

One trial, one boundary!

Validating a rapid task for estimating phonetic category boundaries

Shawn N. Cummings, Xin Xie

UC Irvine

BACKGROUND/INTRODUCTION

Measurement of the 'point of subjective equality' (PSE) between meaningful categories (at the group and/or individual level) is central to theory-building across auditory/speech perception.

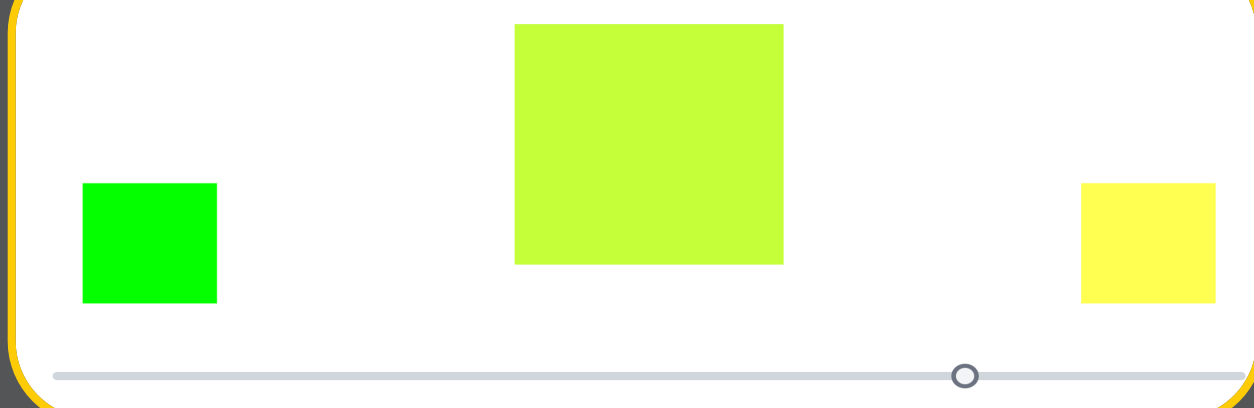
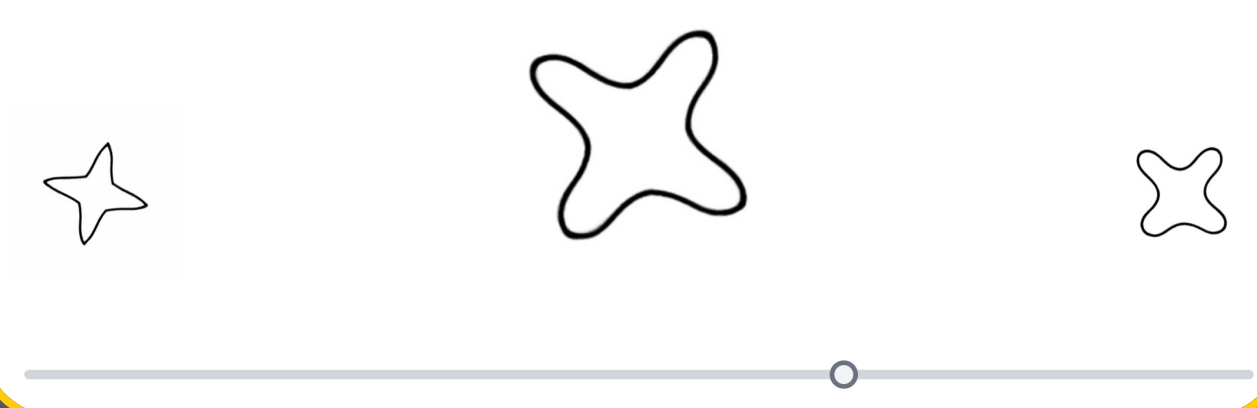
At present, the two-alternative forced choice (2AFC) task serves as the standard tool, despite at least **three known drawbacks**:

1. PSE is *inferred* indirectly
 - 50% response point \leftrightarrow maximal ambiguity
2. Stimulus space is experimenter-determined
 - Unlikely to present *truly* ambiguous acoustics, as most often <10 steps are presented
3. Many repetitions are required
 - Introduces fatigue, potential learning signals, etc.

To ameliorate these, we consider a **new slider-based task**:

Phonetic Boundary Calibration (PBC)

1. Ambiguity is *directly* queried, rather than inferred from p(cat)
2. Listener control the acoustics they hear in a perceptually continuous space (difference between steps < JND)
3. PSE is determinable within a single trial and <30 seconds



ASHI

ASI

Next

PHONETIC BOUNDARY CALIBRATION

- Listeners adjusted a slider between two labeled category endpoints.
- Slider position was sampled continuously (2 Hz) to play tokens from a 50-step perceptually continuous continuum (VOT or %/s/).
- Instructions:
 - move the slider until the sound was "equally like both options."
 - Participants were reminded that the **"perfect blend may not be in the [visual] middle of the slider"**
- Practice trials using blended shapes/colors emphasized that the perceptual midpoint need not align with the visual midpoint.



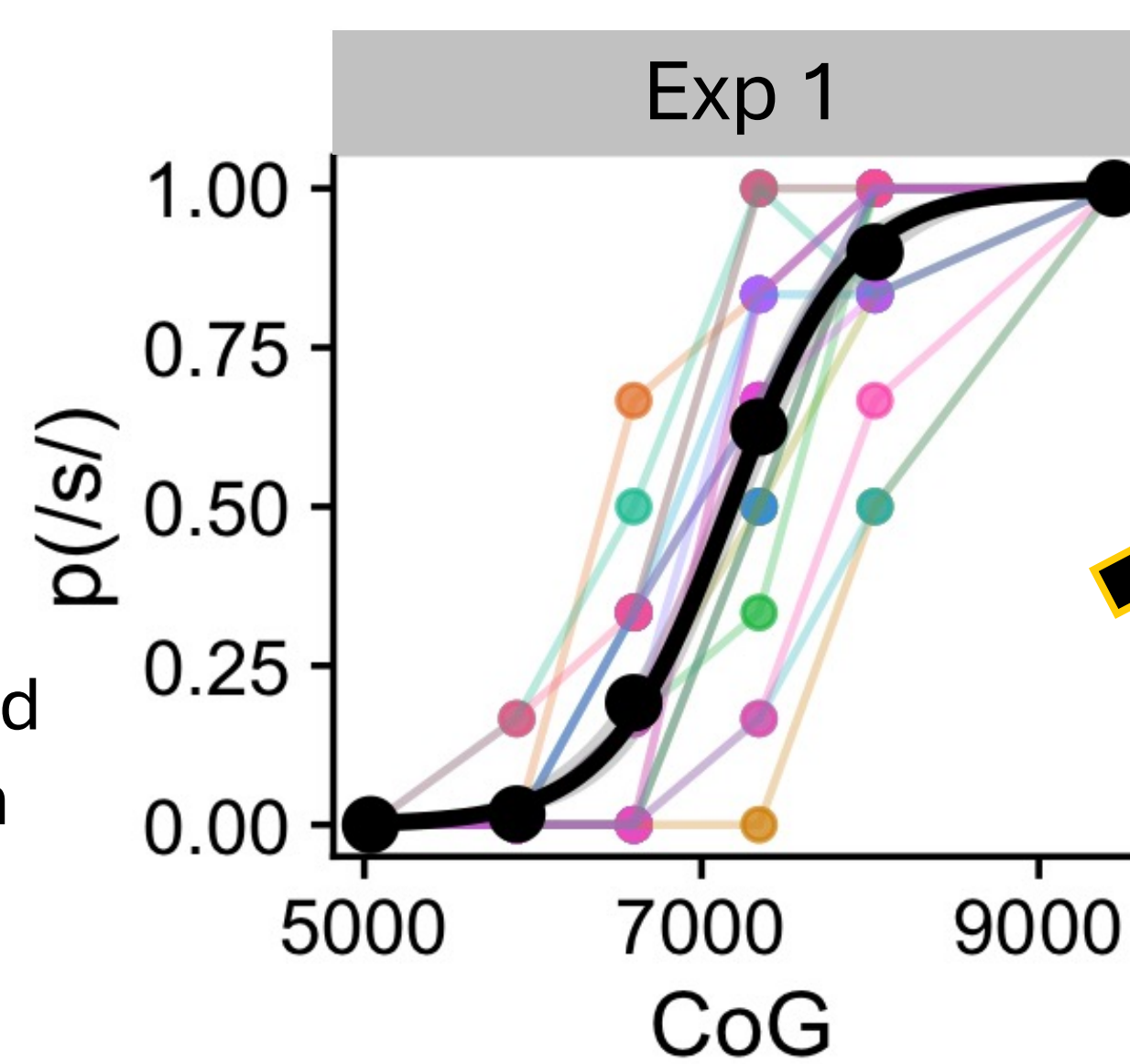
Try the task out!

OVERVIEW OF EXPERIMENTS/METHODS

In each of 3 experiments, listeners performed both the PBC task (5 trials) and standard 2AFC phase (order counterbalanced).

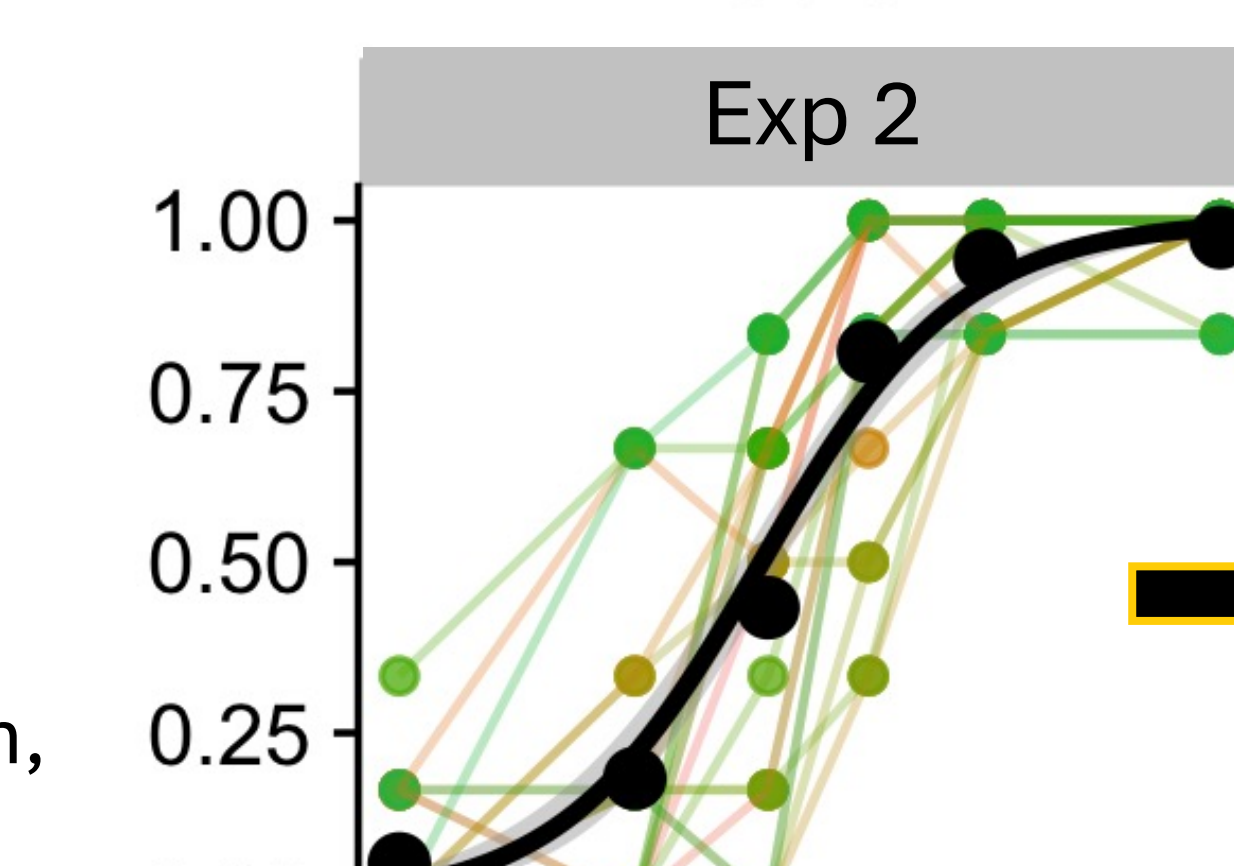
Exp 1 (n=20):

- Fricatives (*asi-ashi*)
- 36 2AFC trials (6 cycles * 6 steps)
 - Tokens from Cummings et al. (2025)
 - Waveform averaged (including duration & amplitude)



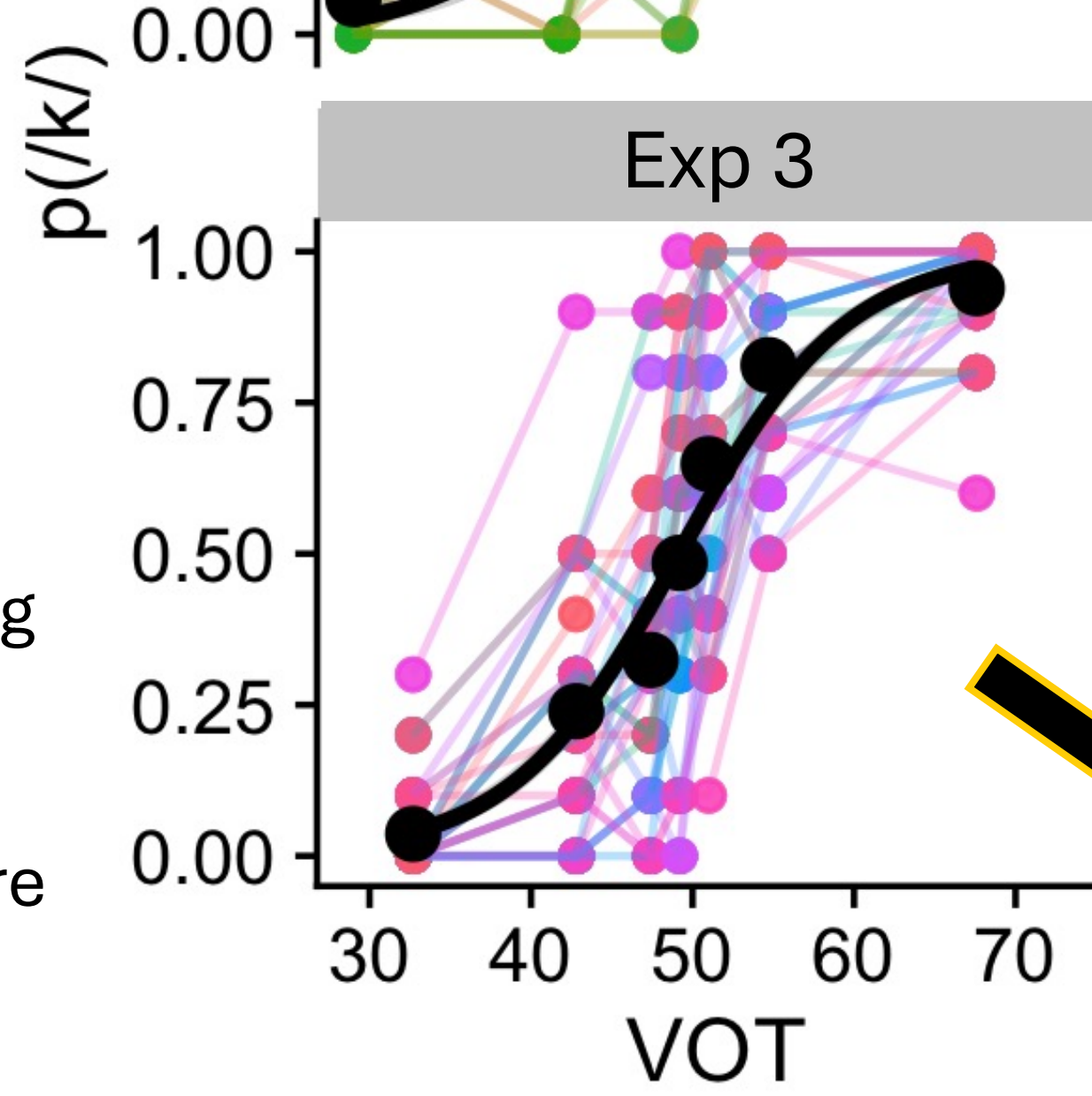
Exp 2 (n=20):

- Stops (*gith-kith*)
- 36 2AFC trials (6 cycles * 6 steps)
 - Tokens from Giovannone & Theodore (2023)
 - VOT-only manipulated (Winn, 2020)

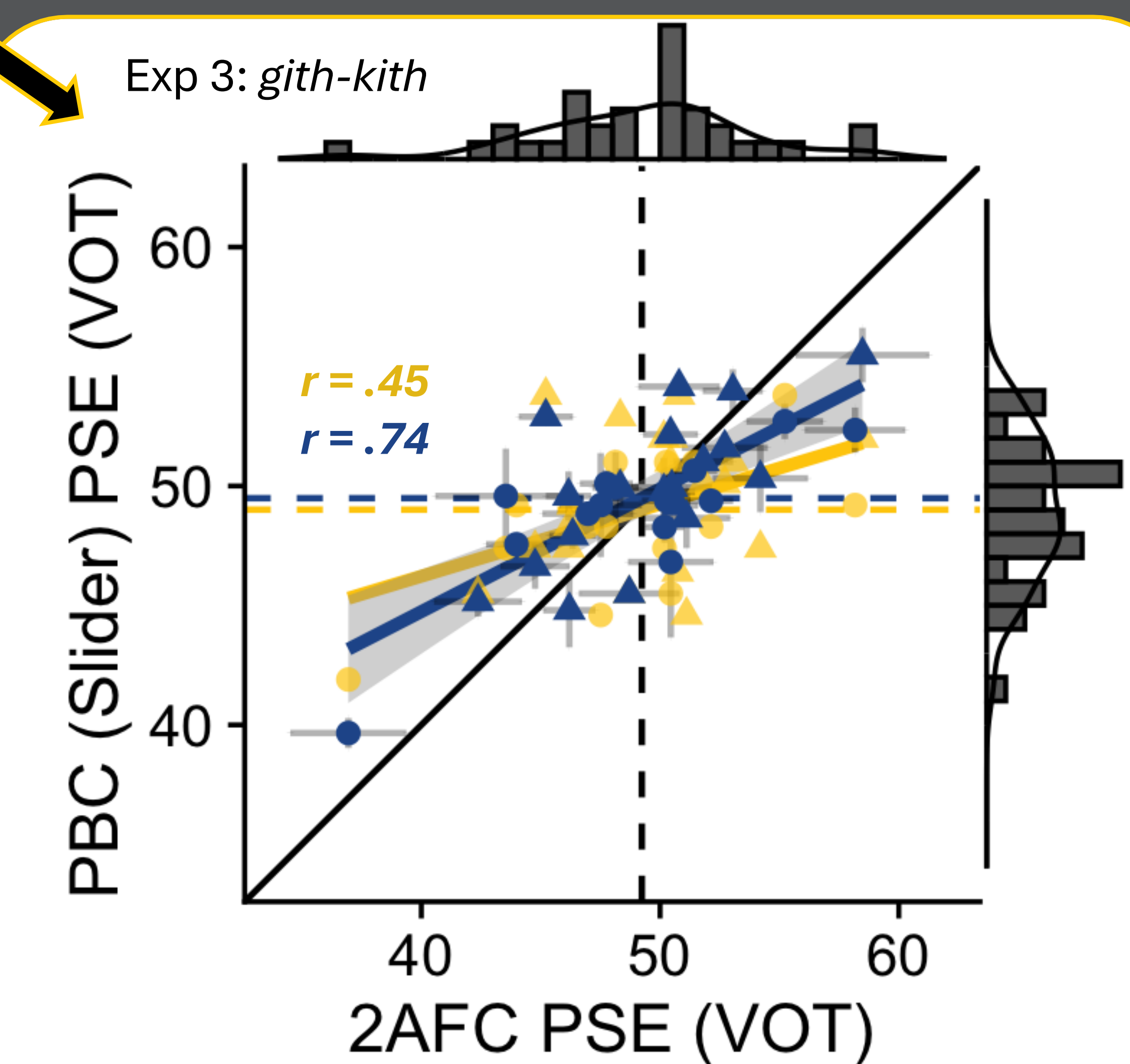
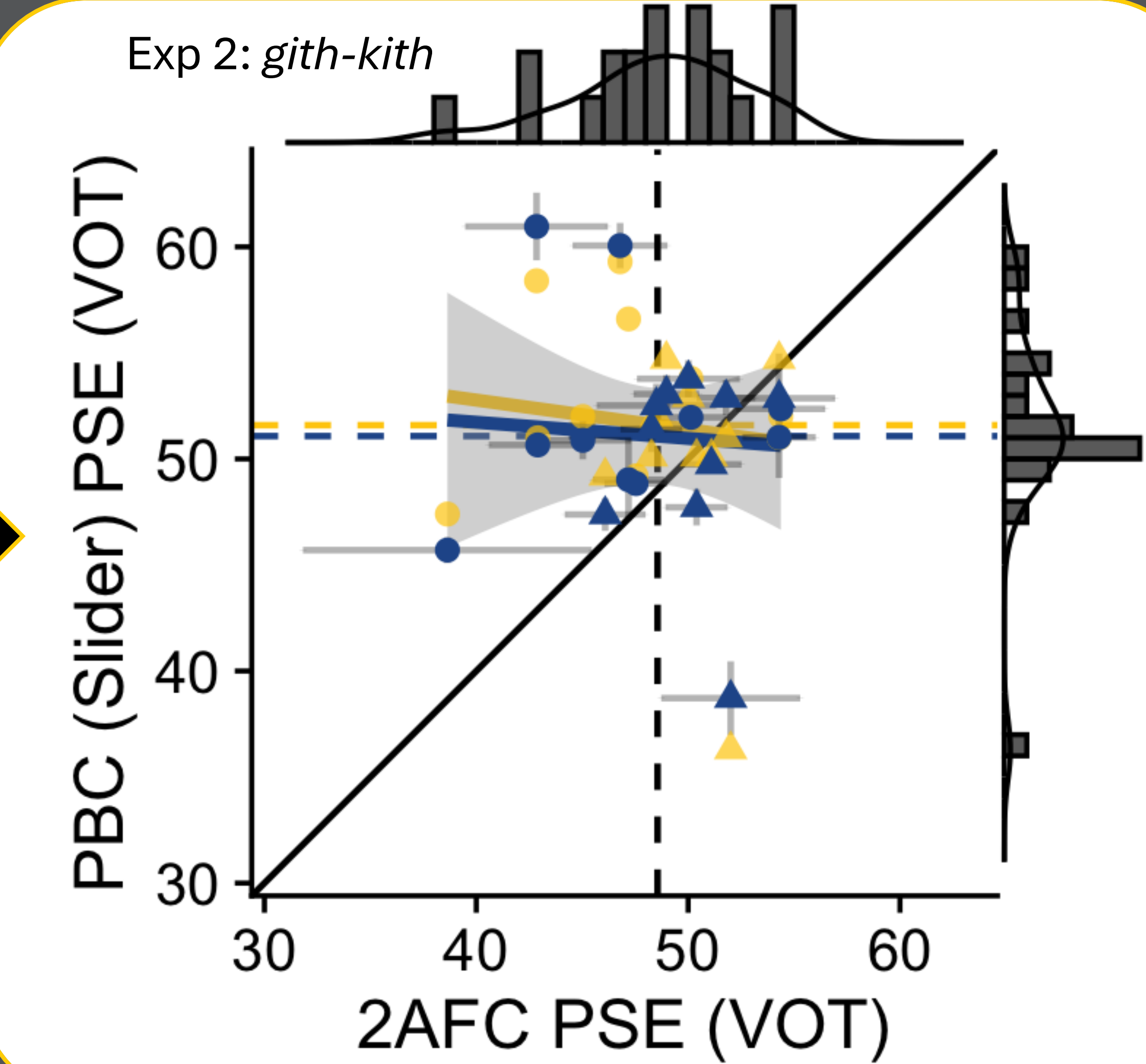
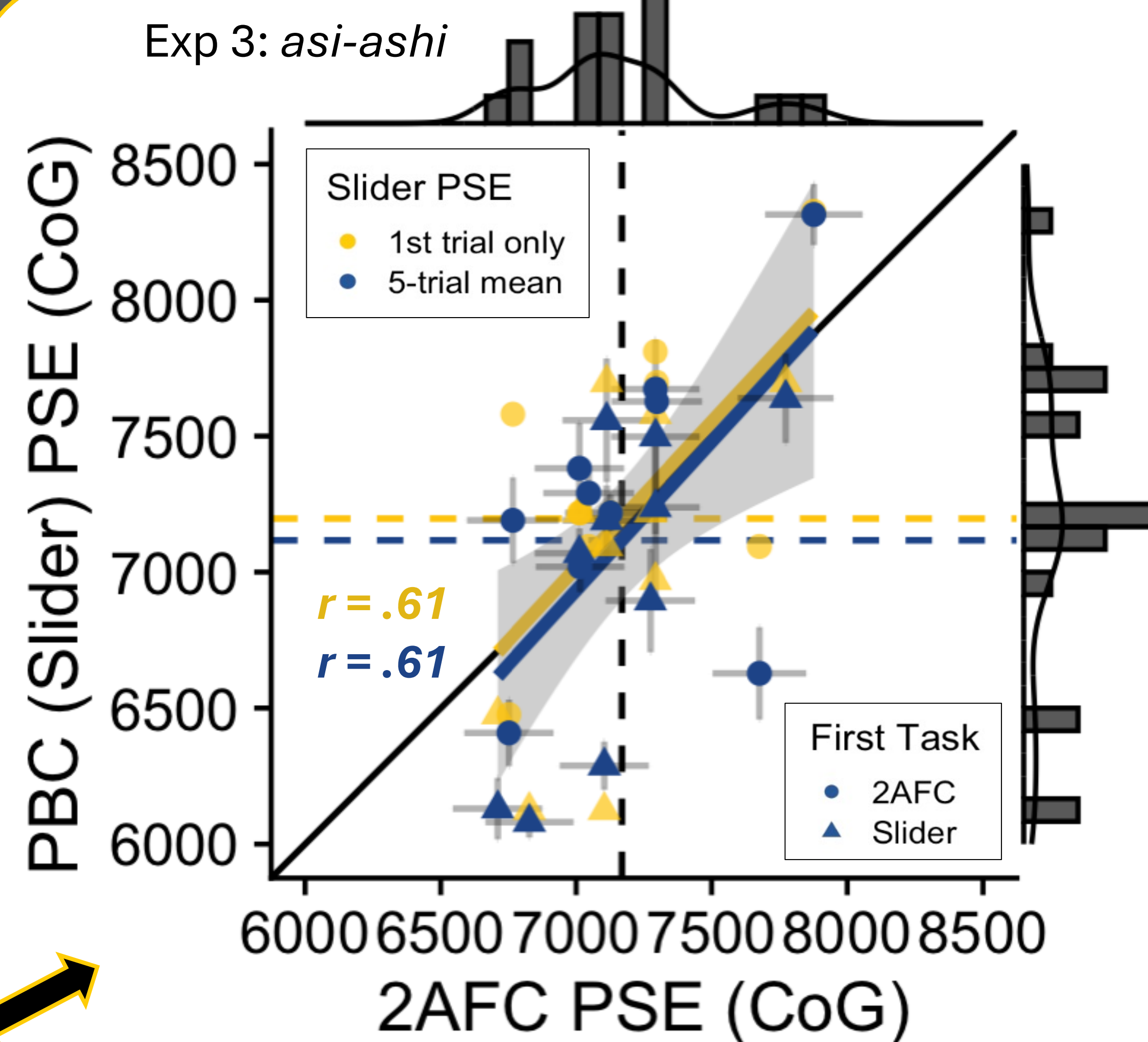
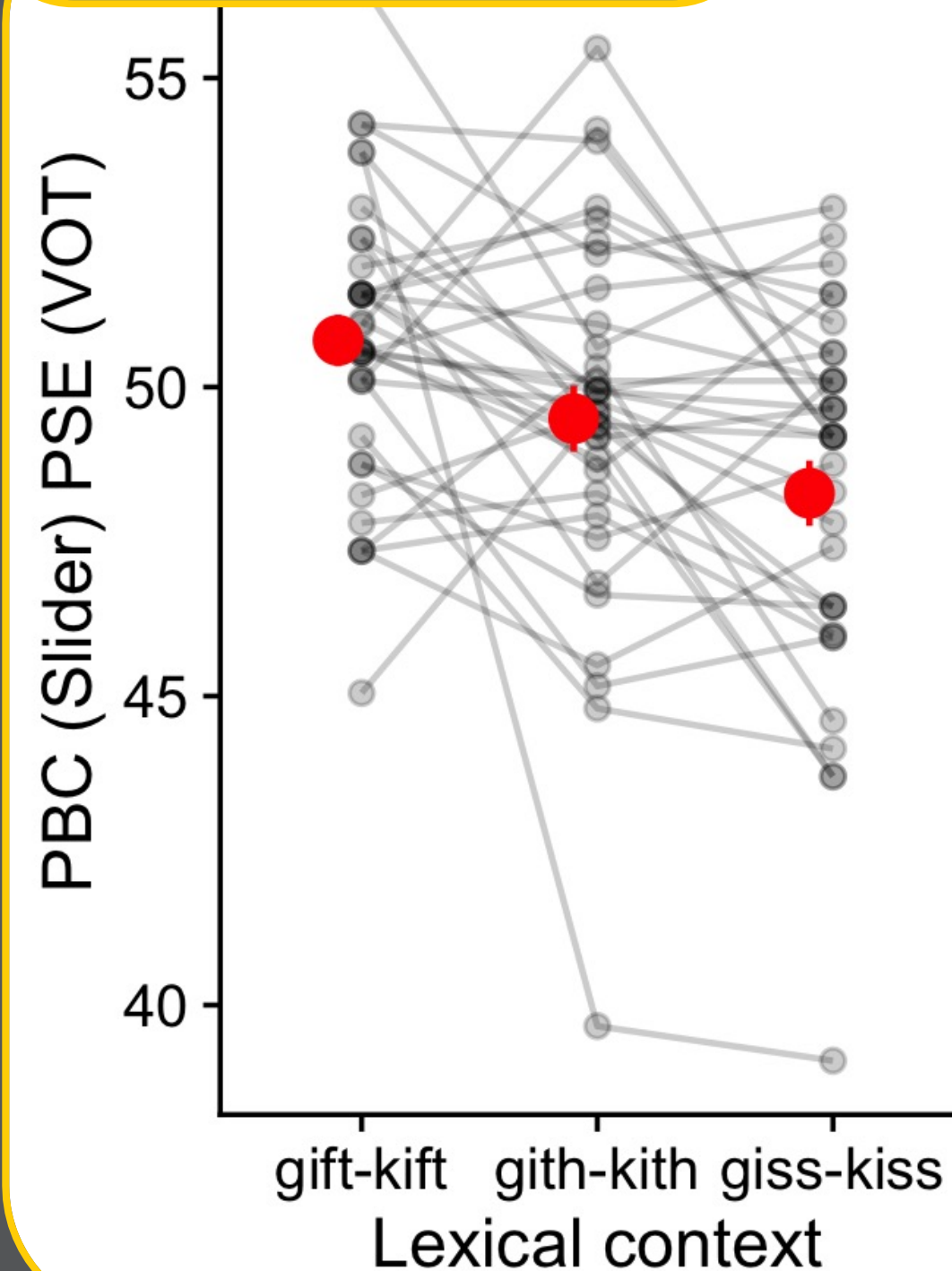


Exp 3 (n=33):

- Stops (*gith-kith*)
- 70 2AFC trials (10 cycles * 7 steps)
 - Added (after both main phases) an exploratory Ganong slider task, identical except that endpoints were *gift-kift* or *giss-kiss* (2 trials each)



GANONG EFFECT



A novel, few (1-5) trial task promises *rapid and high fidelity* assessment of phonetic category boundaries at the group and individual levels!

Across both stop and fricative contrasts, we establish:

- Task reliability
- Slider split half (SB corrected) > 0.92
- 2AFC split half *variable* across studies: E1 = 0.77, E2 = 0.415, E3 = 0.909
- Group-level agreement
- Mean PSE's via the two tasks did not significantly differ via paired t-test
- Estimated differences in PSE: 51 Hz CoG, 2.5 ms VOT, and 0.3 ms VOT
- Individual-level agreement
 - Significant correlations except in the case of poor 2AFC reliability
- Sensitivity to theoretically-important boundary shifts (Ganong)

REFERENCES & ACKNOWLEDGEMENTS

Cummings, S.; Duda, B.; & Theodore, R. 2025. *Autism Research*
Giovannone, N. & Theodore, R. 2023. *Cognition*.
Li, A.Y. et al. 2020. *JEP:G*.
Winn, M. 2021. *JASA*.

This research was supported by NIDCD R01HD111936