Do new objects capture attention?

Steven L. Franconeri, Andrew Hollingworth, & Daniel J. Simons
Harvard University    University of Iowa                 University of Illinois

The visual system relies on several heuristics to direct attention to important locations and objects. One of these mechanisms directs attention to sudden changes in the environment. Although a substantial body of research suggests that this capture of attention only occurs for the abrupt appearance of a new perceptual object, more recent evidence shows that some luminance-based transients (e.g., motion and looming) and some types of brightness change also capture attention. These findings show that new objects are not necessary for attention capture. The present study tests whether they are even sufficient. That is, does a new object attract attention because the visual system is sensitive to new objects or because it is sensitive to the transients that new objects create? In two experiments using a visual search task, new objects did not capture attention unless they created a strong local luminance transient.

Reflected light carries too much information for our visual system to process at once. Instead, some locations or objects are selectively prioritized at the expense of others (e.g. Posner, Snyder, & Davidson, 1980). A fundamental problem for the visual system is to decide which locations or objects deserve priority. Often, priority is task- or goal-dependent. For example, drivers might preferentially attend to red objects due to the importance of brake lights, stop lights, and stop signs. Bicyclists are wary of looming objects (pedestrians, trees). People walking home late at night are especially attuned to sudden motion. Although top-down expectations and goals help determine the focus of attention, some visual events seem to attract attention regardless of the current task; they capture attention. For example, when directing attention to a book on a shelf, an object that suddenly appears (say, a cat jumping onto the shelf) will draw attention away from the book, capturing attention. Sudden changes in the world are likely to be important and might never be attended if priority were governed solely by top-down goals.

What types of stimuli capture attention? According to the most prominent theory, the only stimulus that captures attention is the appearance of a new visual object (Yantis, 1993; Yantis & Jonides, 1996). Under this new object hypothesis, the locations of visible objects are indexed upon first viewing a scene (Kahneman, Treisman, & Gibbs, 1992). If a new object appears later, it requires a new index, which triggers a shift of attention to the new object (Yantis & Jonides, 1996). Thus, the mechanism that prioritizes new objects operates over higher-level visual representations, after candidate objects have been segregated in the visual input. According to this account, the visual system prioritizes new objects because new objects provide novel information that is often behaviorally relevant.

According to an alternative view, the transient hypothesis, some types of simple luminance and motion transients capture attention, whether or not they are associated with a new object (Franconeri & Simons, 2003; Jonides & Yantis, 1988; Yantis & Jonides, 1984). For example, while the cat jumping onto the shelf is a new object, its appearance also coincides with a strong motion transient. By this account, it is this motion signal that draws attention, not the fact that the cat is a new object in the scene. Such unique transients are easily found in cluttered scenes, as long as they involve changes in local luminance (Theeuwes, 1995). In contrast, the new object hypothesis maintains that attention capture is driven solely by the presence of a new object and not by the luminance or motion transients associated with an object's appearance; transients alone are not sufficient to capture attention (Hillstrom & Yantis, 1994; Yantis & Hillstrom, 1994). A cat that moves into view should capture attention by virtue of its status as a new object, but...
Franconeri, Hollingworth, & Simons

a previously visible cat that introduces an equivalent motion transient should not capture attention.

The new object hypothesis drew support from studies using the irrelevant feature search task in which observers search for a target letter among distractors. One randomly chosen letter is given a singleton property such as a unique color, form, luminance, or movement. The singleton could also be a "new object", appearing later than the other items in a previously unoccupied location. Observers know that the singleton is no more likely to be the search target than any of the distractor items. If the unique item captures attention, it should be searched first. Consequently, the search slope (the function relating the response time to the number of distractors) should be shallower (or flat) when the singleton is the target of the search than when the singleton is a distractor. If the unique item does not capture attention, then these slopes should be identical. The new object hypothesis gained support from evidence that new objects captured attention (Hillstrom & Yantis, 1994; Jonides & Yantis, 1988), but salient singletons in color, form, luminance, and motion did not (Folk & Annett, 1994; Hillstrom & Yantis, 1994; Jonides & Yantis, 1988; Theeuwes, 1990, Yantis & Hillstrom, 1994).1

Despite this evidence that new objects were necessary to capture attention, recent studies show that some luminance based transients capture attention even when they are associated with a previously visible object. For example, earlier studies found no capture by moving objects (Hillstrom & Yantis, 1994), but recent studies using stronger forms of motion or different methodologies have found capture (Abrams & Christ, 2003; Franconeri & Simons, 2003; Thomas & Luck, 2000). Concurrent changes in luminance contrast and contrast polarity also capture attention even though they do not introduce a new object (Enns et al., 2001). Luminance changes to other objects can also attenuate capture by a new object, suggesting that luminance changes draw attention away from the new object (Miller, 1989; Martin-Emerson & Kramer, 1997).

Early evidence also suggested that new objects were sufficient to capture attention even when the new object appeared without creating a luminance transient (Yantis & Hillstrom, 1994). For example, new letters introduced via texture segregation, onset of local motion, or changes in binocular disparity captured attention even though these new letters were designed to be equiluminant with their background. (Yantis & Hillstrom, 1994). However, these methods did not entirely eliminate luminance transients (Gellatly, Cole, & Blurton, 1999; Theeuwes, 1995), and when the new letters produced only a weak transient, they captured only weakly (Gellatly et al., 1999). Other recent studies more successfully eliminated luminance transients by defining the new object primarily by a difference in color. When properly calibrated, these equiluminant new objects do not attract attention in spatial cueing tasks (Lambert, Wells, & Kean, 2003) or visual search tasks (Theeuwes, 1995).2 Although the failure of equiluminant new objects to capture attention appears inconsistent with the new object hypothesis, the system that detects the appearance of new objects might simply be less sensitive to objects defined only by a difference in color.

Other studies avoided the problem of creating equiluminant stimuli by instead using luminance defined objects that always created a transient when they appeared, but that were old or new depending on their history. For example, subjects responded faster to respond to a 'new' search display that appeared suddenly, relative to an 'old' one that had traveled quickly across the screen (Yantis & Jonides, 1996). However, the motion of the "old" search array produced more distracting luminance transients that likely pulled attention away from the final search display location, thereby slowing responses.

1 Capture by abrupt onsets can be overridden by a strong top-down strategy if the subject is given a 100% valid cue to the target's future location (Yantis & Jonides, 1990) or if the target is defined by an easily located singleton color (Folk, Remington, & Johnston, 1992). Furthermore, in other visual search tasks, color and shape singletons can also affect search performance (Theeuwes, 1992; but see Bacon & Egeth, 1994; Leber & Egeth, 2003).

2 Because of the variation in perceived luminance of different colors between observers, these objects are usually calibrated to be perceptually equiluminant instead of physically equiluminant (see Cavanagh, 1991).
Another study showed that search responses were slower for displays containing a 'new' subjective square (a global shape formed by a subset of search items) than for displays containing an 'old' subjective square (that had been present before the start of the search), suggesting that the new square captured attention (Rauschenberger & Yantis, 2001). Because the new square itself did not create a unique luminance transient, these results could be taken as evidence that new objects are sufficient to capture attention. However, this result is ambiguous. In the 'new' square condition, all search items underwent transients as they shed their initial masks, spreading attention across all items. All search items capture attention, leading to inefficient search. But in the 'old' square condition, only those search items which did not form the subjective square changed. Critically, those search items always included the target, so that in the 'old' square condition, transients drew attention to the subset of items with the target, greatly improving performance.

The present experiments were designed to provide a definitive test of whether new objects are sufficient to capture attention. In each experiment, a luminance-defined new object appeared immediately prior to search, but its appearance was not accompanied by a unique luminance transient. The method is illustrated in Figure 1. First, observers saw a small set of "figure 8" placeholders surrounded by an annulus. Then the annulus quickly shrank over the course of 180 ms, passing in front of the placeholders. In Experiment 1, the annulus completely occluded the placeholders for only 10 ms; in Experiment 2, it never covered them completely. At the moment of maximal occlusion, a subset of each placeholder's contours was removed to form letters, and a new letter was added. The annulus shrank further, revealing the array of letters. Participants searched for a target letter (either an 'H' or a 'U'), and the target was either the new letter or an old letter (i.e., one of the letters that corresponded to a previously visible placeholder).

If new objects capture attention, then the new letter should be given search priority. If luminance transients capture attention, then it should not, because the luminance transient produced by the disocclusion of the new letter was equal to those created by the disocclusion of the old letters. In each experiment, we included a control condition in which the annulus passed behind the objects, allowing observers to see the unique onset transient created by the new letter. If luminance transients capture attention, then the new letter should only capture in this control condition. This design manipulates the signal produced by the new object - it either is or is not accompanied by a luminance transient - while roughly equating all other factors across the two conditions.

Together, Experiments 1 & 2 constitute a liberal test of the new object hypothesis, because the new object appeared over a very brief disruption, or no disruption. In Experiment 1, the array was fully occluded for only 10 ms, which is much shorter than 50-100 ms estimated persistence of old object representations supporting new object capture (Yantis & Gibson, 1994). To eliminate memory requirements entirely, the annulus was narrowed in Experiment 2 so that it never fully occluded the array.

Experiment 1

In Experiment 1, a new letter was added while an annulus passed in front of (occlusion condition) or behind (control condition) the object array. If the appearance of a new object captures attention, then the new object should receive search priority in both conditions. If unique transients capture attention, then the new object should only receive priority in the control condition, when its luminance increment is visible to the subject.

Methods

32 University of Iowa undergraduates voluntarily participated in exchange for course credit or pay. They reported normal or corrected-to-normal vision. All stimuli were displayed on a 17-in video monitor operating at 100 Hz. Responses were collected by a serial button box. The experiment was controlled by a Pentium-based computer running E-Prime software. Viewing position was maintained at 80 cm by a forehead rest. Stimuli for all experiments are shown in Figure 1. The background was gray (27.2 cd/m²), and the letters and fixation cross were black (< .01 cd/m²). Letter placeholders were block figure 8’s (1.43° wide by 1.85° high) consisting of 7 line segments (0.17° thick), such that any letter in the search array (H, U, S, P, or E) could be made by subtracting
segments. There was always one target, either an H or a U in the display, and distractors were sampled without replacement from the remaining letters. Three placeholders were arranged at the vertices of an upward or downward pointing equilateral triangle, 4.3° from a central fixation point. New letters appeared in between two of these potential placeholder positions at the same distance from the fixation point. For searches with 2 or 3 letters, either 2 or 1 of the original placeholders was removed during occlusion.

The annulus was 2.28° wide and was the same shade of gray as the background. Its inner and outer edges were marked by a 2-pixel darker gray (24.6 cd/m²) outline. At the start of each trial, the annulus encircled the letters, such that its inner edge was farther from the center (6°) than the outer edges of the placeholders. The annulus contracted, passing in front of or behind the letters until its outer edge was 2.28° from fixation.

Observers searched for a target letter (always U or H) and pressed the corresponding button on the button box as quickly and accurately as possible. Incorrect responses resulted in an error message and a short delay. After a 1500 ms delay between trials, a fixation point appeared for 1200 ms, followed by an array of 3 placeholders. After 1000 ms, the annulus quickly contracted for 180ms over 18 frames of animation, coming to rest inside the letter array. In the occlusion condition, the annulus passed in front of the letters. The letters were fully occluded for just 1 frame (10 ms), and during this instant of complete occlusion, the placeholders were replaced by letters, and a new letter was
added. Consequently, any unique onset transient produced by the new letter was not visible. Response timing began when the new object appeared, even if it was behind the annulus. In the control condition, the annulus passed behind the letters, so that the onset transient made by the new letter was visible. The experiment was a 2 (occlusion/control) X 2 (target new object, old object) x 3 (set size 2, 3, or 4) design.

In each condition, subjects understood that new letters were only the targets on 1/n trials (where n is the number of letters in the search array), so that searching the new letter with priority would not increase search efficiency. Each block contained 20 trials of set size 2, 30 of set size 3, and 40 of set size 4. Subjects received two blocks of each condition (occlusion/control), in interleaved order, with the starting condition counterbalanced between subjects. Blocks began with 3 buffer trials (unused in analyses), which were randomly selected from the current block. Participants completed a total of 360 experimental trials. The experiment lasted approximately 45 minutes.

Results & Discussion

Trials with response times greater than 3s or less than 200ms were removed from the analysis (7 total trials). Mean error rate was 1.4%, and it was not higher than 2.4% in any condition. Accuracy data for the occlusion and control conditions were each submitted to a 2x3 ANOVA with new letter type (target, distractor) and set size as factors. In the occlusion condition, observers were slightly more accurate on set size 2 trials than on set size 3 or 4 trials (F(2, 62)=2.6, p=.08), and in the control condition observers were more accurate when the new letter was a target that when it was a distractor (F(1,31)=13.7, p < .001). The interaction between new letter type and set size was not statistically significant in either condition (both F(2,62)<1.6, p>.2), ruling out any potential speed-accuracy tradeoffs.

Mean response times for both conditions are shown in Figure 2. If the new letter captures attention, then the search rate should be lower when the new letter is the target. In the occlusion condition, when the new letter appeared behind the annulus, search rates were not reliably different when the new letter was the target (31 ms/item) compared to when it was a distractor (21 ms/item) (F(1,31)=1.4, p=.24, η=.04). Overall, response times increased with set size (F(2,62)=26.3, p<.001, η=.46), but responses were no faster when the new letter was the target than when it was a distractor (F< 1).

In the control condition, where the annulus passed behind the letters, the new letter captured attention. Search rates were lower when the new letter was the target (12 ms/item) than when the new letter was a distractor (30 ms/item) (F(1,31)=4.8, p=.036, η=.14). There was an overall increase in response times as the set size increased (F(2,62)=24.5, p<.001, η=.44) and faster responses
when the new letter was the target (F(1,31)=111, 
p<.001, \eta=.78).

Consistent with capture by the new letter in the
control condition but not in the occlusion condition,
the slope difference between target types was larger
in the control condition than in the occlusion
condition (F(1,31)=10.8, p=.003, \eta=.26).  These
results are inconsistent with the new object
hypothesis.  While the new object in the control
condition (which created a unique transient),
captured attention, the new object in the occlusion
condition (which did not create a unique transient)
did not.  These results suggest that transients, not
new objects, capture attention.

Perhaps the visual system is sensitive to the
appearance of new objects, but representations of
old objects are volatile.  If so, the extremely brief
(10ms) occlusion in Experiment 1 might have
interrupted representations of the old objects, making
all objects new after the occlusion.  The 10 ms full
occlusion was significantly shorter than estimates of
old object persistence (50-100 ms; Yantis &
Gibson, 1994), so subjects should have been able to
retain the distractor items across the brief
occlusion.3 However, the system that monitors for
the appearance of new objects might not have
access to these more robust forms of representation.
In Experiment 2, we made the annulus narrower, so
that the letters were never fully covered, thereby
reducing the demands on memory.

**Experiment 2**

**Methods**

32 new University of Iowa undergraduates
donated their time in exchange for course
credit or pay.  They reported normal or corrected
to-normal vision.  Viewing distance was controlled
with a chinrest set at 110cm from the monitor.
Letters were 1.88( high and 1.25( wide, with 0.13(

3 To provide further evidence that participants could
indeed keep track of these objects across occlusion, we
conducted a control experiment (n = 18). The method
was the same as Experiment 1, except that on every trial,
the annulus occluded the array and the new object was
the target. Search slopes (9ms/item) were as shallow as
those in the control condition where the transient was
visible (12ms/item), demonstrating that participants
could reliably prioritize the new object in search when it
was perfectly predictive of the target; they were able to
keep track of old objects across occlusion. Although new
objects did not capture attention in the absence of a
transient (occlusion condition of Experiment 1), they can
still be used to guide search. We thank Jan Theeuwes for
suggesting this experiment.
Results & Discussion

Trials with response times greater than 3s or less than 200ms were removed from the analysis (less than 1% of all trials). Mean error rate was 1.4% and was not greater than 2.3% in any condition. Accuracy data for each condition were again submitted to a 2x3 ANOVA with new letter type (target, distractor) and set size as factors. In the control condition, observers were more accurate when the new letter was the target (F(1,31)=6.7, p=.015). As in Experiment 1, new letter type did not interact with set size in either condition (both F(2, 62)<1.4, p>.25), ruling out any speed-accuracy tradeoffs.

The results of Experiment 2 were identical to those of Experiment 1. In the occlusion condition, where the annulus passed in front of the letters, the appearance of the new letter did not capture attention. Slopes were no shallower on trials where the new letter was the target (22 ms/item) than when the new letter was a distractor (16 ms/item) (F(1,31)=1.5, p=.23, η=.05). Mean response times were also no faster when the new letter was a target (F < 1). Response times only increased with set size, (F(2,62)=26, p<.001, η=.46).

In contrast, in the control condition, where the annulus passed behind the letter array, allowing the observer to see the transient created by the new letter, the appearance of the new letter captured attention. When the new letter was a target, search rates were efficient (2 ms/item) relative to when the new letter was a distractor (23 ms/item), (F(1,31)=12.9, p<.001, η=.29). Response times were affected by set size, (F(2,62)=8.1, p<.001, η=.21) and new letter type, (F(1,31)=109, p<.001, η=.78). Consistent with evidence that new letters only captured attention in the control condition, the slope difference between target types was larger in the control than in the occlusion condition (F(1,31)=15, p=.001, η=.33).

Conclusion

Consistent with the transient hypothesis but not the new object hypothesis, we found evidence for capture only when a new object created a unique transient. When a new letter appeared behind an annulus so that the accompanying onset transient could not be seen, the letter was not prioritized in the search. Even though the letter was new according to its history in the display, it did not capture attention, even when the annulus never fully covered the letters. In contrast, when the new object appeared in front of the annulus, so that the accompanying transient was visible, the new letter captured attention.

These results are inconsistent with the hypothesis that new perceptual objects are sufficient to capture attention (Yantis & Jonides, 1996), instead supporting the claim that luminance-based transients are necessary for capture. These transients include certain kinds of brightness changes (Enns et al., 2001), and several kinds of motion (Abrams & Christ, 2003; Franconeri & Simons, 2003). Such changes may capture because they strongly activate transient channels in the visual system, which are maximally sensitive to abrupt onset, luminance flicker, and rapid motion. These channels may play a role in orienting attention to areas of sudden change in the visual field (Breitmeyer & Ganz, 1976; Jonides & Yantis, 1988; Yantis & Jonides, 1984), leading to more detailed analysis by slower, color-sensitive sustained channels (Breitmeyer & Ganz, 1976). Other results suggest that a more sophisticated mechanism governs orienting to sudden changes. For example, persisting flicker does not persist in capturing attention (Nakayama & Mackeben, 1989) and the mechanism underlying these reflexive shifts can learn to isolate only task-relevant areas of a cue (Kristjansson, Mackeben, & Nakayama, 2001). The capture of attention might not even be limited to luminance change. Equiluminant color changes might also capture attention if the color contrast were made strong enough, which would suggest that a broader collection of sensory changes captures attention (Gellatly et al., 1999; Thomas & Luck, 2000)

Although new objects per se may not capture attention, the importance of detecting new objects might still play a role in explaining why some stimuli capture attention. Perhaps only a limited set of luminance changes capture attention because those changes most reliably predict the appearance of a new object (Enns et al., 2001). Detecting a luminance change is computationally simpler than detecting a new object, and unique luminance changes can be immediately found in visual search (Theeuwes, 1995). This heuristic could be one of many used by the visual system to construct high-
level "object" representations for only a subset of items in the visual field.\textsuperscript{4} However, speculation about this role for new objects relates to the function of capture by luminance transients. The present experiments suggest that the actual cues that capture our attention are luminance-based visual transients, not new objects.

\textbf{References}


\textsuperscript{4} Even if new objects do not capture attention, the visual system might still maintain a representation of each object in a scene for other purposes (see Wolfe & Bennett (1997) for evidence that objects are preattentively segmented). However, results from attention capture provide no direct evidence for the existence of such representations.


