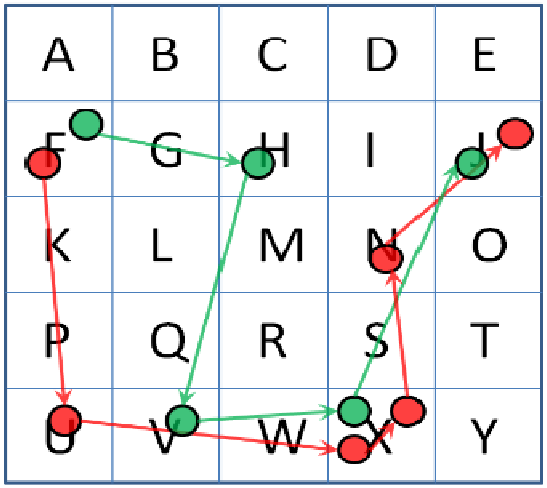
Scanpath similarity depends on how you look at it: Evaluating a 'MultiMatch' comparison algorithm

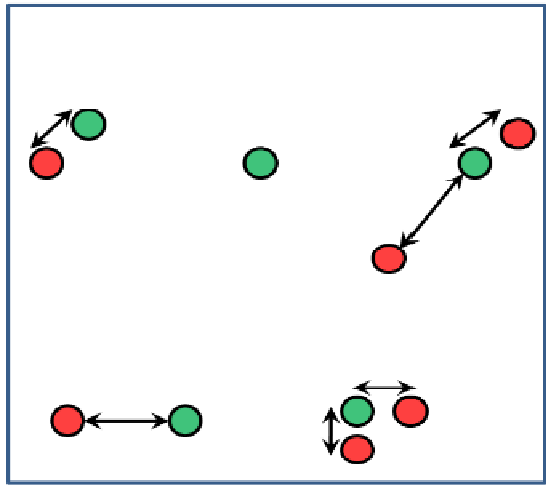
Richard Dewhurst, Marcus Nyström, Halszka Jarodzka & Kenneth Holmqvist

Scanpaths representations often deviate from eye movement data:

AOI-based



Position-based



Becomes → **FUXNJ** / **FHVXJ**

Becomes linear distance between points

AOI-based (e.g. Levenshtein, 1966; Scanmatch, 2010) lose position information

Position-based (e.g. Mannan et al., 1996; Attention maps, see Pomplun, 1996) lose order of eye movements

What we want a (pairwise) scanpath comparison algorithm to preserve:

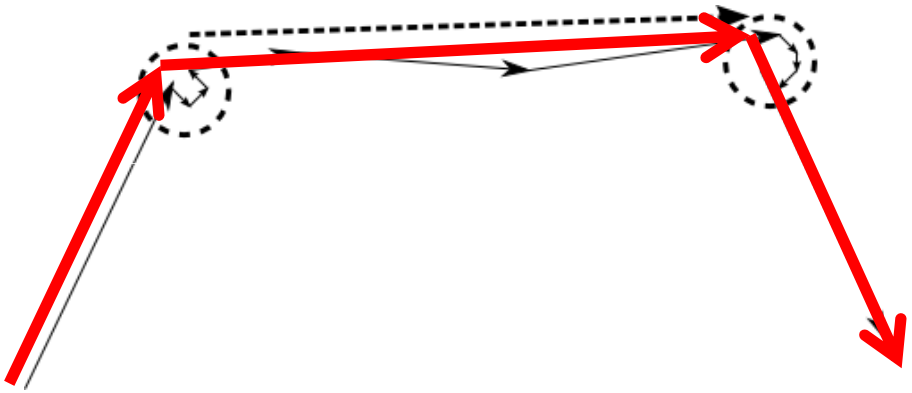
1. **Order** (cf. Levenshtein). To be a path the ordinal sequence of fixations must be reflected in the comparison
2. **Position** (cf. Mannan linear-distance). The representations must reflect locations in x,y space.
3. **Shape** (cf. Mental imagery). Two scanpath visualisations can be similar in shape, proportional geometry between saccade direction and amplitude.
4. **Fixation duration**. This is often omitted, but we know its importance for visual processing

MultiMatch—Implementation details:

Simplification

•Amplitude threshold 10% of screen width

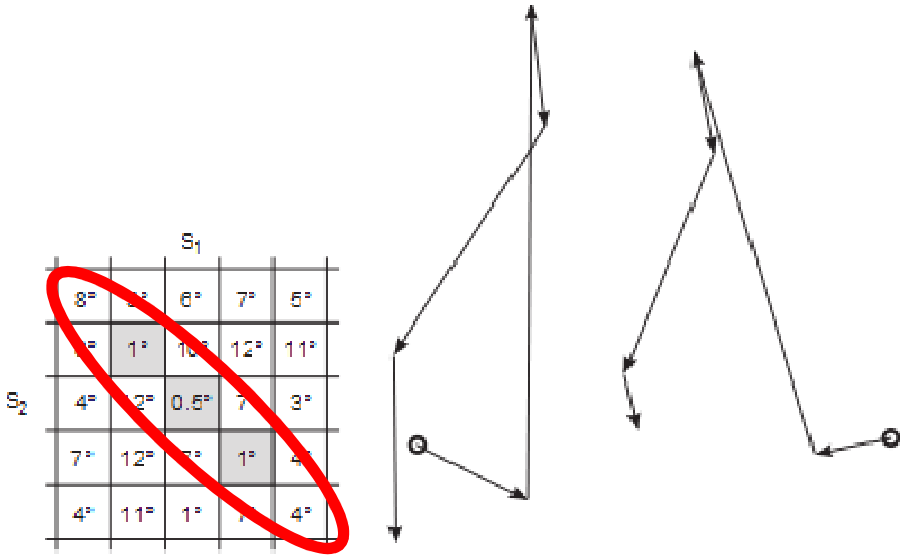
•Direction threshold 45 degrees



Temporal Alignment

•Columns and rows are vectors in order

•Aligned along the shortest path with the Dijkstra algorithm

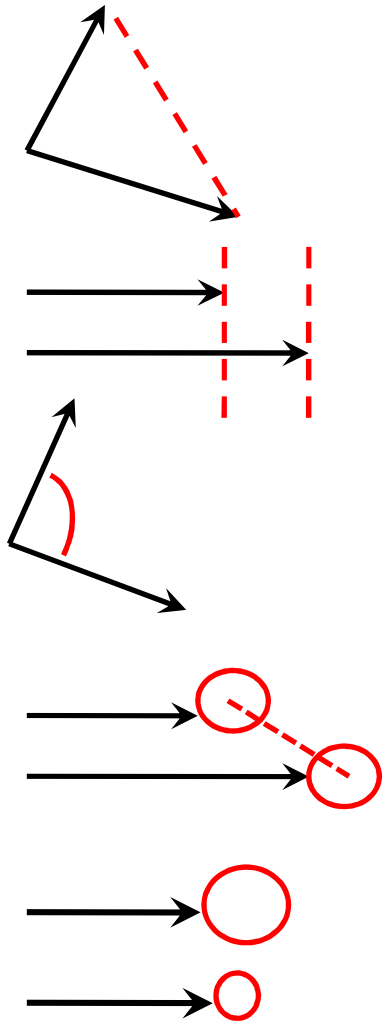


Jarodzka, H.; Nyström, M. & Holmqvist, K. A vector-based, multidimensional scanpath similarity measure. *ETRA,* 2010, 211-218

Comparison

From here, scanpath similarity is a simple matter of subtraction between dimensions for aligned scanpath pairs:

1. Vector difference (shape)
2. Length difference in (saccadic amplitude)
3. Direction difference (angular)
4. Position difference (fixations: $x_1\ y_1 - x_2\ y_2$)
5. Duration difference (fixation duration)

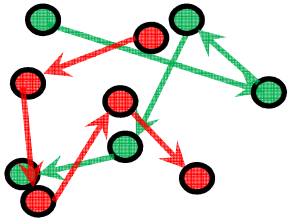


Experiment 1:

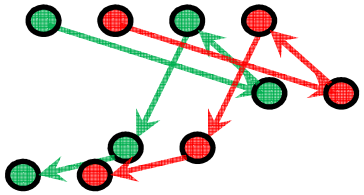
Restricted scanpaths with known similarity dimensions (sample *N* = 20 participants)

Participants viewed sequences of dots of paired scanpaths, randomly presented, while their eyes were tracked

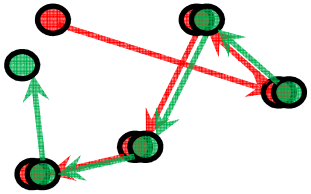
1. Random



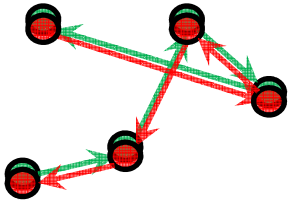
2. Spatial offset



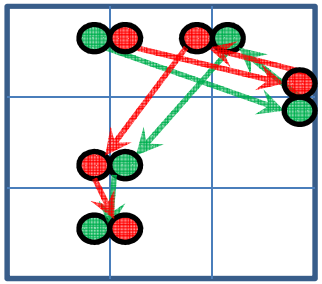
3. Ordinal offset



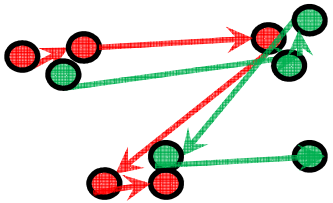
4. Reversed



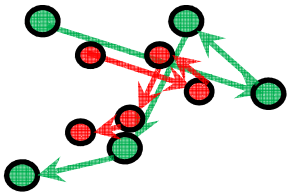
5. AOI border



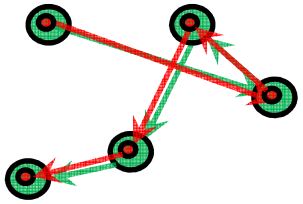
6. Local/Global



7. Scaled



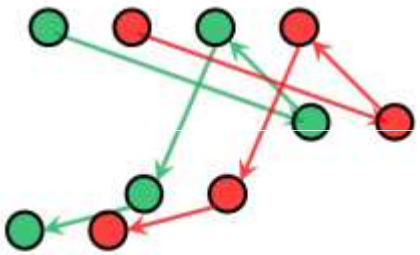
8. Duration



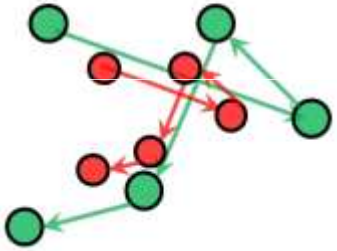
10 versions of each sequence pair, giving 160 sequences (80 pairs)

Results

2. Spatial offset



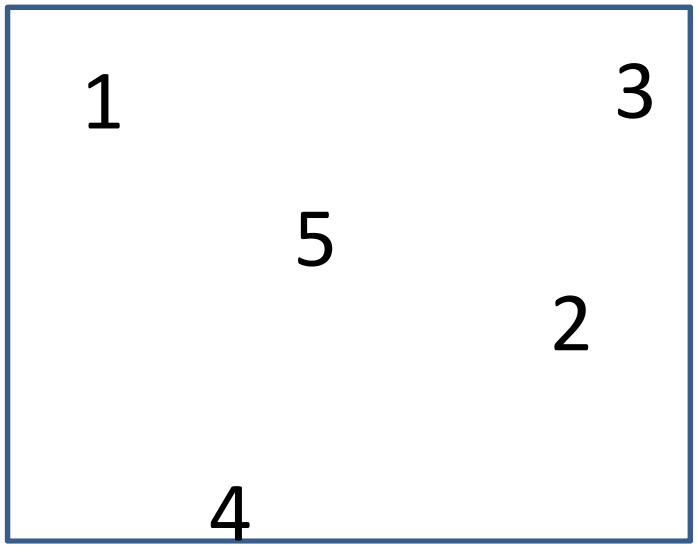
7. Scaled



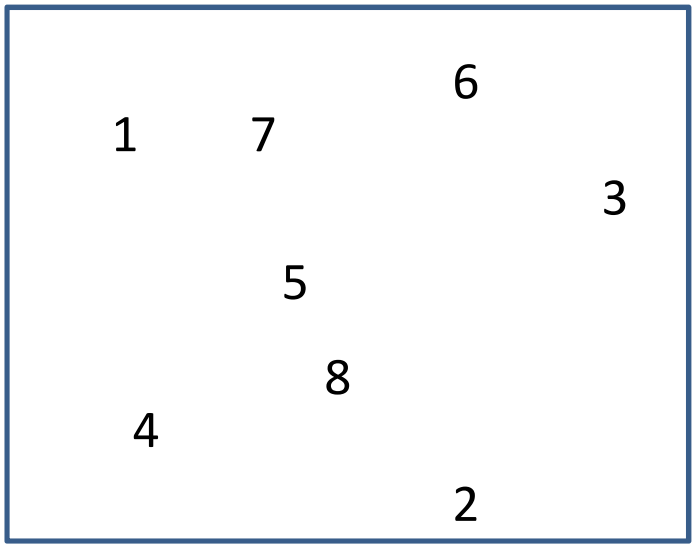
Type of difference in pair	Vector difference	Direction	Length	Position	Duration	ScanMatch
Duration	0.79	0.62	0.84	**0.70**	0.41	0.31
SpatialOffset	**0.89**	**0.81**	**0.90**	**0.78**	0.44	0.28
OrdinalOffset	**0.85**	**0.73**	**0.88**	**0.80**	0.42	**0.46**
Reversed	0.78	0.61	0.81	**0.72**	0.41	**0.36**
AOIborder	**0.85**	**0.80**	**0.85**	**0.82**	0.44	0.32
LocalGlobal	**0.85**	**0.74**	**0.87**	**0.81**	0.45	**0.42**
Scaling	**0.84**	**0.79**	0.76	**0.79**	0.44	0.27
Random	0.75	0.58	0.81	0.63	0.43	0.30

Experiment 2:

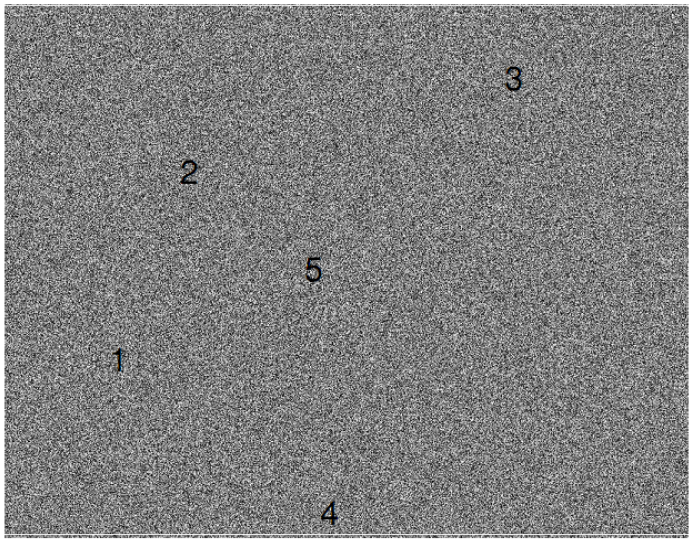
Participants viewed numbers, and their task was to look at the numbers 1-5 in order while their eyes were tracked (sample *N* = 20 participants)



Font size



Number of distractors



Noise level

10 images were generated for each manipulation type and level, yielding the 150 trials

We hypothesised that scanpaths would become less similar as the task became harder.

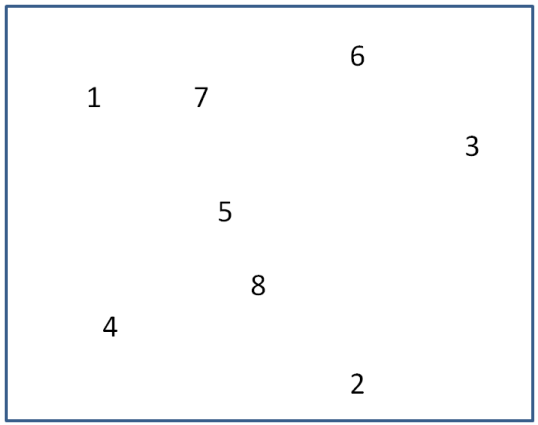
Data Analysis

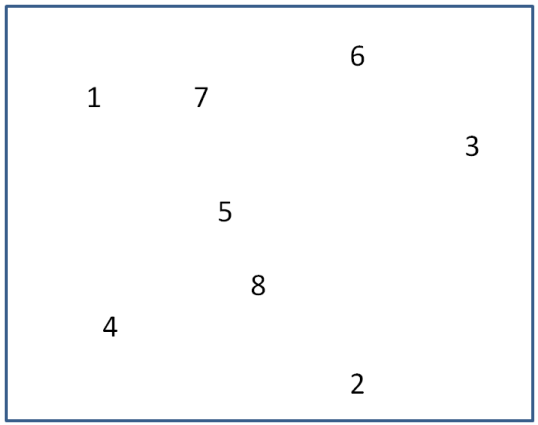
$P1_a \ldots P2_a$

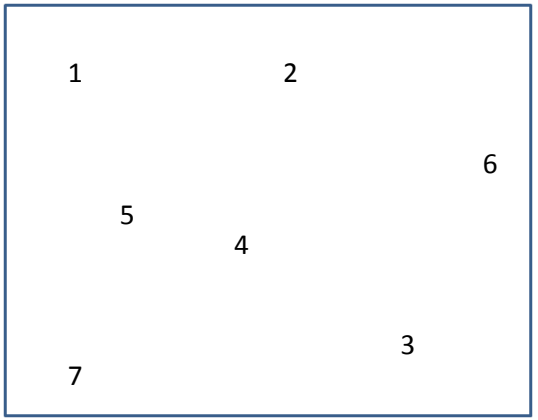
$P1_a \ldots PX_n$

Comparison classification

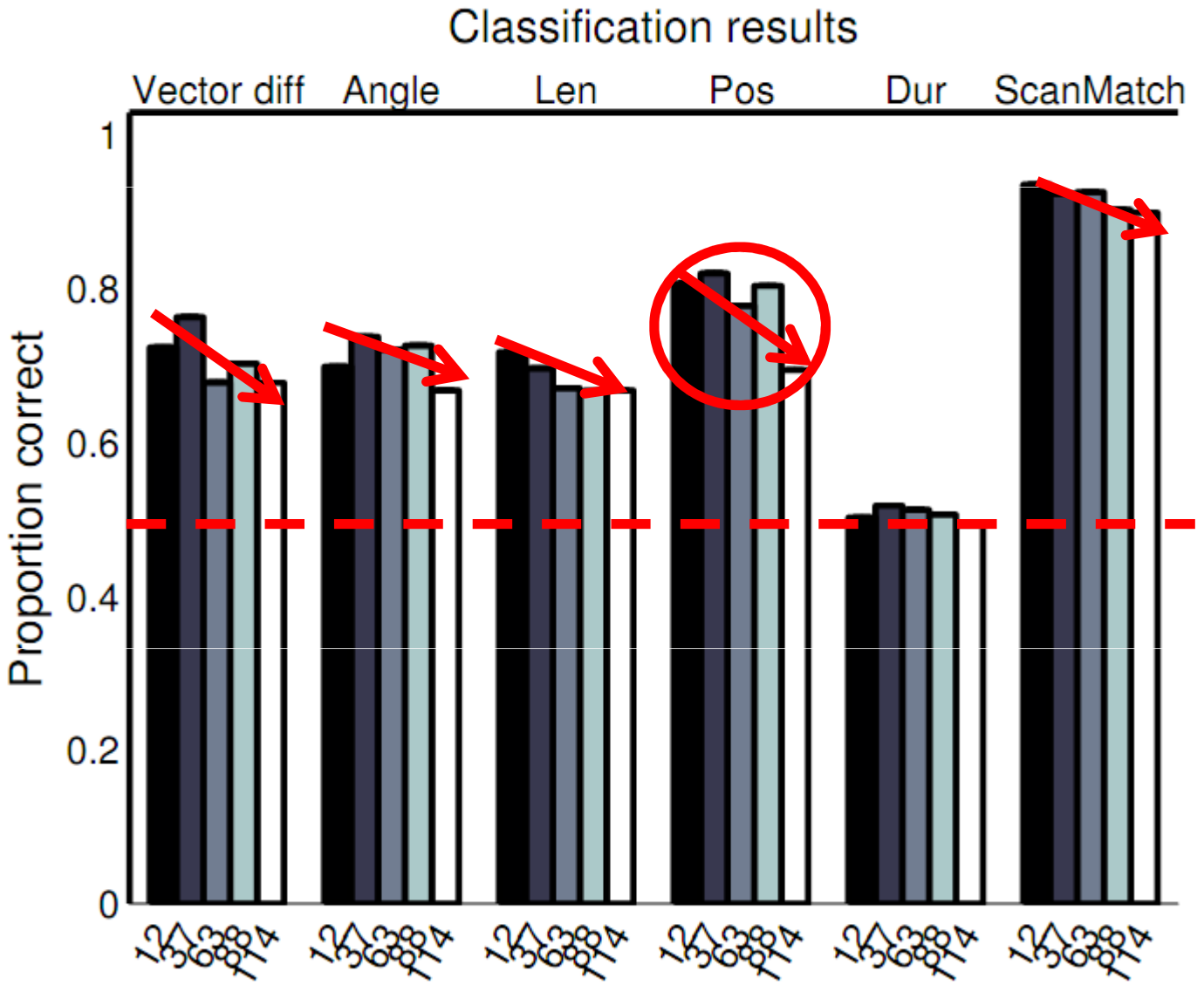
- Participant 1 looking at stimulus 1 vs. participant 2 looking stimulus 1.
- Similarity for this comparison should be higher than for Participant 1 looking at stimulus 1 vs. participant *x* looking at stimulus *n*
- Higher gives a correct classification of 1, otherwise classified as 0 for incorrect.
- This classification was done for all combinations of stimulus types within a condition.



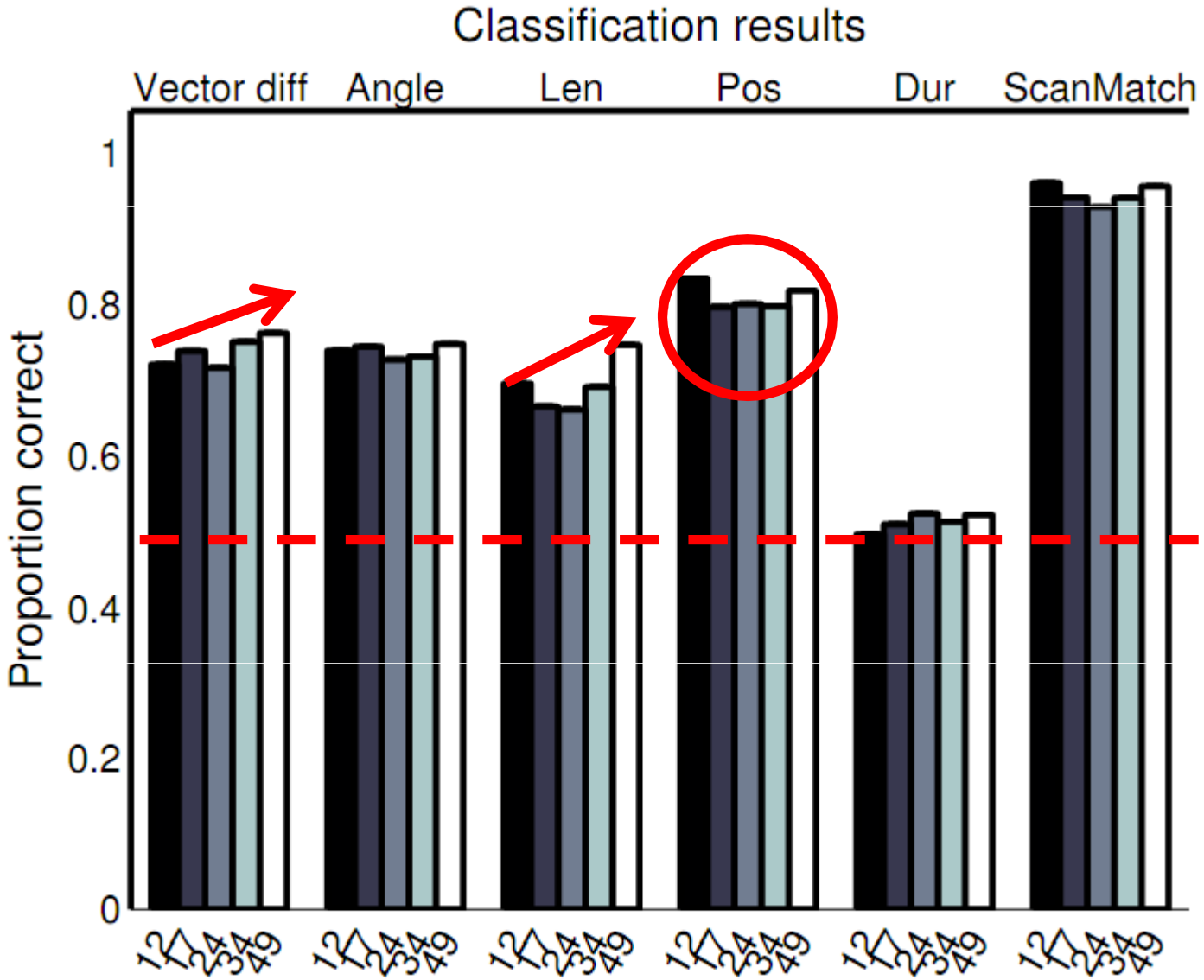




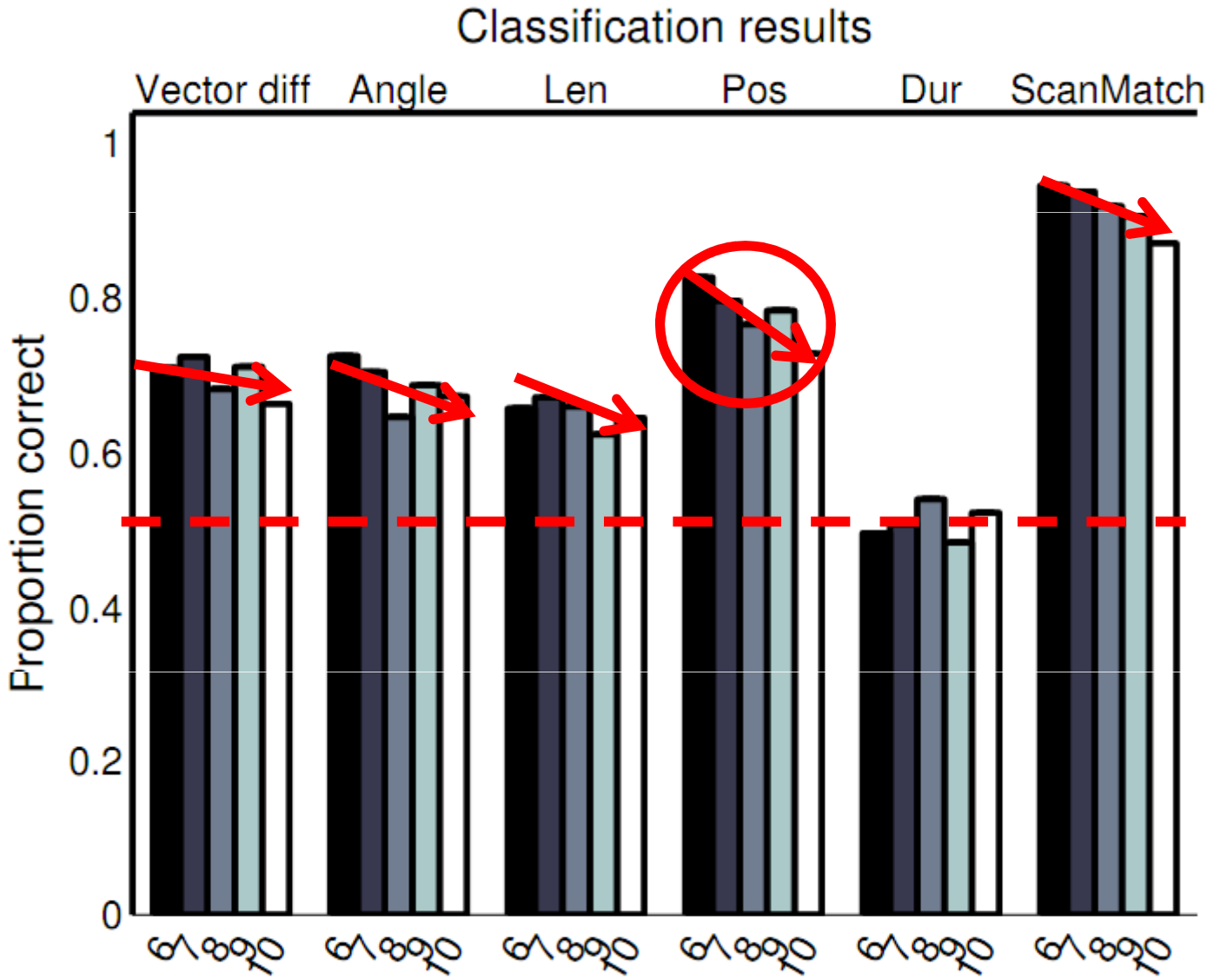
Results: Noise



Results: Font size



Results: Number of distractors



Conclusions, and further directions

"It depends on how you look at it"

- Influence of thresholds?
- Alignment similarity metric?