

Social desirability bias distorts self-reports of mind-wandering



Shivang Shelat^{1,2}, Alexis R. Clubb¹, Jonathan W. Schooler¹, Barry Giesbrecht^{1,2}

¹Department of Psychological & Brain Sciences, University of California, Santa Barbara

²Institute for Collaborative Biotechnologies, University of California, Santa Barbara

Background

Mind-wandering is studied by using probes to collect **task-unrelated thought (TUT)** reports.

...but self-reports can be susceptible to **social desirability bias (SDB)**.

TUT correlates with objective attention measures (e.g., **response times**; RTs).

In two experiments, we ask:

Is SDB present in subjective TUT reports?

Does it obscure objective–subjective coupling?

If so, how can we account for it?

How do we measure SDB?

If a participant claims many overly desirable traits, they score high in SDB.

Validated scales¹ get at this bias with lure items.

Example items (true/false):

I have never intensely disliked anyone.

I'm always willing to admit it when I make a mistake.

No matter who I'm talking to, I'm always a good listener.

Previous work also suggests SDB increases under time pressure.²

Analysis

Logistic / linear mixed-effects regression

Subject ID modeled as a random effect

* $p < .05$, ** $p < .01$, *** $p < .0001$

Individual data points are jittered on all figures

Contact:

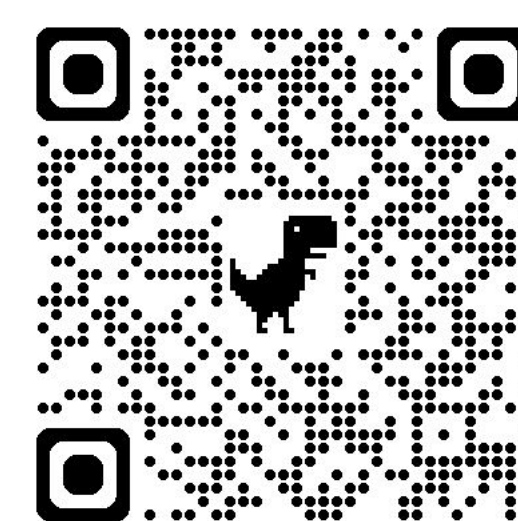
sshelat@ucsb.edu

alexiscclubb@ucsb.edu

jschooler@ucsb.edu

giesbrecht@ucsb.edu

preprint + website



¹Vésteinsdóttir et al. (2017). An item level evaluation of the Marlowe-Crowne Social Desirability Scale using item response theory on Icelandic Internet panel data and cognitive interviews. *Personality and Individual Differences*.

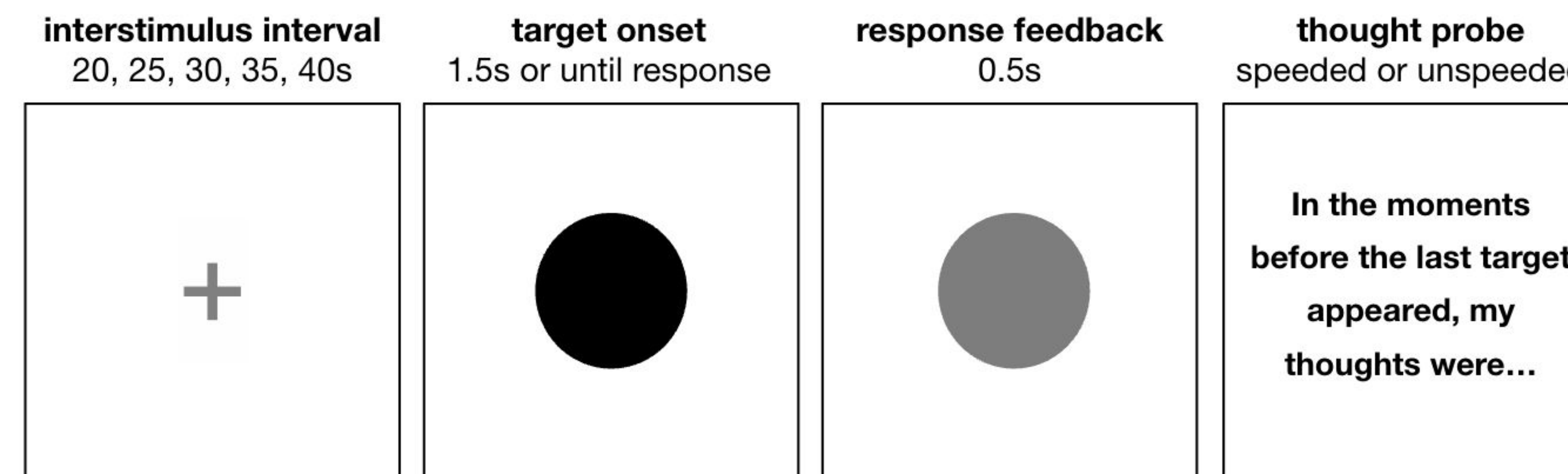
²Protzko et al. (2019). Rushing to appear virtuous: Time pressure increases socially desirable responding. *Psychological Science*.

³Shelat et al. (2024). Predicting attentional lapses using response time speed in continuous performance tasks. *Frontiers in Cognition*.

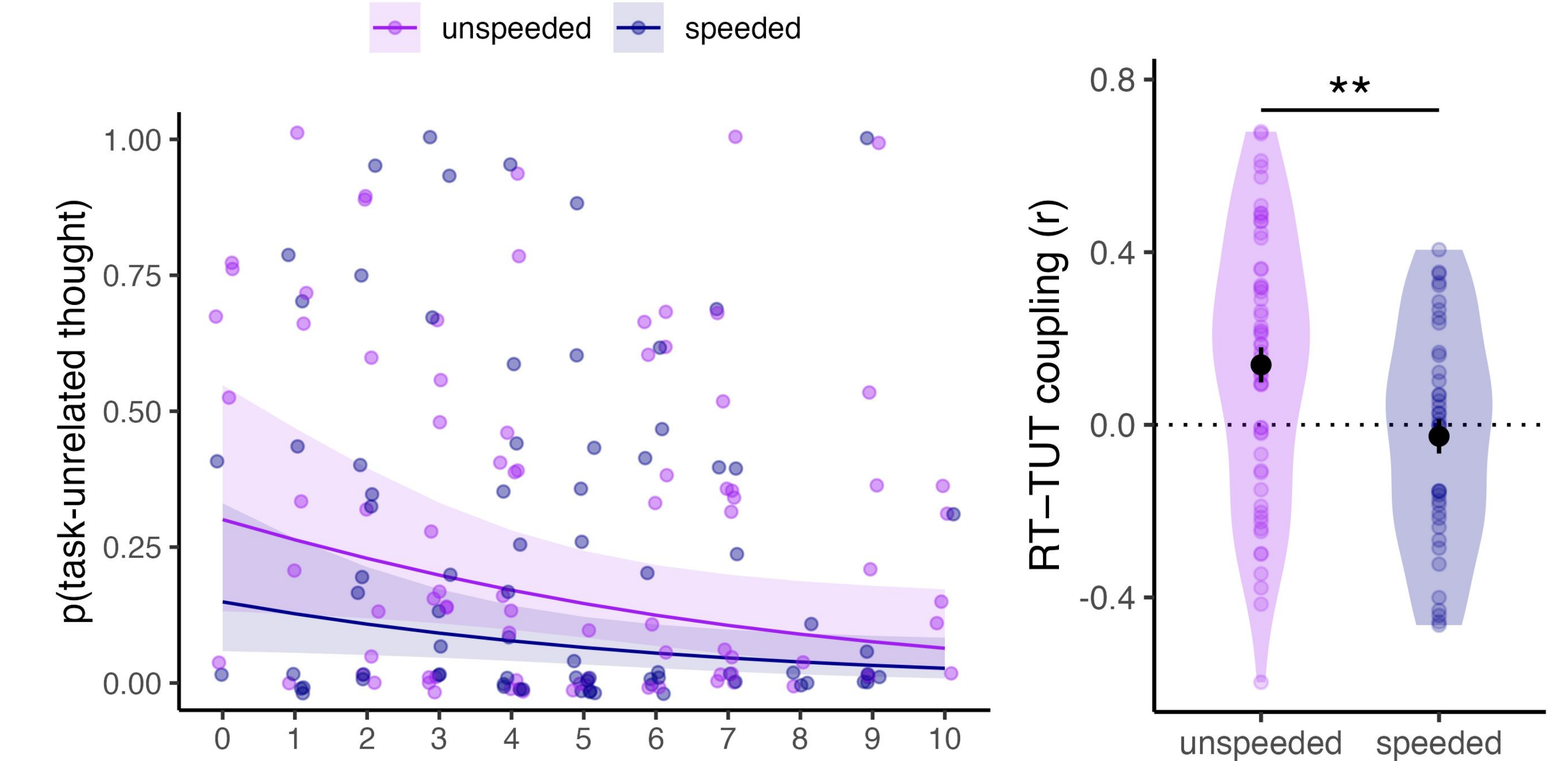
⁴deBettencourt et al. (2019). Real-time triggering reveals concurrent lapses of attention and working memory. *Nature Human Behaviour*.

Experiment 1: Dot Reaction Task

Goal: test the link between SDB and TUT, and explore time pressure as a causal manipulation to amplify SDB²



slow RTs = objective TUT; 1 block × 20 trials, binary thought probes
162 online participants: $M_{age} = 36.1 \pm 10.8$ years, 80 female, 79 male, 3 other

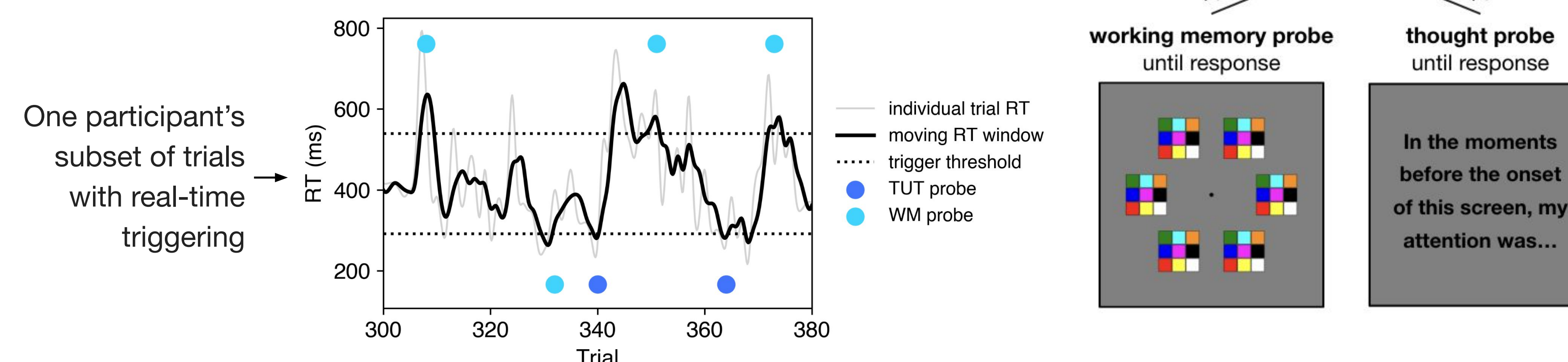
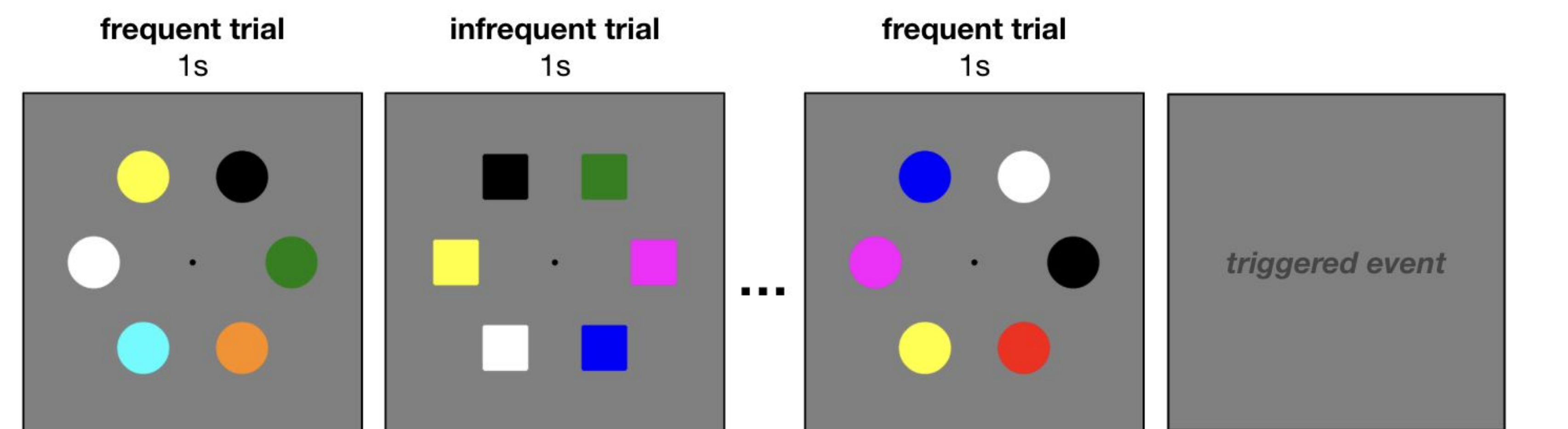


SDB was linked with **fewer TUT reports**, but **not faster RTs**

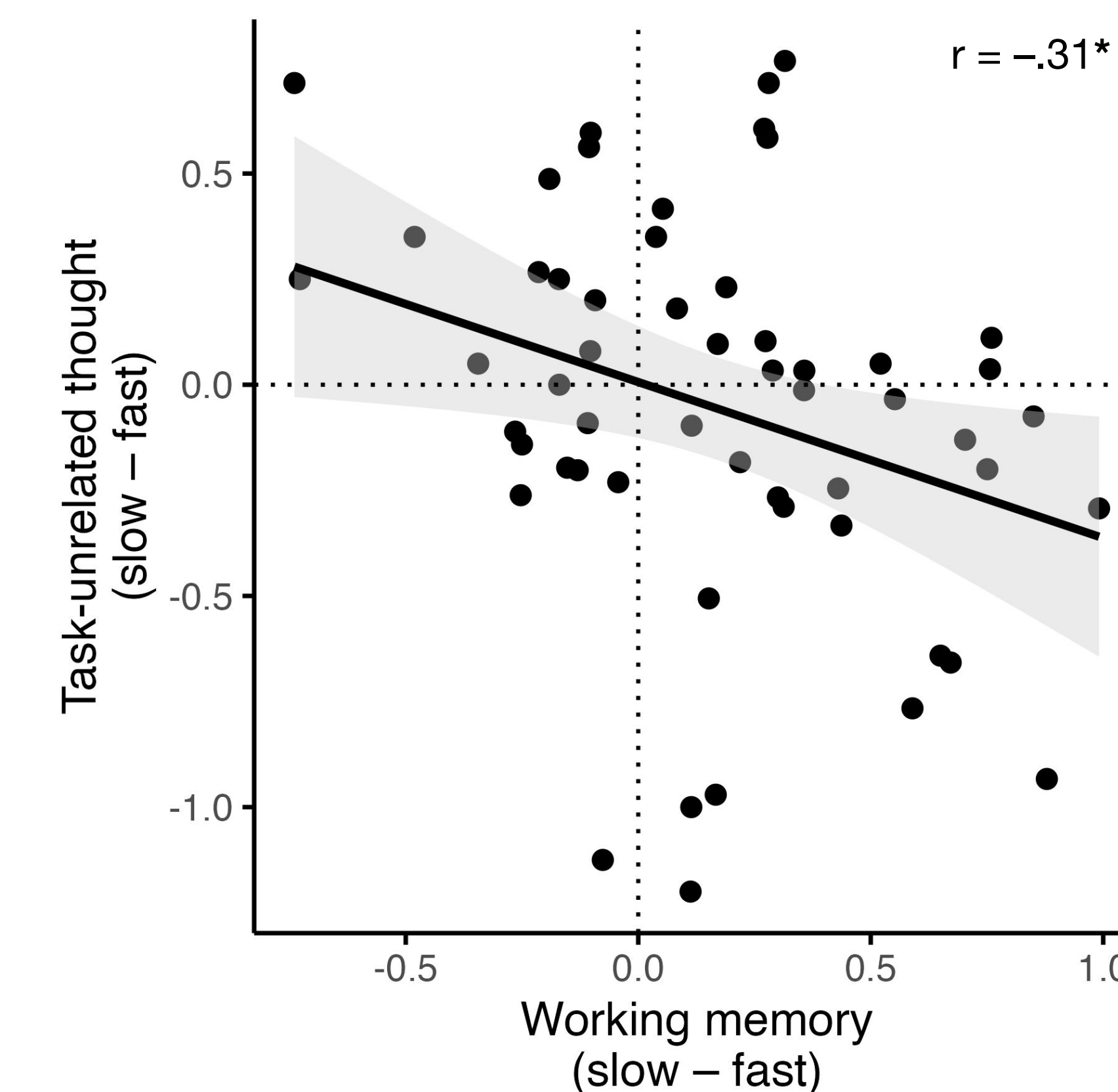
Time pressure **decoupled TUT and RTs**, but no interaction with SDB

Experiment 2: Sustained Attention to Response Task

Goal: test the link between SDB and TUT with sharper sampling across attentional states via real-time triggering^{3,4}



fast triggers = objective TUT; 3 blocks × 600 trials each, 7-point depth thought probes
55 in-person participants: $M_{age} = 19.9 \pm 1.9$ years, 44 female, 8 male, 3 other



Trigger effectiveness for different lapse indices was **moderately correlated**

Fast (vs. slow) triggers captured **deeper TUT**, but only when **controlling for SDB**

Summary

It's there! SDB was linked with suppressed TUT reports, suggesting that self-presentation biases affect self-reports of attentional state.

It's robust! This link was present across two different tasks, samples, settings (online vs. in-person), and probe formats (binary vs. depth).

It's fixable! We can statistically control for SDB and improve thought probe validity by accounting for 10-item survey scores in regression.