The Effect of Lineup Size on Discriminability is Dependent on Filler Similarity and Independent of Encoding Strength

Allan Lam & John Wixted
Department of Psychology, University of California, San Diego

Introduction

A photo lineup, which consists of one suspect and several physically similar fillers, is often used by the police to test an eyewitness's memory. Recent work suggests that including description-matched fillers who match the basic characteristics of the perpetrator (e.g., same age, race, gender) but who are otherwise maximally dissimilar to the suspect optimizes discriminability. However, the optimal lineup size has been found to vary with filler similarity (a retrieval manipulation), with larger lineup sizes increasing or decreasing discriminability depending on whether low- similarity or high-similarity fillers were used, respectively. Here, we investigated whether *encoding* manipulations that affect overall performance also affect how lineup size influences discriminability.

Predictions

There are two competing signal-detection models of eyewitness identifications, the Independent Observations Model and the **Ensemble Model**. The Ensemble model holds that identification decisions are based on how familiar a face is relative to the average familiarity of the faces in the lineup. In terms of memory signals, the decision variable consists of the strongest raw signal minus the average memory signal associated with all faces in the lineup. Because the average familiarity is more precisely determined the more faces there are in the lineup, the Ensemble Model predicts that discriminability increases with lineup size.

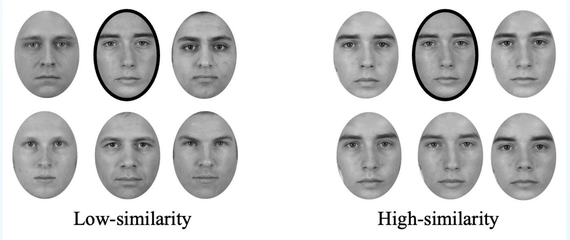
Empirically, recent studies (Akan et al., 2021; Wooten et al., 2020) found that the effect of lineup size plateaus beyond a small number of fillers. Shen et al. (2024) extended this line of research, revealing an interaction between lineup size and filler similarity. Specifically, when fillers were of low similarity to the suspect, increasing lineup size improved discriminability, as predicted by the Ensemble Model. However, when filler similarity was high, increasing lineup size paradoxically reduced discriminability, which no existing model predicts (Ariely, 2001; Mazyar et al., 2012, 2013).

We conducted three experiments investigating whether similar effects are observed when discriminability is manipulated at encoding rather than retrieval.

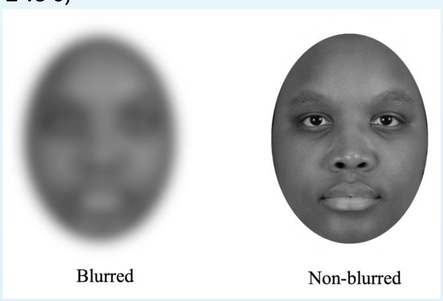
Methodology

Each participant received six trials. Each trial included a 2-second study phase, a 120-second distractor task, and a test phase.

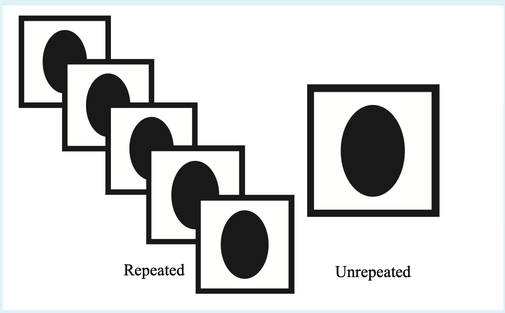
Experiment 1 (*N* = 502)
Between-subject: Filler-similarity (high-similarity vs. low-similarity)
Within-subject: Lineup type (TA vs TP) & Lineup size (1 vs 2 vs 6)



Experiment 2 (*N* = 553)
Between-subject: Target blurriness (blurred vs. non-blurred) Low-similarity fillers ONLY at retrieval
Within-subject: Lineup type (TA vs TP) & Lineup size (1 vs 2 vs 6)



Experiment 3 (*N* = 501)
Between-subject: Target Repetition (Repeated vs. Unrepeated) High-similarity fillers ONLY at retrieval
Within-subject: Lineup type (TA vs TP) & Lineup size (1 vs 2 vs 6)



Results

We estimated d'ig (discriminability between innocent and guilty suspects) by fitting the IO model to the data. Similar estimates (not shown) were obtained by fitting the Ensemble Model to the data.

Similarity	Size	d'_{IG}
Low-Similarity	1	2.07
	2	2.23
	6	2.27
High-Similarity	1	2.15
	2	1.86
	6	1.94

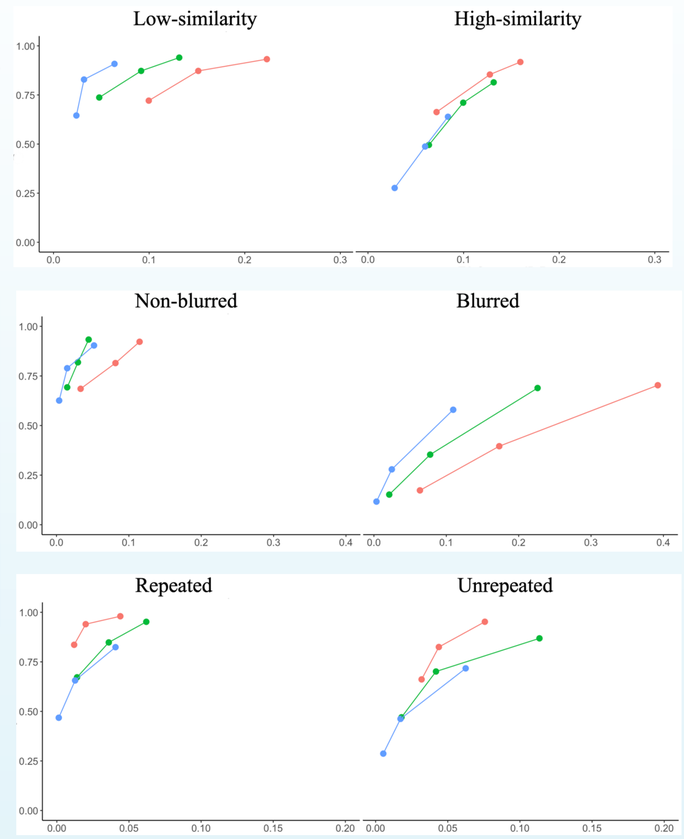
Condition	Lineup Size	d'_{IG}
Non-blurred	1	2.46
	2	2.70
	6	3.08
Blurred	1	0.73
	2	0.99
	6	1.41

Repetition	Size	d'_{IG}
Repeated	1	3.61
	2	3.24
	6	3.19
Unrepeated	1	2.81
	2	2.58
	6	2.56

Summary of Empirical Findings

Overall, **Experiment 1** replicated previous findings, reinforcing the interaction between the effects of filler similarity and lineup size. When low-similarity fillers were used, increasing lineup size enhanced discriminability, but the opposite effect was observed when high-similarity fillers were used. **Experiment 2** found that when low-similarity fillers were used, increasing lineup size consistently improved discriminability, regardless of whether encoding conditions were favorable or unfavorable. This finding aligns with predictions from the Ensemble model. **Experiment 3** revealed that when high-similarity fillers were used, increasing lineup size consistently decreased discriminability, regardless of encoding conditions.

Confidence-Based ROC



Conclusions

The overall pattern of results is not predicted by any model. We therefore offer one possible interpretation to guide future research. The ensemble (average) representation reflects shared features in the lineup, which are nondiagnostic of the perpetrator. However, as the faces in the lineup become increasingly similar, the ensemble face becomes more complete (# non-diagnostic features increases; # diagnostic features decreases). It becomes even more precise as well as lineup size increases. When that occurs, it may serve to command attention to the shared features. It is possible that the ensemble representation, in addition to providing a convenient means of subtracting away the effect of shared features (enhancing d'), also biases perception toward the shared facial features in the lineup (reducing d').