

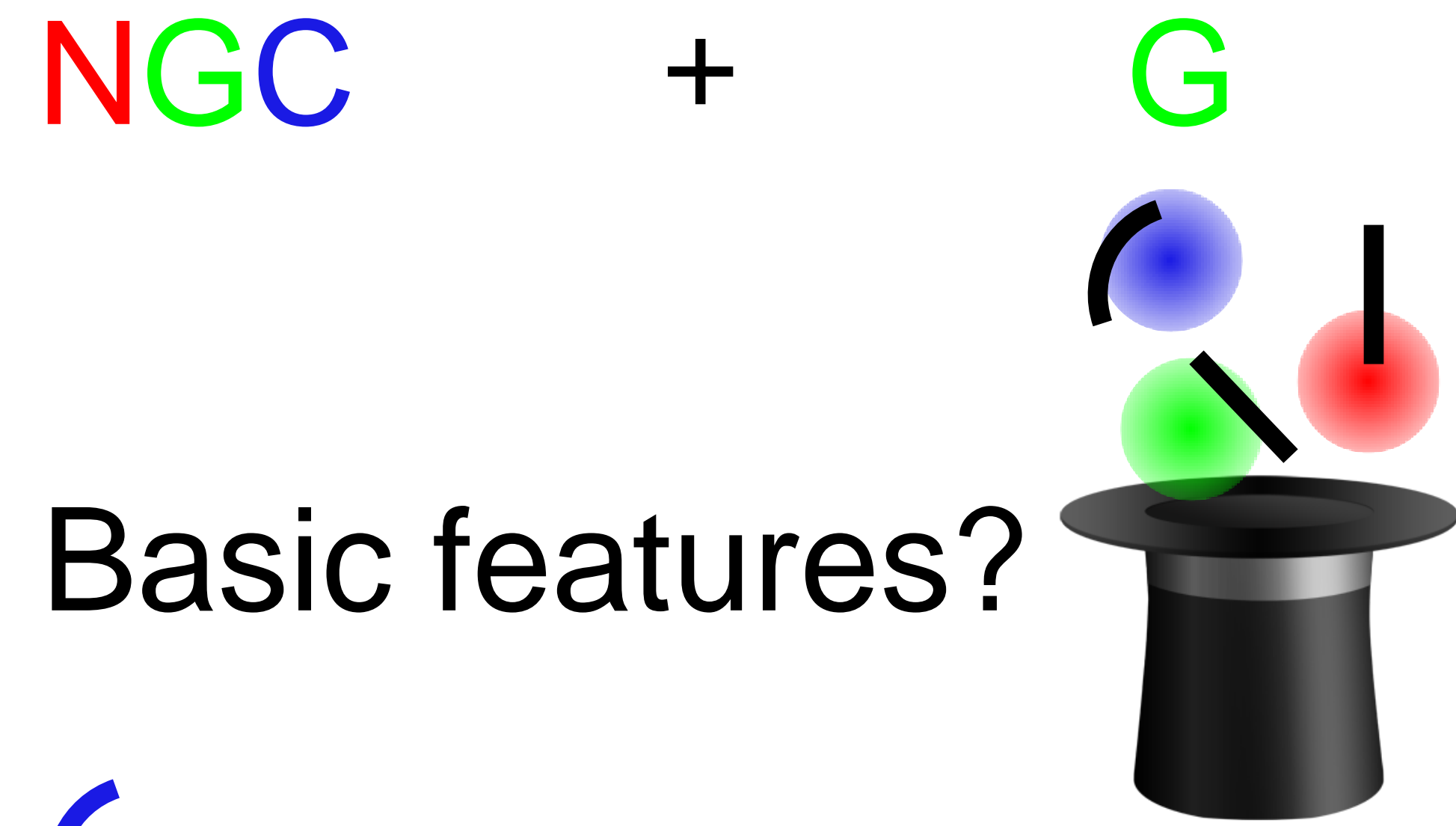
A high-dimensional pooling model accounts for seemingly conflicting substitution effects in crowding



Shaiyan Keshvari¹ & Ruth Rosenholtz^{1,2}

¹Department of Brain and Cognitive Sciences, MIT
²CSAIL, MIT

What does crowding mix up?



Basic features?



Bound features?



Whole objects?

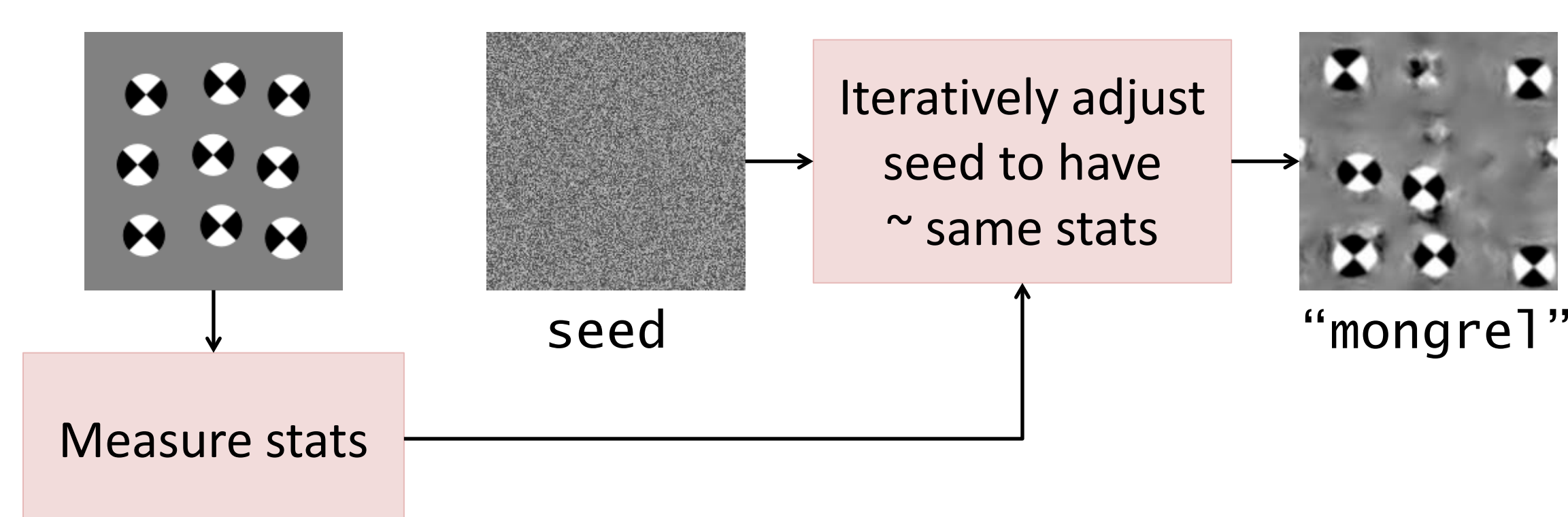
What about a texture representation?

A rich set of statistics of low-level image features

- Autocorrelation
- First three moments of luminance
- Cross-correlations of responses of V1-like cells across location, orientation, & scale
- Phase correlation

(Portilla & Simoncelli, 2000)

- Proposed by Balas et al. as a model for visual crowding

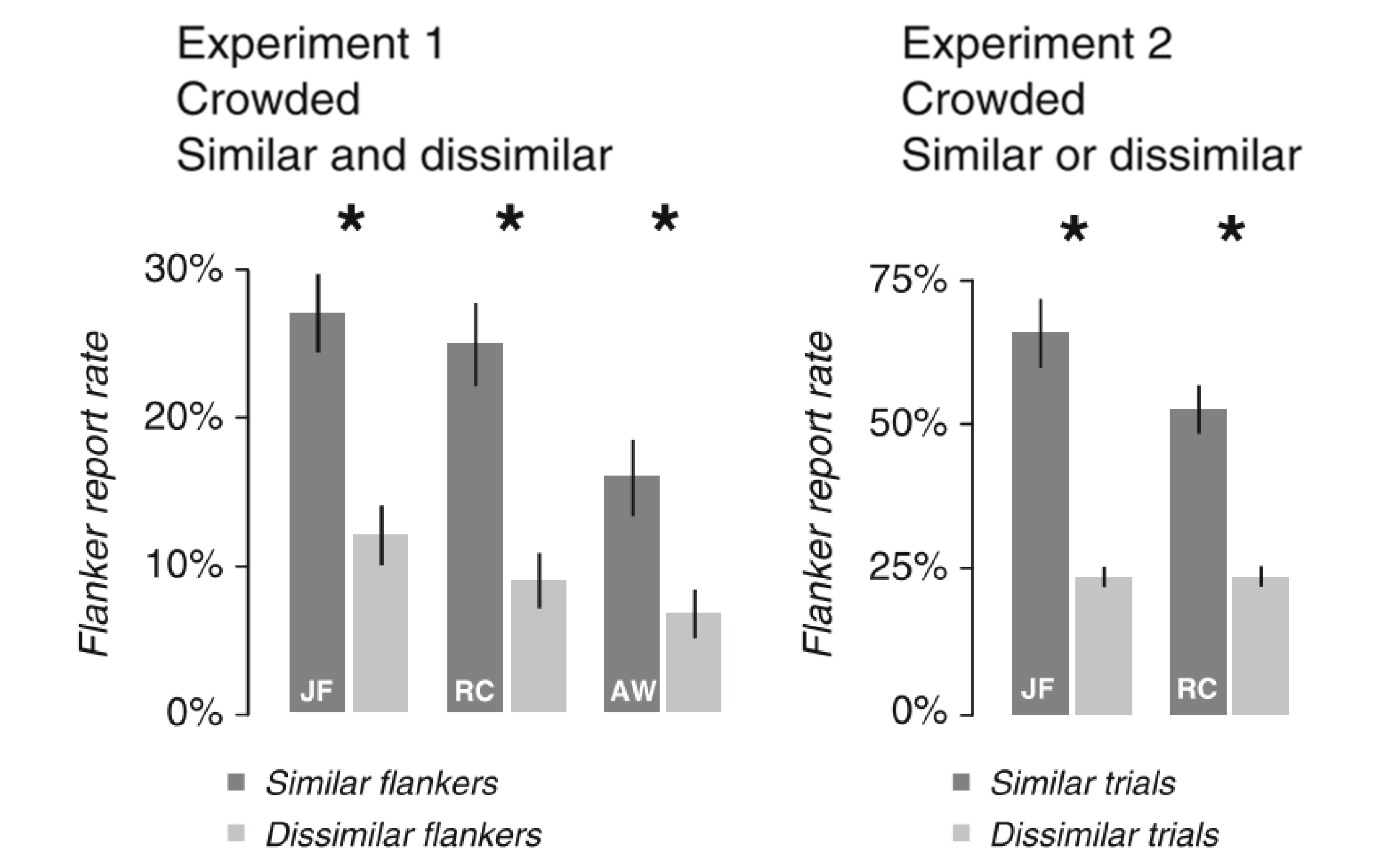


(Balas et al., 2009)

Substitution vs. pooling

- Freeman et al., 2012
- Substitution isn't the whole story
- Observers report "similar" flankers more often than expected by a pure substitution model

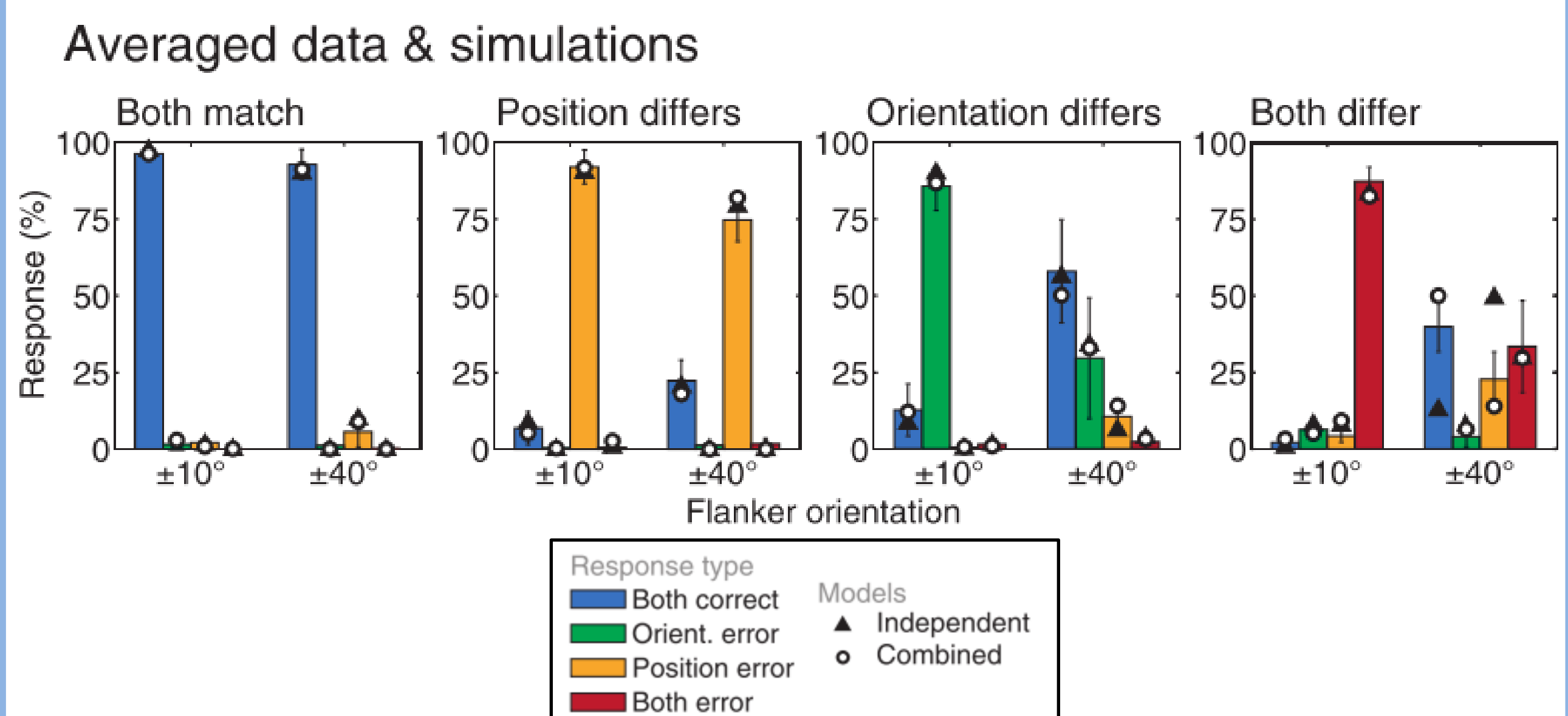
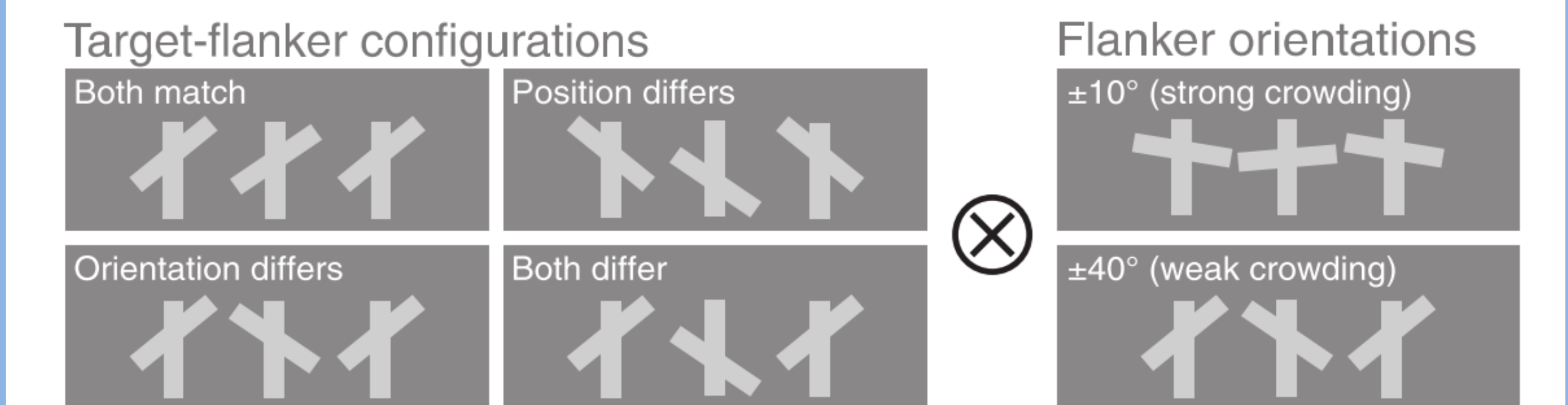
| | |
|------------|---|
| Target | A B C D E F G H I J K L M N O P Q R S T U V W X Y Z |
| Similar | N P G U L T C F B A X I E W Q R O K Z H D Y M S V J |
| Dissimilar | C K W F V Q N U R S Y M D G T B Z P O J E X I A L H |



(Freeman et al., 2012)

Late crowding

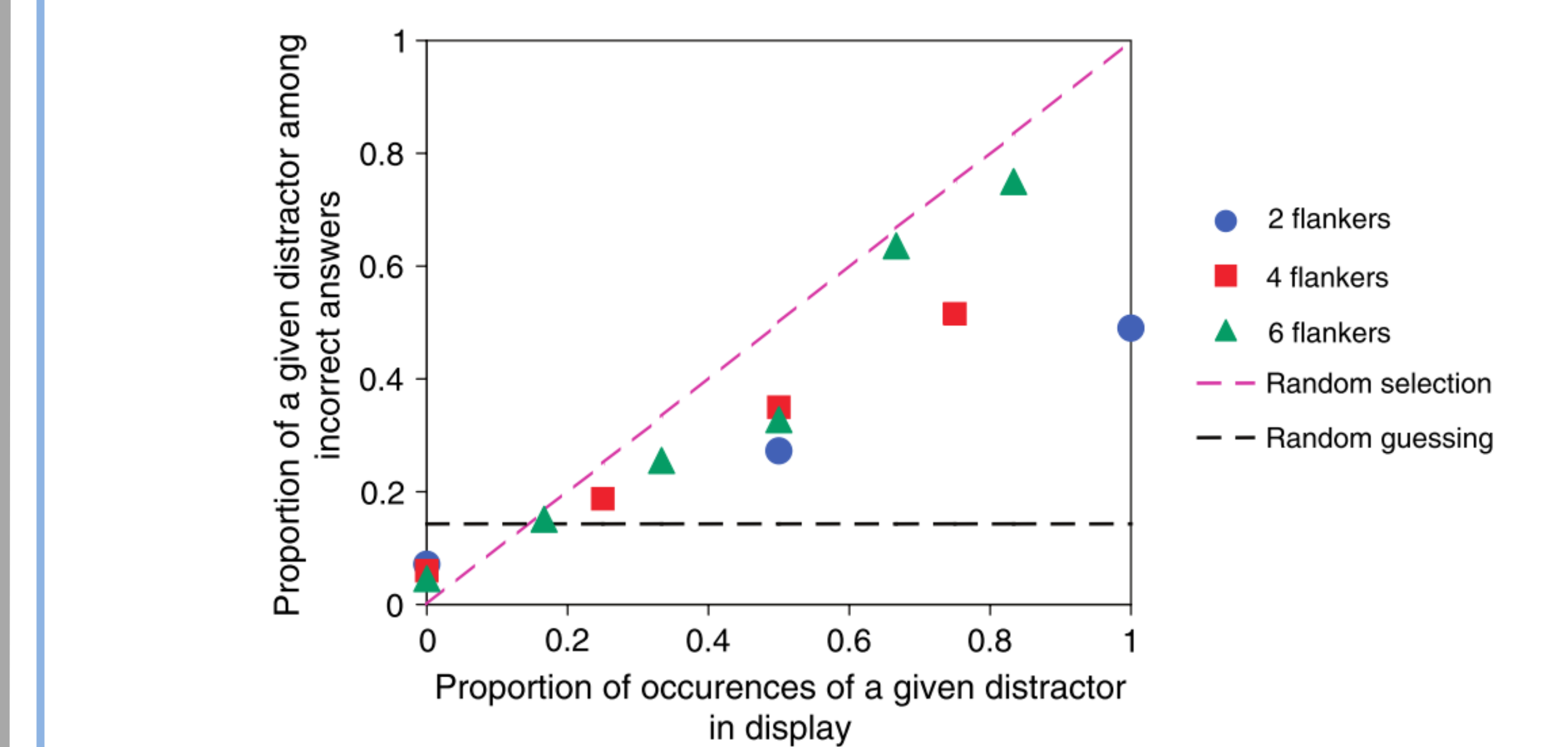
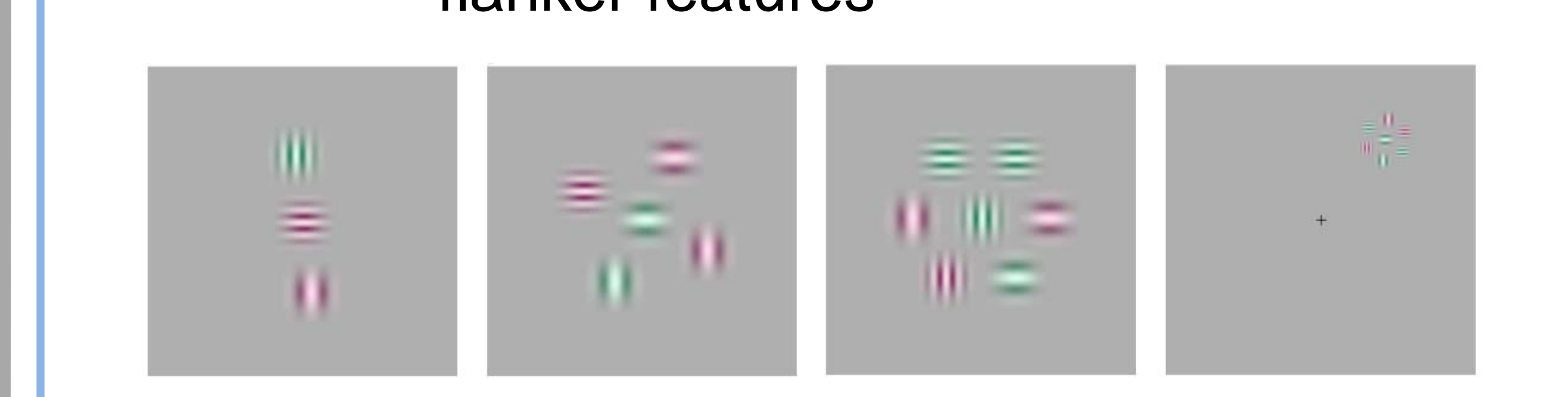
- Greenwood et al., 2012
- Individual features get encoded noisily, but crowding happens after binding
- Decreasing crowding in one feature dimension affects crowding in the other



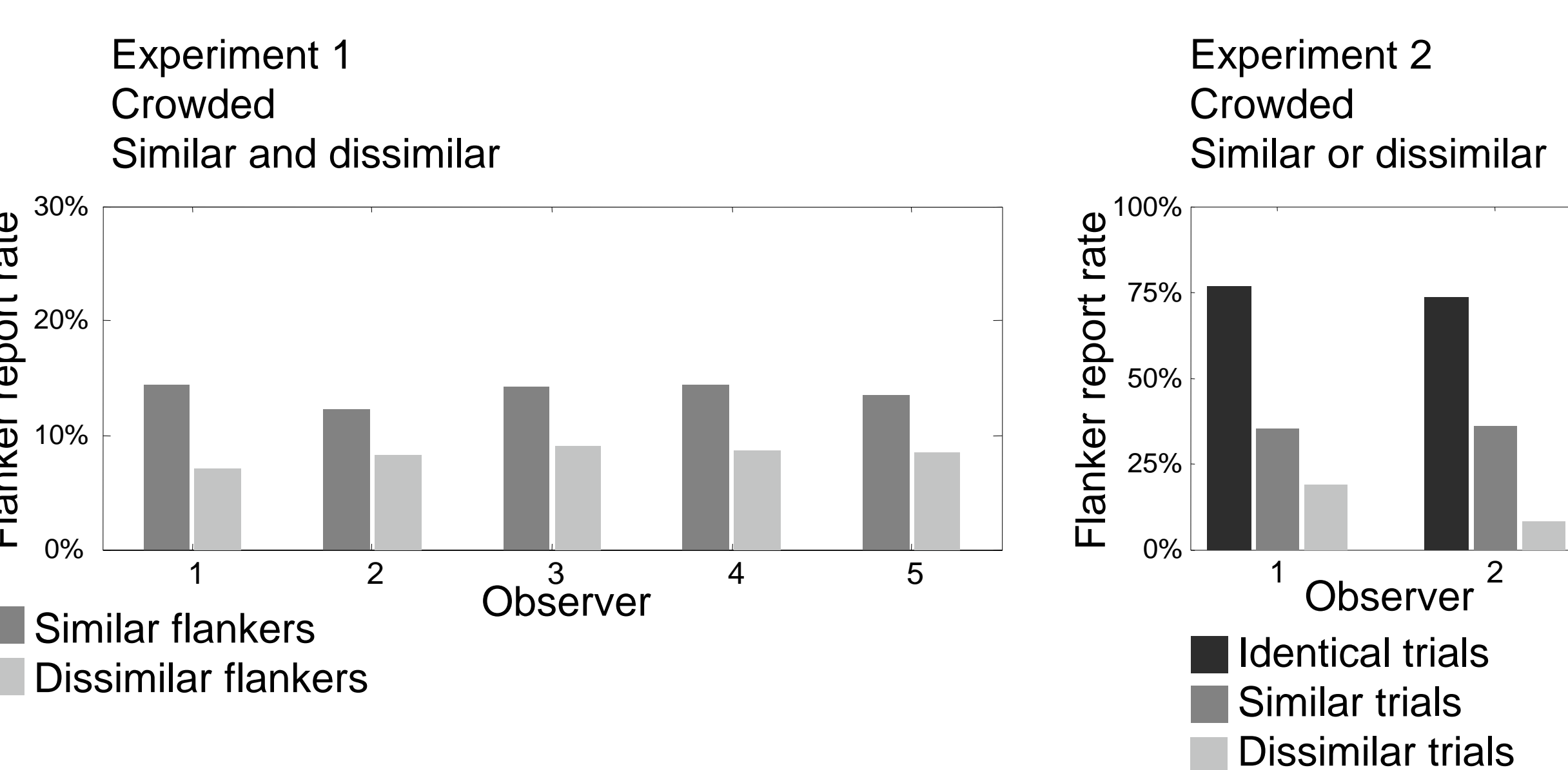
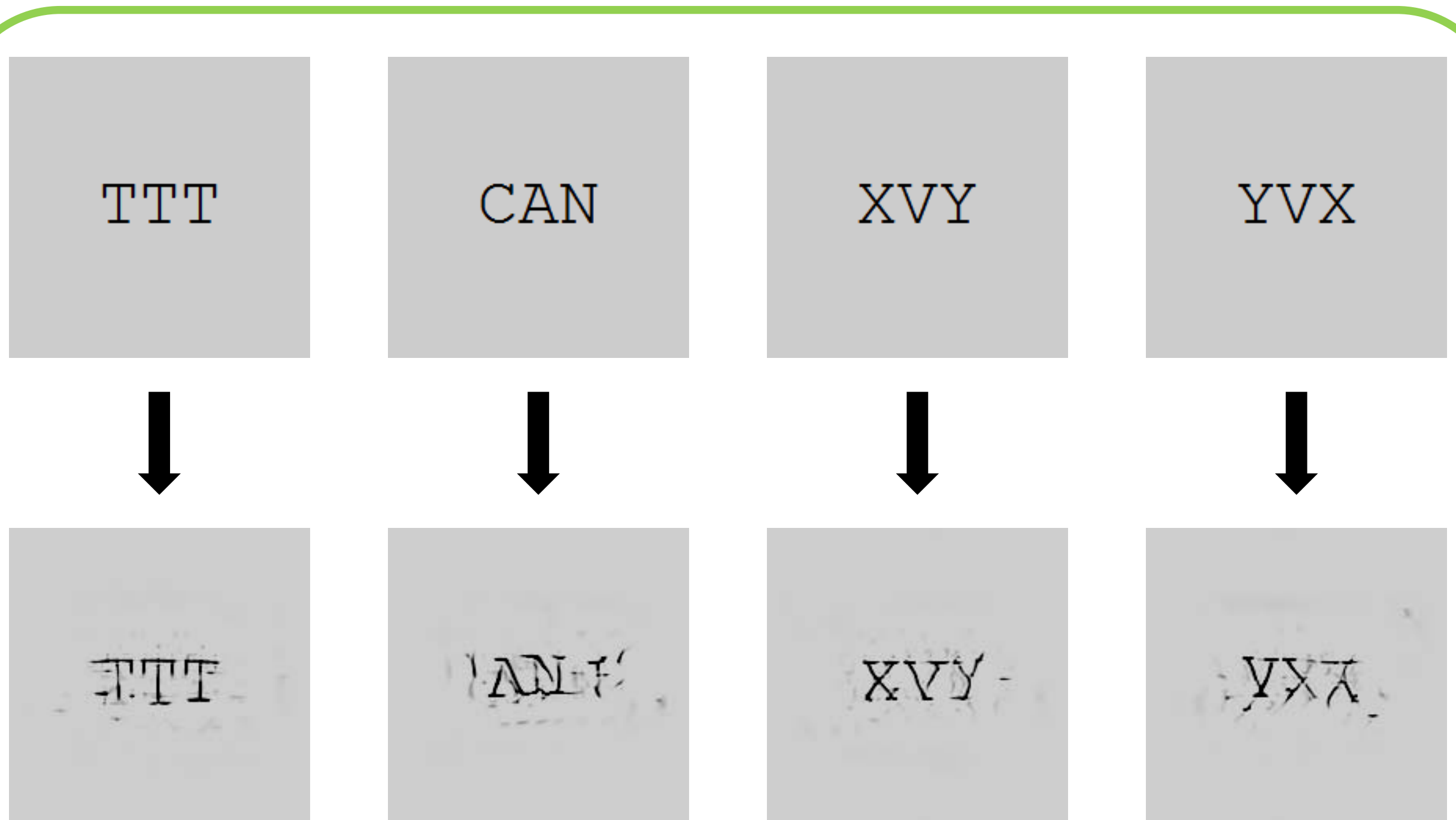
(Greenwood et al., 2012)

Early crowding

- Pöder & Wagemans, 2012
- Crowding causes both mis-bindings between flankers and target features as well as target mislocalizations
- Feature errors are biased towards the flanker features



(Pöder & Wagemans, 2012)

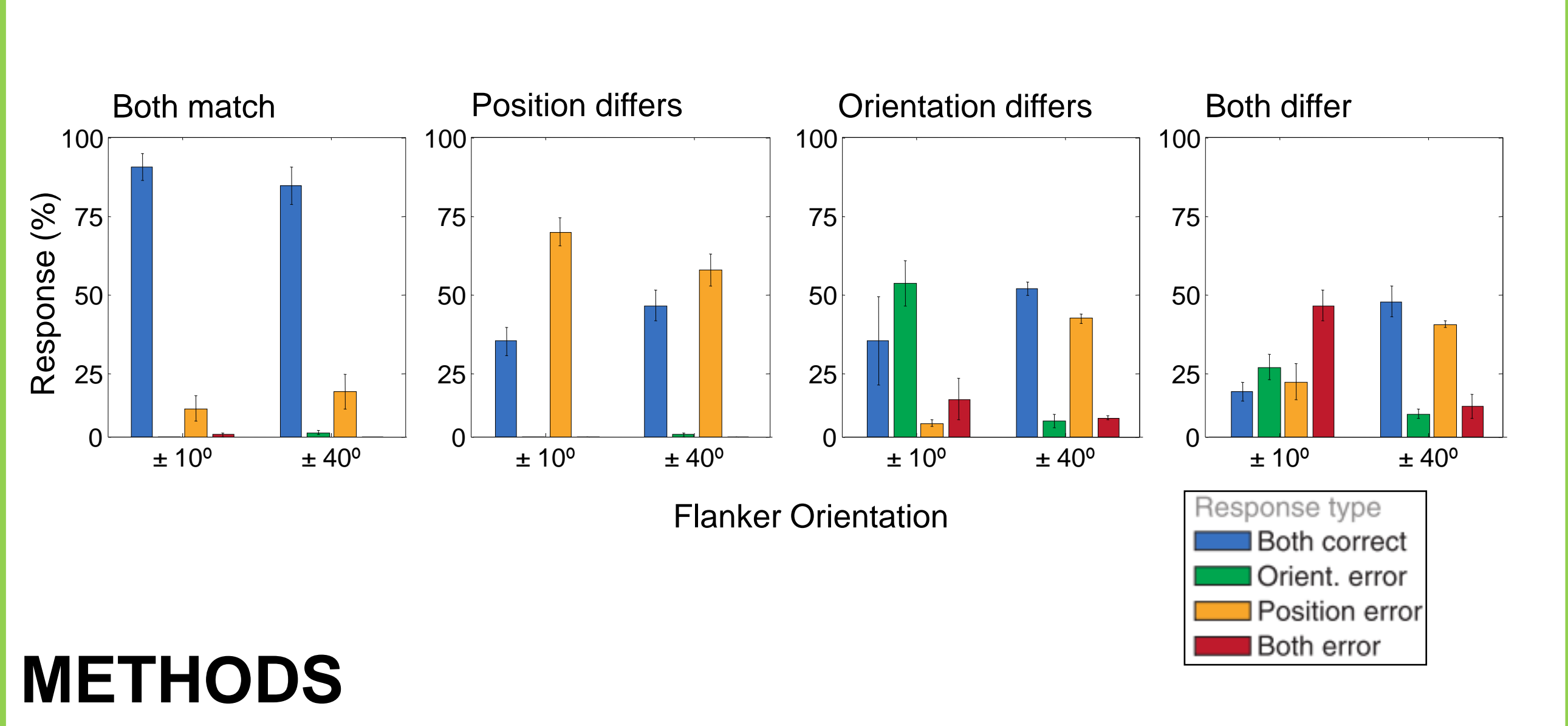
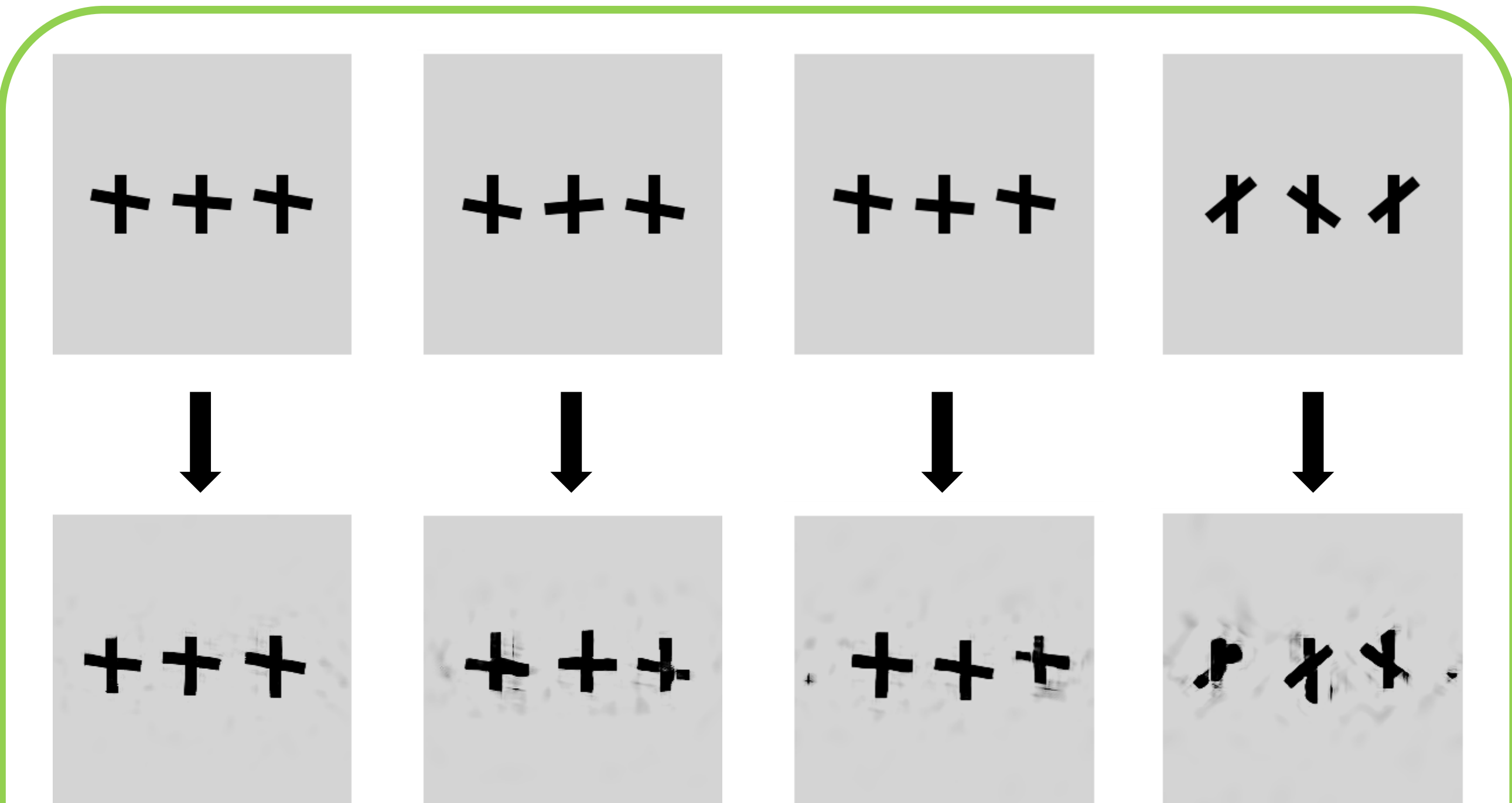


METHODS

- Experiments 1 & 2 used 10 synthesized stimuli per stimulus type in original experiments 1 & 2
- Observers reported three ordered letters

RESULTS

- Observers in the mongrel task also report the similar flanker more often than the dissimilar flanker
- Specific letter combinations show different effects

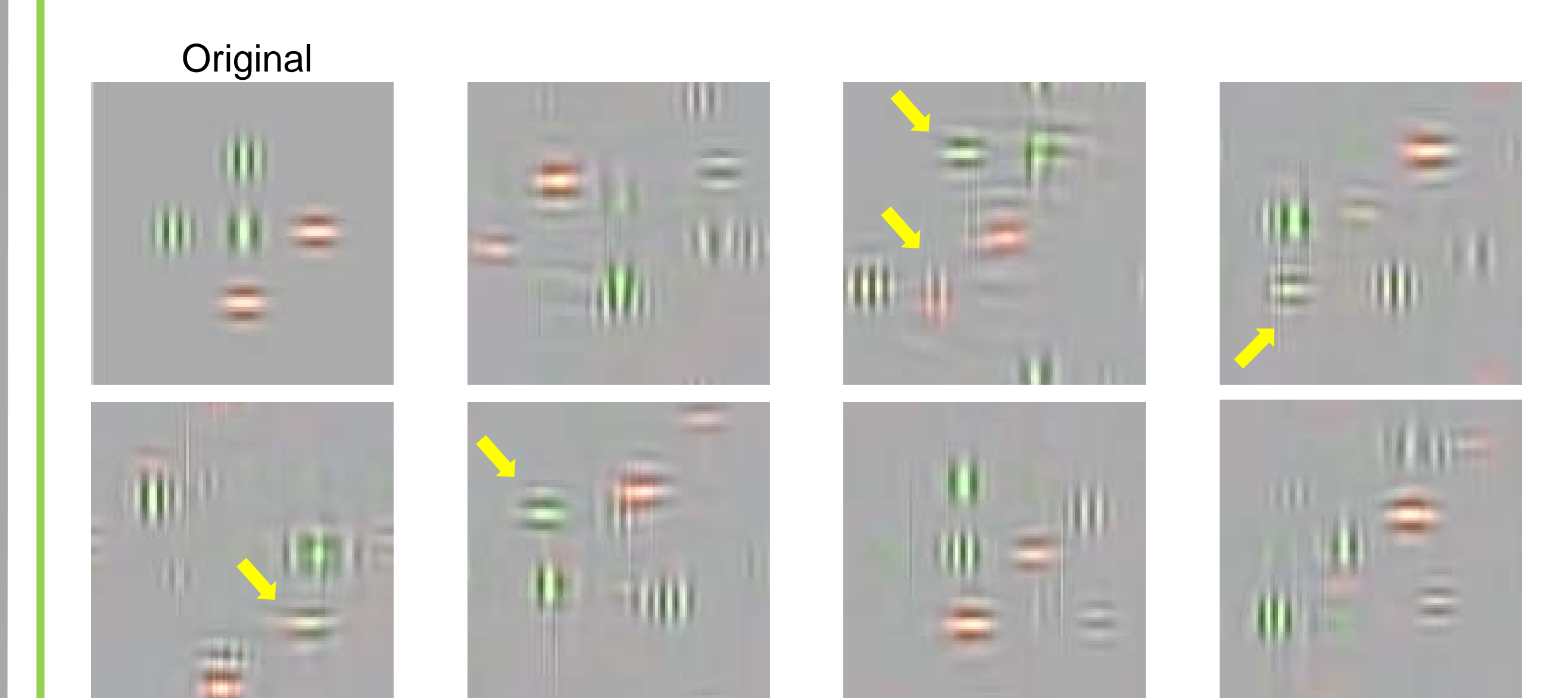
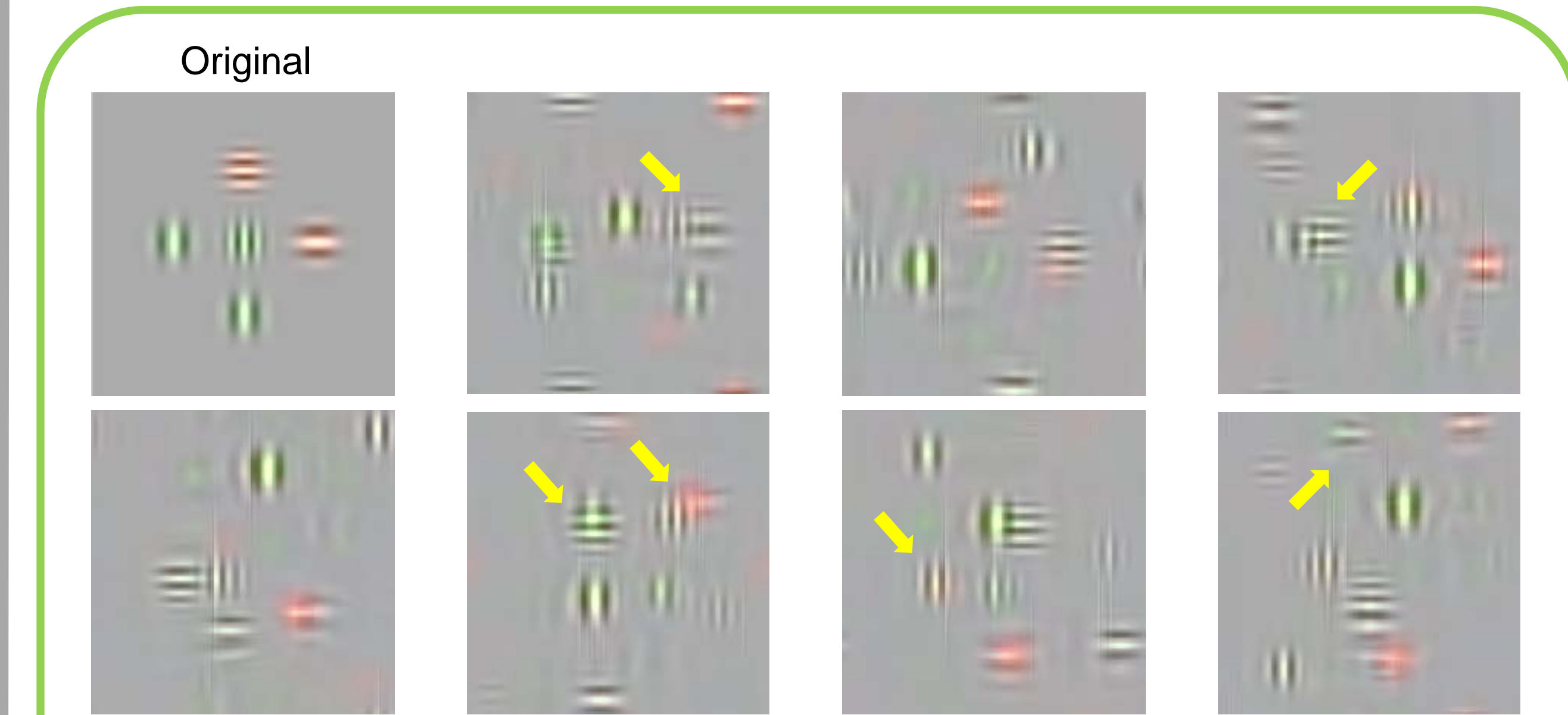


METHODS

- Experiment used 10 synthesized stimuli per stimulus type in original experiment 4
- Observers reported the position and orientation of the unique tilted bar (4 AFC task)

RESULTS

- Observers in the mongrel task also show a relative decrease in position errors when crowding is "released", indicating feature binding



- Color channels separated by PCA, steerable pyramid computed independently in channels, then statistics computed as in Portilla & Simoncelli, 2000
- Evidence for color-orientation misbindings

CONCLUSIONS

- A visual texture model can capture some disparate effects observed in the crowding literature while making predictions about arbitrary images
- A more descriptive pooling model can lead to different binding and substitution effects