One object with two identities: The rapid detection of face pareidolia in face and food detection tasks

Greta Stuart1,2, Blake W. Saurels1, Amanda K. Robinson2\*, Jessica Taubert1\*.

1. The University of Queensland, School of Psychology. 2. The University of Queensland, Queensland Brain Institute.

\*Equally contributing authors

Corresponding author: Greta Stuart, [g.stuart@uq.edu.au](mailto:g.stuart@uq.edu.au)

Abstract

Humans are so sensitive to faces and face-like patterns in the environment that sometimes we mistakenly see a face where none exists – a common illusion called “face pareidolia”. Examples of face pareidolia, “illusory faces”, occur in everyday objects such as trees and food, and contain two identities: an illusory face, and an object. In this study, we studied illusory faces in a Rapid Serial Visual Presentation (RSVP) paradigm over three experiments to explore the detectability of illusory faces under various task conditions and processing speeds. The first experiment revealed the rapid and reliable detection of illusory faces even at an extremely limited processing time of 34 ms, suggesting that face pareidolia arises from an error in rapidly detecting faces. Experiment 2 demonstrated that illusory facial structures within food items did not interfere with the recognition of the object's real identity, affirming that examples of face pareidolia maintain their object identity. Experiment 3 directly compared behavioural responses to illusory faces under different task demands, indicating that when processing time is extended, the object identity dominates perception. From a behavioural context, the findings revealed that illusory faces maintain concurrent identities as both faces and objects. The study sheds light on the brain's processing of objects with dual identities, contributing to our understanding of visual stimulus recognition. Future research could explore the neural representation of these unique stimuli under varying circumstances and attentional demands, providing deeper insights into the perception of face pareidolia.

Keywords: illusory faces, RSVP, face detection

This research was supported by the Australian Research Council (FT200100843 to J.T. and DE200101159 to A.K.R).

1. Introduction

Face detection mechanisms are thought to emerge very early in human development (Goren et al., 1975; Tsao & Livingstone, 2008) and are shared with other primates (Taubert et al., 2017). We are so sensitive to faces and face-like patterns in the environment that sometimes we mistakenly see a face where none exists – a common illusion called “face pareidolia”. Examples of face pareidolia, herein referred to as “illusory faces”, frequently occur in ordinary, inanimate objects such as tree trunks, cars, and food items such as sliced capsicums and cups of coffee (see Figure 1A). Studies measuring time-resolved neural signals have suggested that illusory faces are initially processed in similar ways to real faces, engaging face detection mechanisms in the human and macaque brain (Decramer et al., 2021; Wardle et al., 2020; Rekow et al., 2022; Taubert et al., 2022). This error of face detection is then quickly corrected by the brain and illusory faces are eventually encoded as objects (Wardle et al., 2020, Decramer et al., 2021; Taubert et al., 2020). In this study we investigated whether these two apparently distinct stages of neural processing are reflected in human behaviour.

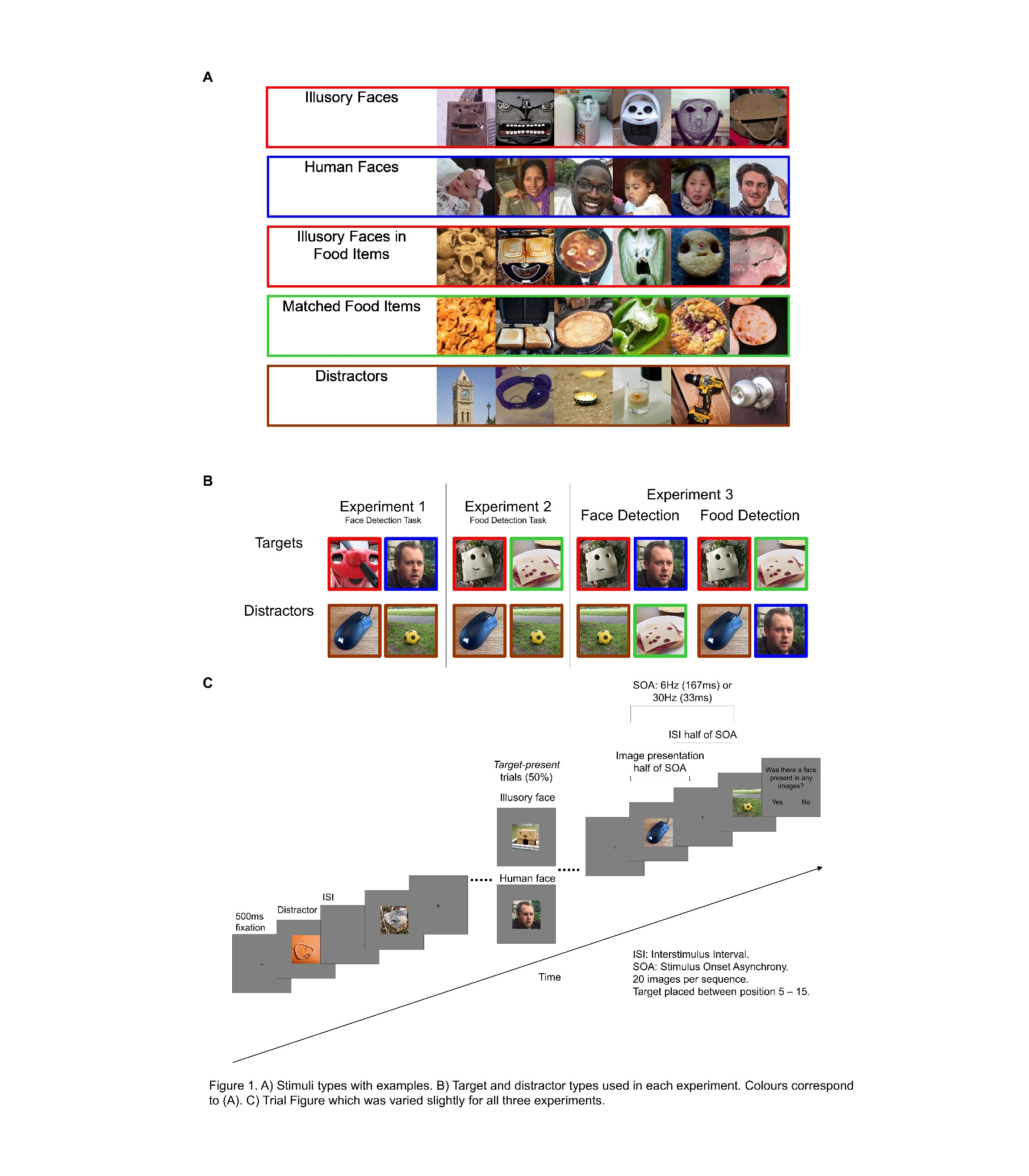
From a behavioural perspective, it remains unclear whether our perception of illusory faces arises from our capacity to rapidly detect faces and face-like patterns (Keys et al., 2021) or our capacity to reinterpret visual ambiguity as being “face-like” (Liu et al., 2014). In previous tasks designed to measure how similar illusory faces are to human faces, researchers have often instructed the participants to explicitly look for facial structure (Wardle et al., 2022; Omer et al., 2019; Uchiyama et al., 2012; Ephihova et al., 2022; Taubert et al., 2023). In other behavioural experiments the participants have been tasked with judging facial attributes such as expressions, gaze direction or age (Palmer & Clifford, 2020; Wardle et al., 2022; Alais et al., 2021). These approaches might have inadvertently drawn attention towards any face-like cues available within the larger image - which typically include their natural backgrounds - that could be interpreted as face-like, and therefore encouraged a particular response profile. Further, perception of the illusory face as an object has never been probed. Therefore, our aim was to measure the detectability of illusory faces under different task and attentional demands. To this end, we used a Rapid Serial Visual Presentation (RSVP) paradigm to determine whether illusory faces could be accurately detected in a speeded detection task based on either their face (Experiment 1 and 3) or object (Experiment 2 and 3) identity.

In this study addressed three specific questions: First, can illusory faces be detected in a RSVP face detection task (Experiment 1)? Secondly, can illusory faces found in food items be detected in an RSVP food detection task (Experiment 2)? And finally, are illusory faces easier to detect as faces or food, and how does this vary with degree of image processing (Experiment 3)?We compare behavioural performance in face and food detection tasks because both food items and faces form separate classes of visual stimuli that are biologically relevant to humans. In other words, being able to recognize the things we can eat and the people we can talk to could be considered central to our survival (Jain et al., 2023; Khosla et al. 2022). Further, images of food, like images of faces, form a class of visual stimuli that are known to drive activity in multiple brain areas (Avery et al., 2021, van der Laan et al., 2011), including stimulus-selective visual areas in the fusiform gyrus (Jain et al., 2023; Kanwisher et al., 1997). Given the parallels between food and faces, and the observation that face pareidolia occurs often in food items (Wardle et al., 2022), we chose food as an appropriate target category for Experiments 2 and 3.

For both the face and object detection tasks (Experiments 1 - 3) we used a RSVP paradigm where, on every trial, we presented a sequence of twenty naturalistic images in quick succession. In Experiment 1 participants completed a face detection task; the participants were instructed to press a button if they detected a face, which was present on 50% of trials. In this experiment, half of the face targets were real human faces and the other half were illusory faces. Previous work has indicated faces are processed optimally at 6Hz (Alonso-Prieto et al., 2013; Gentile & Rossion, 2014), but other work has shown participants are able to derive meaning from images in as little as 13ms per image (approximately 75Hz; Potter et al. (2014). Therefore, we included two presentation speeds in the experimental design, 6 Hz (167 ms) and 30 Hz (33 ms), designed to bias processing to later and earlier stages of processing, respectively (Collins, Robinson & Behrmann, 2018; McKeeff, Remus & Tong, 2007). If the experience of face pareidolia emerges from rapid errors of face detection, rather than the slow reinterpretation of the visual ambiguity, then we expected that participants would be able to detect illusory faces as faces even at high presentation speeds (30 Hz).

In Experiment 2 participants completed a food detection task; the participants were instructed to press a button if they detected food. In this experiment, half of the targets were food items and the other half were food items with illusory faces. Thus, we were able to test whether the object identity of illusory faces was spared at high presentation speeds, when image processing was limited. Further, if face pareidolia arises from an error of rapid face detection, then the expectation was that illusory faces would be detected more easily than food items without illusory faces because, even when task irrelevant, their face identity would summon attentional resources (Jakobsen et al. 2023).

In Experiment 3, we combined the face- and food-detection tasks in a blocked, repeated measures experiment to distil any task-based differences in the detection of illusory faces. We ensured that the only difference between the two task blocks was the instructions given to the participants about what constituted a target. The same illusory faces were, thus, targets in both the face and food detection tasks, allowing us to directly compare behaviour towards these images under different task demands. Assuming that face pareidolia is the consequence of erroneous face detection that occurs before the brain encodes the object identity (Wardle et al., 2020), we expected that illusory faces would be more easily detected in face detection task than the food detection task, when visual processing was limited (30Hz). Whereas, in the corresponding 6Hz condition, we expected that the exact same illusory face stimuli would be easier to detect when participants were looking for food, than when they were looking for faces because the low-level visual cues in illusory faces are more consistent with the object identity than the face identity (Keys, et al., 2021).

Figure 1. Experimental design. (COLOUR)

*Note*. A) Example stimuli for each face, food and distractor condition used in Experiments 1-3. B) Specific examples of targets and distractors for each experiment. Colours correspond to categories in A. C) Example trial timeline for Experiment 1. Images were presented in succession with a blank gap between each image. Presentation rate was 6Hz or 30Hz. A face (Experiments 1 and 3) or food target (Experiments 2 and 3) was present on 50% of trials, displayed between distractor images.

2. Experiment 1

2.1. Method

Experiment 1 was designed to probe face detection performance for real human and illusory faces during rapid image presentation. Images were shown at two different presentation rates to assess how the degree of image processing influenced face detection for human and illusory faces.

2.1.1

*Participants*

A power analysis was run on pilot data to determine the sample size.Data from the first participant was analysed in a one-way repeated measures Analysis of Variance (ANOVA) using trial as the unit of analysis. The observed effect size for F (η²p = 0.32) was used to compute the required sample. Accordingly, twenty-seven participants (*Mage*= 23.6 years, 13 male) were recruited in this study. All participants gave consent to participate and reported normal or corrected-to-normal vision. Participants were recruited from a paid participant pool (the School of Psychology Research Participation Scheme, SONA) and received a $20 voucher for their time. All participants were debriefed at the conclusion of the experiment.

2.1.2

*Stimuli*

The stimulus set included 300 distractor images (ordinary objects with no faces), 100 face pareidolia images, and 100 human face images. All stimuli were obtained from Google Images or Pinterest and were uploaded to Open Science Framework (<https://osf.io/hvuwd/?view_only=b6fb73936179491c89765505bdc2b9c9>). For the distractor images, the focal objects were in naturalistic settings with complex backgrounds to match the conditions the illusory faces were found in. All 500 images in the stimulus set were cropped to be 400 x 400 pixels in size and were displayed on a grey background at a size of 4.9 x 4.9 degrees of visual angle.

2.1.3

*Design*

The experimental protocol included 400 target present trials and 400 target absent trials. The experimental manipulations included both *target type* (human face, illusory face), and *frequency* (6Hz and 30Hz).

2.1.4

*Procedure*

Before beginning the experiment, participants were shown examples of a *human face,* an *illusory face,* and a *distractor object.* The examples shown were not part of the experimental stimulus set. Participants were instructed to detect face images within a sequence of distractor object images. Participants were told both human faces and illusory faces counted as “faces” in the experiment.

During the experiment, participants were required to focus on the centred fixation cross at the beginning and for the duration of each trial. After an initial fixation period (500 ms), a sequence of 20 images were rapidly presented in the centre of the screen, behind the fixation cross. The images were presented at one of two frequencies: 6 Hz (every 166.67 ms) or 30 Hz (every 33.33 ms). A blank interstimulus interval (ISI) was displayed between each image.

A target - an illusory face or a real human face - was presented on 50% of trials (*target present* vs *target absent*). Targets could appear between the 5th and 15th distractor image. This constraint was to reduce the effects of primacy and recency on performance. At the conclusion of each trial, participants were asked (with text on the screen): “Was there a face present in any images?”, with answer options: “yes” or “no.” If participants answered “yes,” a secondary question appeared asking, “What type of face did you see?” either an “illusory face” or “real face.” Participants responded using a handheld 5-button response box (RESPONSEPixx with custom white buttons, https://vpixx.com/products/handheld/) using the left and right buttons to indicate “yes” or “no.” No feedback regarding performance was given. Target absent trials were identical to target present trials except no face stimulus was presented.

There were 800 trials in total, and targets were present in 50% of trials. Targets were presented twice each, once in a 6Hz trial and once in a 30Hz trial. The order of images in each trial was randomised, and the order of trials was randomised for each participant (see Figure 1C for trial structure).

2.1.5

*Apparatus*

The experiment was run using MATLAB R2021b ([www.mathworks.com](http://www.mathworks.com)) and the PTB toolbox (<http://psychtoolbox.org/>). The stimuli was presented on a CRT replacement monitor purchased from Vpixx technologies (<https://vpixx.com/products/viewpixx-eeg/>; Model number: VPX-VPX-2006B). The resolution of the monitor was set to 1920 x 1080 pixels and its refresh rate was 60Hz. Participants viewed the trials 720mm away from the monitor with their head placed on a chin rest.

2.1.6

*Analysis*

For every condition per participant, we computed a sensitivity index (d’) for each of the four unique experimental conditions because a valid measure of “detection performance” needs to take into account both hit rate and false alarm rate. Hit rate was calculated as the percent of correctly detected targets in each condition (6Hz human face, 6Hz illusory face, 30Hz human face, 30Hz illusory face). False alarm rate was calculated using percent of non-target trials per frequency (6Hz and 30Hz) for which participants reported a target was present. Notably, the same false alarm rate was used for the human and illusory faces per frequency condition. The d’ value is undefined when false alarm or hit rates are equal to 0% or 100%. Therefore, for participants with zero false alarms, the false alarm rate was set to 0.5%, and for participants with 100% accuracy in target present trials, the hit rate was set to 99.5%.

2.2. Results

Analysis of the target detection sensitivity showed performance varied across target type and presentation frequency (Figure 2A). A 2 (frequency: 6Hz, 30Hz) X 2 (target type: human face, illusory face) Repeated Measures Analysis of Variance (ANOVA) revealed a main effect of frequency, *F*(1, 26) = 138.40, p < .001, *η²p* = 0.842, indicating that the participants were better at detecting target faces in the 6Hz condition (*M* = 2.6, *SD* = 0.8) compared to the 30Hz condition (*M* = 1.4, *SD* = 0.6). There was a significant main effect of target type, indicating that the participants were better at detecting human faces (*M* = 2.7, *SD* = 0.7) than illusory faces (*M* = 1.3, *SD* = 0.6). There was a significant interaction between frequency and target type, *F*(1, 26) = 8.71, *p* < .001, *η²p* = .007.  Paired sample t-tests were conducted to compare detection sensitivity for human and illusory faces at each presentation frequency. In 6Hz trials, participants were significantly better at detecting human faces (*M* = 3.4, *SD* = 0.9) compared to illusory faces (*M* = 1.9, *SD* = 0.8), t(26) = 15.74, p < .001. Similarly, in 30Hz trials, human faces (*M* = 2.0, *SD* = 0.6) were significantly more detectable than illusory faces (*M* = 0.8, *SD* = 0.5), *t*(26) = 16.3, *p* < .001.

One-sample *t* tests were also conducted to determine if detection sensitivity was above chance in each condition (i.e., d’ > zero). In 6Hz trials, detection was significantly above chance performance for human faces (*t*(26) = 19.24, *p* < .001) and illusory faces (*t*(26) = 11.50, *p* < .001). In 30Hz trials, detection was significantly above chance performance for human faces (*t*(26) = 15.08, *p* < .001) and illusory faces, *t*(26) = 7.73, *p* < .001.

3. Experiment 2

In Experiment 1 we uncovered evidence that the illusory faces in examples of face pareidolia can be reliably detected as faces, even at fast presentation speeds (30Hz). This finding is consistent with the neural correlates described by Wardle et al. (2020). However, examples of face pareidolia have two identities: an illusory face *and* an object. Is detection at 30Hz limited to the illusory face, or is the object identity also detectable at high presentation speeds?

In Experiment 2, we examined rapid detection of objects using the same task described in Experiment 1. Targets in Experiment 2 consisted of food items (such as fruits, cakes and chopped vegetables), with half of the targets also containing illusory faces. If, at high presentation speeds (30Hz) examples of face pareidolia are perceived and detected as faces, not objects, then performance with illusory face targets should be impaired. On the other hand, if examples of face pareidolia are perceived and detected as both faces and objects then there should be no effect of target type, even in the 30Hz condition.

3.1. Method

3.1.1

*Participants*

Twenty-five participants (Mage = 22.8, 19 Females, 6 Males) were recruited by the UQ School of Psychology First Year Research Participation Scheme or via word of mouth. Participants either received course credit or received a $20 voucher as compensation. Participants gave consent to participate and reported having normal or corrected-to-normal vision. Two participants had previously completed Study 1.

3.1.2

*Stimuli*

The stimulus set in Experiment 2 included 234 distractor images (ordinary inedible objects, for example, cars, pens), 50 illusory faces found in edible objects (hereafter referred to as illusory faces), and 50 matched edible objects with no facial structure (matched food). Illusory faces found in food items were naturally occurring examples of face pareidolia in consumable items (including food, drinks, cooking ingredients). For every illusory face in food target, a matched food image was found, which was an image of the same food item with similar visual properties but without a face (See Figure 1A). All distractor images were used in Experiment 1, but those that contained food were removed for Experiment 2.

3.1.3

*Design*

The experimental protocol included 200 target present trials and 200 target absent trials. The experimental manipulations included both *target type* (illusory face, matched food), and *frequency* (6Hz and 30Hz).

3.1.4

*Procedure*

The task was to detect food items amongst a sequence of distractor objects. Before beginning the experiment, participants were shown illustrative examples of what matched-food targets and distractor images would look like. It was explained that “food” counted as all consumables, which included drinks and ingredients involved in cooking. Participants were told to ask for clarification from the experimenter if they were unsure about whether an image counted as ‘food’. Examples of illusory face images were not presented as illustrative examples, because the presence of illusory faces was not discussed until the conclusion of the experiment. This was to discourage participants from expecting and looking for both food items and illusory faces.

The trial structure of Experiment 2 was replicated from Experiment 1, but only 400 trials were presented, which was judged to be enough power to detect the effects of interest while accounting for less availability of illusory faces in food stimuli. Targets were shown twice in the experiment, once in a 6Hz trial and once in a 30Hz trial. At the conclusion of each trial, participants were asked on the screen “Was there food present in any images?” with answer options “yes” or “no.” If participants answered “yes,” a secondary question appeared asking, “Did you like the type of food you saw?” with answer options “yes” or “no.” This secondary question was designed to keep the trial structure of Experiment 2 comparable to that of Experiment 1 but the results of this question were not analysed. No feedback regarding performance was given. The order of images in each trial was randomised, and the order of trials was randomised for each participant. At the conclusion of the experiment, participants were asked verbally “Did you notice anything interesting about the food images you saw?” which was followed up with “Did you notice any faces in the food images?”

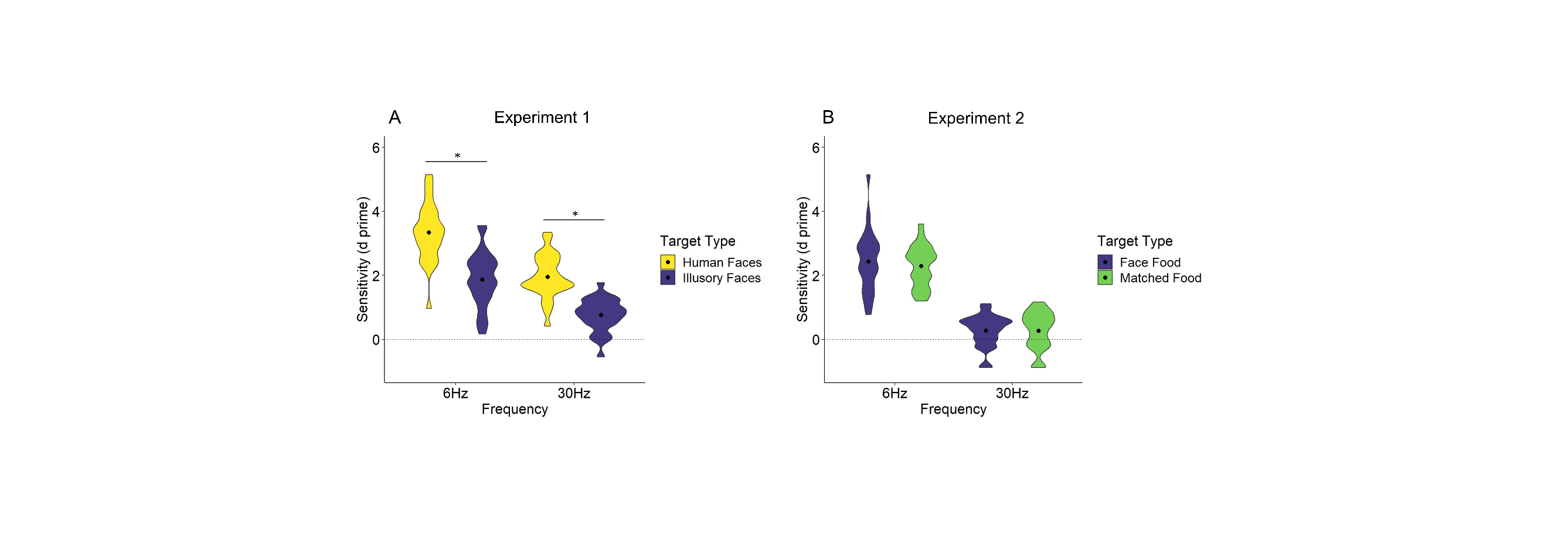
3.2. Results

Twenty-two out of the 23 participants reported seeing faces in food items when questioned at the conclusion of the experiment.

A 2 (frequency: 6Hz and 30Hz) x 2 (target type: illusory face, matched food) repeated measures ANOVA was performed (see Figure 2B). The main effect of frequency was significant, *F* (1, 24) = 185.09, *p* < .001, such that participants were better at detecting targets in the 6Hz frequency (*M* = 2.4, *SD* = 0.8) compared to the 30Hz frequency (*M* = 0.3, *SD*  = 0.5) . However, we found no evidence of a main effect of target type, *F* (1, 24) = 0.829, *p* = .372, η²p = .033. The interaction between frequency and target type was also not significant, *F* (1, 25) = 0.983, *p* = .331, η²p = .039.

Since there was no significant difference between food and illusory face targets, a Bayesian analysis was conducted to test for evidence in favour of the null hypothesis. We conducted a set of repeated measures Bayesian *t*-tests using default priors using the BayesFactor package in R. While we found no substantial evidence for the null hypothesis in the 6Hz frequency condition (*Bayes Factor* = 0.49, *p* = .181), we found substantial evidence for no difference in the 30Hz frequency (*Bayes Factor* = 0.21, *p* = .974). In sum, our data indicate that there was no difference in detection performance for matched food targets and face-food targets in the 30Hz condition.

One-sample *t* tests were also conducted to determine if detection sensitivity was above chance performance (zero) in each condition. In 6Hz trials, detection sensitivity was significantly above chance performance for matched food (*t*(24) = 17.78, *p* <.001) and illusory face, *t*(24) = 12.70, *p* < .001. In 30Hz trials, detection sensitivity was significantly above chance performance for matched food (*t*(24) = 2.30, *p* = .030) and for face-food, *t*(24) = 2.81, *p* = .010.

Figure 2: Target detection performance for Experiment 1 and 2. (COLOUR)

*Note.* Violin plots showing detection sensitivity for A) Experiment 1: Face Detection Task (N = 27) and B) Experiment 2: Food Detection Task (N = 25).

4. Experiment 3

Collectively, the results of Experiments 1 and 2 demonstrate that examples of face pareidolia have two identities that can be recognized even at high presentation speeds. More specifically, when the pareidolia illusion occurs by happenstance in food items, the object can be recognized as a face and as a food item. This suggests that examples of face pareidolia have two recognizable identities that co-occur even when seen very briefly. However, it remained unclear whether examples of face pareidolia were more easily recognized in a face detection task or a food detection task because different stimuli were employed in Experiment 1 and 2. Therefore, in Experiment 3, we combined both participant tasks into a within-subjects experimental design and used the exact same stimuli in both tasks. The illusory faces were occurrences of face pareidolia in food and acted as targets in both the face detection task and food detection task. In accordance with how illusory faces are represented in the brain during the first 200 ms (Wardle et al., 2020), we predicted that detection performance for illusory faces would be higher in the face detection task than in the food detection task when presentation time was limited in the 30Hz condition. In contrast, detection performance for illusory faces would be more equivalent across tasks when participants have more time to view the stimuli (i.e., in the 6Hz condition).

4.1 Method

4.1.1

*Participants*

Forty-two participants (29 females, 13 males) took part in the experiment. Two participants were excluded from analysis due to scores in one condition falling more than 3 standard deviations below the group mean. Therefore, the final analysis included 40 participants (27 females, 13 males; Mage = 22.8 years). Participants reported normal or corrected-to-normal vision and gave informed consent to participate. Participants were compensated with a $20 voucher or course credit.

4.1.2

*Design*

As in Experiments 1 and 2, the experiment included target present (400 trials) and target absent trials (400 trials). The experimental design included three factors in a repeated measures design; (1) task type (*face, food*) (2) frequency (*6Hz, 30Hz*) and (3) target type (*illusory face target, non-illusory face target*).

The face detection and food detection tasks were organised as blocks and these two blocks of trials were counterbalanced across participants such that half of the participants completed the face detection task first whereas the others completed the food detection task first. In the face detection task, the targets were faces (real human faces or food items with illusory faces). In the food detection task the targets were food items (real food items or food items with illusory faces). Consequently, the exact same illusory face stimuli served as targets in both tasks, while the real human faces and real food items functioned as targets in their respective task blocks, but as distractors in the opposing task block. Trials were either 6Hz or 30Hz, and every target was presented once in each frequency condition in each task block.

On an individual participant basis, the order of each trial sequence was allocated at random for Block 1 and targets placed amongst distractor images. Then in Block 2, each trial sequence was kept intact and only the trial order was reshuffled. Therefore, the exact same trials occurred in Block 1 and Block 2 with the only difference being the instructions to participants. This was to ensure that targets placed amongst weaker or stronger distractors in Block 2 could not explain performance. In the face detection task, after viewing a trial sequence, participants were asked whether a face was present among the images - whereas, in the food detection task after viewing a trial sequence, participants were asked whether a food item was present among the images. There was no secondary question in this experiment, to ensure the only difference between the task-type blocks was the instructions of targets.

In target-present trials, we also included some trials where both a human face and a food item was present. Depending on the task type block, one would be a target and the other a distractor. This is because participants may have learnt that the presence of one target type could indicate the absence of the other target type. For example, participants who had completed the face detection block and were onto the food detection block, may have learnt that when a human face was present in a trial, a food item would not be. They may then see a human face in a trial and disengage from the trial, knowing a food item target would not appear.

To protect against this happening, when a non-illusory face target (i.e. human face or ordinary food item) was present, 25% of the time, it was paired with an opposing target. Therefore, participants would not form a belief that only one type of target would ever appear in a sequence. In trials where two targets appeared, the faces and foods appeared at least 3 positions apart, between position 5 and 15.

*Stimuli*

The stimuli for Experiment 3 were 50 human faces, 50 food items with illusory faces (illusory faces), 50 matched food items without illusory faces (hereafter referred to as food items), and 229 distractor objects. Twelve of the illusory faces from Experiment 2 were replaced with different stimuli in Experiment 3. This was to improve the detectability of the food identity as some of the original stimuli were thought to be too ambiguous (e.g., stimuli depicting raw chicken, dipping oil or empty drinking glasses). The 50 human faces were randomly selected from the set that had been employed in Experiment 1.

4.1.4 *Participant Instructions*

Participants were given examples of targets and distractors in each block. The second task was not mentioned until the participant had reached the end of the first block of trials and was preparing to begin the second block of trials. At no time were the participants given explicit instructions about how to treat examples of face pareidolia.

4.2 Results

The results of Experiment 3 revealed a distinction in detection performance across the food and face tasks (Figure 3). A three-way ANOVA between task type, target type, and frequency was conducted. A main effect of task type was revealed, *F*(1, 39) = 93.83, *p* < .001, with participants better at detecting targets in the face detection task (*M* = 2.2, *SD* = 1.4) compared to the food detection task (*M* = 1.6, *SD* = 1.3). As expected, we also found a main effect of frequency, *F*(1, 39) = 948.52, *p* < .001, such that detection sensitivity was higher in 6Hz trials (*M* = 3.0, *SD* = 0.9) compared to 30Hz trials (*M* = 0.8, *SD* = 0.8). A main effect of target type was revealed, *F*(1,39) = 270.31, *p* < .001, such that participants were better at detecting non-illusory faces (human faces and food items) compared to illusory faces.

The three-way interaction between frequency, task type, and target type was significant, *F*(1, 39) = 8.91, *p* = .005. The two-way interaction between task type and frequency was significant, *F*(1, 39) = 7.28, *p* = .010. To follow up the significant three-way interaction, six Bonferroni corrected pairwise comparisons between discrete conditions were conducted. When participants were detecting faces and the presentation speed was slow (6Hz; Figure 3A, left), detection sensitivity was significantly higher for human face targets (*M* = 4.0, *SD* = 0.7) compared to illusory face targets (*M* = 2.5, *SD* = 0.5), *t*(39) = 15.25, *pBonferroni* < .001. Similarly, when the presentation speed was fast (30Hz; Figure 3A, right), detection sensitivity was still significantly higher for human face targets (*M* = 2.0, *SD* = 0.42), compared to illusory face targets (*M* = 0.4, *SD* = 0.4), *t*(39) = 23.69, *pBonferroni* < .001. Consistent with Experiment 1, when looking for faces, human faces were more easily detected than illusory faces at both slow and fast presentation speeds.

However, we found a different pattern of results when participants were looking for food items. When the presentation speed was slow (6Hz), sensitivity was significantly higher for illusory face targets (M = 2.9, *SD* = 0.9) compared to food targets without illusory faces (*M* = 2.6, *SD* = 0.8), *t*(39) = -4.2, *pBonferroni* = .004 (see Figure 3B, left). However, when the presentation speed was fast (30Hz), there was no significant difference in sensitivity between illusory face targets (*M* = 0.3, *SD* = 0.3) and food targets (*M* = 0.5, 0.3), *t*(39) = 3.21, *pBonferroni* = .074 (see Figure 3B, right). So, when looking for food, participants were better at detecting food items with illusory faces than those without, but only at slower presentation speeds.

To determine how illusory face perception varied across tasks, we compared the illusory face detection performance across the face- and food- detection tasks (Figure 3C). As expected, in the 6Hz frequency condition, participants were more sensitive to illusory faces when looking for food than when looking for faces, *t*(39) = -3.49, *pBonferroni =* .034 (see Figure 3C). On the other hand, in the 30Hz condition, participants were equally sensitive to illusory faces, regardless of the task at hand, *t*(39) = 1.01, *pBonferroni* = 1.00 (See Figure 3C).

A group of colorful graphics

Description automatically generated with medium confidenceFigure 3. Target detection performance for Experiment 3. (COLOUR)

*Note.* Violin plots showing target detection performance for each condition in Experiment 3 (N = 40). A) Performance in the face detection task. B) Performance in the food detection task. C) Comparisons between illusory face detection performance across task type.

5. Discussion

This study investigated the rapid detection of illusory faces under different task demands. Experiment 1 revealed that participants could detect illusory faces in a face detection task, even when processing time was extremely limited. This capacity to reliably detect illusory faces, even when the presentation time was reduced to less than 34 ms, supports the claim that the experience of face pareidolia is the consequence of an error in the processes responsible for rapidly detecting faces (Keys et al., 2021; Jakobsen et al., 2023; Omer et al., 2019; Taubert et al., 2018; Taubert et al., 2020; Taubert et al., 2022; Wardle, 2020). Detection performance was higher when the targets were real human faces rather than illusory faces, an observation which is consistent with a number of previous behavioural findings (Keys et al., 2021; Jakobsen et al., 2023; Alias et al., 2021) and suggests that illusory faces are more heterogeneous than human faces in terms of their low-level visual properties (Caruana & Seymour, 2022; Epihova et al., 2022; Weibert et al., 2018).

In Experiment 2, we measured behavioural responses to objects with illusory faces in a task focused on object attributes rather than face attributes. To this end, we built a food detection task and collected examples of food items, such as a sliced capsicum, with and without illusory faces. The results revealed that food items could be accurately detected in both the 6Hz and 30Hz conditions. This is consistent with the idea that food items, like faces, form a class of biologically-relevant and visually-salient stimuli that can be detected, even under extreme processing limitations. However, we found no evidence of a difference in detection performance between food items and illusory faces. This finding suggests that, when faces are irrelevant to the task at hand, the presence of illusory facial structure does not interfere with the recognition of an object’s real identity. In other words, this finding provides proof that examples of face pareidolia maintain their canonical identity as objects.

The specific motivation for Experiment 3 was to directly compare behavioural responses to a set of illusory faces under different task conditions. Thus, we combined the face detection task with the food detection task in a repeated measures design. The results demonstrate that illusory faces are easier to detect as objects than as faces when presented for over 160 ms. Again, as in Experiment 2, this suggests that illusory faces do not interfere or diminish the object’s real identity; a piece of cake is still recognised as a piece of cake, even when we mistakenly see a face. However, the advantage for illusory faces in the food task was not observed when the presentation time was reduced to less than 34 ms; in the 30 Hz condition there was no evidence of a difference in detection performance for illusory faces across the two tasks. Collectively, these findings suggest that when processing time is limited, both of the identities in examples of face pareidolia are perceived. Yet, when processing time is increased, the object identity starts to dominate perception. These findings are consistent with recent neurophysiological and neuroimaging data showing that the mental representation of illusory faces changes over time with illusory faces initially represented as more similar to faces but, later, they are represented more similarly to objects (Wardle 2020).

The tasks employed in Experiment 1 - 3 involved the detection of stimuli from broad categories (faces and food) amongst a set of highly-variable distractor stimuli. We show that participants could detect targets in both tasks, even at the fastest presentation rate, indicating that illusory faces have two robust, yet concurrent identities: they are both food items and faces. Poorer performance at the faster 30Hz presentation speed supports previous findings that faster RSVP streams lead to lower fidelity image representations due to image masking (Potter et al., 2014; Maguire & Howe 2016; Robinson et al 2019). Thus, one intriguing finding is the enhancement of food target detection performance for illusory faces relative to other food items at 6Hz, but not 30Hz (Figure 3B). The same pattern of results was evident in Experiment 2, though non-significant (Figure 2B) possibly owing to a number of ambiguous stimuli that were used as food targets in Experiment 2, such as oil and raw chicken. Nonetheless, the observation that illusory faces were easier to detect than food items in Experiment 3, when participants were directed to look for food items, indicates that illusory faces spontaneously promoted food detection. This is consistent with the idea that illusory facial features are often incorporated into architecture, art and commercial products because they automatically summon attention (Martinez-Conde et al., 2015; Wang et al., 2022; Delbaere et al., 2011; DiSalvo & Gemperle, 2003; Wodehouse et al., 2018). Further, previous work has shown that attentional enhancement occurs 150 ms after image presentation (Moerel et. al., 2022). Thus, the benefit of having a superfluous illusory face should be pronounced when an image can be continually processed for longer than 150 ms, as was the case in Experiment 3. In sum, by measuring behaviour towards illusory faces in a food detection task, we found that illusory facial features confer attentional benefits for object targets, that are independent of task demands.

At present, representations of visual stimuli in the primate brain are thought to be organised according to whether the stimulus is an animate object, like a face, or an inanimate object, like a food item (Kiani et al., 2007; Kriegeskorte et al., 2008; Long et al., 2018; Downing et al., 2006). However, the factors that give rise to this apparent organisation are unknown (Taubert et al., 2020; Op de Beeck et al., 2019). Understanding how the brain processes objects with two identities, one animate and one inanimate, as is the case with examples of face pareidolia, will be important for obtaining a complete understanding of how the brain recognizes visual stimuli in the real world. But, first, we need to understand how the perception of face pareidolia changes across time and circumstance. In this study we took the first steps towards understanding how illusory faces are perceived under different time limitations and task demands. Our data reveal that examples of face pareidolia have two stable, detectable identities - even when presentation time is limited to less than 34 ms. This makes ambient examples of face pareidolia, as a class of visual stimuli, conceptually distinct from ambiguous faces (Liu et al., 2014; Summerfield et al., 2006; Zimmermann et al., 2019), mooney faces (Palmer, Goddard, and Clifford, 2022; Taubert & Parr, 2011) or bi-stable stimuli that are sometimes faces (Wang et al., 2013; Tong et al., 1998). Future research can now focus on characterising the neural representation of these unique stimuli under different circumstances by manipulating attentional and task demands.

Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Alais, D., Xu, Y., Wardle, S. G., & Taubert, J. (2021). A shared mechanism for facial expression in human faces and face pareidolia. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1954), 20210966. https://doi.org/10.1098/rspb.2021.0966

Alonso-Prieto, E., Belle, G. V., Liu-Shuang, J., Norcia, A. M., & Rossion, B. (2013). The 6Hz fundamental stimulation frequency rate for individual face discrimination in the right occipito-temporal cortex. *Neuropsychologia*, *51*(13), 2863–2875. https://doi.org/10.1016/j.neuropsychologia.2013.08.018

Avery, J. A., Liu, A. G., Ingeholm, J. E., Gotts, S. J., & Martin, A. (2021). Viewing images of foods evokes taste quality-specific activity in gustatory insular cortex. *Proceedings of the National Academy of Sciences*, *118*(2), e2010932118. https://doi.org/10.1073/pnas.2010932118

Caruana, N., & Seymour, K. (2022). Objects that induce face pareidolia are prioritized by the visual system. *British Journal of Psychology*, *113*(2), 496–507. https://doi.org/10.1111/bjop.12546

Collins, E., Robinson, A. K., & Behrmann, M. (2018). Distinct neural processes for the perception of familiar versus unfamiliar faces along the visual hierarchy revealed by EEG. *NeuroImage*, *181*, 120–131. https://doi.org/10.1016/j.neuroimage.2018.06.080

Decramer, T., Premereur, E., Zhu, Q., Paesschen, W. V., Loon, J. van, Vanduffel, W., Taubert, J., Janssen, P., & Theys, T. (2021). Single-unit recordings reveal the selectivity of a human face area. *Journal of Neuroscience*, *41*(45), 9340–9349. https://doi.org/10.1523/JNEUROSCI.0349-21.2021

Downing, P. E., Chan, A. Y., Peelen, M. V., Dodds, C. M., & Kanwisher, N. (2006). Domain specificity in visual cortex. Cerebral cortex, 16(10), 1453-1461. https://doi.org/10.1093/cercor/bhj086

Gentile, F., & Rossion, B. (2014). Temporal frequency tuning of cortical face-sensitive areas for individual face perception. *NeuroImage*, *90*, 256–265. https://doi.org/10.1016/j.neuroimage.2013.11.053

Goren, C. C., Sarty, M., & Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, *56*(4), 544–549.

Jain, N., Wang, A., Henderson, M. M., Lin, R., Prince, J. S., Tarr, M. J., & Wehbe, L. (2023). Selectivity for food in human ventral visual cortex. *Communications Biology*, *6*(1), Article 1. https://doi.org/10.1038/s42003-023-04546-2

Jakobsen, K. V., Hunter, B. K., & Simpson, E. A. (2023). Pareidolic faces receive prioritized attention in the dot-probe task. *Attention, Perception, & Psychophysics*, *85*(4), 1106–1126. https://doi.org/10.3758/s13414-023-02685-6

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*(11), 4302–4311. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997

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

Khosla, M., Ratan Murty, N. A., & Kanwisher, N. (2022). A highly selective response to food in human visual cortex revealed by hypothesis-free voxel decomposition. *Current Biology*, *32*(19), 4159-4171.e9. https://doi.org/10.1016/j.cub.2022.08.009

Kiani, R., Esteky, H., Mirpour, K., & Tanaka, K. (2007). Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. *Journal of Neurophysiology, 97*(6), 4296-4309. https://doi.org/10.1152/jn.00024.2007

Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., ... & Bandettini, P. A. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron, 60*(6), 1126-1141. https://doi.org/10.1016/j.neuron.2008.10.043

Liu, J., Li, J., Feng, L., Li, L., Tian, J., & Lee, K. (2014). Seeing Jesus in toast: Neural and behavioral correlates of face pareidolia. *Cortex*, *53*, 60–77. https://doi.org/10.1016/j.cortex.2014.01.013

Long, B., Yu, C. P., & Konkle, T. (2018). Mid-level visual features underlie the high-level categorical organization of the ventral stream. *Proceedings of the National Academy of Sciences, 115*(38), E9015-E9024. https://doi.org/10.1073/pnas.1719616115

Maguire, J. F., & Howe, P. D. L. (2016). Failure to detect meaning in RSVP at 27 ms per picture. *Attention, Perception, & Psychophysics*, *78*(5), 1405–1413. https://doi.org/10.3758/s13414-016-1096-5

McKeeff, T. J., Remus, D. A., & Tong, F. (2007). Temporal limitations in object processing across the human ventral visual pathway. *Journal of Neurophysiology*, *98*(1), 382–393. https://doi.org/10.1152/jn.00568.2006

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

Palmer, C. J., & Clifford, C. W. G. (2020). Face pareidolia recruits mechanisms for detecting human social attention. *Psychological Science*, *31*(8), 1001–1012. https://doi.org/10.1177/0956797620924814

Palmer, C. J., Goddard, E., & Clifford, C. W. (2022). Face detection from patterns of shading and shadows: The role of overhead illumination in generating the familiar appearance of the human face. *Cognition, 225*, 105172. https://doi.org/10.1016/j.cognition.2022.105172

Potter, M. C., Wyble, B., Hagmann, C. E., & McCourt, E. S. (2014). Detecting meaning in RSVP at 13 ms per picture. *Attention, Perception, & Psychophysics*, *76*(2), 270–279. https://doi.org/10.3758/s13414-013-0605-z

Rekow, D., Baudouin, J.-Y., Brochard, R., Rossion, B., & Leleu, A. (2022). Rapid neural categorization of facelike objects predicts the perceptual awareness of a face (face pareidolia). Cognition, 222, 105016. https://doi.org/10.1016/j.cognition.2022.105016

Robinson, A. K., Grootswagers, T., & Carlson, T. A. (2019). The influence of image masking on object representations during rapid serial visual presentation. *NeuroImage*, *197*, 224–231. https://doi.org/10.1016/j.neuroimage.2019.04.050

Summerfield, C., Egner, T., Mangels, J., & Hirsch, J. (2006). Mistaking a house for a face: neural correlates of misperception in healthy humans. *Cerebral cortex, 16*(4), 500-508. https://doi.org/10.1093/cercor/bhi129

Taubert, J., & Parr, L. A. (2011). Geometric distortions affect face recognition in chimpanzees (Pan troglodytes) and monkeys (Macaca mulatta). Animal cognition, 14, 35-43. https://doi.org/10.1007/s10071-010-0341-x

Taubert, J., Wardle, S. G., Flessert, M., Leopold, D. A., & Ungerleider, L. G. (2017). Face pareidolia in the rhesus monkey. *Current Biology*, *27*(16), 2505-2509.e2. https://doi.org/10.1016/j.cub.2017.06.075

Taubert, J., Wardle, S., Flessert, M., Basile, B., Koele, E., James, C., Murray, E., & Ungerleider, L. (2018). Amygdala damage eliminates monkeys’ viewing preference for real and illusory faces. *Journal of Vision*, *18*(10), 1232. https://doi.org/10.1167/18.10.1232

Taubert, J., Wardle, S. G., & Ungerleider, L. G. (2020). What does a “face cell” want?’. *Progress in Neurobiology*, *195*, 101880. https://doi.org/10.1016/j.pneurobio.2020.101880

Taubert, J., Wardle, S. G., Tardiff, C. T., Koele, E. A., Kumar, S., Messinger, A., & Ungerleider, L. G. (2022). The cortical and subcortical correlates of face pareidolia in the macaque brain. *Social Cognitive and Affective Neuroscience*, *17*(11), 965–976. https://doi.org/10.1093/scan/nsac031

Taubert, J., Wally, S., & Dixson, B. J. (2023). Preliminary evidence of an increased susceptibility to face pareidolia in postpartum women. *Biology Letters, 19*(9), 20230126.

Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron, 21*(4), 753-759.

Tsao, D. Y., & Livingstone, M. S. (2008). Mechanisms of face perception. *Annual Review of Neuroscience*, *31*, 411–437. https://doi.org/10.1146/annurev.neuro.30.051606.094238

Uchiyama, M., Nishio, Y., Yokoi, K., Hirayama, K., Imamura, T., Shimomura, T., & Mori, E. (2012). Pareidolias: Complex visual illusions in dementia with Lewy bodies. *Brain*, *135*(8), 2458–2469. https://doi.org/10.1093/brain/aws126

van der Laan, L. N., de Ridder, D. T. D., Viergever, M. A., & Smeets, P. A. M. (2011). The first taste is always with the eyes: A meta-analysis on the neural correlates of processing visual food cues. *NeuroImage*, *55*(1), 296–303. https://doi.org/10.1016/j.neuroimage.2010.11.055

Wang, M., Arteaga, D., & He, B. J. (2013). Brain mechanisms for simple perception and bistable perception. Proceedings of the National Academy of Sciences, 110(35), E3350-E3359. https://doi.org/10.1073/pnas.1221945110

Wang, C., Yu, L., Mo, Y., Wood, L. C., & Goon, C. (2022). Pareidolia in a built environment as a complex phenomenological ambiguous stimuli. *International Journal of Environmental Research and Public Health, 19*(9), 5163. https://doi.org/10.3390/ijerph19095163

Wardle, S. G., Paranjape, S., Taubert, J., & Baker, C. I. (2022). Illusory faces are more likely to be perceived as male than female. *Proceedings of the National Academy of Sciences*, *119*(5), e2117413119. https://doi.org/10.1073/pnas.2117413119

Wardle, S. G., Taubert, J., Teichmann, L., & Baker, C. I. (2020). Rapid and dynamic processing of face pareidolia in the human brain. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-18325-8

Zimmermann, K. M., Stratil, A. S., Thome, I., Sommer, J., & Jansen, A. (2019). Illusory face detection in pure noise images: The role of interindividual variability in fMRI activation patterns. *PLoS One, 14*(1), e0209310. https://doi.org/10.1371/journal.pone.0209310