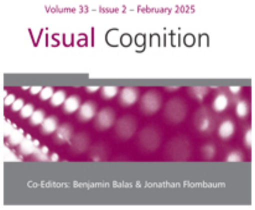

Visual Cognition

ISSN: 1350-6285 (Print) 1464-0716 (Online) Journal homepage: www.tandfonline.com/journals/pvis20

Face detection in the periphery and at fixation

Alice Nevard, Graham J. Hole & Markus Bindemann

VISUAL COGNITION
https://doi.org/10.1080/13506285.2025.2550777

Face detection in the periphery and at fixation

Alice Nevard [a,b], Graham J. Hole[c] and Markus Bindemann[a]

[a]School of Psychology, University of Kent, Canterbury, UK; [b]Department of Psychology, Durham University, Durham, UK; [c]School of Psychology, University of Sussex, Brighton, UK

ABSTRACT
Human faces are located easily in peripheral vision, but it is unclear which facial information drives this process. We employed a gaze-contingent paradigm to investigate face detection in the periphery and at fixation. Participants searched complex scenes for faces. As they made a saccade towards a peripherally-presented face, it sometimes changed in appearance before reaching foveal view. Changes in facial identity (Experiment 1) or the addition of internal features (e.g., eyes, nose, mouth) to an initially-featureless face (Experiment 2) did not affect face detection, yet performance was hindered by the removal of internal features before fixation occurred. While the presence of internal features in peripheral faces could be perceived, the organization of these features was difficult to extract (i.e., when internal features were rotated 90 degrees) (Experiments 3 and 4). These findings point to a two-stage model of face detection. In the initial stage of encountering faces in the periphery, where visual acuity is low, observers are insensitive to internal facial features and their configuration. Once a face is located and can be viewed in high-acuity regions of central vision, a second stage may use detail from internal features to confirm that a face has been detected.

ARTICLE HISTORY
Received 11 November 2024
Accepted 15 August 2025

KEYWORDS
Cognitive templates; detection; face; peripheral; visual search

Faces provide information about a person's identity and their emotional and attentive state, which plays an important role in guiding human behaviour. Crucially, however, a face first must be noticed in the visual field for this information to be accessed. In this sense, *detection* underlies all other tasks with faces. While identification, emotion and gaze perception have been studied extensively in Psychology, these studies typically present single faces at the centre of a blank background. In this situation, the problem of detection is largely solved by the mode of presentation itself. Real-life face detection requires finding faces in complex visual displays, and studies using centrally-presented faces or faces on blank backgrounds tell us little about how this is achieved. For example, when observers are asked to search for faces in visual arrays or complex scenes, a viewpoint effect emerges whereby frontal faces are detected faster and more accurately than profile faces (Bindemann & Lewis, 2013; Burton & Bindemann, 2009; Nevard et al., 2023). In contrast, this viewpoint effect disappears when faces are presented in the screen centre or in otherwise blank displays (Bindemann & Lewis, 2013; Nevard et al., 2023).

By using array and scene paradigms to study detection, it has therefore become clear that this process places highly specialized demands on the perceptual system, which are quite different from other tasks with faces. For example, whereas identification relies heavily on the internal features of a face, such as the eyes, nose and mouth (e.g., Ellis et al., 1979; Moscovitch & Moscovitch, 2000; Nachson & Shechory, 2002; Young et al., 1985), detection appears to proceed unhindered when these features are removed (Hershler & Hochstein, 2005; Nevard et al., 2023). It is clear that detection must rely on visual characteristics common to a wide range of faces – otherwise only some faces would be detected but not others. Therefore, a detection template that operates independently of identity-relevant information has utility. In turn, detection is affected by transformations that leave person identification intact. Geometric distortions of faces, for example,

CONTACT Alice Nevard alice.o.nevard@durham.ac.uk Department of Psychology, Durham University, Durham, DH1 3LE, UK

whereby these stimuli are stretched selectively in either a vertical or horizontal dimension, impair detection (Pongakkasira & Bindemann, 2015) but not recognition (Bindemann et al., 2007; Hole et al., 2002). Detection also operates best with whole rather than part faces and when realistic colours are used (Bindemann & Burton, 2009; Burton & Bindemann, 2009). Collectively these findings imply a face detection template that is attuned to overall face-shape and colour, rather than being based on the internal facial features that are so important for person identification. Yet, detection is also facilitated by cognitive templates that combine information from the internal features of a variety of faces, such as statistical averages of previously encountered faces (Prunty et al., 2024), while blurring, obscuring and reconfiguring of internal features can impair face detection too (e.g., Hershler & Hochstein, 2005; Lewis & Edmonds, 2003). This leaves uncertainty about how these different sources of information are used during detection, and as to why differences between detection and other tasks with faces arise.

In contrast to tasks such as face identification, which are typically resolved with faces at fixation, a variable that may be influential in detection is visual acuity. This is not distributed evenly across the visual field, but is highest for stimuli presented in the centre of the visual field (the fovea) and declines in the periphery (Kalloniatis & Luu, 2007; Westheimer, 1965). It seems likely that peripheral vision holds particular importance in detection processes, such as the search for faces in the environment, as it is relied upon for viewing complex scenes. For example, when search displays are masked in the periphery, inspection times increase, indicating a reliance on peripheral vision for scene processing (Van Diepen & d'Ydewalle, 2003). Searching for non-face targets such as objects or letters in real-world scenes is also optimal when peripheral vision is available and decreases when only foveal vision is accessible (Nuthmann & Canas-Bajo, 2022). Peripheral vision is therefore vital for the spatial search for non-face targets and the processing of complex displays.

As faces often appear in areas of the visual field where acuity is low, this must therefore also put limits on the information that can be extracted from these stimuli in the periphery and used for detection. This implies that observers may only perceive basic low-level information (e.g., colour and shape; see Bindemann & Burton, 2009; Burton & Bindemann, 2009; Prunty et al., 2022) to support this process in the periphery, and this may differ from more detailed information in faces that can be extracted at fixation. However, the existing evidence is mixed in this regard. There is evidence, for example, that detection declines as faces are presented at greater eccentricities in the visual field (Melmoth et al., 2000; Rousselet et al., 2005). Additionally, brain responses such as the N170 component and activity in the fusiform face area are stronger when faces are presented centrally compared to the periphery (De Lissa et al., 2019; Levy et al., 2001). This indicates that face detection might rely on detail-specific processing that deteriorates in the periphery. On the other hand, faces appear to hold a detection advantage over other stimuli in the visual periphery. For example, observers are more likely to saccade towards faces than objects when these are presented at opposite ends of the visual field (Crouzet et al., 2010; Little et al., 2021), and faces retain this advantage over objects even at extreme eccentricities in the visual field and in "crowded" conditions (Boucart et al., 2016). This indicates that it is important to dissociate processing in the periphery and at fixation to understand face detection, but still leaves open the question of which facial characteristics drive this process.

In this study, this is examined directly by manipulating the facial characteristics that contribute to the detection process at different locations in the visual field. A constraint of previous studies is that detection performance was typically measured *indirectly*, by recording observers' response times and accuracy (e.g., Bindemann & Burton, 2009; Bindemann & Lewis, 2013; Burton & Bindemann, 2009). Measures of response times and accuracy alone cannot reveal the locus at which specific visual information becomes important for detection, or whether such information is used during the detection stages of face processing at all. For example, when considering differences in detection performance with frontal and profile faces (Bindemann & Lewis, 2013; Burton & Bindemann, 2009; Nevard et al., 2023), it is unclear whether these viewpoint effects arise during the search for potential face candidates in the visual field or during later decision-making stages to determine whether a looked-at stimulus is, in fact, a face.

The current research dissociates these possibilities. As well as obtaining overt measures of detection performance (i.e., RT and accuracy measures), we also examined face detection while observers' eye movements were tracked. This allows for more direct measures of face detection, such as the speed at which a face is first located (i.e., by measuring the time to first fixations on faces), which can be contrasted with observers' response times as an index of when a detection decision has been made. In this way, we can obtain more clarity as to the stages of detection. In the current study, this methodology is exploited further through a new gaze-contingent paradigm. In this paradigm, participants' eye movements are tracked during the search for faces in scenes. The trajectory of their eye movements (i.e., saccades) is then used to manipulate the stimulus display immediately *after* a face has been located but *before* a detection decision is registered. Crucially, these changes to stimuli can be made in milliseconds during a saccade. This method can create changes in a stimulus face between the point at which the face is first noticed outside of foveal vision and when it is fixated (Houtmans & Sanders, 1983). This allows for the independent manipulation of face content during the visual *search* for potential face candidates (i.e., when an eye movement to the location of a face is initiated) and the final *decision* that a now-fixated region really does contain a face.

In this paradigm, three key measures are used to characterize performance. The first measure refers to the speed with which faces are first fixated (i.e., the latency of a saccade to a face) and indicates how quickly potential face candidates are found in scenes. These latencies should be lower for faces that preserve key information for detection in the periphery, and equivalent for faces with and without information that does not contribute to the search process. The second and third measures to characterize face detection performance are *response time*, which reflects the latency of a conscious decision that a face is present or absent in a scene, and the *accuracy* of that conscious decision. These measures serve as an index of the information that drives definitive detection decisions when faces are presented at the fovea, so latencies should be lowest and accuracy highest for faces displaying the most diagnostic information.

A fourth measure also presents additional insight into the facial information that is accessible in the periphery and at the fovea. The gaze-contingent paradigm provides dwell times (i.e., fixation durations) once a face has been fixated. This provides insight into whether the face detection system computes the change in a stimulus from when it was found in the periphery and processed at the fovea. If such changes are noticed, then this should lead to longer dwell times (i.e., fixation durations) once a face has been fixated, to resolve the conflict between the discrepant percept *before* and *after* a saccade was initiated to the location of a face.

These premises are investigated across a series of four experiments. To determine whether changes to faces are perceived in peripheral vision, Experiment 1 used the gaze-contingent paradigm. Participants viewed a series of scenes. For each scene, they had to decide whether or not a face was present. On face-present trials, as participants made a saccade towards a face, it was either removed from the scene, changed in identity or left unchanged. If changes to faces are undetected in peripheral vision, then mid-saccade changes in identity should not affect participants' detection decisions. Indeed, as face detection must operate on *any* face, regardless of who it belongs to, face identity should be irrelevant to face detection. In contrast, removal of a face should cause a decline in performance at the point at which a detection decision is made. In Experiment 2, the presence or absence of internal facial features was manipulated to determine their importance in the detection of faces. If these features do not drive detection, then their removal or addition during a saccade to a face's location should not affect performance. Yet if such features are important for face detection in the periphery, then faces with the internal features initially removed should produce a decline in initial search performance (i.e., increased fixation latency), and if these features are also important to make conscious detection decisions, then their removal should also increase response times and detection accuracy for faces once these are presented at fixation. Finally, Experiments 3 and 4 examine which facial information observers can perceive in the periphery when they are instructed explicitly to categorize intact faces from faces in which internal features are removed or rotated. This provides an important comparison to determine whether any gaze-contingent changes in Experiments 1 and 2 affect face detection because visual information from internal features is not *seen* or because it is not *used*.

Experiment 1

The task in this experiment is to decide as quickly but as accurately as possible whether a face is present in each scene. On half of the trials, a face appears (only) in peripheral vision. By manipulating the characteristics of the face while the participant saccades to it, the facial information that is used to trigger detection and initiate a saccade to its location is investigated, and whether this differs from the information that is available from a face once it is fixated. In Experiment 1, this gaze-contingent paradigm is applied to exchange a face in a scene for the face of another person, during the interval between target detection and fixation. By contrasting this with a condition in which no changes to faces are made, it is possible to examine whether the detection process in the periphery is sensitive to visual information such as identity or relies on more general characteristics that are shared across different faces. In addition, another condition was included in which a peripheral face that was initially present in a scene, was removed as participants saccaded to its location. This condition imposes unusual demands on participants in a detection task, as a face is *both* present and absent in a trial. Here, this condition serves as a baseline for establishing whether gaze-contingent changes are noticed by participants as the stimuli are moved from peripheral to foveal vision.

We predicted that all three conditions should show the same latency to *initiate* a saccade to the face locations, because at the moment the saccade is triggered, all three conditions show the same face. In addition, detection accuracy should be reduced when the face is removed during the saccade to its location, due to the ambiguity this condition creates about the presence of a face. However, if these changes are noticed by participants, then this should increase the duration of fixations on the location from which a face has disappeared, to reflect the processing of this change. Similarly, if observers are sensitive to the change in identity, then this should increase the duration of fixations on faces in this condition, by creating an additional processing load between the initial detection of a face in the periphery and its classification at fixation.

Methods

Participants

Sixty participants completed the experiment (5 male, 54 female, 1 non-binary). A power analysis was conducted with G*Power based on a repeated-measures ANOVA (within factors) with a medium effect size (Cohen's f = .25 / $\eta_p^2 = .06$, power = .95, number of groups = 1, and number of measurements = 3) and an alpha threshold of $p = .05$, leading to a suggested sample size of 43. This was adopted as the minimum sample size, and we advertised the experiment online for 4 weeks, with a maximum sign-up of 60. The final sample of 60 participants represents all sign-ups that occurred in this time period. This sample size was then maintained across all experiments reported here. These participants ranged in age from 18–36 years, with a mean age of 19.5 (*SD* = 2.38). Participants were undergraduate students at the University of Kent who participated for course credit. All experiments in this study were carried out in accordance with the Declaration of Helsinki for experiments involving human participants. All reported normal or corrected-to-normal vision.

Stimuli

As participants' eye movements were recorded during this experiment, the stimuli were displayed using SR-Research Experiment Builder Software (version 1.1.0), with the dimensions of 1024 × 768 pixels, on a 21inch monitor that was connected to an EyeLink 1000 eye-tracking system, running at a 1,000 Hz sampling rate. The face and scene stimuli in the experiments were sourced from previous detection studies (Bindemann et al., 2009; Bindemann & Burton, 2009; Bindemann & Lewis, 2013). The face stimuli consisted of eight full frontal faces, comprising of four males and four females. These faces were cropped to remove any extraneous background, with only the face and hair visible, and sized to a height of 53 pixels at a resolution of 72 pixels per inch (ppi), with some variation in width due to slight differences in face and hair shape. These faces were presented on a set of 72 scenes, which depicted settings such as living rooms, cafes and restaurants. The scenes were sized to 1024 (w) × 768 (h) pixels at a resolution of 72 ppi. Faces would appear in one of six locations in the periphery of the scene. The locations were determined by a 5 (w) × 3 (h) grid, in which faces would appear in the centre of the top, middle and bottom locations to the left and right sides of the grid. Only the outer locations were used to ensure faces were initially detected in peripheral vision, thus necessitating visual search by the participant. Example stimuli are provided in Figure 1.

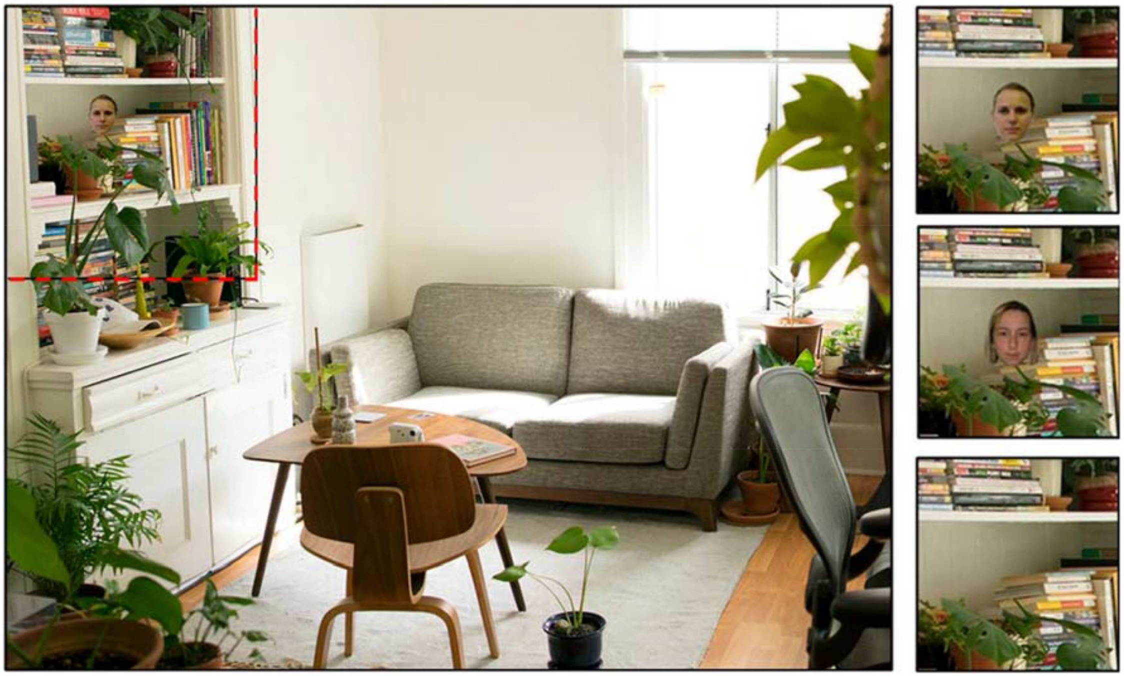

Figure 1. Example of a scene with a face present. The images to the right depict an initial face identity (top), a changed identity (middle) and no face once it has disappeared (bottom). The red dashed line represents the 5 degree VA boundary where changes occurred once that threshold was saccaded to.

The set of 72 scenes was divided into three conditions, corresponding to *change*, *disappear* and *no-change* trials. On change trials, as participants saccaded towards a face, it switched to a different identity. The exchanged faces were matched on general dimensions, such as gender and hair colour. On disappear trials, the face was removed from the scenes as participants saccaded towards it. These visual changes were triggered when saccade trajectories hit a boundary that fell on a face region that spanned approximately 5 degrees of the visual angle (170 pixels) away from the face within the scene (see Figure 1). The size of these regions was calculated using visual angle (VA). The eye-tracking chin rest was positioned 69 cm away from the display screen and central vision is considered to be around 3–5 degrees (Berencsi et al., 2005; Osaka, 1994; Polyak, 1941). Therefore, changes were triggered by eye movements that moved within 5 degrees of VA of the outer height and width dimensions of the faces (corresponding to approximately 6 cm or 170 pixels onscreen). Finally, in the no-change condition, the stimulus display remained the same, regardless of participants' eye movements. In addition, 72 face-absent scenes were also created, in which no faces were present. Of these scenes, 24 were from a new stimulus set, and 48 were repeated from the face-present conditions.

Procedure

In the experiment, each participant's right eye was typically tracked. To track eye movements, participants' heads were stabilized using a chinrest located 69 cm from the monitor. Participants' eye movements were initially calibrated using nine fixation dots on the display monitor. This was then repeated to validate the calibration using the standard EyeLink procedure. Calibration occurred at the beginning of the experiment, and after every 36 trials.

Each trial started with a fixation dot, which was used for drift correction. This dot appeared in the centre of the screen and appeared in-between trials. Once participants fixated this dot, a trial would begin. Participants were told to indicate whether a face was present or absent in the scene by pressing the "F" key for face present and the "J" key for face absent on a standard computer keyboard. Participants were instructed to respond as quickly and as accurately as possible, and were not informed of the changes that would occur.

Each scene was displayed until a response was registered. In the change and disappear conditions, displays would switch with millisecond timing once the participants' saccades hit the five degree boundary, and the second display would remain onscreen until the participant made a response. Trials were presented in four blocks of 36, to give participants breaks and to allow re-calibration. Each participant took part in 144 trials, comprising of 24 trials in the change condition, 24 trials in the disappear condition, 24 trials in the no-change condition, and 72 face-absent trials. The trial order was randomized for each participant.

After the experiment, a questionnaire was administered to determine the extent to which participants were aware of the onscreen changes. The first two questions were open-ended so as not to prime participants and were: "Did you notice anything unusual during the experiment? If so, please describe it." and "Did you notice anything unusual about the faces during the experiment? If so, please describe it." These were then followed by three more specific questions, comprising of "In any of the scenes, did you see faces suddenly disappear?" "In any of the scenes, did you see any of the faces change appearance while you were trying to look at them?" and "In any of the scenes, did you see any faces suddenly appear where there had been none a brief moment ago?" Participants gave written answers to each of the questions and answered each in the presented order.

Results

Detection times and accuracy

The data for all experiments reported in this study can be accessed at https://osf.io/vqe4n/?view_only=9dc572756c0f4e418c880f29cd2f41d6. Five participants were excluded from the analysis due to accuracy scores that fell more than two standard deviations below the sample mean (i.e., scoring below 77.8% accuracy) or reaction times that fell more than two standard deviations above the sample mean (i.e., responses above 3314 ms), leaving a sample of 55 participants for analysis. For the remaining participants, the percentage accuracy scores and the mean correct response times for face-present and face-absent trials were calculated for all conditions. We note that in the disappear condition, faces were both present (at stimulus onset) and absent (by the time the face location was fixated). For the purpose of this experiment, the disappear condition trials were coded as "face-present" trials: i.e., we designated the "correct" response to be to indicate that a face was present, to determine if these faces could clearly be identified in peripheral vision.

In the first step of the analysis, overall performance on face-present trials (i.e., collapsed across face type) and face-absent trials was analyzed to determine whether participants were adhering to task demands. Participants were faster to respond on face-present trials ($M = 1169$ ms, $SD = 215$) than on face-absent trials ($M = 2334$ ms, $SD = 927$), $t(53) = 10.66$, $p < .001$, $d = 1.45$, as the search for faces ends once these stimuli are found. Participants were also less accurate on face-present trials ($M = 81.1\%$, $SD = 8.8$) than on face-absent trials ($M = 97.6\%$, $SD = 3.4$), $t(54) = 12.06$, $p < .001$, $d = 1.63$, indicating that faces were more likely to be missed than falsely detected.

The mean response times for the experimental conditions (change, disappear, no-change) were then analyzed and are provided in Table 1. For reaction times, one participant was additionally excluded from this analysis due to missing data. A repeated-measures ANOVA revealed no effect of experimental condition on reaction times, $F(2,106) = 1.58$, $p = .21$, $\eta_p^2 = .03$. In contrast, a corresponding ANOVA of the number of face-present responses revealed an effect of condition, $F(2,108) = 118.58$, $p < .001$, $\eta_p^2 = .69$. Tukey HSD tests showed that fewer face-present responses were made when faces disappeared compared to the no-change condition, $p < .001$, and compared to when faces changed identity, $p < .001$. In addition, faces were detected more accurately in the change condition than in the no-change condition, $p = .009$.

Table 1. Detection accuracy (%), reaction times (ms), the percentage of trials on which face regions were fixated (%), first fixation times (ms), and fixation duration (ms) for the no-change (N) changed (C) and disappearing (D) conditions of Experiment 1.

	No Change	Change	Disappear	N − C	N − D	C − D
Accuracy	91.6 (7.71)	94.5 (5.38)	57.1 (23.23)	−2.9	34.5	37.4
RTs	1173 (210)	1155 (234)	1197 (301)	18	−25	−42
Trials Fixated	45.0 (27.39)	45.6 (30.95)	16.0 (19.43)	−0.6	29	29.6
First Fixations	800 (243)	773 (193)	573 (179)	27	227	200
Fixation Duration	232 (47)	246 (57)	305 (175)	−14	−73	−59

Parentheses show the Standard Deviation of the Mean.

Eye-tracking data

In the next step of the analysis, participants' eye movements were examined to determine the speed with which faces were first fixated (i.e., saccade latency, as a measure of detection), and the duration of the first fixation (as a measure of face classification). Eye movements were pre-processed by integrating fixations of less than 80 ms with the preceding or following fixation if that fixation lay within half a degree of visual angle (see Rayner & Pollatsek, 1989). The eye fixation data were then compared with a set of regions of interest (ROIs) which reflected the location of faces in the scenes.

The face regions were fixated on 35.8% of all face-present trials, corresponding to 44.5% (SD = 27.3) of no-change trials, 45.3% (SD = 30.7) of change trials, and only 16.4% (SD = 19.4) of disappear trials. A repeated-measures ANOVA on these data revealed an effect of condition, $F(2,108) = 75.68$, $p < .001$, $\eta_p^2 = .58$. Tukey HSD tests showed that the face regions of no-change and change trials were more likely to be fixated than on disappear trials, *ps* < .001. No difference was found in the percentage of fixations between change and no-change faces, $p = .94$.

The time to the first fixation on the face region was then calculated to determine whether experimental condition influenced fixation time. These data are provided in Table 1. A repeated-measures ANOVA revealed an effect of condition, $F(2,80) = 22.59$, $p < .001$, $\eta_p^2 = .36$, as disappearing faces were fixated faster than faces in the no-change and change conditions, both *ps* < .001. In contrast, no difference in fixation times was found between the no-change and change conditions, $p = .40$, which indicates that changes in face identity did not affect detection speed.

Finally, fixation duration on the face ROIs was also calculated. This showed a main effect of condition, $F(2,80) = 16.06$, $p < .001$, $\eta_p^2 = .29$. Tukey HSD tests showed that the locations of disappearing faces were fixated for longer than faces in the change condition, $p = .002$, and the no-change condition, $p < .001$. No difference in fixation duration was found between changing and not-changing faces, $p = .13$.

Questionnaire analysis

The questionnaire was distributed to participants after the experiment to determine their awareness of the changes in their peripheral vision. Open questions were coded as to whether they mentioned faces disappearing, changing or both. When given the open question "Did you notice anything unusual during the experiment?" 65% of participants spontaneously reported that faces had disappeared, while 2% mentioned only the faces changing in appearance, and 9% reported both that faces were disappearing and changing. When prompted directly with the questions "In any of the scenes, did you see faces suddenly disappear?" and "Did you see any of the faces change appearance while you were trying to look at them?" 100% of participants reported that faces had disappeared, while 29% reported seeing faces change. In addition, 38% of participants also falsely believed that faces had suddenly appeared, while 20% reported seeing "flashing" faces.

Discussion

In this experiment, detection was impaired by the disappearance of faces. This reflects the uncertainty in participants that this condition creates, whereby faces are perceived in the visual periphery and trigger eye movements to their locations, but are no longer visible to observers once their location has moved into foveal view. Accordingly, the reduced number of face-present responses in this condition might well reflect how participants construe the task demands of this unusual experimental situation, in which a stimulus disappears mid-saccade and hence disrupts the acquisition of visual information. Notably, the locations of these disappearing faces were also fixated faster than in the other conditions, and participants' verbal reports indicate that awareness of this manipulation was high. We attribute these results to visual transients that are generated by the stimulus offsets, which can draw attention strongly (e.g., Laubrock et al., 2005; Müller & Rabbitt, 1989; Posner & Cohen, 1984).

In contrast, detection was not impaired by a change in identity. If anything, faces in the change condition were detected more accurately than faces in the no-change condition. This finding indicates that a change in identity is not detrimental to detection, presumably because detection utilizes peripheral vision and therefore *precedes* fixation of the faces – and therefore also precedes the change in identity. This reasoning is also supported by the data on time to first fixation (saccade latency), which demonstrate that faces were fixated with comparable speed in both the change and no-change conditions.

The experiment provides additional information that observers were not sensitive to facial identity in the periphery, to yield converging evidence that this does not influence the detection process. Firstly, fixation durations on the face regions were substantially longer in the disappear than the change and no-change conditions, which suggests that observers had some awareness that a face stimulus had been perceived but was no longer present. In contrast, the fixation durations were comparable for the change and no-change conditions, indicating again that a change in identity information was not registered in the periphery. Second, participants were much more likely to report explicitly that faces had disappeared during the experiment than that they had noticed changes in identity.

Taken together, these findings provide initial evidence that face processing should be dissociated in the periphery and at fixation to understand the process of detection, and that this can provide insight into the facial characteristics that drive this process. In this experiment, the removal of a face but not a change in identity impaired detection. This suggests that participants were unaware that a change in identity had occurred. This implies that identity does not interfere with detection during eye movements that reorient facial information from peripheral to foveal vision.

Experiment 2

Experiment 1 indicates that the information used to distinguish different faces, such as identity, does not affect detection performance. This is consistent with detection being a process that must operate on *any* face, regardless of who it belongs to. However, the question remains as to whether internal characteristics that are shared across faces, such as the general configuration of a pair of eyes above a central nose and mouth, are important for detection. We recently demonstrated that detection is impaired when internal facial features are preserved but external features (hair, face outline) were rotated, whereas detection proceeded unhindered when internal features were rotated and external features were preserved (Nevard et al., 2023). This raises the possibility that detection does not rely on either a full face or the internal features despite their importance for other face tasks, such as the recognition of identity and emotion (see, e.g., Ellis et al., 1979; Toseeb et al., 2012; Wegrzyn et al., 2017). However, these experiments relied on indirect measures of detection, such as response times and accuracy.

In the current experiment, the relevance of internal features for face detection is examined directly using the gaze-contingent paradigm introduced in Experiment 1. Thus, the internal features (e.g., eyes, nose, mouth) were either removed as observers saccaded towards a target, to leave a shape-colour face template with intact external features (e.g., only head shape, skin colour and hair); or the internal features were added as observers saccaded towards the target, so that the stimulus changed from a featureless face to an intact face. By comparing these conditions with the detection of intact faces in which no change occurred, the importance of internal and external features *during* the detection process was assessed. If face detection in the periphery is based on a colour-shape template without internal features such as the eyes, nose and mouth, there should be no difference between the intact face condition and the initial presentation of faces without internal features. In contrast, if these internal features support detection, then this process should be impaired when this information is initially absent from the search displays. In contrast, the presence of internal features might facilitate conscious decisions that a foveated stimulus is face, leading to the faster and more accurate classification of intact faces than those in which the internal features have been removed. Finally, the duration of fixations on faces was compared again across conditions, to determine if observers noticed the gaze-contingent changes in facial information between the initial detection of a face in the periphery and its classification at fixation.

Methods

Participants

Sixty participants completed the experiment (8 male, 50 female, 1 gender fluid, 1 non-binary). These participants ranged in age from 18–45 years, with a mean age of 20.52 years ($SD = 4.79$). Participants were undergraduate students at the University of Kent, who participated for course credit. All reported normal or corrected-to-normal vision.

Stimuli & procedure

The stimuli and procedure remained the same as in the previous experiment except for the following changes. Each face stimulus was manipulated in graphics software (Adobe Photoshop) so that the internal features were removed and replaced with skin tones to create featureless faces (see Figure 2). In 24 trials, the full faces would switch to the featureless faces in the *internal-disappear* condition as participants made a saccade towards them. In another 24 trials, the featureless faces would switch to full faces in the *internal-appear* condition. Finally, in another 24 trials, the full faces would stay the same in the *no-change* condition. Participants were instructed to search for a "human face or head" to remove any confusion on seeing a blank featureless face, as some may only categorize a face as a face if internal features are present.

Results

Detection time and accuracy

Six participants were excluded from the analysis due to accuracy scores that fell more than two standard deviations below the sample mean (i.e., scoring below 87.8% accuracy) or reaction times that fell more than two standard deviations above the sample mean (i.e., responses above 3746 ms), leaving a final sample of 56 participants. In the first step of the analysis, overall performance on face-present trials (i.e., collapsed across the experimental conditions) and face-absent trials were analyzed to determine whether participants were adhering to task demands. As in Experiment 1, participants were faster to respond on face-present trials ($M = 1193$ ms, $SD = 211$) than on face-absent trials ($M = 2829$ ms, $SD = 1189$), $t(53) = 11.15$, $p < .001$, $d = 1.52$, as the search for faces ends once a face is found. Again, participants were also less accurate on face-present trials ($M = 94.7\%$, $SD = 4.2$) than face-absent trials ($M = 98.4\%$, $SD = 2.6$), $t(53) = 4.74$, $p < .001$, $d = .65$, indicating that faces are more likely to be missed than falsely detected.

In the next step of the analysis, the mean response times were compared for the experimental conditions (see Table 2). A repeated-measures ANOVA of these data revealed an effect of condition, $F(2,18) = 6.22$, $p = .003$, $\eta_p^2 = .10$, due to slower responses in the internal-disappear than the no-change condition, $p = .002$. No difference was found between the internal-appear condition compared to the disappear condition, $p = .09$, and no-change conditions, $p = .48$. A corresponding ANOVA of accuracy also revealed an effect of condition, $F(2,108) = 9.84$, $p < .001$, η_p^2

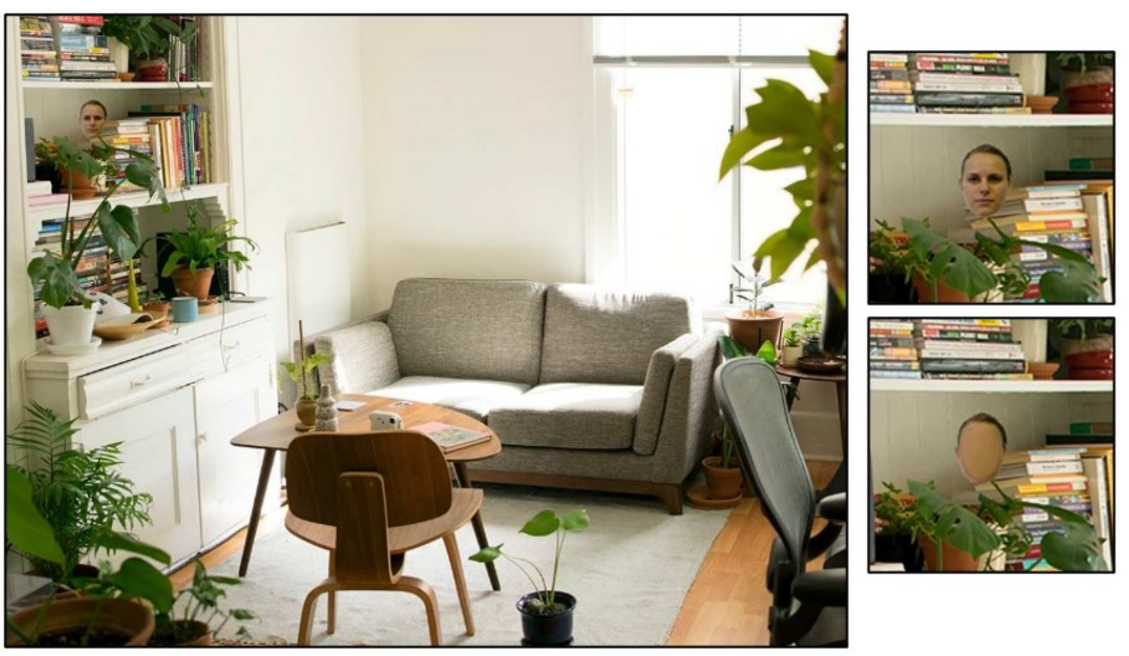

Figure 2. An example of faces where internal features are intact and where the internal features are removed.

Table 2. Detection accuracy (%), Reaction times (*ms*), the percentage of trials on which face regions were fixated (%), first fixation times (*ms*), and fixation duration (ms) for the no-change (*N*), internal-appear (*A*) and internal-disappear (*D*) conditions of Experiment 2.

	No change	Internal appear	Internal disappear	N – A	N – D	A – D
Accuracy	95.5 (5.60)	96.1 (4.50)	92.5 (6.98)	−0.6	3	3.6
RTs	1162 (188)	1186 (254)	1232 (239)	−24	−70	−46
Trials Fixated	50.7 (28.39)	42.90 (34.42)	46.74 (28.97)	8	5.1	−2.9
First Fixations	799 (175)	724 (167)	819 (199)	75	−20	−95
Fixation Duration	230 (56)	225 (57)	240 (65)	5	−10	−15

Parentheses show the Standard Deviation of the Mean.

= .15. Tukey HSD tests showed that accuracy declined in the internal-disappear trials compared to both the no-change and internal-appear conditions, $ps \leq .01$. No difference in accuracy was found between the internal-appear and the no-change conditions, $p = .73$.

Eye-tracking data

Participants' eye-movement data were processed as in Experiment 1. Overall, the face regions were fixated on 50.7% (SD = 28.4) of no-change trials, 42.9% (SD = 34.4) of internal-appear trials, and 46.7% (SD = 29.0) of internal disappear trials. A repeated-measures ANOVA of these data revealed a main effect of condition, $F(2,106) = 7.87$, $p < .001$, $\eta_p^2 = .13$, whereby faces where no-change occurred were more likely to be fixated than faces where the internal features appeared, $p < .001$. No difference was found in the percentage of fixations between internal-disappear faces and internal-appear faces, $p = .14$, or between internal-disappear and no-change faces, $p = .23$.

The time to the first fixation on the face was calculated next to determine whether the experimental conditions influenced how quickly faces were located in the scene displays. A repeated-measures ANOVA revealed an effect of condition on time to first fixation times, $F(2,18) = 5.24$, $p = .007$, $\eta_p^2 = .09$, whereby faces in the internal-disappear condition were fixated more slowly than when internal features appeared during a saccade, $p = .005$. Faces were also fixated faster in the internal-appear than the no-change condition, $p = .04$, while there was no difference in time to first fixations between the internal-disappear and no-change condition, $p = .83$.

The duration of the first fixation on the face regions was also calculated for each condition, as a measure of the decision process to decide that these regions did, in fact, contain a face. A repeated measures ANOVA revealed no effect of condition on fixation duration, $F(2,106) = 2.26$, $p = .11$, $\eta_p^2 = .04$.

Questionnaire analysis

The questionnaire was distributed to participants after the experiment to determine their awareness of the changes to the faces. When given the open question "Did you notice anything unusual during the experiment?" 85% reported not noticing anything, and only 15% mentioned anything to do with the faces changing. However, when prompted directly, 67% of participants reported that faces had disappeared, 79% reported seeing faces change and 24% reported seeing faces appearing.

Discussion

In this experiment, the relevance of internal features for the detection of faces in peripheral vision was examined using the gaze-contingent paradigm introduced in Experiment 1. Thus, the internal features of eyes, nose and mouth were either removed as observers saccaded towards a face to leave only a shape-coloured face template on display, or the shape-coloured face template without internal features was presented and the internal features were added as observers saccaded towards the face. By comparing these conditions with the detection of intact faces in which no such change occurred, the importance of internal and external features *during* the detection process was assessed.

The results demonstrate that the removal of internal features whilst observers are saccading towards a face reduces the speed and accuracy with which these stimuli are classified. In contrast, faces in which the internal features were initially omitted were fixated faster than faces that did not change, and faster than faces that were initially intact but had features removed thereafter. This contrast suggests that two different processes are involved in face detection, and each of these relies on different facial information. The fixation data indicate that internal features are not important for the *initial localization* of a face. This finding appears to make good sense considering the loss of visual acuity in the periphery (see, e.g., Kalloniatis & Luu, 2007; Westheimer, 1965), which would likely reduce the perception of facial detail, such as the eyes, nose and mouth, in the stimuli presented. Indeed, the current results indicate that this localization process might even be hampered by the presence of internal features, as their presence reduced the speed with which faces were fixated. The second process benefits from the presence of internal features and appears to take place closer to fixation. This process might reflect a decision-making process that a located candidate region in a scene is, in fact, a face. As this process draws on information provided by foveal vision, it can use additional visual detail, such as internal facial features, to make accurate decisions.

Experiment 3

Experiment 2 indicates that internal features are not necessary for locating faces in the visual periphery but support decisions that a looked-at face candidate region is, in fact, a face. This raises the question to what extent internal facial features can be processed in the periphery at all in order to drive detection. In Experiment 3, this is examined directly with a visual discrimination paradigm in which observers were presented with pairs of faces, split across the left and right visual field. One of each pair was an intact face. The other was either a featureless face with the internal features removed, or a face in which these features were rotated by 90°. In addition, the difficulty of this task was manipulated in two ways, both of which are important for face detection. First, the eccentricity of these stimuli in the periphery was manipulated by presenting these faces either at 5° or 10° from fixation (as presented in Rousselet et al., 2005), to establish whether sensitivity to featural information is reduced at greater distances. Second, by presenting the faces on blank displays or embedded in scene displays, it is also possible to compare the influence of context on face processing.

The main question of interest here is whether participants can distinguish between the face types by determining in which side of the visual field the intact face is presented. If observers can distinguish between intact faces and faces without internal features in the periphery, then this would suggest that such information is available to the visual system during face detection, even if it is not required to facilitate this process (as in Experiment 2 and Nevard et al., 2023). Moreover, if observers can determine the orientation of these features, then this would indicate that this information can be processed in considerable detail. Thus, this experiment will reveal the extent to which observers are sensitive to both the presence and the organization of facial features in the visual periphery.

Methods

Participants

Sixty participants completed the experiment (5 male, 54 female, 1 non-binary). These participants ranged in age from 18 to 45 years, with a mean age of 19.90 ($SD = 4.06$). Participants were undergraduate students at the University of Kent who participated for course credit. All reported normal or corrected-to-normal vision.

Stimuli

The stimuli consisted of 24 faces in total. Eight full-frontal faces (four male and four female) were shown in three different versions: as intact faces; as featureless faces; and as internally-rotated faces. For the featureless faces, Adobe Photoshop was used to replace the internal features with a uniform patch of skin tone. For the internally-rotated faces, Adobe Photoshop was used to rotate the internal features by 90° with respect to the external face shape As previously, all faces were cropped so that only the face and hair were visible (see Figure 3). The eye-tracking chin rest was always positioned 69 cm away from the display screen, to give a consistent visual angle. Faces were sized so that the height was 3.0° of the visual field, and approximately 1.8° wide, with some variation due to differences in face shape.

Faces were presented on either side of the fixation point, either 5° or 10° from the centre. The gaze contingency was manipulated so that if participants looked away from the fixation cross, the faces would

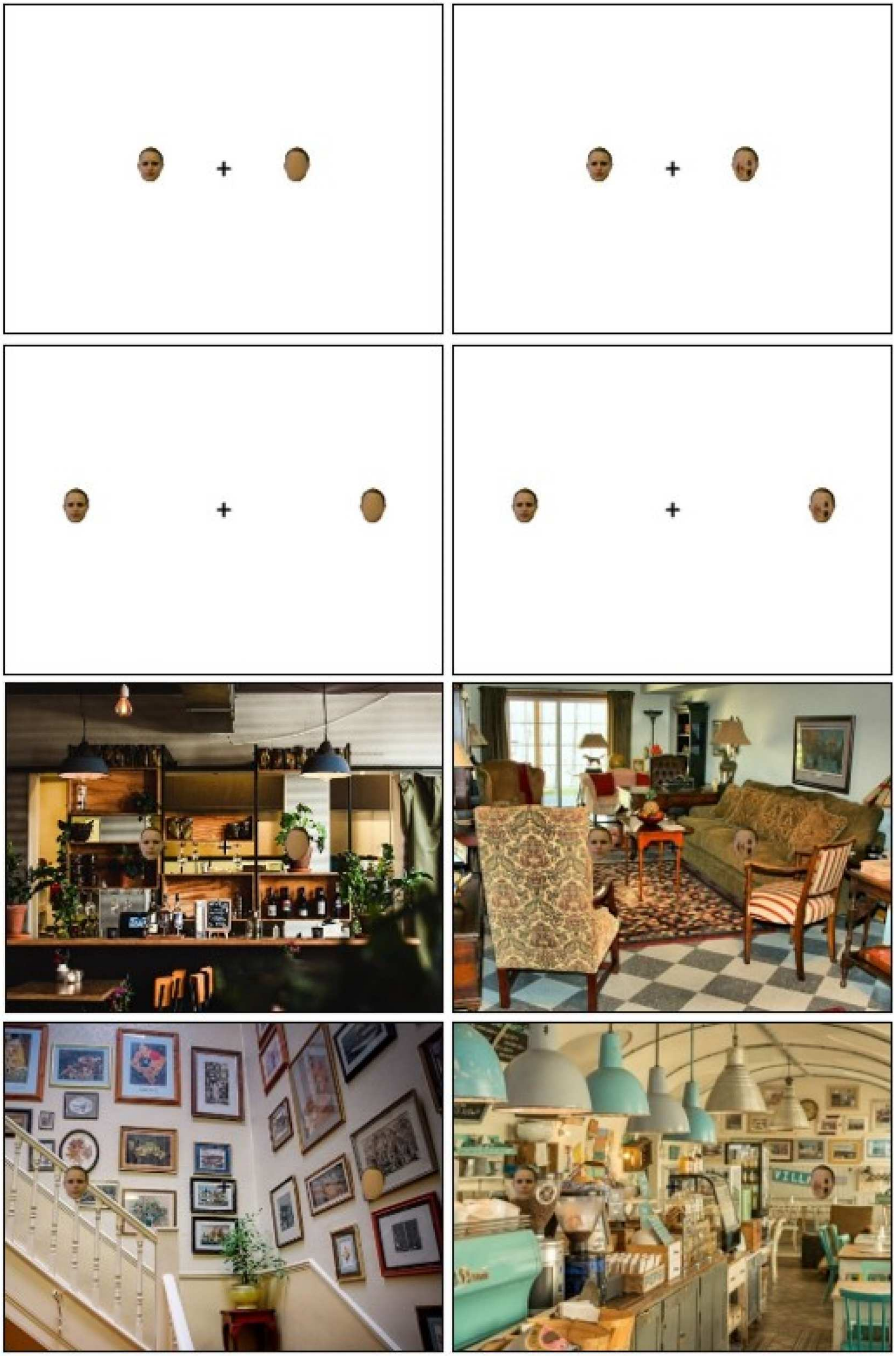

Figure 3. Example stimuli of the featureless face condition (left) and internally rotated face condition (right) presented at 5° (first and third rows) and 10° (second and fourth rows) and on blank (first and second rows) and scene (third and fourth rows) displays.

disappear. This was to ensure that participants were only relying on their peripheral vision to process the faces. The internally rotated or featureless faces were always paired with a full face with matched identity. Over the course of the experiment, each face appeared on both the left and right sides of the fixation point, making 16 trials per condition (see Figure 3).

Faces were also presented in this way on both blank displays and within scenes. The scenes consisted of 64 different images of settings such as living rooms, cafes, restaurants, etc. Scenes were sized to 1024 (w) × 768 (h) pixels at a resolution of 72 ppi, with the faces overlayed on top (for an illustration, see Figure 3).

Procedure

Each trial began with a fixation dot for drift checking, and when fixated, a stimulus display would appear. The display would remain until the participant made a response. Participants were instructed to determine whether the full face appeared on either the left (by pressing the "F" key) or right (by pressing the "J" key) side of the fixation point. Participants were notified of the nature of the other faces and told that if they looked away from the fixation cross, then the faces would disappear. Participants were informed to respond as quickly and accurately as possible. Trials were presented in 2 blocks of 64, either being presented with the blank background trials first or the scene trials. This was counterbalanced across participants. Each participant took part in all 128 trials, and the trial order within each block was randomized for each participant.

Results

Six participants were excluded from the analysis due to either accuracy scores that fell more than two standard deviations below the sample mean (i.e., scoring below 77.8% accuracy) or reaction times that fell more than two standard deviations above the sample mean (i.e., responses above 945 ms), leaving a final sample of 54.

Response times

The percentage accuracy scores and the mean response times were calculated for all conditions and are presented in Table 3. A 2 (face type: featureless, rotated) × 2 (display type: blank, scene) × 2 (eccentricity: 5°, 10°) repeated-measures ANOVA of these data revealed a main effect of face type, $F(1,53) = 171.88$, $p < .001$, $\eta_p^2 = .76$, and an effect of eccentricity, $F(1,53) = 141.51$, $p < .001$, $\eta_p^2 = .73$. These effects were qualified by a two-way interaction between these factors, $F(1,53) = 73.03$, $p < .001$, $\eta_p^2 = .58$. At both distances of 5° and 10°, responses were slower to detect intact faces when these were paired with rotated faces than when they were paired with featureless faces, *ps* < .001. In addition, responses in the featureless and rotated conditions were slower when faces were presented at 10° than at 5° eccentricity, *ps* < .001.

An effect of display type was also found, indicating that face discrimination was faster when faces were presented on blank backgrounds than on scenes, $F(1,53) = 23.45$, $p < .001$, $\eta_p^2 = .31$. There was no interaction between eccentricity and display type, $F(1,53) = 3.02$, $p = .09$, $\eta_p^2 = .05$, between face type and display type, $F(1,53) = 1.37$, $p = .25$, $\eta_p^2 = .03$, or a three-way interaction, $F(1,53) = .03$, $p = .86$, $\eta_p^2 = .00$.

Accuracy

A corresponding ANOVA of response accuracy also revealed main effects of face type, $F(1,53) = 143.49$, $p < .001$, $\eta_p^2 = .73$, eccentricity, $F(1,53) = 213.16$, $p < .001$, $\eta_p^2 = .80$, and display type, $F(1,53) = 12.46$, $p < .001$, $\eta_p^2 = .19$. These effects were qualified by an interaction between eccentricity and face type, $F(1,53) = 123.26$, $p < .001$, $\eta_p^2 = .70$. Tukey HSD test showed that accuracy in the featureless and rotated face conditions was higher at 5° than at 10° eccentricity, *ps* < .001, but intact faces were also discriminated

Table 3. RTs (*ms*) and Accuracy (%) for Featureless (*F*) and Rotated (*R*) Faces in the Blank and Scene Displays, at both 5° and 10° distances in Experiment 3.

		RTs			Accuracy		
		Featureless	Rotated	F – R	Featureless	Rotated	F – R
Blank	5°	530 (102)	595 (113)	−65	99.2 (2.44)	97.9 (5.43)	1.3
	10°	574 (90)	786 (211)	−212	93.3 (1.65)	76.3 (13.21)	17
Scene	5°	573 (119)	658 (150)	−85	99.7 (1.45)	96.8 (6.73)	2.9
	10°	653 (130)	880 (204)	−227	99.2 (2.73)	78.9 (12.87)	20

Parentheses show the Standard Deviation of the mean.

more accurately from featureless than rotated faces at both eccentricities, $p = .05$ for 5° and $p < .001$ for 10°, respectively.

An interaction between eccentricity and display type was also found, $F(1,53) = 20.17$, $p < .001$, $\eta_p^2 = .28$. Faces were discriminated more accurately in both blank and scene displays at 5° than 10° eccentricity, *ps* < .001. At 10°, faces were classified more accurately in scene displays than in blank displays, $p < .001$, however, this was not found at 5°, $p = .77$.

An interaction was also observed between face type and display type, $F(1,53) = 5.07$, $p = .03$, $\eta_p^2 = .09$. Intact faces were distinguished from rotated faces with similar accuracy in blank and scene displays, $p = .89$, whereas featureless faces were distinguished more accurately in scene displays than blank displays, $p < .001$. However, accuracy was generally also higher in the featureless than the rotated face condition, *ps* < .001. Finally, there was no interaction between all factors, $F(1,53) = .60$, $p = .44$, $\eta_p^2 = .01$.

Discussion

This experiment revealed that observers can distinguish the presence and absence of internal features for faces presented in the periphery. This task becomes more challenging as the eccentricity of faces increases from 5° and 10°, though accuracy was generally high (> 93%) and response times were fast (< 660 ms). A different picture emerged when internal features were rotated by 90°. This manipulation increased response times and substantially lowered accuracy to classify faces, to between 76 and 79% across the blank and scene conditions at 10° eccentricity (with a chance level of 50%). Taken together, these findings indicate that observers are sensitive to the *presence* but much less so to the *organization* of facial features in the visual periphery.

These findings have several implications. Since observers can distinguish between intact faces and faces without internal features in the periphery, this suggests that such information is available at some level to the visual system during face detection. However, since the localization of faces is not facilitated by the presence of internal features, as is evident in Experiment 2, the detection process does not appear to depend critically on this information. This could be explained by the second key finding of this experiment, that observers are much less sensitive to the orientation of internal features in the periphery, and that this insensitivity is more pronounced at greater eccentricities.

If the orientation of a basic face schema (e.g., a pair of eyes above a central nose and mouth) cannot be perceived as quickly and as accurately in the periphery as other facial information, then this suggests that these internal features only provide an imprecise detection cue. Indeed, if observers cannot reliably distinguish upright from rotated internal features, then the presence of such features could be potentially misleading, for example, by drawing observers to face-like candidate regions in the visual field that might capture *some* aspects of internal facial features at different rotations, rather than actual face targets. In that case, it would be beneficial to operate with cognitive templates for face detection that do not rely strongly on such featural information. This could explain why the localization process was hampered by the presence of internal features in Experiment 2, where their presence reduced the speed with which faces were fixated. This issue is considered further in the General Discussion.

Experiment 4

Experiment 3 demonstrates that faces in which the internal features have been removed can be discriminated with speed and accuracy from intact upright faces in the periphery, whereas faces in which the internal features are rotated are difficult to distinguish from their intact upright counterparts. However, the experimental design allowed for the unlimited viewing of stimuli, which may have enhanced accuracy in this task by allowing covert attention shifts to the target stimuli to be made. In Experiment 4, this possibility is restricted by limiting the display time of the experimental displays to just 200 milliseconds, to determine the reliability of these findings.

Methods

Participants

Sixty participants completed the experiment (8 male, 50 female, 1 gender fluid, 1 non-binary). These participants ranged in age from 18 to 38 years, with a mean age of 20.10 (*SD* = 3.56). Participants were undergraduate students at the University of Kent who participated for course credit. All reported normal or corrected-to-normal vision.

Stimuli & procedure

The stimuli and procedure remained the same as in the previous experiment except for the following changes. Faces were presented in the periphery for 200 ms before disappearing. This timeframe is based on the time it takes to initiate a saccade (Purves et al., 2001), reducing the opportunity to shift attention to the target faces. The participants would then indicate which side they thought the whole face appeared on, before moving on to the next trial.

Results

Four participants were excluded from the analysis due to accuracy scores that fell more than two standard deviations below the sample mean (i.e., scoring below 69.9% accuracy) or reaction times that fell more than two standard deviations above the sample mean (i.e., responses above 951 ms), leaving a final sample of 56.

Response times

The percentage accuracy scores and the mean response times were calculated for all conditions, and are presented in Table 4. A 2 (face type: featureless, internally rotated) × 2 (display type: blank, scene) × 2 (eccentricity: 5°, 10°) repeated-measures ANOVA was conducted to determine the effect of these factors on reaction times. This revealed main effects of face type, $F(1,55) = 303.79$, $p < .001$, $\eta_p^2 = .85$, eccentricity, $F(1,55) = 114.79$, $p < .001$, $\eta_p^2 = .68$, and display type, $F(1,55) = 26.27$, $p < .001$, $\eta_p^2 = .32$. These effects were qualified by an interaction between eccentricity and face type, $F(1,55) = 39.40$, $p < .001$, $\eta_p^2 = .42$. At both 5° and 10° eccentricity, intact faces were discriminated faster from featureless than rotated faces, *ps* < .001. In addition, response times were faster at 5° than at 10° eccentricity, both in the featureless and the rotated conditions, *ps* < .001.

An interaction between eccentricity and display type was also identified, $F(1,55) = 11.06$, $p = .002$, $\eta_p^2 = .17$. At both 5° and 10° eccentricity, faces were classified faster in blank than in scene displays, *ps* < .001. For both blank displays and scene displays, faces were classified faster at 5° than at 10° eccentricity, *ps* < .001. Finally, no interaction between face type and display type, $F(1,55) = .64$, $p = .43$, $\eta_p^2 = .01$, or between all three factors was found, $F(1,55) = .01$, $p = .91$, $\eta_p^2 = .00$.

Accuracy

A corresponding ANOVA of accuracy also revealed main effects of face type, $F(1,55) = 271.50$, $p < .001$, $\eta_p^2 = .83$, eccentricity, $F(1,55) = 361.41$, $p < .001$, $\eta_p^2 = .87$, and display type, $F(1,55) = 5.33$, $p = .03$, $\eta_p^2 = .09$. These effects were qualified by an interaction between eccentricity and face type, $F(1,55) = 185.73$, $p < .001$, $\eta_p^2 = .77$. At both 5° and 10° eccentricity, accuracy was higher in the featureless than the rotated condition, *ps* < .001. In addition, both featureless and rotated faces were classified more accurately at 5° than at 10° eccentricity, *ps* < .001.

An interaction between face type and display was also found, $F(1,55) = 18.62$, $p < .001$, $\eta_p^2 = .25$. Featureless faces were distinguished from intact faces more accurately in scenes than in blank displays, $p = .03$, whereas rotated faces were distinguished more accurately in blank displays than in scenes, $p = .005$. However, in both blank displays and scenes, featureless faces were distinguished more accurately than rotated faces, *ps* < .001. Finally, no interaction was found between eccentricity and display type, $F(1,55) = 1.69$, $p = .20$, $\eta_p^2 = .03$, or between all three factors, $F(1,55) = 3.34$, $p = .07$, $\eta_p^2 = .06$.

Table 4. RTs (*ms*) and accuracy (%) for featureless and rotated faces in the blank and scene displays, at both 5° and 10° distances in Experiment 4.

		RTs			Accuracy		
		Featureless	Rotated	F – R	Featureless	Rotated	F – R
Blank	5°	527 (105)	624 (111)	−97	99.1 (2.51)	94.2 (8.58)	4.9
	10°	545 (101)	727 (170)	−182	93.4 (2.51)	71.9 (13.59)	21.5
Scene	5°	584 (118)	693 (164)	−109	98.4 (2.98)	90.0 (12.48)	8.4
	10°	641 (146)	839 (257)	−198	96.9 (5.59)	66.7 (10.73)	30.2

Parentheses show the Standard Deviation of the Mean.

Discussion

This experiment replicates the main findings of Experiment 3, by demonstrating that observers can more effectively distinguish the presence and absence of internal facial features in the visual periphery than the orientation of these features. This confirms that observers are sensitive to the presence of facial features but less so to their *organization* when visual acuity is limited in this way. In contrast to Experiment 3, these findings were now observed under conditions in which stimulus display time was limited to just 200 milliseconds, to reduce the possibility of covert attention shifts to these targets. These findings suggest that *some* information about internal features is available during the localization of faces (i.e., their presence), but details such as the orientation of these features or their configuration of two eyes above a nose and mouth are not as accessible. This indicates that the detection process does not depend critically on this information.

The contribution of display background on the discrimination of faces was also considered in Experiments 3 and 4. Both experiments were consistent in their findings that faces are discriminated more efficiently on blank displays. This is consistent with research concluding that detection was easier with decreased display complexity (Nevard et al., 2023). Of significant note, however, is that featureless faces were distinguished more accurately on scene displays than blank displays. This is further evidence that a skin-colour template may be more successful for finding faces in complex displays that represent natural environments. This simple face template seemingly stands out from the detailed display, making it more salient to observers. As this was not found with internal-rotated features, this suggests that such facial detail may be limiting the draw towards the face to an extent, which is consistent with the findings from Experiment 2. This may also account for some of the inconsistencies in face detection research when presenting different faces in different displays, such as the influence of inverting faces, or presenting them in profile view, on blank displays or in scenes (e.g., Bindemann & Lewis, 2013; Nevard et al., 2023).

General discussion

This research introduced a novel gaze-contingent paradigm to dissociate the processes of face detection in the periphery and at fixation over a series of four experiments. In Experiment 1, detection of faces that changed identity between their initial presentation in the periphery and their subsequent fixation, was comparable to detection of faces that did not change. In contrast, response times and face-present responses declined when these stimuli disappeared between peripheral viewing and fixation. In Experiment 2, faces in which the internal features disappeared before they could be fixated, were classified more slowly and less accurately. In contrast, faces that were initially presented without internal features were fixated faster before a detection decision was registered. Finally, in Experiment 3 and 4, featureless faces were distinguished faster and more accurately from intact faces in peripheral vision than faces in which internal features were rotated 90°.

Several insights emerge from these experiments. These experiments suggest that face detection in the periphery is not sensitive to differences between individual faces, such as changes in internal features or holistic information about the overall appearance of a face that is associated with the coding of identity (e.g., Leder & Carbon, 2006; Rhodes et al., 2006; Richler et al., 2009). This is consistent with a process that must operate on *any* face, regardless of who it belongs to. The current experiments go further, to demonstrate that detection is not impaired by the absence of internal features. They also show that this featural information can be *perceived* in peripheral vision, as observers can discriminate intact from featureless faces (Experiment 3 and 4), but is not *used* to facilitate the detection of faces in peripheral vision. At the same time, however, observers could not discriminate well between faces in which the internal features were intact or rotated by 90 degrees. This indicates that they are sensitive to the *presence* of internal features in the periphery, but are not sensitive to their *organization*. In this sense, the detection of internal features in comparison with featureless faces appears to be *non-specific* to faces, but reflects a more general awareness of the presence of *some* visual information in the periphery, compared to its absence. Overall, these findings converge with the previous research, in that faces that preserve a basic colour-shape template can be detected with equal efficiency, regardless of the organization of internal features (Nevard et al., 2023; Pongakkasira & Bindemann, 2015).

The current experiments expand on the previous research by using eye-tracking to dissociate different stages of the detection process. The previous literature identified that faces are detected in peripheral vision (Boucart et al., 2016; Crouzet et al., 2010; Little et al., 2021), and also concluded that manipulations of facial characteristics (such as shape, colour, features configuration) can hinder this process (Bindemann & Burton, 2009; Bindemann & Lewis, 2013; Lewis & Edmonds, 2003; Pongakkasira & Bindemann, 2015; Prunty et al., 2024). What was less clear, however, was the extent to which such features were used *during* detection. Experiment 2 showed that faces with internal features disappearing in peripheral vision were detected less efficiently than intact faces, yet initially-featureless faces that gained features were detected equally, if not more frequently, than intact faces with no changes. This indicates that two processes may be involved in face detection that rely on different information. Firstly, such internal features seem not to be necessary in order to be drawn towards a face stimulus, as initially-featureless faces were detected with equal efficiency to intact, not-changing faces. This is consistent with other research in that such features may not be perceived in the periphery due to limitations in visual acuity (Brown et al., 1997; Melmoth et al., 2000; Rousselet et al., 2005). Secondly, a decision-making process may occur at fixation in order to decide if the stimulus presented matches a cognitive template of a human face. It is during this process that internal features are used, as these can be perceived clearly in foveal vision. This is also reflected in brain activity data in other studies, in which stronger responses to centrally presented faces are attributed to perceived facial detail (Levy et al., 2001).

What is harder to account for in Experiment 2 is the enhanced detection of faces where the internal features were initially absent and then appeared, as well as the finding that observers could more effectively distinguish the presence and absence of internal facial features in the visual periphery than the orientation of these features. One possible explanation for this finding is that featureless faces are easier to detect because of low-level differences that this manipulation might create. For example, whereas faces with intact and rotated internal features might be comparable in terms of mean luminance and visual complexity, featureless faces might have higher luminance or a more even texture that might facilitate their detection.

An intriguing alternative explanation for this finding is that internal features might actually *hinder* the detection process. That is, a basic, skin-coloured template (e.g., a consistent height-to-width ratio, with natural skin tones) may produce a more powerful initial draw than full featural faces (Bindemann & Burton, 2009; Pongakkasira & Bindemann, 2015). As internal features are not a precise face cue in the periphery, it is conceivable that a less complex template is employed so as to not be misled by any "face-like" stimuli. This theorizing draws in part on the phenomenon of face *pareidolia*, whereby non-face objects are sometimes perceived as faces (Keys et al., 2021; Omer et al., 2019; Takahashi & Watanabe, 2015). It is interesting to note that these pareidolia typically reflect the opposite stimulus characteristics that appear to be important in the current experiments, by capturing internal "features" such as two high contrast dots and a horizontal line that might represent a pair of eyes and a mouth (Omer et al., 2019), whereas a typical face outline and skin-colour tones are usually absent. In addition, experiments on the perception of pareidolia typically present these stimuli at or near fixation (Omer et al., 2019; Takahashi & Watanabe, 2015), and when the task requires search for a target in visual arrays, human faces outperformed these face-like objects (Collyer et al., 2024; Keys et al., 2021). Therefore, if our visual system were to be drawn by "face-like" pareidolic stimuli, this could potentially distract us away from the visual information most useful to detect *actual* faces. Internal features might therefore be useful for detection in the absence of other facial information but may not be essential when typical face stimuli are employed. Finally, internal facial features are also subject to skew or distortion when faces appear in peripheral vision at 6°–8° (Balas & Pearson, 2019; Bowden et al., 2019; Tangen et al., 2011). Collectively, the hindrance of internal features in the current research, the evidence that "face-like" internal feature stimuli are detected less efficiently than human faces, and their distortion in the periphery suggests that these features are not the most efficient candidates for face detection. Consequently, the intriguing possibility exists that a colour-shape template, based on circular skin-coloured tones but *without* internal features, might be a better candidate for finding human faces in the environment.

The differentiation of detection processes goes some way in explaining the inconsistent findings from previous studies in the field, with the potential to clarify some of the debate over the contributions of internal and external features in face detection. Manipulations of facial features, such as blurring, obscuring, reconfiguring internal features, or inverting and rotating external features, have given insight into their potential contributions to detection (e.g., Hershler & Hochstein, 2005; Lewis & Edmonds, 2003; Prunty et al., 2024). All of the aforementioned manipulations decrease detection performance, but there is a lack of clarity on which, if any, features are essential for this process. Internal feature configuration does not make the same contribution to detection compared to external configuration (Nevard et al., 2023), but this does not account for why certain manipulations, such as scrambling features, are still found to hinder detection. The current research provides a clearer explanation for this effect, by suggesting that (at least) two distinct processes occur during detection, using different templates. When visual acuity is low in the periphery, observers are not sensitive to the configuration of finer facial details such as internal features, and therefore these details are not used in peripheral detection. Yet, once a face is located and can be viewed in the high-acuity regions of central vision, the additional detail of internal features can be used, and a feature-specific template is used during the decision-making stage in order to classify faces. Therefore, unless both templates are activated by an external stimulus (i.e., basic colour-shape for initial detection or a feature configuration for classification), detection will deteriorate at one step of this process. This starts to build a coherent theory that accounts for diminished performance rates when either height-to-width ratios and colour are manipulated (Bindemann & Burton, 2009; Pongakkasira & Bindemann, 2015), or when internal features are manipulated (Lewis & Edmonds, 2003; Prunty et al., 2024), and why detection is sometimes optimized when both sources of information are intact (Hershler & Hochstein, 2005).

In conclusion, this study shows that the presence of internal features (eyes, nose, mouth) does not facilitate the search for faces in peripheral vision, whereas the classification of faces in foveal vision is enhanced by these internal features. Moreover, while the presence of internal features in peripheral faces could be perceived when observers were instructed to do so, the organization of these features was difficult to extract. These findings point to a two-stage model of face detection. In peripheral vision, where visual acuity is low, observers are not sensitive to internal facial features and their configuration in the early stages of detection. Once faces are located and can be viewed in high-acuity regions of central vision, a second processing stage may use additional detail from internal features to confirm that a face has been detected.

Authorship contributions

Alice Nevard – Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Visualization; Roles/Writing – original draft; Writing – review & editing. Markus Bindemann – Conceptualization; Funding acquisition; Methodology; Visualization; Roles/Writing – original draft; Writing – review & editing. Graham Hole – Conceptualization; Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Supplementary material

The data for all experiments reported in this study can be accessed at: https://osf.io/vqe4n/?view_only=9dc572756c0f4e418c880f29cd2f41d6.

ORCID

Alice Nevard http://orcid.org/0000-0001-6234-6128

References

Balas, B., & Pearson, H. (2019). The flashed face distortion effect does not depend on face-specific mechanisms. *Scientific Reports*, *9*(1), 1612. https://doi.org/10.1038/s41598-018-37991-9

Berencsi, A., Ishihara, M., & Imanaka, K. (2005). The functional role of central and peripheral vision in the control of posture. *Human Movement Science*, *24*(5-6), 689–709. https://doi.org/10.1016/j.humov.2005.10.014

Bindemann, M., & Burton, A. M. (2009). The role of color in human face detection. *Cognitive Science*, *33*(6), 1144–1156. https://doi.org/10.1111/j.1551-6709.2009.01035.x

Bindemann, M., Burton, A. M., Langton, S. R., Schweinberger, S. R., & Doherty, M. J. (2007). The control of attention to faces. *Journal of Vision*, *7*(10), 15–15. https://doi.org/10.1167/7.10.15

Bindemann, M., & Lewis, M. B. (2013). Face detection differs from categorization: Evidence from visual search in natural scenes. *Psychonomic Bulletin & Review*, *20*(6), 1140–1145. https://doi.org/10.3758/s13423-013-0445-9

Bindemann, M., Scheepers, C., & Burton, A. M. (2009). Viewpoint and centre of gravity affect eye movements to human faces. *Journal of Vision*, *9*(2), 7, 1-16. https://doi.org/10.1167/9.2.7

Boucart, M., Lenoble, Q., Quettelart, J., Szaffarczyk, S., Despretz, P., & Thorpe, S. J. (2016). Finding faces, animals, and vehicles in far peripheral vision. *Journal of Vision*, *16*(2), 10–10. https://doi.org/10.1167/16.2.10

Bowden, J., Whitaker, D., & Dunn, M. J. (2019). The role of peripheral vision in the flashed face distortion effect. *Perception*, *48*(1), 93–101. https://doi.org/10.1177/0301006618817419

Brown, V., Huey, D., & Findlay, J. M. (1997). Face detection in peripheral vision: Do faces pop out? *Perception*, *26*(12), 1555–1570. https://doi.org/10.1068/p261555

Burton, A. M., & Bindemann, M. (2009). The role of view in human face detection. *Vision Research*, *49*(15), 2026–2036. https://doi.org/10.1016/j.visres.2009.05.012

Collyer, L., Ireland, J., & Susilo, T. (2024). A limited visual search advantage for illusory faces. *Attention, Perception, & Psychophysics*, *86*(3), 1–14. https://doi.org/10.3758/s13414-023-02833-y

Crouzet, S. M., Kirchner, H., & Thorpe, S. J. (2010). Fast saccades toward faces: Face detection in just 100 ms. *Journal of Vision*, *10*(4), 16–16. https://doi.org/10.1167/10.4.16

De Lissa, P., McArthur, G., Hawelka, S., Palermo, R., Mahajan, Y., Degno, F., & Hutzler, F. (2019). Peripheral preview abolishes N170 face-sensitivity at fixation: Using fixation-related potentials to investigate dynamic face processing. *Visual Cognition*, *27*(9-10), 740–759. https://doi.org/10.1080/13506285.2019.1676855

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception*, *8*(4), 431–439. https://doi.org/10.1068/p080431

Hershler, O., & Hochstein, S. (2005). At first sight: A high-level pop out effect for faces. *Vision Research*, *45*(13), 1707–1724. https://doi.org/10.1016/j.visres.2004.12.021

Hole, G. J., George, P. A., Eaves, K., & Rasek, A. (2002). Effects of geometric distortions on face-recognition performance. *Perception*, *31*(10), 1221–1240. doi:doi.org/10.1068p3252

Houtmans, M. J., & Sanders, A. F. (1983). Is information acquisition during large saccades possible? *Bulletin of the Psychonomic Society*, *21*(2), 127–130. https://doi.org/10.3758/BF03329974

Kalloniatis, M., & Luu, C. (2007). Visual acuity. *Webvision: The Organization of the Retina and Visual System* [Internet].

Keys, R. T., Taubert, J., & Wardle, S. G. (2021). A visual search advantage for illusory faces in objects. *Attention, Perception, & Psychophysics*, *83*(5), 1942–1953. https://doi.org/10.3758/s13414-021-02267-4

Laubrock, J., Engbert, R., & Kliegl, R. (2005). Microsaccade dynamics during covert attention. *Vision Research*, *45*(6), 721–730. https://doi.org/10.1016/j.visres.2004.09.029

Leder, H., & Carbon, C. C. (2006). Face-specific configural processing of relational information. *British Journal of Psychology*, *97*(1), 19–29. https://doi.org/10.1348/000712605X54794

Levy, I., Hasson, U., Avidan, G., Hendler, T., & Malach, R. (2001). Center – periphery organization of human object areas. *Nature Neuroscience*, *4*(5), 533–539. https://doi.org/10.1038/87490

Lewis, M. B., & Edmonds, A. J. (2003). Face detection: Mapping human performance. *Perception*, *32*(8), 903–920. https://doi.org/10.1068/p5007

Little, Z., Jenkins, D., & Susilo, T. (2021). Fast saccades towards faces are robust to orientation inversion and contrast negation. *Vision Research*, *185*, 9–16. https://doi.org/10.1016/j.visres.2021.03.009

Melmoth, D. R., Kukkonen, H. T., Mäkelä, P. K., & Rovamo, J. M. (2000). The effect of contrast and size scaling on face perception in foveal and extrafoveal vision. *Investigative Ophthalmology & Visual Science*, *41*(9), 2811–2819.

Moscovitch, M., & Moscovitch, D. A. (2000). Super face-inversion effects for isolated internal or external features, and for fractured faces. *Cognitive Neuropsychology*, *17*(1-3), 201–219. https://doi.org/10.1080/026432900380571

Müller, H. J., & Rabbitt, P. M. (1989). Reflexive and voluntary orienting of visual attention: Time course of activation and resistance to interruption. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(2), 315–330. https://doi.org/10.1037/0096-1523.15.2.315

Nachson, I., & Shechory, M. (2002). Effect of inversion on the recognition of external and internal facial features. *Acta Psychologica*, *109*(3), 227–238. https://doi.org/10.1016/S0001-6918(01)00058-0

Nevard, A., Hole, G. J., Prunty, J. E., & Bindemann, M. (2023). Understanding face detection with visual arrays and real-world scenes. *Visual Cognition*, *31*(5), 390–408. https://doi.org/10.1080/13506285.2023.2277475

Nuthmann, A., & Canas-Bajo, T. (2022). Visual search in naturalistic scenes from foveal to peripheral vision: A comparison between dynamic and static displays. *Journal of Vision*, *22*(1), 10–10. https://doi.org/10.1167/jov.22.1.10

Omer, Y., Sapir, R., Hatuka, Y., & Yovel, G. (2019). What is a face? Critical features for face detection. *Perception*, *48*(5), 437–446. https://doi.org/10.1177/0301006619838734

Osaka, N. (1994). Peripheral vision. In T. Oyama, S. Imai, & T. Wake (Eds.), *New sensation-perception psychology handbook* (pp. 923–930). Tokyo: Seishin-shobou.

Polyak, S. L. (1941). *The retina*. University of Chicago Press.

Pongakkasira, K., & Bindemann, M. (2015). The shape of the face template: Geometric distortions of faces and their detection in natural scenes. *Vision Research*, *109*, 99–106. https://doi.org/10.1016/j.visres.2015.02.008

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. *Attention and Performance X: Control of Language Processes*, *32*, 531–556.

Prunty, J., Jenkins, R., Qarooni, R., & Bindemann, M. (2024). A cognitive template for human face detection. *Cognition*, *249*, 105792. https://doi.org/10.1016/j.cognition.2024.105792

Prunty, J. E., Jenkins, R., Qarooni, R., & Bindemann, M. (2022). Ingroup and outgroup differences in face detection. *British Journal of Psychology*, *144*(1), 94–111. https://doi.org/10.1111/bjop.12588

Purves, D., Augustine, G. J., Fitzpatrick, D., Katz, L. C., LaMantia, A. S., McNamara, J. O., & Williams, S. M. (Eds.). (2001). Types of eye movements and their functions. In *Neuroscience* 2nd ed., (pp. 361–390). Sunderland, MA: Sinauer Associates.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Prentice Hall.

Rhodes, G., Hayward, W. G., & Winkler, C. (2006). Expert face coding: Configural and component coding of own-race and other-race faces. *Psychonomic Bulletin & Review*, *13*(3), 499–505. https://doi.org/10.3758/BF03193876

Richler, J. J., Mack, M. L., Gauthier, I., & Palmeri, T. J. (2009). Holistic processing of faces happens at a glance. *Vision Research*, *49*(23), 2856–2861. https://doi.org/10.1016/j.visres.2009.08.025

Rousselet, G. A., Husk, J. S., Bennett, P. J., & Sekuler, A. B. (2005). Spatial scaling factors explain eccentricity effects on face ERPs. *Journal of Vision*, *5*(10), 1–1. https://doi.org/10.1167/5.10.1

Takahashi, K., & Watanabe, K. (2015). Seeing objects as faces enhances object detection. *i-Perception*, *6*(5), 2041669515606007. https://doi.org/10.1177/2041669515606007

Tangen, J. M., Murphy, S. C., & Thompson, M. B. (2011). Flashed face distortion effect: Grotesque faces from relative spaces. *Perception*, *40*(5), 628–630. https://doi.org/10.1068/p6968

Toseeb, U., Keeble, D. R., & Bryant, E. J. (2012). The significance of hair for face recognition. *PLoS One*, *7*(3), e34144. https://doi.org/10.1371/journal.pone.0034144

Van Diepen, P., & d'Ydewalle, G. (2003). Early peripheral and foveal processing in fixations during scene perception. *Visual Cognition*, *10*(1), 79–100. https://doi.org/10.1080/713756668

Wegrzyn, M., Vogt, M., Kireclioglu, B., Schneider, J., & Kissler, J. (2017). Mapping the emotional face. How individual face parts contribute to successful emotion recognition. *PLoS One*, *12*(5), e0177239. https://doi.org/10.1371/journal.pone.0177239

Westheimer, G. (1965). Visual acuity. *Annual Review of Psychology*, *16*(1), 359–380. https://doi.org/10.1146/annurev.ps.16.020165.002043

Young, A. W., Hay, D. C., McWeeny, K. H., Flude, B. M., & Ellis, A. W. (1985). Matching familiar and unfamiliar faces on internal and external features. *Perception*, 14(6), 737–746. https://doi.org/10.1068/p140737.