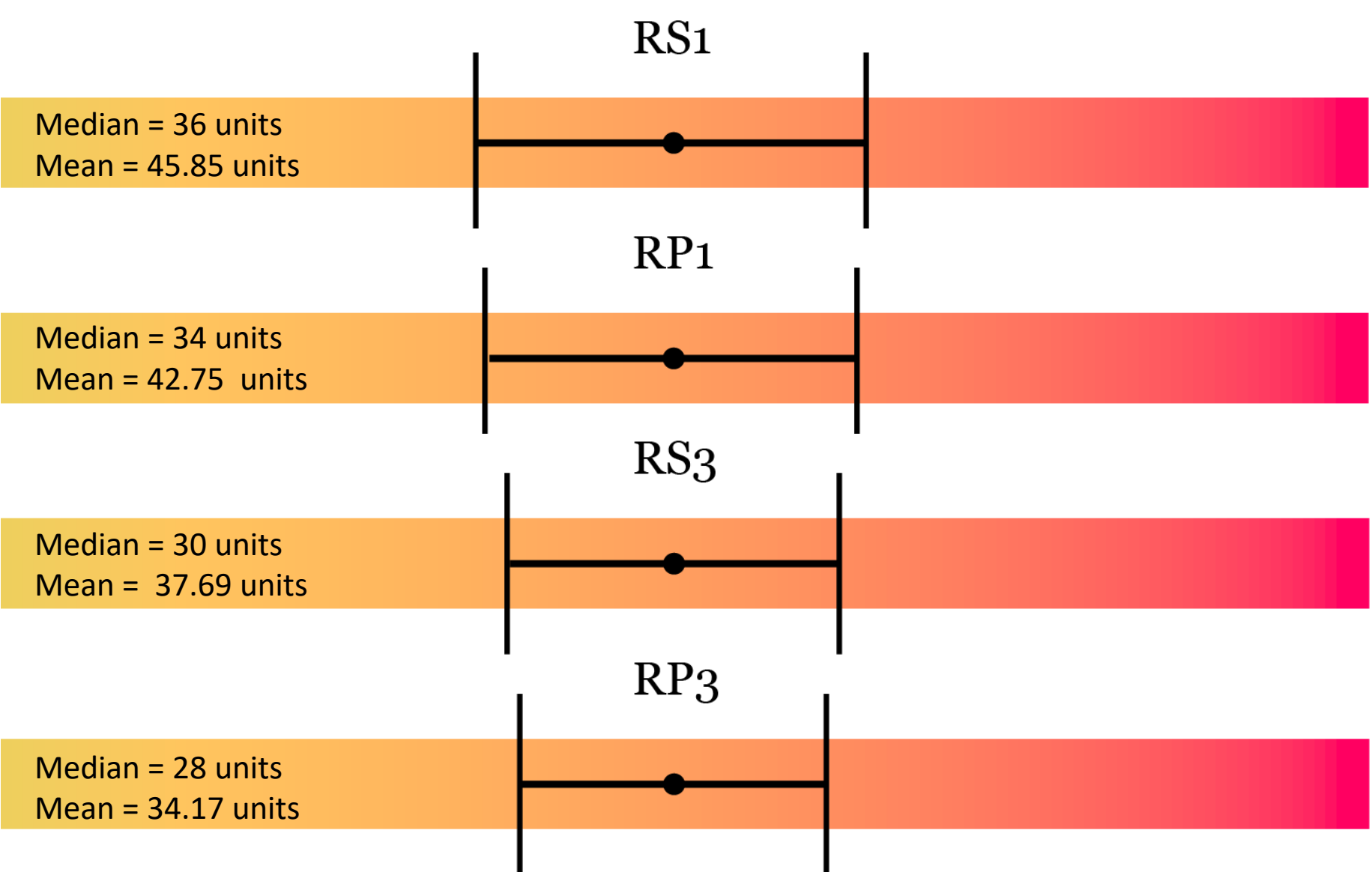


Figure 2. RP ($p = .004$; $d = 0.10$) and greater number of presentations ($p < .0001$, $d = 0.27$) significantly decreased the width of participant generated confidence intervals.

Correct Intervals, Median Width

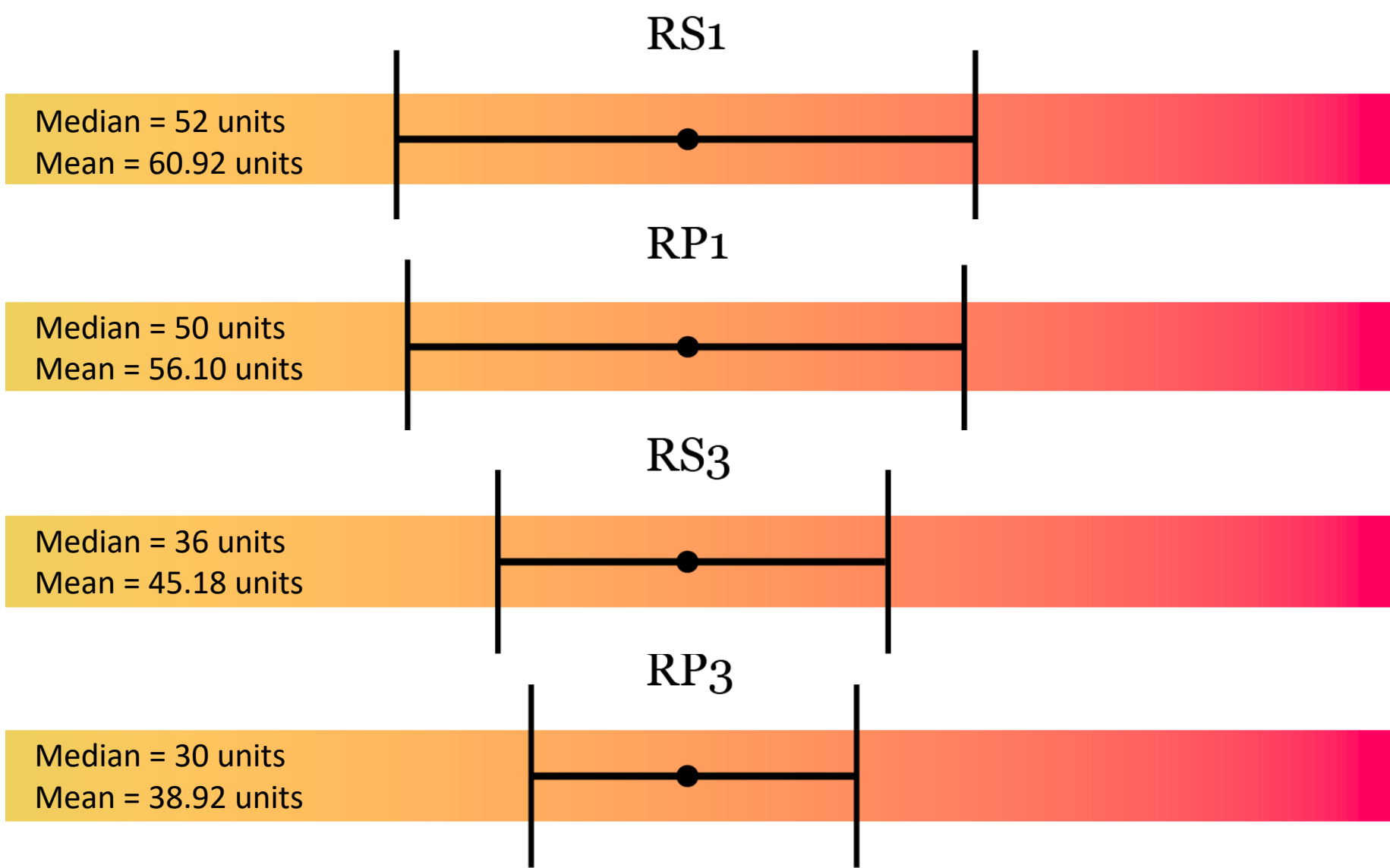


Figure 3. RP ($p = .02$, $d = 0.12$) and greater number of presentations ($p < .0001$, $d = 0.41$) significantly decreased the width of correct participant generated confidence intervals. Intervals were considered correct if the correct answer was located in the given interval.

Proportion Correct

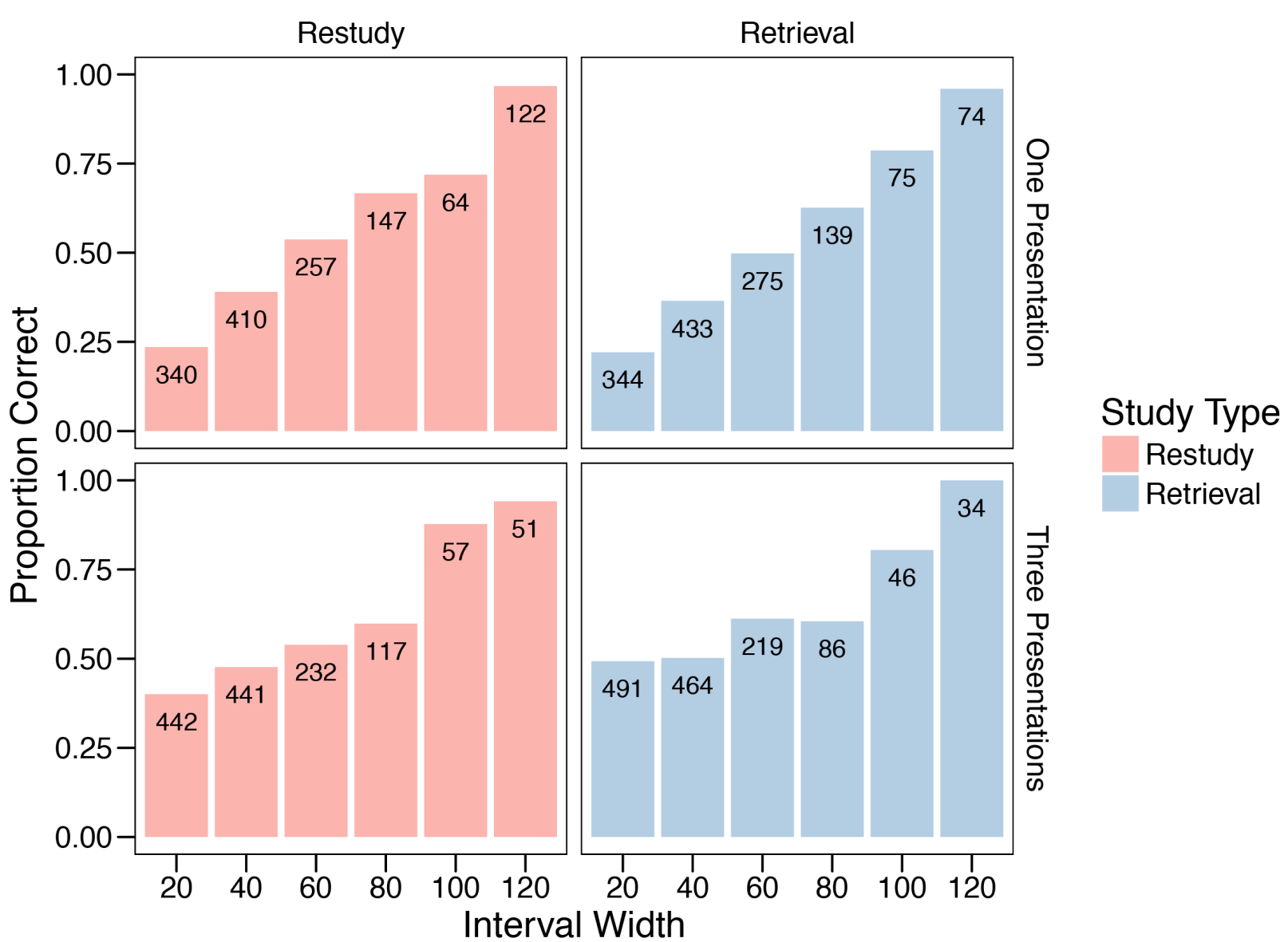


Figure 4. Items in the three-presentation condition were significantly more likely to be assigned correct intervals ($p = .0004$; $d = 0.11$). There was also a significant interaction between study condition and number of presentations ($p = .002$). The numbers within each bar represent the distribution of intervals in each bin. Intervals were considered correct if they contained the correct answer.

Conclusion

Repeated rounds of RP causes decreased error and facilitated the creation of more precise, yet correct intervals. In addition, this research suggests that:

- Learning materials need not be easily verbalizable to be benefitted by RP (c.f. Carpenter, 2011).
- Retrieval practice does not have detrimental effects on participants’ confidence in delayed tests of memory (also shown in a similar experiment by our lab involving response confidence).

Next Steps

Future research will focus on the application of continuous measures to the understanding of metacognition and long-term memory. These studies might involve:

- Retrieval practice without feedback
- Incorrect feedback
- Spatial Stimuli
- Asymmetric Intervals
- Questions involving the estimation of temporal quantities
- Comparison of mixture modeling and metacognitive operationalizations of precision.

Note: Effect sizes (d) were estimated from linear mixed models via the method described in Westfall et al. (2014). Significance test (p) values were computed via Satterthwaite approximation.

References & Acknowledgements

We acknowledge Amanda Chen, Autumn Diggs, Danielle Fang, and Nandita Kasireddy for their help with data collection.

This work was supported by a Kaminsky Undergraduate Research Award.

References

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(6), 1563–1569. <https://doi.org/10.1037/a0017021>

Donkin, C., Nosofsky, R., Gold, J., & Shiffrin, R. (2015). Verbal labeling, gradual decay, and sudden death in visual short-term memory. *Psychonomic Bulletin & Review*, 22(1), 170–178. <https://doi.org/10.3758/s13423-014-0675-5>

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, 103(3), 490.

Sutterer, D. W., & Awh, E. (2016). Retrieval practice enhances the accessibility but not the quality of memory. *Psychonomic Bulletin & Review*, 23(3), 831–841. <https://doi.org/10.3758/s13423-015-0937-x>

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, 143(5), 2020.