Towards a re-evaluation of egocentric behaviour in perspective-taking: the role of stimuli selection
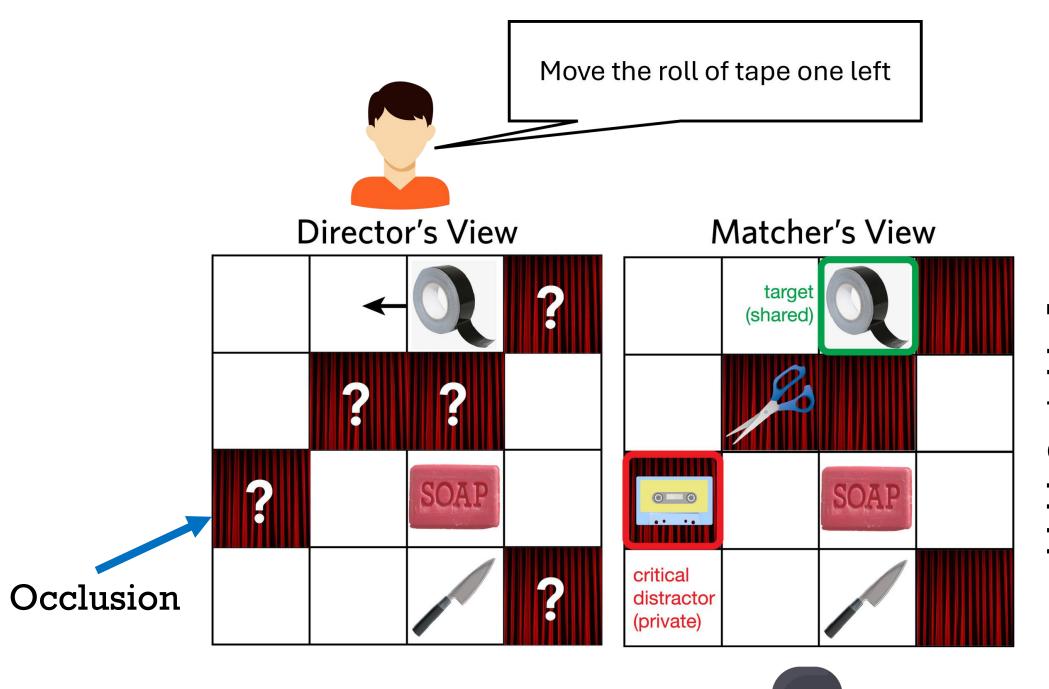
Kastanas Dimitrios[1] & Katsos Napoleon[1]

[1]Department of Theoretical and Applied Linguistics, University of Cambridge, United Kingdom

Introduction

- **Visual Perspective Taking** (VPT) involves inferring visual or spatial properties of a scene relative to another person or position (Keysar et al., 2003).

Figure 1: The Director Task (Keysar et al., 2003). Adapted from Hawkins et al. (2021).

Target: roll of tape
Hidden Distractor: cassette tape
Occlusion: curtains
Director: speaker (male)
Matcher: listener (female)

- The **Target** item is visible to both the speaker and the listener.
- The **Hidden distractor** is visible only to the listener.
- **Occlusions** block the speaker's view, so the items behind the curtains are visible only to the listener.

Previous work

- Eye-tracking, neuroimaging and behavioural data have shown that listeners make a substantial number of errors in this task (Brown-Schmidt & Hanna, 2011).
- Prominent theories:
  - Anchor & Adjustment (Keysar et al., 2003): egocentric processing.
  - Division of labour (Heller et al., 2016): simultaneous activation of one's own and the other's perspective.
- **Limitation**:
- Why do listeners make errors in the first place?
- The Hidden Distractor is consistently a better referent for the target expression (Hanna & Brown-Schmidt, 2011; Heller et al., 2016).
- To our knowledge, no study has systematically investigated this claim.

Methodology

- **Participants**: 33 neurotypical adult native speakers of English recruited from Prolific.
- **Stimuli**: We used the items from one of the influential studies of Keysar and colleagues, which was replicated by Hawkins et al. (2021) in an online setting.

	Instruction/Target expression	Target	Hidden Distractor
1	(Move the) Glasses	Sunglasses	Glasses case
2	(Move the) Bottom block	Block (3rd row)	Block (4th row)
3	(Move the) Tape	Cassette tape	Roll of tape
4	(Move the) Large measuring cup	Medium cap	Large cap
5	(Move the) Brush	Round hairbrush	Flat hairbrush
6	(Move the) Eraser	Board eraser	Pencil eraser
7	(Move the) Small candle	Medium candle	Small candle
8	(Move the) Mouse	Computer mouse	Toy mouse

Figure 2: Critical stimuli and instructions from Keysar et al. (2003).

- **Design**:
  - Participants were asked to indicate the object a speaker would most likely refer to when using an ambiguous word/expression. They also rated their confidence in their options.
  - No Images of the items were shown to participants to tap into their linguistic intuition without influence from visual features.

Results

Distribution of Participant Responses by Ambiguous Item in Keysar et al. (2003) & Hawkins et al. (2021) Critical Stimuli
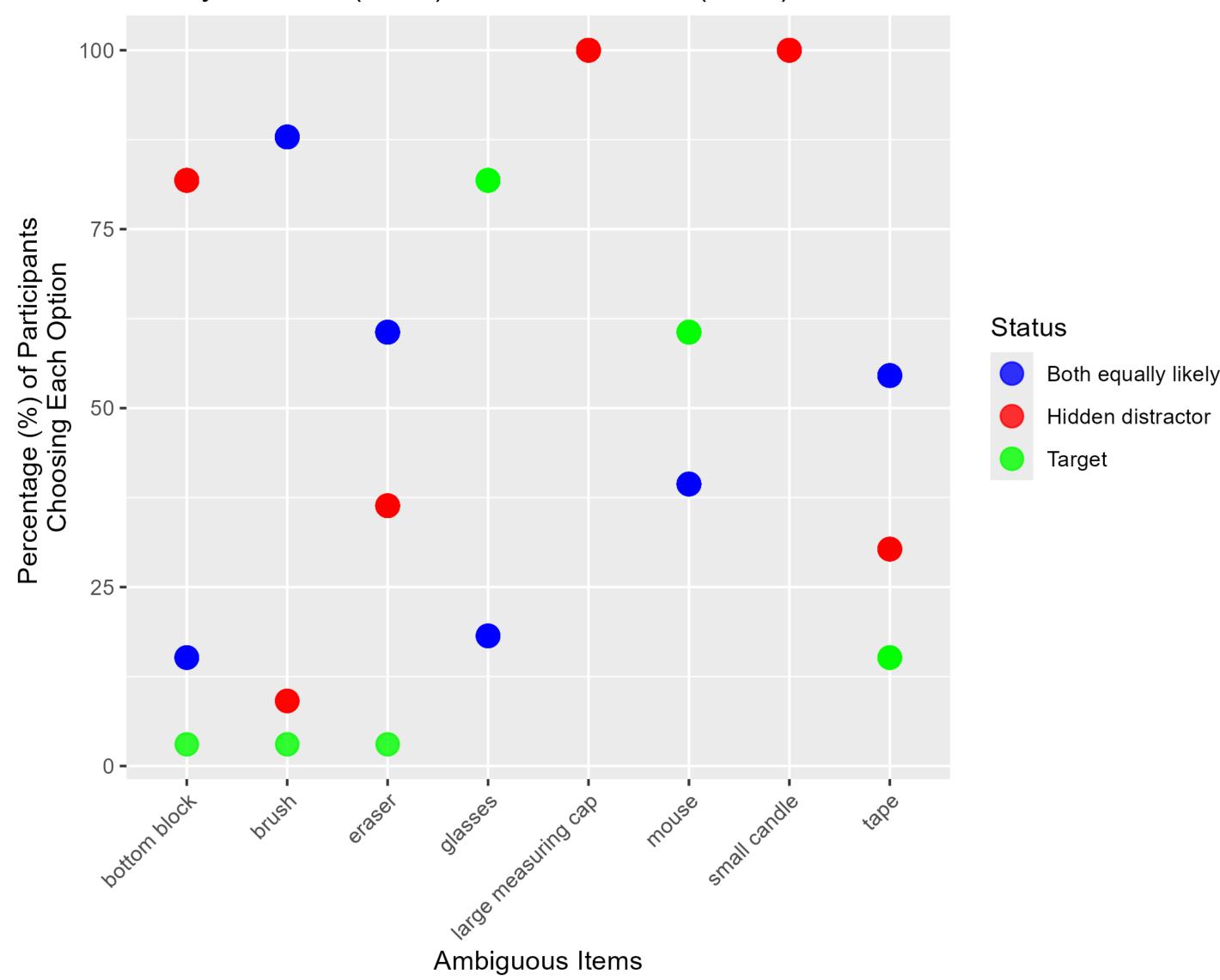
Figure 3: Results from our experiment. **Target** refers to the target item, **Hidden Distractor** refers to the occluded object from the speaker's perspective but accessible to the listener, **Both equally likely** refers to both the Target and Hidden Distractor.

- **Analysis**: Chi-square test of independence.
- Significant associations between Status categories and Ambiguous items were found for bottom block (p<0.001), brush (p<0.001), eraser (p<0.001), glasses (p<0.001), tape (p<0.05), but not for mouse (p=0.223).
- Participants responses were not random but favoured specific interpretations for seven out of eight items and their confidence in their options was over 80%.

Discussion

- Listeners do not make errors because the target expression consistently favours the Hidden Distractor over the Target (cf. Brown-Schmidt & Hanna, 2011). These results highlight a more complex picture regarding the source of listeners' errors.
- Proposed novel taxonomy:
  - **T**: The target expression consistently points to the Target (T).
  - **HD**: The target expression consistently points to the Hidden Distractor (HD).
  - **T+HD**: The target expression consistently points to both the Target and Hidden Distractor (T+HD).
- **Predictions (given that T ≠ HD)**:
  - For items belonging in **T**, listeners will do no errors or very few errors.
  - For items belonging in **HD** and **T+HD**, listeners will do a high number of errors.
  - For items belonging in **T+HD**, the errors will persist even when listeners interact with naïve speakers rather than confederates, unlike items in **HD**.

Limitations & Next Steps

- More participants and stimuli from more VPT studies are needed to refine the proposed classification.
- Analysis of error patterns across neuroimaging, eye-tracking and behavioural paradigms is required for verifying whether our proposed classification is causally related to listeners' performance in VPT tasks.

References

- Brown-Schmidt, S. & Hanna, J. E. (2011), D & D, https://doi.org/10.5087/dad.2011.102
- Hawkins, R. et al. (2021), Cognitive Science, https://doi.org/10.1111/cogs.12926
- Heller, D. et al. (2016), Cognition, https://doi.org/10.1016/j.cognition.2015.12.008
- Keysar, B. et al. (2003), Cognition, https://doi.org/10.1016/s0010-0277(03)00064-7

Acknowledgements

UNIVERSITY OF CAMBRIDGE

SUPPORTED BY
ONASSIS FOUNDATION

CAMBRIDGE COMMONWEALTH EUROPEAN AND INTERNATIONAL TRUST

Dimitrios Kastanas
dk610@cam.ac.uk