

Semantic memories: Image memorability transfers to scene descriptions

Diana Kollenda*, Sophie Schwenger, Sophie Halstenberg & Ben de Haas

*Diana.Kollenda@gmail.com | INDIVISUAL lab: <https://www.individual-perception.com/>

INTRODUCTION

low



„Two people walk along the beach while the sun sets in bright orange on the horizon over the sea.“

The intrinsic image property of **Image Memorability** ... (cf. Bainbridge, 2019)

... is defined as the likelihood that a given individual will later recall a particular stimulus.

... can be predicted using the residual neural network model *ResMem* (Needell & Bainbridge, 2023).

... may be influenced by semantic features (e.g., categories like animals or food) and visual attributes (e.g., color, patterns; cf. Kramer et al., 2023; Lin et al., 2021),

high



„An adult throws a child into the air, while the silhouettes of both can be seen against the sunset.“

Images by Xu et al., 2014

EXPERIMENTS

All experiments were conducted online and followed the same general procedure:

1. Encoding phase: „Is this an indoor or outdoor scene?“

- Participants viewed 40 scenes each with high- ($M = .91$, $SD = .02$) and low- ($M = .63$, $SD = .05$) *ResMem* memorability scores or their corresponding descriptions.

- Participants with accuracies below $M - 2SD$ were excluded from further analysis.

2. Decoding phase: „Do you remember this?“

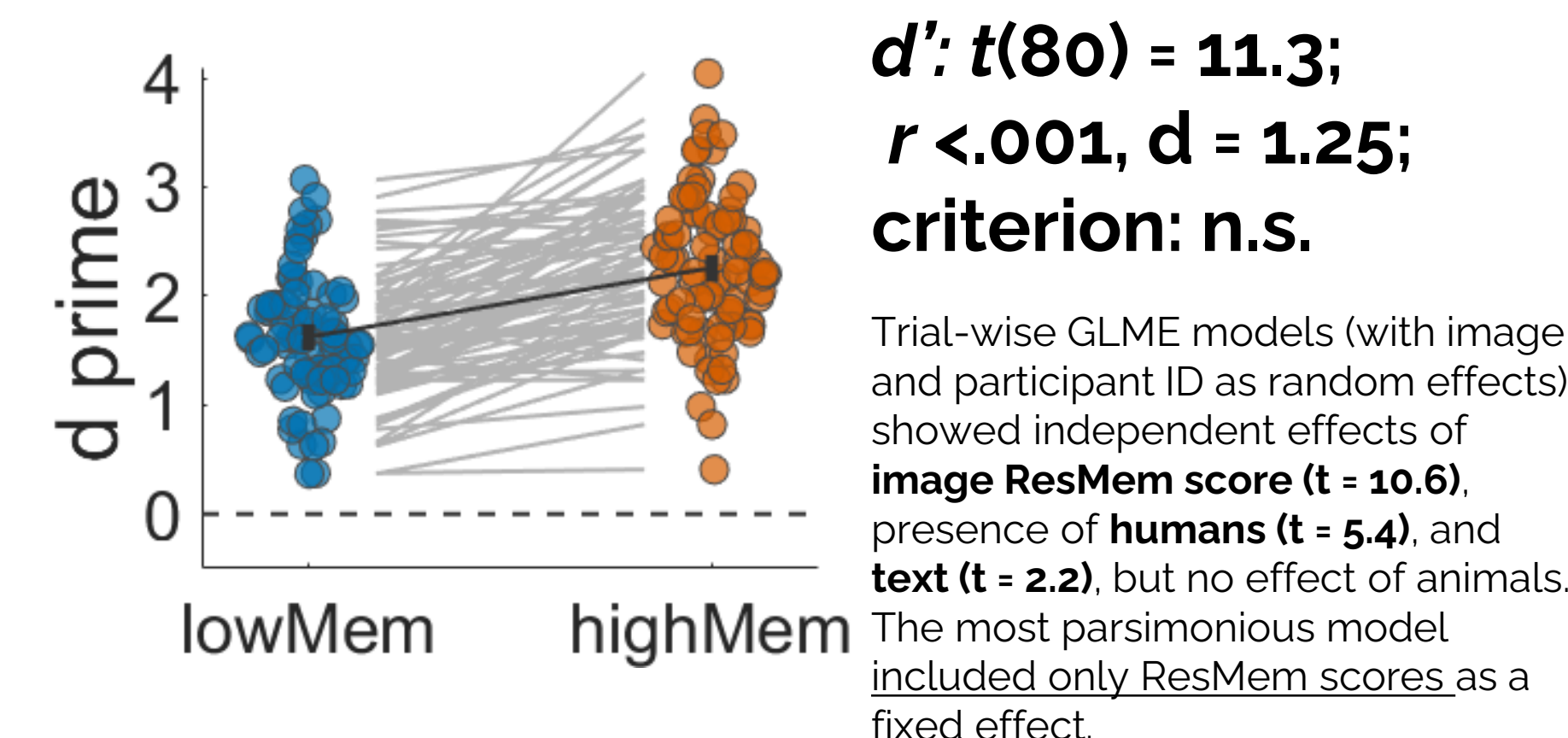
- Participants responded using a four-point scale: yes, rather yes, rather no, no.
- The task included 80 old and 80 new scenes/descriptions, matched for image memorability scores.

3. Memory anchor task: „What helped you decide?“

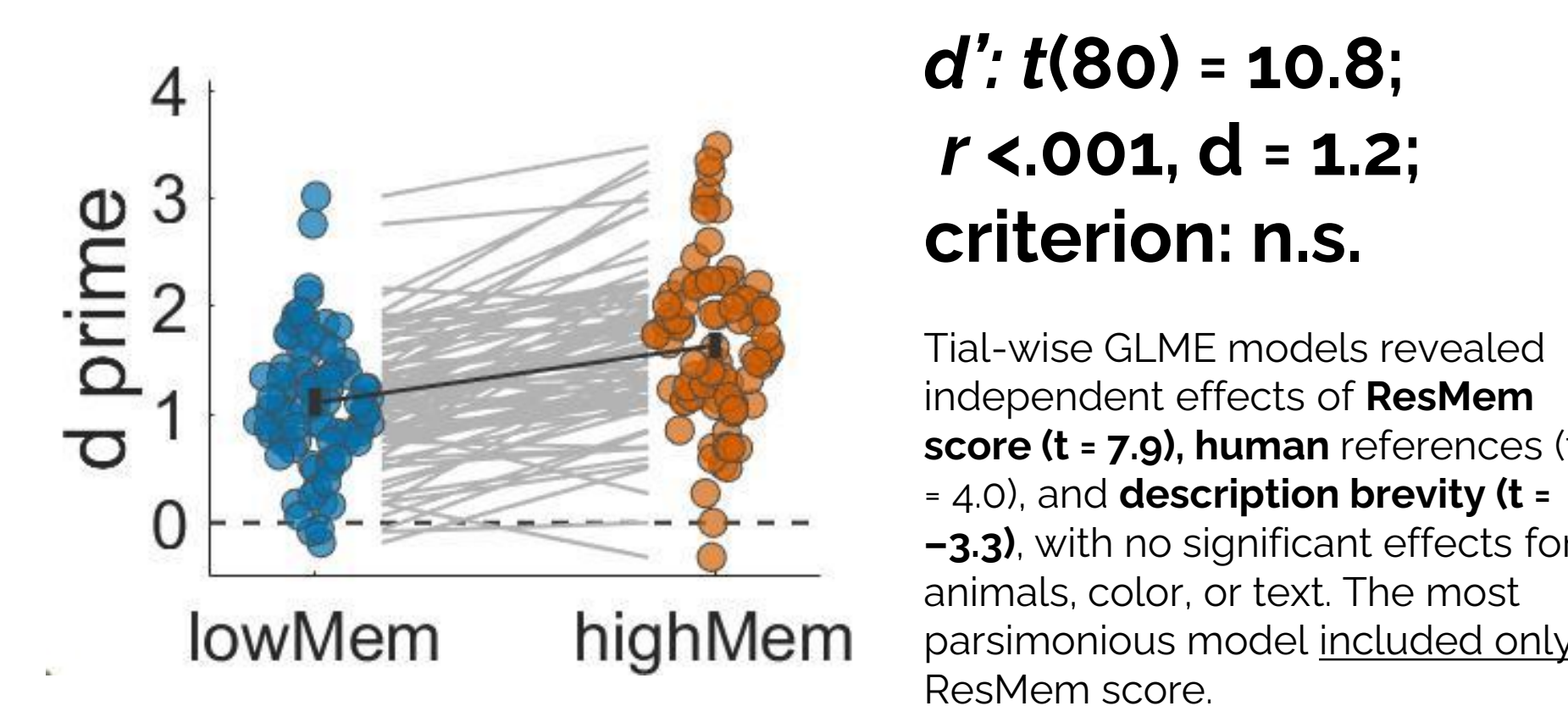
- Free text responses (Exp. 1); single-word selection (Exp. 2a) or multiple-word selection (Exp. 2b)
- Responses were categorized as reflecting either **semantic**, **visual**, **mixed** (cf. Kramer et al., 2023) or **unknown** dimensions.

RESULTS In all experiments N = 81

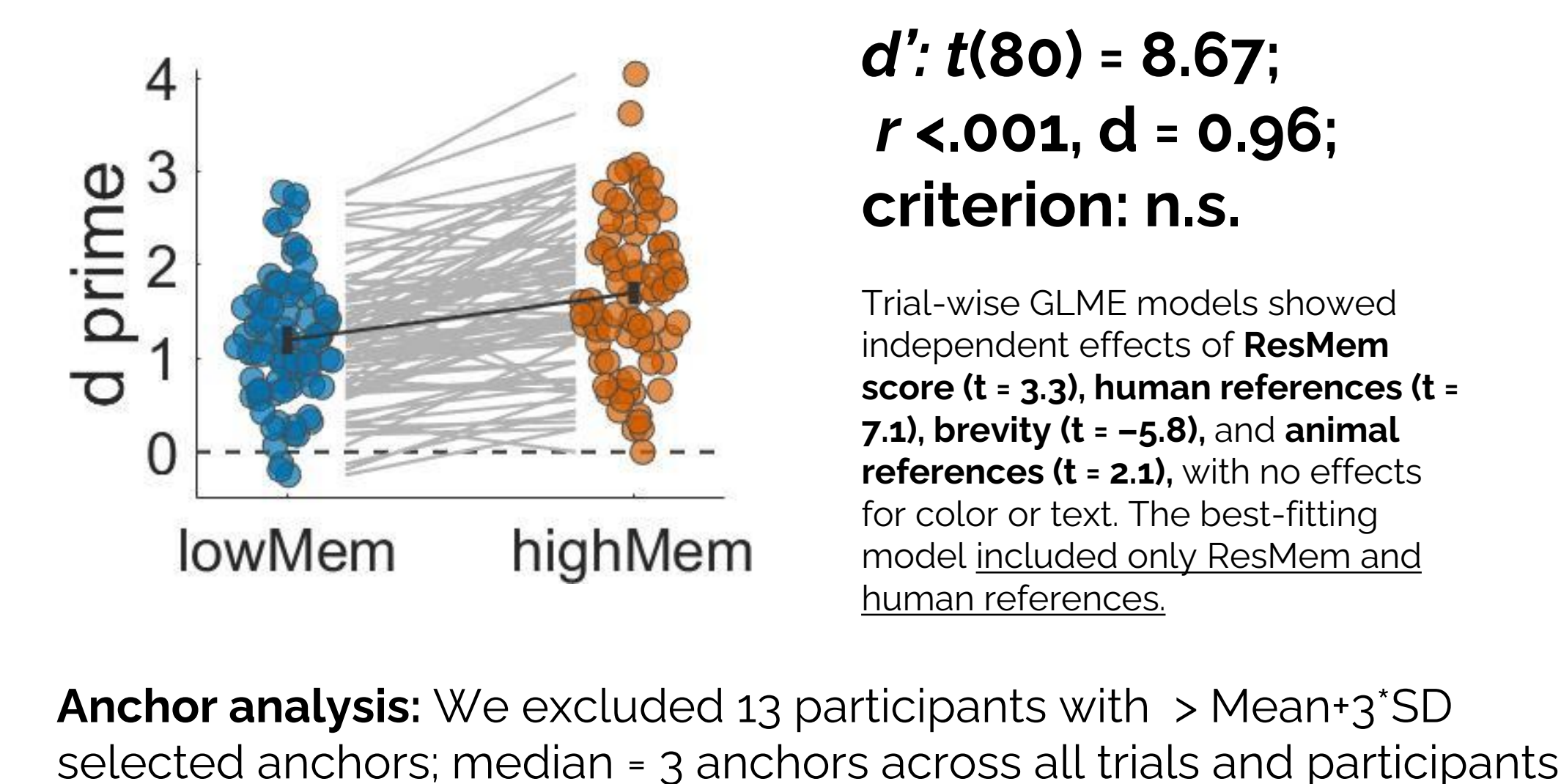
Experiment 1: Scene images & free text responses



Experiment 2a: Human scene descriptions & single-word selection

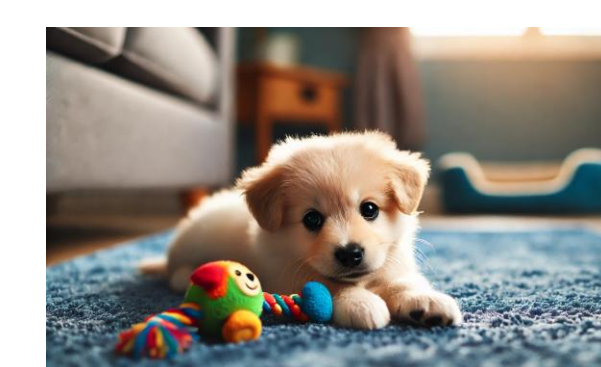


Experiment 2b: Chat-GPT scene descriptions & multiple-word selection



DISCUSSION & CONCLUSION

- Closed-loop AI: We could produce memorable scenes from memorable descriptions, *ResMem* predictions for original images and Gemini-Imagen 3 reproductions: $r = .55$, $p < .001$, and Chat-GPT-DALL.E, $r = .58$, $p < .001$



- Scene descriptions capture the memorability of an image
- Selected memory anchors were mostly semantic
- BUT: Is memorability truly modality-independent and driven by semantics over visual features? Testing this in congenitally blind or aphantasic individuals would be especially informative. If you work with these groups, we'd love to connect—please get in touch!