Journal of Experimental Psychology:
Human Perception and Performance
2010, Vol. 36, No. 5, 1145–1152

0096-1523/10/$12.00 DOI: 10.1037/a0019960

On the Speed of Pop-Out in Feature Search

Massimo Turatto and Matteo Valsecchi
University of Trento

Adriane E. Seiffert
Vanderbilt University

Alfonso Caramazza
University of Trento, and Harvard University

When something unique is present in a scene, this element may become immediately visible and one has the impression that it pops out from the scene. This phenomenon, known as *pop-out* in the visual search literature, is thought to produce the fastest search possible, and response times for the detection of the pop-out target do not vary as a function of the number of nontargets. In this study, we challenge this notion and show that the detection of a given visual feature is faster for multiple targets than for a single pop-out target. However, when the task requires a detailed target analysis, the pop-out condition can be faster than the multiple-target condition. Current models of visual search are discussed in light of the findings.

Keywords: attention, pop-out, visual search

Introspection suggests that when one is looking for something in the visual field, the easiest search occurs when the target one is searching for has a distinct characteristic that differentiates it from the other elements. For instance, people immediately notice the presence of a yellow flower in a bunch of blue flowers. In accordance with what introspection seems to suggest, one of the most uncontroversial notions in the literature on visual attention is that searching for a singleton element (e.g., the red item among green items) is a very fast and efficient (i.e., independent from the number of background distractors) visual search task (Wolfe, 1998). In such conditions, the observer experiences a strong involuntary awareness of the odd target, which literally pops out from the scene. Since the seminal work of Treisman and Gelade (1980), several studies have provided empirical evidence that corroborates this impression by showing that response times (RTs) to detect the target singleton are usually very fast and, crucially, do not increase as the number of nontarget items (set size) in the scene increases. Terms like *pop-out search* or *pop-out target* are used to describe the fact that the RT × Set Size function is usually flat and the compelling phenomenological impression that the target is immediately visible (Nakayama & Joseph, 1998).

According to one class of visual search models, the visual input is processed by simple feature analyzers, which are spatial maps of the visual field arranged in parallel arrays, each one representing a specific basic feature (e.g., color, orientation, contrast). Examples of this class of models are feature integration theory (FIT; Treisman & Gelade, 1980) and guided search (GS; Cave & Wolfe, 1990), in which the pop-out effect is explained by assuming that the presence of a salient item generates in a bottom-up fashion a peak of activation in a feature map. The visual system would then use this peak either to directly emit the detection response, bypassing attention (Treisman & Gelade, 1980), or to allocate attention to the salient location for a fast identification before responding (Cave & Wolfe, 1990; Wolfe, 1994, 2006). Crucially, an item's salience is determined at the bottom-up feature maps level via a mechanism of mutual inhibition for items sharing the same feature (Cave, 1999; Treisman, 1988). GS postulates, in addition, the presence of top-down feature maps in which the activity for each location represents the extent to which any given item matches the feature or features defining the target. Activities in bottom-up and top-down feature maps are then combined in an overall activation map, and the flat RT × Set Size function of pop-out search is explained by assuming that only the highest peak in that map is immediately selected by visual attention (Cave & Wolfe, 1990; Wolfe, 1994).

A different class of visual search models has been proposed by Nakayama (1990) and by Duncan and Humphreys (1989; see also the SEarch via Recursive Rejection (SERR) model of Humphreys & Müller, 1993). A common property of these models is that they do not posit a clear distinction between preattentive and attentive processes in visual search. Instead, they assume that early visual processing is functionally organized into a multilevel, multifeature pyramidal system, in which the visual field is represented at different levels of resolution. At the lowest level, the image is described with a high degree of resolution, in which each visual unit represents a small portion of the visual field. As one moves toward the vertex of the pyramid, the visual field is represented by

---

Massimo Turatto, Center for Mind/Brain Sciences, and Department of Cognitive Sciences and Education, University of Trento; Matteo Valsecchi, Department of Cognitive Sciences and Education, University of Trento; Adriane E. Seiffert, Department of Psychology, Vanderbilt University; and Alfonso Caramazza, Center for Mind/Brain Sciences, University of Trento, and Department of Psychology, Harvard University.

The research reported here was supported by a Progetti di Ricerca di Interesse Nazionale grant from the Ministero dell'Università e della Ricerca to Massimo Turatto and in part by the Cassa di Risparmio di Trento e Rovereto. We would like to thank Steve Yantis and Howard Egeth for helpful discussion on various aspects of the research.

Correspondence concerning this article should be addressed to Center for Mind/Brain Sciences, University of Trento, Corso Bettini 31, 38068 Rovereto (TN), Italy. E-mail: massimo.turatto@unitn.it

a smaller number of units, with the highest level corresponding (ideally) to a single unit representing the whole visual field. Attention can operate at any one of these different layers as a function of the task requirements. At the highest level, attention is distributed over the entire visual field, but at the expense of allowing only a coarse analysis of the information (for a similar idea, see the model proposed by Hochstein & Ahissar, 2002). Visual search is accomplished via a pattern-matching operation, in which the template of the target stored in working memory is compared with the information represented in the unit selected by attention at a given level in the pyramid. To explain the flat RT × Set Size function usually observed in singleton detection tasks, these models do not assume that focused attention is shifted to the singleton location, as GS does (Wolfe, 1994, 2006). Rather, attention operates in a distributed mode, performing the pattern-matching task at the coarsest level of representation, which is sufficient to allow a global analysis of the entire visual field (Bravo & Nakayama, 1992; Turatto, Valsecchi, Tamè, & Betta, 2007), leading to what Hochstein and Ahissar (2002) have termed *vision at a glance*.

One important aspect differentiates this class of models from activation-map–based models (e.g., Cave, 1999; Cave & Wolfe, 1990; Koch & Ullman, 1985; Treisman, 1988; Wolfe, 1994, 2006). According to the latter models, the singleton attracts attention and becomes immediately visible because at the activation-map level, similar (nontarget) elements mutually inhibit each other, leaving a single high peak of activation at the singleton location. In contrast, as outlined earlier, Nakayama's (1990) and Duncan and Humphrey's (1989) models do not rely on the assumption of mutual inhibition of similar distractors to explain the independence of RTs from set size when the target is a singleton. Because of this crucial difference, these models do not assume any special advantage of a target singleton over multiple identical targets (because they are not inhibited), allowing the interesting prediction that feature search might be slower for pop-out than for multiple targets.

That observers can respond faster to multiple targets than to a single target is a well-established phenomenon known as the redundant target effect (RTE; J. W. Todd, 1912). Two different models have been proposed to account for the RTE: the race model (Raab, 1962) and the neural coactivation model (Miller, 1982). Although RTE has been mainly shown using one versus two onset singletons, it has also recently been documented in the case of one versus two feature singletons (e.g., Krummenacher, Müller, & Heller, 2002; Turatto, Mazza, Savazzi, & Marzi, 2004). For example, observers were faster at detecting the presence of two rather than one red pop-out disk among a set of green disks (Turatto et al., 2004).

Given the considerations we have discussed, it is worth exploring whether the advantage of multiple targets over a pop-out target could extend to a condition in which the multiple targets form a homogeneous group of items rather than being singletons.

Preliminary Evidence on the Slowness of Pop-Out: Dupuis and Caramazza (1990)

The possibility that pop-out might be slower than previously believed was originally investigated by Dupuis and Caramazza (1990) in a series of experiments in which observers had to detect the presence of at least one target among a set of homogeneous distractors. Specifically, the task consisted of detecting the presence of one or more lines of a specific orientation (e.g., vertical) in one of three different display types (see Figure 1), which are best described in terms of the target-present displays (Figure 1, top row). Displays of the first type were as in traditional feature search tasks, with one pop-out item among distractors (Figure 1A, singleton target). Other displays consisted of a homogeneous set of lines in the target orientation (Figure 1B, all targets). Also investigated were displays in which all but one of the lines were in the target orientation (Figure 1B, all-targets-but-one distractor). Target-absent trials were analogous to these three conditions (Figure 1, bottom row), in that the displays either were homogeneous (Figure 1F) or had one singleton item that was also a distractor (Figure 1D). Note that the presence of an odd item was independent of the presence of the target orientation, so that accurate target detection demanded identification of the line orientation. Merely detecting an odd item in the display was not sufficient.

In the experiments of Dupuis and Caramazza's (1990) study, each trial began with a single line shown at the center of the screen for 1,000 ms to indicate the target orientation for that trial. A 1,000-ms blank period followed, and then the display appeared for 200 ms. Participants performed a go–no go task: They pressed a button only if the target orientation appeared in the display. False alarms were treated as errors and signaled to the participants. The corresponding trials were presented again later in the experiment to provide the same number of correct RTs in each condition for each participant.

The main findings of the first experiment showed that target detection was fastest in the all-target condition, intermediate in the all-target-but-one distractor condition, and, quite surprisingly, slowest in the singleton-target condition. In other words, the latter condition, which is the typical pop-out search, turned out to be the slowest type of search. In addition, when multiple targets were presented, RTs decreased as the number of targets increased, which was accounted for by RTE.

As originally noted by Dupuis and Caramazza (1990), in the first experiment most of the target-present trials contained multiple targets, so participants may have adopted the strategy of attending

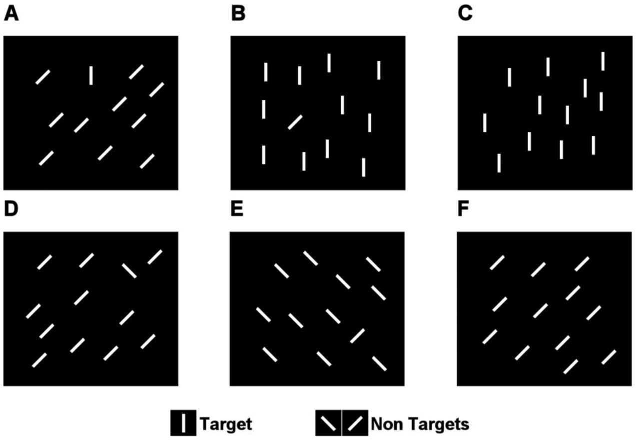

Figure 1. Example of stimuli used in Dupuis and Caramazza's (1990) study. A: singleton target; B: all-target-but-one distractor; C: all targets; D–F: target-absent trials for these conditions.

to the homogeneous group and not to the singleton. To control for strategy effects, in a subsequent experiment they separately tested each of the three display conditions in three different groups of participants. The results were virtually identical to those of the first experiment, suggesting that the participants' strategies, if any, did not contribute to the results. In the singleton-target condition of the control experiment, participants could reliably expect that the singleton item would be the target and therefore were likely to be biased toward the singleton item. However, the singleton-target group still produced slower RTs than the other groups, who viewed multiple-target displays and should not have used the same strategy.

In this study, we aimed to replicate and extend these intriguing findings, pitting pop-out against multiple-target detection. In addition, we addressed how the pop-out phenomenon operates as a function of the attention mode (focused vs. distributed) required by the task. To this end, instead of using stimuli consisting of differently oriented lines, we used colored rings containing letters, which allowed better experimental control of the factors manipulated in the study.

Experiment 1

Dupuis and Caramazza's (1990) findings suggested that in feature search, the presence of multiple targets leads to faster RTs than does the presence of a pop-out target. However, a potential methodological flaw might have affected their results. The stimuli were randomly scattered on the display to maximize participants' uncertainty about the target locations, which may have generated a distance-from-fixation-point search bias in which RTs for target detection tended to increase as a function of target eccentricity (Carrasco, Evert, Chang & Katz, 1995; Carrasco & Katz, 1992). Most crucially, because on average the nearest stimulus to fixation was more likely to be a redundant target than a singleton target, a center-to-periphery bias in scanning the display might explain why RTs were shorter in the all-target than the one-target condition. To avoid the eccentricity problem, in this and the next experiment we arranged the stimuli in a circular array so that they were all at the same distance from the fixation point.

Method

Participants. Ten students or staff members from the University of Trento (Rovereto, Italy; mean age = 29.2, 6 women) volunteered for participation in Experiment 3. All of the participants were naïve as to the purpose of the study and reported normal or corrected-to-normal vision. Informed consent was obtained from all participants, and the experiment was carried out in accordance with the Declaration of Helsinki.

Stimuli. Stimuli (see Figure 2) were colored rings (outer radius = 0.58°; inner radius = 0.28°) containing letters. Colors were red, green, or blue, and the stimuli appeared over a black background (0.5 cd/m^2). On each trial, one of the three possible colors was randomly chosen as the target color, and the remaining two served as distractors. The luminance of the green and blue stimuli (about 3 cd/m^2) was matched to that of red stimuli using a 21.5-Hz flicker-fusion procedure (Wyszecki & Stiles, 1982). Each ring contained a low-luminance capitalized gray letter (height = 0.43°), which could be either a vowel (*A, E, I, O,* or *U*) or a consonant (*B, C, H, N,* or *Z*). On each trial, one of the three possible colors was randomly chosen as the target and one of the remaining two was associated with the distractors. All the rings of the same color also contained the same type of letters (either vowel or consonant), but the specific vowel or consonant associated with each ring was chosen randomly. Stimuli were displayed on a 19-in. (48.3 cm) Iiyama CRT monitor controlled by a Radeon 9550 graphics card. Color depth was set at 32 bits, and screen resolution was 1,024 × 768 pixels. The vertical frame rate of the monitor was 85 Hz. Stimuli were generated using MATLAB and the PsychToolbox (Brainard, 1997).

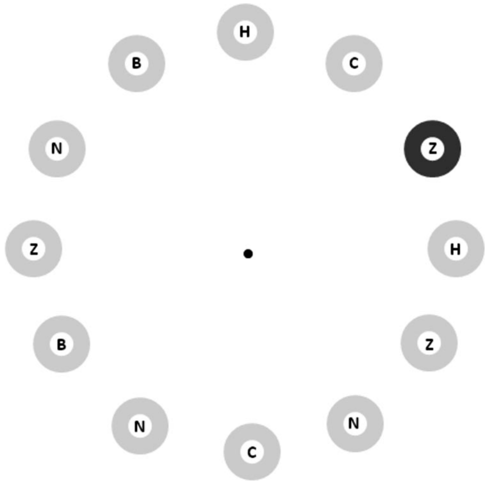

Figure 2. Example of stimuli used in both experiments. The example depicts either the one-target or one-distractor condition. In the all-target and all-distractor conditions all rings had the same color. Stimuli are not drawn to scale; differences in color are represented as differences in gray level.

Procedure. Participants sat approximately 60 cm in front of the monitor in a dimly illuminated room. At the beginning of each trial, a fixation point (0.36°) was presented at the center of the screen for 1,000 ms. Its color indicated to the participants the color of the target or targets. Then, three, six, or 12 rings, each containing a letter, appeared on the screen until a response was made or 2,000 ms elapsed. The compound stimuli (rings + letters) were arranged on an imaginary circle (9°) around the fixation point. The angular distance between the stimuli was constant for each set size (120°, 60°, and 30° for Set Sizes 3, 6, and 12, respectively). Participants had to respond by pressing the spacebar as quickly as possible if at least one of the rings had the target color, while ignoring the letters inside the rings. Participants had to refrain from responding to catch trials. Eye movements were not monitored.

Errors, namely responses on catch trials, misses, or response anticipations (RTs $<$ 100 ms) were signaled via auditory feedback. In this and the following experiment, error trials were reinserted in a random position later in the series. The intertrial interval varied randomly between 1,000 and 2,000 ms.

Design. A 3 × 3 factorial design was used, with condition and set size as factors. Condition had three levels: all targets (all rings had the target color), one target (one ring had the target color and the remaining rings had a different color), and one distractor (all but one ring had the target color). In addition, the all-distractor (all rings had the same nontarget color) and singleton-distractor (all but one ring had the same color, but none of them had the target color) conditions, in which participants had to refrain from responding, were used as catch trials. Set size had three levels: three, six, or 12 rings.

The all-target and all-same-distractor conditions were run in separate sessions. The number of trials per condition was 90, except for the all-distractor condition, which had 45 trials. Catch trial frequency was one in three in both conditions.

Results

False alarms (i.e., responses on catch trials) were very infrequent (1.9%) and were not further analyzed. In this and the following experiment, trials on which RTs were shorter than 150 ms or longer than 2 standard deviations from the mean were treated as outliers and were eliminated before formal data analysis. In this experiment, the outlier-latency criterion removed 4.8% of the data. Correct RTs were entered in a two-way repeated-measures analysis of variance (ANOVA) with condition (all targets, one target, or one distractor) and set size (three, six, or 12) as factors. The effects of condition, $F(2, 18) = 10.365$, $MSE = 2{,}1408.0$, $\eta^2 = .535$, $p < .001$, and the Set Size × Condition interaction, $F(4, 36) = 4.281$, $MSE = 937.1$, $\eta^2 = .322$, $p < .006$, were significant, and the effect of set size ($F < 1$) was not.

Pairwise comparisons (*t* tests) showed that RTs in the all-target condition did not differ significantly from those in the one-distractor condition ($p = .57$), and RTs in the one-distractor condition were faster than those in the one-target condition ($p < .001$; see Figure 3).

The slopes of the RT × Set Size functions were as follows: –1.77 ± 0.4 ms/item in the all-target condition, −1.23 ± 0.6 ms/item in the one-distractor condition, and 1.7 ± 0.5 ms/item in the one-target condition. All but the one-distractor slope were significantly different from 0 (all *p*s < .01).

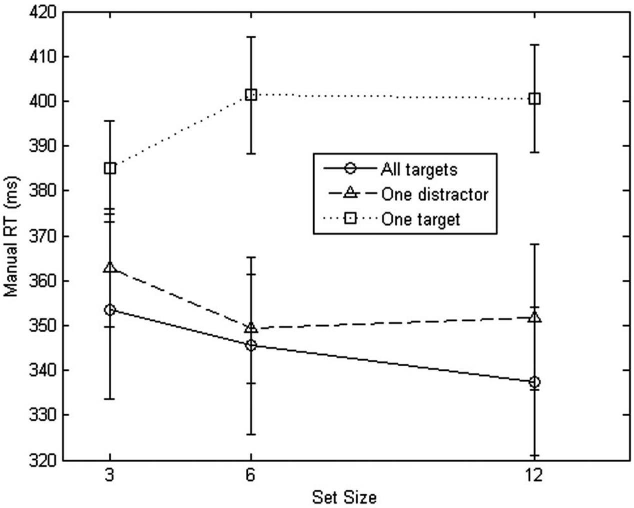

Figure 3. Average manual response times (RTs) in Experiment 1 as a function of set size, separately for the three conditions.

Discussion

The results of Experiment 1 basically confirmed those obtained by Dupuis and Caramazza (1990) with different stimuli and arrangement, showing that during feature search target detection is faster with multiple targets than with a singleton pop-out target. In Experiment 1, the RT × Set Size function in the one-target condition attested to the presence of the pop-out effect because the slope of the function (1.7 ms/item) is compatible with that usually reported for pop-out search (Treisman & Gelade, 1980; Wolfe, 1998). This compatibility is explained by several attention models by assuming that the singleton is detected preattentively or, alternatively, that it immediately attracts attention by virtue of a combined bottom-up (local contrast detector) and top-down mechanism, which should lead to extremely fast detection (Cave & Wolfe, 1990; Treisman 1988; Treisman & Gelade, 1980; Wolfe, 1994, 2006). Yet, pop-out was associated with the slowest responses, with participants being much faster (on average, 50 ms) when multiple targets were present (all-target condition).

Second, feature detection with multiple targets remains faster than with a pop-out target even in the presence of a salient distractor (see Figure 3). On the basis of previous findings by Dupuis and Caramazza (1990); since replicated in our lab), we expected an RT difference between the all-target and the one-distractor condition, with participants being slower in the latter case. Although for all set sizes RTs were indeed numerically smaller in the one-distractor condition, the RT difference was not significant.

Third, we also found that in the all-target condition, RTs decreased significantly as the number of targets increased, indicating that evidence for target detection accumulated more rapidly when multiple targets were present, a result consistent with RTE (J. W. Todd, 1912).

Turatto et al. (2004) have already shown that target detection is faster when more than one target is present in a search display (also see Krummenacher et al., 2002). They found that observers were faster at detecting two red disks among six green disks than one red disk among seven green disks. In that case, however, the multiple-target condition consisted of two target singletons, which were still experienced by the observers as pop-out singletons among a set of multiple nontarget items, and therefore in those experiments the RTE was not pitted against the pop-out effect. Here, instead, the pop-out condition was the only condition in which a target singleton was present because in the redundant condition the display consisted only of multiple targets (ranging from a minimum of three to a maximum of 12). In this way, we had the possibility of evaluating the independent contribution of pop-out versus redundancy in simple visual search. This allowed us to document, for the first time (together with Dupuis and Caramazza, 1990), the slowness of pop-out compared with RTE.

The current findings can be readily accommodated by the models proposed by Duncan and Humphreys (1989) and Nakayama (1990; also see Hochstein & Ahissar, 2002) because they assume that evidence for responding to the presence of a particular feature can accumulate across the whole display, given that similar items are not inhibited by default. This implies that the present task was performed with attention distributed over the entire display (also see Bravo & Nakayama, 1992; Turatto et al., 2007). However, we do not exclude that a target singleton attracts attention (which probably explains our impression of the target popping out from the display), but this would occur after the singleton has been detected during the phase when attention was "distributed" over the entire display. Using the words of Nakayama and Joseph (1998, p. 291), "Pop out, or the narrowing of attention to the odd target, has no direct causal role in detection of the presence of a target."

By contrast, the observed slowness of pop-out compared with RTE is a bit more problematic for those models, such as FIT and GS, that assume that evidence for target detection is collected more or less serially from single items in the display. That is, information is first obtained from the most conspicuous item present in the scene, and similar items tend to inhibit each other. In other words, the pop-out target should be the first item selected because bottom-up processes isolate the salient location by suppressing those items that share the same feature (e.g., Cave & Wolfe, 1990; Treisman, 1988; Wolfe, 1994). A recent variant of the GS model—the GS4 (Wolfe, 2006), a hybrid (parallel–serial) model in which attention does not invariably select and process one item at a time—may be able to account for the results reported here. Because multiple items can accumulate information concurrently in a single attention episode, it is possible that this would compensate for any disadvantage resulting from mutual inhibition of similar items, thus allowing faster detection of multiple targets compared with a singleton target.

Experiment 2

The results of Experiment 1 showed that when the target is defined by a single feature, say its color, evidence for reporting the target's presence is accumulated faster when multiple instances of the feature are present than when a single salient element possesses that feature. However, a different scenario may emerge if observers are required to attend to and select a specific element in the display to perform the task. In this case, the presence of a pop-out target might speed up RTs because both bottom-up and top-down processes would concur in isolating the relevant item.

To test this prediction, in Experiment 2 we used the same set of stimuli as in Experiment 1, but we changed the task so that it required the selection of a single item with focused attention in all conditions tested. To this aim, the participants' task was to determine as quickly as possible whether the letter inside one of the rings of the target color was a consonant or a vowel. Crucially, different letters of the same category (i.e., consonants or vowels) were used for the target and nontarget color rings, which prevented perceptual grouping among the redundant stimuli other than by color. Furthermore, the letters were of lower luminance than the rings, which discouraged parafoveal viewing of the stimuli. Should the manipulation be successful in forcing participants to focus on a single target element to perform the task, we would expect the one-distractor condition, in which multiple targets are presented, to lead to poorer performance compared with the one-target condition because the bottom-up mechanism becomes prominent in guiding target selection. Therefore, when the salient element selected is the target, the next step consisting of letter identification can start, whereas when the salient element selected is the distractor, letter identification cannot proceed, and a new selection of the target among the homogeneous items needs to be implemented.

Method

Participants. Nineteen students or staff members from the University of Trento (mean age = 27.9, 11 women) volunteered for participation in Experiment 2. All of the participants were naïve as to the purpose of the study and reported normal or corrected-to-normal vision. Informed consent was obtained from all participants, and the experiment was carried out in accordance with the Declaration of Helsinki.

Stimuli. The stimuli were the same as those in Experiment 1.

Procedure. The procedure was as in Experiment 1 with the following exceptions: Participants responded by pressing the left arrow on the keyboard with their right-hand index finger to indicate that the target stimulus or stimuli contained a consonant and by pressing the right arrow with their right-hand middle finger to indicate that it contained a vowel. One-target and one-distractor trials were randomly interleaved in one session, whereas all-target trials were administered in a separate session. The order of sessions was counterbalanced across participants.

Results

The outlier-latency criterion removed 4.9% of the data. Correct RTs were entered in a two-way ANOVA with condition (all targets, one target, or one distractor) and set size (three, six, or 12) as factors. The effect of condition, $F(2, 36) = 12.930$, $MSE = 161{,}921.6$, $\eta^2 = .418$, $p < .001$, and the effect of set size, $F(2, 36) = 7.385$, $MSE = 15{,}582.4$, $\eta^2 = .291$, $p < .002$, were significant, whereas the interaction was not significant ($p = .166$).

Results are depicted in Figure 4. Pairwise comparisons (*t* tests) showed that participants were faster in the all-target condition than in the one-target condition ($p < .023$) and in the one-target condition than in the one-distractor condition ($p < .037$). The slopes of the RT × Set Size functions were as follows: 3.4 ± 1.5 ms/item in the all-target condition ($p < .041$), 5.4 ± 1.8 ms/item in the one-distractor condition ($p < .01$), and 2.0 ± 1.2 ms/item in the one-target condition (*ns*).

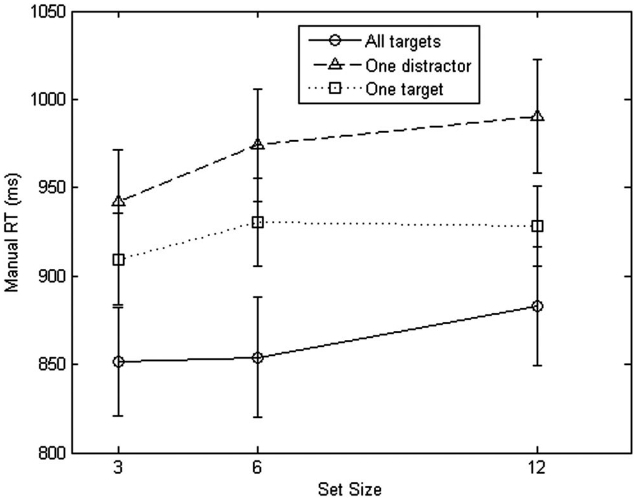

Figure 4. Average manual response times (RTs) in Experiment 2 as a function of set size, separately for the three conditions.

Data from Experiments 1 and 2 were combined in a three-way ANOVA with condition (all targets, one target, or one distractor) and set size (three, six, or 12) as within-participants factors and task (detection or discrimination) as between-participants factors. The effect of condition, $F(2, 54) = 9.266$, $MSE = 83{,}722.661$, $\eta^2 = .255$, $p < .001$; the Task × Condition interaction, $F(2, 54) = 6.197$, $MSE = 55{,}993.246$, $\eta^2 = .187$, $p < .004$; the effect of task, $F(2, 27) = 217.477$, $MSE = 6{,}008{,}257.354$, $\eta^2 = .890$, $p < .001$; and the Task × Set Size interaction, $F(2, 54) = 4.635$, $MSE = 6{,}710.421$, $\eta^2 = .147$, $p < .014$, were all significant.

Discussion

Overall, RTs were much longer in Experiment 2 than in Experiment 1, which is consistent with the fact that participants had to select a given item and discriminate whether the letter it contained was a consonant or a vowel. More important, however, as a comparison between Figures 3 and 4 indicates, the RT pattern in Experiment 2 is strikingly different from that of Experiment 1. Specifically, we found opposite patterns of results for one-target and one-distractor conditions in the two experiments. In Experiment 2, which required participants to focus their attention on a single item, RTs were overall shorter in the one-target condition than in the one-distractor condition. In other words, Experiment 2's results are consistent with the hypothesis that the participants' focus of attention was guided by the singleton, so that RTs for letter identification were shorter when the singleton was the target than when it was a distractor, because in the latter case attention needed to be redeployed to a different item (one of the remaining homogeneous targets). By contrast, when the task involved simple feature detection, as in Experiment 1, which could presumably be accomplished by accumulating evidence for the target-defining feature in parallel from across the entire display, the opposite was true.

When target selection occurred on the basis of top-down mechanisms, as when the display consisted of all targets, results showed that this endogenous selection was negatively affected by multiple targets. Indeed, RTs slowed down as target numerosity increased, probably because the larger the number of targets was, the higher the degree of uncertainty was during item selection. Finally, one may have predicted that target selection might have been faster in the one-target condition than in the all-target condition. However, these two conditions were not directly comparable because the all-target condition was run in a separate session and thus under reduced target uncertainty.

General Discussion

There seems to be something inevitable in the impression of immediately noticing one red apple among lots of green apples or, more generally, a unique element that stands out from a set of similar elements. Although the singleton necessarily owes its status of "uniqueness" to the presence of the same-colored elements, everyone usually has the feeling of seeing the singleton first, and accordingly the singleton is said to pop out.

The perceptual and psychological relevance of the singleton is also attested to by the fact that models of visual search have almost invariably given special prominence to the contrast between cases in which the target is a pop-out element (feature search) versus cases in which the target shares features with the distractors (conjunction search), making it difficult to distinguish from the distractors. Early models treated these two instances as a dichotomy (see FIT; Treisman & Gelade, 1980; Treisman & Souther, 1985), with fast preattentive parallel search in the former case and slow attentive serial search in the latter case. Later models preferred to consider the two types of search as instances in a continuum of search efficiency (see Wolfe, 1998), with pop-out faster than conjunction search because attention is immediately deployed to the singleton element (see GS; Cave & Wolfe, 1990; Wolfe, 1994, 2006). However, virtually all models have addressed simple feature search considering only the case in which the feature defining the target belongs to a single element in the display. Under these conditions, the search has turned out to be extremely fast and independent of the number of distractors (Wolfe, 1998), two results that confirmed the impression that an odd element is indeed immediately visible and detectable.

However, an interesting question often overlooked in previous visual search models is whether the visual system may detect the presence of a given feature faster when it is a singleton, salient element or when it is similar to many elements. Encouraged by the preliminary, unpublished findings of Dupuis and Caramazza (1990), we conducted two experiments to address when a pop-out target is advantageous in feature search.

When the task involved simple feature detection (Experiment 1), we confirmed the results of Dupuis and Caramazza (1990) showing that RTs were faster when multiple targets sharing the same feature were present than when the target was a singleton. In other words, overall Experiment 1 showed that for simple feature detection, pop-out is slower than RTE.[1] Although the idea that pop-out may be a relatively slow phenomenon (it requires approximately 200 ms to produce an efficient guidance of attention) had already been proposed in the literature (e.g., Olds, Cowan, & Jolicoeur, 2000; Olds & Degani, 2003; Palmer, van Wert, Horowitz, & Wolfe, 2006), the present results demonstrate its relative slowness when compared with a condition in which the visual system can accumulate response-relevant evidence from across the whole display.

The results of Experiment 1 seem to be better accounted for by models of visual search like those proposed by Duncan and Humphreys (1989; also see Humphreys & Müller, 1993) and by Nakayama (1990) because according to these models, attention need not be deployed to the singleton to detect its presence. Furthermore, because multiple items sharing the same feature are not inhibited by default, if the feature they share is the one defining the target, evidence could accumulate in parallel across the visual

---

[1] The paradigm we have adopted did not allow us to specify whether the redundancy gain we observed in case of multiple targets was the result of the statistical (Raab, 1962) or neural coactivation model (Miller, 1982). Miller (1982) introduced a method for testing whether the RTE facilitation exceeds the one predicted by a simple race model without coactivation, or whether coactivation needs to be assumed. Unfortunately, this method is only applicable when the RT distribution to a single target presented in isolation is available, which was not the case in the paradigm adopted in these experiments. All we can say is that a redundancy gain is observed when parallel processing of the display is sufficient for responding, whereas a redundancy cost is observed when the task requires focal analysis of the stimuli.

field, leading to faster RTs. In contrast, models such as FIT and GS (also see Cave, 1999) predict the supremacy of pop-out search because of the existence of bottom-up processes that attract attention to the singleton by suppressing all items sharing the same feature. The relevance of the bottom-up component in the pop-out effect was clearly pointed out by Cave and Wolfe (1990) when they claimed that

> the bottom-up component within each feature map does not depend on what target value is expected for that dimension, and will be the same even when nothing is known about what kind of target to expect. We assume that it is the basis for visual pop out. (p. 233)

This is not, however, a fatal problem for this class of model (Cave & Wolfe, 1990; Treisman, 1988; Wolfe, 1994), but they would certainly need to change the relative importance of some of the parameters of the models to accommodate the present findings. An example of such a change is the latest version of GS (GS4; Wolfe, 2006), which is now a hybrid parallel–serial model.

Although we have interpreted the advantage of multiple targets over a single target found in Experiment 1 as being the result of redundancy, one may argue that the same effect could be explained by invoking a sort of priming. Indeed, evidence exists that when the target is repeated (over time), search efficiency is improved, both when the target is defined by a single feature (e.g., Maljkovic & Nakayama, 1994) and when it is defined by two features (e.g., Kristjànsson, Wang, & Nakayama, 2002). In an analogous fashion, one may hypothesize that the presence of multiple elements with the same feature (say, red) acts as a prime for the color red, so that whichever red element is selected, the threshold for detecting its presence is lowered. However, it is not possible to distinguish between these two alternatives on the basis of the present findings because the two hypotheses make the same prediction: namely, faster RTs with multiple targets than with a singleton target.

Interestingly, Experiment 2's results showed that when the observers were to select a specific item to perform the task, pop-out search became faster than RTE. It is reasonable to suppose that in this type of task, in which a potential target needs to be selected, the presence of a salient location could guide attention. Hence, if the salient pop-out item is the target RTs are shorter as compared to when the pop-out item is a distractor, this being a prediction common to all models of visual search. Indeed, in models like FIT and GS, when the singleton is the target its location should receive the strongest activation possible, as both bottom-up and top-down maps converge on the same location. By contrast, when the singleton is a distractor and multiple targets are present, two bottom-up factors would jointly operate in slowing down the target selection. Because of the bottom-up component, the distractor singleton will be signaled as a salient location, while at the same time the multiple target locations will inhibit each other because of their similarity. It follows that deployment of attention to one of the multiple targets should be slowed down compared with when the target is the singleton.

To summarize, the present experiments showed that in a feature search task, namely when the target is defined by a single feature (here, color) and is clearly distinguishable from the distractors, redundant targets dominate pop-out. Models relying on mutual inhibition of items sharing the same feature to explain the pop-out effect have problems accounting for the results of Experiment 1. However, when the task requires the specific selection of one item, bottom-up mechanisms isolating conspicuous locations become dominant for target selection, as predicted by FIT and GS.

Another way of interpreting the present findings is by invoking a difference between distributed and focused attention (Bravo & Nakayama, 1992). In such a framework, the results of Experiment 1 could be interpreted as showing that simple feature detection can take place with attention distributed over the entire display. This attention mode allows parallel pick-up of target information, resulting in faster detection when multiple targets are present instead of a single pop-out target. By contrast, when the search task requires focused attention because of the need for fine-grained analysis, the presence of a pop-out target is advantageous over multiple targets because of the bottom-up guidance of attention by the odd item in the display. This suggests that the visual system can operate in different attention modes, relying on different mechanisms for target detection and identification.

References

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436.

Bravo, M. J., & Nakayama, K. (1992). The role of attention in different visual-search tasks. *Perception & Psychophysics, 51,* 465–472.

Carrasco, M., Evert, D. L., Chang, I., & Katz, S. M. (1995). The eccentricity effect: Target eccentricity affects performance on conjunction searches. *Perception & Psychophysics, 57,* 1241–1261.

Carrasco, M., & Katz, S. M. (1992). Exploring feature asymmetries in visual search: Effects of target position and presentation time. *Proceedings and Abstracts of the Annual Meeting of the Eastern Psychological Association, 63,* 25.

Cave, K. R. (1999). The FeatureGate model of visual selection. *Psychological Research, 62,* 182–194.

Cave, K. R., & Wolfe, J. M. (1990). Modeling the role of parallel processing in visual search. *Cognitive Psychology, 22,* 225–271.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458.

Dupuis, I., & Caramazza, A. (1990). [A novel visual search task reveals the slowness of "pop-out"]. Unpublished raw data.

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron, 36,* 791–804.

Humphreys, G. W., & Müller, H. J. (1993). Search via recursive rejection (SERR): A connectionist model of visual search. *Cognitive Psychology, 25,* 43–110.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology, 4,* 219–227.

Kristjànsson, A., Wang, D., & Nakayama, K. (2002). The role of priming in conjunctive visual search. *Cognition, 85,* 37–52.

Krummenacher, J., Müller, H. J., & Heller, D. (2002). Visual search for dimensionally redundant pop-out targets: Parallel-coactive processing of dimensions is location specific. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1303–1322.

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out. I: Role of features. *Memory & Cognition, 22,* 657–672.

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology, 14,* 247–279.

Nakayama, K. (1990). The iconic bottleneck and the tenuous link between early visual processing and perception. In C. Blakemore (Ed.), *Vision: Coding and efficiency* (pp. 441–422). Cambridge, England: Cambridge University Press.

Nakayama, K., & Joseph, J. S. (1998). Attention, pattern recognition and popout in visual search. In R. Parasuraman (Ed.), *The attentive brain* (pp. 279–298). Cambridge, MA: MIT Press.

Olds, E. S., Cowan, W. B., & Jolicoeur, P. (2000). The time-course of pop-out search. *Vision Research, 40,* 891–912.

Olds, E. S., & Degani, M. D. (2003). Does partial difficult search help difficult search? *Perception & Psychophysics, 65,* 238–253.

Palmer, E. M., van Wert, M. J., Horowitz, T. S., & Wolfe, J. M. (2006). Measuring the timecourse of guidance in visual search. *Journal of Vision, 6,* 443a.

Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Science, 24,* 574–90.

Todd, J. W. (1912). *Reaction to multiple stimuli.* New York, NY: Science Press.

Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology, 40*A, 201–237.

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

Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General, 114,* 285–310.

Turatto, M., Mazza, V., Savazzi, S., & Marzi, C. A. (2004). The role of the magnocellular and parvocellular systems in the redundant target effect. *Experimental Brain Research, 158,* 141–150.

Turatto, M., Valsecchi, M., Tamè, L., & Betta, E. (2007). Microsaccades distinguish between global and local visual processing. *Neuroreport, 18,* 1015–1018.

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review, 1,* 202–238.

Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *Attention* (pp. 13–73). Hove, England: Psychology Press.

Wolfe, J. M. (2006). Guided Search 4.0: Current progress with a model of visual search. In W. D. Gray (Ed.), *Integrated models of cognitive systems.* New York, NY: Oxford University Press.

Wyszecki, G., & Stiles, W. S. (1982). *Color science: Concepts and methods, quantitative data and formulae.* New York, NY: Wiley.

Received August 19, 2008
Revision received February 2, 2010
Accepted February 8, 2010 ■