Effects of Salience Are Short-Lived

Mieke Donk and Wieske van Zoest

Vrije Universiteit Amsterdam

ABSTRACT—A salient event in the visual field tends to attract attention and the eyes. To account for the effects of salience on visual selection, models generally assume that the human visual system continuously holds information concerning the relative salience of objects in the visual field. Here we show that salience in fact drives vision only during the short time interval immediately following the onset of a visual scene. In a saccadic target-selection task, human performance in making an eye movement to the most salient element in a display was accurate when response latencies were short, but was at chance when response latencies were long. In a manual discrimination task, performance in making a judgment of salience was more accurate with brief than with long display durations. These results suggest that salience is represented in the visual system only briefly after a visual image enters the brain.

A salient object tends to attract attention and often initiates a subsequent eye movement. This finding has been replicated often and is well established (Itti & Koch, 2001; Theeuwes, Kramer, Hahn, & Irwin, 1998). It demonstrates that visual selection is very much influenced by the stimulus properties in the visual field. Current computational, functional, and neurophysiological models of visual selection account for stimulus-driven effects by assuming that the brain possesses a salience map, a topographical representation of the relative distinctiveness of all objects in the visual field (Itti & Koch, 2001; Koch & Ullman, 1985; Li, 2002; Treisman & Gelade, 1980; Treisman & Sato, 1990; Wolfe, 1994). The more distinct or conspicuous an object, the greater the corresponding activity in the salience map. The distribution of activity in this map is assumed to determine visual selection such that when there are multiple objects in the visual field, objects are selected in order of decreasing activity. That is, the object with the greatest activity in the salience map is selected first, followed by the next most salient object and so forth (Itti & Koch, 2001).

The notion of a salience map accounts well for the effects of salience in visual selection and can explain why salient objects are selected more often than inconspicuous ones. Salience, however, does not always have an effect. Multiple studies have shown no or only limited influence of salience on visual selection (Folk & Remington, 1998; Folk, Remington, & Johnston, 1992; Ludwig & Gilchrist, 2002; van Zoest, Donk, & Theeuwes, 2004). To explain the finding that visual selection is not exclusively salience driven, theories of visual selection propose that salience-driven activity patterns are subject to modification by voluntary top-down processes. It is believed that salience-driven influences are limited because they may be overruled by the voluntary deployment of attention (Godijn & Theeuwes, 2002; Itti & Koch, 2001; Koch & Ullman, 1985; Trappenberg, Dorris, Munoz, & Klein, 2001; Treisman & Sato, 1990; Wolfe et al., 1989). However, when there is no need for voluntary goal-driven selection to suppress salience-driven activity, stimulus salience is predicted to have a sustained effect on visual selection. In other words, current models of visual selection essentially assume salience to have persistent effects on visual selection unless voluntary goal-driven processes prevent these effects (Desimone & Duncan, 1995; Kastner & Ungerleider, 2001; Reynolds, Chelazzi, & Desimone, 1999; Wolfe, 1994).

The present study tested this key assumption. If salience is continuously represented, it should affect visual selection persistently under conditions in which salience is explicitly task relevant. We performed two experiments in which observers were presented with displays containing two singletons differing in salience relative to the homogeneous background of other elements. In Experiment 1, participants had to make a speeded eye movement to the more salient singleton in the display. We investigated how the proportion of correct eye movements varied as a function of saccade latency. In Experiment 2, display duration was varied, and observers had to indicate which of two simultaneously presented singletons was the more salient.
one. This experiment was performed to investigate how the perception of relative salience varies as a function of time since display onset. Note that there was no reason for goal-driven processes to overrule salience-driven influences in either experiment, as both factors worked in favor of the task requirements. If salience information is represented continuously in the visual system, observers should be able to consistently select the most salient object in the visual field, regardless of the moment in time of their response.

**EXPERIMENT 1**

In Experiment 1, observers viewed displays containing many homogeneously oriented background lines and two singletons, each defined by a different orientation contrast relative to the background lines. Observers were required to make a speeded saccade to the more salient singleton, specified as the one with the larger orientation contrast relative to the background elements. There were two different salience conditions. In the $20'/70'$ condition, the less salient singleton was defined by an orientation contrast of $20'$ relative to the background lines, and the more salient singleton was defined by an orientation contrast of $70'$ relative to the background lines. In the $30'/60'$ condition, the less salient singleton was defined by an orientation contrast of $30'$ relative to the background lines, and the more salient singleton was defined by an orientation contrast of $60'$ relative to the background lines (see Fig. 1). To disentangle the singletons' identities (i.e., specific orientations) from their relative salience, we varied the orientation of the background lines across trials. That is, the background lines were either vertically or horizontally oriented. This manipulation prevented observers from using identity information to infer relative salience and encouraged the use of low-level stimulus-salience information to guide search. Eye movements were registered, and we investigated how the proportion of correct eye movements varied as a function of the time elapsed since the onset of the stimulus display.

**Method**

**Participants**

Twelve paid volunteers (10 women, 2 men; ages 18–36 years, average age = 22 years) participated in Experiment 1. All participants had normal or corrected-to-normal vision.

**Stimuli and Apparatus**

The viewing distance was held constant with a chin rest at 75 cm. The search array consisted of multiple line segments, two of which were oriented differently than the homogenous group of background lines. The background line segments were either all horizontally or all vertically oriented. The line segments were presented in a $17 \times 17$ square matrix with a raster width and height of $17.4' \times 17.4'$ of visual angle. The singletons could be presented at six potential locations, and in each display the circular angle between the two singletons was $180'$. The potential singleton locations were all at the same retinal eccentricity (5.3' of visual angle). All line segments had an approximate length of 0.76' and width of 0.15'. There were two different sets of singletons, yielding two different salience conditions. The orientations of the singletons were $20'$ or $-20'$ and $70'$ or $-70'$ in the $20'/70'$ condition, and $30'$ or $-30'$ and $60'$ or $-60'$ in the $30'/60'$ condition. Thus, each display consisted of multiple homogeneous background lines and two left-tilted singletons (e.g., $-20'$ and $-70'$), two right-tilted singletons (e.g., $20'$ and $70'$), or one left-tilted and one right-tilted singleton (e.g., $-20'$ and $70'$).

---

1It is important to note that salience is a psychological concept. To allow objective quantification of the relative salience of individual singletons, one must define them within a single featural dimension.
Design and Procedure
Participants first viewed a calibration display consisting of nine points in a square array, fixating these points sequentially. The eye-tracking system was calibrated at the start of the experiment and following each block. To start each trial, participants maintained fixation on a central dot (0.25°). They then pressed the space bar in order to apply a drift correction, and the trial began with the presentation of a small fixation point (0.1°) for 500 ms, followed by the search display. The fixation point disappeared the moment the search display was presented. The task was to make a speeded eye movement toward the most salient element upon appearance of the search display. Participants completed one block of 24 practice trials, followed by two blocks of 192 experimental trials. Participants received feedback regarding saccadic latency every 24 trials. The position of the more salient singleton (six possible positions), salience condition (20°/70° or 30°/60°), orientation of the more salient singleton (left- or right-tilted), orientation of the less salient singleton (left- or right-tilted), and orientation of the background lines (horizontal or vertical) were randomly varied within each block of trials.

Analyses
Eye movements were recorded by means of an Eyelink II tracker (SR Research Ltd., Osgoode, Ontario, Canada) with a temporal resolution of 500 Hz and a spatial resolution of 0.2° of visual angle. This system uses an infrared video-based tracking technology to compute the pupil center and pupil size of both eyes. An infrared head-motion tracking system was also used. Saccades were identified by means of a velocity threshold (35°/s) and an acceleration threshold (9500°/s²). The initial saccade was categorized as having landed on a given singleton if its endpoint was within 2.4° of visual angle of that singleton’s position.

All data were subjected to a repeated measures univariate analysis of variance (ANOVA) with saccade-latency quintile and salience condition as within-subject variables.

Results
The proportion of correct eye movements decreased as a function of saccade latency, F(4, 44) = 15.01, p < .001 (see Fig. 2). The proportion of correct eye movements was higher in the 20°/70° condition than in the 30°/60° condition for the shortest-latency eye movements, F(1, 11) = 23.94, p < .001, but not for the longest-latency eye movements, F(1, 11) < 1; this result implies that after some time had elapsed, relative salience differences between conditions did not affect performance. Finally, at the longest saccade latency in both conditions, eye movement performance was at chance level, t(11) = 1.35, p > .1, and t(11) = 1.27, p > .1. Thus, long-latency saccades were equally likely to

![Fig. 2. Proportion of correct eye movements as a function of saccade latency and salience condition in Experiment 1. Error bars indicate standard errors of the means.](image-url)
be directed toward either singleton, irrespective of how much the line segments stood out against the background.2

Discussion
The results of Experiment 1 show that salience condition affected performance only when saccade latencies were short. With long saccade latencies, performance in the 20°/70° condition was equal to that in the 30°/60° condition. More important, eye movements were salience driven only when saccade latencies were short. In contrast, when saccade latencies were long, salience did not affect eye movement behavior. Thus, even though relative salience was explicitly task relevant, observers became progressively less accurate in making an eye movement to the more salient singleton as response latency increased. Indeed, performance decreased to such a dramatic extent that the direction of the eye movements was completely unaffected by relative salience at the longest latency. This suggests that information regarding the relative salience of locations in the visual field is available only transiently. Eventually, the visual system may have information concerning which locations in the visual field contain distinctive features without knowing how distinctive those features are.

To further test the idea that salience may be only transiently represented in the visual system, we conducted Experiment 2. Instead of recording eye movements, we measured the ability of observers to judge the relative salience of two singletons using a manual response. The dependent variable was accuracy in locating the more salient singleton. We investigated the time course of processing by varying presentation duration of the stimulus display and analyzing accuracy as a function of presentation duration.

EXPERIMENT 2
The displays used in Experiment 2 were similar to those used in the 20°/70° condition of Experiment 1. However, this time, one orientation singleton was always presented at the left side of the display, and the other was always presented at the right side, such that the two singletons were at mirror locations relative to the vertical meridian. Participants had to manually indicate on which side of the display the more salient orientation singleton was presented. Display durations were brief and variable, and displays were followed by a mask. A mask was used because it interrupted or terminated further processing of the display, so that we could obtain snapshots of the salience representation at different moments in time (Breitmeyer & Ogmen, 2006).

Method
Participants
Eight paid volunteers (6 women, 2 men; ages 16–34 years, average age = 20 years) participated in Experiment 2. All participants had normal or corrected-to-normal vision.

Stimuli and Apparatus
Experiment 2 was conducted in a dimly lit sound-attenuated chamber. Viewing distance was held constant with a chin rest at 75 cm. The search array used in Experiment 2 was similar to the one in Experiment 1 except for the size of the matrix and the positions of the two singletons. The line segments were presented in a 23 × 17 matrix measuring 23.8° × 17.4° of visual angle. One singleton was always presented at the left side of the screen (at a random position within columns 4–7 and rows 4–14), and the other singleton was always presented at the right side of the screen (within columns 17–20 and rows 4–14). The two singletons were always at mirror positions relative to the vertical meridian. The orientations of the singletons and background lines were the same as in the 20°/70° condition of Experiment 1. Search displays were presented for a variable duration: 25, 42, 83, or 158 ms. Masking displays contained a 23 × 17 matrix of masks, each one consisting of multiple superimposed line segments.

Design and Procedure
In Experiment 2, the task was to indicate the location of the most salient element in the display. Each trial began with the presentation of a central fixation point for 1,000 ms, followed by the presentation of the search display for a variable time interval. Finally, the mask display was presented until the participant responded. Observers pressed the “z” key of a computer keyboard to indicate that the singleton at the left side of the display was the more salient one, and they pressed the “m” key to indicate that the singleton at the right side of the display was the more salient one. Participants completed one block of 16 practice trials and one block of 512 experimental trials. Presentation duration (25, 42, 83, or 158 ms), the orientation of the more salient singleton (left- or right-tilted), the orientation of the less salient singleton (left- or right-tilted), the orientation of the background lines (horizontal or vertical), and the display side containing the more salient singleton (left or right) were randomly varied within each block of trials.

2We found a strong performance drop with increasing saccade latency in both salience conditions. Previous studies have demonstrated that performance does not necessarily decrease with increasing saccade latency (e.g., van Zost & Donk, 2005; van Zost et al., 2004). Nevertheless, it is possible that the negative relationship between accuracy and saccade latency in the present study was not related to short-lived effects of salience, but reflected an overall tendency of saccadic responses to become less accurate as saccade latency increased. To investigate this possibility, we ran a control experiment in which 8 observers (4 women, 4 men; ages 19–30 years, average age = 23 years) were centrally cued to make a speeded eye movement to one of two simultaneously presented, equally salient singletons. The experiment consisted of 24 practice trials and 238 experimental trials. The displays were essentially the same as those used in Experiment 1 except that a central cue consisting of a white arrow was superimposed on the display, and the simultaneously presented singletons were equally salient, with an orientation contrast of 20°, 30°, 60°, or 70° relative to the vertical or horizontal background lines. The overall proportion of correct eye movements increased with saccade latency, F(3, 21) = 11.66, p < .001. Performance in the last quintile was perfect for each of the observers for each of the four contrast conditions.
Analyses
All data were subjected to a repeated measures univariate ANOVA with presentation duration as a within-subjects variable.

Results and Discussion
Observers were less able to indicate the location of the most salient element as presentation duration increased, $F(3, 21) = 8.01, p < .001$ (see Fig. 3). Thus, people were better able to indicate which of the two singletons was more salient when presentation durations were brief than when they were long. These results suggest that salience information was only transiently available and for the most part disappeared during the longer presentations. If salience information had been persistently represented, observers should have been able to indicate the most salient object regardless of how long the stimulus display was presented. Indeed, if anything, the task should only have become easier with longer viewing times, and accuracy should have increased with increasing presentation duration. However, the results suggest that, as in Experiment 1, additional time did not benefit performance in any way.

Previously, we noted that there was no reason for goal-driven processes to overrule salience-driven influences in either experiment, as both factors were supposed to work in favor of the task requirements. However, some researchers have argued that rated salience may be different from low-level measurements of visual distinctiveness (Wright, 2005). For instance, explicit judgments of salience might be based on high-level object or identity properties, rather than low-level properties like luminance or orientation contrast. Possibly, observers judged the singletons used in the present experiments as equally salient despite the physical difference in orientation contrast. If so, the time course of the effects observed in our experiments might reflect not a transient representation of salience, but merely a shift from using bottom-up salience to using top-down (i.e., judged) salience.

To investigate whether or not explicit judged salience corresponded to physical salience, we obtained salience ratings for both the less and the more salient singletons in Experiment 2. Displays were presented until response, with a minimum duration of 4 s, so that observers had ample opportunity to extract information concerning the precise orientations of the individual singletons and background lines. The rating scale ranged from 1 to 5, with 1 corresponding to very inconspicuous, 2 corresponding to inconspicuous, 3 corresponding to neutral, 4 corresponding to conspicuous, and 5 corresponding to very conspicuous. Eight observers (4 women, 4 men; ages 18–31 years, average age = 22 years) were presented with displays similar to those used in Experiment 2 except that each display contained only one singleton (which had an orientation of $20^\circ$, $-20^\circ$, $70^\circ$, or $-70^\circ$) in a background of horizontal or vertical lines. Each observer completed 16 practice trials and 48 experimental trials. All 8 observers rated the physically more salient singleton to be more salient than the physically less salient singleton.

![Fig. 3. Proportion correct as a function of presentation duration in Experiment 2. Error bars indicate standard errors of the means.](image-url)
(sign test, $p = 0.013$; mode of median values for the less salient singleton $= 2$, mode of median values for the more salient singleton $= 5$). These results show that observers judge the more salient singleton to be more salient than the less salient singleton under unlimited viewing conditions. Therefore, the time course observed in our experiments is unlikely to have been due to a shift from using bottom-up salience to using top-down (i.e., judged) salience.

**GENERAL DISCUSSION**

The results of the present study suggest that salience is not persistently represented in the visual system. The brain does not seem to continuously hold information concerning the relative salience of different objects in the visual field. Instead, it appears that after visual input enters the brain, relative salience is represented for a limited time interval. Once this time interval has passed, the visual system no longer holds information concerning the relative salience of objects, but carries information concerning object presence only.

The transience of salience may result from the time-varying nature of neuronal responses within the salience map itself. Indeed, relative salience might be encoded in terms of the order in which neurons within the salience representation start to fire (Thorpe, 1990; VanRullen, Guyonneau, & Thorpe, 2005). For instance, it might be that when an observer is presented with two different singletons, the more salient singleton causes corresponding neurons in the salience map to fire at an earlier point in time than the less salient singleton does. The visual system may be able to discriminate between the two singletons as long as there is differential activity. Once the neurons corresponding to the less salient singleton also fire, the singletons may become indistinguishable in terms of relative salience. From this point on, information regarding relative salience would be lost, and the visual system would hold only information concerning the presence of singletons.

Various studies have demonstrated that the onset latencies of individual neurons are unrelated to their firing rates, which emerge at later points in time. For instance, in one study, temporal-slice analyses revealed that there was no relationship between the onset latencies of individual neurons in monkey striate cortex (V1) and the firing rates of those neurons beyond the first 100 ms after a cell started to fire (Cerletini, Thorpe, Trotter, & Imbert, 1993). In other words, after some time had elapsed, the neurons’ activity no longer reflected their onset latency. In another study, monkeys viewed oriented stimuli that varied in contrast while the responses of 37 striate cortical complex cells were recorded (Gawne, Kjaer, & Richmond, 1996). The results showed that response latency was a function of stimulus contrast, but response rate was not. Together, these findings suggest that information concerning the onset of activity in individual neurons may be lost as time progresses.

The present results strongly suggest that information on salience is lost within a few hundred milliseconds after the onset of a visual display. However, several studies have demonstrated that eye movement behavior may be affected by salience well beyond the first seconds after stimulus onset (Itti, 2005; Kayser, Nielsen, & Logothetis, 2006; Parkhurst, Law, & Niebur, 2002). For instance, Parkhurst et al. (2002) presented 4 observers with natural and artificial images. The observers were instructed to freely view each image for 5 s, while eye movements were recorded. The results demonstrated that even though salience primarily affected the initial eye movements, it also affected later eye movements, even those at the end of the trial. These results appear to be inconsistent with the idea that salience is short-lived. However, it should be noted that the present study revealed salience to be transiently represented within one fixation only. Possibly, under free-viewing conditions, the salience representation dissipates during each fixation, but is reinstated each time an eye movement is made. If this is the case, eye movements that are rapidly emitted after the beginning of a fixation would be expected to be salience driven, whereas those that are initiated after longer fixation durations would be expected to be goal driven, irrespective of whether these eye movements occur after the first or the last fixation on a scene. Note that this hypothesis is completely consistent with the results of the present study, although our results do not allow us to reach a meaningful conclusion regarding its validity. Further studies may be able to provide insight into how the transience of salience affects visual search in the course of a sequence of eye movements.

Visual selection is generally perceived as being the outcome of an interplay between salience-driven and goal-driven processes. Goal-driven processes allow visual selection to be in line with the goals and intentions of an observer. Substantial evidence supports the idea that goal-driven processes need more time to develop than do salience-driven processes (van Zoest & Donk, 2006). The results of the present study suggest that salience is represented only transiently in the human brain. Given the very different time courses of stimulus-driven and goal-driven processes, their interaction in visual selection may be much smaller than is generally assumed.

**Acknowledgments**—We thank Marvin Chun, Daniel Simons, and an anonymous reviewer for their helpful comments on an earlier version of this article.

**REFERENCES**


(Received 6/25/07; Revision accepted 12/23/07)
This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.