

# Employing a Computational Model to Reveal Stimulus-Driven Influences in Interfaces

Proceedings of the Human Factors and Ergonomics Society Annual Meeting  
1–5

Copyright © 2024 Human Factors and Ergonomics Society  
DOI: 10.1177/10711813241260301  
journals.sagepub.com/home/pro



Jeremiah D. Still<sup>1</sup> and Mary L. Still<sup>1</sup> 

## Abstract

It is challenging to design complex interfaces that support efficient visual searches. It requires understanding how numerous sources of information contribute to attentional selection, along with an understanding of the tools and tasks used to study visual search. This work is intended to examine visual salience, a pre-attentive process often overlooked in traditional design approaches. Although the importance of visual salience in task-driven searches has been debated, recent evidence suggests this stimulus-driven property contributes to the guidance of attention. Computational modeling has proven effective in predicting the influence of salience (e.g., N-SEEV and Attentional Priority models). With advances in this area of research, continuing examination of the applications and limitations of computational salience models is warranted. Therefore, the present work also examines the predictive power of computational models in displays like web pages and mobile interfaces. Emphasis is placed on the role salience should play in optimizing interface search efficiency.

## Keywords

salience model, applied attention, interface design, efficiency

## Statement of Purpose

Designers typically want their system displays to be usable. In other words, users ought to be able to interact with systems safely, efficiently, and intuitively. A well-designed interface capitalizes on users' natural perceptual capabilities. This is only achievable by understanding the cognitive processes that underly visual search. Visual search is supported by an array of complex cognitive mechanisms, but it is also characterized by notable resource capacity limitations. Naturally, capacity restrictions result in selection of a subset of available visual information for further processing (see classic review: Johnston & Dark, 1986). Therefore, understanding what factors influence attentional selection becomes critical for predicting search efficiency.

The present work is meant to provide an overview of stimulus-driven influences from a Cognitive perspective, recognize the tradition of salience modeling in Human Factors, and explore how salience can be revealed in displays. An exhaustive review of computational salience models is beyond the scope of this proceeding as that would cover a broad multidisciplinary literature. Instead, we will highlight models developed to support designers. More specifically, those models that can be used to improve search efficiency in interfaces like webpages and mobile interfaces. A discussion of the strengths and weaknesses of these models provides direction for future research.

## Guidance of Visual Attention

### *Role of Stimulus-Driven Influences*

Designers strive to create products that guide users effectively through an interface, but many are unaware of influences arising from pre-attentive stages of processing (Treisman & Gelade, 1980). With the ability to predict these influences, designers could leverage this fast and low-resource stage of processing to improve designs. For instance, research has shown that co-locating task-relevant information with salient visual features reduces search times and often facilitates successful task completion (Wolfe, 2007; Wolfe & Horowitz, 2017). Salient features are those that are visually unique relative to their surroundings (Duncan & Humphreys, 1989; Parkhurst et al., 2002). Some examples of high salience are apparent, like bolded text amongst non-bolded text. Text is typically uniform making bolding a simple way to attract the reader's attention. However, many interfaces are rich with visual media. This obscures points of high salience as multiple feature dimensions are in play.

---

<sup>1</sup>Old Dominion University, Norfolk, VA, USA

### Corresponding Author:

Jeremiah D. Still, Old Dominion University, 5115 Hampton Blvd, Norfolk, VA 23529, USA.

Email: jstill@odu.edu

Given this complexity, designers are often forced to make their best guesses about which regions are salient or depend on a computational model to reveal the salience distribution. Among the features that contribute to salience, transient signals that exogenously attract attention (Yantis & Jonides, 1984) like flashes or sudden onsets are typically ranked the highest in attentional priority. Because these signals are apparent to designers and they do not add insightful value to the usage of salience models, they are not discussed further in this work.

The world we operate in is overwhelming, but attention regulates what enters awareness. Attention is guided by two low-level, early influences to cope with our conscious capacity limitations. One influence, visual saliency (Itti et al., 1998), impacts humans uniformly (an exception being those with vision-related abnormalities like colorblindness). A second influence comes from culturally based schemas learned during previous interaction experiences (c.f., Chun, 2000; Malcolm & Henderson, 2010). These schemas can reflect design conventions (J. D. Still & Dark, 2010, 2013); they act to prioritize certain interface regions over others. Salience models have been developed to accommodate both types of influence; they can be used to identify the highest salient regions likely to be viewed first (J. D. Still, 2018) and to provide a ranking of attentional deployment. The latter of which is much more difficult to achieve.

According to Wickens and McCarley (2008), a “computational model guided solely by bottom-up salience calculation can do a reasonable job of simulating human search behavior (Itti & Koch, 2000), but models incorporating a top-down component perform better (Navalpakkam & Itti, 2005)” (p. 71). We agree that top-down influences have large effects on visual search performance. But, accounting for enough top-down information to make accurate predictions is often not practical, particularly for items associated with a large number of tasks - consider the possible tasks for something as simple as a website homepage. Regardless, salience still guides attention. Even so, the amount of control salience exerts over attention compared to an individual’s task and goals has been the subject of much debate in the Cognitive literature (for seminal discussion, see Luck et al., 2021).

### **Brief Overview of Attentional Capture Theory**

Beyond transient signals (e.g., blinking lights), certain visual features can automatically capture attention and serve as a singleton. Wolfe and Horowitz (2004) identified a variety of these guiding attributes. They suggested that some features were *undoubted* (e.g., color, motion, orientation) while others were *probable* (e.g., shape, break in a line, direction of curvature) in their ability to automatically capture attention. A group of researchers provided evidence that attentional capture can occur regardless of the task, suggesting that salience can exert influence independently of other top-down influences (c.f., Jonides & Yantis, 1988; Theeuwes, 1993).

However, other researchers examined the claims of independence and discovered boundary conditions (c.f., Folk et al., 1992). For instance, they demonstrated that capture was contingent on the task attentional set. After decades of debate, we are starting to see some resolution. At least, we see enough resolution for practical application to move forward.

According to Luck et al. (2021), “the debate has undergone [a] significant transformation, partly as the result of growing evidence that an inhibitory process can sometimes prevent attentional capture even if a stimulus produces a priority signal (Cosman et al., 2018; Gaspelin & Luck, 2018, 2019; Vatterott et al., 2018; Weaver et al., 2017)” (p. 2). Although individual explanations for these effects may vary, all research camps agree that singletons are suppressible when there is a small set size and none of the elements are highly salient. Taking Theeuwes’ viewpoint, singleton suppression can occur in small display searches as a result of learning. Both reactive and proactive suppression result from regular experiences with distractor features and/or spatial locations (Theeuwes, 2010). Observers can learn that a location or feature is irrelevant, which produces rapid disengagement during searches. Wang and Theeuwes (2020) explored the top-down feature suppression theory further. They replicated Gaspelin et al.’s (2015) four item display showing the suppression effect. However, in larger displays with 6 and 10 items, the traditional stimulus-driven effect returned; there was no singleton suppression. From a practical perspective, this means that designers creating typical displays with more than four items, they do not need to worry about inhibitory processes and can follow a traditional understanding of stimulus-driven influences. Simply put, recent findings in this area suggest that in typical displays (those with more than four items) all salient elements, transient or feature-based, influence attention regardless of task demands.

## **Computational Models of Salience**

### ***Historical Implementations in Human Factors***

The Human Factors literature often highlights the importance of considering visual salience, especially, in alert design. For instance, Steelmann and colleagues, integrated a computational salience model into Wickens et al.’s (2003) SEEV model (Steelman-Allen et al., 2009). The SEEV model attempts to account for four dimensions influencing attention: Salience, Effort needed, operator’s Expectancy, and task-relevance (SEEV). The updated N-SEEV includes the processing of static visual saliency by adopting Itti et al. (1998). This goes beyond the traditional SEEV dimension of color to include the feature dimensions of intensity and edge orientation. They showed a good fit between model predictions and pilot’s overt attention.

Still and colleagues have also shown that a saliency model can predict fixations within web pages (Masciocchi & Still, 2013; J. D. Still & Masciocchi, 2010) and mobile interfaces

(Stills et al., 2020). Further, salience continues to guide attentional deployment even during a naturalistic task like e-commerce product searches (J. Still & Still, 2019). Web developers can employ a salience model instead of depending on their understanding of visual hierarchy. Traditionally, designers have relied on heuristics like those provided by Faraday (2000), to gain insight into element visual dominance across an interface. Unfortunately, there is limited empirical evidence supporting the effectiveness of these heuristics within displays (Grier et al., 2007). For example, J. D. Still (2018) showed Faraday's heuristics do a poor job of predicting initial fixations. However, he did reveal that spatial positioning and visual distinctiveness were the best predictors. These considerations can be computationally implemented in a salience model that includes a convention map instead of relying on an individual to interpret and apply the heuristic.

### *Interface-Oriented Models: Web Page and Mobile*

The employment of a salience model can be intimidating. A screenshot of an interface is used to render a corresponding saliency map that provides predictions about where users will fixate. To complicate matters, not all saliency models are created equal and most models do require the employment of MATLAB to render the computation. A seemingly useful web-based implementation has become popular in the last decade; unfortunately, according to J. D. Still et al. (2016), the *saliencytoolbox* implementation does not perform above chance level performance. In our experience, it is best to employ the traditional Itti et al. (1998) implementation; we recommend the simple version provided by Harel et al. (2006). This Itti et al. model produces a saliency map that predicts the relative salience of different regions within an image. It first extracts feature channels -intensity, color, and orientation. Then, the difference of center-surround is calculated for each feature channel and is normalized. These feature maps are combined to form a single saliency map with each pixel value ranging from 0 (unlikely to be salient) to 100 (very likely to be salient). These maps were designed to predict fixations in images of nature scenes, home interiors, urban scenes, and computer-generated fractal images (Parkhurst et al., 2002). The present research focuses on two models that build on the traditional salience model. Other recent examples found in the computer science literature do not fit into this discussion of saliency – they focus on explainability and crowdsourcing data (Cheng et al., 2023; Samuel et al., 2021).

The Attentional Priority model adapts a saliency model (Itti et al., 1998) to web pages by accounting for unique spatial biases. These spatial biases can manifest as conventions, learned from common design patterns, that implicitly guide users' expectations. Conventions are best elicited by observing users' interactions with an interface (Still et al., 2015) as opposed to asking individuals to introspect about them. The convention map for the Attentional Priority model was

generated using Masciocchi and Still's (2013) fixation database. The map reflects a viewing pattern that roughly resembles a combination of top left bias with an overlaid F viewing pattern. Their work suggests that in the final attentional priority map salience should contribute 60% of the weighting while the remaining weighting is assigned to the convention map.

The Mobile Interface Attentional Priority model also adopts the saliency model (J. D. Still & Hicks, 2020) and, like the Attentional Priority model, it includes a convention map. The unique mobile interface convention map resembles a railroad-like pattern. Although mobile displays mostly fall within the fovea and do not require many fixations, surprisingly, saliency still drives attentional selection. According to Still et al. (2017) the saliency model was able to predict the deployment of attention across the first six fixations within a mobile screen. Stills et al. (2020) also demonstrated that target saliency impacts tablet search efficiency. Participants found higher salience targets faster than lower saliency targets on a tablet.

### *Current Models: Strengths, Weaknesses, Future*

In the rare case that a display is uniform, salient elements do appear to pop-out (e.g., a single yellow button among red buttons). In design, it is seldom the case that an interface element is on a uniform background and pops out visually. Instead, we view displays like those in Figure 1 where it would be difficult-to-predict distribution of salience across the display. The highest points of salience are not available through introspection and we cannot simply assume certain characteristics will be salient in all contexts (e.g., red or bold text). Salience is emergent; it is not an independent property associated with an element. The major strength of employing a computational saliency model is to reveal the underlying distribution of salience across a display.

Current salience models are proficient at detecting the most salient regions in a display; thus, predicting early fixations is what the models do best. With this information, designers can at least verify critical elements are, in fact salient. This information can also help designers determine entry points (i.e., where a user will begin their search). However, the ability to predict scan paths—the order in which an individual will search these areas of high salience—within interfaces is still underdeveloped. It is possible that further examination of relative salience could be used to improve these predictions. The models can rank how salient an element is relative to others, perhaps this information can be leveraged.

While salience maps identify specific areas of high salience, they also happen to reveal other areas likely to benefit from attentional resources. For example, J. Still and Still (2019) found that items near an area of high salience benefit from higher search efficiency even though those items are not high salience themselves. With this knowledge designers have a solution in cases where attributes of critical features

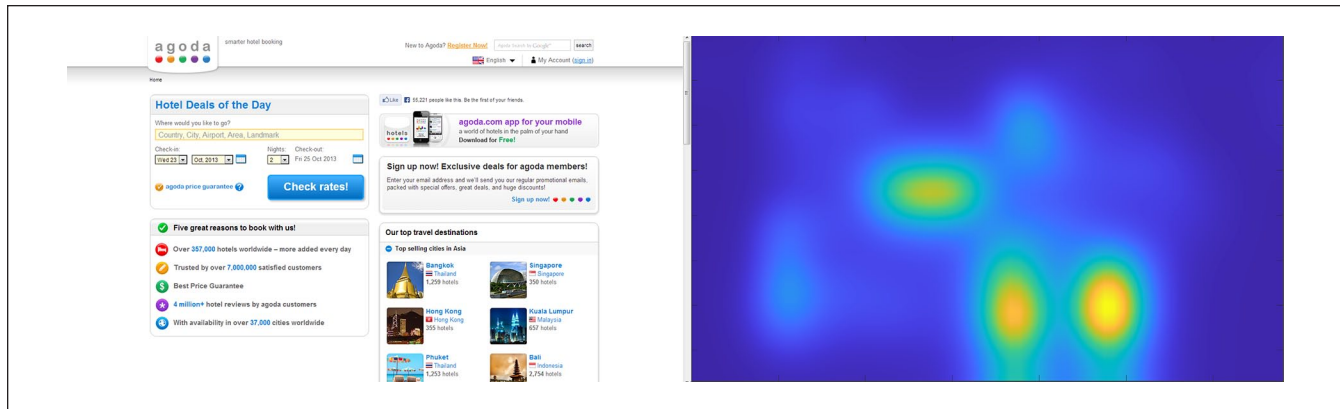


Figure 1. Example of a webpage interface and its associated saliency map.

cannot be changed; they can manipulate the salience of a neighboring item to “drive traffic” to the critical element.

Even with these benefits, a critical limitation of current salience models is that they are context-bound. There is no absolute scale of salience. We lack norming data to understand just what those relative differences mean in terms of search efficiency and noticeability. The more designers employ salience as a measurement, the greater our understanding will grow. Along similar lines, there is the question of when to use a salience model in the design process. It cannot be employed during low-fidelity prototype deployment because there simply is not enough visual data to make accurate computations. Future research is needed to reveal the best practices for employment during the design process. While it is possible that it could be used during mid-fidelity, it has only been assessed for high-fidelity products.

Once salience data have been obtained, it can be difficult to know what to do next. The model is only descriptive, it is not prescriptive. That is, the model output will not help the designer know what specific change would increase the relative salience of an interface element. Therefore, even with the use of a salience model, iteration will be needed to reach design goals. Because of this, future models should provide greater transparency. Designers need to know what feature dimensions to manipulate to change the existing salience distribution. Providing insight into the usage of feature dimensions, such as color, is heavily used, but orientation is also available. These technical insights can empower designer decision-making.

Finally, it is imperative to recognize the significant, practical impact, task information has on search efficiency. Across two experiments, J. Still and Still (2019) asked observers to find a product after only hearing the name or after previewing the exact image. Task information (name or picture) and stimulus-driven salience both contributed to search efficiency. However, providing a picture of the target instead of the name improved search times by 714 msec,

which is a much greater benefit than the 188 msec saved by being near an area of higher visual salience. While additional research is needed to further quantify the relative contributions of task information and salience, this clearly demonstrates the strong contribution of task. From a practical perspective, though, it is worth remembering that to effectively design for task-relevant information you have to know what an individual is looking for. This is not always possible. Salience, by contrast, is a constant influence therefore, designers ought to appreciate the influence of salience and use it to their benefit.

### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

### ORCID iD

Mary L. Still  <https://orcid.org/0000-0002-8511-8280>

### References

- Cheng, S., Fan, J., & Hu, Y. (2023). Visual saliency model based on crowdsourcing eye tracking data and its application in visual design. *Personal and Ubiquitous Computing*, 27, 613–630.
- Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences*, 4, 170–178.
- Cosman, J. D., Lowe, K. A., Zinke, W., Woodman, G. F., & Schall, J. D. (2018). Prefrontal control of visual distraction. *Current Biology*, 28, 1330–1337.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96, 433–458.
- Faraday, P. (2000). *Visually critiquing web pages* [Conference session]. Association for Computing Machinery: Proceedings of the 6th Conference on Human Factors and the Web, Austin, TX.



- Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology Human Perception & Performance*, *18*, 1030–1044.
- Gaspelin, N., Leonard, C. J., & Luck, S. J. (2015). Direct evidence for active suppression of salient-but-irrelevant sensory inputs. *Psychological Science*, *26*, 1740–1750.
- Gaspelin, N., & Luck, S. J. (2018). The role of inhibition in avoiding distraction by salient stimuli. *Trends in Cognitive Sciences*, *22*, 79–92.
- Gaspelin, N., & Luck, S. J. (2019). Inhibition as a potential resolution to the attentional capture debate. *Current Opinion in Psychology*, *29*, 12–18.
- 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). Information Science Reference.
- Harel, J., Koch, C., & Perona, P. (2006). *Graph-based visual saliency* [Conference session]. Association for Computing Machinery: Proceedings of Neural Information Processing Systems, 1–8.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*, 1489–1506.
- Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*, 1254–1259.
- Johnston, W. A., & Dark, V. J. (1986). Selective attention. *Annual Review of Psychology*, *37*, 43–75.
- Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, *43*, 346–354.
- Luck, S. J., Gaspelin, N., Folk, C. L., Remington, R. W., & Theeuwes, J. (2021). Progress toward resolving the attentional capture debate. *Visual Cognition*, *29*, 1–21.
- Malcolm, G. L., & Henderson, J. M. (2010). Combining top-down processes to guide eye movements during real-world scene search. *Journal of Vision*, *10*, 4.1–11.
- 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*, 417–441.
- Navalpakkam, V., & Itti, L. (2005). Modeling the influence of task on attention. *Vision Research*, *45*, 205–231.
- Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, *42*, 107–123.
- Samuel, S. Z. S., Kamakshi, V., Lodhi, N., & Krishnan, N. (2021). Evaluation of saliency-based explainability methods. *ArXiv, abs/2106.12773*.
- Steelman-Allen, K. S., McCarley, J. S., Wickens, C., Sebok, A., & Bzostek, J. (2009). N-SEEV: A computational model of attention and noticing. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *53*, 774–778.
- Still, J. D., & Still, M. L. (2019). Influence of visual salience on webpage product searches. *ACM Transactions on Applied Perception*, *16*, 1–11.
- Still, J. D., Still, M. L., & Grgic, J. (2015). Designing intuitive interactions: Exploring performance and reflection measures. *Interacting with Computers*, *27*, 271–286.
- Still, J. D., Hicks, J., Cain, A. A., & Billman, D. (2017). Predicting stimulus-driven attentional selection within mobile interfaces. *Springer: Proceedings of the 8th International Conference on Cognitive and Neuroergonomics*, (pp. 255–261). Virtual Conference.
- Still, J. D., Hicks, J. M., & Gall, J. (2016, November). *Using computational models to quantify stimulus-driven influences in web pages* [Poster presentation]. 57th Annual Meeting of the Psychonomic Society, Boston, MA.
- Still, J. D. (2018). Web page visual hierarchy: examining Faraday’s guidelines for entry points. *Computers in Human Behavior*, *84*, 352–359.
- Still, J. D., & Dark, V. J. (2010). Examining working memory load and congruency effects on affordances and conventions. *International Journal of Human-Computer Studies*, *68*, 561–571.
- Still, J. D., & Dark, V. J. (2013). Cognitively describing and designing affordances. *Journal of Design Studies*, *34*, 285–301.
- Still, J. D., & Hicks, J. M. (2020). Mobile interface attentional priority model. *Springer Nature: Computer Science*, *1*, 1–11.
- Still, J. D., & Masciocchi, C. M. (2010). *A saliency model predicts fixations in web interfaces*. Association for Computing Machinery: Proceedings of the 5th International Workshop on Model-Driven Development of Advanced User Interactions (pp.25–28). Atlanta, GA.
- Stills, J. D., Hicks, J. M., & Cain, A. A. (2020). Examining the influence of saliency in mobile interface displays. *Journal of AIS Transactions on Human-Computer Interaction*, *12*, 28–44.
- Theeuwes, J. (1993). Visual selective attention: A theoretical analysis. *Acta Psychologica*, *83*, 93–154.
- Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, *135*, 77–99.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136.
- Vatterott, D. B., Mozer, M. C., & Vecera, S. P. (2018). Rejecting salient distractors: Generalization from experience. *Attention, Perception & Psychophysics*, *80*, 485–499.
- Wang, B., & Theeuwes, J. (2020). Saliency determines attentional orienting in visual selection. *Journal of Experimental Psychology Human Perception & Performance*, *46*, 1051–1057.
- Weaver, M. D., van Zoest, W., & Hickey, C. (2017). A temporal dependency account of attentional inhibition in oculomotor control. *NeuroImage*, *147*, 880–894.
- Wickens, C. D., Goh, J., Helleberg, J., Horrey, W. J., & Talleur, D. A. (2003). Attentional models of multi-task pilot performance using advanced display technology. *Human Factors*, *45*, 360–380.
- Wickens, C. D., & McCarley, J. S. (2008). *Applied attention theory*. CRC Press.
- Wolfe, J. M. (2007). Guided Search 4.0: Current progress with a model of visual search. In W. D. Gray (Ed.), *Integrated models of cognitive systems* (pp. 99–119). Oxford University Press.
- Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, *5*, 495–501.
- Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*, Article #0058.
- Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology Human Perception & Performance*, *10*, 601–621.