

# 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...



# ... but sometimes need to read text with annotations (WHY)

## bold font and yellow background

**same**, shedding gallons of tears, until there was a **large** pool all round her, about four inches **deep** and reaching half down the hall. After a time she heard a **little** pattering of feet in the distance, and she **hastily** dried her eyes to see what was coming. It was the **White** Rabbit returning, **splendidly** dressed, with a pair of **white** kid gloves in one hand and a **large** fan in the **other**: he came trotting **along** in a **great** hurry, muttering to himself as he came, 'Oh! the Duchess, the Duchess! Oh! wo**n't** she be **savage** if I've kept her waiting!' Alice felt **so desperate** that she was **ready** to ask help of any one; **so**, when the Rabbit came near her, she began, in a **low**, **timid** voice, 'If you please, sir--' The Rabbit started **violently**, dropped the **white** kid gloves and the fan, and skurried **away** into the darkness **as hard** as he could go. Alice took up the fan and gloves, and, as the hall was **very hot**, she kept fanning herself all the time she went on talking: 'Dear, **dear**! How queer everything is **to-day**! And yesterday things went on **just as usual**. I wonder if I've been changed in the night? Let me think: was I the **same** when I got up this morning? I **almost** think I can remember feeling a **little different**. But if I'm **not the same**, the **next** question is, Who in the world am I? Ah, THAT'S the **great** puzzle!' And she began thinking over all the children she knew that were of the **same** age as herself, to see if she could have been changed for any of them. 'I'm **sure** I'm **not** Ada,' she said, 'for her hair goes in **such long** ringlets, and mine does **n't** go in ringlets at all; and I'm **sure** I ca**n't** be Mabel, for I know all sorts of things, and she, oh! she knows such a **very little**! Besides, SHE'S she, and I'm I, and -- oh **dear**, how **puzzling** it all is! I'll try if I know all the things I used to know. Let me see: four times **five** is twelve, and four times six is **thirteen**, and four times seven is -- oh **dear**! I shall **never** get to twenty at that rate! **However**, the Multiplication **Table** does **n't** signify: let's try Geography. London is the **capital** of Paris, and

## extra spacing 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 n o 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 : 'O u e s t m a c h a t t e ?' 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 a n i m a l's feelings. 'I q u i t t e forgot you d i d n't like cats.' 'N o t like cats!' cried the Mouse, in a h i r 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: 'd o n'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 i t thing to nurse -- and she is s u c h a c a p i t a l 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 wo n't talk about her any m o r e if you'd r a t h e r n o t .' 'W e 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! O u r f a m i l y a l w a y s H A T E D 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
Spaced out font	- 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

Technique	Perf. Rank	Mean/StDev
<b>font size</b>	A	 0.86 (0.12)
<b>border</b>	A B	 0.84 (0.14)
<b>background</b>	B C	 0.78 (0.14)
<b>red</b>	C	 0.76 (0.16)
<b>bold</b>	C	 0.74 (0.15)
<b>shadow</b>	C	 0.71 (0.15)
<b>underlined</b>	D	 0.58 (0.18)
<b>spacing</b>	D	 0.55 (0.23)
<b>italic</b>	E	 0.15 (0.14)

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



did not expect improvements

# Study 2 - results

distractor technique -->

	<b>fs</b>	<b>bo</b>	<b>bg</b>	<b>red</b>	<b>bold</b>	<b>sha</b>	<b>und</b>	<b>spa</b>	<b>it</b>
<b>font size</b>		0.75	0.78	0.82	0.49	0.76	0.64	0.45	0.53
<b>border</b>	0.66		0.79	0.79	0.77	0.76	0.50	0.59	0.53
<b>background</b>	0.69	0.67		0.66	0.73	0.68	0.62	0.56	0.52
<b>red</b>	0.65	0.69	0.78		0.65	0.63	0.58	0.55	0.51
<b>bold</b>	0.44	0.64	0.74	0.77		0.64	0.61	0.57	0.52
<b>shadow</b>	0.59	0.64	0.70	0.70	0.63		0.43	0.57	0.41
<b>underlined</b>	0.47	0.46	0.60	0.63	0.54	0.52		0.42	0.41
<b>spacing</b>	0.35	0.38	0.52	0.53	0.42	0.45	0.38		0.28
<b>italic</b>	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










Technique	Perf. Rank	Mean/StDev	Deviation
<b>border</b>	A	 0.67 (0.22) -0.17 (-20%)	
<b>font size</b>	A B	 0.65 (0.25) -0.21 (-24%)	
<b>background</b>	A B	 0.64 (0.19) -0.14 (-18%)	
<b>red</b>	A B	 0.63 (0.20) -0.13 (-17%)	
<b>bold</b>	B C	 0.62 (0.19) -0.12 (-16%)	
<b>shadow</b>	C	 0.58 (0.22) -0.13 (-18%)	
<b>underlined</b>	D	 0.51 (0.20) -0.07 (-12%)	
<b>spacing</b>	E	 0.41 (0.20) -0.14 (-25%)	
<b>italic</b>	F	 0.22 (0.14) +0.07 (+47%)	

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

	<b>fs</b>	<b>bo</b>	<b>bg</b>	<b>red</b>	<b>bold</b>	<b>sha</b>	<b>und</b>	<b>spa</b>	<b>it</b>
font size		<u>-16.4</u>	-9.7	-9.2	<u>-32.4</u>	<u>-15.4</u>	<u>-17.7</u>	<u>-56.3</u>	<u>-71.5</u>
border	<u>-13.7</u>		<u>-21.9</u>	-5.6	<u>-14.3</u>	<u>-32.5</u>	<u>-77.1</u>	<u>-13.9</u>	<u>-74.9</u>
background	0.5	<u>-13.2</u>		<u>-34.7</u>	-8.1	<u>-37.5</u>	<u>-42.7</u>	<u>-23.8</u>	<u>-64.7</u>
red	3.5	4.5	<u>-31.2</u>		<u>-26.5</u>	<u>-39.2</u>	<u>-34.3</u>	<u>-30.6</u>	<u>-69.4</u>
bold	-13.9	-0.7	-2.6	<u>-23.2</u>		<u>-28.2</u>	<u>-34.7</u>	<u>-35.2</u>	<u>-64.5</u>
shadow	4.8	-12.0	<u>-25.2</u>	<u>-30.0</u>	<u>-23.0</u>		<u>-78.3</u>	<u>-43.3</u>	<u>-105.0</u>
underlined	<u>20.6</u>	<u>-22.3</u>	-6.1	-2.5	-5.6	<u>-45.7</u>		8.9	<u>-97.5</u>
spacing	0.0	<u>25.4</u>	12.7	5.5	-0.4	<u>-11.0</u>	13.6		<u>-73.9</u>
italic	<u>70.1</u>	<u>68.8</u>	<u>68.3</u>	<u>66.6</u>	<u>66.7</u>	<u>56.7</u>	<u>41.9</u>	<u>52.6</u>	

Only underlined + spacing showed improvement over both individually

# Study 3 - results

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

***questions?***