Designing and Evaluating Livefonts

Danielle Bragg
University of Washington
Seattle, WA
dkbragg@cs.washington.edu

Shiri Azenkot
Cornell Tech
New York, NY
shiri.azenkot@cornell.edu

Kevin Larson
Microsoft
Redmond, WA
kevlar@microsoft.com

Ann Bessemans
Hasselt University/PXL-MAD School of Arts
Hasselt, Belgium
ann.bessemans@uhasselt.be

Adam Tauman Kalai
Microsoft Research
Cambridge, MA
adam.kalai@microsoft.com

ABSTRACT
The emergence of personal computing devices offers both a challenge and opportunity for displaying text: small screens can be hard to read, but also support higher resolution. To fit content on a small screen, text must be small. This small text size can make computing devices unusable, in particular to low-vision users, whose vision is not correctable with glasses. Usability is also decreased for sighted users straining to read the small letters, especially without glasses at hand. We propose animated scripts called *livefonts* for displaying English with improved legibility for all users. Because paper does not support animation, traditional text is static. However, modern screens support animation, and livefonts capitalize on this capability. We evaluate our livefont variations' legibility through a controlled lab study with low-vision and sighted participants, and find our animated scripts to be legible across vision types at approximately half the size (area) of traditional letters, while previous smartfonts (static alternate scripts) did not show a significant legibility advantage for low-vision users. We evaluate the learnability of our livefont with low-vision and sighted participants, and find it to be comparably learnable to static smartfonts after two thousand practice sentences.

ACM Classification Keywords
H.5.2 Information Interfaces and Presentation (e.g. HCI): User Interfaces

Author Keywords
Fonts; Reading; Learning; Scripts; Low-vision; Accessibility.

INTRODUCTION
Reading is difficult for the approximately 246 million people[1] with low vision. Low vision can be defined as vision that is not near 20/20 acuity with best correction (e.g. with glasses) and affects everyday life. The small screens of personal computing devices, like smartphones and smartwatches, compound these problems by forcing text to be small and limiting the amount of magnified text that fits on the screen. To improve the usability of personal devices for people with blurry vision, Bragg et al. proposed *smartfonts*, novel scripts for displaying text [5].

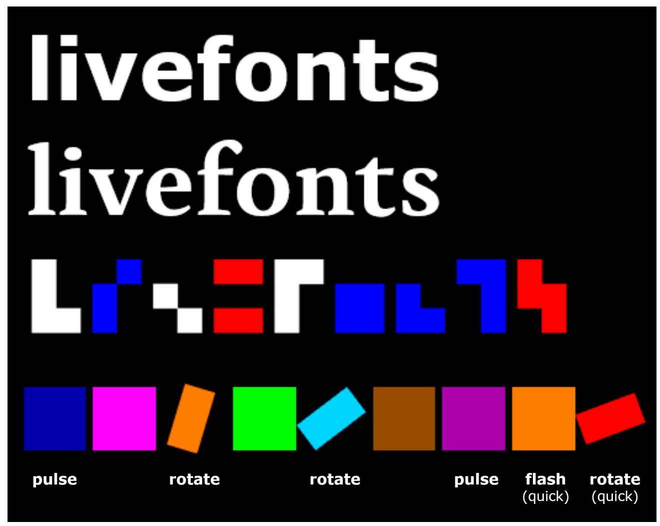


Figure 1: The word "livefonts" rendered in: a standard font (Verdana, top), a font designed for low vision readers (Matilda, second), a smartfont font from Bragg et al. (third), and our livefont (bottom). Note: figures should appear animated when this document is viewed with Adobe Reader.

Bragg et al. showed that by replacing traditional letterforms, smartfonts can improve the reading experience. By redesigning each letterform individually, as in Figure 1, they preserve spelling. Though traditional letterforms are the result of years of refinement, they evolved for display on paper, not screens. Smartfonts challenge our assumption that words should be rendered in traditional letterforms on modern screens. Bragg et al. argue that personal devices have made this challenge possible, by allowing users to adopt new character systems without language reform or mass adoption. Smartfonts can be installed and integrated into existing software systems, e.g.,

[1]This is according to the World Health Organization http://www.who.int/mediacentre/factsheets/fs282/en/.

UIST 2017, October 22–25, 2017, Quebec City, QC, Canada

ACM 978-1-4503-4981-9/17/10. . . $15.00
https://doi.org/10.1145/3126594.3126660

as font files, allowing individuals to change their text displays without impacting anybody else's reading experience. See Bragg et al. [5] for a discussion of additional potential benefits, including aesthetics, privacy, and intellectual challenge.

Past smartfont designs explored *some* opportunities modern screens offer visual text design, namely color, shape, and spacing variations. In this work, we add animation to create *livefonts*. Writing systems have traditionally been limited to what is readily produced by hand: monochromatic characters comprised of lines, curves, and dots. In contrast, screens support virtually limitless colors, textures, and shapes. Past smartfont designs explored colors and shapes impractical to create by hand but easily rendered on screens. However, modern screens also support *live, dynamic* displays, and these past designs were stationary. To the best of our knowledge, we are the first to propose using animation to differentiate characters in a script. Unlike animated text in video, e.g., movie credits, livefonts are novel scripts that use animation itself to define and distinguish characters, animating them individually and systematically. They also integrate into text-based platforms.

Our research questions are: (a) by incorporating animation, how small can we make text to ensure differentiation between characters so that it is still legible to low-vision readers? and (b) can people learn to read animated text? Intuitively, adding animation to text has the potential to compress text while maintaining legibility. In particular, adding one dimension helps remove certain restrictions on other dimensions that previously limited legibility. For example, 26 characters that vary along a single dimension (e.g., shape) cannot be very dissimilar, and thus the smallest legible size is limited. If the 26 characters may vary along two dimensions (e.g., shape and color), pairs of characters can be more similar in one dimension as long as they vary in the other. Hence, we hypothesize that adding animation to the smartfont design space will allow us to generate smartfonts that are legible at significantly smaller sizes than strictly static text, or equivalently significantly more legible at the same size. It is less clear whether or not people would be able to read such animated text.

In this work, we present a livefont to improve legibility by means of recognition for both low-vision and sighted readers. The designs are informed by iterative design and a perceptual study we ran on animation and color. Unlike Bragg et al. ([5]) that evaluated smartfont legibility remotely using blurred text to simulate low vision, we conducted a controlled laboratory study of our livefont's legibility with both low-vision and sighted participants. We find that livefonts are legible at approximately half the size of traditional Latin fonts (lowercase letters typically used for printed English) across vision types. While Bragg et al. also reported a nearly two-fold area decrease in blurred text, our experiments do not reflect a significant advantage for their smartfont among low-vision users, which is consistent with previous work on the inadequacy of disability simulation. Perhaps more surprisingly, livefonts were found to be readable with practice for many people, in an evaluation of learnability similar to that of Bragg et al.

Key contributions of this work are: (a) the idea of livefonts, animated scripts to enhance reading, and specifically to improve legibility for low vision; (b) a controlled in-lab study on livefont legibility with low-vision and sighted participants, using a novel transcription methodology; and (c) an exploration of livefont learnability for low-vision and sighted readers.

RELATED WORK

Work related to livefonts for low-vision readers includes the psychophysics of low-vision reading, animation in relation to reading, and previous work on smartfonts.

Low-vision Reading

Visual perception of text impacts reading. During reading, a retinal image of the displayed text is created in the eye, which is subsequently processed by the brain. This first visual step can impact reading and cause bottlenecks. For example, perceived text size [20] and contrast [23], the difference in luminance between letters and background, impact reading. Because vision is clearest in the fovea, the number of letters identifiable in the periphery is limited. The number of letters recognizable in a single fixation is strongly correlated with reading speed [19, 18]. Livefonts may improve reading, with mastery, if more letters can be recognized in a single fixation.

Low vision is typically characterized by low acuity, making small text illegible and reading slow. Central vision field loss, where the central retinal picture is absent, also leads to difficulty reading [8, 34, 21], and is caused by common diseases such as macular degeneration. Central field vision loss forces people to use peripheral vision to read, which has a lower acuity (increased "blurriness"), and makes reading more difficult. Consequently, low-vision readers often use strong magnification [18], from software such as ZoomText [30], which offers 60x magnification. Magnification makes it particularly hard to fit enough content on small screens to use personal devices effectively [31]. Even on large screens, magnification impedes reading by limiting the window of legible text and requiring panning, which can be so cumbersome that magnification is abandoned altogether [32]. By compressing text display, our livefonts are designed to help address such problems.

Low-vision Fonts

Fonts have been created to improve reading for low vision, including Tiresias [12] and APHont [9]. Due to the wide diversity of low-vision conditions, font personalization is particularly effective for low vision, as evinced by the wide array of fonts low-vision readers created for themselves, given a font personalization tool [1]. Adding animation to smartfont letters, as we do in this work, provides a larger design space for both personalized and general low-vision scripts.

Color and animation can improve legibility for low-vision readers. Color is particularly useful for vision partial to certain light wavelengths [22]. For example, colored lenses can ease or speed up reading [35]. White-on-black text is commonly preferred [29, 32, 37] by readers with a clouded lens, which scatters light and creates glare. Because a black background reduces light and subsequent glare, it often improves reading. Due to this general preference, we use a black background for our designs and experiments. There is evidence that named colors are more easily recognized [36], so our livefonts use colors with distinct English names. Motion has also been leveraged to support visual search for low-vision people [38].

In this work, we compare livefont performance to a state-of-the-art low-vision research-based typeface, Matilda [4, 3]. Characterized by "wide, open, and round letters" with a "friendly feeling" [3], the typeface was designed for low-vision children. The letters are dynamic and solid, constructed and organic, and built on a stable vertical axis. Contrast within the letters is low, to easily enlarge or reduce text. The curves are open, the serifs are asymmetric and emphasized to augment letter individuality and distinctiveness. Based on structured experimentation and design experience, its design is both scientifically rigorous and aesthetically pleasing. In our legibility studies, we use Matilda as an exemplar low-vision font.

Animation and Reading

There are some similarities in motion perception between sighted and low vision people, though the psychophysics of motion perception is an open research area. Typical vision can be more sensitive to peripheral movement than central movement; in contrast, motion detection tends to deteriorate for low vision in the periphery [33]. However, low vision is similarly sensitive to typical vision in central motion detection [33], and in peripheral motion detection with sufficient motion speed [16]. Low vision also exhibits larger variance in motion detection than typical vision. Given the similarities between low vision and typical motion perception, we suspect that animating letters might benefit both groups.

Text animation has been introduced to film through kinetic typography [6]. Kinetic typography is used to add character, engagement, and styling to text in videos. Originally created by film and advertisement companies, various tools have been developed to facilitate kinetic typography more widely (e.g., [24, 10, 14, 17]). Online communities have emerged to share such designs[2]. Unlike livefonts, kinetic typography does *not* create novel scripts. Kinetic typography is typically used to animate choice words or passages for effect, while livefonts animate individual characters systematically and consistently to improve the reading experience. Kinetic typography can be read by anyone familiar with the traditional A-Z, while livefonts are novel scripts that need to be learned. Kinetic typography is also typically constrained to video, while livefonts can integrate into text-based platforms.

Animation has become part of reading on digital devices, as animated emoji and GIFs have integrated into text. Emoji [27] are pictures or animations rendered by text applications, e.g., . Emoji are internationally popular and used for a variety of purposes [15]. They offer richer displays than their predecessor, emoticons, low-tech pictures made of keyboard characters, e.g., :-D. Recently, providers started animating emoji to create even more engaging text displays. Livefonts further this trend of enhancing text through animation.

The successful integration of animated emoji in text demonstrates the technical feasibility of livefonts. A growing Unicode block is reserved for emoji [7], supporting smoother cross-platform rendering, and underscoring their prevalent use. Applications support different renderings of these Unicode characters, and some expand upon this standardized set. For example, Skype provides animated emoticons like a hugging bear, and story-like GIFs they call "Mojis"; and Facebook inserts animated stickers into chat conversations. SMS applications are starting to offer similar animated options. Animated GIFs are also integrated into text-based social media like Twitter, and used by electronic newspapers to introduce or enhance articles. With increased support for animated emoji, livefonts will become similarly technically feasible, and extend this trend towards rich, animated text interfaces.

Smartfonts

Bragg et al. [5] point out that the advent of computer screens presents an opportunity to redesign the way characters look and tailor them to modern screens. Their smartfonts do precisely this, replacing the 26 English lowercase letters with 26 new letterforms. Uppercase letters, which can be demarcated by a single indicator symbol, were left for future work.

Bragg et al. did not evaluate smartfont legibility with low-vision people, instead blurring text artificially. Blurring text uniformly does not adequately simulate low vision, as low-vision conditions vary greatly, and often include limited field of vision. Furthermore, low-vision readers have experience learning how to best use their vision, whereas sighted people typically have no experience with blurry vision. Their results also only show an improvement in legibility for a single level of blur and a small range of sizes. Our evaluation includes low-vision users, an audience that can potentially benefit greatly from new scripts. We also produce stronger results for a range of vision conditions and text sizes. Bragg et al. also evaluated learnability only with sighted participants; our learnability study involves both sighted and low-vision participants.

There is a tradition of unconventional alphabet design predating computer screens. Of particular interest is Green-Armytage's response [13] to the prominent perceptual psychologist Rudolf Anheim's assertion that an alphabet differentiating letters solely through color would be unusable [2]. Green-Armytage compared alphabets comprised solely of different colors, shapes, and faces, and found the color alphabet to be identified most quickly. The idea that constraints on effective alphabet design are looser than we intuitively think is also supported by psychophysical models that posit that visual letter recognition is accomplished through simple feature recognition [28, 11]. Consequently, "any set of characters should do, as long as they contain a sufficient number of simple, detectable features in distinct spatial configurations" [18].

LIVEFONT DESIGNS

We present our livefont in two variations (Figure 2). Since color is more easily recognized in large solid blocks than in detailed strokes, its letters are (animated) squares, which maximize character area. To design these livefonts, we first engaged low-vision readers in an iterative design process to constrict our design space, and then ran a crowdsourced study to fully explore that design space and find the most perceptually distinguishable character sets.

[2](e.g., http://animography.net/collections/typefaces)

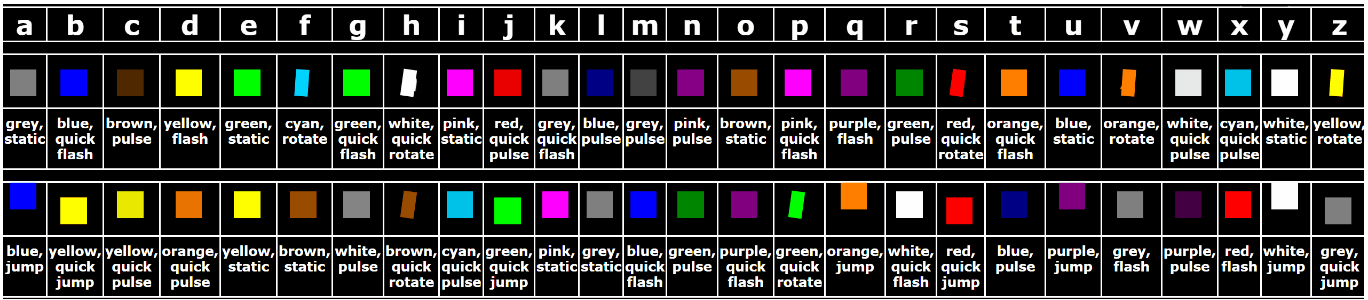


Figure 2: Our livefont versions (top: *Version 1*, bottom: *Version 2*), with descriptions of character color and animation. Characters are colored, animated blocks, chosen to be maximally distinguishable by our color/animation perception study.

Narrowing the Design Space to Color and Animation

We used a participatory design process with low-vision readers to narrow our designs to animated blocks of colors. Because the design space is virtually unconstrained by modern screens, it was important to limit our design space. We chose to involve low-vision readers to best design a livefont that met their needs, as they are a target group who can potentially benefit enormously from improved legibility. We met regularly with local low-vision people, and remotely with a low-vision visual artist, to solicit feedback on designs. At the initial meetings, we conducted informal interviews to learn about their vision and reading. At subsequent meetings, we showed them designs, and gathered feedback and suggestions. Designs explored included sets of colored dots, abstract shapes, various moving gradients, animated/colored traditional letterforms, and traditional letterforms tailored to low-vision. We found that a black background was generally preferable, and that large blocks of color with simple animations were typically easier to perceive, which became our design space.

Our final color palette was: red, orange, yellow, green, cyan, blue, purple, pink, white, grey, and brown. The colors were hand-selected with input from people with low vision, to be discernible and to support clarity on a variety of monitors and personal device screens. They are spaced out in hue and have distinct English names. Our final set of animations were: static, flash, pulse, jump, and rotate, each (besides static) available at two speeds. Pulse is a gradual increase and decrease in color; flash intermittently shows and hides the block's color; jump is an up-and-down shift in position; and rotate is a clockwise rotation. All animations run continuously, and were implemented with CSS animations. These animations involve large area changes over time, which we found to be most discernible during the iterative design process. Further exploration of the color/animation design space is left for future work.

Selecting Alphabet Characters

After narrowing our design space to animated blocks of color, we ran a perception study on the identifiability of characters in our design space at small sizes. This allowed us to choose letters likely to be highly legible. We crowdsourced the study with sighted participants in order to gather sufficient data on the large design space. The study presented a series of animated, colored blocks, and asked participants to identify their color and animation, as shown in Figure 3. Target blocks were presented one at a time. Each participant answered 9 practice questions. They then answered 99 test questions, covering all color/animation combinations.

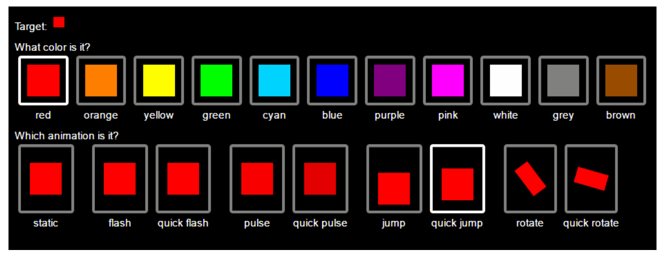


Figure 3: Animation/color perception study task. A target block is shown. The participant is asked to identify the target color (red selected) and animation (jump selected).

The practice target height was 1em (at 14 point), and the test target height was .15em. We wanted to make targets small enough that they were challenging, to gather data at the limits of perception. Because it was a crowdsourced web survey, it was impossible to control for absolute size or visual angle, but the size was commensurate with typical browser font size. In addition, participants were instructed not to zoom in.

We posted the task on Amazon's Mechanical Turk crowdsourcing platform and recruited 50 participants (19 female, 31 male). Ages ranged 24-69 (mean 35.4). Twenty-six owned glasses or contacts, all but three of whom wore them during the study. Three identified as having low vision; nineteen identified as nearsighted; and one as farsighted. Two reported being unsure if they were colorblind, and the remaining forty-eight identified as not colorblind. No participants identified as having a learning disability or as being dyslexic. All participants except one, who dropped out, evaluated all 99 color/animation pairs.

In total, we collected 5386 evaluations, 49-50 for each color/animation combination (49 evaluations for 14 combinations due to our one drop-out). Individual accuracies in identifying the correct color/animation combination ranged from 0.23 to 0.90 (avg. 0.60, dev. 0.18), perhaps due to variance in visual acuity. Mean accuracies in identifying colors and animations are shown in Table 1. Red and blue were the most accurately identified colors, and brown was the least. Out of our animations, static and jump were identified most correctly, with the two rotation speeds least accurately identified.

To obtain our final livefont design from this data, we adopted the optimization procedure from Bragg et al. [5]. Specifi-

(a) Colors

Red	94.44%
Blue	92.87%
Green	88.20%
Cyan	86.89%
Yellow	83.71%
White	76.39%
Orange	75.67%
Purple	75.28%
Grey	71.49%
Pink	63.31%
Brown	48.66%

(b) Animations

Static	97.99%
Jump	96.72%
Quick Flash	89.78%
Pulse	87.80%
Quick Jump	69.22%
Flash	69.03%
Quick Pulse	56.04%
Rotate	50.46%
Quick Rotate	47.64%

Table 1: Color and animation identification accuracy.

cally, we created a 99 x 99 confusion matrix, with rows and columns representing all animation/color combinations. We then select the 26 rows (and corresponding columns) that yield the lowest net confusion. Specifically, let $c_{i,j}$ represent the fraction of times character i was confused for character j. Then we choose the set of 26 characters S that minimizes $\sum_{i,j \in S, i \neq j} c_{ij}$. Because this is an NP-hard problem, we used a branch-and-bound algorithm guaranteed to find an optimal solution, which terminated quickly. This approach does not consider English letter frequency, and we defer deeper explorations into language-dependent optimization to future work. We performed this optimization twice, once including jumping animations, and once without, to produce two variations. We excluded jumping from one to explore if the additional vertical area required for jumping paid off in increased legibility.

LEGIBILITY STUDY

To evaluate the legibility of our livefont, we conducted a controlled laboratory study with both low-vision and sighted participants, unlike previous evaluations using remote sighted participants and simulated blurry vision [5]. Because low-vision readers struggle with legibility, meaning letter and word recognition, our study involves letter and word recognition tests. It does not evaluate readability in terms of comprehension. Future studies are needed to study longitudinal usage, and effects of animating long passages.

The experimental design is within-subjects across a range of traditional fonts and smartfonts. The study had one session, divided into two parts: 1) a transcription task measuring script acuity ($\sim$ 55 min), and 2) a scanning task measuring visual scanning time ($\sim$ 5 min). Participants were compensated $20.

We recruited 25 participants (10 low vision, 15 sighted). Participants varied in age (15-67, mean 34), and gender (15 female, 10 male). Sighted participants were recruited through relevant email lists from the local population. Low-vision participants were recruited from local low-vision mailing lists and support groups. To verify that our participants had low vision, we conducted a brief screening interview. Our low-vision participants had a range of vision conditions including ocular albinism, retinitis pigmentosa, nystagmus, retinopathy of prematurity, and Norrie disease, resulting in a range of reading challenges, in particular difficulty with small letters. Because low vision is very diverse, we did not further categorize low-vision participants by condition, though further improvements may be possible by addressing conditions separately. Participant responses to high-level vision questions are shown in Table 2.

All study procedures were completed using a computer with a standard monitor. All scripts were rendered with a black background. For all scripts other than smartfonts we used white, bold versions, to yield the best results for low vision, and the black characters of Tricolor, the static smartfont we used for comparison (see Figure 11), were made white for visibility on the black background.

Part I: Transcription Methodology

We employed a novel evaluation methodology based on transcribing characters at increasingly small sizes. Evaluating the legibility of smartfonts, scripts that nobody knows how to read, is difficult. Methods for testing acuity typically involve identifying letters by name (e.g., an optometrist's Snellen chart), or reading. These methods do not apply to smartfonts without extensive training. In Bragg et al.'s evaluation, participants identified 5-character strings using multiple-choice options that differ by a single character, at increasingly small sizes. However, this method produces a single piece of information with every task, namely the single mistaken character. Our transcription methodology provides more data with every task, namely which characters were misread as which others.[3]

Because visual acuity varies greatly, especially among low vision, we first calibrated text size for each participant. We presented a list of random[4] sentences in a traditional font, at increasing sizes, and asked participants to select the "smallest readable size," as done with a MNREAD acuity chart [25]. A chin rest was used throughout the study to fix the distance from the screen and control angular text size.

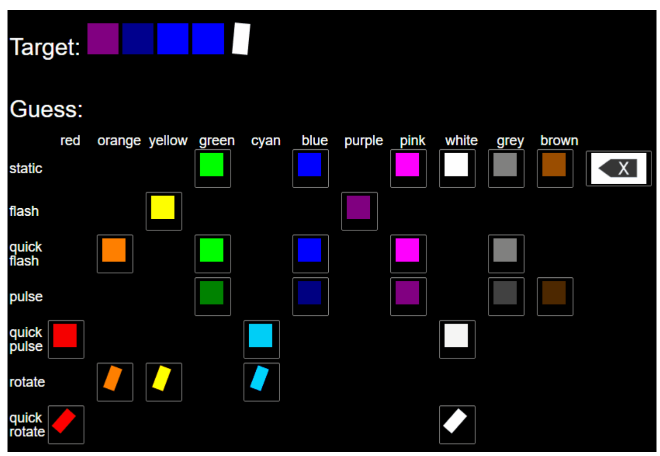


Figure 4: Transcription task with Version 1 – a target string (and partial guess), with a visual keyboard for transcription.

[3]To see the informational advantage, suppose all letters were clear except m was easily confused for n. In order for this to be discovered in a multiple-choice test, a pair of words would have to be generated which differed by an m replaced with a n (or vice-versa), which happens on less than 5% of randomly chosen questions, whereas 32% of random 5-character strings used in transcription tasks would have an m or n, each of which is an opportunity to identify the confusion.

[4]Random sentences from the random sentence generator http://www.randomwordgenerator.com/sentence.php/ (accessed Aug. 2016)

	Owns Glasses		Wore Glasses		Nearsighted			Farsighted			Colorblind		
	Yes	No	Yes	No	Yes	No	IDK	Yes	No	IDK	Yes	No	IDK
Low-vision (10)	9	1	8	2	6	2	2	3	7	0	2	7	1
Sighted (15)	13	2	11	4	11	4	0	11	4	0	0	15	0

Table 2: Self-reported vision descriptions from our legibility study participants, separated into low-vision and sighted groups. Due to variability between low-vision users and even between an individual's two eyes, it can be difficult to answer these questions.

After calibration, participants completed a series of transcription tasks. The task, shown in Figure 4, presented a target string of five randomly chosen characters. Participants transcribed the target characters in order, using an on-screen keyboard. As they clicked on matching characters, their partial guess appeared below the target. The keyboard for the two Latin fonts and for the Latin-esque Tricolor smartfont adopt the standard QWERTY layout. Livefont keyboards were organized into animation-by-color matrices. Rows were organized by animation, and columns by color. Absent characters were left blank. This design supported visual search by color or animation, helping to even the comparison to transcribing traditional letters with the familiar QWERTY keyboard. All keyboards contained a backspace button for corrections.

Each participant completed transcription tasks for all scripts, randomly ordered. Participants transcribed targets from each script at decreasing sizes until failure, when they proceeded to the next script. Each script began at 1.5 times the calibration size, to provide practice before reaching a size where where mistakes were likely due to limited acuity. Each subsequent target was 90% the size (area) of the previous. We operationalized area by normalizing each script's height to yield the same alphabet area, computed as the area of the smallest enclosing rectangle. We stopped participants when they made at least 6 mistakes across two trials to prevent participant frustration and data collection on random guesses. The scripts evaluated were: a traditional font (Verdana), a font specifically designed for low-vision reading (Matilda), our livefont variations, and the "best" static smartfont from previous work (Tricolor).

Part II: Scanning Methodology

After transcription, participants completed scanning tasks. The task presented a random 5-string target, which participants identified in a random pseudo-sentence (Figure 5). The sentence contained 10 strings, one of which was identical to the target. The other 9 strings were generated randomly, with length between 1 and 8. A limitation of this design is that string length can be used as a cue during scanning. They familiarized themselves with each target before viewing the sentence. The time between the sentence's appearance and when they clicked on the match was recorded internally. Selected strings were outlined in white. Corrections could be made by deselecting and selecting a new string. When satisfied, they clicked "Done", and were shown the correct response. Each participant completed five scanning tasks per script.

Figure 5: Scanning task with Latin Verdana Bold. The target string (top), and the selected match outlined in white (bottom).

Script order was randomized. The scripts used were: Verdana, Matilda, Tricolor Braille, Version 2, Version 1, Hebrew, Arabic, Armenian, Devangari, and Chinese. These scripts were chosen for diversity, and taken from previous work [28]. To control for variance in alphabet size, we chose 26 random lowercase characters to represent scripts with more than 26.

Legibility Study Results

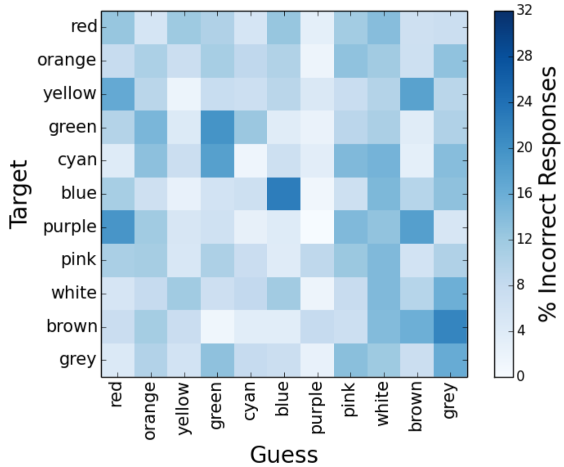
Evaluating our legibility study results requires controlling for variance in script size and eyesight. To compare scripts that vary in character height and width, we use alphabet area as the metric of size. To compare individuals with varied acuity, we normalize individual results for each script by their results for traditional letterforms (Verdana). This yields normalized metrics for both transcription (the *Area Ratio*) and scanning (the *Time Ratio*). Using these metrics, we find evidence that our livefonts are legible much smaller than traditional letterforms, and might support faster scanning with practice. However, due to small sample size and noise, follow-up studies with larger populations are needed to confirm our results.

Part I: Transcription Results

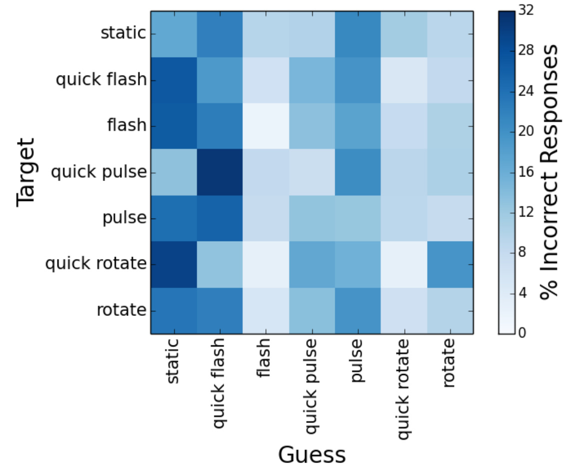
To quantify how small people could make out each script, we define a metric called the *Area Ratio*. As described above, each participant reached a smallest legible size for each script, defined as the first size where they failed to transcribe at least 6 out of 10 characters for that script. To account for differences in acuity across participants, we normalize this failure size with respect to the participant's Latin failure size. We call this ratio their *Area Ratio* for a particular script. The Area Ratio is 1 for any participant with Latin. A value lower than 1 means that the script was more "legible" than Latin, and a value above 1 means that it was less legible. For example, a score of 0.5 means that that script was legible at half the size (area) of Latin, for that participant. We note that an n-fold reduction in area corresponds only to a $\sqrt{n}$-fold reduction in font size according to more standard one-dimensional metrics.

The Area Ratios for our participants are shown in Figure 7.[5] For both low-vision and sighted participants, Version 1 was generally legible at the smallest sizes, at approximately half the size of Latin, with a minority of participants reaching sizes 4-6 times smaller than Latin. We ran one-way ANOVAs with repeated measures and found statistical significance between scripts for both sighted $(F(3,14) = 13.59p << 0.05)$ and low-vision $(F(3,9) = 3.19, p = 0.04)$ groups. Post-hoc paired t-tests with Bonferroni correction show statistical significance $(p < .0083)$ for sighted transcription between Version

[5] In all box plots, the red line is the median. The box lower and upper limits are the 1st and 3rd quartiles. Whiskers extend to 1.5 IQR (interquartile range) in either direction from the 1st and 3rd quartiles.



(a) Color errors for Version 1.



(b) Animation errors for Version 1.

Figure 6: Distribution of errors in identifying (a) colors and (b) animations of Version 1 across all participants in the transcription task. Each row shows the breakdown of mistakes in identifying that color or animation.

1/Matilda and Tricolor/Matilda. These results suggest that livefonts can improve legibility for both low-vision and sighted readers, though follow-up studies are needed for verification.

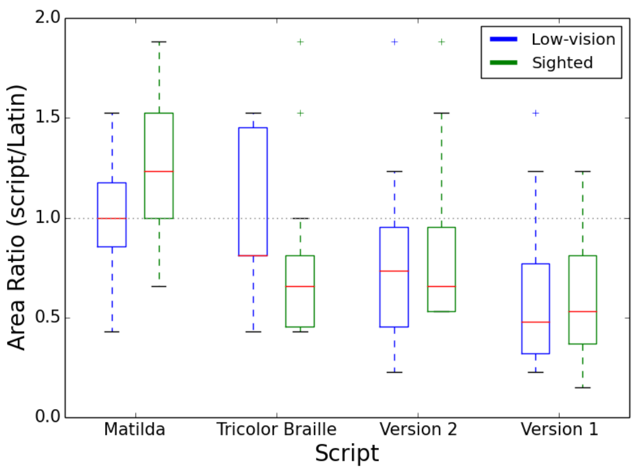


Figure 7: Transcription results. Box-plots of participants' smallest legible size for that script, normalized by their smallest legible Latin size. Lower means more legible.

We also examined the breakdown of transcription errors by color and animation. The average accuracy in identifying color and animation – Version 1: color 74%, animation 74%; Version 2: color 72%, animation 73% – suggests both were salient identifying features. The error distribution for Version 1 is shown in Figure 6. Among colors (Figure 6a), blue characters were most often mistaken for other blue characters. This coincides with participant feedback during the study that the blue characters were hard to see on the black background. Green, white, brown, and grey were also commonly mistaken for other characters of the same color. White characters were often transcribed in place of a variety of colors, perhaps due to the neutrality of white making it a natural random guess. Transcription of white characters for red is due to Version 1 containing both a white and red quick rotate. Among animations (Figure 6b), quick rotate and quick flash were often mistaken for static characters. It is likely that as size decreased, the rotation was lost. Pulse was commonly guessed in place of a variety of animations, possibly due to its sharing properties with many animations (e.g., a similar on-off pattern to quick pulse, flash, and static flash). Similar trends exist for low-vision and sighted groups separately, with sighted errors generally more evenly distributed. Some participants also reported visual fatigue and expressed annoyance at some characters, in particular the blinking ones, while others described the task and scripts as fun.

Part II: Scanning Results

To evaluate our scanning results, we define another normalized metric, the *Time Ratio*. For each participant and every script, we compute a *Time Ratio*, defined as their median scanning time for that font divided by their median Latin scanning time. We normalize by Latin scanning time to account for innate variance in scanning speed. Time Ratios for each script are plotted in Figure 8. We also ran one-way ANOVAs with repeated measures to evaluate statistical significance.

Our livefonts yielded relatively fast scanning times, compared to traditional unfamiliar scripts. Matilda generally produced the fastest scanning times, likely because Matilda uses Latin letterforms, which are easier to identify due to familiarity. Furthermore, those letterforms are tailored to low vision, which likely boosted scanning speed for low-vision participants. For both low-vision and sighted participants, Chinese produced the slowest scanning times, likely due to the absence of additional spacing between Chinese characters in adjacent words and complex Chinese character design. One-way ANOVAs with repeated measures reveal statistical significance between scripts for both sighted $(F(8,14) = 7.39, p << 0.05)$ and low-vision $(F(8,9) = 8.96, p << 0.05)$ groups. Post-hoc paired t-tests with a Bonferroni correction

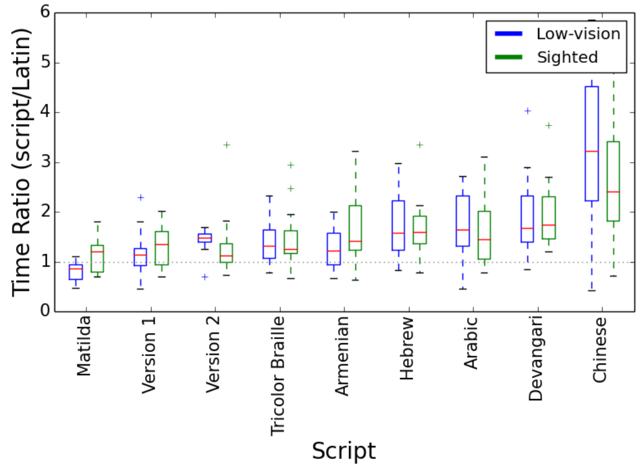


Figure 8: Scanning results. Box-plots of participants' median scanning time finding a 5-character string, normalized by their median time with Latin. Lower means faster.

show significance ($p < .0014$) for sighted scanning between Matilda/Devangari, Matilda/Chinese, Version 1/Chinese, Version 2/Hebrew, Version 2/Devangari; and for low-vision between Version 2/Matilda. Our livefonts' comparable scanning times to the fastest foreign scripts suggest they might yield faster scanning times with practice, though follow-up work is needed to fully explore potential scanning benefits.

LEARNABILITY STUDY

As noted by Bragg et al. [5], for a new character system to be useful, it must be learnable. To evaluate our livefonts' learnability, we adopted Bragg et al.'s evaluation design, where participants learn to read smartfonts through encoded practice questions online. Unlike their evaluation with only sighted people, we recruit both low-vision and sighted participants.

Learnability Study Methodology

We recruited 15 participants (7 sighted and 6 low-vision) for our learning study through Amazon Mechanical Turk. Recruiting low-vision participants was a two-step process. First, we ran a survey on the platform to gather information on people's vision, without any hint of future work. The survey consisted of 10 questions to probe whether or not they were low-vision. This included questions on how they identified (typically sighted, blind, or low-vision), what vision conditions they have been diagnosed with if any, what visual aids they use, and whether their vision is correctable with glasses. 543 people responded, 12 of whom we identified as low-vision. Second, we advertised our learning study to these 12. Sighted participants were recruited from the general Mechanical Turk population by offering 12 workers direct access to our learning study. Our survey showed that only about 2% of workers are low-vision, so the probability of obtaining low-vision workers in the general recruitment is very small.

During the study, participants visited a website that taught them to read Version 1. We chose to study Version 1 over Version 2 due to its better performance in our legibility study. The site teaches the user the new script through several components: 1) an introductory tutorial explaining the livefont structure and providing the alphabet 2) encoded yes/no questions, and 3) flashcards to drill individual character meanings. We used the same 2739 crowdsourced questions as Bragg et al. [5], a supplemented set from MindPixel [26].

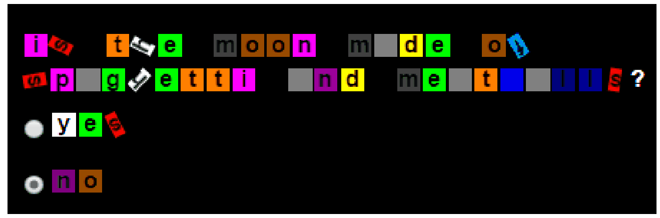


Figure 9: Sample yes/no practice question (*Is the moon made of spaghetti and meatballs? Yes/No*). Here, some letters are overlaid with Latin letterforms to ease the learning curve.

The yes/no questions, pictured in Figure 9, were the primary teaching tool, and response time was the primary metric we used to evaluate learning. A cheatsheet was available upon demand during the yes/no practice questions, showing the alphabet and including mnemonics we designed to help memorability. The cheatsheet overlaid the current question, forcing the participant to remember what they learned from the cheatsheet in order to answer the question. Every tenth question was not encoded (in plain English), for a control comparison.

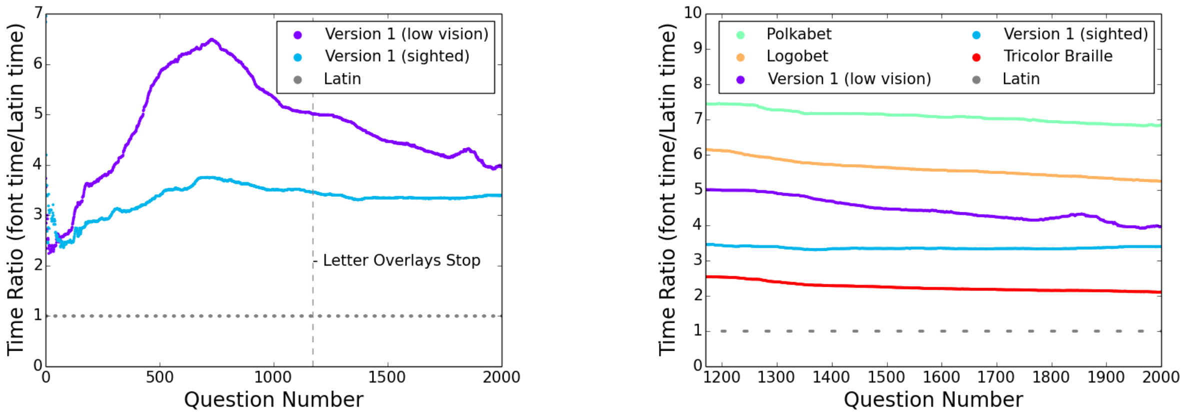
To ease learning, we initially overlaid livefont characters with their traditional Latin representations, and gradually removed the overlays. At the start of the study, all characters were overlaid with Latin. Every 45 encoded questions, another letter's overlay was removed, in alphabetical order, so that after 1170 encoded questions (1300 total questions), no characters were overlaid with Latin, and participants were forced to rely entirely on their memory, plus the supplemental learning aids. This differs from earlier experiments by Bragg et al. [5], where learning was upfront based upon rote memorization and mnemonics, and this difference should be taken into consideration when comparing results across studies.

Participants were paid \$5 for the first 10 questions. After that, they were paid on a per-question basis. They were not paid directly for their flashcard use, though flashcard drills could improve their hourly rate by improving their speed. If they reached the end of the study, they received a \$50 bonus. Because low-vision reading is typically slower than sighted reading, we paid low-vision participants 7 cents per yes/no question, and sighted participants 5 cents per yes/no question. The site was in operation for 10 days.

Several days after the learning study closed, we distributed a survey to obtain feedback and gauge how much learning had converted to longer-lasting memory. The survey quizzed participants on the animation and color of randomly chosen letters, asked participants to rate usefulness of site resources, and gathered open-ended feedback. Participants were paid \$5.

Learnability Study Results

Our primary metrics of learning are time spent answering the yes/no questions, and accuracy. Reading time is a preferred



(a) Average learning curves for Version 1, with an initial letter-overlay aid. (b) Tail of livefont Version 1 learning curves, compared to static smartfonts.

Figure 10: Average livefont response time (a) for Version 1 alone and (b) compared to static smartfonts. Results are normalized by individual average Latin time. Each point is the median of a sliding window of averages across participants to remove outliers.

metric in psychophiscs research [18], and accuracy reflects content understanding.

Learning Accuracy

All participants maintained a high level of accuracy through the experiment. Average (mean) accuracy was 98.23% (min 97.59%, *SD*=0.50%) among low-vision participants, and 97.79% (min 96.67%, *SD*=0.75%) among sighted participants. Given that with random guessing the expected accuracy would be 50%, it is safe to assume that participants were processing question content. The difference between each group's livefont and Latin accuracy was statistically significant (p<0.001, Kruskal-Wallis). Accuracy and response time were significantly correlated for our low-vision group (r=-0.0226, p=0.0062, Pearson), but not for our sighted group (r=0.0027, p=0.7842, Pearson). Interestingly, the correlation for low-vision participants is opposite what would naively be expected – increased time is associated with a *decreased* accuracy (or vice versa). It is possible that while time might help the eyes focus and gather more information, additional time is predominately indicative of difficulty or frustration.

Learning Speed

The fast initial reading speed and subsequent slowdown for Version 1 for both low-vision and sighted participants, as shown in Figure 10a, is attributable to the overlaid Latin letters we initially provided. The learning curves for both low-vision and sighted participants peak well before all letters are hidden, at 1170 questions. It is likely that providing overlays for the letters at the end of the alphabet (e.g., *x*, *y*, and *z*) did not have an impact because these letters are rare, especially in simple sentences like those used in the experiment. The fast initial speed, and low peaks of about 3.5 and 6.5, compared to average starting times of up to 25 times slower than Latin in previous work [5], suggests that overlaying Latin letters can significantly reduce the effort required to start learning a smartfont. It could lower the barrier to learning smartfonts, and make them a more practical option for more people. However, a comparison between these two learning methodologies on the same font would be necessary to verify this conjecture.

The learning curve was initially steeper for low-vision participants, compared to sighted ones. It is possible that this difference is due to an increased effect of removing the overlaid letters for low-vision participants. Removing the overlaid letters forced participants to gradually rely on their color/animation perception alone, which might have been an easier transition for sighted participants. Nonetheless, the normalized speed of our low-vision readers approaches that of the sighted participants as they approached 2000 practice questions. The difference after 2000 questions was not statistically significant, according to an unpaired t-test ($t(6498) = -1.0221, p = 0.3067$). If low-vision and sighted participants continue this trend past 2000 questions, the average low-vision participant might reach or surpass the average sighted participant.

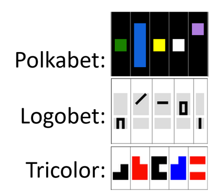
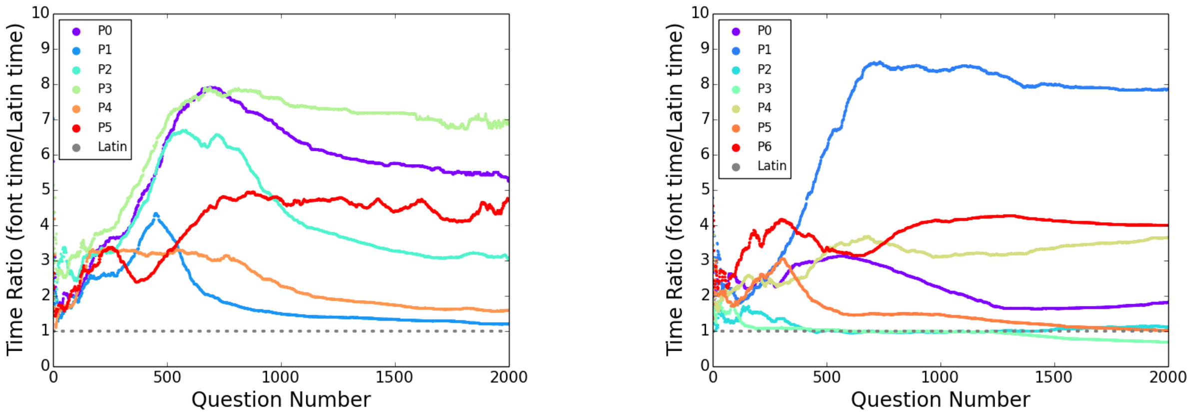


Figure 11: A sample (letters A-E) of the static smartfonts to which we compare livefont Version 1's learnability.

We also compare Version 1 learnability to stationary (non-animated) smartfonts (Figure 10b). The stationary smartfonts (Figure 11) were produced by the same experimental setup [5]. However, in that experiment, no Latin overlays were provided, so we start the comparison where our overlays finished. As shown, reading speed with our livefont is comparable to other smartfonts after 2000 practice sentences. Reading speed for both low-vision and sighted participants was faster than all smartfonts except Tricolor. We ran unpaired t-tests to determine statistical significance at 2000 questions, and found statistical significance between low-vision Version 1 times and each static smartfont.[6] Normalized response times for

[6] Polkabet: ($t(6998) = -17.00, p < 0.0001$), Tricolor: ($t(5998) = 17.94, p < 0.0001$), Logobet: ($t(7498) = -13.60, p < 0.0001$).



(a) Low-vision participants.

(b) Sighted participants.

Figure 12: Individual (a) low-vision and (b) sighted participant learning curves. Results are normalized by individual average Latin time. Each point is the median of a sliding window, for smoothness.

the last 500 questions were our measures for each participant. No statistical significance was found between sighted Version 1 and static smartfont response times.[7] Note that the static smartfont results were produced with sighted participants. It is possible that the significant difference for low vision is due to this difference in vision rather than smartfont design.

Individual learnability of Version 1 varied greatly among both participant groups, as shown in Figure 12. Among low-vision participants (Figure 12a), the livefont was particularly learnable for P1 and P4. These two participants almost reach their Latin reading speed with 2000 practice sentences. On the other hand, Version 1was not very learnable for some participants, in particular P3, who made very little improvement in reading speed and whose Version 1 speed was over 5 times slower than Latin. A similarly wide range in learning is exhibited by our sighted participants (Figure 12b). Bragg et al. [5] also reported a large variance between participants in their learning experiment. While it is difficult to attribute this learnability disparity to visual or cognitive differences, the spread suggests the need for personalized smartfonts or a wider range of smartfonts from which to choose. Alternatively, the potential smartfont user population might not include everyone.

Post-Study Survey

Four low-vision participants and five sighted participants completed the survey we distributed after the learning study closed. Up to a week after practicing, they identified character color with 58% accuracy[8] and animation with 46% accuracy.[9] Given that with random guessing we would expect 9% color and 14% animation accuracy, it seems that participants did commit characters to memory. Interestingly, they remembered more character colors than animations, though there were more colors than animations to confuse. A number of possible explanations could account for this: our mnemonics were more helpful for colors, our colors were more memorable than our animations or had more memorable names, or colors are simply easier to remember than animations.

Participants' average evaluation of resource usefulness (on a scale of 1-5), in order, were: 1) overlaid letters (4.6, std. 0.7); 2) tutorial (4.3, std. 1.1); 3) cheatsheet (3.9, std. 1.1); 4) flashcards (3.7, std. 1.5). They typically found the overlaid letters most helpful. One participant summarized the benefit, "It gave me a lot of help by learning to read the words step by step. By omitting characters this way, it's less of a shock than them disappearing all of a sudden." Participants also found the study generally fun and stimulating. As one participant concluded, "I liked progressing. That was fun."

DISCUSSION AND FUTURE WORK

Livefonts offer exciting possibilities of improved legibility for (small) screen devices, especially for low-vision readers. Magnification helps low-vision readers distinguish letters, but the accompanying loss of visual context and required panning are inconvenient at best. Increased legibility from livefonts can potentially help reduce or eliminate the magnification needed to identify letters. Sighted users can also benefit, especially people reading small text on small screens, those who wear glasses but do not always have them at hand, and people who need glasses but cannot afford them.

While we present the first animated scripts, this work has several limitations. First, we do not claim to have created an optimal animated script. There is a virtually unlimited design space for livefonts, and we only examine two possibilities in this space. Our experiments also have limitations. They do not evaluate readability comprehensively, but rather legibility in terms of character and word identification, and learnability in terms of understanding short sentences. While letter and word identification are fundamental to reading, we do not measure the legibility of long excerpts of text. We also have not studied the long-term impact of reading animated smartfonts. Given the small sample sizes of our studies, larger studies with diverse users are needed to confirm our results and better understand the research space.

[7] Polkabet: $(t(7498) = 0.03, p = 0.9739)$, Tricolor: $(t(6498) = 1.66, p = 0.0964)$, Logobet: $(t(7998) = 0.60, p = 0.4874)$

[8] 64% accuracy for low-vision, 53% accuracy for sighted participants

[9] 36% accuracy for low-vision, 54% accuracy for sighted participants

Livefont design is a rich space for future work. The use of color and animation can potentially distract or annoy the user and inhibit reading, in particular for color or motion insensitive readers. Ideally, users would choose from livefonts with varying color and motion patterns to best suit their sensitivities. For practical considerations, the present work focused on designing and evaluating two options. Long-term studies, beyond this paper's scope, are needed to understand and design livefonts to mitigate these effects. Interactions between adjacent animations or colors can also impair or aid legibility. A thorough understanding of such effects, combined with data on bigram and trigram frequencies, would make it possible to pair animated characters to English characters so as to minimize characters dominating their neighbors. Animating traditional Latin letterforms could also improve text legibility without requiring readers to learn a full set of new characters.

The effects of spacing and timing on livefonts also offers rich opportunities for study. Animations can be sped up or slowed down, and it would be interesting to study which speeds best suit which types of vision, and to see how many distinct speeds of a single animation can be distinguished – in this work we only use two. Synchronization across characters can also yield powerful effects. For example, characters blinking in unison create a unifying effect across the page, whereas staggering can help blinking characters blend in. A "wave" effect can also be made by slightly offsetting adjacent letters, which might impact reading speed by guiding the eyes through the text.

CONCLUSION

In this work, we introduce animated scripts, and present two possible design variations. Though these designs were created with low-vision readers in mind through a structured design process, they are clearly suboptimal. Rather, we have shown that it is possible to learn to read animated scripts, and that this animation can lead to improved legibility.

We evaluated livefont legibility through an in-lab study with transcription and scanning tasks. Unlike previous smartfont evaluations, which simulate low-vision reading using online participants, we used both low-vision and sighted participants in a controlled environment. We also evaluated learnability through encoded practice questions. Unlike previous evaluations, we used both low-vision and sighted learners. Our results suggest that animation can make text legible at significantly smaller sizes than traditional letterforms or even prior smartfonts, and can still be learned with practice.

Livefonts align with user demand for rich reading experiences. Emoticons and emoji have already integrated into text applications, and language experts claim that emoji have become part of the English language. Recently, a growing set of animated emoji, stickers, and GIFs are emerging and integrating into text. Though these animated pictures live next to letters, until now animation has not been considered when designing letters themselves. We have provided evidence that animating letterforms can be useful, and encourage other researchers, typographers, and designers to consider animating scripts.

REFERENCES

1. Aries Arditi. 2004. Adjustable typography: an approach to enhancing low vision text accessibility. *Ergonomics* 47, 5 (2004), 469–482.
2. Rudolf Arnheim. 1956. *Art and visual perception: A psychology of the creative eye*. Univ of California Press.
3. Ann Bessemans. 2016. Matilda: a typeface for children with low vision. *Digital Fonts and Reading* 1 (2016), 19.
4. Ann Bessemans and others. 2012. *Letterontwerp voor kinderen met een visuele functiebeperking*. Ph.D. Dissertation. Leiden University.
5. Danielle Bragg, Shiri Azenkot, and Adam Tauman Kalai. 2016. Reading and Learning Smartfonts. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. ACM, 391–402.
6. Barbara Brownie. 2014. *Transforming Type: New Directions in Kinetic Typography*. Bloomsbury Publishing.
7. M Davis and P Edberg. 2016. *Unicode emoji-unicode technical report# 51*. Technical Report. Technical Report 51 (3).
8. Donald C Fletcher, Ronald A Schuchard, and Gale Watson. 1999. Relative locations of macular scotomas near the PRL: effect on low vision reading. *Journal of Rehabilitation Research and Development* 36, 4 (1999), 356–364.
9. American Printing House for the Blind Inc. 2004. APHont™: A Font for Low Vision. `http://www.aph.org/products/aphont/`. (2004). (Accessed 2017-03-12).
10. Jodi Forlizzi, Johnny Lee, and Scott Hudson. 2003. The kinedit system: affective messages using dynamic texts. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 377–384.
11. Wilson Geisler and Richard Murray. 2003. Cognitive neuroscience: Practice doesn't make perfect. *Nature* 423, 6941 (2003), 696–697.
12. John Gill and Sylvie Perera. 2003. Accessible universal design of interactive digital television. In *Proceedings of the 1st European conference on interactive television: from viewers to actors*. Citeseer, 83–89.
13. Paul Green-Armytage. 2010. A colour alphabet and the limits of colour coding. *JAIC-Journal of the International Colour Association* 5 (2010).
14. Jun Kato, Tomoyasu Nakano, and Masataka Goto. 2015. TextAlive: Integrated design environment for kinetic typography. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 3403–3412.
15. Ryan Kelly and Leon Watts. 2015. Characterising the inventive appropriation of emoji as relationally meaningful in mediated close personal relationships. *Experiences of Technology Appropriation: Unanticipated Users, Usage, Circumstances, and Design* (2015).

16. Joseph S Lappin, Duje Tadin, Jeffrey B Nyquist, and Anne L Corn. 2009. Spatial and temporal limits of motion perception across variations in speed, eccentricity, and low vision. *Journal of Vision* 9, 1 (2009), 30–30.

17. Johnny C Lee, Jodi Forlizzi, and Scott E Hudson. 2002. The kinetic typography engine: an extensible system for animating expressive text. In *Proceedings of the 15th annual ACM symposium on User interface software and technology*. ACM, 81–90.

18. Gordon E Legge. 2007. Psychophysics of reading in normal and low vision. In *OSA Noninvasive Assessment of the Visual System, 1993, Monterey; Portions of this research (MNREAD acuity charts) were presented at the aforementioned conference.* Lawrence Erlbaum Associates Publishers.

19. Gordon E Legge, J Stephen Mansfield, and Susana TL Chung. 2001. Psychophysics of reading: XX. Linking letter recognition to reading speed in central and peripheral vision. *Vision research* 41, 6 (2001), 725–743.

20. Gordon E Legge, Denis G Pelli, Gar S Rubin, and Mary M Schleske. 1985. Psychophysics of reading - I. Normal vision. *Vision research* 25, 2 (1985), 239–252.

21. Gordon E Legge, Julie A Ross, Lisa M Isenberg, and James M Lamay. 1992. Psychophysics of reading. Clinical predictors of low-vision reading speed. *Investigative Ophthalmology & Visual Science* 33, 3 (1992), 677–687.

22. Gordon E Legge and Gary S Rubin. 1986. Psychophysics of reading. IV. Wavelength effects in normal and low vision. *JOSA A* 3, 1 (1986), 40–51.

23. Gordon E Legge, Gary S Rubin, and Andrew Luebker. 1987. Psychophysics of reading-V. The role of contrast in normal vision. *Vision research* 27, 7 (1987), 1165–1177.

24. Jason E Lewis and Alex Weyers. 1999. ActiveText: a method for creating dynamic and interactive texts. In *Proceedings of the 12th annual ACM symposium on User interface software and technology*. ACM, 131–140.

25. JS Mansfield, SJ Ahn, GE Legge, and A Luebker. 1993. A new reading-acuity chart for normal and low vision. *Ophthalmic and Visual Optics/Noninvasive Assessment of the Visual System Technical Digest* 3 (1993), 232–235.

26. Chris McKinstry, Rick Dale, and Michael J Spivey. 2008. Action dynamics reveal parallel competition in decision making. *Psychological Science* 19, 1 (2008), 22–24.

27. Mayumi Negishi. 2014. Meet Shigetaka Kurita, the Father of Emoji. *Wall Street Journal* (2014).

28. Denis G Pelli, Catherine W Burns, Bart Farell, and Deborah C Moore-Page. 2006. Feature detection and letter identification. *Vision research* 46, 28 (2006), 4646–4674.

29. Louise L Sloan. 1977. Reading aids for the partially sighted. *American Journal of Optometry and Physiological Optics* 54, 9 (1977), 646.

30. Ai Squared. 2017. ZoomText. `http://www.aisquared.com/products/zoomtext/`. (2017). (Accessed 2017-03-13).

31. Sarit Szpiro, Yuhang Zhao, and Shiri Azenkot. 2016b. Finding a store, searching for a product: a study of daily challenges of low vision people. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 61–72.

32. Sarit Felicia Anais Szpiro, Shafeka Hashash, Yuhang Zhao, and Shiri Azenkot. 2016a. How People with Low Vision Access Computing Devices: Understanding Challenges and Opportunities. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 171–180.

33. Duje Tadin, Jeffrey B Nyquist, Kelly E Lusk, Anne L Corn, and Joseph S Lappin. 2012. Peripheral Vision of Youths with Low Vision: Motion Perception, Crowding, and Visual Search Peripheral Function in Youths with Low Vision. *Investigative ophthalmology & visual science* 53, 9 (2012), 5860–5868.

34. Stephen G Whittaker and JAN Lovie-Kitchin. 1993. Visual requirements for reading. *Optometry & Vision Science* 70, 1 (1993), 54–65.

35. Arnold J Wilkins, Nirmal Sihra, and Andrew Myers. 2005. Increasing reading speed by using colours: issues concerning reliability and specificity, and their theoretical and practical implications. *Perception* 34, 1 (2005), 109–120.

36. Jonathan Winawer, Nathan Witthoft, Michael C Frank, Lisa Wu, Alex R Wade, and Lera Boroditsky. 2007. Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences* 104, 19 (2007), 7780–7785.

37. Yuhang Zhao, Sarit Szpiro, and Shiri Azenkot. 2015. Foresee: A customizable head-mounted vision enhancement system for people with low vision. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility*. ACM, 239–249.

38. Yuhang Zhao, Sarit Szpiro, Jonathan Knighten, and Shiri Azenkot. 2016. CueSee: exploring visual cues for people with low vision to facilitate a visual search task. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 73–84.