Search for the Unknown: Guidance of Visual Search in the Absence of an Active Template

Oryah C. Lancry-Dayan¹, Matthias Gamer², and Yoni Pertzov¹
¹Department of Psychology, The Hebrew University of Jerusalem, and ²Department of Psychology, University of Würzburg

Abstract
Can you efficiently look for something even without knowing what it looks like? According to theories of visual search, the answer is no: A template of the search target must be maintained in an active state to guide search for potential locations of the target. Here, we tested the need for an active template by assessing a case in which this template is improbable: the search for a familiar face among unfamiliar ones when the identity of the target face is unknown. Because people are familiar with hundreds of faces, an active guiding template seems unlikely in this case. Nevertheless, participants (35 Israelis and 33 Germans) were able to guide their search as long as extrafoveal processing of the target features was possible. These results challenge current theories of visual search by showing that guidance can rely on long-term memory and extrafoveal processing rather than on an active search template.

Keywords
eye movements, memory, visual search, template, open data, open materials, preregistered

We spend approximately 10 min a day searching for things, which adds up to roughly 153 days of our lifetime (Miller, 2012). Since visual search is such a central task, it comes as no surprise that the neurocognitive research community has been investing substantial effort to understanding the mechanisms that enable it to function efficiently. Emerging from the seminal ideas of feature-integration theory (Treisman & Gelade, 1980) and guided search (Wolfe et al., 1989), current theoretical views are predicated on the need for a guiding template (Wolfe, 2020) for visual search to be efficient. As an active representation of the properties of the search target, the guiding template resides in visual working memory and guides search by being compared with the visual input and directing search to the probable whereabouts of the target. Indeed, studies have shown that gaze is attracted to locations that share the target's low-level characteristics (Alexander et al., 2019; Findlay, 1997; Luria & Strauss, 1975; Motter & Belky, 1998), even if the target is not present (Tavassoli et al., 2009). Search can be efficiently executed even when the guiding template is specified by an altered depiction of the target (Bravo & Farid, 2009; Hout & Goldinger, 2015; Vickery et al., 2005), a name cue (Malcolm & Henderson, 2009; Spotorno et al., 2014; Yang & Zelinsky, 2009; Zelinsky et al., 2013), or the target's function (Castelhano & Witherspoon, 2016; Humphreys & Riddoch, 2001). All of these studies suggest that the active template can vary in its level of abstraction but is nevertheless crucial for guiding search toward plausible locations.

However, one everyday experience completely challenges the notion that an active guiding template is a prerequisite for efficient search. Imagine that you go to a conference and decide to look for someone familiar to sit next to. Suddenly, you notice a colleague whom you did not expect to see. Clearly, you could not have
had an active guiding template of this colleague’s face in mind because you were not expecting to see them. In addition, because this search involved looking for a familiar face among many other faces, even unspecific templates (such as category-based templates) could not help you distinguish the target face from other face distractors. Therefore, in this type of search, people do not have an active guiding template but, rather, a very large range of representations in long-term memory (LTM) that are target candidates. Given that statistical models (Gelman, 2013; Gurevitch, 1961; McCormick et al., 2010) estimate that a person knows hundreds of people, it is hard to believe that this massive number of representations in LTM can be combined into one or a handful of active templates—that is, to depict a set of visual properties that are compared with the visual input until the target is found. If an active guiding-template mechanism is not probable in this case, what are the cognitive processes underpinning this type of search?

Obviously, it could be argued that a search for a familiar face can be achieved only through unguided search, in which the person must serially scan the people at the conference until someone familiar is found. Although this is possible, some of us feel that a familiar face can attract our attention and simply “jump” into view. This intuitive feeling suggests that despite not having an active guiding template, search can still be guided by LTM representations. That is, the familiar stimulus could be processed to some extent (whether fully recognized or only eliciting a familiarity signal) and thus direct the search to the whereabouts of the target. This implies that the matching between an active guiding template and the visual input is not the sole way to conduct a guided search. In the current study, we suggest an alternative mechanism to guided search that relies on the ability to process LTM representations through extrafoveal vision. For that purpose, we tracked Israeli and German participants’ eye movements during a visual search task in which they were asked to look for the face of a familiar person from among five faces, without knowing whom they were searching for. The findings provide compelling evidence that search for an unexpected familiar face can successfully be guided, thus suggesting that a revision of current theoretical visual search frameworks would be warranted and highlighting the importance of LTM-based extrafoveal processing.

**Method**

The experimental procedure was preregistered at https://osf.io/kxpdlg and https://osf.io/wf5K2 for the Israeli and German groups, respectively. This study was approved by the Ethics Committee of the Faculty of Social Sciences at The Hebrew University of Jerusalem.

**Participants**

Thirty-six and 41 participants were recruited for the Israeli and German groups, respectively. We eliminated participants with more than 50% disqualified trials (see exclusion criteria below), resulting in a final sample of 35 Israelis (10 men; M = 24.8 years) and 33 Germans (six men; M = 28.5 years). Prior to the experiment, we conducted a power analysis based on the null hypothesis for the mean ordinal number in the revealed block (for more details regarding the analysis, see below). This power analysis indicated that to obtain a power of at least .8 for detecting a medium effect size (d = 0.5), approximately 32 participants were needed (for more information, see the preregistrations).

All participants had normal or corrected-to-normal vision and took part in the experiment in return for course credits or payment (Israelis: 30 shekels; Germans: 7.5 euros). To encourage correct performance, we told participants that they would receive a bonus payment (Israelis: 10 shekels; Germans: 2.5 euros) if they started the search on time, found the familiar face quickly, and correctly recognized which face was the familiar one.

**Statement of Relevance**

Searching for something, whether it is your keys or a familiar face, is a frequent everyday activity. Under some circumstances, such as in security settings, it even carries life-saving implications. Until now, it was widely believed that in order to find what they are looking for, people need to know at least some aspects of what they are trying to find. However, this assumption is inconsistent with common human experiences, such as suddenly finding a friend in a crowd although there was no prior expectation of seeing them. Here, we demonstrate that despite the many people that each person knows, we can all find a familiar face embedded in unfamiliar ones, even without knowing the identity of that face in advance. This calls for a modification of current theories and makes clear that the cognitive system can utilize information from a large area of the visual environment to guide search, even if it is unclear what the target looks like.
Apparatus and stimuli

The stimuli were displayed on a 24-in. monitor, with a 1,920 × 1,080 pixel screen resolution and a 120-Hz (Israelis) or 144-Hz (Germans) refresh rate. Monocular gaze position was tracked at 1,000 Hz with a tower-mounted EyeLink 1000+ eye tracker (SR Research, Kanata, Ontario, Canada). The distance from the center of the screen to the participants' eyes was 53 cm.

After collection of face stimuli from Google images, all images were cropped, resized, and transformed to black and white. Then, the images were divided into 100 sets; each set consisted of five faces that included one face of an Israeli celebrity and one face of a German celebrity. We used MATLAB (The MathWorks, Natick, MA) code to make sure that the mean luminance of all the images in each set was equal. In addition, we used Amazon Mechanical Turk (MTurk) questionnaires to ensure that (a) the familiar faces did not differ from the unfamiliar faces in a trait other than familiarity, (b) the familiarity status of the faces was correct (i.e., no familiar face was mistakenly tagged as unfamiliar), and (c) the familiar Israeli and German faces were not universally familiar. For further details on the processing of the stimuli and the MTurk questionnaires, see the Supplemental Material available online.

Design

Participants completed a visual search task in which they were asked to look for a familiar face among five faces (see Fig. 1). The array of faces included one Israeli celebrity, one German celebrity, and three nonfamous faces. The area of each face was scaled to a circle with a radius of 4 degrees of visual angle (DVA). The center of each face was at an equal distance of 9 DVA from the center of the screen. The distance between the center of each pair of adjacent faces and each pair of nonadjacent faces was 10.6 DVA and 17.1 DVA, respectively. Before each trial, a drift correction was carried out, allowing for a deviation of only 2 DVA between the predicted gaze position and the actual fixation point. After the drift-correction procedure, a cross appeared in the center of the screen (for 500 or 1,500 ms), and participants were asked to fixate on the cross and start the search only when it disappeared. In instances in which they shifted their gaze from the cross before it disappeared, an error beep was played and the trial continued as usual (although these trials were later excluded from analysis). Participants were asked to find the target and press the space bar when they were looking at it. If participants pressed the space bar without fixating on the target, a message appeared on
the screen informing the participants that they had not looked at the target when they responded.

This task was carried out in two separate blocks. In the revealed block (100 trials), all five faces were visible from the start of the search task. In contrast, in the gaze-contingent block (10 trials), participants saw five circles serving as placeholders for the faces. The face was displayed only when participants shifted their gaze to one of these placeholders. Obviously, in this condition, no stimulus was presented in the periphery of the visual field, and thus no guidance of gaze by familiar faces was expected. Accordingly, our aim in this block was to confirm that the pattern of results in the revealed block was due to the processing of familiar faces by extrafoveal vision. Each face in the revealed condition appeared only once throughout the experiment. The faces in the gaze-contingent block had already appeared in the preceding revealed block.

To familiarize participants with the task and to provide an example of a task that would be strongly based on extrafoveal vision, we administered a color block prior to the main task. In this block, the five stimuli were not faces but, rather, four circles in one color and a single circle in another color (colors switched across trials). This block was composed of 10 trials that were repeated until participants completed more than 70% of them correctly (i.e., they started the search only after the cross had disappeared and responded only when looking at the target). Beyond this training, participants also engaged in four training trials with faces, prior to the initiation of the revealed block. At the beginning of the experiment (prior to the color block), each participant went through the standard 9-point calibration and validation procedure provided with the eye tracker.

Data-exclusion criteria

Trials were excluded when one or more of the following conditions occurred: (a) Participants shifted their gaze before the cross in the center had disappeared, (b) participants pressed the space bar without looking at the familiar face, and (c) eye tracking was of bad quality (more than 20% of the eye samples were missing). This procedure led to the removal of 21% and 36% of the trials in the Israeli and German samples, respectively. In addition, we excluded participants when more than 50% of their trials were disqualified according to these criteria. To ensure that the results are robust and representative even without these strict exclusion criteria, we ran the same analysis without removing any trials. The results of this analysis were qualitatively similar to the results described below and are further discussed in the Supplemental Material.

Results

The ordinal number

The spatiotemporal dynamics of gaze behavior during the visual search task provided information on the order in which the faces were scanned during the trial, thus making it possible to assign an ordinal number for each face. If guidance was present, the mean ordinal number of the familiar faces should be lower than (a) the expected value under the null hypothesis, (b) the mean ordinal number of the familiar faces of the other group, and (c) the mean ordinal number of the familiar faces in the gaze-contingent block. Below, we elaborate on each of these comparisons. For further information on the computation of the ordinal number, see the Supplemental Material.

Under the null hypothesis (i.e., the search could not be guided toward the familiar face), we would expect that the familiar face would not be privileged in terms of scanning order. Thus, each ordinal number of the familiar face would be equally probable, implying that the ordinal number should follow a discrete uniform distribution with a probability of .2 for each possibility and an expected value of 3. However, if participants could guide their search and direct their gaze according to face familiarity, we would expect them to direct their gaze toward the familiar face earlier in the trial, resulting in a mean ordinal number of less than 3. Our analysis showed exactly this pattern, where the mean ordinal number of the familiar face was less than 3 in both groups—Israeli sample: $t(34) = -5.26, p < .001, d = 0.89, 95\%$ confidence interval (CI) of $d = [0.5, 1.29]$; German sample: $t(32) = -6.83, p < .001, d = 1.19, 95\%$ CI $= [0.75, 1.66]$ (see Fig. 2).

Moreover, to control for visual distinctiveness of the famous faces, we compared the ordinal number of the familiar face of one national group with the ordinal number of the familiar face of the other national group. That is, for the Israeli observers, we compared the ordinal number of the Israeli celebrity with the ordinal number of the German celebrity, and vice versa for the German participants. This analysis enabled us to compare the same sets of faces, where familiarity was the only factor differentiating these sets between the groups. The ordinal number of the familiar face was lower than the ordinal number of the familiar face of the other group in both groups—Israeli sample: $t(34) = -4.55, p < .001, d = 0.77, 95\%$ CI $= [0.39, 1.16]$; German sample: $t(32) = -3.34, p < .001, d = 0.58, 95\%$ CI $= [0.21, 0.96]$ (see Fig. 2). Finally, we also compared the ordinal number of the familiar face in the revealed block with the ordinal number of the familiar face in the gaze-contingent block. This comparison showed
that in the gaze-contingent block, when extrafoveal vision was not available, the ordinal number of the familiar face was higher in comparison with the revealed block for both groups—Israeli sample: $t(32) = 2.15, p = .019, d = 0.36, 95\% \text{ CI} = [0.02, 0.71]$; German sample: $t(32) = 2.08, p = .023, d = 0.36, 95\% \text{ CI} = [0.01, 0.72]$. Note that the color block provided a case in which task completion could easily be achieved by extrafoveal vision. Indeed, the mean ordinal number of the target in this block was very close to 1 (1.03 in both samples).

To further probe what caused the decrease in the ordinal number of the familiar face, we examined the distribution of the ordinal numbers across all participants. This analysis demonstrated that the distribution of the ordinal numbers departed from a uniform distribution in both groups—Israeli sample: $\chi^2(4, N = 2,754) = 51.84, p < .001$; German sample: $\chi^2(4, N = 2,283) = 80.1, p < .001$. The discrepancy from the uniform distribution was mainly caused by the high proportion of trials in which the ordinal number of the familiar face was 1. Consistent with this analysis, results showed that the mean proportion of trials in which the ordinal number of the familiar face was 1 was significantly higher than .2, which is the expected mean proportion under the null hypothesis—Israeli sample: $t(34) = 4.54, p < .001, d = 0.77, 95\% \text{ CI} = [0.39, 1.16]$; German sample: $t(32) = 7.43, p < .001, d = 1.29, 95\% \text{ CI} = [0.84, 1.78]$ (see Fig. 2).

Next, we examined whether more time to process the extrafoveal information would improve participants’ search performance. To test this assumption, we validated that, as instructed, participants started their search only 500 ms or 1,500 ms after the search array appeared. The results did not show a substantial improvement associated with additional display time before the search began. Participants fixated on the familiar face earlier when they had more time before they were allowed to move their gaze, but this effect was not statistically significant in either group—Israeli sample: $t(34) = -1.51, p = .070, d = 0.26, 95\% \text{ CI} = [-0.08, 0.6]$; German sample: $t(32) = -1.34, p = .095, d = 0.23, 95\% \text{ CI} = [-0.12, 0.59].
**Scanning pattern**

The ordinal-number analysis (see Fig. 2) indicated a high proportion of trials in which participants fixated on the familiar face before all other faces. To explore whether there was guidance by familiarity when participants did not find the familiar face right away and scanned the array more extensively, we further analyzed the scanning pattern after the first fixation on the faces.

First, we characterized the scanning pattern and investigated whether participants scanned the array of faces in a systematic manner (i.e., clockwise or counterclockwise). For that purpose, we analyzed the gaze transitions between faces and scored a transition as 1 if it adhered to a clockwise or counterclockwise scanning pattern (see Fig. 3a). Then, we summed these scores separately in two counters for clockwise and counterclockwise patterns. The final scanning score was the maximum value between these two sums divided by the total number of transitions. Thus, the scanning score reflected the proportion of transitions that followed a consistent clockwise or counterclockwise scanning pattern. A permutation procedure was used to quantify the chance level of the scanning scores (for additional details, see the Supplemental Material).

As can be seen in Figure 3a, both groups had scanning scores above chance level—Israeli sample: \( t(34) = 34.26, p < .001, d = 5.79, 95\% \text{ CI} = [4.44, 7.27] \); German sample: \( t(32) = 30.67, p < .001, d = 5.34, 95\% \text{ CI} = [4.05, 6.75] \). Interestingly, there was also a cultural difference between the two samples, in that the German participants had a more systematic pattern of scanning than the Israeli sample: \( t(66) = -6.05, p < .001, d = 1.47, 95\% \text{ CI} = [0.93, 2] \).

The strong tendency to scan the array in a clockwise or counterclockwise manner implies that when participants reached the first face, they usually shifted their gaze between adjacent faces. However, shifts between nonadjacent faces were still present. These “long-distance” shifts (see Fig. 3b) toward a familiar face are a valid indicator of guidance of search because they show that the participant deviated from his or her typical scanning pattern, even during active search. Accordingly, we examined whether these long-distance shifts were more probable toward the familiar face than toward the familiar face of the other group. For that purpose, we computed the percentage of trials in which long-distance shifts were made toward the familiar face (Israeli sample: \( M = 5.1\%, SD = 3.9\% \); German sample: \( M = 6.8\%, SD = 5.3\% \)) in comparison with the familiar face of the other group (Israeli sample: \( M = 3.8\%, SD = 3.6\% \); German sample: \( M = 5.5\%, SD = 4\% \)). These differences were significant—Israeli sample: \( t(34) = 2.54, p = .008, d = 0.43, 95\% \text{ CI} = [0.08, 0.79] \); German sample: \( t(32) = 1.74, p = .046, d = 0.3, 95\% \text{ CI} = [-0.05, 0.66] \).

**Saccade amplitude**

The higher proportion of trials with long-distance shifts toward the familiar face suggests that guidance toward the familiar face would also manifest in the amplitude of the saccade (i.e., the length of the movement of the eye between two relatively stable fixations). Specifically, if gaze was shifted toward familiar faces from longer distances than to unfamiliar faces, we would predict that the amplitude of the saccades preceding the first fixation on the familiar faces would be, on average, longer than the amplitude of the saccades preceding the first fixation to the unfamiliar face. As predicted, the analysis confirmed this pattern of gaze behavior (see Fig. 4a): Larger saccade amplitudes preceded the first fixation on the familiar face than the first fixation on the face that was familiar to the other nationality—Israeli sample: \( t(34) = 3.22, p = .001, d = 0.54, 95\% \text{ CI} = [0.19, 0.91] \); German sample: \( t(32) = 3.4, p < .001, d = 0.59, 95\% \text{ CI} = [0.22, 0.97] \). Because the distance between the central cross and the center of the faces (9 DVA) was smaller than the distance between the centers of the adjacent faces (10.6 DVA), we also verified that this effect could not be attributed to the first saccade from the central cross to the first face. For that purpose, we excluded the saccade that preceded the fixation on the first face and repeated the analysis. The same pattern of results was obtained: Longer saccades preceded the fixation on the familiar face in comparison with the familiar face of the other group—Israeli sample: \( t(34) = 2.7, p = .005, d = 0.46, 95\% \text{ CI} = [0.11, 0.81] \); German sample: \( t(32) = 2.17, p = .019, d = 0.38, 95\% \text{ CI} = [0.02, 0.74] \).

To better characterize the source of this effect, we compared the distributions of saccades preceding the familiar face with the saccades preceding the familiar face of the other group. We plotted the quantiles of one distribution against the quantiles of the other distribution (i.e., quantile-quantile [QQ] plots). QQ plots provide an intuitive visualization of the comparison between distributions according to their quantiles; if the distributions are similar, the values of the quantiles of the two distributions should be approximately equal. Therefore, when we plot the values of the quantiles against each other, they should lie on the cardinal diagonal (i.e., \( y = x \)). Any deviation from that diagonal line reflects a difference between the two distributions. As
Fig. 3. Summary of scanning-pattern analysis. The distribution of average scanning scores across all trials (a) is shown separately for the Israeli and German participants. The gray area is the distribution of scores following permutations of the order of faces and captures the chance level of scanning scores across both groups. The vertical line represents the mean of each distribution. On the left, we display the scanning patterns that were considered systematic scanning. The percentage of trials in which long-distance shifts were made toward the familiar Israeli faces (bars with blue and white stripes) and the familiar German faces (bars with black, red, and yellow stripes) is shown in (b), separately for Israeli and German participants. Error bars indicate 95% confidence intervals. Asterisks indicate significant differences between whether faces were famous in participants’ native country or the comparison country (*p < .05, **p < .01 for paired-samples t-test comparisons). At the top, we display an example of long-distance shifts.
highlighted in Figure 4b, the QQ plots of the two groups diverged from the diagonal line in the range of approximately 12.5 to 17 DVA. Specifically, in this range, the values that corresponded to the quantiles of the distribution of the saccades preceding the familiar face were larger. This finding indicates that in this range, the saccades that preceded the familiar face were longer in comparison with the saccades that preceded the familiar face of the other group. Because the distance between the centers of two adjacent faces was 10.5 DVA (the distance between the near and far edges was 2.5 and 18.5 DVA, respectively; see the Method section), having longer saccades in the range of 12.5 to 17 DVA constituted another indication for saccades between pairs of nonadjacent faces or between relatively distant locations in adjacent faces. Accordingly,
the discrepancy between the two distributions of saccades specifically in the described range provides further support for shifts in gaze toward the familiar face based on extrafoveal processing of familiarity (because the fovea lies within a radius of approximately 1 DVA from the center of the retina; Strasburger et al., 2011).

Discussion

Two groups of participants, in two different labs and two different continents, were exposed to the same sets of faces and were asked to look for the familiar face in each set. The familiar face of one group was an unfamiliar face for the other group. Nevertheless, gaze behavior was impressively similar in the two groups: Participants fixated on the familiar face earlier than expected by chance and typically before gaze was directed toward the familiar face of the other group. This effect was mainly driven by a higher probability that the familiar face would be the first face selected during the trial. However, guidance was also evident when participants did not find the familiar face on their first attempt. During the active search, there was a higher probability for the participants to shift their gaze toward the familiar face from nonadjacent faces. Gaze shifts before the first fixation on the familiar face were significantly larger, on average, than the shifts targeting the familiar face of the other group. These indices of guidance were found despite participants' tendency to systematically scan the face array in a clockwise or counterclockwise manner.

Importantly, the design of the task (in which participants saw the same faces but with different familiarity tagging) makes it possible to distill the effect of familiarity on gaze behavior and rule out alternative explanations for these results (e.g., visual saliency). Moreover, because familiarity differentiated the target from the distractors and because of the large number of faces that are familiar to each individual, it seems improbable that any active template was available to guide search in this case. In addition, each face was displayed only once, hampering the construction of an active guiding template throughout the experiment. Therefore, our findings suggest that long-term familiarity was sufficient to guide gaze toward the target, even in the absence of an active search template. It should be emphasized that our findings not only are consistent with the ability to process familiarity through extrafoveal vision but also provide strong evidence that this ability can be further used by the cognitive system to facilitate the deployment of visual attention during search.

How could the search for the familiar face be guided even though the participants did not know whom they were looking for? This seemingly paradoxical situation led previous researchers to assume that an active guiding template is mandatory for successful completion of a search (Bravo & Farid, 2009, 2014; Hout & Goldinger, 2015; Madrid et al., 2019; Malcolm & Henderson, 2009; Reeder et al., 2015; Vickery et al., 2005; Wolfe, 2020; Yang & Zelinsky, 2009; Zelinsky et al., 2006). This assumption is apparently inconsistent with visual search tasks in which participants search for the odd-one-out target, sometimes without knowing in advance what the target is (Friedman-Hill & Wolfe, 1995; Maljkovic & Nakayama, 1994; Müller et al., 1995; Wang et al., 1994). However, in these tasks, the odd-one-out target is defined by the visual features of the distractors (e.g., color and orientation), so that an active guiding template can easily be constructed after exposure to two distractors (e.g., two green distractors mean that the target color is different). Accordingly, although such tasks might start without an active template, it can be constructed rapidly after the onset of the trial. Importantly, although searching for a familiar face among unfamiliar ones may be conceptualized as a search for the odd-one-out target, the exposure to the distractors (i.e., the unfamiliar faces) is not informative in any way as to the target in question (i.e., the familiar face). Thus, although searching for the familiar face resembles the odd-one-out search in some ways, it differs from it on a crucial point: When participants search for the familiar face, no active guiding template could be generated by processing the distractors.

A recent theoretical review (Hulleman & Olivers, 2017) suggests that the centrality of the search template in previous studies and frameworks derives from theories that emphasize individual items as the unit of search. According to these item-based theories, search proceeds on the basis of selection of individual items that are either rejected as distractors or recognized as the target. Therefore, it is reasonable to view search as a process in which items are compared with a template specifying the target. In contrast, Hulleman and Olivers put forward an alternative framework of visual search in which fixations (and not individual items) are the central unit of search. Specifically, during visual search, a functional viewing field is defined around each fixation according to the area of the visual field that can be processed through this fixation. The functional viewing field is not fixed but changes as a function of processing difficulty, that is, the more easily a target can be discriminated from distractors through extrafoveal vision, the larger the functional viewing field should be and the faster the target can be found. The reverse would be true when the target is more difficult to discriminate. In line with this framework, having an active guiding template could make the target more discriminable and facilitate search, but it is not indispensable
to the search process; other features of the target may also enhance its discriminability. Specifically, our results demonstrate that familiarity is one such feature because we observed guidance of search toward the familiar face even when participants did not know whom they were looking for. Importantly, when participants could not benefit from extrafoveal vision (the gaze-contingent condition in the current study), this type of guidance vanished.

When an active guiding template does exist, search is presumed to be guided to places in the visual field that generate a match to the template (Bravo & Farid, 2014). When an active template is not available, search must be guided by multiple LTM representations rather than a few active ones. Computational models of memory suggest that encoding and retrieval processes (including face recognition) may be performed by attractor networks (Amit, 1992; Farah et al., 1993; Tank & Hopfield, 1987). These models are composed of several units that are linked through weighted connections and are updated during encoding until the network incorporates several stable states (i.e., attractors) that represent memories. Subsequent to a new input, the network settles on one of the attractors on the basis of the input and the weights. Notably, even a partial input can activate the correct memory, as suggested by theories of pattern-completion processes and empirical findings on the robustness of face recognition despite substantial changes to the visual input (Bruce et al., 2001; Bruck et al., 1991; Ramon, 2015). The attractor network may be a plausible model underlying the guidance of search based on LTM, as observed in our study. In some cases, the information perceived through extrafoveal vision was sufficient to lead the network toward an established attractor, thus guiding search toward the target without necessarily matching the input to an active template.

Early studies of visual search (Treisman & Gelade, 1980; Wolfe et al., 1989) used search arrays with different set sizes to examine how the slope of the reaction times (i.e., time until the target is detected) changes as a function of the number of distractors. This slope estimates the cost in search time for additional distractors and was conceptualized as an indication of how efficiently attention can be guided toward the target. Interestingly, certain findings indicated that the overall number of fixations during the search closely mirrored the reaction time (Binello et al., 1995; Young & Hulleman, 2013; Zelinsky & Sheinberg, 1995, 1997). Moreover, there is considerable evidence that gaze shifts are preceded by covert shifts of attention toward the same locations (Deubel & Schneider, 1996; Henderson, 1992; Henderson et al., 1989; Hoffman & Subramaniam, 1995; Kowler et al., 1995). Taken together, gaze shifts may be another, presumably more overt, indicator of attentional guidance. Therefore, despite some accounts that view eye movements as an epiphenomenon of visual search (Treisman, 1982; Wolfe, 1998), eye movements have become increasingly more common and crucial to visual search studies (Beutter et al., 2003; Deubel & Schneider, 1996; Hout & Goldinger, 2015; Rao et al., 2002; Tavassoli et al., 2009; Young & Hulleman, 2013; Zelinsky et al., 2013) and theories (Zelinsky, 2008; Zelinsky et al., 2006). Thus, in the current study, we used eye movements rather than reaction time slopes to shed light on the dynamics of attention in visual search.

It is no wonder that visual search, being one of the most common human activities, has become one of the most frequently investigated topics in the cognitive sciences. In different frameworks, the active search template tends to be considered essential for guiding search. In the current study, we showed that search can be guided even in the absence of an active visual template through extrafoveal processing of familiarity. These findings corroborate new theoretical accounts (Hulleman & Olivers, 2017) that claim that search performance is determined by the discriminability of the target through extrafoveal vision. Hence, we believe that to decipher the underlying mechanisms of visual search, researchers should focus their efforts on understanding the processing of target information in extrafoveal vision rather than the ability to maintain an active guiding template.

**Transparency**

**Action Editor:** Krishnankutty Sathian

**Editor:** Patricia J. Bauer

**Author Contributions**

All the authors developed the study concept and contributed to the design and the required analysis methods. O. C. Lancry-Dayan analyzed and interpreted the data under the supervision of Y. Pertzov and M. Gamer. All the authors wrote the manuscript and approved the final version for submission.

**Declaration of Conflicting Interests**

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

**Funding**

This research was supported by an Israel Science Foundation grant (2414/20) to Y. Pertzov and a German Israel Foundation grant (I-2416-105.4/2016) to Y. Pertzov.

**Open Practices**

All data, analysis code, and materials have been made publicly available via OSF and can be accessed at https://osf.io/bkuwt. The design and analysis plans for the Israeli and German groups were preregistered at https://osf.io/kxpdg and https://osf.io/wf5k2, respectively. This article has received the badges for Open Data, Open Materials,
References

Alexander, R. G., Nahvi, R. J., & Zelinsky, G. J. (2019). Specifying the precision of guiding features for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 45*(9), 1248–1264.

Amit, D. J. (1992). *Modeling brain function: The world of attractor neural networks*. Cambridge University Press.

Beutter, B. R., Eckstein, M. P., & Stone, L. S. (2003). Saccadic and perceptual performance in visual search tasks. I. Contrast detection and discrimination. *Journal of the Optical Society of America A, 20*(7), 1341–1355.

Binello, A., Ruddock, K. H., & Mannan, S. (1995). The characteristics of eye movements made during visual search with multi-element stimuli. *Spatial Vision, 9*(3), 343–362.

Bravo, M. J., & Farid, H. (2009). The specificity of the search template. *Journal of Vision, 9*(1), Article 34. https://doi.org/10.1167/9.1.34

Bravo, M. J., & Farid, H. (2014). Informative cues can slow search: The cost of matching a specific template. *Attention, Perception, & Psychophysics, 76*(1), 32–39.

Bruce, V., Henderson, Z., Newman, C., & Burton, A. M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology: Applied, 7*(3), 207–218.

Bruck, M., Cavanagh, P., & Ceci, S. J. (1991). Fortysomething: Recognizing faces at one’s 25th reunion. *Memory & Cognition, 19*(3), 221–228.

Castelhano, M. S., & Witherspoon, R. L. (2016). How you use it matters: Object function guides attention during visual search in scenes. *Psychological Science, 27*(5), 606–621.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36*(12), 1827–1838.

Farah, M. J., O’Reilly, R. C., & Vecera, S. P. (1993). Dissociated overt and covert recognition as an emergent property of a lesioned neural network. *Psychological Review, 100*(4), 571–588.

Findlay, J. M. (1997). Saccade target selection during visual search. *Vision Research, 37*(5), 617–631.

Friedman-Hill, S., & Wolfe, J. M. (1995). Second-order parallel processing: Visual search for the odd item in a subset. *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 531–551.

Gelman, A. (2013, February 19). The average American knows how many people? *The New York Times*, D7. https://www.nytimes.com/2013/02/19/science/the-average-american-knows-how-many-people.html#:%3E%3D:text=The%20average%20American%20knows%20about%20600%20people.

Gurevitch, M. (1961). *The social structure of acquaintance-ship networks* [Doctoral thesis]. Massachusetts Institute of Technology. https://dspace.mit.edu/handle/1721.1/11312

Henderson, J. M. (1992). Visual attention and eye movement control during reading and picture viewing. In K. Rayner (Ed.), *Eye movements and visual cognition* (pp. 260–283). Springer.

Henderson, J. M., Pollatsek, A., & Rayner, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Perception & Psychophysics, 45*(3), 196–208.

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics, 57*(6), 787–795.

Hout, M. C., & Goldinger, S. D. (2015). Target templates: The precision of mental representations affects attentional guidance and decision-making in visual search. *Attention, Perception, & Psychophysics, 77*(1), 128–149.

Hulleman, J., & Olivers, C. N. (2017). The impending demise of the item in visual search. *Behavioral and Brain Sciences, 40*, Article e132. https://doi.org/10.1017/S0140525X15002794

Humphreys, G. W., & Riddoch, M. J. (2001). Detection by action: Neuropsychological evidence for action-defined templates in search. *Nature Neuroscience, 4*(1), 84–88.

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research, 35*(13), 1897–1916.

Luria, S. M., & Strauss, M. S. (1975). Eye movements during search for coded and uncoded targets. *Perception & Psychophysics, 17*(3), 303–308.

Madrid, J., Cunningham, C. A., Robbins, A., & Hout, M. C. (2019). You’re looking for what? Comparing search for familiar, nameable objects to search for unfamiliar, novel objects. *Visual Cognition, 27*(1), 8–20.

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition, 22*(6), 657–672.

McCormick, T. H., Salganik, M. J., & Zheng, T. (2010). How many people do you know? Efficiently estimating personal network size. *Journal of the American Statistical Association, 105*(489), 59–70.

Miller, A. (2012, March 20). *Lost something already today? Mislplaced items cost us ten minutes a day*. Daily Mail. https://www.dailymail.co.uk/news/article-2117987/Lost-today-Mislplaced-items-cost-minutes-day.html

Motter, B. C., & Belky, E. J. (1998). The guidance of eye movements during active visual search. *Vision Research, 38*(12), 1805–1815.

Müller, H. J., Heller, D., & Ziegler, J. (1995). Visual search for singleton feature targets within and across feature dimensions. *Perception & Psychophysics, 57*(1), 1–17.
Ramón, M. (2015). Perception of global facial geometry is modulated through experience. *PeerJ*, 3, Article e850. https://doi.org/10.7717/peerj.850

Rao, R. P., Zelinsky, G. J., Hayhoe, M. M., & Ballard, D. H. (2002). Eye movements in iconic visual search. *Vision Research*, 42(11), 1447–1463.

Reeder, R. R., van Zoest, W., & Peelen, M. V. (2015). Involuntary attentional capture by task-irrelevant objects that match the search template for category detection in natural scenes. *Attention, Perception, & Psychophysics*, 77(4), 1070–1080.

Spotorno, S., Malcolm, G. L., & Tatler, B. W. (2014). How context information and target information guide the eyes from the first epoch of search in real-world scenes. *Journal of Vision*, 14(2), Article 7. https://doi.org/10.1167/14.2.7

Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision*, 11(5), Article 13. https://doi.org/10.1167/11.5.13

Tank, D. W., & Hopfield, J. J. (1987). Collective computation in neuronlike circuits. *Scientific American*, 257(6), 104–115.

Tavassoli, A., van der Linde, I., Bovik, A. C., & Cormack, L. K. (2009). Eye movements selective for spatial frequency and orientation during active visual search. *Vision Research*, 49(2), 173–181.

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance*, 8(2), 194–214.

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

Vickery, T. J., King, L.-W., & Jiang, Y. (2005). Setting up the target template in visual search. *Journal of Vision*, 5(1), Article 8. https://doi.org/10.1167/5.1.8

Wang, Q., Cavanagh, P., & Green, M. (1994). Familiarity and pop-out in visual search. *Perception & Psychophysics*, 56(5), 495–500.

Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *Attention* (pp. 13–74). Psychology Press.

Wolfe, J. M. (2020). Visual search: How do we find what we are looking for? *Annual Review of Vision Science*, 6, 559–562.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 15(3), 419–433.

Yang, H., & Zelinsky, G. J. (2009). Visual search is guided to categorically defined targets. *Vision Research*, 49(16), 2095–2103.

Young, A. H., & Hulleman, J. (2013). Eye movements reveal how task difficulty moulds visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 39(1), 168–190.

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review*, 115(4), 787–835.

Zelinsky, G. J., Peng, Y., & Samaras, D. (2013). Eye can read your mind: Decoding gaze fixations to reveal categorical search targets. *Journal of Vision*, 13(14), Article 10. https://doi.org/10.1167/13.14.10

Zelinsky, G. J., & Sheinberg, D. L. (1995). Why some search tasks take longer than others: Using eye movements to redefine reaction times. In J. M. Findlay, R. Walker, & R. W. Kentridge (Eds.), *Studies in visual information processing: Vol. 6: Eye movement research: Mechanisms, processes and applications* (Vol. 6, pp. 325–336). Elsevier.

Zelinsky, G. J., & Sheinberg, D. L. (1997). Eye movements during parallel–serial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 23(1), 244–262.

Zelinsky, G. J., Zhang, W., Yu, B., Chen, X., & Samaras, D. (2006). The role of top-down and bottom-up processes in guiding eye movements during visual search. In Y. Weiss, B. Schölkopf, & J. Platt (Eds.), *Advances in Neural Information Processing Systems* (Vol. 18, pp. 1569–1576). MIT Press.