



Usability News is a free web newsletter that is produced by the Software Usability Research Laboratory (SURL) at Wichita State University. The SURL team specializes in software/website user interface design, usability testing, and research in human-computer interaction.

[Barbara S. Chaparro](#), Editor

## Examining Legibility of the Letter "e" and Number "0" Using Classification Tree Analysis

by [Doug Fox](#), [Barbara S. Chaparro](#), & [Ed Merkle](#)

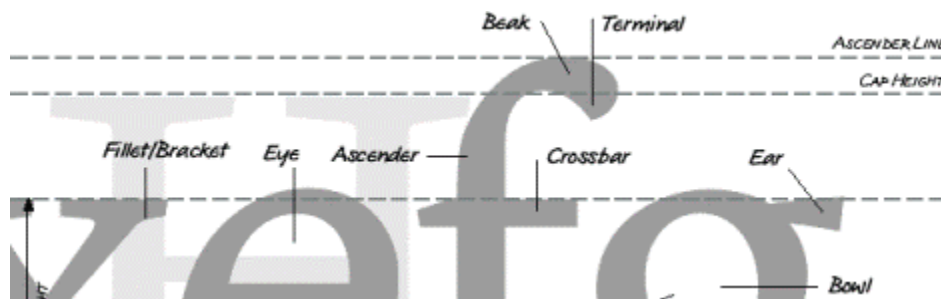
**Summary.** This study investigated the legibility of onscreen typefaces and the influence of individual character features on correct identification. Specific attributes of alphanumeric characters and symbols shown to be the least legible were measured and analyzed using a statistical method called classification tree analysis. Results from this analysis for the letter "e" and the number zero are discussed.

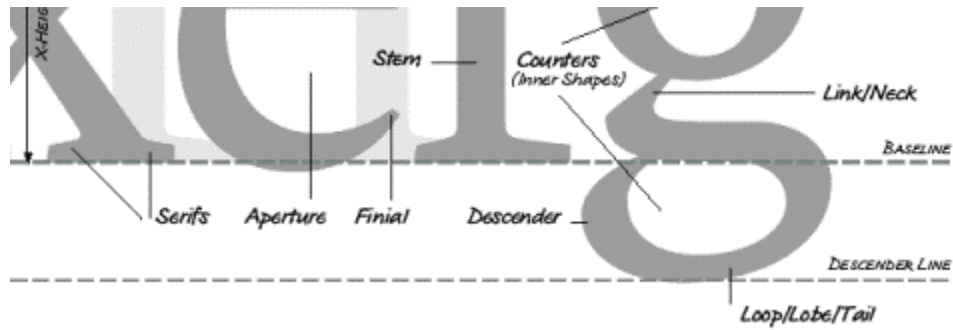
### INTRODUCTION

Typeface legibility is becoming an important issue as reading shifts from print to onscreen (Shaikh & Chaparro, 2004). The common forms of print reading such as newspapers, books, magazines, etc. are being replaced with websites, electronic books, and ezines. This shift of reading methods has resulted in an increased demand for optimized legibility. Lupton (2004) noted "the rise of the Internet as well as cell phones, hand-held video games, and PDAs have insured the continued relevance of pixel-based fonts as more and more information is designed for publication directly onscreen" (p. 27).

There have been many studies that have focused on the legibility of print (Jha & Daftuar, 1981; Mansfield, Legge, & Bane, 1996; Roethlein, 1912; Sanocki, 1988; Woods, Davis, Scharff, 2005) which have resulted in recommendations for different types of printed material. Unfortunately there has not been much research investigating onscreen legibility. Chaparro, Shaikh, and Chaparro (2006) studied the onscreen legibility of six ClearType™ typefaces, developed by Microsoft, which take advantage of sub-pixel rendering. This study found that two of these typefaces (Cambria and Constantia) were more legible than the popular Times New Roman.

Some suggestions have been made as to what specific features of a typeface contribute to the legibility of the characters themselves. Each character is made up of certain attributes (see Figure 1) that distinguish it from other characters within the same typeface and from the same character in other typefaces. These differences are often very subtle but contribute to the overall theme of the typeface. Few empirical studies have been done to determine how these features influence legibility.





**Figure 1. Features that are used for the design of characters** (<http://gmunch.home.pipeline.com/typo-L/faq/anat.htm>).

Tinker (1928) suggested that the size, simplicity of outline, serif style, shading, area of whitespace, and delineation of distinguishing characteristics influence the legibility of lowercase characters. Benjamin Bauemeister developed a program called PANOSE as a guide to measuring typeface attributes including family type, serif style, weight, proportion, contrast, stroke variation, arm style, letterform, midline, and x-height (<http://www.panose.com/ProductsServices/pan2.aspx>).

Pelli, Burns, Farell, & Moore-Page (2006) state that we have commonly based letter identity on visual detection of independent features. However, Pelli et al. (2006) argue that it is much more than just feature detection and the notion of complexity is what actually predicts whether a letter is legible or not. Complexity is measured as a character's inside-and-outside perimeter which is then squared and divided by the "ink" area. For stroke characters complexity is measured by taking four times the length of the stroke divided by the width of the stroke. Pelli et al. (2006) chose complexity as a measure because "it tends to capture how convoluted a character is, and is easily computed, independent of size" (p. 4648).

The purpose of this study was to determine what features of alphanumeric characters and symbols contribute to their onscreen legibility. First, legibility was measured for 47 characters from 20 different typefaces using a classification technique. Participants were asked to identify each character after a very brief exposure time. From this, confusion matrices were created to identify those characters with the poorest legibility. Then, features were measured for each of these characters (see Table 1) and their influence on legibility was analyzed using a statistical method called classification tree analysis.

This article reviews the results of this analysis for two characters, the letter "e" and number zero. The character "e" is commonly confused for other characters such as the letter "c" or "o" (Roethlein, 1912). The zero is one of the most commonly confused numerical characters due to its resemblance to the letter "o". However, there are some typefaces where these characters are confused more than others. Thus, it is important to consider what features of these two characters contribute most to their legibility.

**Table 1. Features used in the study and their definitions.**

Feature	Definition
area	Character width multiplied by character height; represents the total space occupied by a single character.
complexity	Inside-and-outside perimeter <sup>2</sup> divided by "ink" area; for stroked characters it is roughly four times the length of the stroke divided by the width of the stroke or four times the aspect ratio of the untangled stroke.
contrast	Ratio between the thickest point on the stroke and the narrowest point the narrowest stroke weight divided by the widest stroke weight
height	Vertical space covered by a character; measured from the top to bottom of a character including ascenders and descenders.

midline*	Ratio of the height of the baseline to the center horizontal of the character to the overall height of the character
perimeter	This measures the length of the boundary between black and white (the 'perimeter'). Perimeter squared divided by black area is complexity.
stroke variation	A more detailed measurement of contrast that by describing the kind of transition that occurs as the stem thickness changes on rounded glyph shapes.
weight	Letter height over the stem thickness taken from a stem at the letter's midpoint. Weight in general is the darkness (blackness) of a typeface, independent of its size (Bringhurst, 1992).
width	Horizontal space covered by a character; measured from outside pixels on both sides of the character and includes both the strokes and the whitespace.

\*this measure only pertained to the "e"

## METHOD

### Participants

Ten participants (6 male, 4 female) between the ages of 18-25 volunteered for this study. All participants had at least 20/20 vision. All were compensated \$50 for their participation.

### Materials

A Dell Core 2 Duo laptop with a TrueLife™ display and ClearType™ font rendering was used. The screen resolution was 1920 x 1080 with 147 dpi. A program was written in C# to display the character in 10-point font size, and the laptop was positioned at a distance so that the characters, regardless of typeface, were viewed at a visual angle of .08°. Participants used a chinrest to stabilize their head and maintain a constant distance from the monitor. The 20 typefaces investigated are shown in Table 2.

Twenty-six lowercase letters were used in combination with the digits 0-9 and 11 symbols frequently found in mathematical or scientific documents. Tinker's (1928) study on the relative legibility of letter, numbers, and symbols provided guidance in character selection. The symbols used included: ÷ = + ? % ± \$ # @ & !. The chosen symbols had a character height similar to the lowercase letters.

**Table 2. The 20 typefaces examined in this study.**

Typefaces	
Baskerville	DIN 1451 Mittelschrift
Bodoni	Franklin Gothic ITC
Calibri	Frutiger
Cambria	Futura
Candara	Garamond
Centaur	Georgia
Clearview Text	Gulliver
Consolas	Helvetica
Constantia	Rockwell
Corbel	Verdana

### Procedure

Participants were shown a sample of what the characters looked like before each trial started. Each trial started with a "\*" to indicate where the characters were going to appear on the monitor and ended with a "\*" to indicate that the trial was over. Each participant participated in four trials for each typeface with the first trial being excluded from data collection as practice. Each trial consisted of all 47 characters being exposed briefly (34 ms) one at a time with a blanking time of 1.5 second between each exposure. The order the characters and the typefaces was randomized. Characters were displayed in black type on a white background. Character identification was read aloud by the participants and accuracy was recorded by the experimenter.

## RESULTS

Percent correct for each character in each typeface was calculated for all characters; however, only the results for the letter "e" and zero are discussed here (see Table 3 and Table 4). For the letter "e", participants performed the worst with the Garamond typeface with only 10% of the "e" presentations correctly identified. Performance was the best with Clearview Text and Verdana, where the letter "e" was correctly identified 100% of the time. The mean percentage correct across all 20 typefaces for the letter "e" was 87.68% (SD=6.96). For the number zero, participants performed the worst with Constantia (6.7% correct identification) and the best with Centaur and Rockwell (both 100% correct identification). The overall mean percentage correct for the number zero across all 20 typefaces was 69.84% (SD=32.36).

**Table 3. Percentage of correct identifications for the character "e".**

Typeface	Character	Percentage Correct
Garamond	e	10.00
Centaur	e	73.30
Constantia	e	76.70
DIN 1451 Mittelschrift	e	86.67
Baskerville	e	86.70
Bodoni	e	90.00
Futura	e	90.00
Helvetica	e	90.00
Rockwell	e	90.00
Franklin Gothic ITC	e	93.33
Frutiger	e	93.33
Georgia	e	93.33
Corbel	e	96.67
Gulliver	e	96.67
Calibri	e	96.70
Cambria	e	96.70
Candara	e	96.70
Consolas	e	96.70
Clearview Text	e	100.00
Verdana	e	100.00

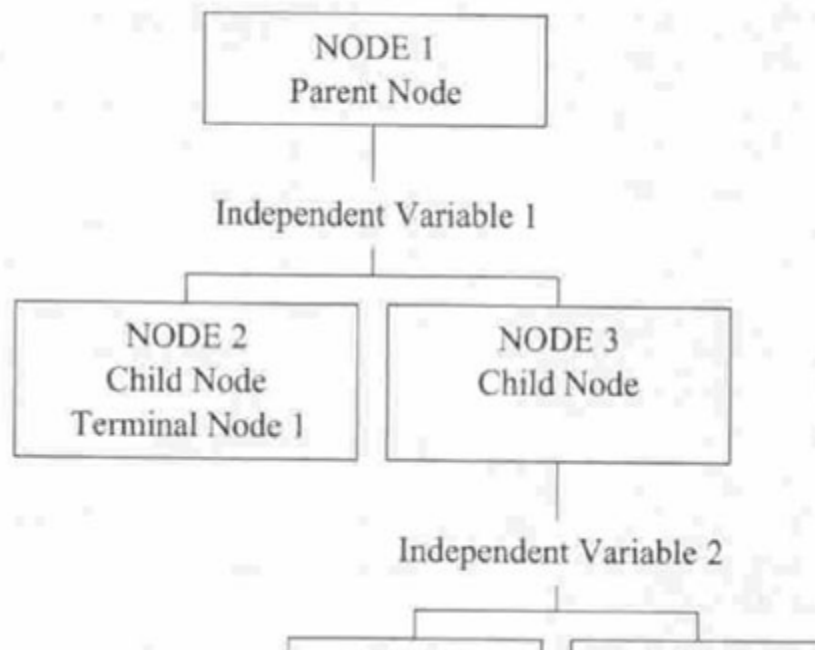
**Table 4. Percentage of correct identifications for the zero.**

Typeface	Character	Percentage correct
Constantia	0	6.70
Corbel	0	10.00
Frutiger	0	10.00

Fruiger	0	10.00
Georgia	0	16.67
Candara	0	16.70
Consolas	0	60.00
Garamond	0	70.00
Bodoni	0	83.30
Baskerville	0	86.70
Cambria	0	86.70
Calibri	0	90.00
Helvetica	0	90.00
Clearview Text	0	93.30
Futura	0	93.33
Gulliver	0	93.33
DIN 1451 Mittelschrift	0	96.67
Franklin Gothic ITC	0	96.67
Verdana	0	96.67
Centaur	0	100.00
Rockwell	0	100.00

Classification tree analysis was used in this study to determine which features (IVs) of the characters were the most influential in correct identification. Classification tree analysis is a nonparametric statistical procedure that identifies homogenous subgroups (nodes) to accurately predict a dependent variable (DV) chosen by the researcher. The subgroups created share common characteristics that influence the DV. A unique visual structure (see Figure 2) that links the nodes together is the result of the classification tree analysis. The branching that is made is based on the analysis of all the independent variables (IVs). The IV that creates the most differentiating groups is selected as the branch IV. This procedure is repeated creating more nodes until there are no other IVs that influence the DV. (Lemon, Roy, Clark, Friedmann, & Rakowski, 2003)

In this study, the DV measure was the number of misclassifications; thus, the more times a character was misclassified the less legible it was. The classification tree analyzed all of the features (area, complexity, contrast, height, midline, perimeter, stroke variation, weight, and width) to determine which feature contributed the most to a misclassification.



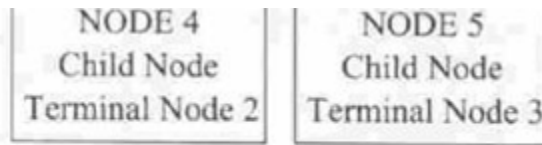


Figure 2. Example of classification tree (Lemon et al., 2003).

### The Letter "e"

The results of the classification tree analysis for character "e" (see Figure 3) suggest that the midline is the feature that influences misclassification the most. Typefaces with a ratio midline value greater than or equal to .54 pixels represent the trials that were more likely to be confused or skipped; conversely, the typefaces that had a ratio midline value of less than .54 pixels represents the trials that were less likely to be confused or skipped. The numbers within the nodes represent the number of incorrect trials and number of correct trials, so for midline values greater than or equal to .54 pixels, there were 49 incorrect trials and 101 correct trials (67% correct identification rate), compared to midline values of less than .54 pixels which contained 25 incorrect trials and 425 correct trials (94% correct identification rate). Figure 4 shows the midline of the worst (Garamond) and the best (Verdana) letter "e". The classification tree analysis did not identify any other feature that influenced misclassification.

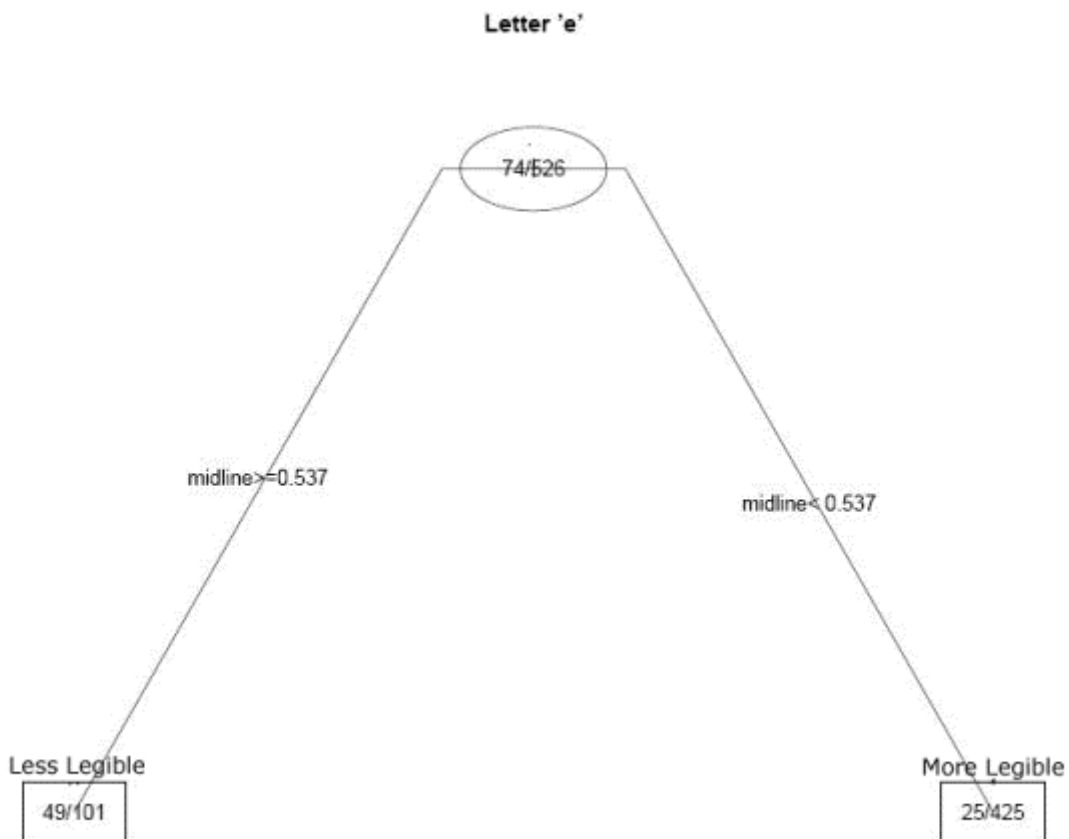
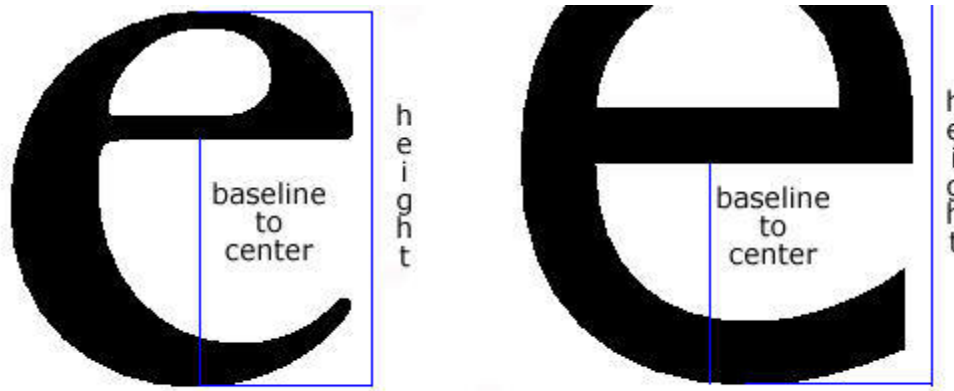


Figure 3. Classification tree for the "e".





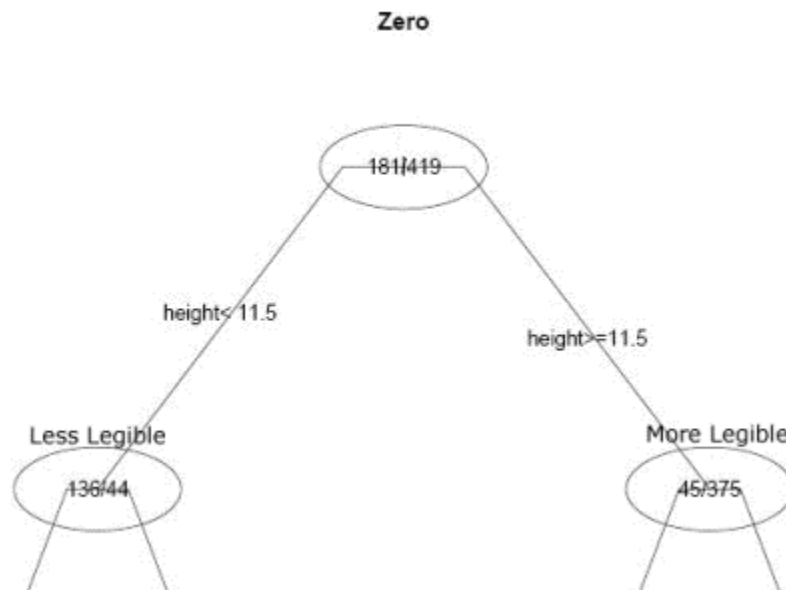
**Figure 4. The Garamond "e" (left) has a much higher midline ratio since the bottom of the eye is higher than that of the Verdana "e" (right). Classification tree results suggest that "e" characters with a higher midline ratio were confused more.**

### The Number Zero

The classification tree results for the number zero are a little more complex and contain more nodes (see Figure 5). The primary branch was for height, suggesting that the height of the zero was the most influential feature out of all the IVs. The zeros with heights equal to or greater than 11.5 showed a better classification ratio of only 45 misclassifications to 375 correct identifications (89% correct identification rate) while the typeface zeros with heights less than 11.5 had a ratio of 136 misclassifications to 44 correct identifications (24% identification rate) (see Figure 6).

When the height was greater than or equal to 11.5 pixels then weight was the next most influential feature. When weight was greater than or equal to 2.82 there was a higher correct identification rate of 24 misclassifications to 336 correct identifications (93% correct identification rate) while a weight less than 2.82 had 21 misclassifications to 39 correct identifications (65% correct identification rate).

When the height was less than 11.5 pixels then the perimeter was the next most influential feature. When perimeter was greater than or equal to 179 there was a higher ratio of correct identifications with 4 misclassifications to 26 correct identifications (87% correct identification rate). If the perimeter was less than 179 there was a ratio of 132 misclassifications to 18 correct identifications (12% correct identification rate). There were no other influential features for the zero according to the classification tree results.



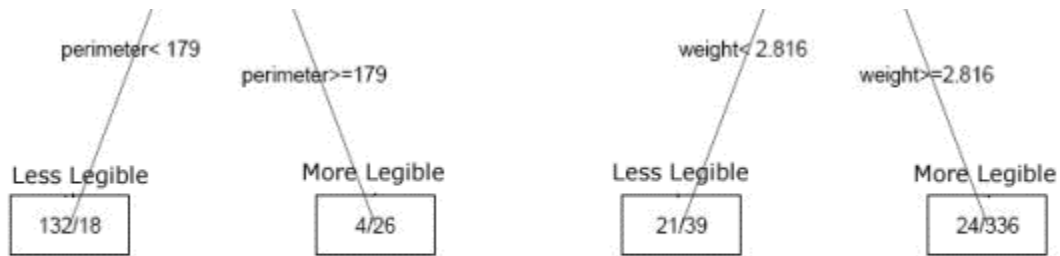


Figure 5. Classification tree for the zero.

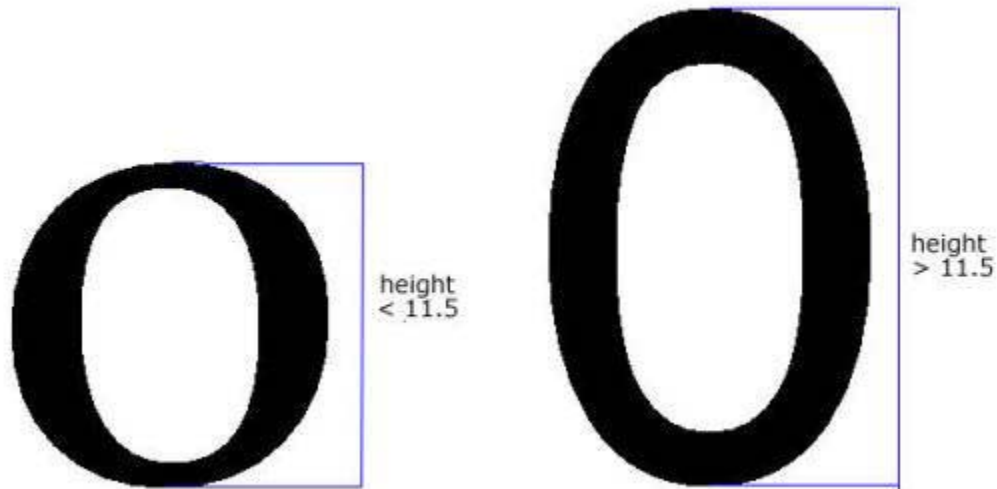


Figure 6. The Constantia zero (left) has a height that is below 11.5 pixels while Rockwell zero (right) has a height above 11.5 pixels.

## DISCUSSION

Results from the classification tree analysis show that the midline was an important feature for the legibility of the character "e". The "e" characters that had a ratio midline of .54 or greater were less legible than those with a ratio midline less than .54. This is supported further by the rank order of percentage correct identifications (see Table 5). Midline is the ratio of the height of the baseline to the center horizontal of the character to the overall height of the character (see Figure 4). The measure of the center horizontal to baseline is affected by the bottom part of the eye of the "e"; the higher the bottom line of the eye is, the greater the midline ratio. Notice that the Garamond (least legible) "e" bottom line of the eye is much higher than that of Verdana (most legible) "e".

Table 5. Rank order of the "e" based on percentage correct. The typefaces above .54 tended to be the least legible.

Typeface	Char.	% correct	Midline
Garamond	e	10.00	0.67
Centaur	e	73.30	0.55
Constantia	e	76.70	0.56
DIN 1451 Mittelschrift	e	86.67	0.40
Baskerville	e	86.70	0.56
Bodoni	e	90.00	0.56
Futura	e	90.00	0.40

Helvetica	e	90.00	0.40
Rockwell	e	90.00	0.40
Franklin Gothic ITC	e	93.33	0.40
Frutiger	e	93.33	0.40
Georgia	e	93.33	0.50
Corbel	e	96.67	0.44
Gulliver	e	96.67	0.52
Calibri	e	96.70	0.44
Cambria	e	96.70	0.44
Candara	e	96.70	0.45
Consolas	e	96.70	0.40
Clearview Text	e	100.00	0.40
Verdana	e	100.00	0.45

For the character zero, height was found to be the most influential feature for correct identification. A zero that had a height equal to or greater than 11.5 pixels was more likely to be identified correctly than a zero with a height less than 11.5 pixels (see Table 6). The shorter zeros (commonly associated with old style numerals) were more likely to be confused with the letter "o". While not surprising, the classification tree analysis provides designers with a specific cutoff of how much height is necessary to improve legibility.

**Table 6. Rank order of the zero based on percentage correct. Each typeface above the bold line were the least legible and had a height measure of less than 11.5 pixels.**

Typeface	Char.	% correct	Height
Constantia	o	6.70	9
Corbel	o	10.00	9
Frutiger	o	10.00	10
Georgia	o	16.67	11
Candara	o	16.70	10
<b>Consolas</b>	<b>ø</b>	<b>60.00</b>	<b>13</b>
<b>Garamond</b>	<b>0</b>	<b>70.00</b>	<b>12</b>
<b>Bodoni</b>	<b>0</b>	<b>83.30</b>	<b>13</b>
<b>Baskerville</b>	<b>0</b>	<b>86.70</b>	<b>13</b>
<b>Cambria</b>	<b>0</b>	<b>86.70</b>	<b>13</b>
<b>Calibri</b>	<b>0</b>	<b>90.00</b>	<b>13</b>
<b>Helvetica</b>	<b>0</b>	<b>90.00</b>	<b>14</b>
<b>Clearview Text</b>	<b>0</b>	<b>93.30</b>	<b>13</b>
<b>Futura</b>	<b>0</b>	<b>93.33</b>	<b>15</b>
<b>Gulliver</b>	<b>0</b>	<b>93.33</b>	<b>15</b>
<b>DIN 1451 Mittelschrift</b>	<b>0</b>	<b>96.67</b>	<b>14</b>
<b>Franklin Gothic ITC</b>	<b>0</b>	<b>96.67</b>	<b>14</b>
<b>Verdana</b>	<b>0</b>	<b>96.67</b>	<b>15</b>
<b>Centaur</b>	<b>0</b>	<b>100.00</b>	<b>12</b>
<b>Rockwell</b>	<b>0</b>	<b>100.00</b>	<b>14</b>

In addition to height, the classification tree analysis showed that the weight of the zero also influenced the legibility of the taller zeros. For zero heights greater than 11.5 pixels, weight values greater than or equal to 2.82 resulted in better legibility (see Table 7). This implies that the darker the stroke of the zero, the better it was identified.

**Table 7. Rank order of the zero based on percentage correct. Each typeface above the bold**

line were the least legible and had a weight measure of less than 2.82.

Typeface	Char.	% correct	Weight
Consolas	ø	60.0	1.86
Constantia	o	6.7	2.25
Frutiger	o	10.00	2.50
Georgia	o	16.67	2.75
<b>Bodoni</b>	0	83.3	2.89
Corbel	o	10.00	3.00
Garamond	o	70.00	3.00
Baskerville	0	86.7	3.21
Cambria	0	86.7	3.25
Clearview Text	0	93.3	3.25
Candara	o	16.7	3.33
DIN 1451 Mittelschrift	0	96.67	3.50
Helvetica	0	90.00	3.50
Rockwell	0	100.00	3.50
Futura	0	93.33	3.75
Gulliver	0	93.33	3.75
Verdana	0	96.67	3.75
Centaur	o	100.0	4.00
Franklin Gothic ITC	0	96.67	4.00
Calibri	0	90.0	4.33

For the shorter zeros (less than 11.5 pixels) the classification tree analysis identified the perimeter to influence legibility. Perimeter values greater than or equal to 179 resulted in better legibility (see Table 8). This was interesting because the perimeter is one of the measures that Pelli et al. (2006) suggests is influential to legibility. (It should be noted that the exceptions to this rule, Centaur and Garamond, were characters with a height > 11.5 pixels.)

**Table 8. Rank order of the zero based on percentage correct. Each typeface above the bold line were the least legible and had a perimeter measure of less than 179.**

Typeface	Char.	% correct	Perimeter
Corbel	o	10.00	130
Constantia	o	6.7	141
Candara	o	16.7	150
Frutiger	o	10.00	156
Centaur	o	100.0	161
Georgia	o	16.67	174
Garamond	o	70.00	176
<b>Baskerville</b>	0	86.7	179
Bodoni	0	83.3	181
Calibri	0	90.0	184
Cambria	0	86.7	184
Consolas	ø	60.0	184
DIN 1451 Mittelschrift	0	96.67	188
Clearview Text	0	93.3	190

Helvetica	0	90.00	196
Rockwell	0	100.00	196
Franklin Gothic ITC	0	96.67	200
Verdana	0	96.67	212
Futura	0	93.33	216
Gulliver	0	93.33	217

The use of classification tree analysis appears to be a promising method of determining which features of a character influence legibility. In addition, these results can be used for creating recommendations on specific design elements. While this article only addressed the results for the letter "e" and the number zero, this technique can be used to analyze any character which tends to be easily confused. Optimizing the legibility of individual characters is not only important for general onscreen reading but also for settings in which users must identify codes or single characters to complete a task; such as an air traffic controller reading symbols on a display.

*Acknowledgment: The authors would like to thank Kevin Larson, Ph.D., for his comments on this article. This study was funded by a grant from the Advanced Reading Technology team at Microsoft Corporation.*

## REFERENCES

- Bringhurst, R. (1992). *The elements of typographical style*. Vancouver: Hartley & Marks, Publishers.
- Chaparro, B. S., Shaikh, A. D., & Chaparro, A. (2006). The legibility of two new ClearType fonts. *Usability News*, 8(1). Retrieved June 30, 2007, from <http://psychology.wichita.edu/surl/usabilitynews/81/legibility.htm>
- Jha, S. S. & Daftuar, C. N. (1981). Legibility of type faces. *Journal of Psychological Researches*, 25(2), 108-110.
- Lemon, S. C., Roy, J., Clark, M. A., Friedmann, P. D., & Rakowski, W. (2003). Classification and regression tree analysis in public health: Methodological review. *Annals of Behavioral Medicine*, 26(3), 172-181.
- Lupton, E. (2004). *Thinking with type: A critical guide for designers, writers, editors, & students*. New York: Princeton Architectural Press.
- Mansfield, J. S., Legge, G. E. & Bane, M. C. (1996). Psychophysics of reading: XV. Font effect in normal and low vision. *Investigative Ophthalmology and Visual Science*, 37(8), 1492-1501.
- Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, 46, 4646-4674.
- Roethlein, B. E. (1912). The relative legibility of different faces of printing types. *The American Journal of Psychology*, 23(1), 1-36.
- Sanocki, T. (1988). Font regularity constraints on the process of letter recognition. *Journal of Experimental Psychology*, 14(3), 472-480.
- Shaikh, A. D., & Chaparro, B. S. (2004). A survey of online reading habits of Internet users. *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting*, 875-879.
- Tinker, M. A. (1928). The relative legibility of the letters, the digits, and of certain mathematical signs. *Journal of General Psychology*, 1, 472-496.
- Woods, R. J., Davis, K., & Scharff, L. F. V. (2005). Effects of typeface and font size on legibility for children. *American Journal of Psychological Research*, 1(1), 86-102.
-

SUBSCRIBE to *Usability News*!