Poster walkthrough

Understanding the role of linguistic distributional knowledge in cognition

Cai Wingfield & Louise Connell
Lancaster University

DUCOG 2021
18–21 May 2021

Poster layout

1. Introduction

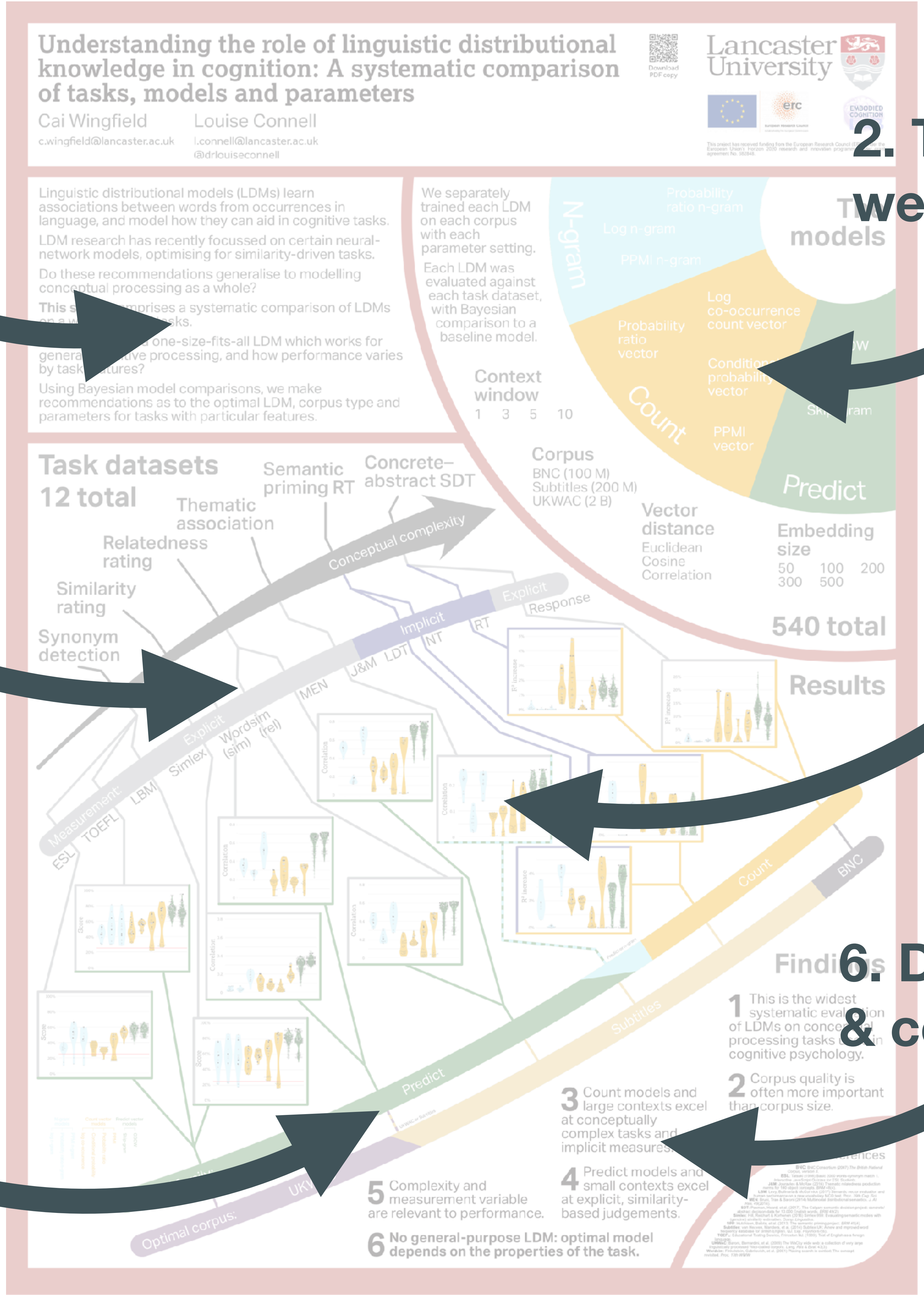
2. The models we considered

3. The datasets we modelled

4. All the results

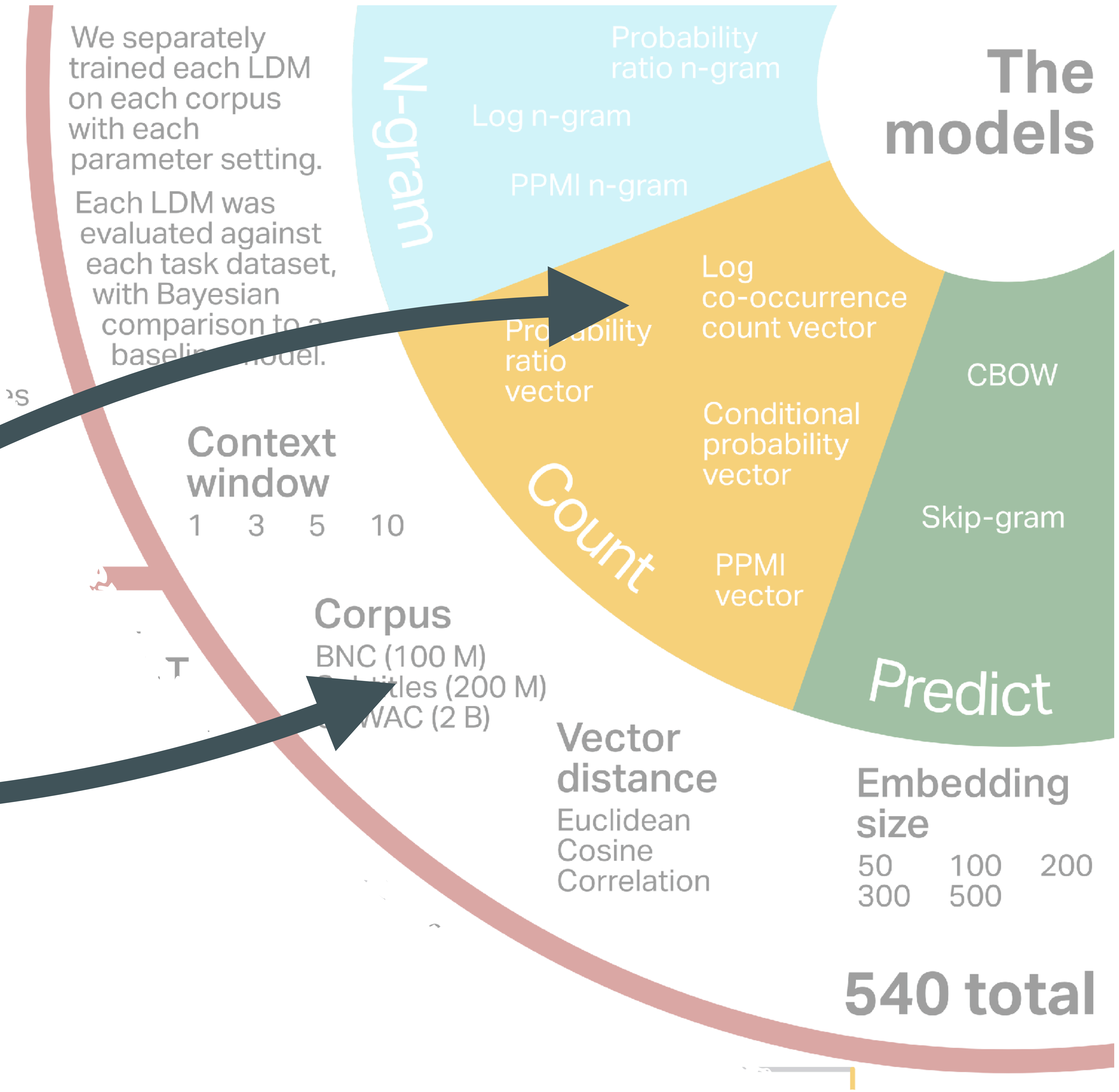
5. Summary of results

6. Discussion & conclusion



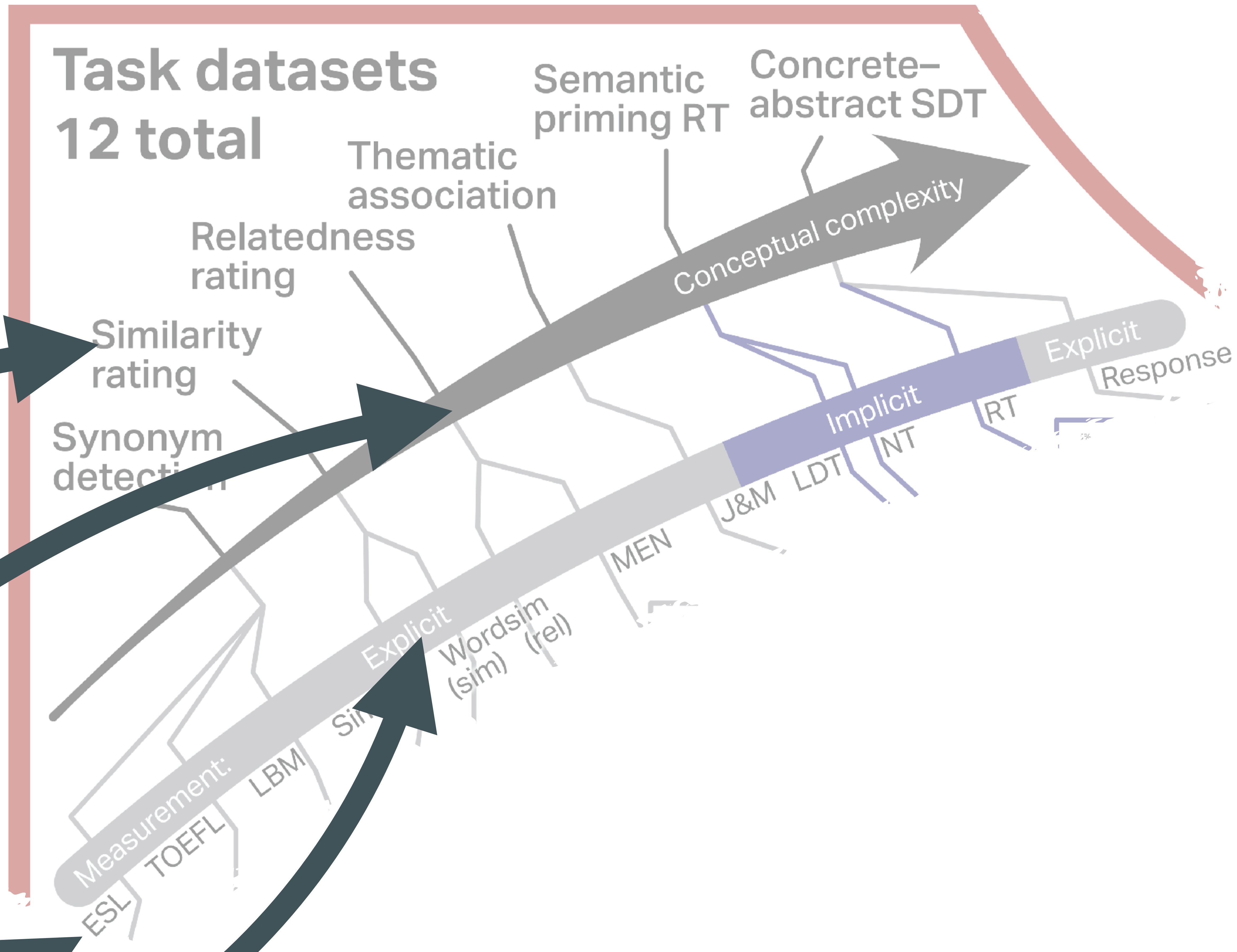
The models

- We trained **linguistic distributional models from three families**: n-gram, count-vector and predict-vector. Nine types of model in total.
- We varied **parameters for each model**: training corpus and context-window size for all models, vector distance for all vector models, and embedding size for all predict models.



The task datasets we modelled

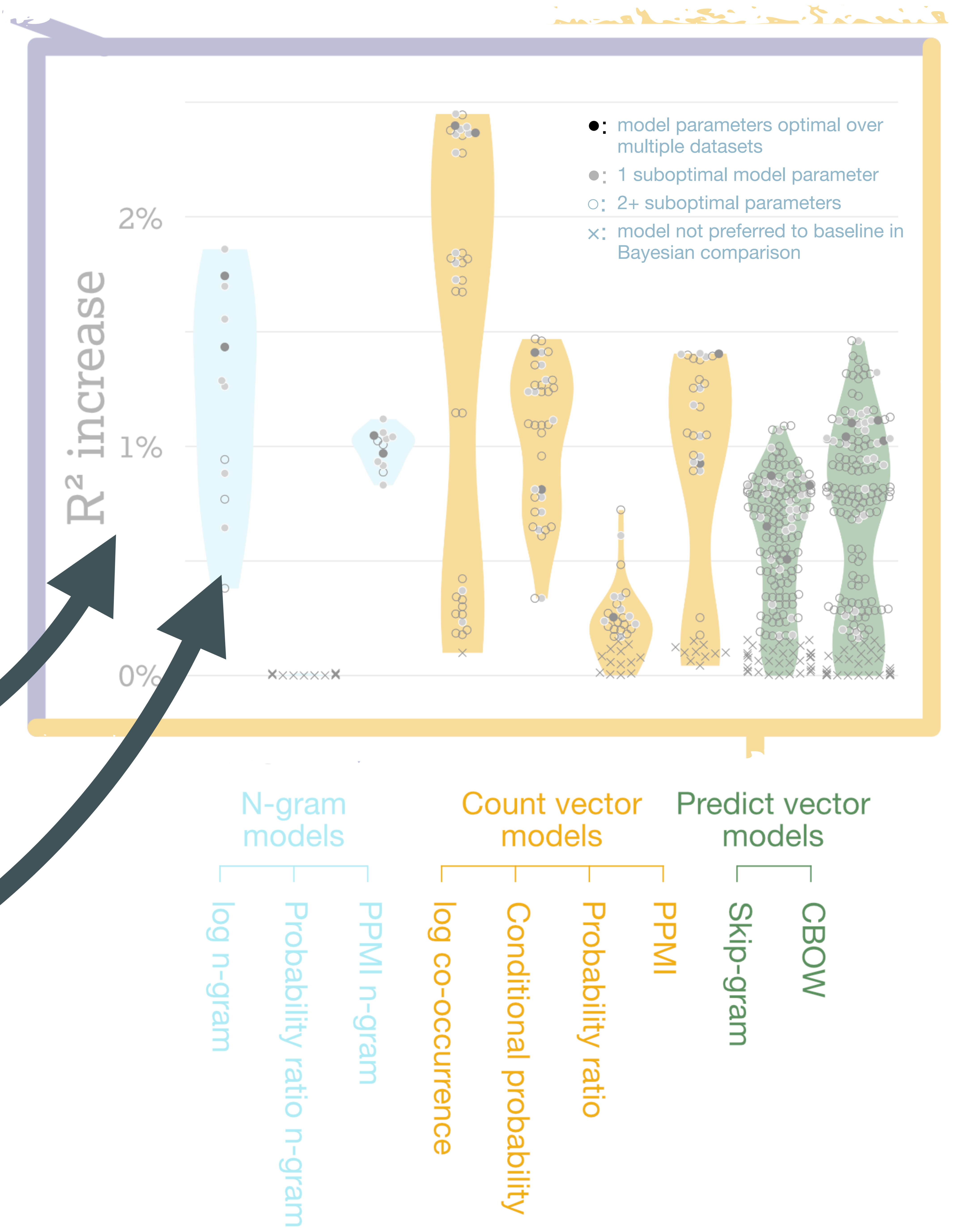
- We looked at **6 tasks** used in cognitive psychology.
- These tasks varied in **conceptual complexity**, from low (e.g. detecting synonyms) to high (e.g. making judgements about concept abstractness).
- We instantiated the tasks using **12 example datasets**.
- Some dependent variables were **explicit** measures of conceptual processing (e.g. rating similarity of word pairs), others were **implicit** (e.g. lexical decision reaction times).



ESL: *English as a Second Language quiz* (Tatsuki, 1998)
TOEFL: *Test of English as a Foreign Language* (Educational Testing Service, 1989)
LBM: *Vocabulary MCQ test* (Levy et al., 2017)
Simlex: *Similarity judgements* (Hill et al., 2016)
Wordsim: *Similarity (**sim**) & relatedness (**rel**) judgements* (Finkelstein et al., 2001)
MEN: *Relatedness judgements* (Bruni et al., 2014)
J&M: *Thematic relatedness norms* (Jouravlev & McRae, 2016)
LDT: *Lexical decision semantic priming* (Hutchison et al., 2013)
NT: *Word naming semantic priming* (Hutchison et al., 2013)
RT: *Concrete/abstract semantic decision reaction times* (Pexman et al., 2017)
Response: *Concrete/abstract semantic decision responses* (Pexman et al., 2017)

Reading the results graphs

- For each dependent variable in each task, we compared the relative performance of different model types, and different parameters. Here is **an example: word naming RT**.
- The test statistic is on **the vertical axis**. For example here, it's R^2 increase for the model predictor. For other tests, it's correlation or prediction accuracy.
- Each graph has **one distribution for each type of model**, showing performance for each parameter setting.



Summarising the results

- We found **a general trend**:
    - Explicitly measured, conceptually simple tasks were best predicted by predict models trained on large, low-quality corpora.
    - Implicitly measured and conceptually more complex tasks were best predicted by count-based models trained on smaller, higher-quality corpora.
- Upper coloured bar: **recommendations for the optimal model family**, reflecting consistently superior performance using consistently good parameters.
- Lower coloured bar: **recommendations for optimal corpus**, reflecting consistently superior performance over a range of models and parameters.
- Overall, the optimal model, and parameters was highly task-dependent, with no candidate performing consistently well in all scenarios.