Single Anchor Sorting of Visual Appearance as an Oriented Graph

Nathan Moroney, Ingeborg Tastl and Melanie Gottwals

Print Adjacencies & 3D Lab, HP Labs, HP Inc.

Michael Ludwig[1,2] and Gary Meyer[1]

University of Minnesota[1]

Google[2]

IS&T Color Imaging Conference, CIC26

Vancouver, BC

November 2018

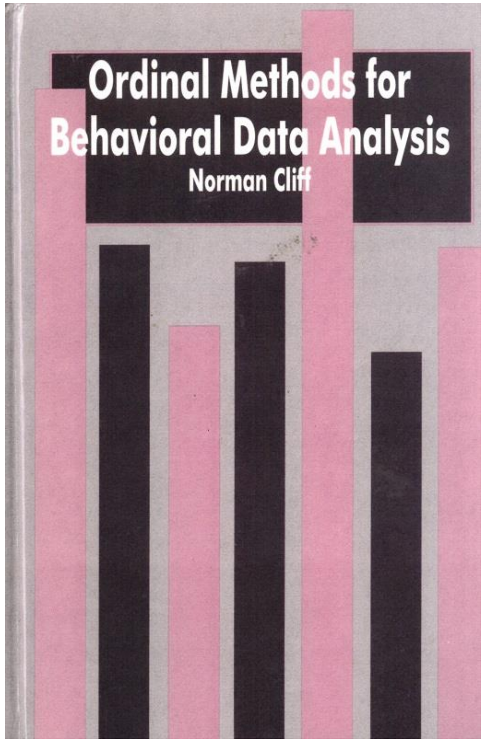
Outline

TL;DR – perceptual similarity as a graph

- Motivation and context
- Ranking stimuli relative to an anchor
- Two experiments
  - Web-based ranking of 3DP complex appearance
  - Laboratory ranking of solid colors
- Representation and analysis
  - Graph in which nodes are stimuli and edges are the rank sequence
  - Thickness of edges is proportional to rank for all observers
  - Rank aggregation using Schulze voting method
- Evaluating models
- Future directions

Motivation and Context

Ordinal Methods for Behavioral Data Analysis was an interesting read

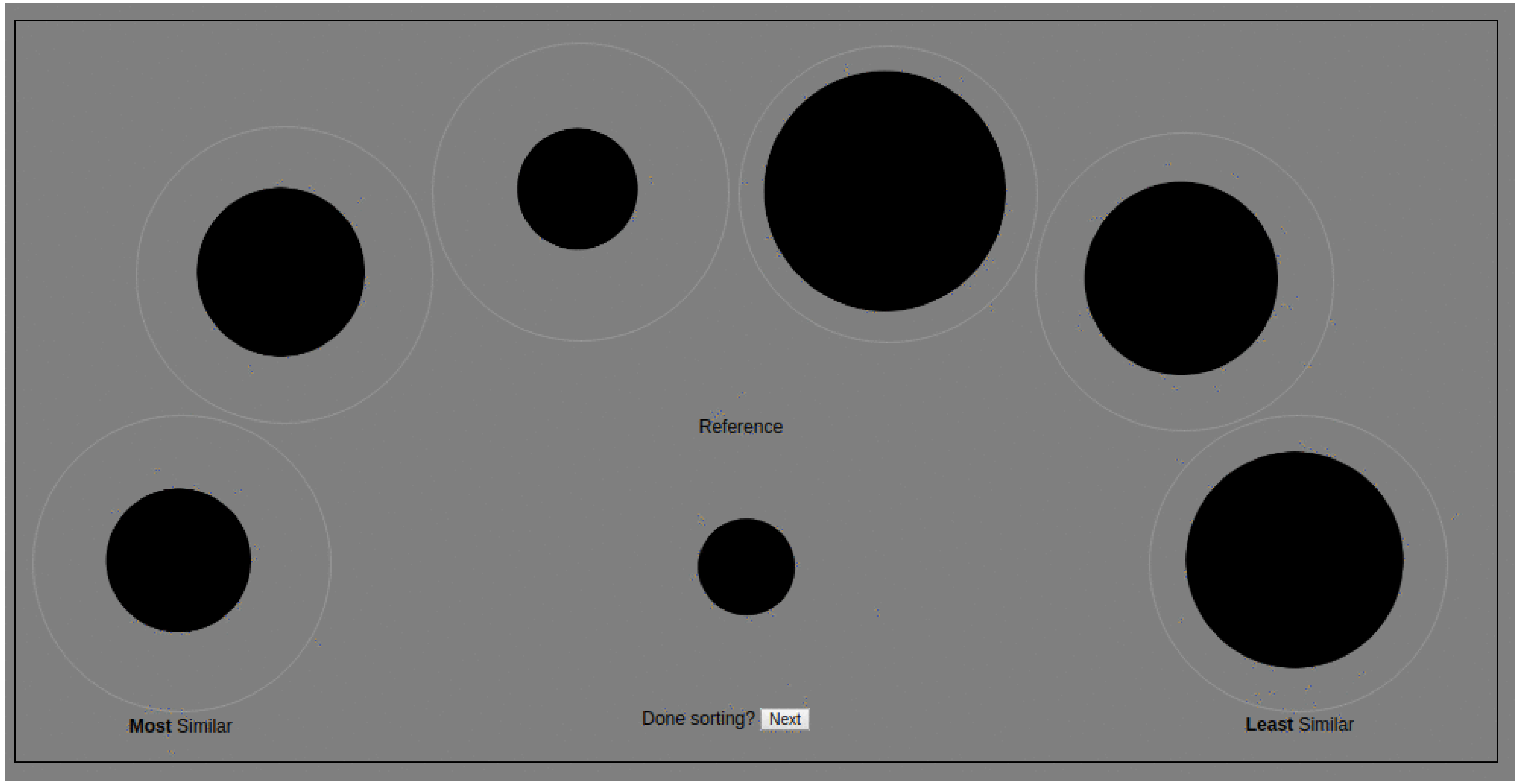


- Even efficient methods of paired-comparison can be time consuming and volunteers may complain
- Interested in rank order methods and metrics
- Perhaps most familiar example in the field, Farnsworth-Munsell 100 hue test
- Engledrum describes ordinal methods and averaging ranks, for a single *-ness*
- Progress made during Michael's summer internship and subsequent collaboration with Gary Meyer:
  - Ludwig, Michael, et al. "Perceptual Appearance Uniformity in 3D Printing." *Electronic Imaging* 2018.8 (2018): 1-12.
- Recent results in web-scale data collection for perceptual metrics:
  - Zhang, Richard et al. "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric", CVPR (2018).

Anchored Ranking

Rank perceptual similarity relative to a single reference

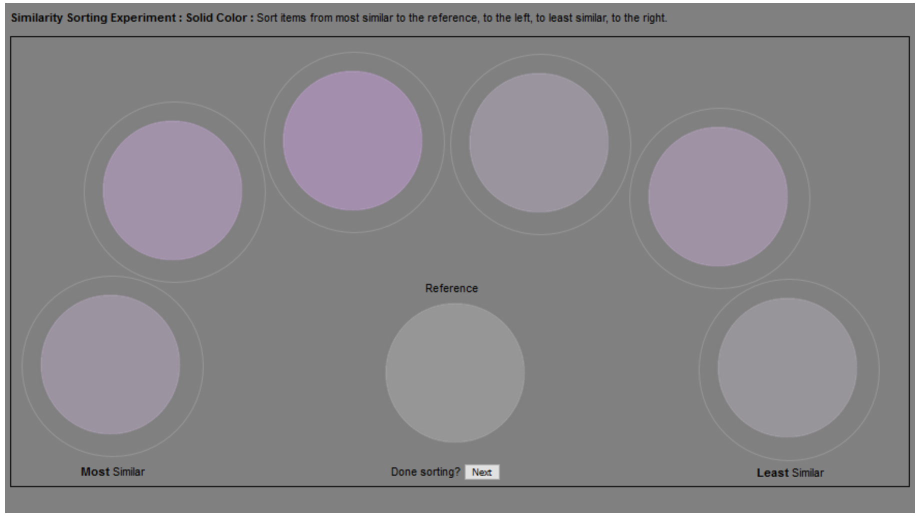
- Simple example varying only the diameter
- Drag-and-drop as interface
- **Center** is reference
- **Left** is most similar
- **Right** is least similar
- Arc layout is used to allow larger stimuli and roughly equal distance for comparisons to reference stimuli

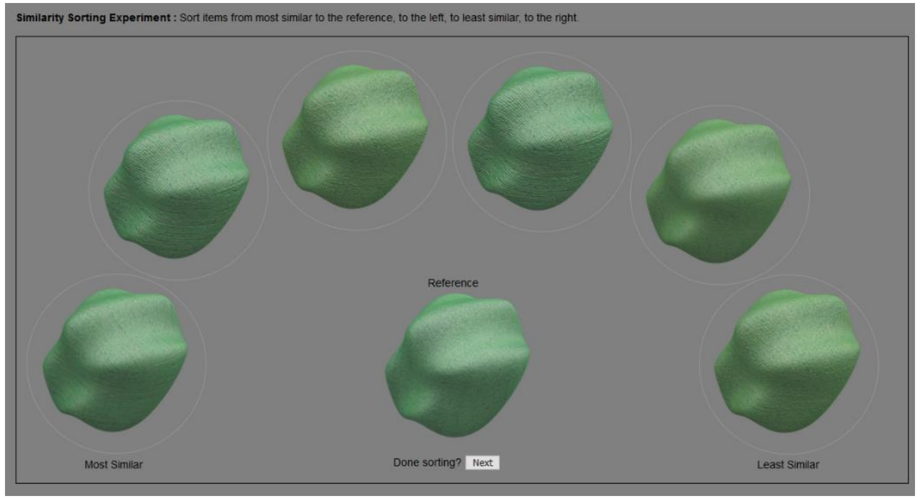


Two Visual Experiments

In all cases, one reference and six stimuli to sort

- Solid colors
  - Generated with a target color difference between neighbors
  - Conducted in laboratory with calibrated display and viewing conditions
- Simulated complex appearance of 3D printed blob
  - Generate varying two rendering parameters per sequence
  - Conducted online with unknown displays and viewing conditions
- Both experiments started with two test sequences
  - A Deuteronomic color sampling
  - A sequence of abstract shapes

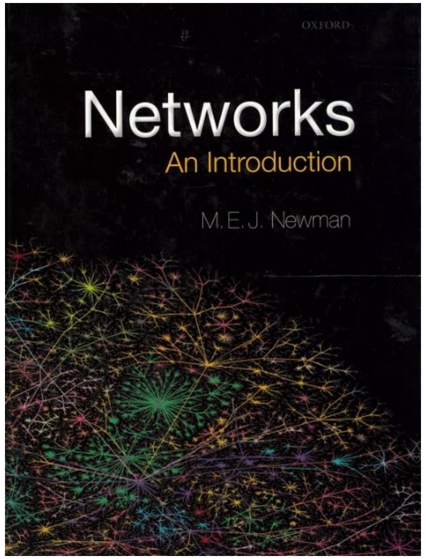




Graphing Similarity

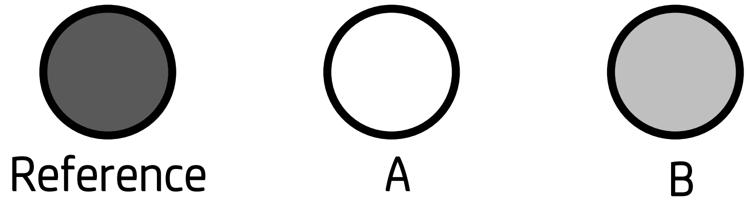
Networks: An Introduction was also an interesting read

- Perceived similarity as a directed graph
  - **Nodes** : stimuli
  - **Edges** : sort sequence neighbors (or rank order)
- In the proceedings referred to as a sort-sequence graph
- For simplicity will show as undirected graph after this slide

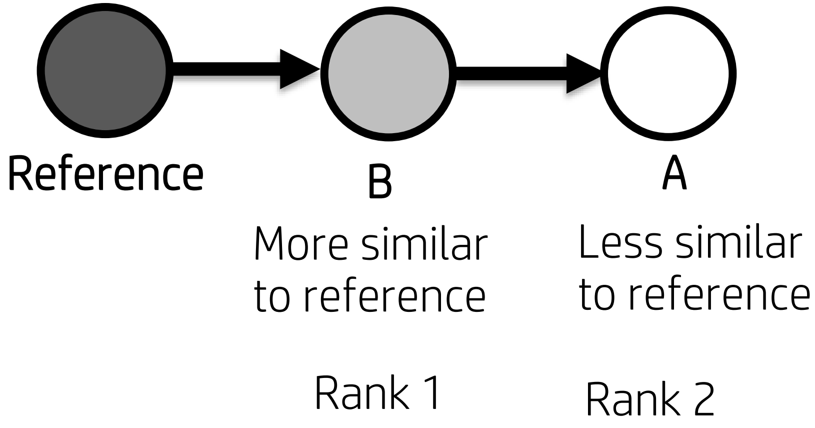


Lots of metrics, tools and concepts in graph theory and network analysis to possibly leverage for perception

Stimuli



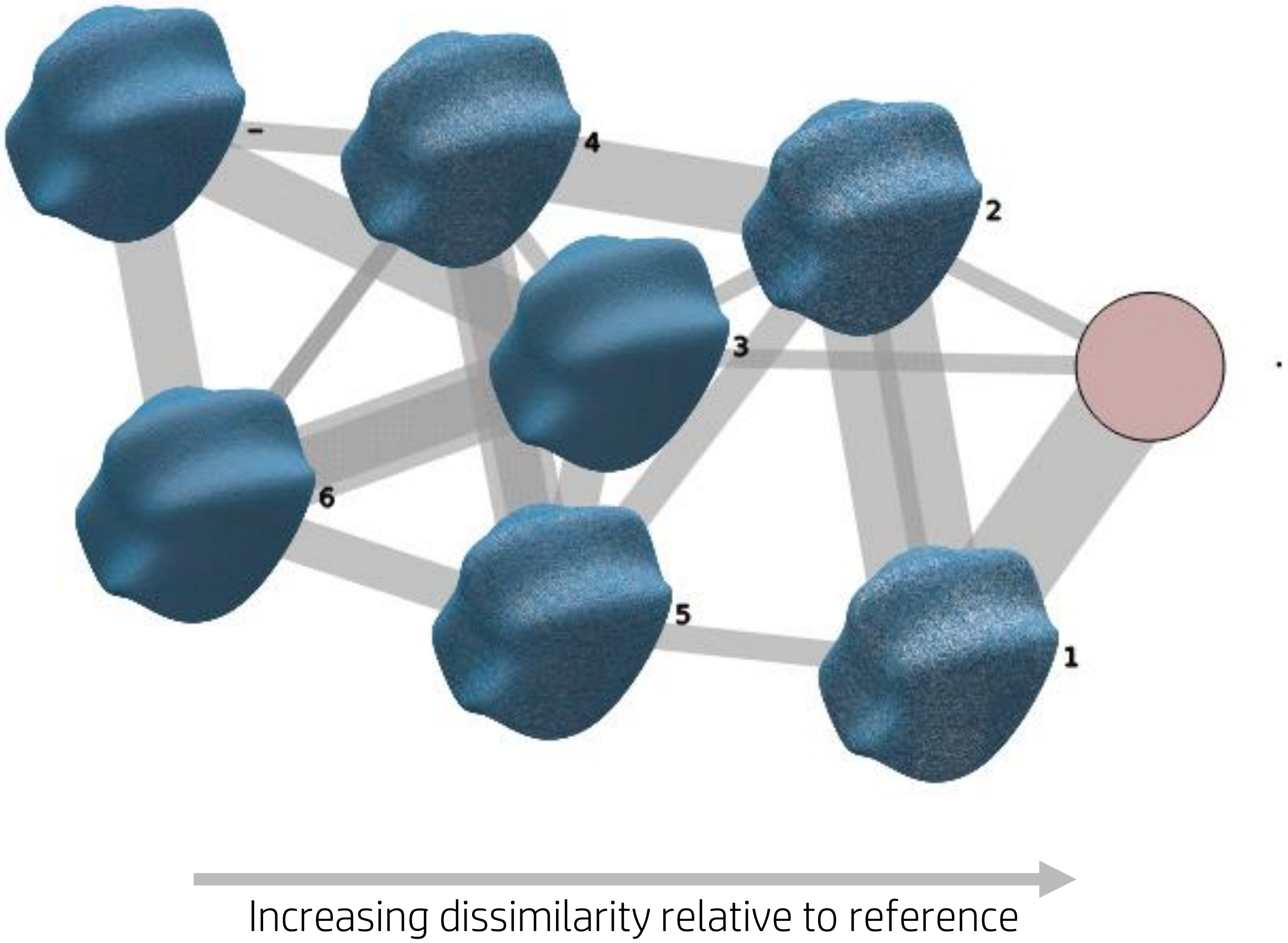
Sorted Stimuli



Oriented Graphs

Establishing a convention for "reading" the graphs

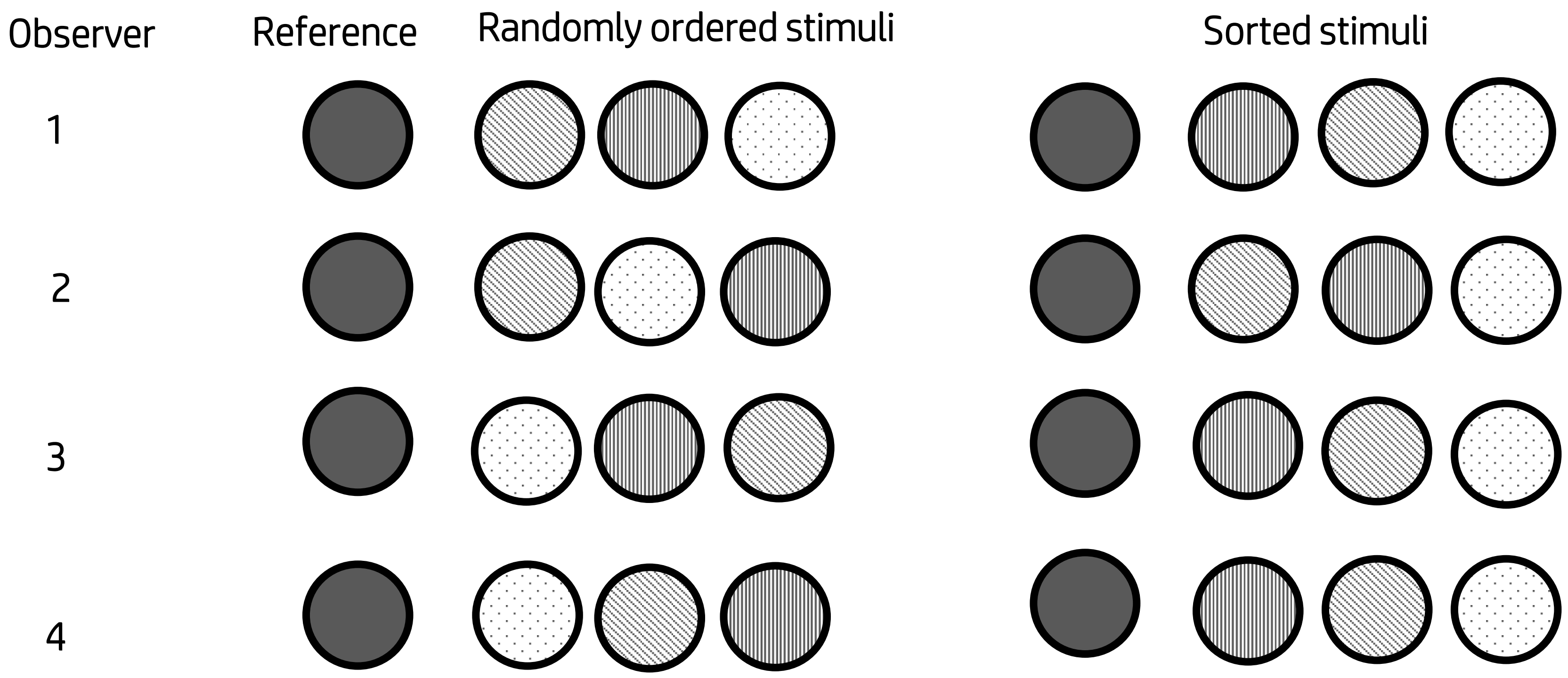
- Algorithms for graph layout is its own research area
  - D3.js force-directed layout for initial layout
  - Manually orient increasing dissimilarity from left to right
  - Similar to the number line in mathematics and matches the experimental task instructions
- Note "oriented" is not an established term in the terminology of graph theory
- Additional conventions
  - Dash or – as label for reference
  - Period or . with empty node for "end of graph"



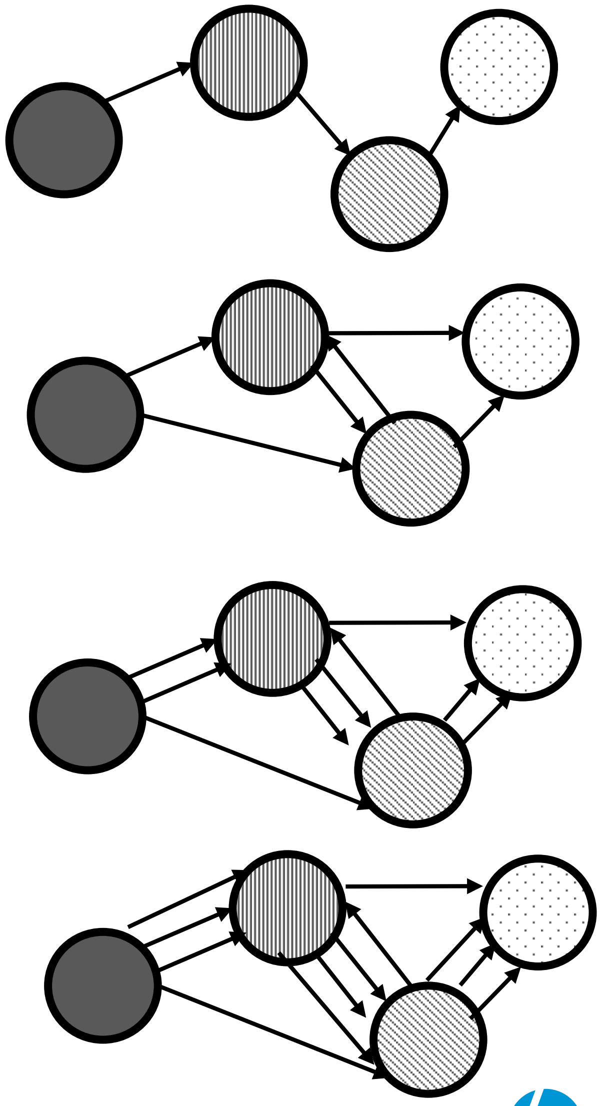
Weighted Graph

Consider using when the dissimilarity matrix just looks like a table of numbers

- Thickness is proportional to number of times a node rank order occurs



Accumulating weights

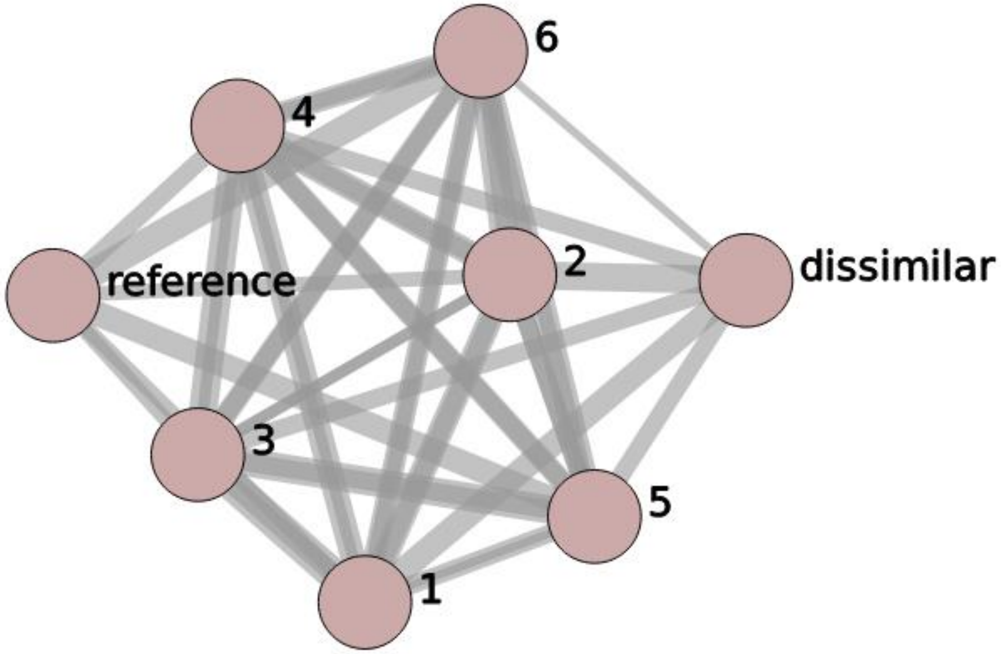


For each observer, connect corresponding neighboring ranks.
Convert number edges to thickness. Above pattern fills only an example set of stimuli.

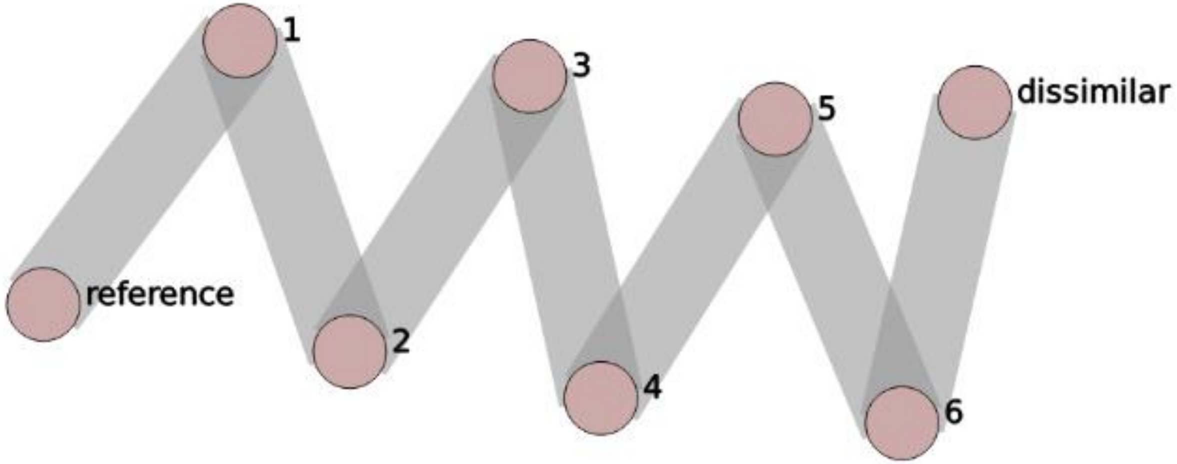
Computational Graphs

Graphs that can be directly simulated and used for comparison purposes

Two examples shown below, other processes and models possible.
But a useful example of one pair of generative endpoints.



Uniform random sorting

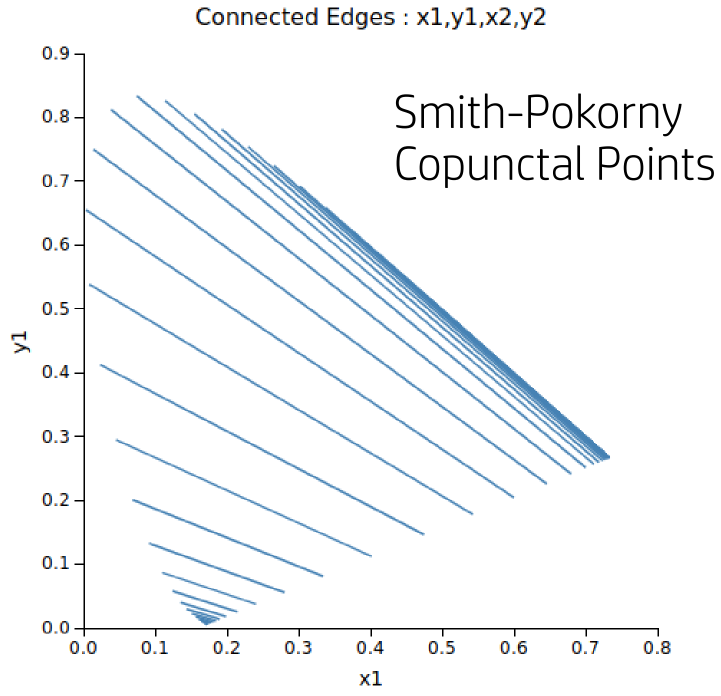


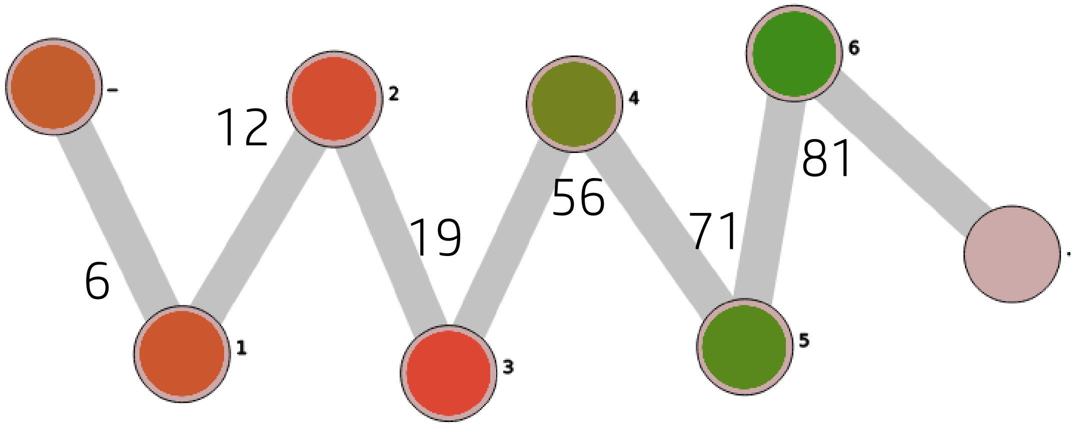
Perfect order sorting

Test Sequence One: Deuteranopic Confusion Colors

Ramp of equal luminance solid colors from a reddish to a greenish color

- 70% of 150+ participants provided a sort in perfect agreement based on ΔE^*_{ab} assuming sRGB as display, shown at bottom right
- 23% of participants had only within-color term disagreements (i.e. only oranges out of order or only greens
- 7% of participants have cross-color term disagreement (i.e. mixing oranges & greens)
- The above 7% can be deuteranopic or 'adversarial' (I've estimated ~4% adversarial rates for past experiments)



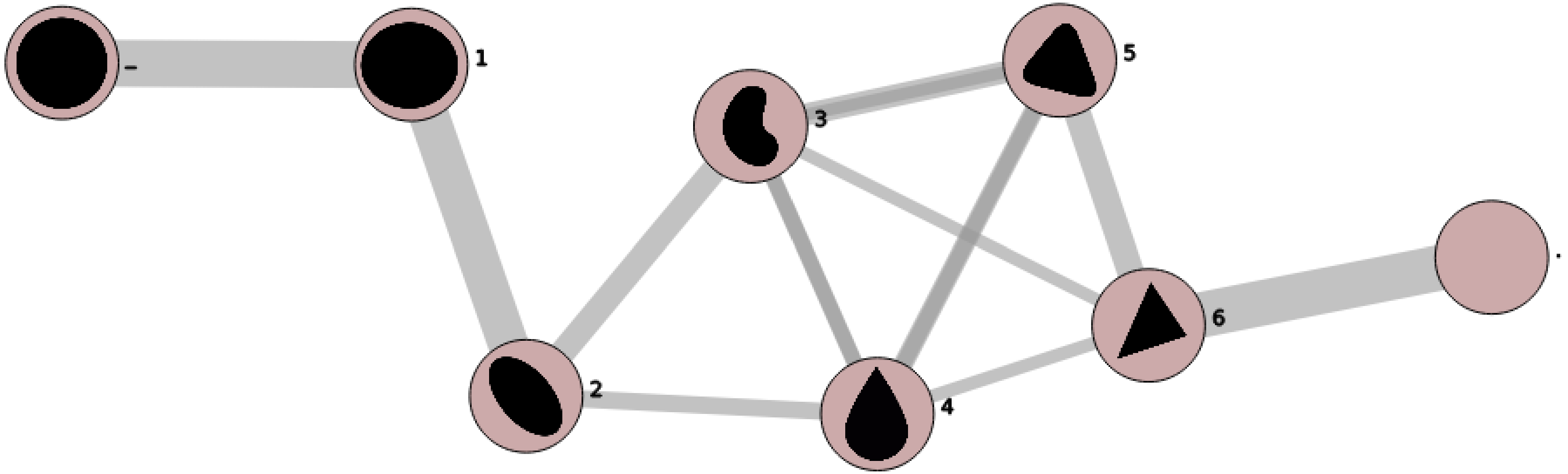


Larger numbers above are ΔE^*_{ab} w.r.t. the reference.

Test Sequence Two: Simple Shape Sequence

Ramp of solid black shapes from circular to triangular

This test (and the previous) are useful for assessing instructional clarity and/or identifying adversarial participants



Rank Aggregation

Schulze voting method or widest path

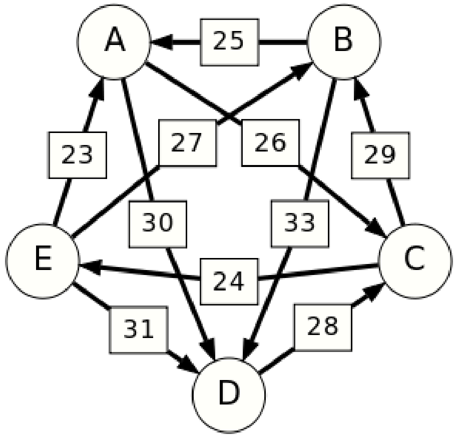
number of voters	order of preference
5	$ACBED$
5	$ADECB$
8	$BEDAC$
3	$CABED$
7	$CAEBD$
2	$CBADE$
7	$DCEBA$
8	$EBADC$

45 voters provide rank preferences for 5 candidates

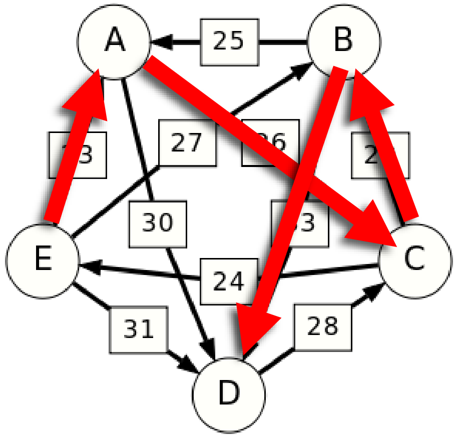
Matrix of pairwise preferences

	$d[*,A]$	$d[*,B]$	$d[*,C]$	$d[*,D]$	$d[*,E]$
$d[A,*]$		20	26	30	22
$d[B,*]$	25		16	33	18
$d[C,*]$	19	29		17	24
$d[D,*]$	15	12	28		14
$d[E,*]$	23	27	21	31	

Directed preference graph



Variant of Floyd–Warshall algorithm



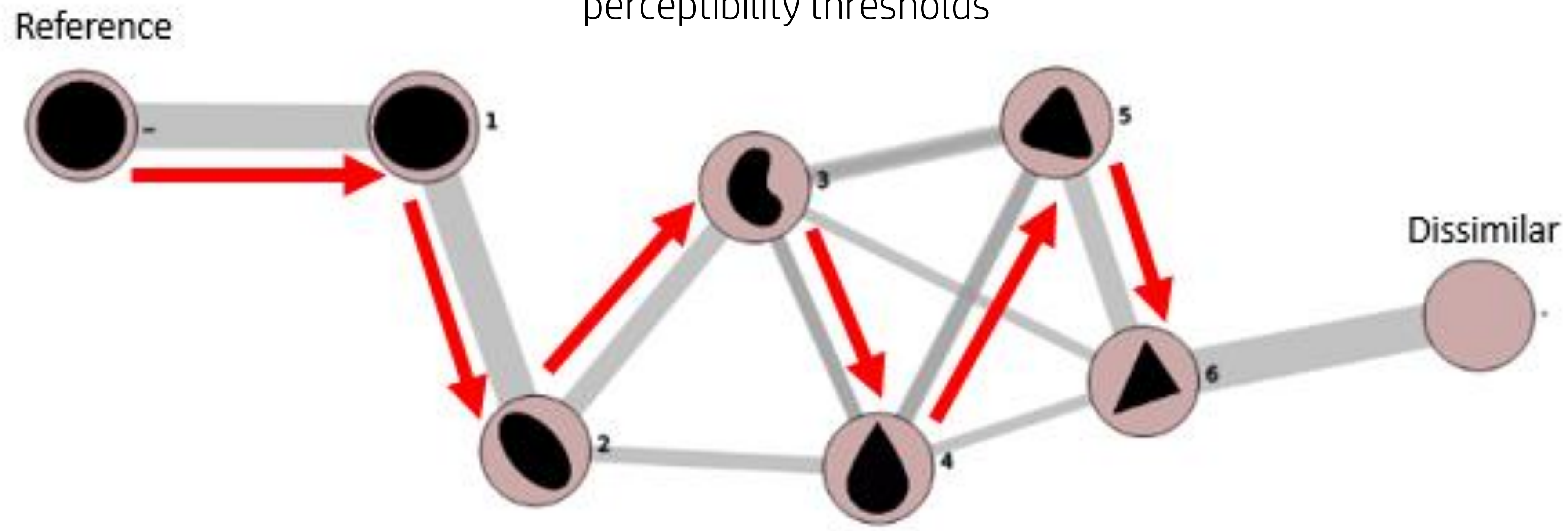
Aggregated rank or widest path is shown as red sequence

Graphics on this slide are from *https://en.wikipedia.org/wiki/Schulze_method*

Test Sequence Two: Simple Shapes Widest Path

Circle, ellipse, bean, tear drop, rounded triangle and finally triangle

Variations here are likely due to multiple cognitive sorting criteria, and not approaching perceptibility thresholds

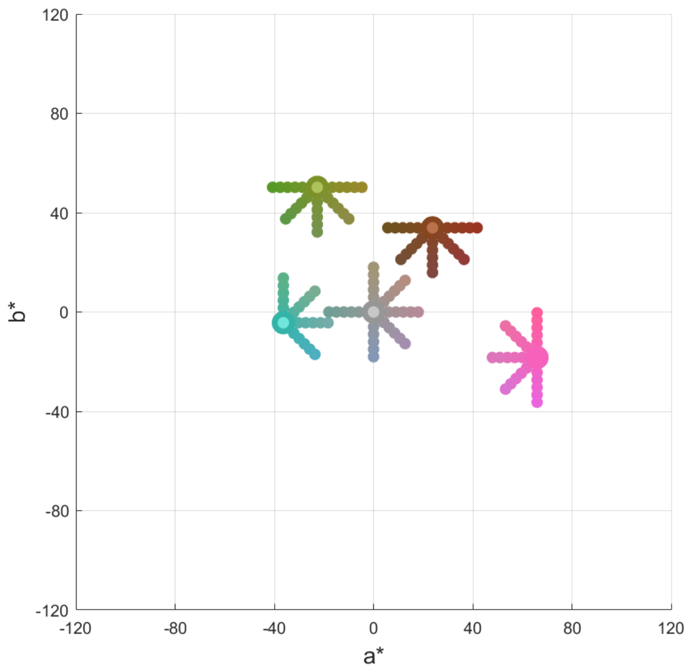


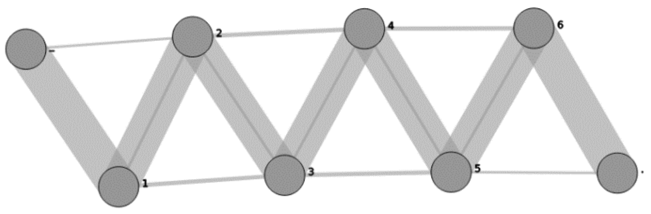
The inverse ordering of this path (the "wrongest path") is of interest as a possible method for outlier detection.

Solid Color Sorting

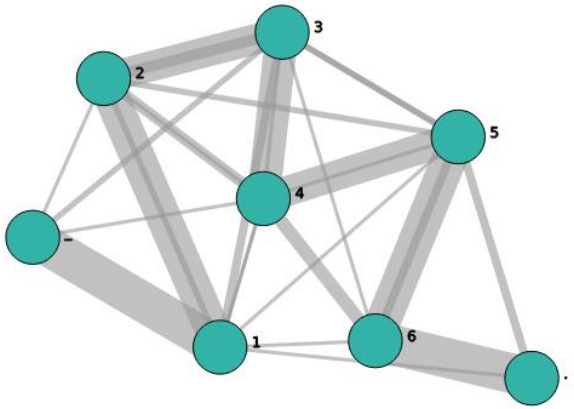
Sorting is easy enough, repetitions are feasible

- Five nominal centroids : olive, brown, pink, teal and gray
- 25 different directions (in sRGB gamut)
- Four ΔE^*_{ab} step sizes : 3, 2, and 1
- Use sRGB display in dark surround, four observers
- Reference is centroid
- Rank sequence is relative to stimulus set that is monotonic with color difference

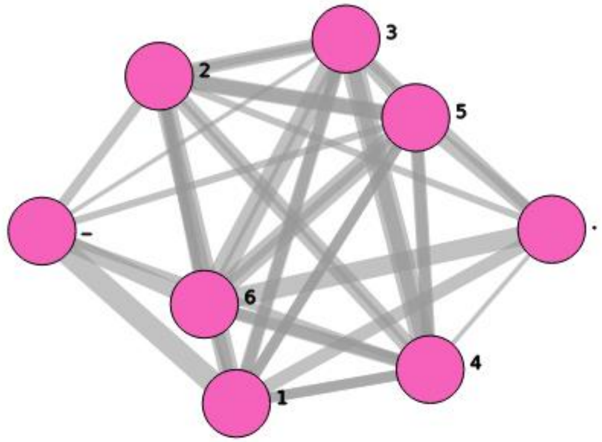




1 ΔE^*_{ab} : all grays, all trials



1 ΔE^*_{ab} : all teals, all trials

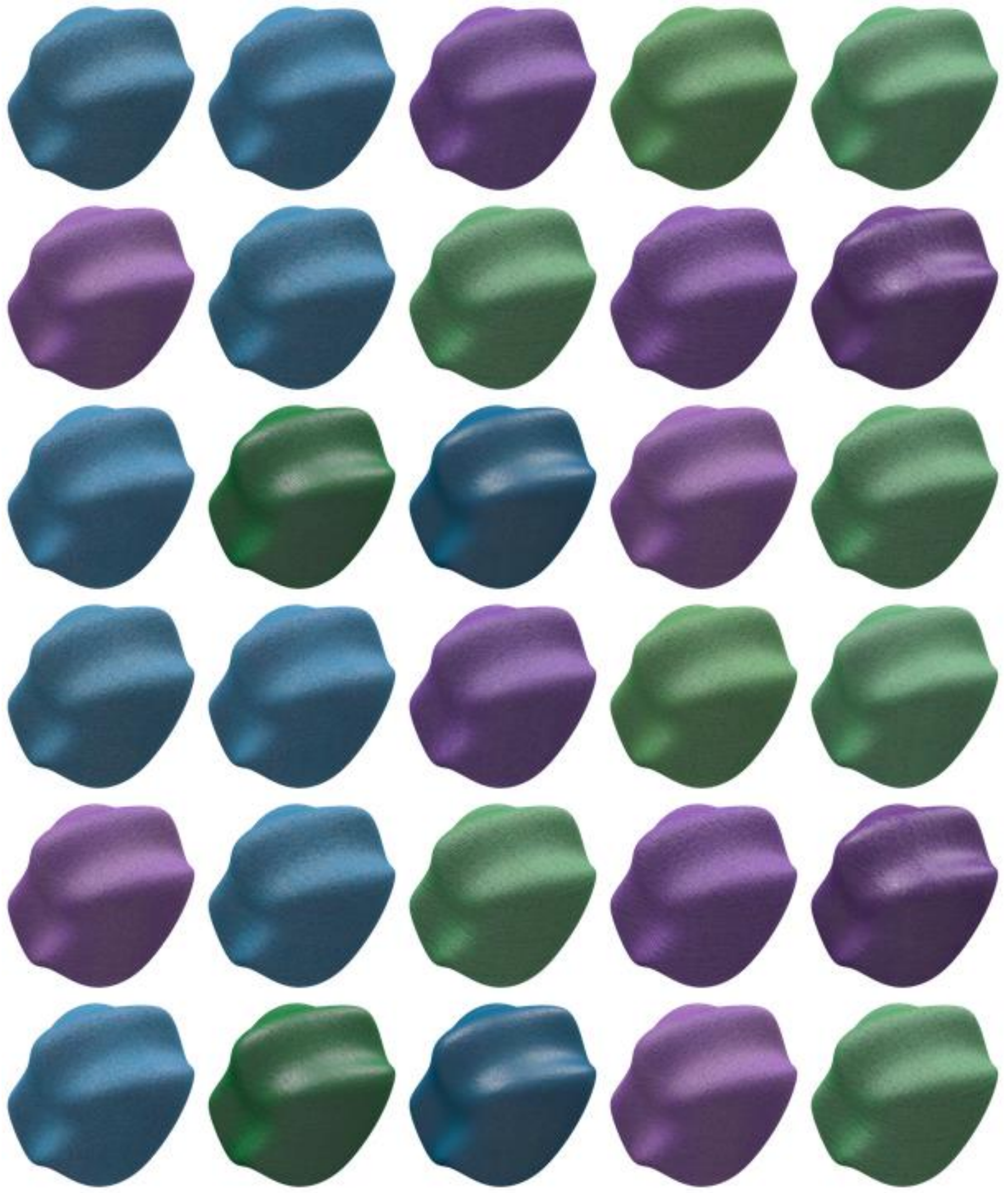


1 ΔE^*_{ab} : all pinks, all trials

Sorting Complex Appearance

Add subtitle

- HP MultiJetFusion or MJF 3D printed tiles were scanned using an X-Rite TAC 7
- Resulting AxF[tm] or Appearance Exchange Format data post-processed to magnify or diminish one or more rendering parameters
- Qualitatively the diffuse color, gloss, roughness and graininess
- Generate 30 sequences that have bivariate or two parameter variations in the rendering
- Randomly select 6 sequences for each observer as an all volunteer web-based experiment

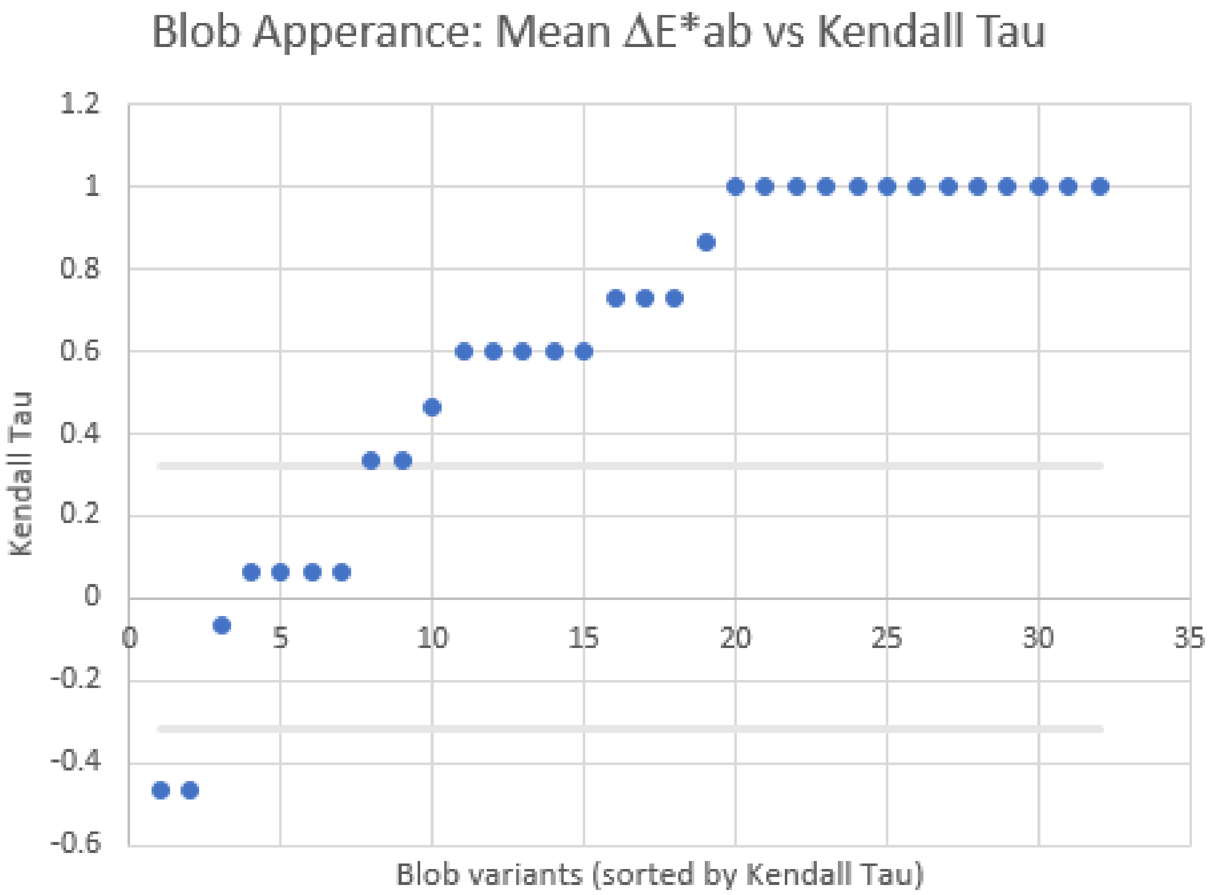


Blob object via the X-Rite Pantora™ software with acknowledgements to Prof. Roland W. Fleming's design

Evaluating Models

How good is simply using ΔE^*_{ab}?

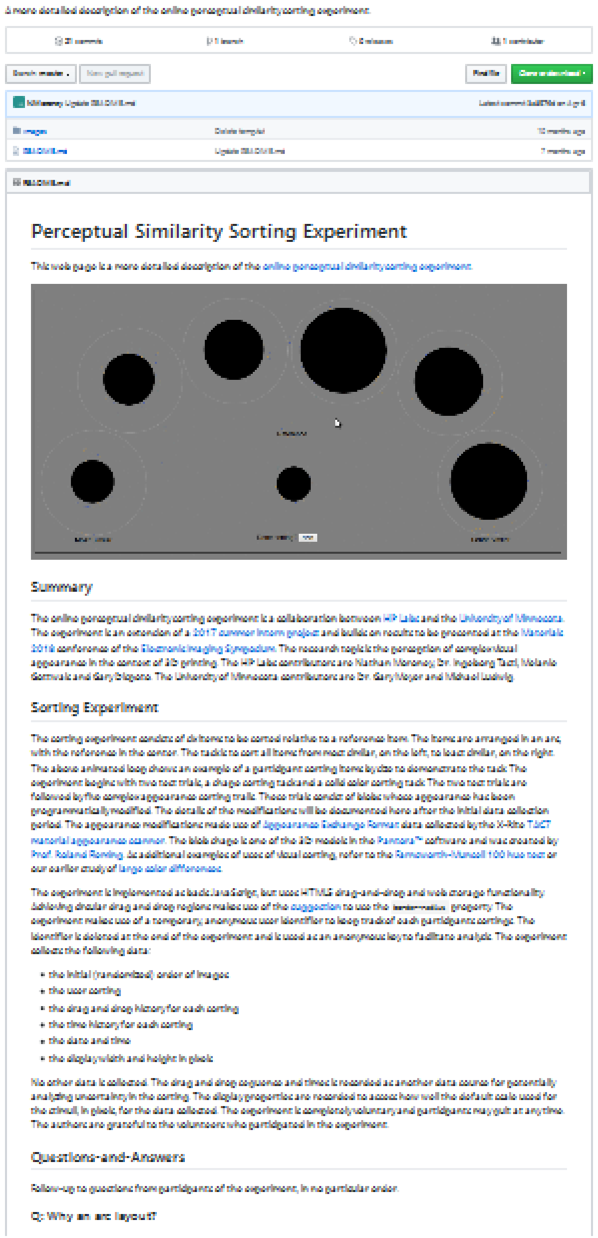
- Use variant a of Kendall's Tau or T_a to compare aggregated observer ranks versus the rank order as predicted by average pixel DE*ab
  - Simple baseline model for future comparisons
- Sort the sequences by T_a and plot the resulting ordered trails
- The gray bars are the plus/minus 1 standard deviation for a "random ranker"
- Roughly 2/3rds have T_a value of > 0.6



Future Directions

Obligatory plug for a related web site

- Graph-based similarity analysis; kernels and classification
- Using the rank data directly, in fact this is something Michael has done
  - Ludwig, Michael, et al. "An appearance uniformity metric for 3D printing." *Proceedings of the 15th ACM Symposium on Applied Perception*. ACM, 2018.
- A focused investigation of color differences at the extremes of large gamut displays
- For JND situations, implement a “logistic” ranker for comparison purposes
- Revisit our large color difference experimental sorting data from CIC22
- All data is accessible on public github page
  - https://github.com/NMoroney/SimilaritySorting



Summary

Now would be a good time to think of a question or two

- Sorting stimuli with respect to a single anchor can be represented as an oriented weighted directed graph
- Rank aggregation can be performed using the Schulze voting method or the widest-path graph algorithm
- Model ranks can be compared to aggregated ranks using Kendall's Tau a or T_a
- Graph visualizations can be used to supplement dissimilarity matrices and tabular data
- Ranking can also be simulated, i.e. random or perfect
- ΔE^*_{ab} of 1 ranges from nearly perfect sorting for grays to nearly random sorting for pinks
- Average pixel ΔE^*_{ab} provides a baseline prediction of MJF 3D print appearance with $2/3^{rds}$ $T_a >= 0.6$
- Thanks!


keep reinventing