A Similarity Measure for Large Color Differences

Nathan Moroney, Ingeborg Tastl and Melanie Gottwals
HP Labs

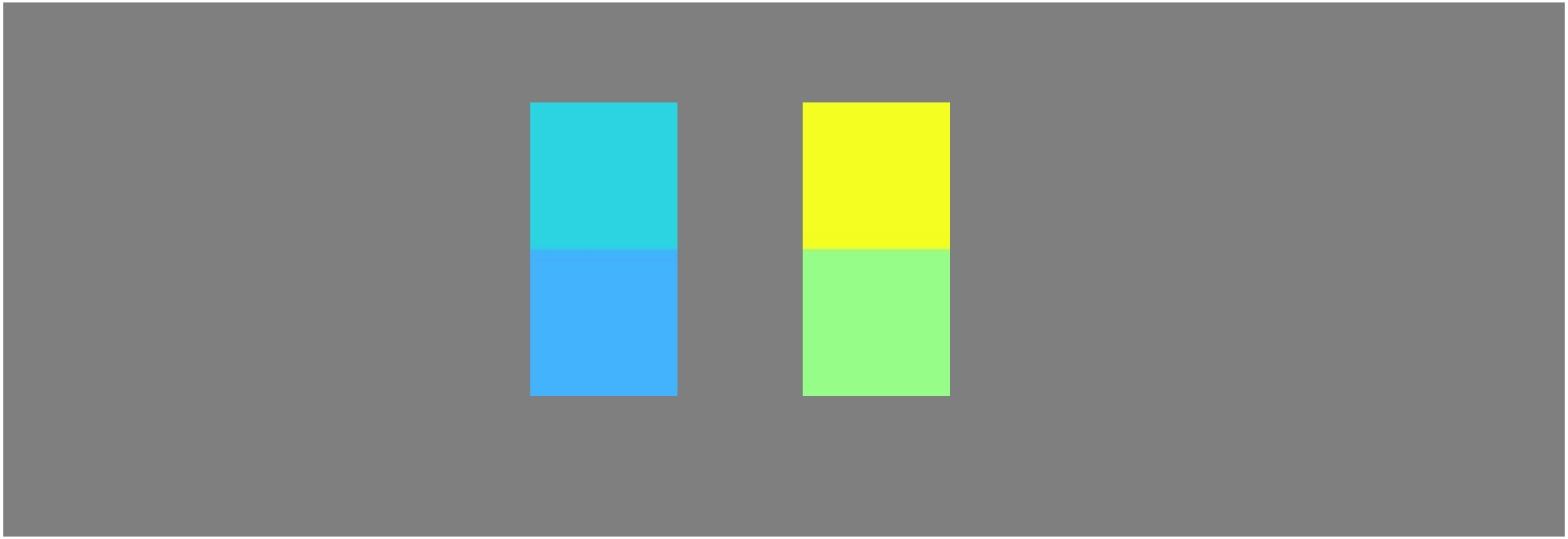
22nd IS&T Color and Imaging Conference, November 2014, Boston MA

Outline

- **Motivation**
  - Large vs. Small Differences
- **Experiments**
  - Hundreds of pairs of stimuli with ΔE_{00} of 20
  - Sorting of 9 pairs of differences from smallest to largest
  - Web and laboratory
- **Results**
  - Observers rank pairs having equal ΔE_{00} as being different
- **Similarity metric**
  - Cosine similarity given categorical vectors
  - Cosine similarity is small within a category & large across categories
- **Discussion and Speculation**

Motivation

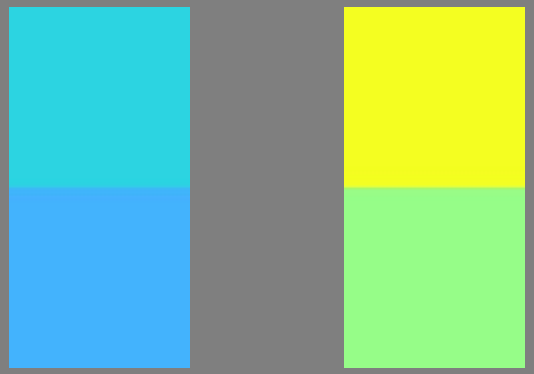
Which of these two pairs, has the larger ΔE_{00}?



Motivation

Using ΔE_{00} they are equal

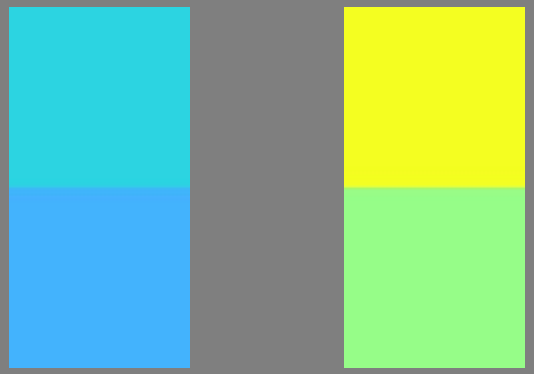
Both are 20 ΔE_{00}

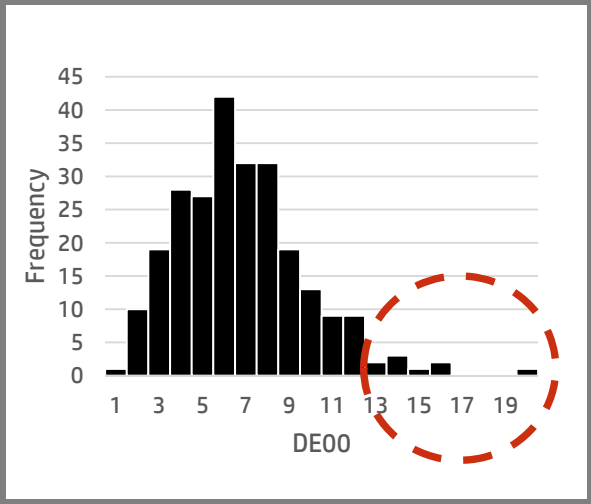


Motivation

But isn't ΔE_{00} only for small differences ? OK, so how to measure maximum error?

Both are 20 ΔE_{00}





Large vs. Small Color Differences

We'll fill this in by the end of the talk

	Small Differences	Large Differences
Application	Just Noticeable Difference	
Central Question	Do 2 colors match?	
Metrics*	ΔE^*_{ab}, ΔE_{94}, ΔE_{00}	
Underlying Metric	Euclidean distance, with weighting schemes	
Input	CIELAB coordinates, weights	
Output	"Geometric" distance, where a JND is approximately < 1	

* Small color differences metrics not recommended for large differences or greater than ~5.

Experiments

Sorting of Non-Repeating Random Vectors

- **Stimuli**
  - Non-repeating random vectors
  - Similar to a farthest-point sampling of vector endpoints
  - All vectors within 0.0001 of 20 ΔE_{00},

Not regular or centroid plus offsets sampling

- **Task**
  - Sorting of multiple color differences
  - Observer sorts 9 color differences from smallest to largest
  - 21 blocks of 9 random color difference pairs

Not forced choice paired comparison of 2 pairs

- **Web-Based Experiment**
  - 285 participants; perform sorting of 1 block of differences; unknown displays
- **Laboratory Experiment**
  - 12 participants; sort all 21 blocks of differences; sRGB mode HP DreamColor Z27x Display

Non-Repeating Random Vectors

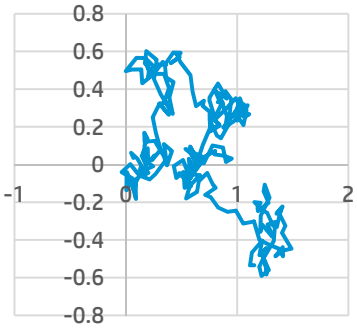
(Insert comment about Jim King here)

```
NonRepeatingRandomVectors(n, dist, t1, t2)
  i = 0
  while (i < n)
    test_ave  = F
    test_ends = F
    start = RandomVector()
    end   = start
    while ((Distance(start, end) - dist) < epsilon)
      end = RandomStep()

    ave = Average(start, end)
    for (i = 0 to number_vectors)
      if ((ave - ave[i]) < t1) test_ave = T

      d11 = start - start[i]
      d22 = end   - end[i]
      d12 = strt  - end[i]
      d21 = end   - start[i]
      min_diff   = Min(d11, d22, d12, d21)
      if (min_diff < t2) test_ends = T

    if ((test_ave & test_ends) = T)
      AddVector(start, end)
      ++I
```

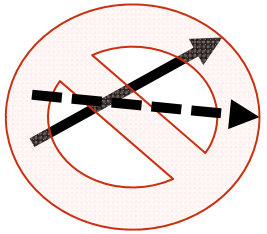
Non-Repeating Random Vectors

```
NonRepeatingRandomVectors(n, dist, t1, t2)
  i = 0
  while (i < n)
    test_ave  = F
    test_ends = F
    start = RandomVector()
    end   = start
    while ((Distance(start, end) - dist) < epsilon)
      end = RandomStep()

    ave = Average(start, end)
    for (i = 0 to number_vectors)
      if ((ave - ave[i]) < t1) test_ave = T

      d11 = start - start[i]
      d22 = end   - end[i]
      d12 = strt  - end[i]
      d21 = end   - start[i]
      min_diff   = Min(d11, d22, d12, d21)
      if (min_diff < t2) test_ends = T

    if ((test_ave & test_ends) = T)
      AddVector(start, end)
      ++i
```

Random walk

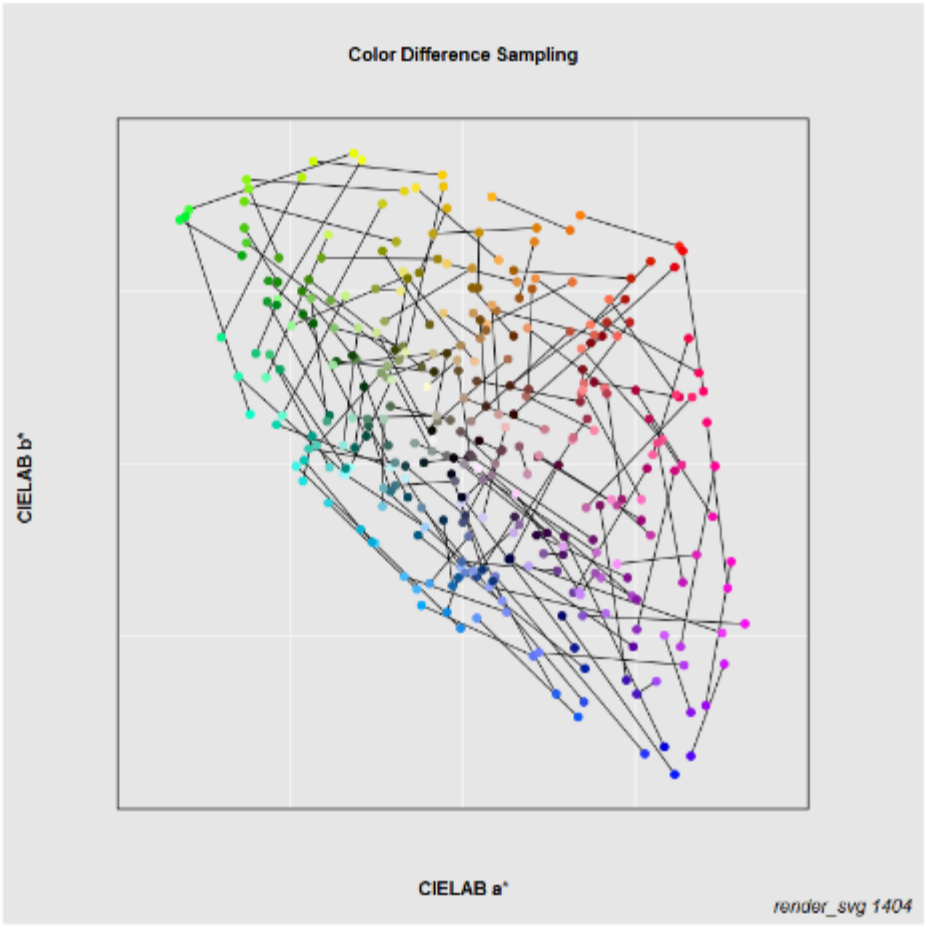


Don't duplicate vectors



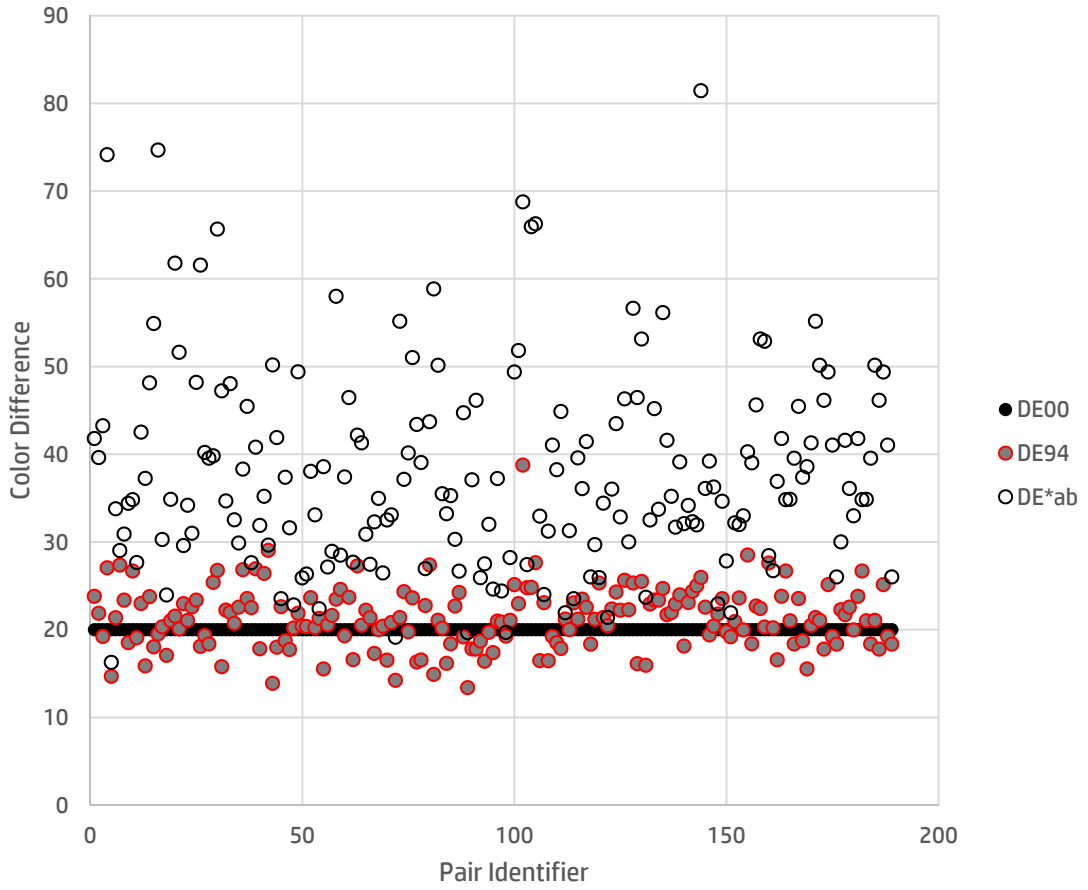
Non-Repeating Random Vectors

21 blocks rendered as RGB patches and as a CIELAB a* versus b* plot



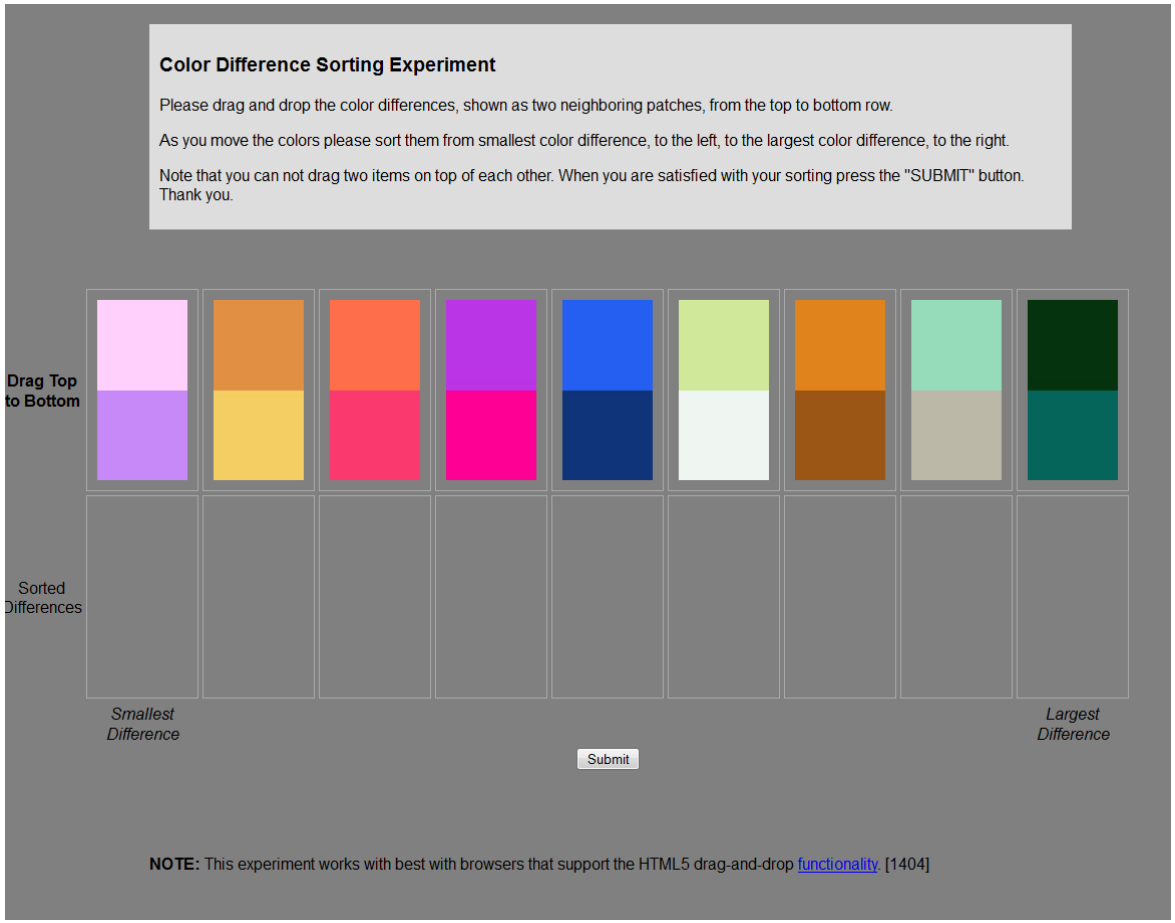
Non-Repeating Random Vectors

20 ΔE_{00} and the corresponding ΔE_{94} and ΔE^*_{ab} color differences



Color Difference Sorting

HTML5 drag-and-drop interface with color patches rendered to 90x90 pixels in size



Results: *Block 118*

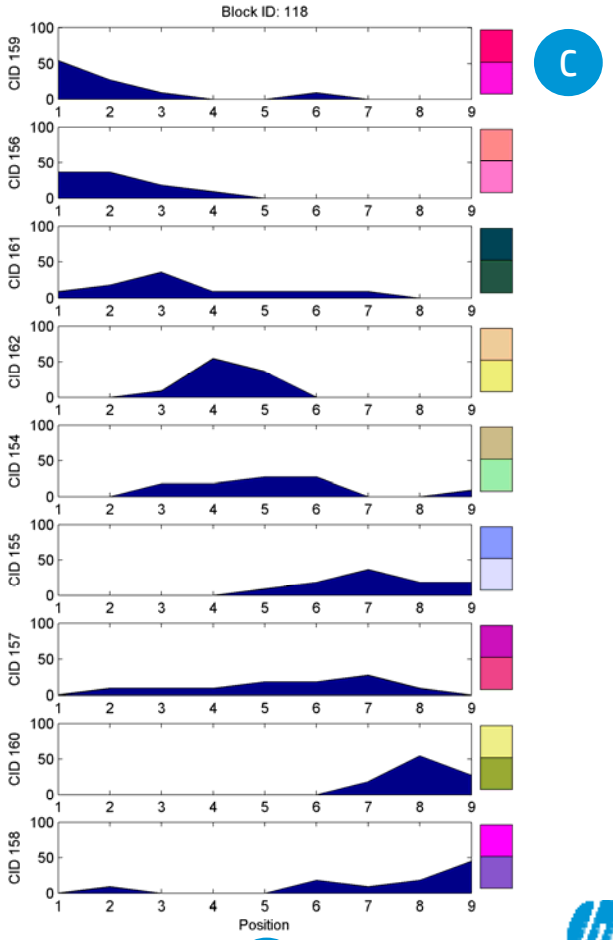
- **Initial visualizations**
  - Each set of 9 sub-plots is for 1 block

A **X-axis is sorted rank**
  - Left: smallest difference
  - Right: largest difference

B **Y-axis is relative frequency**
  - Larger: observers consistently used this rank
  - Zero: observers did not use this rank

C **Approximate color rendering**
  - Shown to the right of each sub-plot

- **Approximate sorting top to bottom**



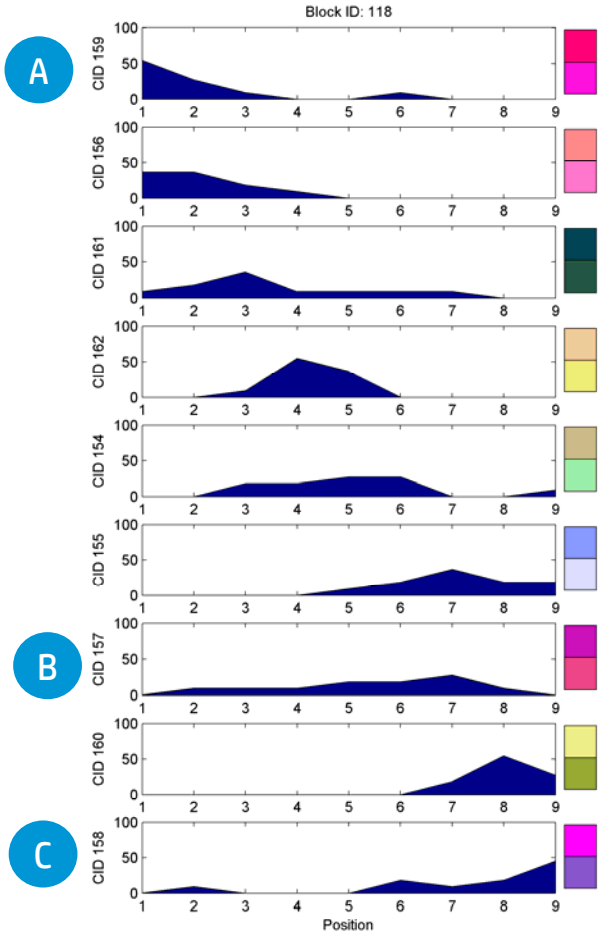
Results: *Block 118*

A **Smallest difference**
- Relatively narrower distribution

B **Intermediate differences**
- Relatively wider distributions

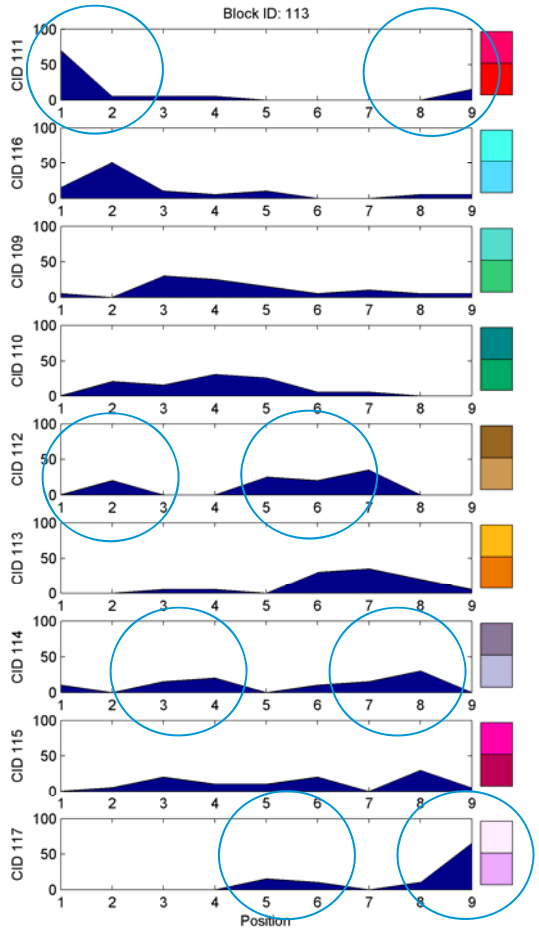
C **Largest difference**
- Relatively narrower distribution
- Qualitatively, at crossing of naming boundaries?

- **For ideal set of 9 equal differences**
  - Flat histograms for each pair
  - Not seen in the experimental data…



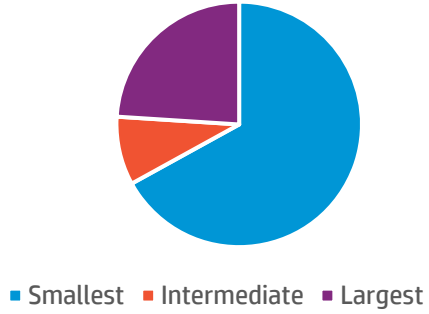
Results: *Block 113*

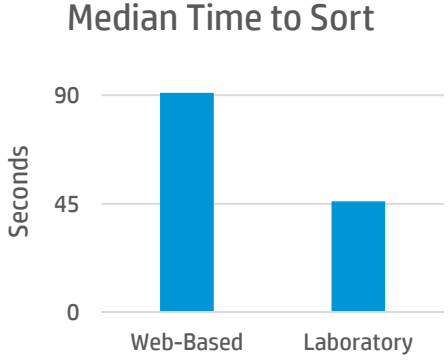
- **Similar results seen across other blocks**
- **Similar results for web-based & laboratory**
- **But also multi-modal distributions?**
  - Complicates even basic analysis
  - Can occur for any ranking, shown circled to right
  - Almost as if there were multiple criterion for sorting….

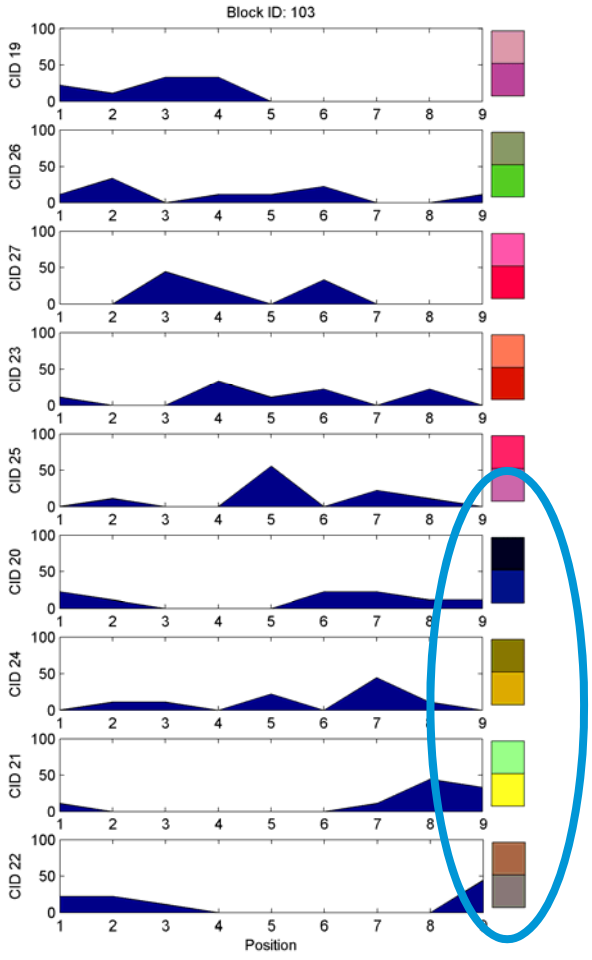


Results: *Block 103*

- **Name boundary crossing**
  - Tends to result in a mode with a larger rank or sorted color difference
  - Shown circled
- **Other interesting results**

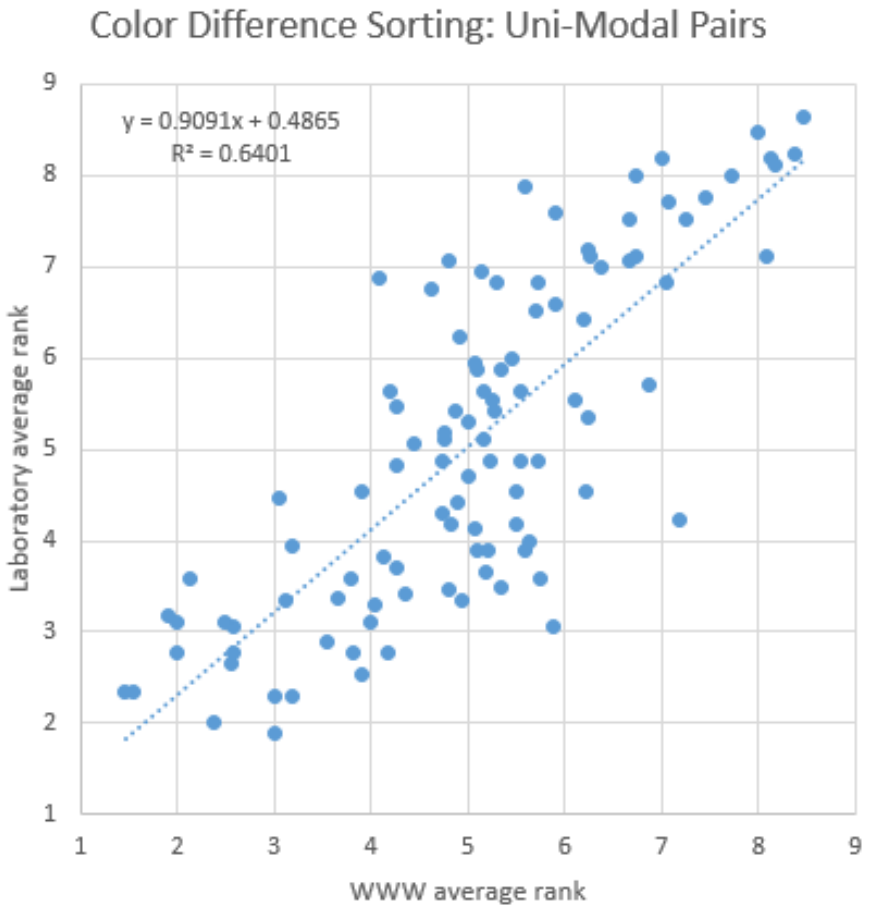






Correlation of Web-Based and Laboratory Experiments

Limited to the approximately 2/3^{rds} of pairs with uni-modal distributions* for both experiments



* F. Schwaiger, H. Holzmann, and S. Vollmer, "bimodalitytest: Testing for bimodality in a normal mixture", R package version 1.0, (2013)

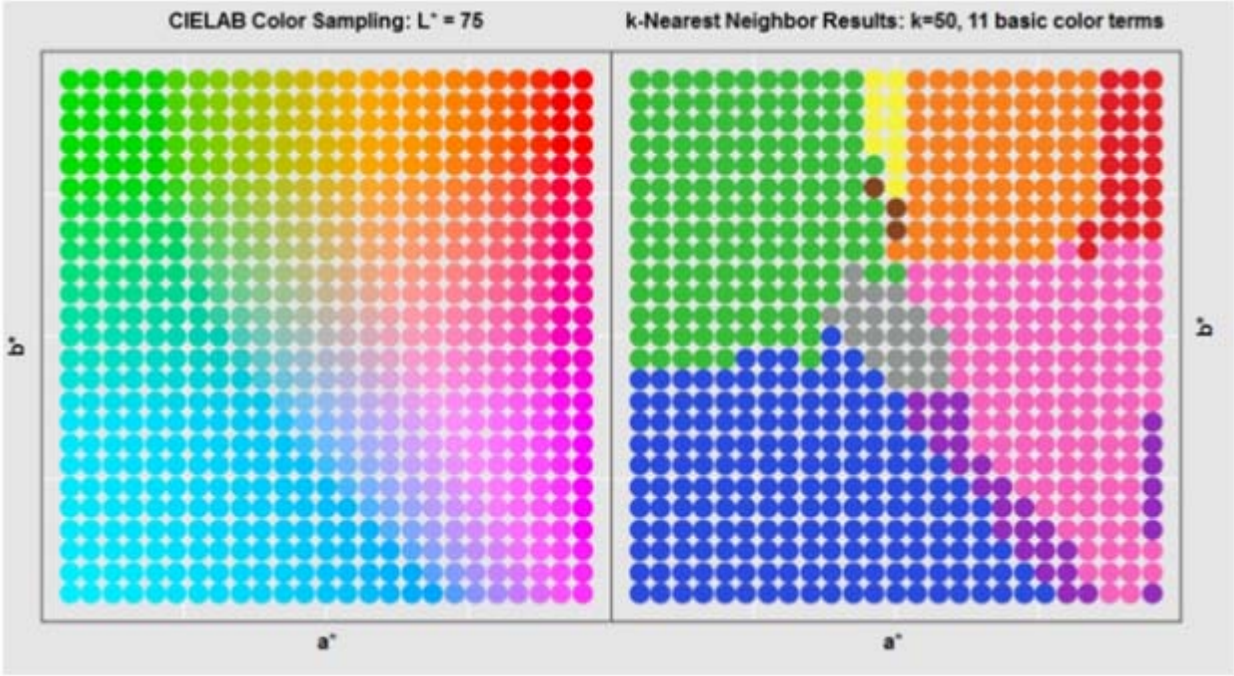
Similarity Metric

- **Given these results can a similarity metric be designed that is:**
  - Smaller within a color category
  - Larger across color categories
  - Not unlike results seen in Categorical Perception from vision science
- **Not based on weighted Euclidean distances?**
- **Similar to other similarity metrics?**
  - Value of 1 for identical
  - Value of 0 for dissimilar

Similarity Metric

Use Categorical Vectors

- **K-nearest neighbors used to transform input CIELAB data, on left, to categorical counts, maximum shown color coded on right**
- **Start with basic 11 terms as the vocabulary**
- **Similar processing used for document classification**



Similarity Metric

Use Cosine Similarity of Categorical Vectors

- **K = 100**
- **Same example from the motivation slides earlier**
- **These two have ΔE_{00} differences of 20…**

$$\Delta S = \frac{\sum A \cdot B}{\sqrt{\sum A^2 \cdot \sum B^2}}$$

	Top	Bottom	Top	Bottom
sRGB	43, 212, 224	67, 179, 253	244, 254, 32	150, 253, 137
Gray	3	1	0	0
Blue	96	98	0	0
Green	0	1	0	83
Pink	0	0	0	0
Black	1	0	0	0
Green	0	0	0	0
Orange	0	0	0	1
Purple	0	0	0	0
Red	0	0	0	0
White	0	0	0	0
Yellow	0	0	100	16
Similarity	0.99		0.18	

Large vs. Small Color Differences

We'll fill this in by the end of the talk

	Small Differences	Large Differences*
Application	Just Noticeable Difference	Consistently Describable Difference(s)
Central Question	Do 2 colors match?	When do 2 colors stop looking similar?
Metrics*	ΔE^*_{ab}, ΔE_{94}, ΔE_{00}	ΔS
Underlying Metric	Euclidean distance, with weighting schemes	Cosine similarity
Input	CIELAB coordinates, weights	Categorical or lexical vectors
Output	Geometric distance, where a JND is approximately < 1	Similarity measure, where 0 is completely dissimilar

* Large color differences also relevant to image segmentation, analysis & retrieval.

Discussion & Speculation

- **Large color differences do not necessarily have a single sorting**
  - Can be consistently sorted by multiple criterion
  - More sophisticated analysis needed to detect systematic trends in multi-modal sorting
- **Small difference metrics probably not a good idea for maximum errors**
  - At a minimum probably want to visualize
- **Cosine similarity of categorical vectors is a promising metric**
  - Initial stages of optimization but already useful in ways that differ from the ΔE's
  - Training data and algorithms are key aspects of the metric
- **Experiment and data are public and ongoing**
  - Same data could also be used to investigate more uniform color spaces
  - Would like to *expert-source* additional analysis & related experiments