Taking The **Human Decision Maker** into Account: A Critical Evaluation of **Sorting Algorithms** for Large Stimuli Sets and Guidelines for Improvement

Guang YANG · Renell RODRIGUES · David YU · Frans A.J. VERSTRATEN

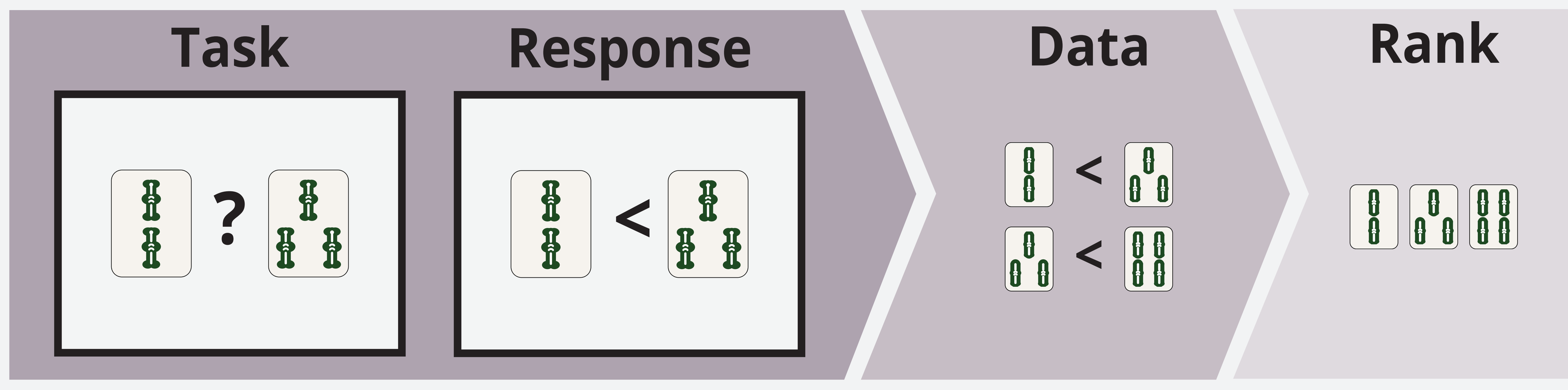
THE UNIVERSITY OF SYDNEY

Scan to track development progress

Introduction

Efficiently ranking many items along specific dimensions can be a challenge across science, especially psychology. Exhaustively comparing all possible item pairs of a specific feature, is often impractical and time consuming.

Traditional sorting algorithms reduce comparisons by strategically scheduling them to allow inference (e.g., if A < B and B < C, then A < C). However, such **algorithms do not account for human cognitive limitations: Humans are prone to making errors (e.g. 2 > 3).**



Research Question

1. How well do existing sorting algorithms perform when human cognitive constraints are taken into account?

2. Can an enhanced algorithm that is both efficient and human-compatible be developed?

We developed an enhanced algorithm - *Emerge Sort* - and benchmarked its performance against traditional methods.

Method

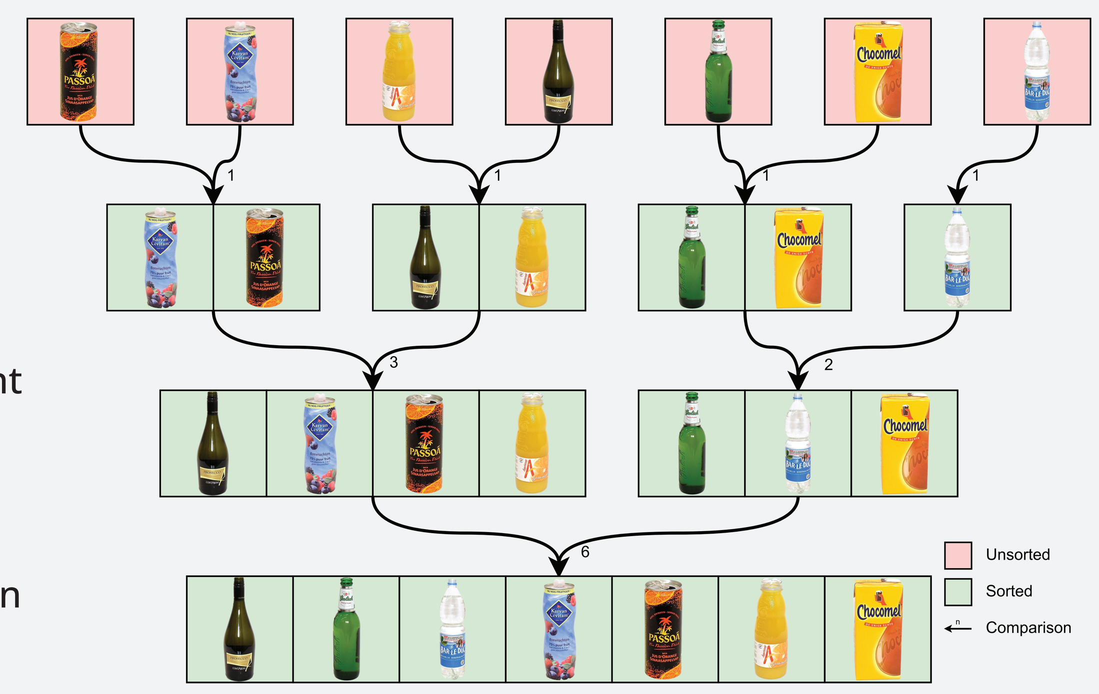
Monte Carlo simulations were conducted to evaluate sorting algorithms.

Example of Merge Sort for Perceptual Salience

The figure shows an example of a sorting algorithm for a perceptual task where participants judged the visual salience of images.

The algorithm recursively merges subsequence stimuli based on participant judgments, adaptively reducing the number of comparisons required.

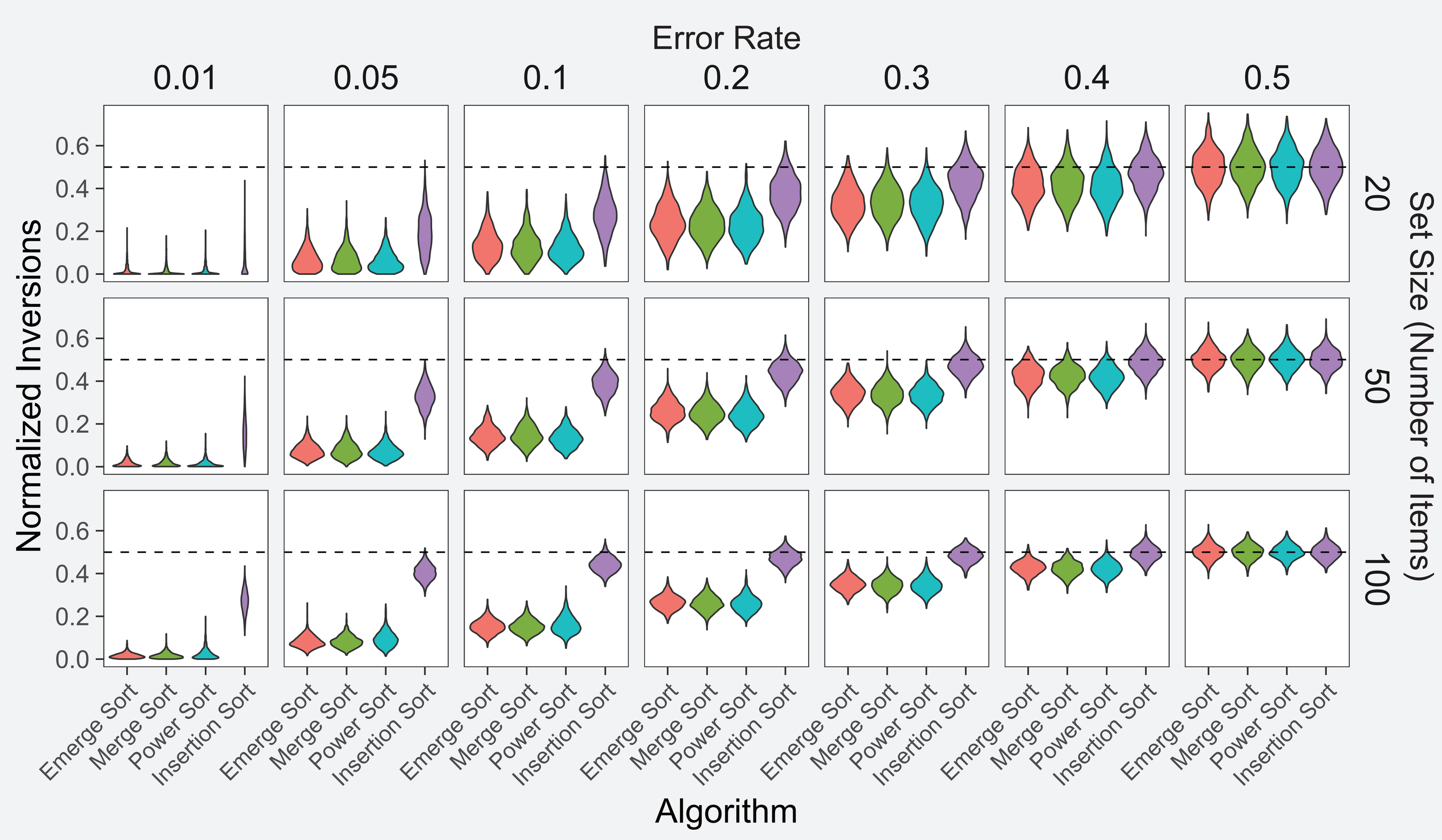
Only 14 comparisons were required, compared to 21 in a full pairwise design.



Result

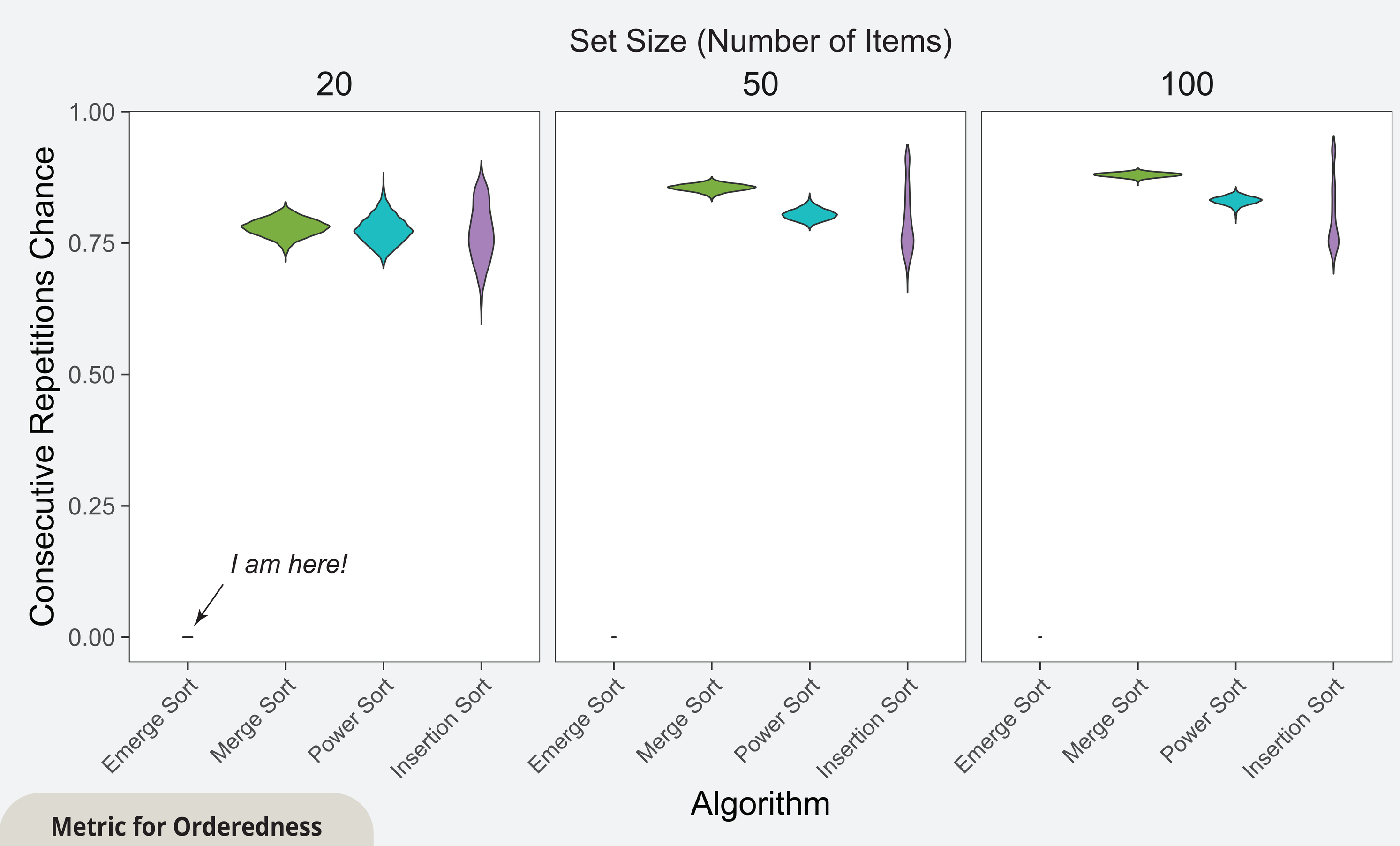
Fault Tolerance

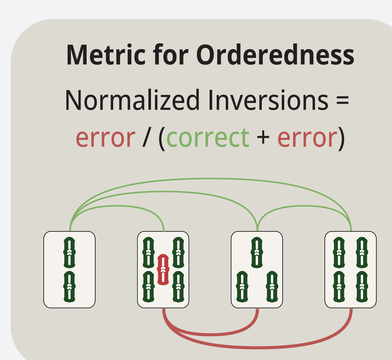
a single error should have limited impact.



No Consecutive Item Repetition

to reduce fatigue and practice effects.



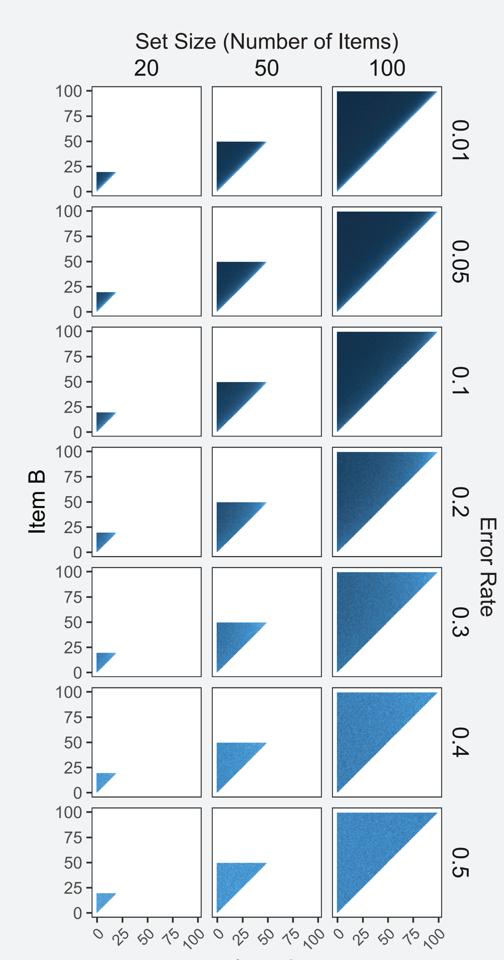


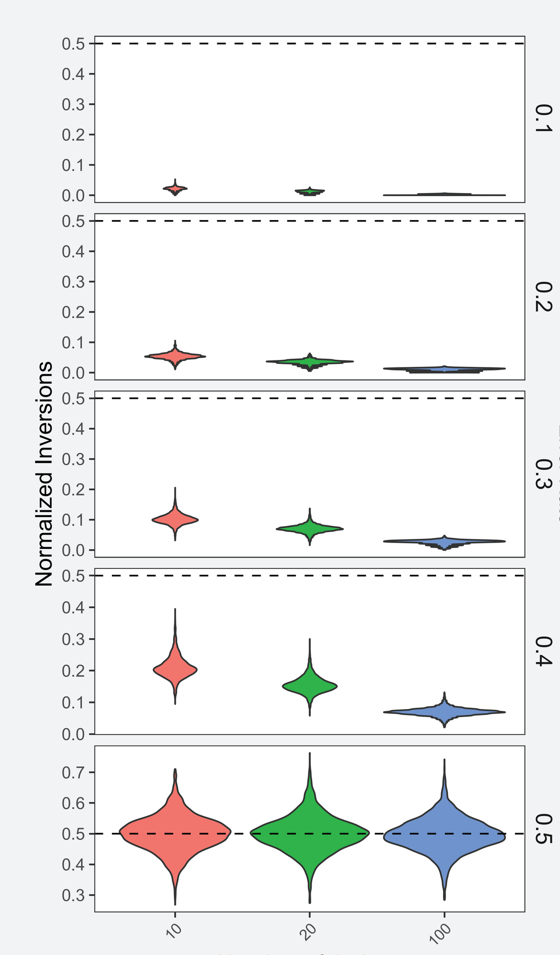
Emerge Sort

As illustrated on the right, Emerge Sort demonstrates two key characteristics:

- Resilient evidence accumulation under noise: As the error rate increases, the algorithm tends to explore the comparison space more thoroughly.
- Improved consensus with larger sample size for judgement task.

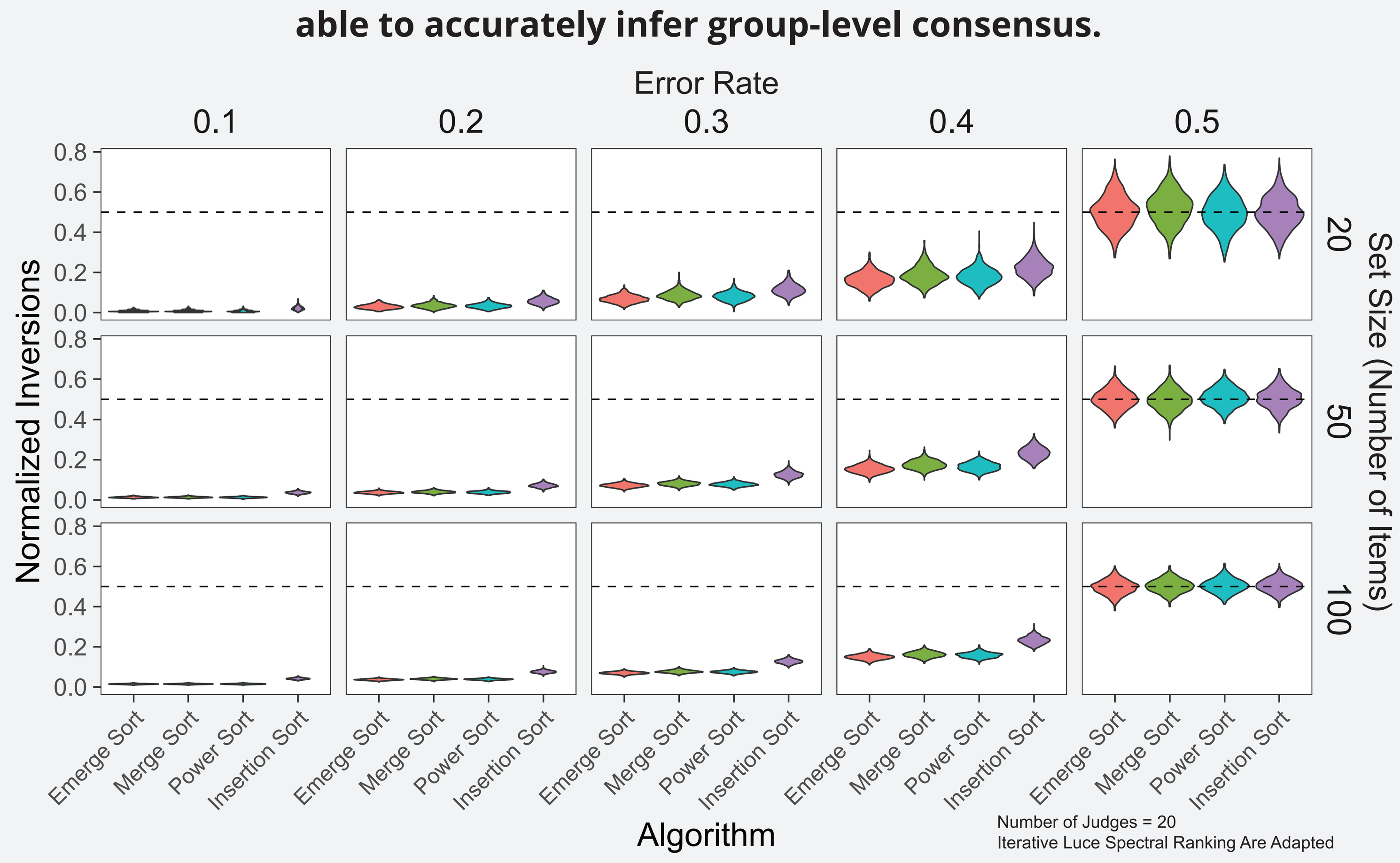
Together, these properties make Emerge Sort well-suited for robust group-level ranking.





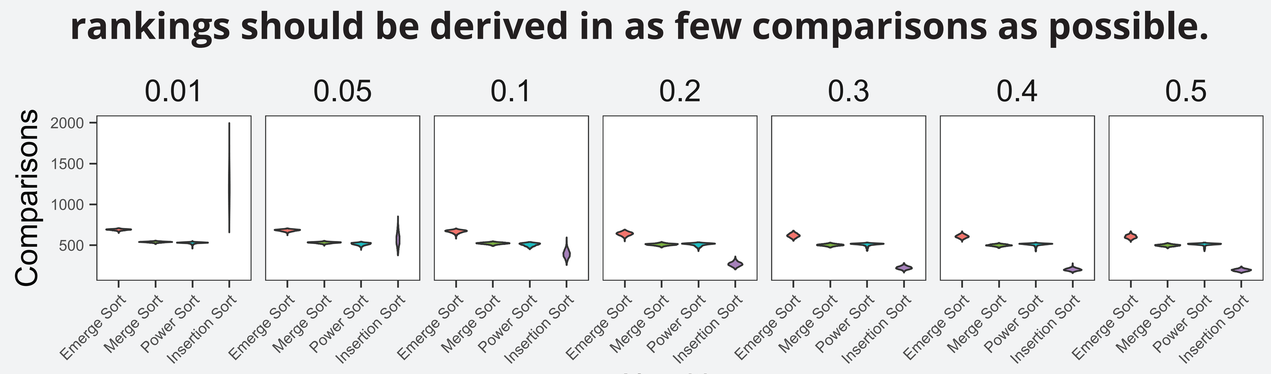
Group-level Ranking Inference

able to accurately infer group-level consensus.



Efficiency

rankings should be derived in as few comparisons as possible.



Conclusion

If one needs to sort a large database of stimuli along a particular dimension, the traditional sorting algorithms are not really suitable for human-based decision making.

The alternative, so called merge-sort-based methods are a better choice, where our proposed variant – Emerge Sort – performs best.