

# Speech exposure familiarizes listeners with talkers' vocal identity

Jayden J. Lee & Tyler K. Perrachione

Department of Speech, Language, & Hearing Sciences, Boston University



## Introduction

•Talker identification experiments typically explicitly train listeners to identify voices.<sup>[1-4]</sup> Ecologically, however, listeners learn to identify talkers without explicit practice.

•Exposure sans practice has been shown to be effective for perceptual learning<sup>[5][6]</sup> but it is unknown whether the same applies for higher-order learning, like voice identification.<sup>[7]</sup>

### Research Questions:

- Can listeners gain familiarity with voices even when they are not explicitly identifying them?
- When exposed to the same speech, do listeners still learn vocal information if they are differentially directing their attention between talkers' vocal identity and verbal content?

## Methods
















**Participants:** Native speakers of American English (N = 96; 74 female, 22 male; ages 18-31; mean = 20.5 years) who reported no history of speech, hearing, or language disorder

**Stimuli:** Recordings of ten digits (0-9) presented in five-digit strings by 10 female and 10 male American adult speakers. Identical across all exposure task conditions and test.

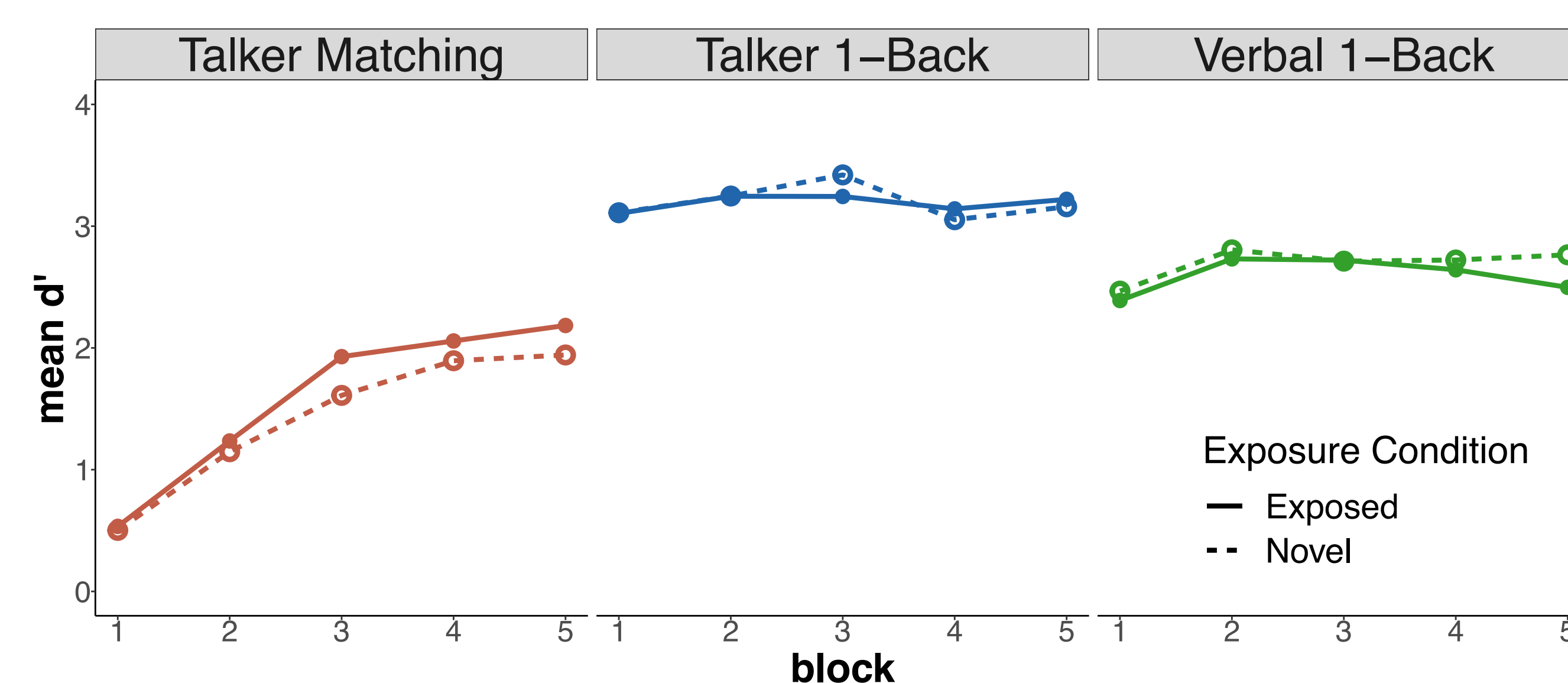
**Procedure:** Each participant was randomly assigned to one of three task conditions for the exposure phase (200 trials of five-digit sequences) with feedback. All participants were then tested for talker identification ability (50 trials of five new digits with feedback).

- **Talker Matching Task:** Identify voices in an active training paradigm (matching voice heard with avatar)
- **Talker 1-Back Task:** Indicate whether the talker on each trial was the same as the talker from the previous trial
- **Verbal 1-Back Task:** Indicate whether the middle digit on each trial was the same as the middle digit from the previous trial

### Task Design:

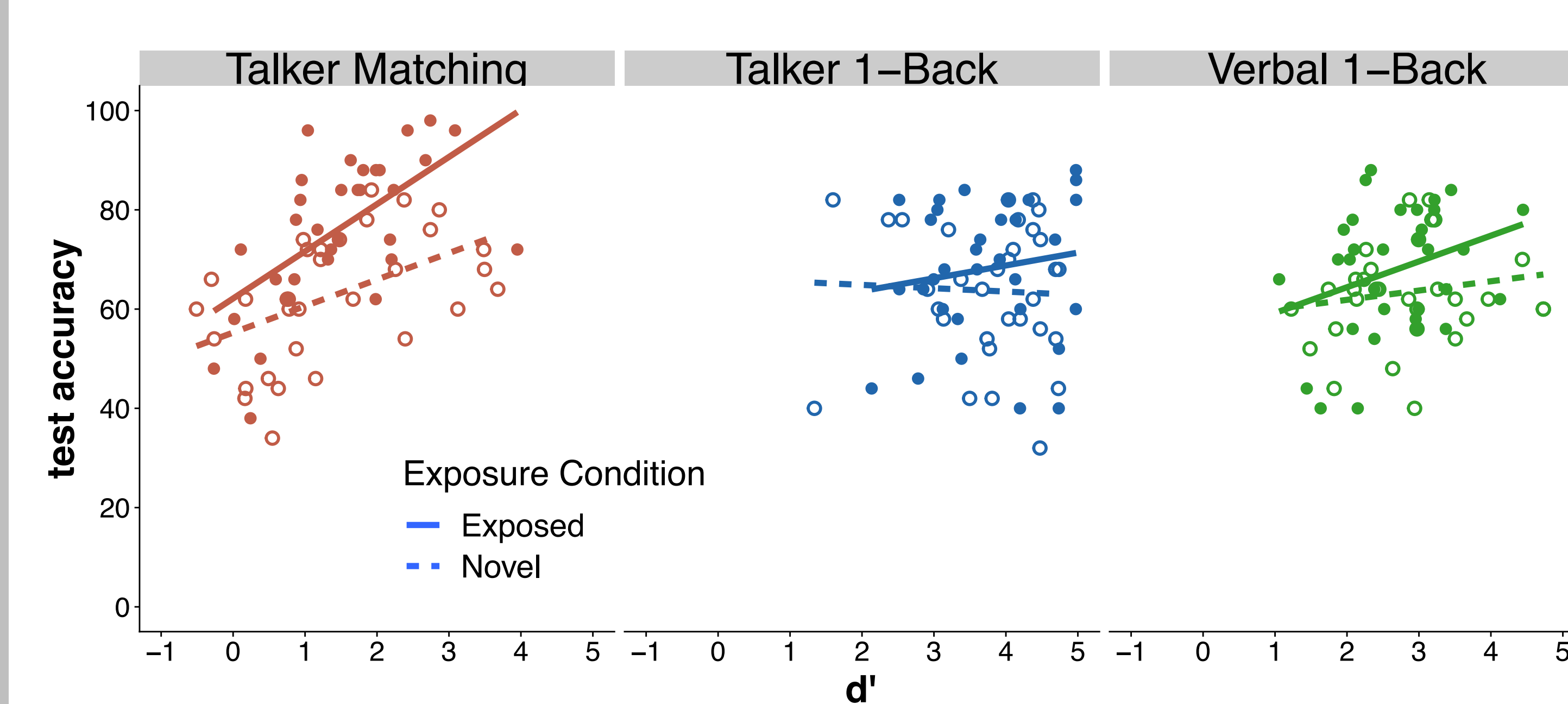
Trial 1	 "8, 3, 1, 4, 7" 	Response
		Talker Matching 
		Talker 1-Back 
Trial 2	 "3, 7, 4, 8, 1" 	Response
		Talker Matching 
		Talker 1-Back 
Trial 3	 "1, 7, 4, 3, 8" 	Response
		Talker Matching 
		Talker 1-Back 
.	 	Response
		Verbal 1-Back 

## Exposure Phase Results



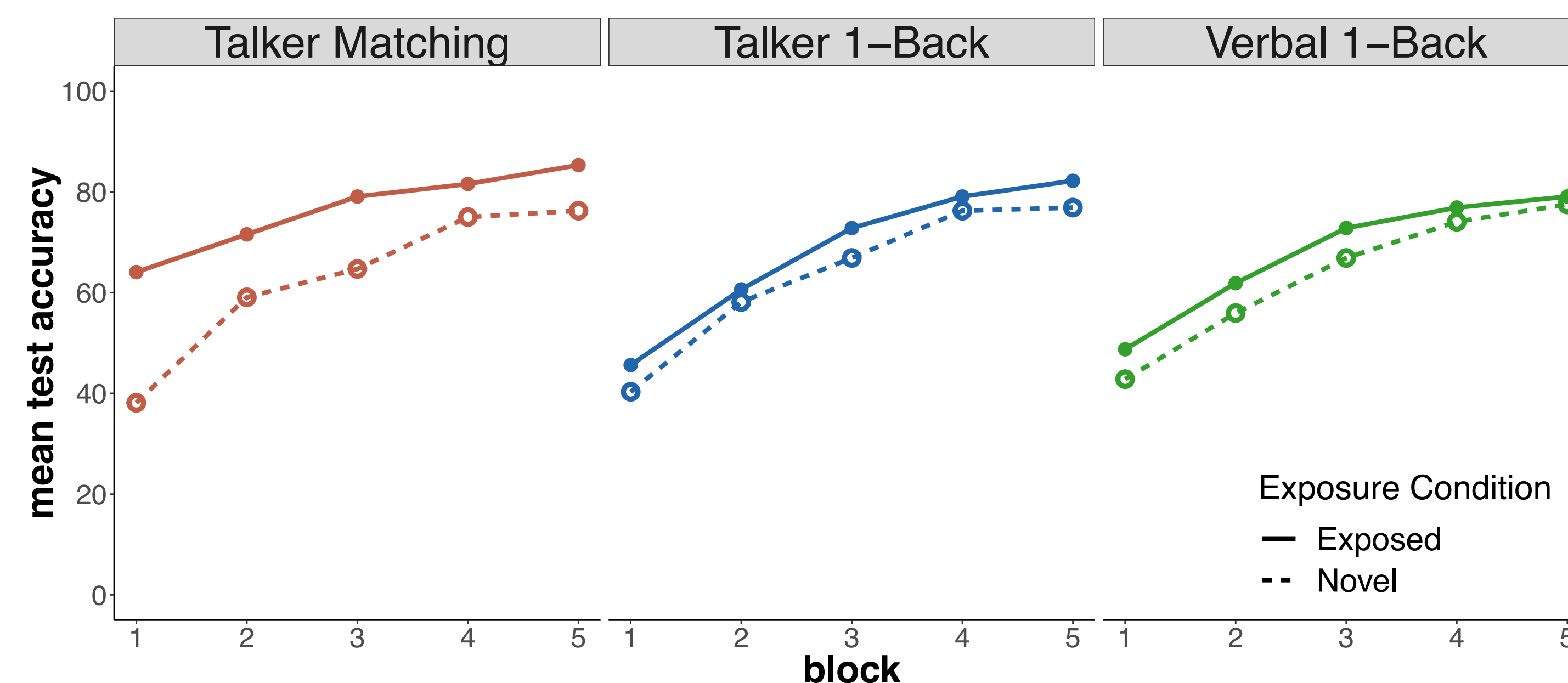
**Fig. 1:** Mean task performance accuracy ( $d'$  score) for all exposure task conditions, divided into blocks of 40 trials (200 total). Main effects of *task* ( $\chi^2(2) = 124.70, p < 0.0001$ ) and *block* ( $\chi^2(1) = 99.72, p < 0.0001$ ). The main effect of *exposure condition* was not significant ( $\chi^2(1) = 0.07, p = 0.79$ ). Significant *task*  $\times$  *block* interaction ( $\chi^2(2) = 155.49, p < 0.0001$ ) as well.

## Correlations

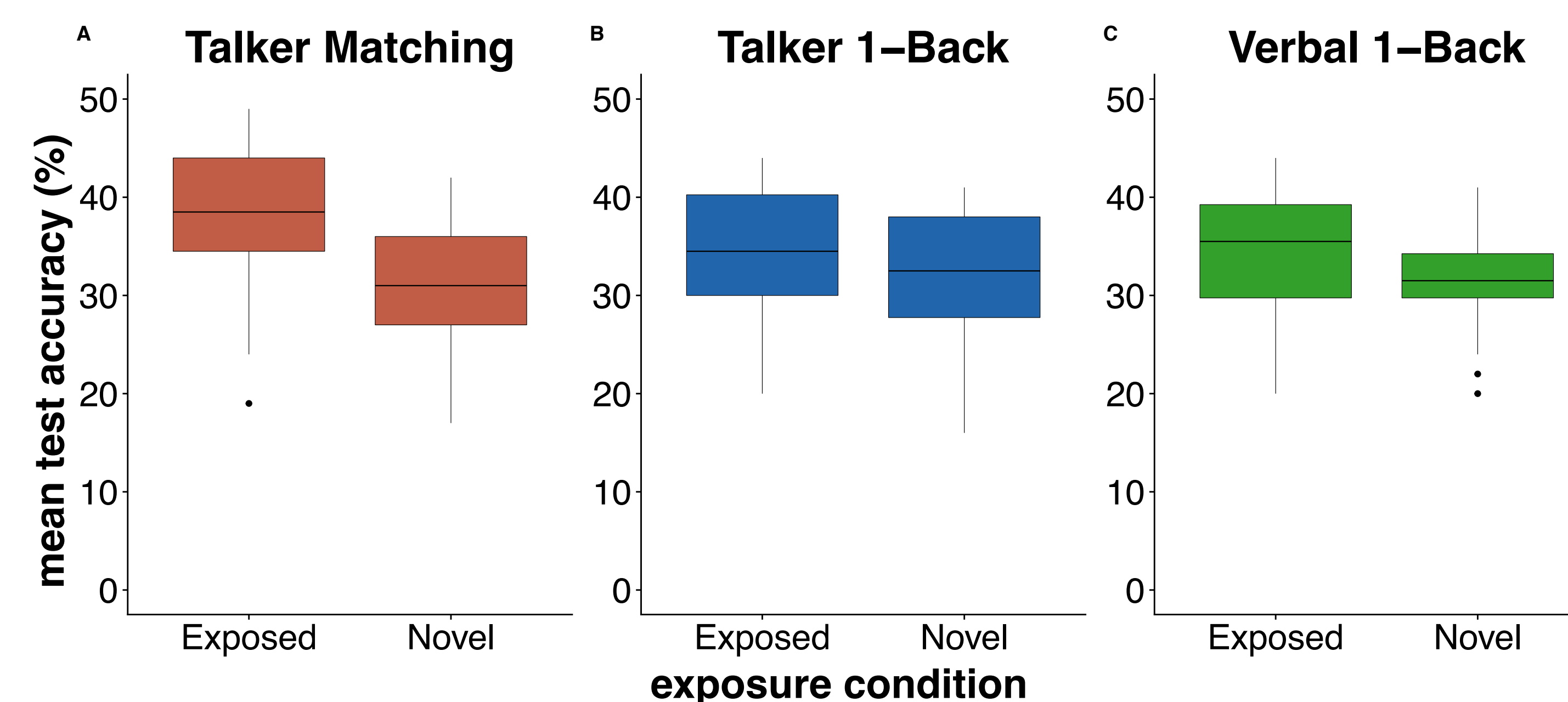


**Fig. 4:** Relationship between exposure task performance ( $d'$  score) and test accuracy (percentage correct). The only significant correlations between exposure and test performance were for the Talker Matching task, for which the Pearson's correlation coefficients were  $r = 0.61$  ( $p < .001$ ) for the exposed voices and  $r = 0.49$  ( $p = 0.004$ ) for the novel. Pairwise correlations for both the Talker 1-Back ('exposed'  $r = -0.04$  ( $p = 0.81$ ), 'novel'  $r = 0.15$  ( $p = 0.41$ )) and Verbal 1-Back ('exposed'  $r = 0.31$  ( $p = 0.09$ ), 'novel'  $r = 0.16$  ( $p = 0.37$ )) tasks were not significant.

## Test Phase Results



**Fig. 3:** Mean test performance accuracy (50 response questions) for all exposure task conditions, divided into blocks of 10 trials. Type-III ANOVA on gLME for binomial data found significant main effects of *exposure condition* ( $\chi^2(1) = 28.62, p < 0.0001$ ) and *block* ( $\chi^2(1) = 594.76, p < 0.0001$ ) and significant *task*  $\times$  *exposure condition* interaction ( $\chi^2(2) = 9.64, p < 0.009$ ).



**Fig. 4:** Mean test accuracy across the exposure task conditions (boxplots). For the talker matching task, there was a significant effect of *exposure condition* ( $\chi^2(1) = 29.30, p < 0.0001$ ). For the talker 1-back task, *exposure condition* did not account for a significant amount of the variance ( $\chi^2(1) = 2.88, p = 0.09$ ). This factor was also significant in the verbal 1-back task ( $\chi^2(1) = 4.06, p < 0.05$ ).

## Discussion

- Prior exposure to voices enhances accuracy in talker identification.
- Listeners were significantly more accurate at identifying talkers if they had been previously exposed to those talkers versus novel talkers.
- Those who were assigned the Talker Matching task during the exposure phase were the most accurate among the three task conditions.
- Since the group that practiced identifying talkers during the exposure phase was the most accurate, affirming that explicit talker ID practice was the most effective task.
- The Verbal 1-Back group also performed significantly better at the talker identification test when given exposed voices rather than novel talkers, even though listeners were attending to the speech content.
- Interestingly, the Talker 1-Back group did not show any significant improvement at talker identification when given exposed voices.
- These results suggest that it is possible for listeners to unknowingly learn talkers' vocal identity during speech perception while focusing on a separate task and without explicit practice.

## References

- [1] Nygaard & Pisoni (1998). *Perception & Psychophysics*, 60(3): 355-376.
- [2] Perrachione & Wong (2007). *Neuropsychologia*, 45: 1899-1910.
- [3] Bregman & Creel (2014). *Cognition*, 130(1): 85-95.
- [4] Orena, Theodore, & Polka (2015). *Cognition*, 143: 36-40.
- [5] Wright et al. (2010) *Journal of Neuroscience*, 30(38): 12868-12877.
- [6] Szpiro, Wright, & Carrasco (2014). *Vision Research*, 101(2014): 118-124.
- [7] Wright et al. (2015) *Journal of Acoustical Society of America*, 138(2): 928-937.

## Acknowledgements

We thank Jessica Tin, Yaminah Carter, Terri Scott, Ja Young Choi, & Sung-Joo Lim for their help as well as every research assistant who assisted in running the experiment and collecting data. This research was supported by the NIH T32DC013017 training grant and NIH NIDCD R03DC014045 to T.P.