RESEARCH REPORT

Evidence for a theta-band behavioural oscillation in rapid face detection

Xiaoyi Liu[1] | **Elio Balestrieri[2,3]** | **David Melcher[1]**

[1]New York University Abu Dhabi, Abu Dhabi, United Arab Emirates

[2]Institute of Psychology, University of Münster, Münster, Germany

[3]Otto-Creutzfeldt-Center for Cognitive and Behavioral Neuroscience, University of Münster, Münster, Germany

Correspondence
David Melcher, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates.
Email: david.melcher@nyu.edu

Edited by: Ali Mazaheri

Abstract

Theories of rhythmic perception propose that perceptual sampling operates in a periodic way, with alternating moments of high and low responsiveness to sensory inputs. This rhythmic sampling is linked to neural oscillations and thought to produce fluctuations in behavioural outcomes. Previous studies have revealed theta- and alpha-band behavioural oscillations in low-level visual tasks and object categorization. However, less is known about fluctuations in face perception, for which the human brain has developed a highly specialized network. To investigate this, we ran an online study ($N = 179$) incorporating the dense sampling technique with a dual-target rapid serial visual presentation (RSVP) paradigm. In each trial, a stream of object images was presented at 30 Hz and participants were tasked with detecting whether or not there was a face image in the sequence. On some trials, one or two (identical) face images (the target) were embedded in each stream. On dual-target trials, the targets were separated by an interstimulus interval (ISI) that varied between 0 to 633 ms. The task was to indicate the presence of the target and its gender if present. Performance varied as a function of ISI, with a significant behavioural oscillation in the face detection task at 7.5 Hz, driven mainly by the male target faces. This finding is consistent with a high theta-band-based fluctuation in visual processing. Such fluctuations might reflect rhythmic attentional sampling or, alternatively, feedback loops involved in updating top-down predictions.

KEYWORDS
behavioural oscillations, face perception, RSVP, theta/alpha rhythm

1 | INTRODUCTION

The brain is confronted with a constant influx of sensory input and yet is able to form a stable ongoing visual experience. Despite the seemingly continuous operation, evidence has shown that sensory sampling works in a periodic way, with moments of high and low responsiveness to external stimulations interleaving with each other. This temporal structure of visual processing is reflected in ongoing neural oscillations, where the phase and power of prestimulus theta- (3–7 Hz) and alpha-band (8–12 Hz) activity have been shown to predict perceptual

Abbreviations: FFA, (face fusiform area); FFT, (Fast Fourier Transform); HR, (hit rate); ISI, (interstimulus interval); RSVP, (rapid serial visual presentation); TIW, (temporal integration window).

outcomes in a variety of visual tasks, including near-threshold target detection (Busch et al., 2009; Dugué et al., 2011b; Hanslmayr et al., 2007; Mathewson et al., 2009; Romei et al., 2010; van Dijk et al., 2008), visual search (Dugué et al., 2015; Merholz et al., 2019), and rapid segregation of two visual stimuli into separate percepts (Ronconi et al., 2017; Wutz et al., 2014).

Recently, the dense sampling technique has allowed researchers to further examine the temporal dynamics of the perception and behaviour with high temporal resolution. This line of research typically involves a visual event (e.g., a flash, cue, or prime) to reset the perceptual rhythm, followed by a target whose onset is densely distributed across time (Fiebelkorn et al., 2011). By analysing the rhythmic fluctuations in performance (i.e., behavioural oscillations), researchers have found that visual attention, even when participants are explicitly asked to attend one location, alternates between multiple spatial locations at about 8 Hz, reflecting an alpha-band rhythmic perceptual sampling and a theta-band attentional modulation (Dugué et al., 2016; Fiebelkorn et al., 2013; Gaillard & Hamed, 2020; Landau & Fries, 2012; Michel et al., 2020). Similarly, a 4-Hz behavioural oscillation in the priming effect was found when participants were asked to discriminate the pointing direction of a probe arrow (Huang et al., 2015). Using a dual-task paradigm, Balestrieri et al. (2021) directly tested the hypothesis that these behavioural oscillations can also occur for the alternation of attentional allocation between external visual stimulation and internal representations in a working memory task. They found that near-threshold target detection performance fluctuated over time during the memory maintenance period, at a slower rhythm when visual working memory load was high (5 Hz) and a faster rate (7.5 Hz) when there was little demand for internal information maintenance. These results suggest the existence of a common theta-band-based oscillatory modulation from a central attention network to early visual cortex.

Another intriguing implication of rhythmic sampling theories is that perception of sequences of stimuli may differ depending on whether those stimuli fall within the same perceptual sample. In the simple case of two visual flashes, for example, the probability of perceiving only one flash is higher when the interstimulus interval is brief. This is thought to reflect a "temporal integration window" (TIW) during which sensory input is combined into a single coherent percept. A number of studies have reported both behavioural (Freschl et al., 2019; Sharp et al., 2019; Wutz, Drewes, & Melcher, 2016) and neural (Wutz, Muschter, et al., 2016; Ronconi et al., 2017; Samaha & Postle, 2015) evidence for such temporal integration windows in processing simple visual stimuli

such as flashes, linked to specific sampling rhythms (Battaglini et al., 2020; Ronconi et al., 2017; Wutz & Melcher, 2014).

Here, we ask if the same oscillatory temporal organization found in near-threshold target detection and temporal integration/segregation generalizes to high-level visual perception. Is the theta- and/or alpha-band oscillation merely a characteristic of the functioning of the early visual cortex, or a general mechanism through which the brain organizes all types of incoming visual inputs? Congruent with the latter hypothesis, subdural recordings have linked behavioural oscillations in different visual tasks to periodic fluctuations in cortical excitability across the whole frontoparietal network (Helfrich et al., 2018).

Categorical perception, such as recognition of objects or faces, is a particularly interesting topic to investigate. Is such high-level processing also subject to neural and behavioural oscillations? Some evidence for this idea comes from neuroimaging studies (fMRI) showing a temporal processing capacity of around four to six items per second in category-selective areas such as the fusiform face area (FFA) (Gentile & Rossion, 2014; McKeeff et al., 2007; Stigliani et al., 2015). Specifically, the processing of each individual face seemed to be optimal when the faces were presented at a rate consistent with a theta-band rhythm. The temporal limitations of the cortical areas at this relatively high level of the visual processing hierarchy open up the possibility that high-level visual recognition may also be susceptible to periodic modulation. Indeed, by asking participants to categorize animal/vehicle targets embedded in a rapid serial visual presentation (RSVP) stream, Drewes et al. (2015) found a behavioural oscillation of 5 Hz in accuracy as a function of the duration between a flash event and the target. It is interesting to note that the behavioural oscillation was driven by the "animal" trials. Because animal and vehicle recognitions are supported by different brain regions as evident in category selectivity literatures (Huth et al., 2012; Peelen & Downing, 2017), this finding raises the question whether the exact oscillation frequency depends on task demands and stimulus properties.

The goal of the current study is to explore the temporal dynamics of rapid face perception. The face is a special stimulus, because it carries important social information critical to survival, and humans have evolved an extensive network to process faces efficiently and differently from other non-face objects (Farah et al., 1998; Kanwisher & Yovel, 2006). A number of studies using rapid serial presentation of faces provide evidence for optimal face processing when the stimuli are presented at a rate within the theta-band (Gentile &

Rossion, 2014; McKeeff et al., 2007; Stigliani et al., 2015). If faces are presented faster than about 6 Hz, the measured neural activity seems to blend together, suggesting that the sampling rhythm is slower than 4–6 Hz. However, those studies looked at a neural index of the "saturation" (subadditive activity) of face processing areas, rather than face perception itself. Using a classification image technique, Blais et al. (2013) found an alpha-band (10 Hz) oscillation in information use when participants performed a face identification task for face images embedded in a dynamic noise mask. The difference in oscillatory frequency in Blais et al. (2013) and the neuroimaging studies (Gentile & Rossion, 2014; McKeeff et al., 2007; Stigliani et al., 2015) might due to the nature of the different tasks that they used. In Blais et al. (2013), the task was identification (not detection) and face images remained visible to participants throughout the trial. Their results are consistent with prior studies showing a 10-Hz fluctuation in detectability to simple visual stimuli, like phosphenes (Dugué et al., 2011b; Romei et al., 2008), dots (Mathewson et al., 2009; Romei et al., 2010), and letters (Hanslmayr et al., 2007), and thus might not be specific to faces per se.

To measure comparable behavioural oscillations at the perceptual level, we adopted a dual-target RSVP paradigm as in Drewes et al. (2015). On the key trials of interest, the face target was repeated twice in each RSVP stream (33 ms/30 Hz) of non-face object images. We manipulated the interstimulus interval (ISI) between the two identical targets with the dense sampling technique. Participants were asked to detect a face and, if present, to discriminate its gender. Interestingly, Drewes et al. (2015) also found evidence for a temporal integration window, in that performance was higher when both stimuli fell within a window of around 150 ms. That pattern of results would also be predicted by a high theta-band sampling window of around 6–8 Hz.

In the current study, our first hypothesis was to see oscillatory fluctuations in hit rate (HR) in both face detection and gender discrimination tasks. Because gender discrimination inherently involves additional processing than simple detection, and thus could occur more slowly, we predicted that the frequency of oscillation might be faster for the easier (detection) task. We also examined the effect of stimulus properties and participant gender on behavioural oscillations by comparing performance when the target was a female versus male face.

Second, we used the pattern of performance as a function of ISI to test whether there was a temporal integration window (TIW) of around 120–170 ms, as might be expected if face processing (as indexed by face categorical responses, such as the N170/M170: see Rossion &

Jacques, 2011, for example) integrates information within a single sample. To test this, we estimated the convergence level (i.e., the stabilized HR during the last 204-ms ISI interval) at which the largest effect of ISI largely plateaued. The corresponding estimated time of convergence was used as an indicator of TIW, because targets presented within which would tend to be perceived as an integrated percept. We predicted that the TIW would correspond to one cycle of the behavioural oscillation, as have been found in neural oscillation studies (Ronconi et al., 2017; Samaha & Postle, 2015).

At the same time, our method also took into account the alternative hypothesis that fluctuations in detection would be more closely aligned with alpha oscillations (8–12 Hz). A number of studies have provided behavioural (Ronconi & Melcher, 2017) and neural (Dugué et al., 2011b; Mathewson et al., 2009) measures of visual detection linked to the alpha rhythm, including the Blais et al. (2013) study of face identification mentioned above. If so, then perhaps detection of faces, although they are higher-level stimuli than typically used in previous studies, might also be characterized by a fluctuation at around 10 Hz, the typical peak alpha frequency in young adults (Benwell et al., 2019; Grandy et al., 2013).

2 | MATERIALS AND METHODS

2.1 | Participants

We recruited 10 participants (one female) for the pilot study (see procedure) and 179 participants (87 females) for the experiment from the online experiment platform Prolific (www.prolific.co). The demographics of the participants (50% female, around 25% each self-identified as "Asian," "Black," "Latino," or "White" on the Prolific platform) matched that of the face stimuli (see below). These ethnic categories were self-defined by participants as part of signing up for Prolific studies. We recruited from a diverse group of participants in terms of self-identified ethnicity and used a diverse set of face images.

All participants gave informed consent by checking the relevant boxes on a consent form webpage and received 4.00 GBP as compensation. The experiments were approved by the New York University Abu Dhabi Human Research Protection Program Internal Review Board (IRB).

In the pilot study, two participants were excluded because they failed to see any faces on 75% of the trials. For the main experiment, we excluded 31 participants who did not run the experiment on a computer with a

60-Hz refresh rate (see Section 2.4.1). We excluded a further two participants (1.35% of all included participants) for not being able to complete the task as instructed. An additional six participants (two females; 4.05%) were classified as "atypical" participants based on their performance (see Data analysis) and were excluded from the frequency analysis. The remaining 140 participants used in the main data analyses were an average age of 26.31 (±6.80) years old. We had relatively equal distributions of gender (75 females) and self-identified ethnicity groups (39 Asian, 39 Black, 22 Latino, and 40 White). The exact proportions varied from expected due to a discrepancy between the initial self-identification registered by the participants on Prolific and their response to the demographic questionnaire during the actual experiment, in which they indicated their age, gender, and self-identified ethnicity. According to the self-registered demographic information in Prolific, the participants were from at least 37 different countries of birth and at least 23 countries of current residence (five unknown).

2.2 | Materials

Images of non-face objects (baseline) and faces (target) were presented in a RSVP paradigm. Baseline images were selected from the database used by Konkle et al. (2010). Images that contained animals or face-like objects (e.g., eyeglasses) were excluded, leaving a set of 3910 images of inanimate objects. A separate set of 504 face images were selected from the Academic Dataset by Generated Photos (https://generated.photos/datasets) (Karras et al., 2019). The faces in this database, which was designed to support machine learning, were generated using professional photographers, lighting and make-up, and features a diverse set of faces in terms of demographics, facial expressions and poses. The images used in this study were actually synthetic faces generated out of tens of thousands of face images. To ensure diversity, an equal number of female and male face images from that database were selected from four ethnicity groups (defined by the Dataset as: Asian, Black, Latino, and White). The age, expression, view, and size of the face varied across images. All images were converted to greyscale and resized to 256 pixels in width and height with Adobe Photoshop graphics software. The hair part of the face images was cropped out, such that only face contour was left. All the stimuli were histogram-matched for luminance through the SHINE toolbox (Willenbockel et al., 2010). In addition, we added 20% salt-and-pepper noise background to all the stimuli in the main experiment (see below for details).

2.3 | Procedure

2.3.1 | Pilot study

A pilot study was conducted to determine the stimulus type and presentation rate. We randomly selected 1376 objects images and 80 face images and superimposed them onto four different background conditions (i.e., the plain background, white noise background, and 10% and 20% salt-and-pepper noise backgrounds) (Figure 1a). In the second condition, all stimuli were superimposed onto a white noise background. In the last two conditions, we added 10% and 20% salt-and-pepper noise on the stimuli from the second condition using MATLAB *imnoise* function.

The RSVP paradigm was employed in the pilot study. In each trial, participants were presented with an RSVP sequence of 20 images, among which a single target image would appear randomly at any positions except the first and last one. Participants were asked to indicate the presence of a face image by pressing a key (face detection task). If they reported face present, they would indicate its gender by pressing another key (gender discrimination task). The presentation rate of the sequence was 60 Hz in the plain, white noise, and 10% salt-and-pepper background conditions, and was 30 Hz in the 20% salt-and-pepper background condition. There were 30 trials for each stimulus type condition and 24 catch trials. Each block contained one background condition and 6 catch trials. Participants performed below chance level in the face detection task at all 60-Hz conditions, but were able to reliably report the presence of a face in the 30-Hz condition. In light of this pattern, we used the 20% salt-and-pepper background and a presentation rate of 30 Hz in the formal experiment.

2.3.2 | Main experiment

The experiment was coded using the *jsPsych* library (de Leeuw, 2015), embedded in an HTML environment. Stimuli were presented on participants' desktop or laptop computer on a browser of their choice. Most participants used Chrome, with the exception of a few participants who used Firefox (eight participants) or Opera (8). Refresh rate was double checked after the experiment (see Data screening), with most participants (84.85%) running the study on a computer of 60-Hz refresh rate. Only these participants were included in data analysis.

We adopted a dual-target RSVP paradigm, in which a target face image was typically repeated twice in each image stream. As shown in Figure 1b, each trial began with a central fixation cross for 500 ms, followed by an

FIGURE 1 (a) Sample stimuli used in the pilot. The background conditions from left to right columns are the plain, white noise, 10% and 20% salt-and-pepper noise backgrounds. (b) Illustration of a dual-target trial in the main experiment. Each trial started with a 500-ms central fixation, followed by an RSVP sequence of 30 image (30 Hz/33 ms) and another 500-ms fixation. The task was to indicate the presence/absence of face target(s) and to report the gender if present. There were also trials where only the first target was presented (the one-target trials) and trials without any targets (the catch trials).
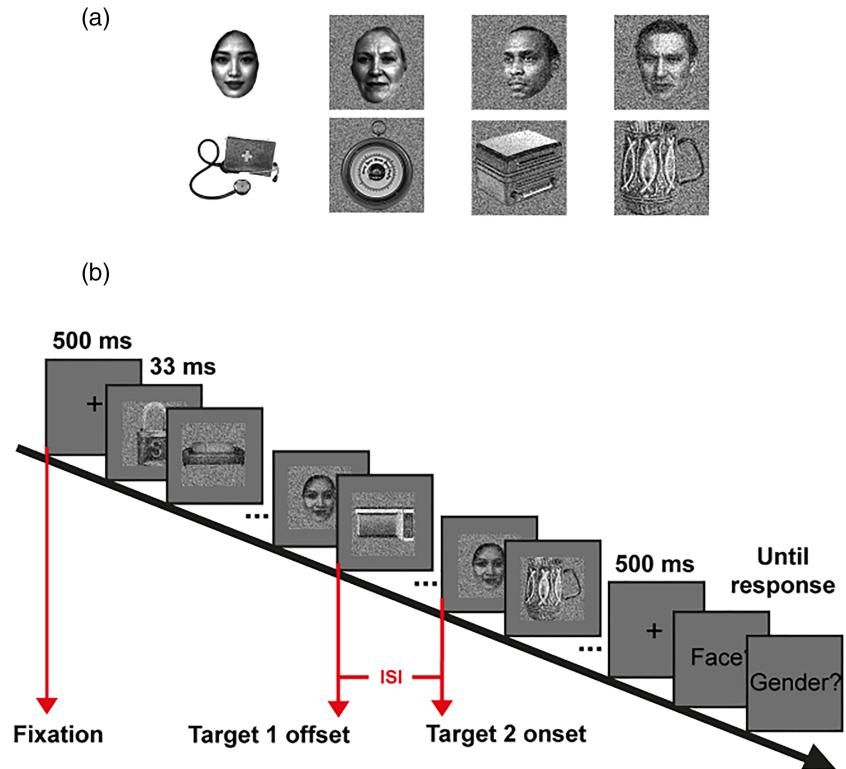
RSVP sequence of 30 images presented at 30 Hz. We coded our own RSVP plugin in the *jsPsych* library using the *Window.requestAnimationFrame()* function such that each image was presented for 2 refresh frames. The first target image appeared randomly in the stream at the 8th, 9th, or 10th position. In 80% trials (i.e., the dual-target trials), a second target image appeared after a randomly selected ISI from 0 to 633 ms in steps of 33 ms. Each ISI condition was repeated 24 times. We also included 24 one-target trials (4% of all trials), which included only the first target image and served as a baseline measure, and 96 catch trials (16% of all trials), where no target was presented. The image stream ended with a 500-ms fixation frame. Participants were then prompted to complete the face detection task and a gender discrimination task, as in the pilot study. There were a total of 600 trials, divided in 20 blocks. Before the experiment, participants completed seven practice trials to familiarize themselves with the task. The practice was the same as the experiment, except that the first and second target images were fixed in the image stream at the 10th and 13th position and colourful object images and cartoon face images were used as stimuli.

2.4 | Data analysis

Behavioural performance was analysed with MATLAB R2019b (The MathWorks, Inc). We first calculated each participant's mean HR in the one-target trials and in each ISI condition of the dual-target trials. Participants who had a mean HR lower than 50% in the 0-ms ISI condition in the face detection task were excluded from further analyses. In this condition, the second target was presented immediately following the first target, making the target twice as long as the other stimuli in the image stream. Detection in this case should be the easiest condition. If participants could not reliably report the presence of the target in this condition, then we reasoned that they were not able to complete the task as instructed.

After screening the data, we focused the analysis on the dual-target trials. There was a decreasing trend in the raw percent correct data as the ISI increased, because when two targets were presented within the same TIW the task would be easier (Figure 2a). We computed the convergence level by averaging the HR across the last 204-ms interval (i.e., from 433- to 633-ms ISI conditions) when the performance was relatively stable. The time of convergence, as an indicator of the TIW, was defined as the ISI at which the hit rate decreased to 10% (or 5%) above the convergence level (Drewes et al., 2015). The convergence level was analysed using a 2 * 4 * 2 * 4 (Target Gender * Target Race * Participant Gender * Participant Race) analysis of variance (ANOVA) with target gender (female, male) and target ethnicity (Asian, Black, Latino, White) as within-subjects variables and participant
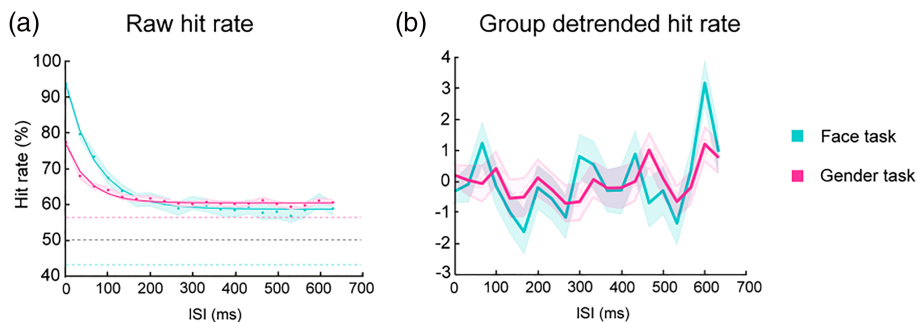
FIGURE 2 (a) Raw HR as a function of ISI. Dots represent group average HR at each ISI condition. Shaded areas represent SEM. Solid lines show the exponential functions fit to the raw data. Coloured dashed lines show the average HR in the one-target trials. Black dashed line shows performance at chance level. (b) Detrended results after the exponential function is removed from the raw data. Solid lines represent detrended HR. Shaded area represent SEM.

gender (female, male) and participant self-reported ethnicity (Asian, Black, Latino, White) as between-subjects variables.

The typical, strong effect of temporal integration on the time course of HR was observed as an exponential decay function of the ISI (Drewes et al., 2015). In order to disclose the underlying oscillatory structure of the data, we detrended the HR series by first fitting, to each series, an exponential decay function in the form:

$$f(x) = a_0 e^{\frac{-x}{\tau}} + c$$

where x is the ISI, a_0 the starting point, τ the time constant (that is, the variable describing the steepness of the decay), and c the right-hand asymptote, or the baseline performance. The fitted trends were then subtracted from the original HR series. Figure 2 shows the average across participants of HR series before (a) and after (b) detrend.

We then performed a fast Fourier transform (FFT) on the detrended HR time course for each participant to convert the results from time domain to frequency domain (Figure 3). We assessed the presence of oscillatory components across participants by computing the phase-locked sum (Balestrieri et al., 2021). This method takes into account both the amplitude and phase of the oscillations composing the original time series by summing together the vectors originating from the FFT in the complex domain, as follows:

$$PLS_f = \left(\frac{1}{n}\right) \sum_{k}^{n} A_{(f,k)} e^{i2\pi\theta_{(f,k)}}$$

where PLS_f is the phase-locked sum for frequency f, n is the number of participants, and A and θ are, respectively, the amplitude and the phase angle of the complex vector,

for each frequency f and participant k. This computation is a proxy for the presence of oscillations in a time series because it tests the degree to which non-zero oscillations are consistently in phase across participants, a phenomenon often observed in dense sampling studies (Landau & Fries, 2012; Michel et al., 2021). In order to assess the significance of the peaks observed in the spectra, we performed a permutation test by shuffling without replacement the ISI labels of the individual detrended data for 10,000 permutations, and for each surrogate dataset obtained in this way, we applied the aforementioned computation of phase-locked sum. Statistical threshold for selecting peaks of interest was defined as the 95 percentile of the phase-locked sum computed on the surrogate datasets. Bonferroni correction for the whole spectra was applied in order to account for multiple comparisons. A recent article (Brookshire, 2022) cast some doubts on this randomization procedure for statistical testing, suggesting an alternative based on autoregressive models. Given those concerns, we also tested for statistical significance based on this new pipeline in order to confirm the main results (see Supporting Information). To verify the phase concentration at the peak frequency, we ran a Rayleigh test of non-uniformity for circular data with the MATLAB CircStat toolbox (Berens, 2009).

In order to seek confirmation of sinusoidal patterns in the grand average data, we fitted a sinusoidal model in the form:

$$f(x) = A\sin(2\pi x f + \theta)$$

where A is the amplitude of the underlying oscillation, f is its frequency, and θ is the phase angle. The model was fitted using MATLAB via non-linear least squares, and adjusted R^2 was computed in order to evaluate the goodness of fit by taking the number of coefficients into
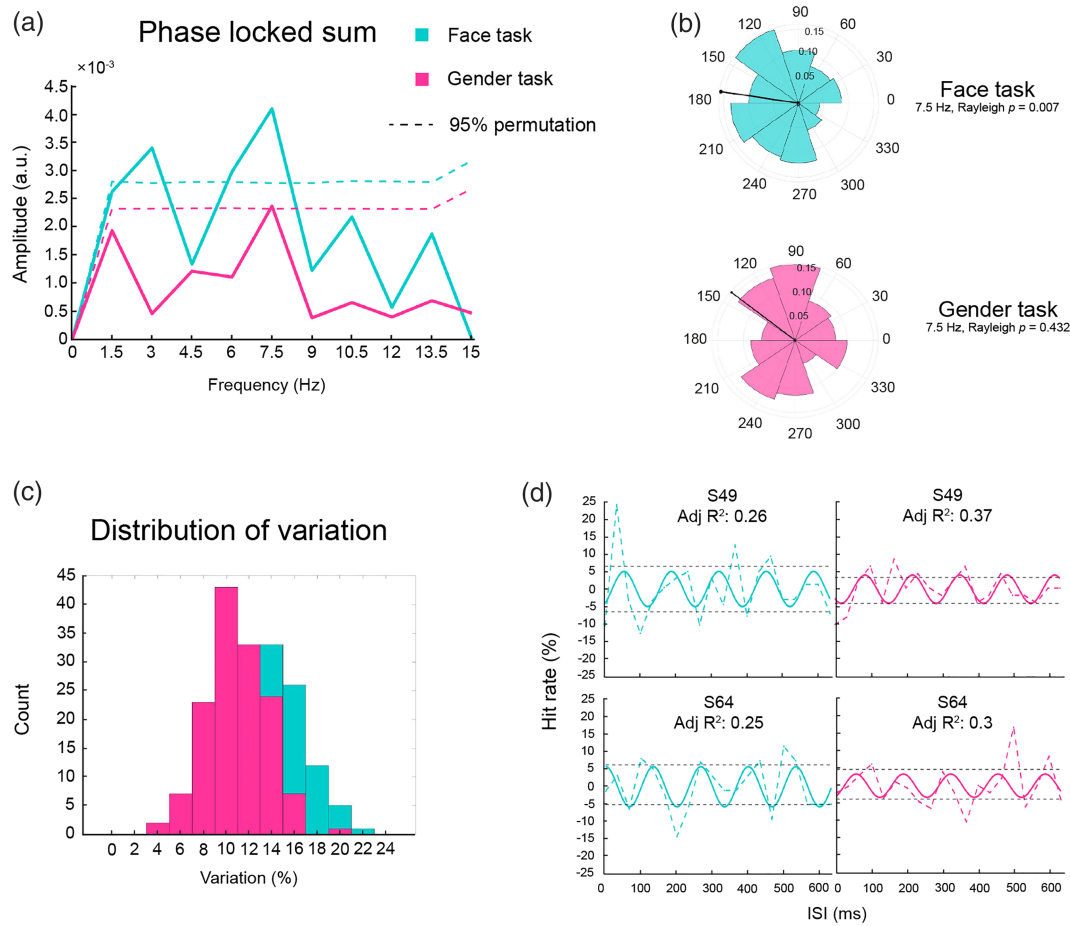
FIGURE 3 (a) Phase-locked amplitude of group spectra in the face (blue) and gender (gender) tasks. The dashed lines represent 95% percentile of permutations used to define significance of the peaks before multiple correction. (b) Polar histogram of the phase angle distribution for the highest peak of the face and gender tasks. (c) Distributions of variation (i.e., the distance between the means of positive and negative individual detrended hit rates across ISIs) in face and gender task. (d) Dashed colour lines represent the detrended hit rate of representative participants (S49 and S64). Solid lines represent the best sinusoidal fit at 7.5 Hz to the individual detrended data. The distance between the dashed black lines represent the variation.

account. Significance of the goodness of fit was evaluated non-parametrically by a permutation test: for 10,000 surrogate datasets, obtained as described above, the sinusoidal model was fitted and a corresponding adjusted R^2 was extracted. The fit was defined as significant if the original adjusted R^2 exceeded the 95% of permutation values (Balestrieri et al., 2021; Ronconi & Melcher, 2017).

2.4.1 | Data screening

As described above, we screened the participants by the refresh rate of the computer they used to run the experiment. The refresh rate was first collected by self-reporting through a separate online survey after the experiment. Actual refresh rate was then computed using the timing

information (*time_elapsed*) recorded automatically by *jsPsych* during the experiment. Most participants' actual refresh rate matched the self-report one with five exceptions. Of all 179 participants, 148 participants had a refresh rate of 60 Hz. We included only these participants in the main data analysis.

Participants who had a mean HR lower than 50% in the 0-ms ISI condition for the detection task were excluded prior to any further analyses. Among the eight participants flagged by the screening procedure, two showed low hit rates across all ISI conditions and thus were excluded from any further analysis. Another six participants showed an unexpected behavioural pattern in which their hit rate in the face detection task increased as the ISI increased. This pattern is unexpected in that it is opposite to the behavioural pattern found in the previous study adopting a similar dual-target RSVP paradigm

(Drewes et al., 2015) and in the other 140 typical participants. In the final data analysis, only the 140 typical participants were included.

3 | RESULTS

3.1 | A general oscillatory effect was found in the face detection task

As described in the data analysis section, an FFT analysis was performed on the detrended data to determine the oscillatory effects in both tasks. There was a significant peak at 7.5 Hz in the face task ($p_{(corrected)}$ = 0.017) (Figure 3a), which accounted for around 6.4% variation in the detrended data (S2). The pattern is consistent with the high theta oscillations found in other types of visual perception tasks (Balestrieri et al., 2021; Huang et al., 2015). The result of Rayleigh test for non-uniformity shows a significant phase coherence across participants at the same frequency ($z = 4.90$, $p = 0.007$), meaning that the participants were uniformly good or bad at the face detection task at a given ISI (Figure 3b). To check the effect size of the behavioural oscillation, we computed the distance between the means of positive and negative individual detrended hit rates across the ISIs (i.e., variation; Figure 3c). The average variation in the face task was 13.29% ± 3.43%. Figure 3d shows the detrended data of two representative participants. The variation was represented by the distance between the two dashed black lines. The result was confirmed by the sinusoidal fitting analysis on the group data showing that the best-fitting frequency was 7.52 Hz, although this
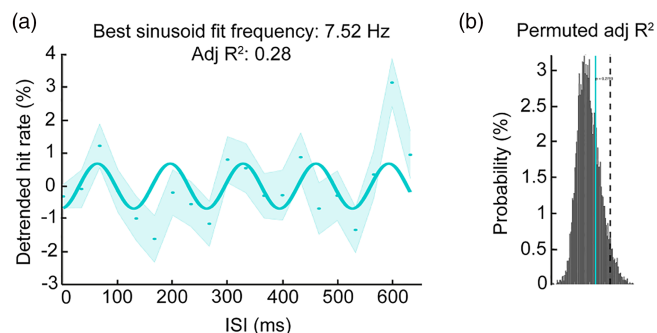
fit was not statistically significant in our permutation test (adj $R^2 = 0.28$, $p = 0.27$) (Figure 4).

In terms of the gender task, the nature of our task led to a large number of missing trials because if the participants did not report seeing any face, there would not be a gender task in that trial. To compensate for different proportions of missing trials in different ISI conditions, we replaced them with a guessing parameter (i.e., 50%). However, we still did not find any significant rhythms in the gender task. The null result could be explained by a significantly smaller variation in the gender task (10.96% ± 2.63%) than the face task, $t(139) = 6.02$, $p < 0.001$, as shown in Figure 3c,d. The lack of phase coherence (Figure 3b) at 7.5 Hz also likely contributed to the null result. It is possible that participants only reported face present when they were confident about its gender, and thus any oscillatory effect in the gender task was masked. In future work, it might be interesting to test gender discrimination on every trial in order to uncover whether face perception is all-or-nothing (perceiving a face also results in gender perception), but that remains an open question.

3.2 | The observed oscillation was driven by male target faces

We then tested how the oscillation in the face task differed between female and male target faces. Trials were grouped based on the gender of the targets. The same analyses were implemented for each data subset as for the pooled data. Figure 5a shows the detrended HR and the amplitude of phase locked sum in female and male face detection tasks. There was a significant peak at 7.5 Hz in male face detection tasks only ($p_{(corrected)}$ = 0.042), suggesting that the previous finding was mainly driven by the male faces (Figure 5b). Rayleigh test of non-uniformity confirmed the phase concentration ($z = 5.30$, $p = 0.005$) (Figure 5b). The sinusoidal fitting analysis also showed a best-fit frequency at 7.57 Hz but failed to reach statistical significance with this smaller number of trials (adj $R^2 = 0.17$, $p = 0.52$). A separate FFT analysis on trials where the target was a female face showed two non-significant peaks at 3 and 10.5 Hz (Figure 5b).

We also tested whether female and male participants contributed to the 7.5 Hz oscillation in the male face detection task differently. It is possible, for example, that people perceive same-/different-gender faces in a different way. We divided the dataset into female and male participant groups. The FFT results showed a peak at 7.5 Hz in both male ($p = 0.046$) and female ($p = 0.057$) participant groups (Figure 6), although neither reached significance after multiple corrections. The reduced sample size might explain the non-significant results.

FIGURE 4 Detrended HR as a function of ISI, with related to sinusoidal fitting. Left: Dots represent detrended HR at each ISI condition. Shaded area represents SEM. Solid line represents the sinusoidal function fitted to the data at 7.52 Hz. Right: Histogram of adj R^2 computed from each permutation ($n = 10,000$). Dashed vertical line represents 95% percentile of permutation distribution as threshold for significance. Blue solid line represents the observed adj R^2.
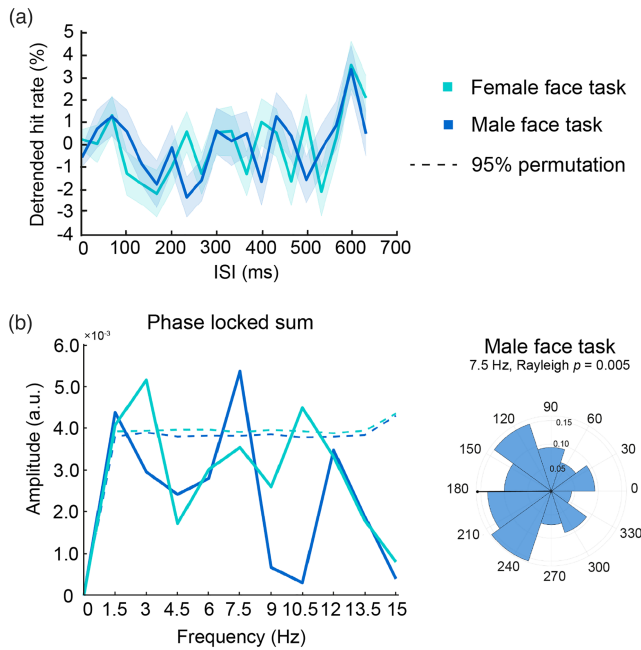
FIGURE 5 (a) Detrended HR as a function of ISI in female face detection task (light blue) and male face detection task (dark blue). Solid lines represent detrended HR. shaded areas represent SEM. (b) Left: Phaselocked amplitude of group spectra. The dashed lines represent 95% percentile of permutations used to define significance of the peaks before multiple correction. Right: Polar histogram of the phase angle distribution for the highest peaks.
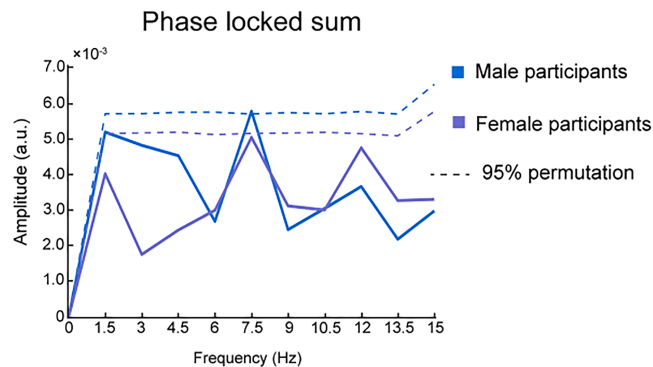

FIGURE 6 Phase-locked amplitude when female (purple) and male (blue) participants viewing male target faces. The horizontal dashed lines represent 95% percentile of permutations used to define significance of the peaks before multiple correction. The data from both female and male participants showed a peak at 7.5 Hz.

3.3 | Convergence level varied across conditions

Consistent with our hypothesis, detection performance was best when both stimuli were presented near to each other in time. As expected, peak performance was found

when the two faces were presented sequentially with no intervening distractor, because this effectively doubled the duration of the stimulus. However, this benefit was not limited to the zero ISI condition but was also found when there were one or more intervening stimuli, reducing as a function of ISI up to a limit (thus it was not just probability summation; see Drewes et al., 2015). In general, as the ISI increased, participants showed a large decline in their HRs and eventually converged at the level of 58.60% in the face detection task. There was a significant difference between the convergence level and the baseline performance ($42.86\% \pm 21.30\%$) in the one-target trials (paired t test, $p < 0.001$). We found a time of convergence of between 129.5 ms (10% above convergence level) and 181.1 ms (5% above convergence level). Interestingly, the time of convergence found here is consistent with previous estimates of a temporal integration window of around 100–150 ms (Wutz & Melcher, 2014).

An ANOVA was performed to check the convergence levels in different experimental conditions. There was a main effect of target gender, $F(1, 132) = 8.37$, $p = 0.004$, $\eta_p^2 = 0.060$, with higher convergence level in the male face detection task ($59.31\% \pm 22.09\%$) than the female face detection task ($57.88\% \pm 21.98\%$). In other words, participants were slightly better at detecting male faces. The participant gender did not show a significant main effect, nor did it interact with the target gender main effect.

A main effect of target ethnicity, $F(3, 396) = 6.25$, $p < 0.001$, $\eta_p^2 = 0.05$, and a main effect of participant ethnicity, $F(3, 132) = 2.83$, $p = 0.041$, $\eta_p^2 = 0.06$, were found. The Latino face stimuli were the easiest to detect ($60.43\% \pm 22.20\%$), followed by Black faces ($59.32\% \pm 23.25\%$), White faces ($57.35\% \pm 22.31\%$), and Asian faces ($57.30\% \pm 22.91\%$). Thus, there was a small (around 3%) difference in performance for the different face stimulus ethnicity groups. At the between-subject level, participants who identified as Latino showed the highest performance level ($65.21\% \pm 19.34\%$), followed by those who identified in the Black group ($64.15\% \pm 20.85\%$), the Asian group ($55.17\% \pm 24.26\%$), and the White group ($52.89\% \pm 19.98\%$). In other words, the participants that self-identified as Latino were about 12% better at detecting the RSVP face targets than the group self-identified as White. No other main effects or interaction effects were found.

4 | DISCUSSION

The current study aimed to test (1) whether the theta-band oscillatory temporal organization, repeatedly reported in low-level vision tasks, generalizes to more

complex high-level visual perception, and (2) whether the exact frequency of oscillation differs based on task demands and stimulus properties. We adopted the dual-target RSVP paradigm and manipulated the ISI between two identical face target images with the dense sampling technique. The main finding was that face detection performance fluctuated strongly as a function of ISI, with a variability of around 13% in hit rate. This variability reflected a strongly phase-locked 7.5-Hz behavioural oscillation in the dual-target trials in the face detection task as a function of ISI. Behavioural oscillations have been demonstrated in a variety of visual tasks, including low-level visual feature perception (Dugué et al., 2016; Fiebelkorn et al., 2013; Landau & Fries, 2012), predictive coding (Huang et al., 2015; Wang & Luo, 2017), and rapid object categorization (Drewes et al., 2015). Our finding adds to the existing literature and suggests a general theta-band oscillatory temporal organization underlying visual perception.

One interesting difference with previous studies is that most have used a flash event to reset the attentional rhythm (Drewes et al., 2015; Landau & Fries, 2012). In the present study, we did not include any additional events yet still found a robust phase consistency across participants. Given the great importance of faces to humans, the first face target might have been so salient among the baseline object images that it induced a strong attention capture, leading to a strong reset on ongoing brain oscillations.

It is interesting to note that our high theta-band (7.5 Hz) behavioural oscillation differs somewhat from the alpha-band (10 Hz) oscillation found in Blais et al. (2013). There are multiple possible reasons for this discrepancy in frequency. Perhaps the most important difference is that the previous study did not investigate the presence of a 7.5-Hz oscillation. Blais et al. (2013) probed the presence of oscillatory patterns in face discrimination by parametrically varying the faces' signal-to-noise ratio (SNR) at different frequencies and phases. This constrained the searchlight to the specific set of frequencies a priori chosen (5, 10, 15, 20 Hz), which does not include the frequency (7.5 Hz) in which we found significant phase coherence across participants. Another possibility is that Blais et al. (2013) may have tested a different sampling rhythm due to the difference in task. In their study, the face stimulus was visible to participants throughout the trial, and so their task posed a different challenge to the face-selective area compared to our study and other previous neuroimaging studies using RSVP and varying stimulus types (Gentile & Rossion, 2014; McKeeff et al., 2007; Stigliani et al., 2015). As a result, the behavioural oscillation in Blais et al. (2013) might reflect a sensory sampling rate (VanRullen, 2016), which is

presumably driven by the occipital alpha activities and is relatively stable for different kinds of stimuli (Gaillard & Hamed, 2020).

What is the origin of the observed fluctuation in rapid detection performance as a function of ISI? A possible mechanism is suggested by a recent investigation of the neural substrates of behavioural oscillations in the priming effect (Guo et al., 2020). The authors in that study presented pairs of congruent or incongruent primes and probes (faces/houses) separated by densely sampled ISIs during fMRI recording. They revealed a 5-Hz rhythm in classification accuracy of multivoxel activity patterns in the FFA and the parahippocampal place area (PPA). Those results support a functional link between behavioural oscillations in visual object priming and the dynamic activation patterns at the category-selective areas of the brain. In terms of the present study, there may have been competing interpretations of the stream of events, both in terms of face versus non-face stimuli and between male and female faces (which was task relevant). In this case, the first target could have served as a prime that generated two opposite predictions (face present/absent) in high-level areas. The predictions were then carried to the occipito-temporal visual areas by feedback signals in an interchanging way and gave rise to rhythm in detection accuracy.

On the other hand, it should also be noted that the frequency of the oscillation in our study (7.5 Hz) differs from the 5 Hz finding in Guo et al. (2020) and Wang and Luo (2017), both of which had faces as stimuli (as well as the Drewes and colleagues 5-Hz rhythm for animal/vehicle judgment). One possibility is that the face detection task that we used relies on somewhat different brain areas, and/or requires fewer communications between the areas, than the identification or categorization task used in those other studies using faces. The simpler task (detection) might require less specialized face processing that could be resolved, for example, by the occipital face area (OFA). There is MEG evidence that the OFA is activated much earlier (around 100–125 ms) than the fusiform face neurons (around 160–170 ms) (Halgren et al., 2000; Yovel, 2016). Thus, face detection related processing (and the prediction/interpretation of the stimulus) might occur faster and earlier if it is supported by the OFA, with its impact on the second target also occurring earlier.

Another potential issue is the assumption, ubiquitous in dense sampling studies, that events at each ISI are instantaneous points in the time domain. This is also true for our study, where we computed the behavioural oscillation frequency as a function of the ISI and did not include the two 33-ms face stimuli on either side of the ISI. If the "unit" of face-ISI-face is considered as a whole,

rather than just the ISI, the peak frequency would be around 5 Hz. Thus, it might be reasonable to argue that our results are consistent with a behavioural oscillation between 5.0–7.5 Hz, which overlaps with the 4–6 Hz estimates in previous studies of face perception (Gentile & Rossion, 2014; McKeeff et al., 2007; Stigliani et al., 2015).

Another difference between this study and the previous reports is that although the prime face was presented briefly in their studies, the neural processing of the face was sustained and uninterrupted by any backward mask (e.g., Wang & Luo, 2017). In contrast, the neural signals of the face image in our study were more transient due to the nature of the RSVP paradigm. Therefore, the observed oscillation could rise from local recurrent signals within the ventral visual pathway instead of a slower long-range feedback processing. In line with this theory, Mohsenzadeh et al. (2018) incorporated the RSVP paradigm with a combined MEG-fMRI analysis to track the processing of a face target among other object images. They found that in addition to an early activation that corresponded to an initial feedforward sweep, early visual cortex (EVC) activated again about 10 ms after the face target signal reached the inferior temporal cortex. Although they did not test the behavioural implication of the reactivation, it could lead to an increased detection if there were a second face target presented at the same time. For a 7.5-Hz effect, this would correspond nicely to a reactivation at 133.3 ms, which fits plausibly into the reported timing of re-entrance feedback from OFA to EVC (Eick et al., 2020).

Alternatively, the current finding could be interpreted in terms of rhythmic attentional sampling, which has been suggested as involved in behavioural oscillations in low-level visual tasks. This is supported by a series of studies applying TMS at the fronto-parietal attention regions and V1 at multiple delays (Dugué et al., 2011a, 2015, 2019). They suggest that these two regions are periodically involved in visual search tasks. As proposed in theories of rhythmic attention, the attention network at the fronto-parietal regions of the brain exerts periodic top-down control over the EVC through theta-band oscillatory neural activities, creating windows that favour either external or internal processing (Dugué et al., 2019; Fiebelkorn & Kastner, 2019; Hanslmayr et al., 2011; Jia et al., 2017).

Importantly, these theories also indicate that the frequency of behavioural oscillations differs depending on task demands. When the task is more difficult and requires more attention resources for internal processing, the attention sampling rhythm will be slower (Balestrieri et al., 2021; Chen et al., 2017). Because face detection is easier than face identification or categorization, this could predict a faster oscillation than that found by Guo et al. (2020) and Wang and Luo (2017). This could also account for the difference in oscillatory frequency between the current study (7.5 Hz) and a previous study using a similar dual-target RSVP paradigm with animals and vehicles as targets (5 Hz) (Drewes et al., 2015). Because humans are great experts at detecting faces, perhaps more than animals or vehicles, in an ever-changing world, face detection might demand fewer attention resources.

Another interesting finding is that the 7.5-Hz behavioural oscillation in the face detection task appeared to be mainly driven by male faces. This effect (and the 7.5-Hz peak) did not differ across participant genders. However, the female face detection task only evoked two non-significant peaks at 3 and 10.5 Hz. This result, while perhaps difficult to interpret, is not completely novel. In a recent study, Bell et al. (2020) manipulated the delay at which a face image (female, male, or androgynous) was presented after participants pressed a button to start the trial. In a subsequent gender discrimination task, participants' response to the androgynous faces oscillated at different frequencies when the trial was preceded by a female versus male face trial. They suggest that female and male face recognition might rely on different facial features (e.g., eyes, nose, or mouth) as cues, which could be characterized by different oscillatory frequencies (Liu & Luo, 2019). In our study, participants were also prompted to attend to gender information and participants were significantly better at detecting the male faces. It might be that the male face stimuli were intrinsically more salient or easier to detect, or that the participants chose them as a sort of "default" category in the judgment (similar to the interpretation of Drewes and colleagues in terms of the differences in detecting animal and vehicle stimuli). This interpretation is supported by the convergence level analysis, which showed a significantly higher performance in male versus female face detection tasks.

Compared with the one-target trials, face detection accuracy was generally enhanced, as expected by the presentation of two targets. In the dual-target trials, most participants exhibited a decreasing trend in HR as the ISI increased. Performance stabilized at around 130 ms, consistent with a TIW of around 100–150 ms (Wutz & Melcher, 2014). It is interesting to note that this TIW would correspond to one cycle of the 7.5-Hz behavioural oscillation (i.e., around 133 ms). Whereas two targets presented in separate oscillatory cycles facilitate performance by simple probability summation (see Treisman (1998) for example), those presented within the same cycle may be temporally integrated together as one salient target with a higher total accumulated signal.

It is important to reiterate that this data was collected as an online study. On the one hand, the online platform provided an opportunity to collect a more diverse sample of participants in terms of demographic variables such as countries of birth and current residence. On the other hand, the fact that participants all ran the study on their own computer inevitably introduced larger variability in the testing environment and less control over some experiment parameters. Although participants were instructed to block any environmental interference and stay an arm's length away from the screen, it was not possible to verify compliance with these instructions. In future, it would be useful to replicate these findings in laboratory settings.

In summary, the present study shows a 7.5-Hz fluctuation in rapid face detection performance, extending previous findings of behavioural oscillations in low-level visual perception to high-level vision tasks. These results are consistent with suggestions of a general theta-band oscillatory temporal organization for visual perception. Alternatively, the fact that our peak frequency was faster than that found in other studies with faces and other high-level categories (Drewes et al., 2015; Wang & Luo, 2017) suggests that the exact frequency of oscillation may depend either on specific task demands (e.g., difficult/easy) and/or stimulus properties (e.g., male/female faces or other categories).

CONFLICTS OF INTEREST
The authors declare no competing interests.

PEER REVIEW
The peer review history for this article is available at https://publons.com/publon/10.1111/ejn.15790.

DATA AVAILABILITY STATEMENT
Data and code for analyses are available at: https://github.com/xiaoyiliuXL/face-rsvp-datashare.

ORCID
Xiaoyi Liu https://orcid.org/0000-0001-5129-0683
Elio Balestrieri https://orcid.org/0000-0001-8090-1957
David Melcher https://orcid.org/0000-0003-0926-585X

REFERENCES
Balestrieri, E., Ronconi, L., & Melcher, D. (2021). Shared resources between visual attention and visual working memory are allocated through rhythmic sampling. *European Journal of Neuroscience*, 55, 3040–3053. https://doi.org/10.1111/EJN.15264

Battaglini, L., Mena, F., Ghiani, A., Casco, C., Melcher, D., & Ronconi, L. (2020). The effect of alpha tACS on the temporal resolution of visual perception. *Frontiers in Psychology*, 11, 1765. https://doi.org/10.3389/FPSYG.2020.01765

Bell, J., Burr, D. C., Crookes, K., & Morrone, M. C. (2020). Perceptual oscillations in gender classification of faces, contingent on stimulus history. *IScience*, 23(10), 101573. https://doi.org/10.1016/j.isci.2020.101573

Benwell, C. S. Y., London, R. E., Tagliabue, C. F., Veniero, D., Gross, J., Keitel, C., & Thut, G. (2019). Frequency and power of human alpha oscillations drift systematically with time-on-task. *NeuroImage*, 192, 101–114. https://doi.org/10.1016/J.NEUROIMAGE.2019.02.067

Berens, P. (2009). CircStat: A MATLAB toolbox for circular statistics. *Journal of Statistical Software*, 31(1), 1–21. https://doi.org/10.18637/JSS.V031.I10

Blais, C., Arguin, M., & Gosselin, F. (2013). Human visual processing oscillates: Evidence from a classification image technique. *Cognition*, 128(3), 353–362. https://doi.org/10.1016/j.cognition.2013.04.009

Brookshire, G. (2022). Putative rhythms in attentional switching can be explained by aperiodic temporal structure. *Nature Human Behaviour*, 1–12. https://doi.org/10.1038/s41562-022-01364-0

Busch, N. A., Dubois, J., & VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *Journal of Neuroscience*, 29(24), 7869–7876. https://doi.org/10.1523/JNEUROSCI.0113-09.2009

Chen, A., Wang, A., Wang, T., Tang, X., & Zhang, M. (2017). Behavioral oscillations in visual attention modulated by task difficulty. *Frontiers in Psychology*, 8, 1–9. https://doi.org/10.3389/fpsyg.2017.01630

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, 47(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Drewes, J., Zhu, W., Wutz, A., & Melcher, D. (2015). Dense sampling reveals behavioral oscillations in rapid visual categorization. *Scientific Reports*, 5(1), 1–9. https://doi.org/10.1038/srep16290

Dugué, L., Beck, A. A., Marque, P., & VanRullen, R. (2019). Contribution of FEF to attentional periodicity during visual search: A TMS study. *Eneuro*, 6(3), 1–10. https://doi.org/10.1101/414383

Dugué, L., Marque, P., & VanRullen, R. (2011a). Transcranial magnetic stimulation reveals attentional feedback to area V1 during serial visual search. *PLoS ONE*, 6(5), e19712. https://doi.org/10.1371/JOURNAL.PONE.0019712

Dugué, L., Marque, P., & VanRullen, R. (2011b). The phase of ongoing oscillations mediates the causal relation between brain excitation and visual perception. *Journal of Neuroscience*, 31(33), 11889–11893. https://doi.org/10.1523/JNEUROSCI.1161-11.2011

Dugué, L., Marque, P., & Vanrullen, R. (2015). Theta oscillations modulate attentional search performance periodically. *Journal of Cognitive Neuroscience*, 27(5), 945–958. https://doi.org/10.1162/jocn_a_00755

Dugué, L., Roberts, M., & Carrasco, M. (2016). Attention reorients periodically. *Current Biology*, 26(12), 1595–1601. https://doi.org/10.1016/j.cub.2016.04.046

Eick, C. M., Kovács, G., Rostalski, S. M., Röhrig, L., & Ambrus, G. G. (2020). The occipital face area is causally involved in identity-related visual-semantic associations. *Brain*

Structure and Function, *225*(5), 1483–1493. https://doi.org/10.1007/S00429-020-02068-9/FIGURES/3

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. (1998). What is "special" about face perception? *Psychological Review*, *105*(3), 482–498. https://doi.org/10.1037/0033-295X.105.3.482

Fiebelkorn, I. C., Foxe, J. J., Butler, J. S., Mercier, M. R., Snyder, A. C., & Molholm, S. (2011). Ready, set, reset: Stimulus-locked periodicity in behavioral performance demonstrates the consequences of cross-sensory phase reset. *Journal of Neuroscience*, *31*(27), 9971–9981. https://doi.org/10.1523/JNEUROSCI.1338-11.2011

Fiebelkorn, I. C., & Kastner, S. (2019). A rhythmic theory of attention. *Trends in Cognitive Sciences*, *23*(2), 87–101. https://doi.org/10.1016/j.tics.2018.11.009

Fiebelkorn, I. C., Saalmann, Y. B., & Kastner, S. (2013). Rhythmic sampling within and between objects despite sustained attention at a cued location. *Current Biology*, *23*(24), 2553–2558. https://doi.org/10.1016/j.cub.2013.10.063

Freschl, J., Melcher, D., Kaldy, Z., & Blaser, E. (2019). Visual temporal integration windows are adult-like in 5- to 7- year-old children. *Journal of Vision*, *19*(7), 5–5. https://doi.org/10.1167/19.7.5

Gaillard, C., & Hamed, S. B. (2020). The neural bases of spatial attention and perceptual rhythms. *European Journal of Neuroscience*, *55*, 3209–3223. https://doi.org/10.1111/EJN.15044

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

Grandy, T. H., Werkle-Bergner, M., Chicherio, C., Schmiedek, F., Lövdén, M., & Lindenberger, U. (2013). Peak individual alpha frequency qualifies as a stable neurophysiological trait marker in healthy younger and older adults. *Psychophysiology*, *50*(6), 570–582. https://doi.org/10.1111/PSYP.12043

Guo, B., Lu, Z., Goold, J. E., Luo, H., & Meng, M. (2020). Fluctuations of fMRI activation patterns in visual object priming. *Human Behaviour and Brain*, *1*, 78–84. https://doi.org/10.37716/HBAB.2020010601

Halgren, E., Raij, T., Marinkovic, K., Jousmäki, V., & Hari, R. (2000). Cognitive response profile of the human fusiform face area as determined by MEG. *Cerebral Cortex*, *10*(1), 69–81. https://doi.org/10.1093/CERCOR/10.1.69

Hanslmayr, S., Aslan, A., Staudigl, T., Klimesch, W., Herrmann, C. S., & Bäuml, K. H. (2007). Prestimulus oscillations predict visual perception performance between and within subjects. *NeuroImage*, *37*(4), 1465–1473. https://doi.org/10.1016/J.NEUROIMAGE.2007.07.011

Hanslmayr, S., Gross, J., Klimesch, W., & Shapiro, K. L. (2011). The role of alpha oscillations in temporal attention. *Brain Research Reviews*, *67*(1–2), 331–343. https://doi.org/10.1016/J.BRAINRESREV.2011.04.002

Helfrich, R. F., Fiebelkorn, I. C., Szczepanski, S. M., Lin, J. J., Parvizi, J., Knight, R. T., & Kastner, S. (2018). Neural mechanisms of sustained attention are rhythmic. *Neuron*, *99*(4), 854–865. https://doi.org/10.1016/j.neuron.2018.07.032

Huang, Y., Chen, L., & Luo, H. (2015). Behavioral oscillation in priming: Competing perceptual predictions conveyed in alternating theta-band rhythms. *Journal of Neuroscience*, *35*(6), 2830–2837. https://doi.org/10.1523/JNEUROSCI.4294-14.2015

Huth, A. G., Nishimoto, S., Vu, A. T., & Gallant, J. L. (2012). A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron*, *76*(6), 1210–1224. https://doi.org/10.1016/J.NEURON.2012.10.014

Jia, J., Liu, L., Fang, F., & Luo, H. (2017). Sequential sampling of visual objects during sustained attention. *PLoS Biology*, *15*(6), e2001903. https://doi.org/10.1371/journal.pbio.2001903

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *361*(1476), 2109–2128. https://doi.org/10.1098/rstb.2006.1934

Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 4396–4405. https://doi.org/10.1109/CVPR.2019.00453

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology: General*, *139*(3), 558–578. https://doi.org/10.1037/a0019165

Landau, A. N., & Fries, P. (2012). Attention samples stimuli rhythmically. *Current Biology*, *22*(11), 1000–1004. https://doi.org/10.1016/j.cub.2012.03.054

Liu, L., & Luo, H. (2019). Behavioral oscillation in global/local processing: Global alpha oscillations mediate global precedence effect. *Journal of Vision*, *19*(5), 1–12. https://doi.org/10.1167/19.5.12

Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., & Ro, T. (2009). To see or not to see: Prestimulus α phase predicts visual awareness. *Journal of Neuroscience*, *29*(9), 2725–2732. https://doi.org/10.1523/JNEUROSCI.3963-08.2009

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

Merholz, G., VanRullen, R., & Dugué, L. (2019). Oscillations modulate attentional search performance periodically. *Journal of Vision*, *19*(10), 279b. https://doi.org/10.1167/19.10.279B

Michel, R., Dugué, L., & Busch, N. A. (2020). Theta rhythmic attentional enhancement of alpha rhythmic perceptual sampling. BioRxiv. https://doi.org/10.1101/2020.09.10.283069

Michel, R., Dugué, L., & Busch, N. A. (2021). Distinct contributions of alpha and theta rhythms to perceptual and attentional sampling. *European Journal of Neuroscience*, *55*, 3025–3039. https://doi.org/10.1111/EJN.15154

Mohsenzadeh, Y., Qin, S., Cichy, R. M., & Pantazis, D. (2018). Ultra-rapid serial visual presentation reveals dynamics of feedforward and feedback processes in the ventral visual pathway. *eLife*, *7*, 1–23. https://doi.org/10.7554/elife.36329

Peelen, M. V., & Downing, P. E. (2017). Category selectivity in human visual cortex: Beyond visual object recognition. *Neuropsychologia*, *105*, 177–183. https://doi.org/10.1016/j.neuropsychologia.2017.03.033

Romei, V., Brodbeck, V., Michel, C., Amedi, A., Pascual-Leone, A., & Thut, G. (2008). Spontaneous fluctuations in posterior α-band EEG activity reflect variability in excitability of human visual areas. *Cerebral Cortex*, *18*(9), 2010–2018. https://doi.org/10.1093/cercor/bhm229

Romei, V., Gross, J., & Thut, G. (2010). On the role of prestimulus alpha rhythms over occipito-parietal areas in visual input regulation: Correlation or causation? *Journal of Neuroscience*, *30*(25), 8692–8697. https://doi.org/10.1523/JNEUROSCI.0160-10.2010

Ronconi, L., & Melcher, D. (2017). The role of oscillatory phase in determining the temporal organization of perception: Evidence from sensory entrainment. *Journal of Neuroscience*, *37*(44), 10636–10644. https://doi.org/10.1523/JNEUROSCI.1704-17.2017

Ronconi, L., Oosterhof, N. N., Bonmassar, C., Melcher, D., & Heeger, D. J. (2017). Multiple oscillatory rhythms determine the temporal organization of perception. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(51), 13435–13440. https://doi.org/10.1073/pnas.1714522114

Rossion, B., & Jacques, C. (2011). The N170: Understanding the time-course of face perception in the human brain. In *The Oxford handbook of event-related potential components* (pp. 115–142). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0064

Samaha, J., & Postle, B. R. (2015). The speed of alpha-band oscillations predicts the temporal resolution of visual perception. *Current Biology*, *25*(22), 2985–2990. https://doi.org/10.1016/j.cub.2015.10.007

Sharp, P., Melcher, D., & Hickey, C. (2019). Different effects of spatial and temporal attention on the integration and segregation of stimuli in time. *Attention, Perception, & Psychophysics*, *81*(2), 433–441. https://doi.org/10.3758/s13414-018-1623-7

Stigliani, A., Weiner, K. S., & Grill-Spector, K. (2015). Temporal processing capacity in high-level visual cortex is domain specific. *Journal of Neuroscience*, *35*(36), 12412–12424. https://doi.org/10.1523/JNEUROSCI.4822-14.2015

Treisman, A. (1998). Feature binding, attention and object perception. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *353*(1373), 1295–1306. https://doi.org/10.1098/RSTB.1998.0284

van Dijk, H., Schoffelen, J.-M., Oostenveld, R., & Jensen, O. (2008). Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *Journal of Neuroscience*, *28*(8), 1816–1823. https://doi.org/10.1523/JNEUROSCI.1853-07.2008

VanRullen, R. (2016). Perceptual cycles. *Trends in Cognitive Sciences*, *20*(10), 723–735. https://doi.org/10.1016/j.tics.2016.07.006

Wang, Y., & Luo, H. (2017). Behavioral oscillation in face priming: Prediction about face identity is updated at a theta-band rhythm. *Progress in Brain Research*, *236*, 211–224. https://doi.org/10.1016/bs.pbr.2017.06.008

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods*, *42*(3), 671–684. https://doi.org/10.3758/BRM.42.3.671

Wutz, A., Drewes, J., & Melcher, D. (2016). Nonretinotopic perception of orientation: Temporal integration of basic features operates in object-based coordinates. *Journal of Vision*, *16*(10), 3. https://doi.org/10.1167/16.10.3

Wutz, A., & Melcher, D. (2014). The temporal window of individuation limits visual capacity. *Frontiers in Psychology*, *5*, 952. https://doi.org/10.3389/fpsyg.2014.00952

Wutz, A., Muschter, E., van Koningsbruggen, M. G., Weisz, N., & Melcher, D. (2016). Temporal integration windows in neural processing and perception aligned to saccadic eye movements. *Current Biology*, *26*(13), 1659–1668. https://doi.org/10.1016/J.CUB.2016.04.070

Wutz, A., Weisz, N., Braun, C., & Melcher, D. (2014). Temporal windows in visual processing: "Prestimulus brain state" and "poststimulus phase reset" segregate visual transients on different temporal scales. *Journal of Neuroscience*, *34*(4), 1554–1565. https://doi.org/10.1523/JNEUROSCI.3187-13.2014

Yovel, G. (2016). Neural and cognitive face-selective markers: An integrative review. *Neuropsychologia*, *83*, 5–13. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2015.09.026

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.