Transactions on Human-Computer Interaction

Volume 12 Issue 1

3-2020

Examining the Influence of Saliency in Mobile Interface Displays

Jeremiah D. Still

Old Dominion University, Department of Psychology, jstill@odu.edu

John M. Hicks

Old Dominion University, Department of Psychology, j1hicks@odu.edu

Ashley A. Cain

Old Dominion University, Department of Psychology, acain001@odu.edu

Recommended Citation

Still, J. D., Hicks, J. M., & Cain, A. A. (2020). Examining the influence of saliency within mobile interface displays. *AIS Transactions on Human-Computer Interaction*, *12*(1), pp. 28-44.

DOI: 10.17705/1thci.00127

Available at http://aisel.aisnet.org/thci/vol12/iss1/2

Transactions on Human-Computer Interaction

Research Paper DOI: 10.17705/1thci.00127 ISSN: 1944-3900

Examining the Influence of Saliency in Mobile Interface Displays

Jeremiah D. Still, John M. Hicks, & Ashley A. Cain

Old Dominion University, Department of Psychology, USA

[jstill; j1hicks; acain001]@odu.edu

Abstract:

Designers spend more resources to develop better mobile experiences today than ever before. Researchers commonly use visual search efficiency as a usability measure to determine the time or effort it takes someone to perform a task. Previous research has shown that a computational visual saliency model can predict attentional deployment in stationary desktop displays. Designers can use this salience awareness to co-locate important task information with higher salience regions. Research has shown that placing targets in higher salience regions in this way improves interface efficiency. However, researchers have not tested the model in key mobile technology design dimensions such as small displays and touch screens. In two studies, we examined the influence of saliency in a mobile application interface. In the first study, we explored a saliency model's ability to predict fixations in small mobile interfaces at three different display sizes under free-viewing conditions. In the second study, we examined the influence that visual saliency had on search efficiency while participants completed a directed search for either an interface element associated with high or low salience. We recorded reaction time to touch the targeted element on the tablet. We experimentally blocked high and low saliency interactions and subjectively measured cognitive workload. We found that a saliency model predicted fixations. In the search task, participants found highly salient targets about 900 milliseconds faster than low salient targets. Interestingly, participants did not perceive a lighter cognitive workload associated with the increase in search efficiency.

Keywords: Mobile Interface, Saliency Model, Visual Search, Cognitive Engineering, Efficiency.

Dirk Hovorka was the accepting senior editor for this paper.

1 Introduction

We have experienced a shift from personal computing to ubiquitous computing (see Alan Kay's Third Paradigm of Computing). In recent history, individuals met their computing needs with a personal computer. However, users now depend on numerous computers to complete daily activities. According to StatCounter (2018), from June, 2017, to June, 2018, across the world, individuals used mobile devices (M = 52% monthly, SD = 1%) to access the Internet more than any other device, while desktops came in second place (M = 44% monthly, SD = 1%). Users' interaction with mobile devices has become pervasive, and, as a result, designers now spend more resources and effort to develop better mobile experiences than ever before. According to Punchoojit and Hongwarittorrn (2017), early mobile design research focused on developing interaction styles or prototyping techniques. But, since 2010, the literature has started focusing more on the usability of design patterns and user factors (i.e., age, accessible). Indeed, the touch screen has impacted mobile design more than any other feature; it removed tactile feedback and introduced gesture interactions, which resulted in placing higher demands on visual attention (Punchooji & Hongwarittorrn, 2017).

Mobile devices present unique usability challenges due to limited screen size, display resolution, connectivity, processing power, and data-entry methods (Ismail, Ahmad, Kamaruddin, & Ibrahim, 2016). Further, one needs to consider differences in use contexts compared to traditional stationary desktop computers (Kjeldskov & Stage, 2004). When designing mobile interfaces, designers must apply guidelines and principles developed for stationary computers carefully (Dunlop & Brewster, 2002). Indeed, the activities and dynamic environment surrounding mobile system interactions can demand greater working memory resources and compete for users' attention. Ahmad, Rextin, and Kulsoom (2018) performed a systematic literature review and reduced the existing 359 smartphone application usability guidelines to seven categories. One such category focused on cognitive load, which refers to how much cognitive effort individuals needs make to process information in their working memory. Research has established that minimizing cognitive load can lead to improvements in user performance and user satisfaction. For instance, Ahmad et al. (2018) suggest that designers "provide little and homogenous information in modules to avoid cognitive load" and provide users with a visual hierarchy of content.

Providing a visual hierarchy follows the logic that one can use it to guide users through a complex display and, thereby, facilitate visual searches. However, designers often do not know the influences on visual searches that arise from pre-attentive information-processing stages—influences such as salience (Treisman & Gelade, 1980). Pre-attentive processing occurs automatically and requires little if any resources; capitalizing on this type of processing is beneficial because the deliberate component of the cognitive information processing system is restricted by a limited working memory capacity (Cowan, 2000). Attention constitutes one mechanism by which the cognitive can accommodate a large amount of information that exceeds working memory capacity. For instance, attention "selects" specific information to undergo additional processing while leaving other information unanalyzed (Johnson & Dark, 1986). Researchers have often described attention as a spotlight (Fernandez-Duque & Johnson, 2002) that moves through an interface. Pre-attentive processing will guide the spotlight to regions in a scene that contain visually unique features (Wolfe, 2007) or objects that will likely support ongoing task needs (Norman & Shallice, 1986). Interfaces present users with an overwhelming amount of visual elements to sort through, which makes it critical that designers have a way to predict where users will deploy their attention. One may characterize searching for a target as easy or effortful depending on how much of the selection process an individual can accomplish pre-attentively.

Efficiency measures used in the human-computer interaction literature represent one method to quantify the cognitive effort associated with a search task (Hornbaek, 2006). Traditionally, researchers used either time to complete a search task or the perceived cognitive effort associated with the search task to represent efficiency. If individuals deploy attention directly to the target, they will have an efficient experience. Therefore, designers need to know what guides users' attention in an interface to improve efficiency. Generally, designers can guide attention with goal-directed influences (i.e., top-down processing) such as familiarity or stimulus-driven influences (i.e., bottom-up processing) such as salience. We discuss these psychological concepts to provide insight into how to improve search efficiency.

Designers often describe attentional deployment in a display as scanpaths (Rayner, 1998). As expected, users fixate on some Web elements and ignore others altogether. The attentional bias to some elements can impact search efficiency (Hong, Thong, & Tam, 2005). Researchers have often used Faraday's (2000) webpage visual hierarchy model to predict content classification from most to least distinct. The first stage of the model (the search phase) focuses on predicting salient interface entry points. Designers reflectively examine the visual elements that comprise the interface and rank order them using the guidelines from what should be most to least influential (e.g., motion, size, images, color, text style). For instance, if an interface lacks motion but includes a large picture, then the picture is the most salient element (i.e., the entry point). The second stage of the model describes further deploying attention between visual elements following an entry point. Many researchers have developed and used a predictable visual hierarchy (e.g., Coa, 2015; Jones, 2011; Soegaard, 2016). Bradley (2015) suggests that designers can create entry points by having a single dominant element (e.g., large picture). However, based on data from Still (2018), even under near-perfect conditions, Faraday's visual hierarchy model could not predict fixations in webpages. Thus, it appears emphasizing element size and an image's presence does not predict individuals' attention. Other researchers (Grier, Kortum, & Miller, 2007) have also come to the same conclusion that Faraday's visual hierarchy fails to predict entry points in the first stage.

The challenges encountered with reflective or heuristic models could result, in part, from the aspects of the selection process that occur pre-attentively. As such, researchers recommend that designers use a computational saliency model (e.g., Itti, Koch, & Niebur, 1998) to predict attentional deployment in an interface (Still, 2018). Every element in an interface is naturally associated with a different amount of visual salience. Accordingly, one can use a computational model to reveal salience by examining contrast differences across key feature dimensions (i.e., intensity, color, edge orientation). Previous research has shown that a saliency model can predict fixations in webpages (Still & Masciocchi, 2010). Further, some evidence shows that visual salience still influences specific product searches (those that rely on more involved top-down processes) in an e-commerce page (Still & Still, 2019). Even with these results, researchers have not previously examined the saliency model in a mobile interface context.

The rapid shift from personal computing to ubiquitous computing often leaves designers to depend on tools not validated for their development needs. For instance, Masciocchi and Still (2013) have shown a saliency model to help designers determine the saliency of interface elements. Designers can use this knowledge to facilitate users' searches. However, it remains unclear whether a saliency model can predict fixations in small mobile interface displays with touch screens. Across two studies, we explored whether a saliency model could predict fixations in small displays and explored the impact that saliency had on tablet search efficiency and subjective workload. In the first study, we examined a saliency model's ability to predict how users deployed attention in small mobile interfaces. In the second study, we examined the influence that visual salience had on search efficiency when users needed to physically touch a target on a mobile device. We did so by co-locating searched elements with higher amounts of salience to see if, as previous research has predicted, search efficiency would improve and user subjective experience of cognitive workload would decrease.

1.1 Influence of Bottom-up and Top-down Processes on Search Performance

When an interface guides a user's attention to irrelevant elements, the search can take longer to complete. Designers can avoid poor search efficiency by understanding and recognizing both the top-down (knowledge of target) and bottom-up (salience) influences guide users' searches. Users often have something they search for in an interface. Their knowledge of the target influences searches from a top-down perspective (e.g., knowing it is red in color and square in form). This research explores the influence of salience during a top-down guided search. Theeuwes (1992, 2004) has shown saliency has a strong influence on visual search performance. For instance, the presence of a salient distractor slowed the response time to a target. This finding suggests that users cannot ignore a salient element in a display even when it lacks relevance to the task (cf. Kim & Cave, 1999). However, designers can find predicting saliency in complex displays difficult.

One may determine visual saliency in simple displays to some degree using reflection (cf. Healey, Booth, & Enns, 1996) or by using a computational model. In either case, uniqueness across a dimension of interest would indicate salience. For instance, a yellow square is salient when surrounded by red squares. Further, users can easily and quickly find the salient yellow square (Treisman & Gelade, 1980). They can immediately find the yellow square no matter the number of red squares, which demonstrates that users have a clear and unconscious bias towards selecting salient objects. Unfortunately, interfaces' complexity and numerous visual elements make it difficult to determine salient display areas by simply looking at them. Our working memory lacks the capacity to consciously examine the interactions between multiple feature dimensions. Therefore, predicting salience in complex displays requires one to employ a computational model.

Researchers have studied visual search in realistic images and found that many of the same findings from basic visual search still apply in more visually complex displays (e.g., Nuthmann, 2014, Wolfe, Võ, Evans, & Greene, 2011). Evidence also suggests that these basic findings apply to artificial complex displays. Still and Still (2019) examined the influence top-down and bottom-up had on product searches in webpages. The authors provided participants with an image of a target product and directed them to report its price. In this way, they created a heavy top-down influence by providing optimal familiarity with the product. Further, they placed the target product at higher or lower salience regions in order to manipulate bottom-up influence. The authors found that a heavy top-down search did not override salience's influence. Subjects found the target product near higher salience regions more quickly than they found the target product near lower salience regions—even in situations without ambiguity about the product's visual appearance. Burke, Hornof, Nilsen, and Gorman (2005) further demonstrate the influence of salience under conditions with heavy top-down influence. They found that salient distractors in the form of banner ads in websites increased search times for the target. Users directed attention to the ads even though they did not pertain to the task. An increase in subjective workload also accompanied the increase in search times. Typically, increases in workload hinder effortless interactions (Krug, 2006). Past studies have shown that workload can affect visual search behavior as well (Faure, Lobjois, & Benguigui, 2016; Recarte, Pérez, Conchillo, & Nunes, 2008; Van Orden, Limbert, Makeig, & Jung, 2001). For designers to create good interfaces, they need to employ this knowledge.

Note that, even though some cases where one can infer salience exist (e.g., the yellow square among red squares), those cases are atypical. According to Johansen and Hansen (2006), designers cannot predict where users will fixate on a webpage. Others recommend that one employ guidelines to predict users' scanpath in webpages (cf. Cao, 2015; Faraday, 2000). Still (2018) empirically examined Faraday's visual hierarchy guidelines for predicting initial fixations during the "search phase". He found the guidelines failed to predict points of entry into a webpage as the size of a Web element or the presence of an image did not predict entry points. Therefore, determining salience with traditional visual hierarchy guidelines does not appear to be an effective approach, and asking a designer to determine the influence of saliency without a computational model will likely not be effective. According to Still (2018), designers ought to employ a computational saliency model to predict saliency influences in interfaces.

1.2 Computational Saliency Model

One cannot easily account for visual saliency (bottom-up influence) in complex and rich images (e.g., interfaces or nature scenes). Indeed, individuals cannot readily introspect about the cognitive processes that determine the distributions of visual salience across a scene. As an analogy, consider what it is like to recognize a coffee cup. It is quick and simple to identify a coffee cup, but we cannot describe the processes we went through to retrieve it from memory, and we may not even know how we first learned about a coffee cup. This inability to introspect about some cognitive processes that guide visual search likely contributes to the ineffectiveness of visual hierarchy heuristics, which require a designer to perform a reflective analysis. The cognitive science community has provided insight into pre-attentive processes by creating biologically inspired computational models to measure salience (e.g., Itti et al., 1998). Itti et al.'s (1998) model determines a saliency level for each pixel by detecting local feature contrasts across three key channels (i.e., intensity, color, and orientation). Differences in local contrast appear to be an effective means to predict fixations. Parkhurst, Law, and Niebur (2002) showed the model could predict fixations across a wide class of images (home interiors, buildings, natural scenes, fractal images). Still and Masciocchi (2010) suggested that the saliency model could also predict fixations in webpages (i.e., visually complex homepages) as they provided evidence that participants tend to look at higher regions of salience first. Others have also shown that a saliency model can predict fixations in webpages with low clutter (Hicks, Cain, & Still, 2017). However, research has not yet explored the impact that display size has on the model's predictive performance.

1.3 Mobile Display Design

Mobile device usage continues to grow in the ubiquitous computing era. We now spend more time interacting with small displays rather than the traditional large displays for various activities (e.g., social media, shopping, banking, emails, and scheduling). Creating usable mobile interfaces is difficult in part due to small screen size (Chittaro, 2011; Dunlop & Brewster, 2002; Ziefle, 2010). Further, portable devices carry other unique limitations such as service quality, display resolution, processing power, and data entry (Harrison, Flood, & Duce, 2013; Zhang & Adipat, 2005). The various contexts surrounding mobile interactions (i.e., user, environment, activity) also constitute a complicating factor (Coursaris & Kim, 2011). Ahmad, Rextin, and Kulsoom (2018) systematically reviewed the literature on mobile applications to form a comprehensive list of usability guidelines. In doing so, they focused on including multiple technology platforms and application genres. They divided the guidelines into seven distinct categories: navigation, content, error handling, input method, equitable use, cognitive load, and design. For our purposes here, we care particularly about the impact that cognitive load has on mobile application usability. Cognitive load is associated with the amount of perceived effort or working memory an interface consumes. To lower cognitive load, interfaces should provide minimal visual information organized into related modules or chunks (Ahmad et al., 2018). Further organization can involve hierarchies that comprise similar modules.

Designers face the difficult task of making critical information easy to find with fewer and often simpler interface elements. Interface display limitations such as screen size can amplify usability issues (Jones, Marsden, Mohd-Nasir, Boone, & Buchana, 1999; Raptis, Tselios, Kjeldskov, & Skov, 2013). In fact, some researchers have shown that screen size can have a greater impact on performance than task complexity (Christie, Klein, & Watters, 2004). Others have found that users find targets on the periphery of small screens more slowly than targets closer to the center (Lim & Feria, 2012). Biggs, Adamo, and Mitroff (2014) found that individuals detected high salient items more often than low salient items in a mobile game that simulated an airport X-ray scanning task. However, they determined salience as an item's average response time rather than using a computational model. The literature offers other mobile-specific human-centered design considerations based on usability investigations (Garcia-Lopez, Garcia-Cabot, & de-Marcos, 2015; Hayes, Hooten, & Adams, 2011). As for visual saliency, we are interested in facilitating the usability dimension efficiency (Hornbaek, 2006).

Designers can make searching easier from a top-down perspective. They can provide clues that users can use to find the correct path to follow (Pirolli, 1997). St. Amant, Horton, and Ritter (2007) examined search efficiency in menu hierarchies on a mobile phone. They found that customizing menu items according to user profiles reduced search times by up to one-third. Besides considering top-down factors that influence phone usability, researchers have focused greatly on exploring the impact that walking has on it (e.g., Lin, Goldman, Price, Sears, & Jacko, 2007). However, there is little research that can help designers increase search efficiency in small interfaces exists. Ahmad et al. (2018) suggested lessening cognitive workload by keeping menu structure simple, minimizing visual information, and employing a visual hierarchy. But, again, researchers have shown that a visual hierarchy approach does not successfully predict fixations (Still, 2018).

The vision science community has primarily explored the saliency model's predictability by employing images displayed at 43.18 centimeters or larger (e.g., Parkhurst et al., 2002; Still & Masciocchi, 2010). Gutwin, Cockburn, and Coveney (2017) examined pop-out effects for large displays and found that motion draws individuals' attention to stimuli even when they view it at a wide visual angle. A small mobile display size (i.e., that measured 20 centimeters along the diagonal) could possibly impact the visual search's efficiency and the visual saliency model's performance. Recent research has found that Fitts' (1954) law does not predict search behavior in mobile displays as well as it does in traditional desktops (Hayes, Steiger, & Adams, 2016), which suggests that some traditional methods of analyzing visual search may not translate to smaller displays as well. Xu et al. (2012) showed that touch saliency (i.e., popular screen touch regions) can predict fixations on mobile devices. However, these authors focused on touch interactions in photographs and not mobile interfaces. Of course, small displays mostly fall in our fovea and, thus, require only a few fixations to gather a complete high-resolution representation. Therefore, eye-tracking measurements might not reflect the subtle differences. However, saliency should continue driving attentional deployment whether individuals need to move their eyes or not. Accordingly, we conducted two studies in which we explored whether saliency could predict fixations in small mobile application interfaces and whether relative target saliency impacted tablet search efficiency and subjective cognitive workload.

2 Study 1: Attentional Deployment

2.1 Method

2.1.1 Participants

Twenty undergraduate students participated in the study (females: 14; English is native language: 18; owns a tablet: 9; daily usage: M = 1.5 hours; age: 16 between 18-23 years and 4 between 26-41 years) in exchange for course research credit. Our university's institutional board approved all study procedures.

2.2 Materials and Apparatus

The stimuli comprised 50 screenshots from NASA's Playbook mobile application (Marquez et al., 2013). Playbook facilitates an astronaut's ability to schedule operations in the larger context of ongoing spaceship activities. The screenshots of interfaces displayed a uniform appearance across the entire set with colored blocks and text labels. We randomly assigned playbooks images to the three possible sizes and were scaled accordingly : 17 images at 675 x 900 pixels (the 100% size) with a visual angle of 17.04° x 22.54°, 17 images at 540 x 720 pixels (the 80% size) with a visual angle of 13.63° x 18.03°, and 16 images at 405 x 540 pixels (the 60% size) with a visual angle of 10.22° x 13.52°). We recorded fixations with a Tobii X3-120 eye tracker. The system collects binocular data and performs tracking using both dark and bright pupil data at a 120 hertz sampling rate. The monitor took up 38.83° of visual angle by 22.26° in the visual field. We displayed images on a monitor 43 centimeters wide and 24 centimeters high with the resolution at 1600 x 900 pixels. Participants viewed the monitor from 61 centimeters away. The smallest mobile interface condition was 42 percent of the presentation size of typical stimuli that the vision science literature has examined (e.g., Parkhurst et al., 2002). We did not constrain participants' head movements. The system had the following tracking accuracy: $M = 0.49°$, $SD = 0.21$. We used Tobii Studio (3.4.6) to define fixations and present the study.

We produced the saliency maps by employing Harel, Koch, and Perona's (2006) MATLAB implementation of the Itti et al. (1998) model. This biologically inspired computational model focuses on processing the feature dimensions color, light intensity, and orientation. The model provides a saliency value for every pixel in the provided image. We normalized these values by dividing all map values by the maximum value for each map and multiplying by 100. A value near 100 signifies high salience (likely to be fixated), while values near 0 signifies low salience (unlikely to be fixated). We used the fixation locations x, y to determine saliency map bin locations for value extraction.

2.2.1 Procedure

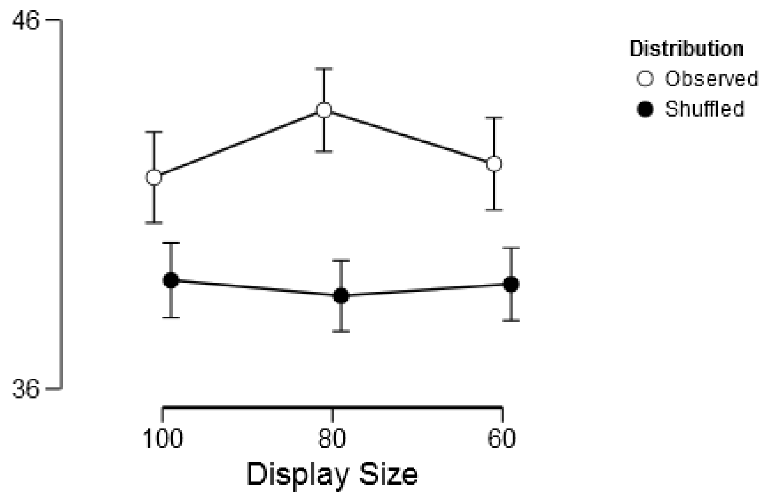
We simply instructed participants to free-view a series of Playbook app screenshots. We presented stimuli on a desktop monitor to allow for optimal eye-tracking accuracy. We filled the empty padding around the stimuli with solid gray. The study started with calibration sequences to set up the tracking system and calculate tracking accuracy. Participants viewed 50 images in three variable sizes (100%, 80%, and 60%), which appeared in random order. Each trial began with a central fixation cross that remained until the participant pressed the spacebar, which signaled they were ready to view an image. Each image appeared for three seconds after which the next trial appeared. This study took approximately eight minutes.

2.2.2 Results

We conducted two-tailed statistical tests with an alpha level of 0.05. We created two distributions to determine whether the saliency model could predict attentional deployment above chance. We formed the observed distribution by extracting values at fixation coordinates (x, y) from the saliency maps (x, y). We formed the shuffled distribution by extracting saliency values at fixation locations (x, y) across every map except for the one to which the fixations belonged. We employed this bootstrapping technique to build a conservative chance performance distribution. It includes top-left and center spatial biases common across the stimuli. These data formed a within-subjects "shuffled" distribution.

We employed a three (display size: 100%, 80%, 60%) by two (distribution: observed, shuffled) by six (fixation: 1, 2, 3, 4, 5, 6) repeated measures ANOVA with Greenhouse-Geisser corrections to determine whether the difference between distributions varied by fixations and/or display size. The observed distribution ($M = 42.47$, $SD = 5.73$) was significantly higher than the shuffled distribution ($M = 38.77$, $SD = 5.27$), $F(1, 19) = 124.58$, $p < .001$, $\eta p2 = .87$). We found an interaction between display size and distribution ($F(2, 32) = 6.87$, $p = .005$, $\eta p2 = .27$) (see Figure 1). Paired samples t-tests confirmed that the observed distribution was significantly higher than shuffled for each display type (100%: $t(19) = 7.39$, $p < .001$; 80%: $t(19) = 8.55$, $p < .001$; 60%: $t(19) = 6.59$, $p < .001$). These findings suggest that one can employ a saliency model to predict fixations in small displays. Further, by using the shuffled distribution as a comparison, we show that the model produces better predictions than known visual search biases (such as the center of the display or top-left). A greater difference between observed and shuffled at the 80 percent display size compared to the difference at the 100 and 60 percent display sizes appears to have driven the interaction between display size and distribution. We did not find an interaction between distribution and fixations (F(3,

63) = 0.09, p = .971, ηp2 = .01) or between display size, distribution, and fixations (F(5, 94) = 0.77, p = .575, ηp2 = .04).



The figure error bars represent 95% confidence intervals.

Figure 1. Mean Saliency Values across the First Six Fixation Locations for the Observed and Shuffled Databases by 100%, 80%, and 60% Display Sizes

3 Study 2: Search Efficiency

In the second study we conducted, participants visually searched Playbook interfaces on a tablet after hearing a target description. We directed their search with a goal, but we placed their search target alongside a region associated with either high or low salience. In the study, we examined whether visual saliency could account for improved search efficiency when participants had to touch their target following a directed search. Further, we explored whether participants subjectively experienced a change in cognitive workload based on target saliency.

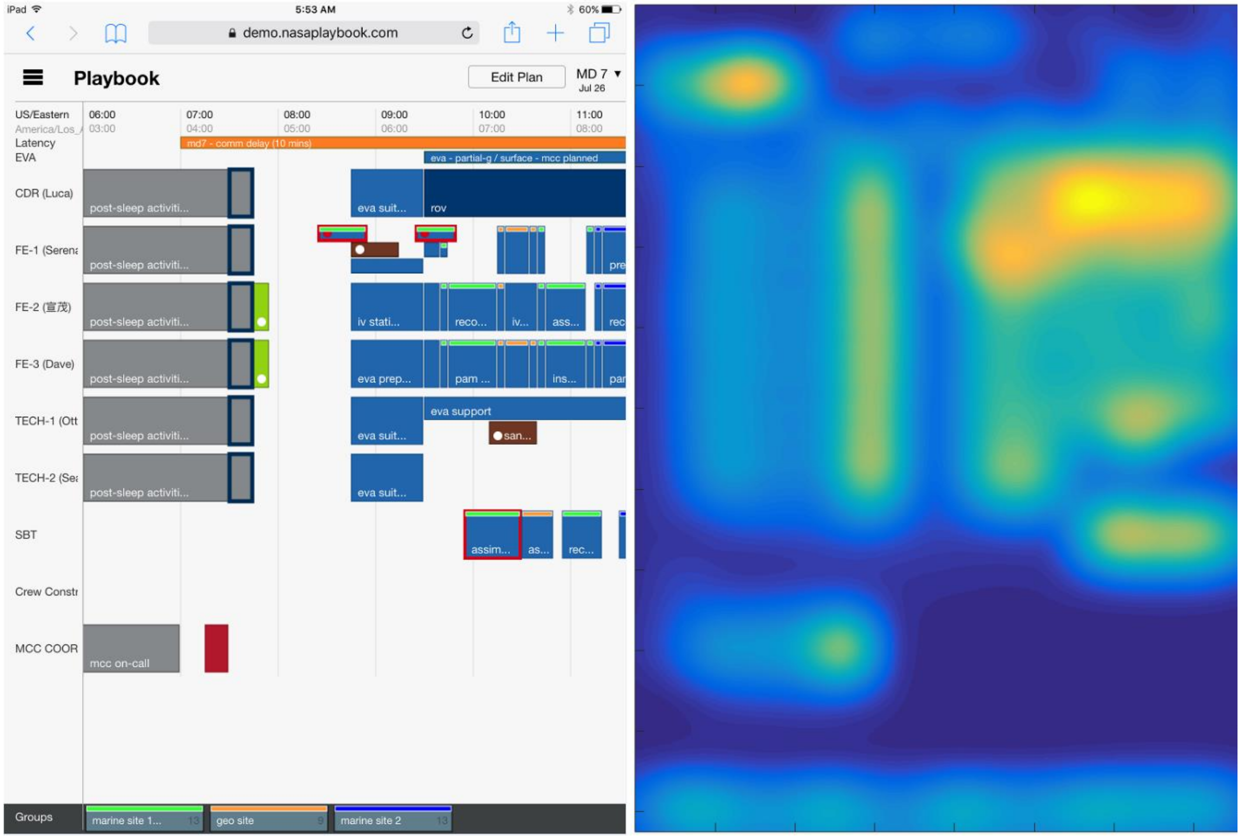
3.1 Method

3.1.1 Participants

Fifty undergraduate students participated (females: 36; English is native language: 45; owns a tablet: 27; daily usage: M = 1.5 hours; age: 37 between 18-23 years and 13 between 24-41 years) in exchange for course research credit. Our university's institutional board approved all study procedures.

3.1.2 Materials and Apparatus

We presented the stimuli and recorded responses using the experimental software Paradigm. By doing so, we could precisely collect participants' reaction time on an iPad Mini (a screen that measured 20 centimeters along the diagonal). We employed the saliency model to process the 18 Playbook screenshots (see Figure 2). Using the saliency heat maps, we determined two unique elements for each screenshot: one associated with high saliency (hot region) and the other with low saliency (cold region). We distributed the search target's spatial location across the interface in an effort to prevent favored search locations. The target element's label became the search target.



Hotter regions in the map reflect areas likely to be salient, while colder regions reflect areas likely to be not salient. We asked the participants to “select CDR rov” (high saliency) or “select MCC on-call” (low saliency).

Figure 2. A Playbook Screenshot Displayed with its Corresponding Saliency Map

3.1.3 Procedure

We presented participants with an iPad Mini and introduced them to the NASA Playbook application. For their task, we told them to find a target and tap it as quickly and as accurately as possible while seated. Previous work has examined the impact that sitting or walking has on subjective cognitive workload while interacting with a mobile device and found no significant effect on subjective cognitive load (Alshehri, Freeman, & Freeman, 2013). Participants needed to correctly tap on the search target to advance. Each trial began with a screen that instructed participants to wait until the researcher told them the target. Once they were ready, participants tapped on the screen to bring up the target display. Participants had one practice trial and then did two blocks of nine trials each. We blocked targets with either high or low saliency. We counterbalanced block order across participants. After each block participants completed the NASA task load index (TLX; Hart & Staveland, 1988)—a subjective measure of workload that contains six scales (physical demand, mental demand, temporal demand, effort, performance, and frustration level) that range from low to high. This study took approximately 12 minutes.

3.2 Results

We conducted two-tailed significance tests with an alpha level of 0.05.

3.2.1 Target Selection Reaction Time

We filtered the data to remove abnormally long responses. We defined outliers as being two standard deviations above the mean of all reaction times in the study. This procedure identified 36 trials (4% of the total data) for removal. We missed data for 26 trials (3% of the total data)—typically due to the participant accidentally starting a trial before we read the search target aloud.

A paired samples t-test revealed that participants found the high saliency targets (M = 5191 milliseconds, SD = 1846) faster than the low saliency targets (M = 6096 milliseconds, SD = 2075; t(49) = -2.77, p = .008, d = -0.39).

3.2.2 NASA TLX

We calculated the normalized mean TLX ratings by equally weighting each subscale. A paired samples t-test did not reveal a statistical difference between high saliency targets (M = 23.94, SD = 18.70) and low saliency targets (M = 26.22, SD = 23.94; t(49) = 1.86, p = .068, d = 0.26). We show mean ratings in Table 1.

Table 1. Mean TLX Ratings (Standard Deviations)

	Scale						
Saliency	Mental demand	Physical demand	Temporal demand	Performance	Effort	Frustration level	Total
Low	35.00 (18.67)	10.82 (7.06)	34.28 (19.10)	25.56 (16.55)	32.75 (19.49)	18.88 (17.72)	26.22 (23.94)
High	31.01 (18.81)	9.80 (7.21)	34.06 (20.19)	21.93 (14.49)	29.77 (18.93)	17.07 (17.45)	23.94 (18.70)
Total	33.01 (18.76)	10.31 (7.12)	34.17 (19.56)	23.75 (15.59)	31.26 (19.17)	17.97 (17.52)	25.08 (18.90)

4 Discussion

Users' interactions with mobile devices have become ubiquitous. Designers today rapidly develop applications that help users complete daily activities. These mobile applications appear on smaller displays compared to the previous personal computer displays (Chittaro, 2011; Dunlop & Brewster, 2002; Ziefle, 2010). Further, reseachers encourage designers to create visually simpler interfaces (Ahmad et al., 2018). Accordingly, mobile interface designers work hard to develop better user experiences. According to Punchooji and Hongwarittorrn (2017), with the removal of tactile feedback and the introduction of invisible gesture interactions, mobile technology today places greater demands on visual attention than ever before. The time or effort required to perform a task often determines visual search efficiency (Hornbaek, 2006). Previous research has shown a computational saliency model can predict attentional deployment in traditional webpages (Masciocchi & Still, 2013). Further, co-locating search targets with higher salience regions can improve an interface's efficiency (Still & Still, 2019). However, the model remains untested in the key mobile technology design dimensions such as small displays and touch screens. Accordingly, across two studies, we explored whether we could extend a saliency model to smaller mobile application interfaces.

In the first study, we examined using a saliency model in the context of small mobile displays. Research has extensively explored saliency as a bottom-up guiding influence during visual search under basic research conditions (e.g., Kim & Cave, 1999; Treisman & Gelade, 1980). And, while it may be possible to predict the influence of saliency in simple displays (cf. Healey & Booth, 1996), one cannot easily use reflective heuristics in visually complex interfaces (Still, 2018). Therefore, researchers have employed computational models to capture saliency (e.g., Still & Masciocchi, 2010). We started by examining the influence of saliency under the conventional free-viewing task while recording eye movements to capture attentional deployment in interface displays. The results confirmed that a saliency model could predict fixations in mobile interfaces. Further, from a practical perspective, the model's predictability does not appear to depend on display size. However, we did not replicate one common finding associated with the saliency model: we did not find that the model accounts for the first fixation better than the other fixations (cf. Parkhurst et al., 2002).

The cognitive science literature (e.g., Foulsham, Chapman, Nasiopoulos, & Kingstone, 2014) remains unclear about whether saliency would influence a naturalistic task such as searching in a tablet's calendar for an upcoming activity. However, Still and Still (2019) have provided evidence that a strong top-down driven search in e-commerce webpages does not overcome salience's influence. Individuals found targets near higher salience regions faster than those near lower salience regions. However, we note that the authors ended participants' visual search task when they fixated on the searched product, which perhaps limited the opportunity to observe the influence of top-down effects.

In the second study, we had participants find and touch a specific interface element on a mobile device. We found that saliency did impact search efficiency in a directed task that required a touch response. When searching for a target higher in saliency, participants performed almost a full second more quickly compared to when they searched for a target lower in saliency. Interestingly, participants did not subjectively view these high saliency target searches as having a lower cognitive workload than low saliency target searchers, which may have resulted from the ease of the search task. Grier (2015) explored over 237 publications employing the NASA TLX and categorized reported workload ranges by task type. Publications exploring visual search tasks had global workload values between 29 and 79. Our average rating of 25 was below the minimum range of value. Clearly, our visual search tasks for high and low saliency targets required little cognitive workload. Future work could introduce additional mental workload while participants completed the search task. In the real world, we rarely would only think about scheduler interactions and have the exact element name in mind. The additional mental workload might make an increase in efficiency more noticeable to users.

Given this study constitutes the first to explore visual saliency in tablet interactions, we highlight some limitations with the knowledge that future investigations will more richly explain saliency's influence on mobile interactions. We employed NASA's Playbook interface, which offered naturalistic mobile application interface stimuli (i.e., simple and highly consistent). Thus, we did not directly manipulate saliency—we captured it only with a saliency model (a conventional practice in the vision science literature). Further, we used novices as participants rather than expert users. Individuals highly familiar with Playbook might possess top-down knowledge that affects their search behaviors in a way that decreases saliency's influence. However, search times might also depend on the interfaces element variability over time. That is, the display might be too dynamic for a search to become an automatic process (Shiffrin & Schneider, 1977) and, thus, prevent individuals from developing the long-term memory structures that they need to successfully overcome saliency's influence. The current study cannot answer these questions.

5 Conclusion

Our studies provide empirical evidence that one can extend Itti et al.'s (1998) saliency model to predict how individuals will deploy their overt attention in small mobile interface displays. Further, we found that associating a target with higher salience improved search efficiency by 900 milliseconds for mobile searches that require touch input. Surprisingly, participants did not experience a reduced subjective workload in response to greater search efficiency—possibly due to the task's overall ease and individuals' automatically processing saliency's bottom-up guiding influence. We encourage designers with an interest in increasing their users' search efficiency to consider employing a computational saliency model (cf. Still & Still, 2019) rather than a conventional visual hierarchy approach to predict attentional deployment (Still, 2018). The computational model makes saliency's bottom-up influence visible, which allows an individual interface element's relative saliency to be apparent. This knowledge about salience's varying degrees in an interface should help designers improve their interfaces' usability.

6 Acknowledgments

We thank Steven Hillenius and Dorrit Billman for providing access to the Playbook software through a demo site. We also thank the Virginia Space Grant Consortium for financially supporting this research.

References

Ahmad, N., Rextin, A., & Kulsoom, U. E. (2018). Perspective on usability guidelines for smartphone applications: An empirical investigation and systematic literature review. *Journal of Information and Software Technology, 94*, 130-149.

Alshehri, F., Freeman, M., & Freeman, A. E. (2013). Are you smart enough for your smartphone? A cognitive load comparison. *A proceedings of the 24th Australasian Conference on Information Systems*, 1-11.

Biggs, A. T., Adamo, S. H., & Mitroff, S. R. (2014). Rare, but obviously there: Effects of target frequency and salience on visual search accuracy. *Acta Psychologica, 152,* 158-165.

Bradley, S. (2015). Design principles: Dominance, focal points and hierarchy. *Smashing Magazine*. Retrieved from https://www.smashingmagazine.com/ 2015/02/design-principles-dominance-focal-points-hierarchy/.

Burke, M., Hornof, A., Nilsen, E., & Gorman, N. (2005). High-cost banner blindness: Ads increase perceived workload, hinder visual search, and are forgotten. *ACM Transactions on Computer-Human Interaction, 12,* 423-445.

Cao, J. (2015). *The 5 pillars of visual hierarchy in Web design.* Retrieved from https://thenextweb.com/dd/2015/04/30/the-5-pillars-of-visual-hierarchy-in-web-design/

Christie, J., Klein, R. M., & Watters, C. (2004). A comparison of simple hierarchy and grid metaphors for option layouts on small-size screens. *International Journal of Human-Computer Studies, 60*, 564–584.

Chittaro, L. (2011). Designing visual user interfaces for mobile applications. In *Proceedings of the 3rd ACM International Symposium on Engineering Interactive Computing Systems.*

Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 154-176.

Coursaris, C. K., & Kim, D. J. (2011). A meta-analytical review of empirical mobile usability studies. *Journal of Usability Studies, 6*, 117-171.

Dunlop, M., & Brewster, S. (2002). The challenge of mobile devices for human computer interaction. *Journal of Personal and Ubiquitous Computing, 6*, 235-236.

Faraday, P. (2000). Visually critiquing web pages. In *Proceedings of the 6th Conference on Human Factors and the Web.*

Faure, V., Lobjois, R., & Benguigui, N. (2016). The effects of driving environment complexity and dual tasking on drivers' mental workload and eye blink behavior. *Transportation Research Part F, 40,* 78-90.

Fernandez-Duque, D., & Johnson, M. L. (2002). Cause and effect theories of attention: The role of conceptual metaphors. *Review of General Psychology, 6*, 153-165.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47, 381-391.

Foulsham, T., Chapman, C., Nasiopoulos, E., & Kingstone, A. (2014). Top-down and bottom-up aspects of active search in a real-world environment. *Canadian Journal of Experimental Psychology, 68*, 8-19.

Garcia-Lopez, E., Garcia-Cabot, A., & de-Marcos, L. (2015). An experiment with content distribution methods in touchscreen mobile devices. *Applied Ergonomics, 50*, 79-86.

Grier, R. A. (2015). How high is high? A meta-analysis of NASA-TLX global workload scores. In *Proceeding of the 59th Human Factors and Ergonomics Society Conference*.

Grier, R., Kortum, P., & Miller, J. (2007). How users view web pages: An exploration of cognitive and perceptual mechanisms. In P. Zaphiris, & S. Kurniawan (Eds.), *Human computer interaction research in Web design and evaluation* (pp. 22-41). Hershey, PA: Idea Group.

Gutwin, C., Cockburn, A., & Coveney, A. (2017). Peripheral popout: The influence of visual angle and stimulus intensity on popout effects. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*.

Harel, J., Koch, C., & Perona, P. (2006). Graph-based visual saliency. In *Proceeding of Neural Information Processing Systems.*

Harrison, R., Flood, D., & Duce, D. (2013). Usability of mobile applications: Literature review and rationale for a new usability model. *Journal of Interaction Science, 1*(1), 1-16.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.) *Human mental workload.* Amsterdam: North Holland Press.

Hayes, S. T., Hooten, E. R., & Adams, J. A. (2011). Visually-cued touch gestures for accurate mobile interaction. In *Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting.*

Hayes, S. T., Steiger, J. H., & Adams, J. A. (2016). Modelling touch-interaction time on smartphones. *Behaviour & Information Technology, 35*, 1022-1043.

Healey, C. G., Booth, K. S., & Enns, J. T. (1996). High-speed visual estimation using preattentive processing. *ACM Transactions on Human Computer Interaction, 3*(2), 107-135.

Hicks, J. M., Cain, A. A., & Still, J. D. (2017). Visual saliency predicts fixations in low clutter web pages. In *Proceedings of the Human Factors and Ergonomics Society International Conference.*

Hong, W., Thong, J. Y., & Tam, K. Y. (2005). The effects of information format and shopping task on consumers' online shopping behavior: A cognitive fit perspective. *Journal of Management Information Systems, 21*(3), 149-184.

Hornbaek, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies, 64*(2), 79-102.

Ismail, N. A., Ahmad, F., Kamaruddin, N. A., & Ibrahim, R. (2016). A review of usability issues in mobile application. *Journal of Mobile Computing & Application, 3*(3), 47-52.

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based fast visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 20*(11), 1254-1259.

Johansen, S. A., & Hansen, J. P. (2006). Do we need eye trackers to tell where people look? In *Proceeding of Computer-Human Interaction Extended Abstracts on Human Factors in Computing Systems*.

Johnson, W. A., & Dark, V. J. (1986). Selective attention. *Annual Review of Psychology*, *37*, 43-75.

Jones, B. (2011). Understanding visual hierarchy in *web* design. *Envato.* Retrieved from http://webdesign.tutsplus.com/articles/understanding-visualhierarchy-in-web-designewebdesign-84

Jones, M., Marsden, G., Mohd-Nasir, N., Boone K., & Buchanan, G. (1999). Improving web interaction on small displays. *International Journal of Computer and Telecommunications Networking, 31,* 1129–1137.

Kim, M. S., & Cave, K. R. (1999). Grouping effects on spatial attention in visual search. *General Psychology, 126*(4)*,* 326-352.

Kjeldskov, J., & Stage, J. (2004). New techniques for usability evaluation of mobile systems. *International Journal of Human-Computer Studies, 60*(5-6), 559-620.

Krug, S. (2006). *Don't make me think: A common sense approach to Web usability*. Berkeley, CA: New Riders.

Lim, J. J., & Feria, C. (2012). Visual search on a mobile device while walking. In *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services*.

Lin, M., Goldman, R., Price, K. J., Sears, A., & Jacko, J. (2007). How do people tap when walking? An empirical investigation of nomadic data entry. *International Journal of Human Computer Studies*, *65*(9), 759-769.

Marquez, J. J., Pyrzak, G., Hashemi, S., Ahmed, S., McMillin, K., Medwid, J., Chen, D., & Esten, H. (2013). Supporting real-time operations and execution through timeline and scheduling aids. In *Proceeding of the 43rd International Conference on Environmental Systems*.

Masciocchi, C. M., & Still, J. D. (2013). Alternatives to eye tracking for predicting stimulus-driven attentional selection within interfaces. *Journal of Human-Computer Interaction*, *28(5)*, 417-441.

Nuthmann, A. (2014). How do the regions of the visual field contribute to object search in real-world scene? Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 40*(1)*,* 342-360.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (vol. 4, pp. 1-18). New York, NY: Plenum Press.

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, *42*(1), 107-123.

Pirolli, P. (1997). Computational models of information scent-following in a very large browsable text collection. In *Proceeding of the Human Factors in Computing Systems Conference*.

Punchoojit, L., & Hongwarittorrn, N. (2017). Usability studies on mobile users interface design patterns: A systematic literature review. *Journal of Advances in Human-Computer Interaction,* 1-22.

Raptis, D., Tselios, N., Kjeldskov, J., & Skov, M. (2013). Does size matter? Investigating the impact of mobile phone screen size on users' perceived usability, effectiveness and efficiency. In *Proceedings of 13th Annual MobileHCI Conference*.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372-422.

Recarte, M. A., Pérez, E., Conchillo, A., & Nunes, L. M. (2008). Mental workload and visual impairment: Differences between pupil, blink, and subjective rating. *The Spanish Journal of Psychology, 11*(2)*,* 374-385.

Shiffrin, R. N., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic, attending, and a general theory. *Psychological Review, 84*(2), 127-190.

Soegaard, M. (2016). Visual hierarchy: Organizing content to follow natural eye movement patterns. *Interaction Design Foundation.* Retrieved from https://www.interaction-design.org/literature/article/visual-hierarchy-organizing-content-to-follow-natural-eye-movement-patterns

St. Amant, R., Horton, T. E., & Ritter, F. E. (2007). Model-based evaluation of expert cell phone menu interaction. *ACM Transactions on Computer-Human Interaction, 14*(1), 1- 24

StatCounter. (2018). *Desktop vs mobile vs tablet market share worldwide* [data from June, 2017, to June, 2018]. Retrieved from http://gs.statcounter.com/platform-market-share/desktop-mobile-tablet/worldwide

Still, J. D. (2018). Web page visual hierarchy: Examining Faraday's guidelines for entry points. *Journal of Computers in Human Behavior, 84*, 352-359.

Still, J. D., & Masciocchi, C. M. (2010). A saliency model predicts fixations in Web interfaces. In *Proceeding of the 5th International Workshop on Model-Driven Development of Advanced User Interfaces*.

Still, J. D., & Still, M. L. (2019). Influence of visual salience on webpage product searches. *ACM Journal of Transactions on Applied Perception, 16*(1), 1-11.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics, 51*(6), 599-606.

Theeuwes, J. (2004). Top-down search strategies cannot override attentional capture. *Psychonomic Bulletin & Review, 11*(1), 65-70.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97-136.

Van Orden, K. F., Limbert, W., Makeig, S., & Jung, T. (2001). Eye activity correlates of workload during a visuospatial memory task. *Human Factors, 43*(1)*,* 111-121.

Wolfe, J. M. (2007). Guided search 4.0: Current progress with a model of visual search. In W. Gray (Ed.), *Integrated models of cognitive systems* (pp. 99-119). New York, NY: Oxford.

Wolfe, J. M., Võ, M. L. H., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences, 15*(2)*,* 77-84.

Xu, M., Ni, B., Dong, J., Huang, Z., Wang, M., & Yan, S. (2012). Touch saliency. In *Proceedings of the 20th ACM International Conference on Multimedia.*

Zhang, D., & Adipat, B. (2005). Challenges, methodologies, and issues in the usability testing of mobile applications. *International Journal of Human-Computer Interaction, 18*(3), 293-308.

Ziefle, M. (2010). Information presentation in small screen devices: The trade-off between visual density and menu foresight. *Journal of Applied Ergonomics, 41*(6), 719-730.

About the Authors

Jeremiah Still is an Assistant Professor at Old Dominion University. He received the prestigious Earl Alluisi Award from the American Psychological Association's Division 21 for outstanding achievements made by an engineering psychologist within ten years of earning their PhD. His Psychology of Design laboratory explores the relationship between human cognition and technology; specifically, he is focusing on usable cybersecurity, visual attention, and intuitive design. He earned a PhD from Iowa State University in Human-Computer Interaction.

John Hicks is a doctoral candidate in the Human Factors Psychology program at Old Dominion University. His research interests include visual attention and memory. Specifically, he focuses on examining factors that influence attentional guidance in complex displays, most notably in web pages. He earned his Master's in Experimental Psychology from Villanova University.

Ashley Cain is a Human Factors Engineer. She currently performs research to ensure the safety of robotic technology used for surgery. She completed her PhD in human factors psychology at Old Dominion University focusing on the human side of cybersecurity. She earned her master's degree from San Jose State University in Experimental Psychology.