Understanding language comprehension: the challenge of measuring individual differences within experimental designs

Lena Blott¹, Anna Gowenlock¹, Kate Nation², & Jennifer Rodd¹
¹ Department of Experimental Psychology, University of Oxford; ² Department of Experimental Psychology, University of London

Background
- Individual differences in language comprehension are important for theory development, and can help answer questions about classic debates in psycholinguistics, e.g. relative contributions of domain-general vs domain-specific factors to processing (see e.g. Kidd et al., 2016).
- Findings at the group-level suggest that people sometimes find it hard to access relevant meanings of ambiguous words, especially when they are subordinate (Overy et al., 1990; Foxe, Siver, & Siver, 1995; Rodd, Johnsrude, & Davis, 2010, and many more…).
- We know less about individual differences in using context to resolve lexical-semantic ambiguity (but see e.g. Norbury, 2005; Nation & Snowling, 1998; Khanna & Billard, 2010).

Methods
- **Test of disambiguation ability:** listen to a short narrative and choose appropriate picture (See e.g. Kidd et al., 2018). Each participant encountered three conditions:
  - Ambiguous (N=66): Even with his binoculars, she couldn’t find the source of the hooting noise. He waited a few minutes. Then he finally saw the heron.
  - Matched control (N=66): Even with his binoculars, she couldn’t find the source of the hooting noise. He waited a few minutes. Then he finally saw the crane.
  - Unmatched control (N=66): The baby didn’t want to play with any of her own toys. Everything else was more interesting. She managed to steal her mum’s picture.

Results
- **Comparison of Ambiguous vs matched condition allows for statistical inferences about the effect of ambiguity, whilst controlling for narrative- or picture-based factors, at the group-level.**
- **In within-participant individual differences design, we need to compare Ambiguous vs a condition that is unmatched.**

Participants
- **Can we develop a measure that reliably captures individual differences in the ability to use context to resolve lexical-semantic ambiguity?**

Can we develop a measure that reliably captures individual differences in the ability to use context to resolve lexical-semantic ambiguity?

Reliability
- Difference scores between Ambiguous and Unmatched control
- Using 5000 random splits, Spearman-Brown corrected reliability estimates are relatively low: 0.43-0.63

Recommendations for measuring individual differences within experimental designs
- Isolate process of interest, e.g. with a comparison between well-matched conditions
- Design items to tap into the same construct and cover a range of difficulty
- Avoid adding noise that is unrelated to the process of interest, e.g. by presenting the same list of items in the same order to all participants
- Aim for the highest N of trials you can get away with
- Pilot the paradigm and desired outcome measures carefully, and check that your measure is sensitive to individual differences
- Report reliability of your measure(s) in each sample you use them on
- Make use of trial-level data
- If possible, use a latent variable approach to minimise measurement error