Guidelines for Effective Usage *of* Text Highlighting **Techniques**

Hendrik Strobelt, Daniela Oelke, Bum Chul Kwon, Tobias Schreck, Hanspeter Pfister

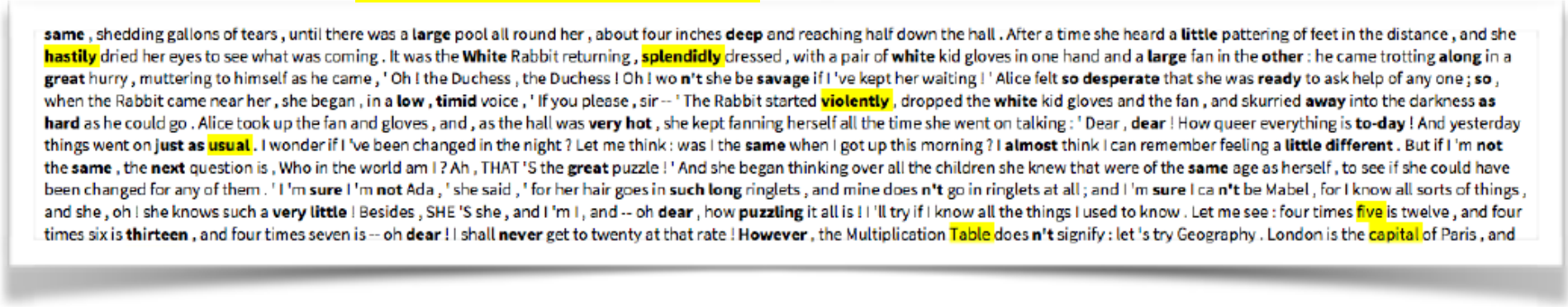
presented by Jordon Johnson

Many text vis tools…

http://textvis.lnu.se/

… but sometimes need to read text with annotations (WHY)

bold font and yellow background



e x t r a s p a c i n g and *italics*

, and seemed to her to wink with one of its l i t t l e eyes , but it said nothing . ' P e r h a p s it does n ' t understand English , ' thought Alice ; ' I daresay it 's a F r e n c h mouse , come over with William the Conqueror . ' (For , with all her knowledge of history , Alice had no v e r y c l e a r notion how l o n g ago anything had happened .) S o she began a g a i n : ' Ou est ma chatte ? ' which was the f i r s t sentence in her F r e n c h lesson-book . The Mouse gave a s u d d e n leap out of the water , and seemed to quiver all over with fright . ' Oh , I beg your pardon ! ' cried Alice *h a s t i l y* , a f r a i d that she had hurt the p o o r *animal* 's feelings . ' I q u i t e forgot you did n ' t like cats . ' ' N o t like cats ! ' cried the Mouse , in a s h r i l l , p a s s i o n a t e voice . ' Would YOU like cats if you were me ? ' ' W e l l , p e r h a p s n o t , ' said Alice in a s o o t h i n g tone : ' don ' t be a n g r y about it . And y e t I wish I could show you our cat Dinah : I think you 'd take a fancy to cats if you could *o n l y* see her . She is s u c h a d e a r q u i e t thing , ' Alice went on , half to herself , as she swam *l a z i l y* a b o u t in the pool , ' and she sits purring s o *n i c e l y* by the fire , licking her paws and washing her face -- and she is s u c h a n i c e s o f t thing to nurse -- and she 's s u c h a *capital* one for catching mice -- oh , I beg your pardon ! ' cried Alice a g a i n , for this time the Mouse was bristling a l l o v e r , and she felt c e r t a i n it must be *r e a l l y* offended . ' We won ' t talk about her any m o r e if you 'd r a t h e r n o t . ' ' We i n d e e d ! ' cried the Mouse , who was trembling d o w n to the end of his tail . ' As if I would talk on such a subject ! Our *family* a l w a y s HATED cats : n a s t y , l o w , v u l g a r

Design study...-ish

- Elicits requirements from domain experts
  - separate interviews with 5 NLP experts
- Carries out user studies to evaluate techniques
- All evaluated techniques have been in use for decades
  - similar to a study of the relative effectiveness of different marks and channels

Requirements (WHAT)

Annotations can be:

- **statistical**
  - word length
- **syntactic**
  - parts-of-speech
- **semantic**
  - sentiment tags
- **structural**
  - page margins
- **domain-specific**
  - proper names

- **categorical**
- **ordered**
- **quantitative**
- **boolean**

- **of any textual scope**
- **overlapping**

Pop-out is key

Characters/words are marks that are fairly densely packed and regularly spaced, and that already make use of some visual channels

To make highlighting detectable, need to **maximize pop-out**

Common highlighting techniques (HOW)

Technique	Use	Typical variations	Used in our studies
Font color	c q	Saturation, luminance, hue	Red color (`rgb(227,26,28)`)
Background color	c q	Saturation, luminance, hue	Bright yellow (`rgb(255,255,50)`)
Underlined	c q	Styles, thicknesses	Single underline
Font size	- q	% increase	150% increase
Font style	- -	Italics, subscript,...	Italics
Font weight	- -	Font weight	bold font
Rectangular border	c q	Styles of border, lines, thickness	Single border
S p a c e d o u t f o n t	- q	Letter spacing	5px spacing
Text shadow	- -	Offset, intensity,...	CSS: `text-shadow: 4px 4px 3px rgb(50, 50, 50);`
Font family	(c) -	Sans-serif, Times, Helvetica,..	—
CAPITALIZATION	- -	Small caps, large caps	—
~~Strike-through~~	- -	True, false	—
* Blinking *	- -	True, false	—

- Each technique can also encode boolean features (scope of paper limited to this consideration)
- 9 techniques used in user studies

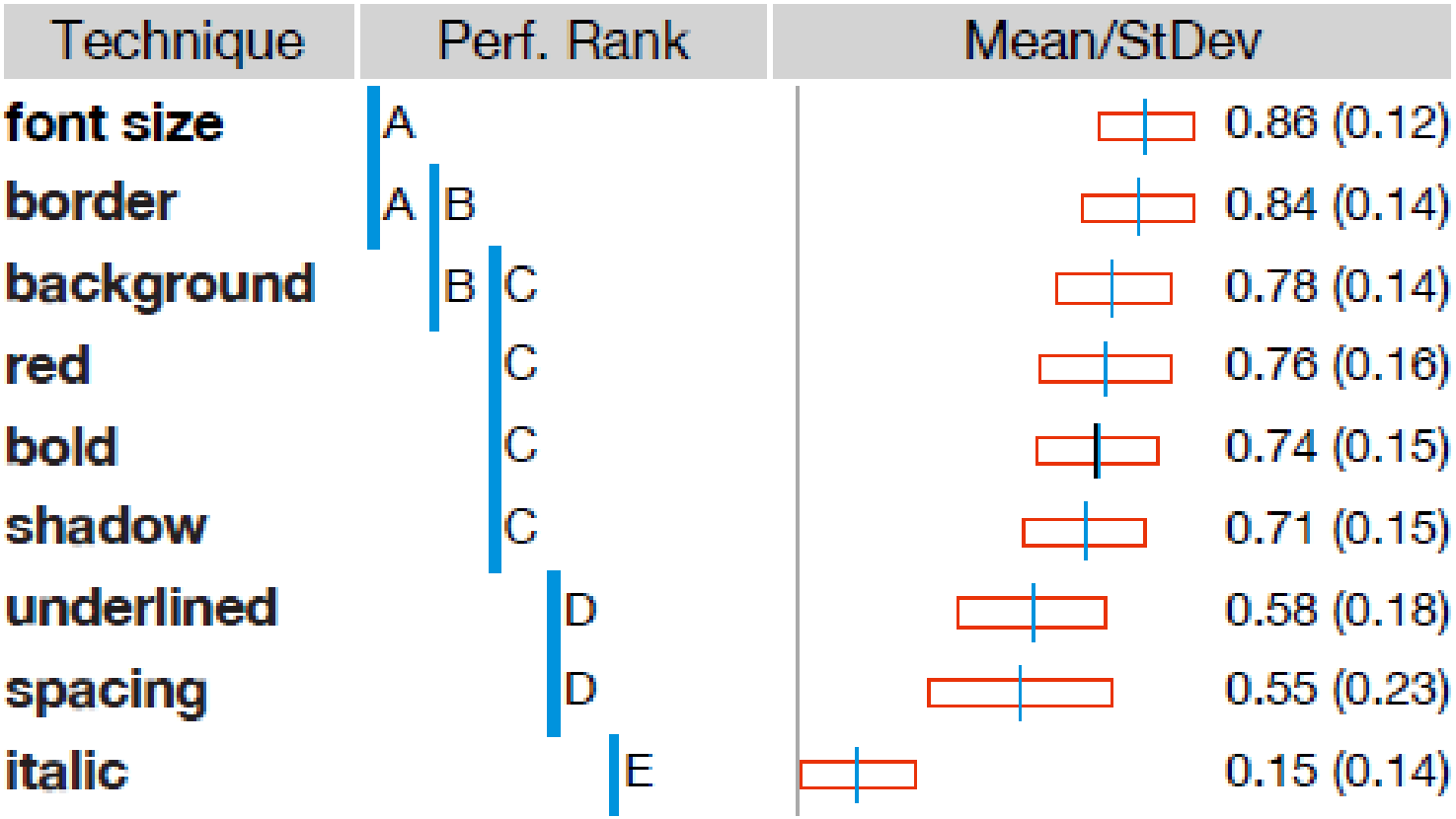
3 User Studies

- Performed using Amazon Mechanical Turk
- **Analysis techniques: ANOVA** and **Tukey HSD**
- **Unwanted variation**
  - **Individual difference:** normalized each participant's responses with respect to their performance range
  - **Learning curve:** discarded first trials in first study, added training trials in others
  - **Fatigue effects:** not observed

Study 1: Ranking Techniques

- **Goal: rank** techniques with respect to pop-out
- 673 words, 20 randomly highlighted
  - Find as many highlighted words as possible within a time limit
- 45 participants
- 3 trials per technique (27 trials total) per participant
  - trials ordered randomly

Study 1 - results



Study 1 - discussion

Possible explanations of strong results:

- Increased font size: sticks out from cap line, fill white space
- Border: makes the target appear larger
- Colour: strong pop-out effect
  - background may outperform text colour because coloured area is larger

Study 1 - discussion

Possible explanations of weak results:

- Letter spacing: already a n o r m a l feature of text
- Italics: slanted character *features* already found in text

Study 2: Search with Distractor

- **Goal:** determine how different techniques (A,B) **interfere** when used in the same text
  - Is relative strength of techniques a factor?
- 20 highlighted words for each of A, B, A+B
  - must choose words highlighted only with A
- 30 participants
- All pairs of techniques tried (72 trials total) per participant

Study 2 - results

weaker techniques

distractor technique -->

	fs	bo	bg	red	bold	sha	und	spa	it
font size		-15.4	-10.1	-4.6	-74.8	-12.5	-33.5	-92.9	-62.1
border	-27.1		-6.3	-5.8	-8.8	-10.6	-66.4	-42.8	-59.9
background	-13.5	-16.0		-17.5	-6.8	-14.5	-26.1	-40.0	-50.0
red	-17.2	-9.7	2.7		-16.5	-19.9	-30.5	-39.4	-48.8
bold	-68.6	-15.5	0.3	3.3		-15.1	-21.1	-29.9	-43.2
shadow	-20.1	-10.4	-1.7	-1.3	-13.4		-65.4	-23.8	-73.3
underlined	-22.8	-25.5	3.0	7.3	-6.9	-10.6		-37.3	-40.4
spacing	-56.0	-45.3	-6.4	-4.5	-30.3	-21.8	-44.6		-97.3
italic	23.2	35.6	48.3	37.6	31.9	28.8	15.6	2.5	

did not expect improvements

Study 2 - results

	distractor technique -->								
	fs	**bo**	**bg**	**red**	**bold**	**sha**	**und**	**spa**	**it**
font **s**ize		0.75	0.78	0.82	0.49	0.76	0.64	0.45	0.53
border	0.66		0.79	0.79	0.77	0.76	0.50	0.59	0.53
bground	0.69	0.67		0.66	0.73	0.68	0.62	0.56	0.52
red	0.65	0.69	0.78		0.65	0.63	0.58	0.55	0.51
bold	0.44	0.64	0.74	0.77		0.64	0.61	0.57	0.52
shadow	0.59	0.64	0.70	0.70	0.63		0.43	0.57	0.41
underlined	0.47	0.46	0.60	0.63	0.54	0.52		0.42	0.41
spacing	0.35	0.38	0.52	0.53	0.42	0.45	0.38		0.28
italic	0.20	0.23	0.29	0.24	0.22	0.21	0.18	0.15	

Fig. 8: Absolute performance values of Study 2 (referenced as Matrix M2).

Study 2 - results

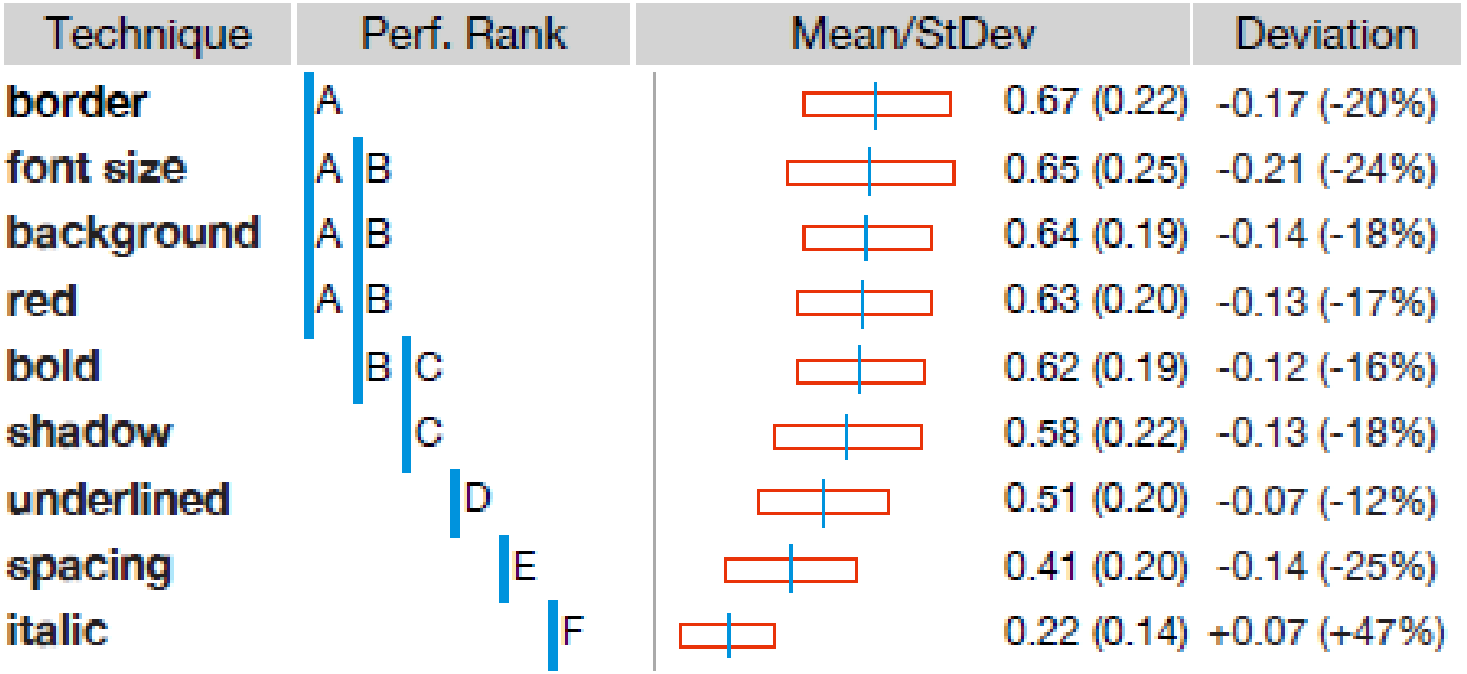


Fig. 5: Performance rank of target highlighting with a distractor (Study 2). The column *Deviation* reports the Deviation of the Mean Score from Study 1 (Percentage Change of Mean Score from Study 1). See caption of Figure 3 for how to read the *Perf. Rank* column.

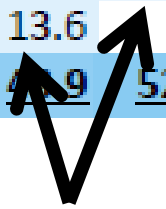
Study 3: Visual Conjunctive Search

- **Goal:** How strong is a combination of techniques (A,B) compared to each alone?
- 20 highlighted words for each of A, B, A+B
  - must choose only A+B
- 24 participants
- All pairs of techniques tried (36 trials total) per participant

Study 3 - results

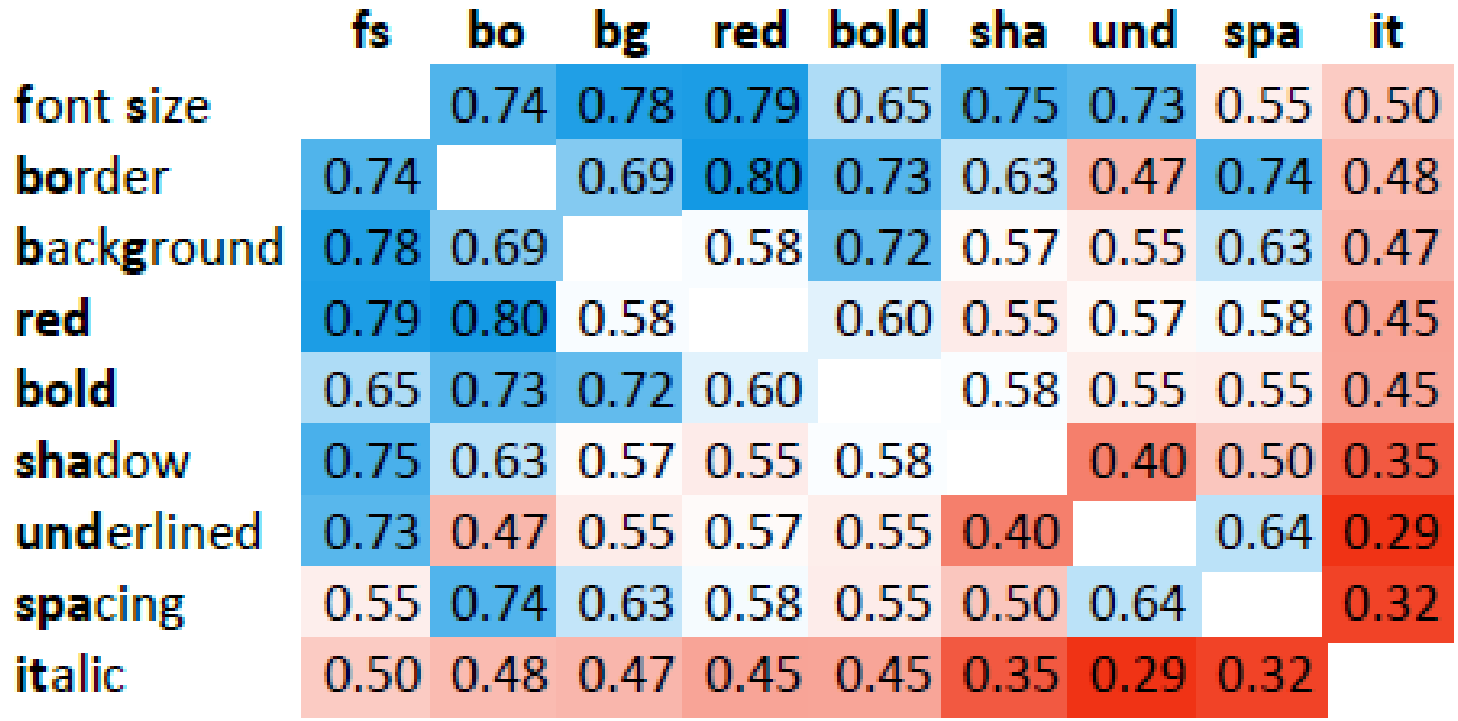
results similar to study 2

	fs	bo	bg	red	bold	sha	und	spa	it
font size		-16.4	-9.7	-9.2	-32.4	-15.4	-17.7	-56.3	-71.5
border	-13.7		-21.9	-5.6	-14.3	-32.5	-77.1	-13.9	-74.9
background	0.5	-13.2		-34.7	-8.1	-37.5	-42.7	-23.8	-64.7
red	3.5	4.5	-31.2		-26.5	-39.2	-34.3	-30.6	-69.4
bold	-13.9	-0.7	-2.6	-23.2		-28.2	-34.7	-35.2	-64.5
shadow	4.8	-12.0	-25.2	-30.0	-23.0		-78.3	-43.3	-105.0
underlined	20.6	-22.3	-6.1	-2.5	-5.6	-45.7		8.9	-97.5
spacing	0.0	25.4	12.7	5.5	-0.4	-11.0	13.6		-73.9
italic	70.1	68.8	68.3	66.6	66.7	56.7	[illegible]9	52.6	



Only underlined + spacing showed improvement over both individually

Study 3 - results



	fs	bo	bg	red	bold	sha	und	spa	it
font size		0.74	0.78	0.79	0.65	0.75	0.73	0.55	0.50
border	0.74		0.69	0.80	0.73	0.63	0.47	0.74	0.48
background	0.78	0.69		0.58	0.72	0.57	0.55	0.63	0.47
red	0.79	0.80	0.58		0.60	0.55	0.57	0.58	0.45
bold	0.65	0.73	0.72	0.60		0.58	0.55	0.55	0.45
shadow	0.75	0.63	0.57	0.55	0.58		0.40	0.50	0.35
underlined	0.73	0.47	0.55	0.57	0.55	0.40		0.64	0.29
spacing	0.55	0.74	0.63	0.58	0.55	0.50	0.64		0.32
italic	0.50	0.48	0.47	0.45	0.45	0.35	0.29	0.32	

Fig. 9: Absolute performance values of Study 3 (referenced as Matrix M3).

Guidelines

Scenarios:

- Only one feature should be highlighted
- Both features should have the same visibility; conjunctive visual search is not important
- Conjunction of features is more important than each individually
- One feature is significantly more important than the other
- Both features should have the same visibility; their conjunction should be easy to see

Only one feature

Choose a technique with strong pop-out

Examples:

- Font size
- Borders
- Yellow background

Same visibility; conjunction unimportant

Choose techniques with strong pop-out that do not significantly interfere with each other

Examples:

- Bold + yellow background
- Border + red
- Font size + yellow background
- Font size + border

Conjunction of features more important than each individually

Choose techniques that scored high in visual conjunction test

Examples:

- Border + red
- Font size + red
- Font size + yellow background

One feature significantly more important than the other

Choose techniques such that one has significantly higher pop-out

Examples:

- Yellow background + spacing
- Font size + underlined
- Border + italics

Same visibility, easy-to-see conjunction

Choose techniques with strong pop-out that do not significantly interfere with each other, whose conjunction is easy to see

Examples:

- Border + red
- Font size + yellow background
- Yellow background + bold

Discussion/Future Work

Increase **scope**

- Combinations of more than two techniques
- Include more techniques (eg. different colour combinations
- Include categorical/ordered/quantitative data
- Include tasks that require context/analysis
- Consider overlay visualizations

Comments/Critiques

- The guidelines for some scenarios are **very** similar, and multiple examples cover multiple scenarios
  - 3 studies for 5 scenarios
  - Some scenario refactoring would not be amiss
- I would have liked to see a larger scope
  - The authors don't misrepresent the scope
  - A larger scope would be a lot more work
  - **BUT** a larger set of matrices might reveal more clusters to fit the scenarios better

Comments/Critiques

- I would have liked to see a statement of expected results, based on existing understanding of marks and channels

Are there any

q u e s t i o n s ?