



Perceptual foundation and extension to phase tagging for rapid invisible frequency tagging (RIFT)

Eelke Spaak, Floortje G. Bouwkamp, Floris P. de Lange

Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, The Netherlands

Corresponding Author: Eelke Spaak (eelke.spmaak@donders.ru.nl)

ABSTRACT

Recent years have seen the emergence of a visual stimulation protocol called Rapid Invisible Frequency Tagging (RIFT) in cognitive neuroscience. In RIFT experiments, visual stimuli are presented at a rapidly and sinusoidally oscillating luminance, using high refresh rate projection equipment. Such stimuli result in strong steady-state responses in visual cortex, measurable extracranially using EEG or MEG. The high signal-to-noise ratio of these neural signals, combined with the alleged invisibility of the manipulation, make RIFT a potentially promising technique to study the neural basis of visual processing. In this study, we set out to resolve two fundamental, yet still outstanding, issues regarding RIFT; as well as to open up a new avenue for taking RIFT beyond frequency tagging per se. First, we provide robust evidence that RIFT is indeed subjectively undetectable, going beyond previous anecdotal reports. Second, we demonstrate that full-amplitude luminance or contrast manipulation offer the best tagging results. Third and finally, we demonstrate that, in addition to frequency tagging, phase tagging can reliably be used in RIFT studies, opening up new avenues for constructing RIFT experiments. Together, this provides a solid foundation for using RIFT in visual cognitive neuroscience.

Keywords: rapid invisible frequency tagging (RIFT), magneto-encephalography (MEG), steady-state visual evoked potential/field (SSVEP/F), subjective invisibility, phase tagging

1. INTRODUCTION

The use of periodic visual stimuli to elicit robust stimulus-specific neural activity has a long history in visual cognitive neuroscience (Adrian & Matthews, 1934; Norcia et al., 2015; Vialatte et al., 2010). Typically, individual stimuli, ranging from simple LEDs to complex visual patterns, are flashed to participants at a particular presentation frequency (most commonly between 10 and 30 Hz). This reliably results in a steady-state visual evoked potential or field (SSVEP/SSVEF) at the presentation frequency over occipital regions, measurable by EEG or MEG, respectively. The SSVEP/F is known to have a high signal-to-noise ratio and such ‘tagging’ studies have

therefore proven a valuable addition to the visual cognitive neuroscientist’s toolkit (Vialatte et al., 2010).

Recently, driven by technical innovations in presentation equipment, several researchers have pointed out the distinct advantages of presenting rhythmic stimuli not as flashes at relatively low frequency, but as sinusoidally modulated stimuli at higher frequencies. This technique is often referred to as Rapid Invisible Frequency Tagging (RIFT) (Drijvers et al., 2021; Marshall et al., 2024; Seijdel et al., 2022; Zhigalov et al., 2019). In classical SSVEP/F studies using slower frequencies, the rhythmic modulation is highly apparent to the participants; i.e., they can clearly see the stimuli flicker. Furthermore, rhythmic stimulation at moderate frequencies is known to interact with

Received: 22 December 2023 Revision: 4 June 2024 Accepted: 21 June 2024 Available Online: 10 July 2024



The MIT Press

© 2024 The Authors. Published under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

Imaging Neuroscience, Volume 2, 2024
https://doi.org/10.1162/imag_a_00242

endogenous neural oscillations in the same frequency range (de Graaf et al., 2013; Haegens, 2020; Keitel et al., 2019; Mathewson et al., 2012; Obleser & Kayser, 2019; Spaak et al., 2014). This poses problems if one wishes to study these endogenous oscillations and their interaction with incoming stimuli (Seijdel et al., 2022). RIFT employs sinusoidally modulated stimuli around 60 Hz, corresponding to somewhat of a ‘sweet spot’ on the frequency axis. This frequency range is high enough that the resulting tagging is thought to be invisible to participants, yet low enough to still result in an appreciable peak in the spectrum of electromagnetic brain activity (resonant amplitude drops with increasing frequency; (Herrmann, 2001; Minarik et al., 2023)). The apparent invisibility allows RIFT to be employed in considerably more naturalistic settings than traditional SSVEP/F (in nature, strong and visible flicker is rare). Furthermore, narrowband modulated stimuli around 60 Hz do not entrain endogenous neural oscillations (Duecker et al., 2021), allowing the study of the latter in stimulated circumstances.

The promises (and existing accomplishments) of RIFT notwithstanding, whether these modulations are truly invisible has so far never been objectively tested. Going forward, RIFT needs a stronger foundation than the subjective, anecdotal, assertion of invisibility by researchers. Therefore, the first goal of the present study is to perform an objective test of the (in)visibility of RIFT. Furthermore, one can apply a sinusoidal modulation of stimulus material in several ways, each of which has a different consequence for the appearance of the stimulus on the screen, and possibly for the neural response as well. As a second objective, we set out to compare several such stimulation protocols. Finally, RIFT is typically focused on tagging different stimuli with different frequencies, as this allows a relatively straightforward analysis of the corresponding brain responses. Given the narrow ‘sweet spot’ of RIFT frequencies, it would be desirable if one could additionally manipulate tagging *phase* and reliably measure the resulting neural response as well. This would considerably increase tagging bandwidth. As a third objective, we therefore test the feasibility of this hitherto unexplored approach.

2. METHODS

We set out to answer the above questions by means of an experiment in which we presented RIFT-tagged stimuli to participants, tested their behaviour, and recorded MEG data.

2.1. Participants

We recruited 14 healthy participants (8 female; 6 male; approximate age 26 ± 6 years (mean \pm s.d.; only year of

birth recorded for privacy reasons), range 22–44 years) from the Radboud University participant pool. All participants had normal or corrected-to-normal vision. Participants received € 15 compensation for their participation. The study was approved by the local ethics committee (CMO Arnhem-Nijmegen, Radboud University Medical Center) under the general ethical approval for the Donders Centre for Cognitive Neuroimaging (‘Imaging Human Cognition’, CMO 2014/288). Participants provided written informed consent prior to the experiment.

Data from one participant was contaminated by a strong high-frequency drifting artifact due to MEG equipment malfunction. Data from another participant was of poor quality due to excessive head motion. Both these participants are excluded from further analyses, leaving a total of 12 participants on which the reported analyses are based.

2.2. Stimuli and task

Participants completed two different tasks: a discrimination task and a passive viewing task. In both tasks, stimuli were circular aperture square wave gratings, presented at a randomly chosen orientation of either -45° or $+45^\circ$. In passive viewing task blocks, the grating radius was 4 dva (degrees of visual angle), while in discrimination blocks, the radius was 2 dva. Gratings had a small circular central cutout (radius 0.4 dva in passive blocks, 0.2 dva in discrimination blocks) to facilitate fixation. In passive viewing blocks, a fixation dot was presented at all times in the center of the screen.

Each trial began with a fixation dot presented in isolation for a random duration between 500 and 800 ms. In the passive viewing task, this was followed by either one grating presented centrally or two gratings presented, one in either visual hemifield (Fig. 1B). Gratings were presented for 1200 ms. Before approximately every five trials (interval randomly chosen between 4 and 6 inclusive), the fixation dot would dim, indicating to participants that they had to press a button, in order to keep them alert and fixating. Additionally, participants were instructed that they should blink their eyes during these button press events, to reduce the number of blinks during trials of interest. Data from button press events was not included in any analyses.

In the discrimination task, the fixation dot was followed by three gratings presented on the screen (Fig. 1A), identical in all stimulus parameters (size, orientation, spatial frequency, etc.). The left and right grating served as ‘templates’ (i.e., response options), the center grating was a ‘target’. The left and right templates differed only in the tagging that was applied (e.g., 60 Hz vs no tagging), and the target had (no) tagging applied identically to

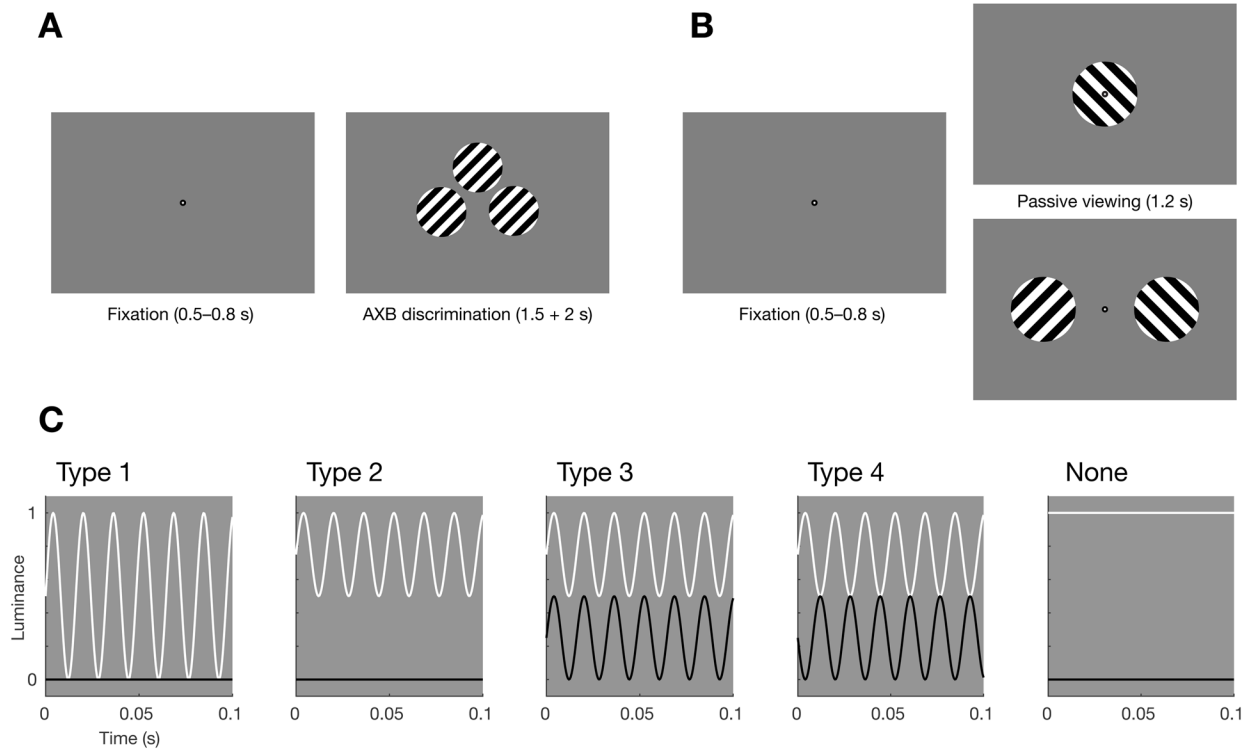


Fig. 1. Experimental design and tagging types. (A) Setup of the discrimination task. After a brief fixation period, participants were presented with two templates (left and right) and one target (middle top) and had to indicate with a button press whether the target was identical to either the left or right template. (B) Setup of the passive viewing task. After a brief fixation period, participants were presented with either a single central stimulus or two stimuli to either side of fixation. An occasional fixation dot dim (not shown) indicated that a button press was required, in order to maintain attention. (C) The four different tagging types, illustrated by plotting the luminance over time for the white and black bars of the grating stimuli. See text for details.

either the left or right template. Participants were instructed that the center grating was in some way identical to either the left or the right template grating, and their task was to press a left or right button to indicate which of these two templates they believed was identical to the target. Flicker was not explicitly mentioned as a feature to either attend or disregard. (Relevant instruction text: ‘You will see three stimuli. The middle one is in some way identical to either the leftmost or the rightmost (though you may not be able to see this). [...] Do not worry if this is (very) difficult, just try to make a well-informed guess if you can.’) Gratings were presented for 1500 ms, and participants had an additional 2000 ms after that during which to provide their response. Such a discrimination task is sometimes referred to as an ‘AXB’ task: the objective is to compare a target (X) to two provided templates (A/B) (Greenaway, 2017). Such a task should be more sensitive to pick up any perception of tagging in participants than a simple two-stimulus ‘pick the tagged one’ task, since at any time participants could compare the target to simultaneously presented reference templates. In contrast to the passive viewing task, there was no fixation dot presented with the stimuli during

the discrimination task, and participants were instructed (before the experimental session) that keeping fixation was not essential during the discrimination task.

2.3. Tagging manipulation

We applied RIFT according to four protocols or ‘tagging types’, which we refer to simply as types 1–4 (Fig. 1C). Type 1 corresponds to full amplitude luminance tagging: the white bands (i.e., those bands that are white in the untagged grating) of the grating oscillate between 0% and 100% luminance, while the black bands remain at 0% constant. Note that this leads to a time-average luminance of the white bands of 50%, or mid-grey. That means that the ‘white’ bands (including, importantly, the edges of the grating, which were constructed to always be white) will become invisible against a typical mid-grey background. This does not pose an issue for the tagging per se, but the stimulus would then appear clearly different from the other tagging type conditions (specifically, only black bars against a grey background). Therefore, to keep appearance more uniform (though, unavoidably, not entirely identical), we presented Type 1 trials always

against a full-black (0%) background, while the other tagging types are presented against a mid-grey background. Type 2 tagging corresponds to the white bands oscillating between 50% and 100% luminance (half amplitude luminance tagging), with black bands again at a constant 0% black. This has the advantage that a mid-grey background is again useable, since the ‘white’ bands are now at a time-average 75% grey, clearly distinguishable from the 50% background. Type 3 tagging corresponds to the white bands again oscillating between 50% and 100% luminance, while the black bands oscillate between 0% and 50% luminance in-phase. This protocol is what would naturally result if one constructs an arbitrary monochrome stimulus at 50% dynamic range and applies the tagging as a uniform additive luminance modulation on top. Type 4 tagging corresponds again to the white bands oscillating between 50% and 100% luminance, and the black bands between 0% and 50% luminance, but with the important difference that here the tagging oscillations are in anti-phase. In other words, Type 4 tagging corresponds to *contrast* tagging, rather than luminance tagging. This protocol corresponds to constructing an arbitrary monochrome stimulus at full dynamic range and manipulating its contrast (at full amplitude) over time.

Tagging was applied at 60 Hz, unless specified otherwise (see below). Typically, RIFT manipulation is performed by keeping tagging phase constant across trials, allowing for a straightforward analysis. Since one of the goals of this study was to test whether phase can reliably be reconstructed as well, the passive viewing task included both blocks with a constant phase across trials, as well as blocks where tagging phase was random across trials.

2.4. Experimental design

Participants first completed nine blocks of passive viewing trials, in random order. Tagging type and stimulus condition (one or two stimuli) were manipulated block-wise. Four blocks were single-stimulus (corresponding to tagging types 1–4), fixed phase; two blocks were single-stimulus (tagging types 1 and 4), random phase; two blocks were two-stimulus, with one stimulus at a random phase (sampled uniformly, without binning, between 0 and 360°), and the other at a 90° phase offset. In one block (single stimulus), no tagging was applied at all. Three additional passive viewing blocks were recorded, but not analyzed in the present study: two stimuli tagged at different frequencies (60/66 Hz) for two tagging types (1 and 4), and one block of two stimuli with no tagging. All single-stimulus passive viewing blocks consisted of 50 trials, while two-stimulus blocks consisted of 80 trials.

After the passive viewing blocks, participants engaged in two blocks that included an attentional manipulation (two tagged stimuli on the screen, spatial attention cue). Data from these blocks are not analyzed in the present study. (Any data not analyzed here is still included in the open data release that accompanies this study.)

Finally, participants completed three blocks (50 trials each) of the discrimination task. In two of these (tagging types 1 and 4), one template (left or right) stimulus was always not tagged, while the other was tagged at 60 Hz. In the third discrimination block, both templates were tagged (type 1), with one at 60 Hz and the other at 66 Hz. The target stimulus (presented simultaneously with the two templates) was always randomly chosen to be identical to either the left or right template.

2.5. Apparatus

Stimuli were presented using Matlab (The Mathworks, Inc., Natick, Massachusetts, United States) and custom-written scripts using the Psychophysics Toolbox (Brainard, 1997), and back-projected onto a translucent screen (48 × 27 cm) using a ProPixx projector (VPixx Technologies, Saint-Bruno, Québec, Canada) driven at a resolution of 960 × 540 and a refresh rate of 1440 Hz. Participants were seated at a distance of 85 cm from the screen.

Brain activity was recorded using a 275-channel axial gradiometer MEG system (VSM/CTF Systems, Coquitlam, British Columbia, Canada) in a magnetically shielded room. During the experiment, head position was monitored using three fiducial coils (nasion/left ear/right ear). Whenever participants’ head movement exceeded ~5 mm, the experiment was manually paused and the head position was shown to the participant, who would subsequently reposition (Stolk et al., 2013). Eye position and blinks were recorded using an Eyelink 1000 eye tracker (SR Research Ltd., Mississauga, Ontario, Canada). A tagging signal, if present, was always applied not only to the experimental stimuli, but also to a small square drawn in the bottom left corner of the experiment screen. This square was covered in a light-sensitive diode, and data from this sensor was recorded along with the MEG data. All data were on-line low-pass filtered at 300 Hz and digitized at a sampling rate of 1200 Hz. Immediately prior to the MEG session, participants’ headshape and the location of the three fiducial coils were digitized using a Polhemus 3D tracking device (Polhemus, Colchester, Vermont, United States) in order to facilitate subsequent source analysis.

2.6. Discrimination task data analysis

We analyzed the data from the three discrimination task blocks (60 Hz vs no tagging, Type 1; 60 Hz vs no tagging,

Type 4; 60 Hz vs 66 Hz, Type 1) separately. We analyzed the single trial data using a Bayesian hierarchical logistic regression model (Bernoulli family, logit link), modelling the single-trial response (left or right) as a function of the actually correct response (left or right), plus intercept. Both the intercept and sensitivity (i.e., the regression coefficient for the actually correct response) were included as subject-random factors. Priors for all parameters were chosen to be weakly informative (i.e., wide spread, with regression coefficients' priors centered on zero). Modelling and prior construction was done using the Bambi toolbox (Yarkoni & Westfall, 2016), with underlying Markov-chain Monte Carlo (MCMC) estimation done using PyMC (Salvatier et al., 2016). We report evidence for the null or alternative hypothesis as Bayes factors, BF_{01} (in favour of null) or BF_{10} (in favour of alternative). Trials with no response due to timeout were excluded from all analyses (2.3% of trials, median across participants). For graphical purposes (Fig. 2A), we additionally compute simple accuracy as the proportion of correct responses per participant. Before computing accuracy, we removed data from one participant, in one condition, since they only responded on one trial in that condition. These data (i.e., this single trial) is still included in the Bayesian analysis.

To put the Bayes factors for the observed data in perspective, we additionally conducted a sensitivity analysis by simulation. For a number of possible true underlying accuracies (from 0.4 to 0.6, inclusive, in steps of 0.01), we simulated a random dataset of identical size (number of participants, number of trials per participant) and spread (s.d. of observed accuracies across participants) as the observed data. We then subjected those simulated

datasets to the same analysis pipeline as above, and recorded the resulting BF_{01} . Per accuracy level, we conducted 300 such simulations with a different random initialization.

2.7. MEG data preprocessing and trial selection

All MEG preprocessing, spectral analyses, and source modelling were performed using custom-written scripts and the FieldTrip toolbox (Oostenveld et al., 2011). MEG data from the passive viewing blocks were segmented from -0.4 to $+1.2$ s around each stimulus display onset. Data were screened for outliers including eye blinks or eye movements, MEG SQUID jumps, and muscle artifacts, using a semi-automatic routine (FieldTrip's `ft_rejectvisual`), rejecting artifactual segments and/or excessively noisy MEG channels. This led to the removal of 72 ± 26 trials (mean \pm s.d. across participants) out of 750 trials total per participant ($9.6 \pm 3.4\%$). After artifact rejection, data were downsampled to 600 Hz (after applying an anti-aliasing filter and baseline correction relative to a window -0.3 to -0.1 s pre-stimulus) to speed up subsequent analyses. Infomax Independent Component Analysis (ICA; (Bell & Sejnowski, 1995)) was then used to clean the data of artifacts caused by ongoing cardiac activity and any residual eye movements.

2.8. MEG spectral analysis

Fourier spectra for the passive viewing trials were obtained by segmenting the passive viewing trials from 0.2 to 1.2 s after stimulus onset, yielding 1 s-long data segments and a corresponding 1 Hz frequency resolution.

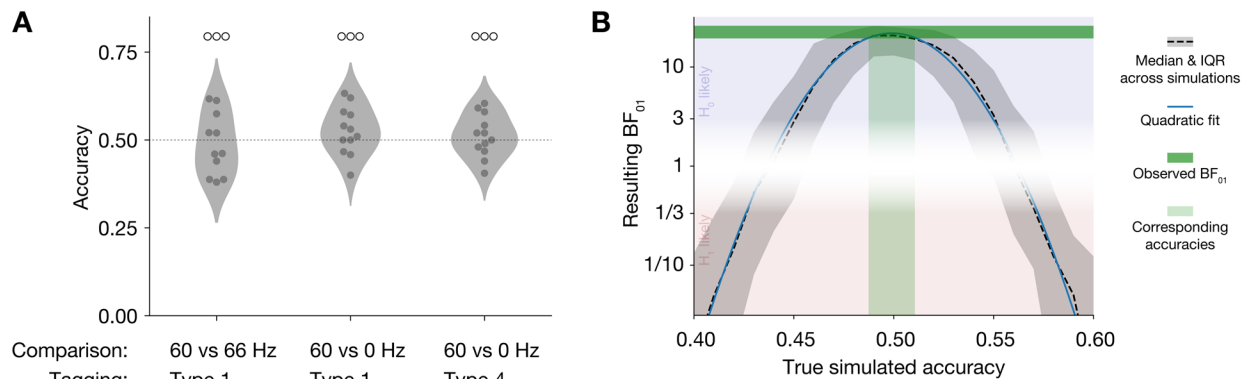


Fig. 2. Discrimination task behavioural results. (A) Accuracy for the three conditions of the AXB discrimination task. Dots reflect individual participants. Open circles above the density plots reflect strength of evidence for the null hypothesis of random guessing in hierarchical Bayesian logistic regression. $\circ\circ\circ$: $10 < BF_{01} < 30$. (B) Results for sensitivity analysis. BF_{01} are shown for varying true simulated observer accuracy levels. Y-axis is on a log₂-scale for interpretability. Dashed line shows median across simulations, shading is the inter-quartile range (IQR). Blue line shows a quadratic fit to the median data. Dark green band indicates the minimum and maximum (across the three conditions) observed BF_{01} ; light green band indicates those (median) accuracies that correspond to these BF_{01} .

A boxcar (flat) taper was used to minimize power smearing across neighbouring frequencies. Fourier spectra were then transformed to power spectra or inter-trial coherence (ITC) using standard methods, or to brain-to-tagging coherence. While ITC (Delorme & Makeig, 2004) measures the consistency of the phase of the MEG data across trials (which we report in Fig. 4 and corresponding analyses, for comparison), with brain-to-tagging coherence (labelled ‘phase-corrected’ in Fig. 4) we measure the consistency of the phase difference between a corrected tagging signal and the MEG data.

To obtain this corrected tagging signal, we use: (1) the stored tagging signal as constructed during the experimental session (i.e., the generated sinusoid that was used to tag the stimuli, and also stored to disk), (2) the data from the light sensor attached to the MEG projection screen, and (3) a recording of any missed ‘flips’ (i.e., graphics card buffer updates) during the experimental session (Brainard, 1997). The corrected tagging signal is constructed per trial by first identifying the time point at which the light sensor data ramps up above baseline levels. The stored tagging signal is aligned to that zero point. Then, for any flip that was missed by more than 1 ms, we introduce a matching delay in the corrected tagging signal. (We note that future experiments may employ better (monotonic, ideally linear) analog light sensors which may make this correction procedure redundant (one can then rely on a perfectly measured tagging signal), but the point stands that the stored tagging signal by itself is not optimal for brain-to-tagging coherence.)

To compute signal-to-noise ratio (SNR) from power or coherence spectra, we compared the power or coherence at the frequency of interest (foi) of 60 Hz to the average of neighbouring frequencies up to 5 Hz away, excluding the frequency bins immediately adjacent (i.e., neighbours are $[foi - 5, foi - 2] \cup [foi + 2, foi + 5]$), using decibel units:

$$SNR = 10 \times \log_{10} \frac{x_{foi}}{x_{neighbours}}$$

2.9. Disentangling orthogonal tagging phases

To analyze the data from the passive viewing task with two simultaneously presented stimuli, tagged at identical frequencies, but at a relative phase shift of 90°, we adopted the following four-step approach: (1) identify an optimal brain-to-tagging phase lag per participant; (2) project the data onto that phase lag and the corresponding 90° shifted lag; (3) determine the optimal single-dipole pattern explaining the phase-projected data; (4) test for differences between the resulting dipoles (Fig. 5A). We elaborate on each of the four steps below.

(1) We used the single-stimulus, random phase, data to estimate an optimal tagging-to-brain phase lag per participant. We first corrected the MEG data Fourier coefficients for the actual tagging phase ($F_{k,corr} = F_{k,uncorr} / F_{tagging}$, where k indexes channel), to take into account the random stimulation phase used across trials. Then, we projected these complex Fourier coefficients onto several possible phases (0 to 2π in steps of $\pi/16$). We here take ‘projection onto a phase ϕ ’ to mean the dot product of the Fourier coefficient (for a particular channel) with a phase factor:

$$\begin{aligned} A_{k,proj} &= F_{k,corr} \cdot e^{i\phi} \\ &= \text{Re}\{F_{k,corr}^* \times e^{i\phi}\} \\ &= \text{Re}\{F_{k,corr}\} \cos\phi + \text{Im}\{F_{k,corr}\} \sin\phi \end{aligned}$$

Different projected phases result in different projection strengths across MEG channels (Fig. 5A, top). The amount of total projected MEG variance (across channels; also known as global field power) is a measure of how strongly the brain responds to the stimulus at any particular phase lag. This allows the determination of the optimal tagging-to-brain phase lag per participant (Fig. 5A, middle). (2) Then, we applied the same tagging correction procedure to the two-stimulus, random phase, data, and projected the corrected complex Fourier coefficients onto the previously estimated optimal phase, and the relative 90° shifted phase.

(3) The two resulting phase-projected channel topographies (Fig. 5A, bottom) were then subjected to equivalent current dipole (ECD) source modelling (Hämäläinen et al., 1993) (Fig. 5B), with a single dipole per topography as a reasonable model for a response in early visual cortex (which we confirmed by inspecting the resulting explained variances). Starting location of the fit was always the central calcarine sulcus (MNI x, y, z coordinates (0, -88, -9) mm). Forward models for the source modelling were constructed by nonlinearly warping a template brain (and scalp surface) to the individual participants’ recorded Polhemus headshape data and constructing a single shell volume conduction model (Nolte, 2003). (4) The resulting dipole locations were compared in terms of their left/right displacement (MNI x-coordinates) as a critical test. If the phase tagging were successful, left visual stimuli should result in peaks in right visual cortex, and vice versa.

For illustration (Fig. 5D), we furthermore computed a 3D kernel density of resulting dipoles across participants, separately for the 0° (left stimulus) and 90° (right stimulus) cases, using a Gaussian kernel with (isotropic) s.d. 15 mm. We subtracted the two (normalized) densities and projected the difference onto a template MNI brain.

2.10. Statistical testing of spectral analysis and source modelling

To assess differences in SNR and ECD position and explained variance, we used hierarchical Bayesian linear models with intercepts as subject-random factors. Models and priors were specified using Bambi (Yarkoni & Westfall, 2016) and MCMC fits were performed using PyMC (Salvatier et al., 2016). Model convergence was verified by inspecting the traces (Kumar et al., 2019) and the Gelman–Rubin statistic \hat{r} .

3. RESULTS

3.1. Confirming the perceptual invisibility of RIFT

To assess whether participants could observe any difference between tagged and non-tagged stimuli, we adopted an ‘AXB’ paradigm. Participants were presented with two templates (A and B) and a target (X) and had to judge whether the target was identical (‘in some way’) to either A or B (Fig. 1A). This arguably provides a more sensitive test of subjective perception than merely asking participants to identify which of two stimuli is tagged (or whether a single stimulus is tagged or not) (Greenaway, 2017). There were three comparisons: no tagging versus 60 Hz tagging (Type 1, luminance tagging, see Figure 1C and Methods section), no tagging versus 60 Hz tagging (Type 4, contrast tagging), and 60 Hz versus 66 Hz tagging (Type 1). We analyzed these data using hierarchical Bayesian logistic regression, with both the intercepts and sensitivity (i.e., the amount by which the actually correct stimulus influences the response) as subject-random factors. In all three cases, we found strong evidence that participants performed at a level equivalent to random guessing (Fig. 2A; 60 vs 66 Hz tagging: $BF_{01} = 19.46$, 60 Hz vs no tagging (Type 1): $BF_{01} = 19.45$, 60 Hz vs no tagging (Type 4): $BF_{01} = 26.05$).

To put the observed evidence for the null hypothesis in perspective, we conducted a sensitivity analysis. For a range of possible true underlying detection accuracies, we simulated random data sets of identical size and spread as the observed data, and subjected these to the identical analysis pipeline as above. Results are shown in Figure 2B (dashed line; blue line indicates quadratic fit). As expected, the evidence for the null hypothesis of guessing (as expressed in BF_{01}) peaks when true performance is indeed around chance level (0.5), while a true performance that is clearly above or below chance results in evidence against the null hypothesis ($BF_{01} < 1/3$). The sensitivity analysis reveals that the BF_{01} for the actually observed data (horizontal green band in Fig. 2B) corresponds to a true (median) detection accuracy of between 0.49 and 0.51 (vertical green band). We thereby confirm

that the high observed BF_{01} likely corresponds to chance performance.

Finally, we note that at the end of the discrimination blocks, we verbally asked participants whether they noticed anything in particular about the stimuli while they made their decisions. All participants indicated they were entirely guessing; none reported any objective basis for making their left/right decisions. Taking the objective evidence and subjective reports together, we therefore conclude that people indeed cannot see whether RIFT is applied or not.

3.2. Full luminance and contrast tagging result in the strongest neural response modulation

We next focused on comparing the different tagging types (Fig. 1C; see also Methods section) in terms of the resulting neural responses when there is a single tagged (fixed phase) stimulus on the screen which is passively attended (Fig. 1B, top). We assessed the neural response modulation both in terms of power and in terms of brain-to-tagging coherence. Power spectra for the different tagging types are shown in Figure 3A. Clearly identifiable are peaks in the alpha range (~10 Hz), power line noise (50 Hz) and, importantly, at the tagging frequency of 60 Hz. The peak at 60 Hz is absent for the no-tagging condition (most visible in the inset). Coherence spectra (Fig. 3B) similarly show a clear peak at 60 Hz for all tagging types.

To make direct comparisons not just among the tagging types, but also between coherence and power, we expressed the tagging as signal-to-noise ratio (SNR) by contrasting the activity at the frequency of interest to neighbouring frequencies. Average results are shown in Figure 3C. Brain-to-tagging coherence analyses provided a considerably higher SNR than power (2.79 ± 0.21 dB difference; posterior mean \pm s.d.; $BF_{10} > 10^4$). For power, tagging types 1, 2, and 4 showed demonstrable tagging responses (vs no tagging: $BF_{10} > 10^4$), while tagging type 3 resulted in inconclusive evidence ($BF_{10} = 1.87$). For coherence, all tagging types showed clear tagging responses (all $BF_{10} > 10^4$). For both power and coherence, tagging types 1 and 4 resulted in the highest SNR, though with not all pairwise comparisons presenting conclusive evidence (see bars in Fig. 3C). Tagging types 1 and 4 resulted in equally high coherence ($BF_{01} = 59.54$), while their difference was inconclusive for power ($BF_{10} = 1.15$). Overall, we conclude that brain-to-tagging coherence analysis is preferable over power analysis. Furthermore, tagging types 1 (full luminance tagging) and 4 (contrast tagging) should be preferred over types 2 and 3, when possible. The average sensor topographies (Fig. 3D) corroborate this conclusion.

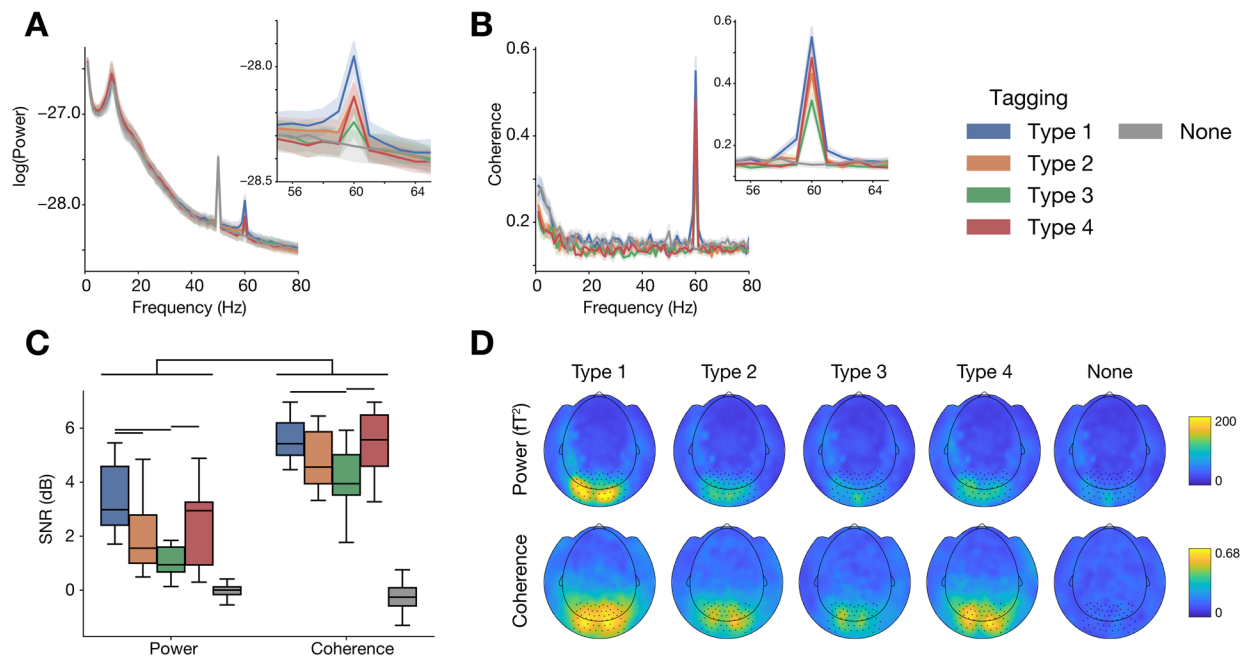


Fig. 3. Tagging results for different tagging types (fixed phases, single stimulus on screen). (A) Power spectra (log₁₀-transformed) for the four tagging types and the no-tagging condition. Noticeable are peaks in the alpha frequency range (~10 Hz), power line noise (50 Hz) and, importantly, the tagging frequency (60 Hz; only for the tagging conditions). Shading reflects standard error of the mean across participants. Inset shows a zoomed-in version of the spectrum around 60 Hz. Spectra reflect a large set of parieto-occipital sensors, highlighted in panel D. (B) Spectra of brain-to-tagging coherence, colours and conventions as in panel A. (C) Power and coherence at the frequency of interest (60 Hz), expressed as signal-to-noise ratio (SNR). Box plots reflect the median and inter-quartile range across participants. Horizontal bars indicate differences for which the comparison yields $BF_{10} > 3$. (D) Topography of MEG responses at 60 Hz. Highlighted sensors are those averaged over for the other panels.

3.3. Brain-to-tagging coherence allows the use of random tagging phases

To date, RIFT studies have always employed a fixed tagging phase across trials (Drijvers et al., 2021; Zhigalov et al., 2019). This enables the use of a straightforward coherence-based analysis, namely inter-trial coherence (ITC). One of the goals of the present study was to assess whether phase manipulation could be used for tagging, in addition to frequency. By definition, phase manipulation precludes the use of fixed phases. Therefore, we first tested whether non-constant (i.e., random) tagging phases across trials during the experiment would still allow a robust readout of the tagging signal. We expect the classical ITC analysis to break down here (since it relies on a constant phase). Therefore, we analyzed the data using brain-to-tagging phase corrected coherence, taking the exact tagging phase into account (see Methods section for details), and compared the results to ITC. We again focus on the single-stimulus trials here. Given the established equivalence between tagging types 1 and 4, we aggregate across these for the analysis.

There was a clear interaction between tagging phase (fixed/random) and analysis method (ITC/phase-corrected; Fig. 4; $BF_{10} > 10^4$). As expected, ITC analysis

does not detect any tagging signal when phase is randomized across trials (posterior mean SNR \pm s.d.: 0.10 ± 0.25 dB, $BF_{01} = 54.95$), but performs well with constant phases (3.73 ± 0.24 dB, $BF_{10} > 10^4$). The evidence for a decrease in ITC with phase randomization was strong ($BF_{10} = 42.31$). Using phase-corrected brain-to-tagging coherence analysis, tagging is clearly present for both the fixed-phase trials (5.52 ± 0.24 dB, $BF_{10} > 10^4$) and the random-phase trials (4.72 ± 0.24 dB, $BF_{10} > 10^4$), with no evidence for a difference between the two ($BF_{10} = 1.38$). From this, we conclude that it is indeed feasible to use a random phase across trials in RIFT studies, and that the neural tagging response can reliably be read out, as long as one takes into account the actual phase that was used in stimulation.

Interestingly, even with a fixed phase across trials, including a phase correction factor substantially increased SNR compared to simple ITC (difference 1.80 ± 0.26 dB, $BF_{10} > 10^4$). This is presumably due to subtle but non-constant delays in the experimental apparatus. Measuring the stimulation phase and compensating the tagging reference signal for this allows the recovery of higher SNR, so would in general be recommended, even with constant phases across trials.

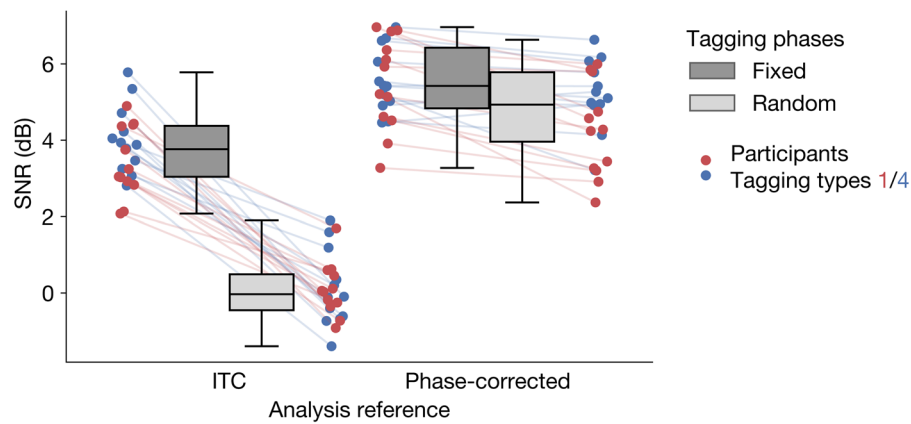


Fig. 4. Tagging signal-to-noise ratio (SNR) for fixed and random phase trials, analyzed using inter-trial coherence (ITC) or brain-to-tagging phase corrected coherence.

3.4. Orthogonal phase tagging disentangles multiple same-frequency stimuli

Having established that the use of arbitrary phases across trials for the tagging signal is feasible, we next asked whether phase can be used to uniquely tag different stimuli on the screen. To test this, we presented two stimuli simultaneously to participants (Fig. 1B). Both stimuli were tagged at 60 Hz. The phase of the two tagging signals was randomly chosen per trial, with the constraint that the right stimulus always was tagged at a 90° shift relative to the left stimulus.

Since the tagging-to-brain phase lag is different for different participants, properly extracting phase information requires an estimate of this lag. We used the single-stimulus, random phase trials described previously as independent data to perform this estimate (see Methods section and Fig. 5A for details). Briefly, projecting the single-stimulus data onto various phases allowed an estimation of the peak tagging-to-brain phase lag for each participant. We then projected the two-stimulus data onto that same phase lag, as well as onto the relative 90° shifted phase. By hypothesis, the left stimulus tagging signal should be most strongly visible when projecting at the (independently estimated) peak phase lag, while the right stimulus tagging signal should be dominant when looking at the 90° shifted phase. Taking advantage of the known contralateral projection of visual hemifields, we reasoned that if this hypothesis were true, the non-shifted signal (i.e. the left stimulus) should result in a peak in the right visual cortex, while the 90°-shifted signal (i.e. the right stimulus) should peak in the left visual cortex. We tested this prediction by performing equivalent current dipole (ECD) modelling on the phase-projected channel topographies.

Across participants, dipole fits were of high quality (explained variance $r^2 = 0.78 \pm 0.034$, posterior

mean \pm s.d.), with no difference between the peak and 90° shifted phases ($\Delta r^2 = 0.053 \pm 0.037$, $BF_{01} = 6.46$). The resulting dipoles for an example (representative) participant are shown in Figure 5B. Visible are peaks in bilateral occipital cortex, with the left signal peaking in the right hemisphere and vice versa. Critically, also across participants, comparing the MNI x-coordinates for the fits corresponding to the different phase lags revealed that right stimuli peaked in left visual cortex, and vice versa ($\Delta x = 16.39 \pm 4.0$ mm, $BF_{10} = 448.51$; Fig. 5C, D). We can thus conclude that two stimuli presented simultaneously, tagged with RIFT at an identical frequency but with orthogonal phases, can reliably be disentangled.

4. DISCUSSION

4.1. Overall conclusions

We set out to establish a solid empirical foundation for the novel experimental technique of RIFT. First, we confirmed that the ‘I’ in ‘RIFT’ is appropriate: sinusoidally modulated stimuli around 60 Hz are indeed indistinguishable from stimuli that are not modulated at all; the tagging is invisible. Second, we compared four different possible tagging protocols and found that two of them (full-amplitude luminance tagging and contrast tagging) clearly outperform the others. Third, in a direct quantitative comparison, we found that coherence analyses result in higher SNR than power, while power analysis nonetheless still yields an appreciable tagging response. Fourth, we demonstrated that a fixed phase across trials (as typically used) is not a hard requirement for RIFT. Fifth, and finally, we leveraged this potential for non-fixed phases to lay the groundwork for applying phase tagging in future RIFT studies, thereby considerably expanding the potential tagging signal bandwidth.

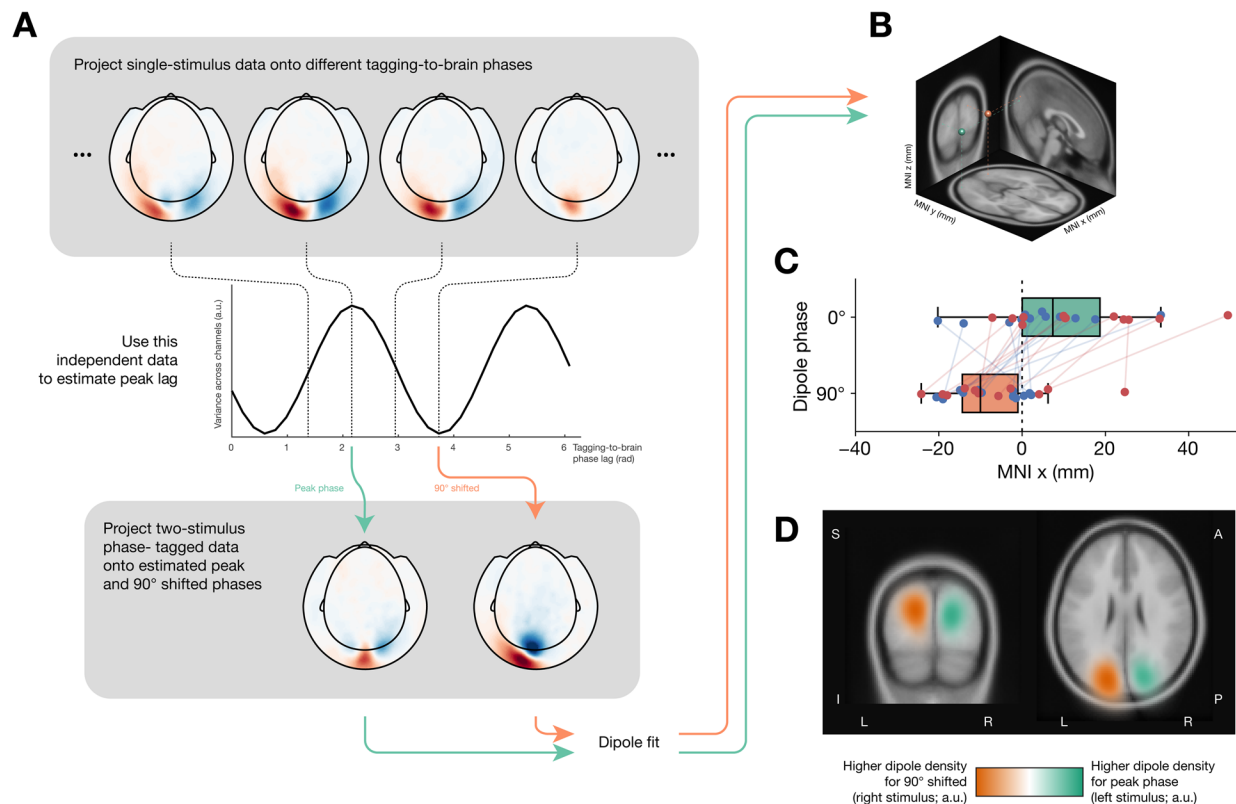


Fig. 5. Using RIFT phase tagging. (A) Analysis pipeline. Per-channel complex coherency for the single-stimulus (random phase) data is projected onto different phase lags (top). This yields a profile of tagging-to-brain phase lag (middle). The coherency for the two-stimulus data (both stimuli tagged at 60 Hz, always at 90° phase relative to each other) are projected onto both the peak phase lag estimated in this fashion, and the phase 90° away. The resulting projected topographies (bottom) are then subjected to equivalent current dipole (ECD) fitting. (B) The two resulting dipoles for a single representative participant. Green corresponds to the fit for the stimulus left on the screen (phase ϕ), orange corresponds to the stimulus right on the screen (phase $\phi + 90^\circ$). (C) X-coordinates in Montreal Neurological Institute (MNI) space for the ECD fits for all participants. Left stimulus (green) ends up predominantly in the right hemisphere ($x > 0$), while the right stimulus (orange) ends up predominantly in the left hemisphere ($x < 0$). Red/blue dots are individual participants, tagging types 1 (red) and 4 (blue). (D) Volumetric density plot of all ECD fits, difference between the left and right stimulus, projected onto the template MNI brain. Peaks for the stimuli are clearly shown in contralateral occipital cortex.

4.2. RIFT's subjective invisibility

Concluding subjective invisibility from objective task performance is a notoriously difficult problem in psychophysics (Schmidt, 2015; Vadillo et al., 2022). We believe the conclusion of RIFT's invisibility is justified here for several reasons. First, we deliberately used an experimental paradigm where a target had to be compared to two templates that were all on the screen simultaneously (an 'AXB' task, see (Greenaway, 2017)). This setup eliminates the problematic role of conservative or liberal biases that are present in a more typical one-stimulus setting (where the question may have been 'is this flickering or not?'). Second, relatedly, discrimination performance is known to be higher in cases where stimuli are presented simultaneously (Gibson, 1969; Mundy et al., 2009), making the null evidence we found more striking. Third, although we report accuracies in Figure 2, the main evi-

dence for invisibility stems from hierarchical Bayesian logistic regression, in which sensitivity (i.e., the amount by which the actually correct option influences the response) can be well isolated from any remaining bias. Fourth, simulations demonstrated that this analysis pipeline indeed would only reveal very strong evidence in favour of invisibility (as we observed) when actual performance is indistinguishable from chance level. Fifth, and finally, all participants indicated at debriefing that they had no idea on what to base their responses and were entirely guessing. It is noteworthy that the discrimination task was performed at the end of the experimental session, after participants had already experienced RIFT-tagged stimuli for hundreds of trials. Taking all this evidence together, we conclude that RIFT indeed is invisible.

We reach the above quantitative conclusion on the basis of testing only one particular stimulus class, i.e.,

gratings. We cannot rule out that other, more complex, types of stimuli would result in above-chance ability to detect RIFT, if tested in a similar discrimination task. However, we believe that possibility unlikely for two reasons. First, there is no reason to assume that the present results would *not* generalize to more complex stimuli. If anything, the subtle nature of the RIFT manipulation may be more likely to go unnoticed when embedded in richer stimuli than when applied to an otherwise uneventful grating. Second, previous (anecdotal) reports of RIFT's invisibility apply not just to studies involving gratings, but also those involving richer naturalistic stimuli (Zhigalov et al., 2019), including videos (Drijvers et al., 2021). Nonetheless, future studies may wish to explicitly test whether the presentation of tagged complex stimuli indeed retains subjective invisibility.

Although we were here concerned with RIFT's *subjective* discriminability, of further interest is the extent to which high-frequency flickering stimuli may *implicitly* affect (attentional) performance. It has previously been shown that visual flicker at frequencies that render it subjectively invisible is also ineffective in exogenously cueing spatial attention (Alais et al., 2016), corroborating RIFT's invisibility.

4.3. Comparing tagging types

Perhaps unsurprisingly, applying luminance modulation at 100% amplitude (Type 1) resulted in the strongest neural response. A drawback of this type of tagging is that any stimulus that appears white (100%) before applying the modulation is rendered as grey when tagging is applied (time-averaged 50%). Since 50% grey is an often-used background colour in psychophysical experiments, this may not be ideal. Furthermore, how to apply this type of tagging with more complex stimuli than black and white gratings is not well-defined. Finally, we note that, to keep stimulus appearance comparable across conditions, tagging Type 1 was tested against a black background, unlike the other tagging types (see [Methods](#) section). This could potentially have played a role in yielding the strongest neural response for this type of tagging. We examined two flavours of half-amplitude modulation (Types 2 and 3) which may alleviate some of the above issues, but found that these resulted in considerably weaker neural tagging responses. Fortunately, we found that contrast modulation (Type 4) produces a tagging response that is equally strong as full-amplitude luminance modulation, while lending itself well to more diverse stimuli. We would recommend future studies to use contrast-modulated tagging in general. Only in specific cases, where stimuli happen to lend themselves well to full-amplitude luminance tagging, would we recom-

mend the use of that tagging type. One interesting use case may be to tag part of a uniform 50% grey background using Type 1 tagging (0-100% modulation). This would result in an invisibly tagged patch of background, without an apparent stimulus, in essence exploiting the limitation we sketched above for positive effect.

4.4. Coherence versus power

The majority of RIFT studies to date has relied on some form of coherence analysis to extract the tagging responses (Drijvers et al., 2021; Zhigalov et al., 2019). By computing signal-to-noise ratios (SNR), we were able to directly compare coherence and power estimates of tagging. This way, we confirmed that, indeed, coherence is a more sensitive measure than power. However, it is worth noting that power analyses still resulted in clear tagging responses. Power analyses are more flexible than coherence; for example, they are easily applied at the single-trial level, unlike coherence. Therefore, when such flexibility is required, it is feasible to resort to power analysis (see, e.g., (Bouwkamp et al., 2023)).

4.5. Leveraging phase variability

Past RIFT studies have relied on a constant stimulation phase across trials to facilitate subsequent analysis, either through inter-trial coherence (ITC; also known as phase-locking factor; (Delorme & Makeig, 2004)) or through computing the power spectrum of the event-related field (ERF). While this indeed makes analyses straightforward, it limits the potential bandwidth by which one can apply RIFT; specifically, it precludes using phase as a tagging information channel. We confirmed that, indeed, these typical analyses break down when phase is non-constant. Importantly, we demonstrated that a relatively straightforward extension of the fixed-phase analysis method (which we call 'phase-corrected coherence' or 'tagging-to-brain coherence') is able to recover strong neural tagging signals, even in the case of non-constant phases. Interestingly, this technique resulted in higher SNR than ITC even with constant tagging phases, possibly due to non-constant delays in the presentation equipment, leading us to recommend its use in general.

One clear use case for non-constant phases in RIFT is to tag multiple stimuli at the same frequency, but with different phases. Since there is a narrow 'sweet spot' on the frequency axis for RIFT (high enough to be invisible, but low enough to result in good neural SNR), such a doubling of tagging bandwidth can be highly desirable. Relying on our discovery that random phases are feasible, we next demonstrated that, indeed, two stimuli simultaneously presented and tagged at the same frequency, but at

orthogonal phases, can reliably be disentangled in the neural response. It would be interesting to see future work examine the feasibility of exploiting the phase axis even further, beyond two orthogonal phases. Furthermore, our current approach relied on using independent data to estimate an individual's tagging-to-brain phase lag before disentangling the neural responses to two stimuli. The amount of data required for such an estimate is low (just a few minutes should be sufficient), so the extra experimental burden for researchers wishing to go down this route is light. Nonetheless, future work may explore different avenues for estimating the optimal phase lag, possibly without the use of independent (in essence 'phase localizer') data.

4.6. Spatial extent of neural entrainment?

Recent MEG studies employing RIFT have reported that the peak of the tagging response is located in primary visual cortex (V1) (Duecker et al., 2021; Zhigalov & Jensen, 2020), which is in line with the present results (Fig. 5B, D). The cited studies show no clear propagation of the tagging response to higher-order (visual) cortex. If true, this limits the applicability of RIFT to settings where the excitability of specifically early visual cortex is of interest. However, we note that this evidence relies on beamformer source reconstruction, which is geared towards isolating peak responses. The spread of beamformer-identified source maps depends strongly on the SNR of the neural recording (Gross et al., 2001), so a narrow peak in V1 may reflect either the absence of spread to other regions, and/or a high-SNR source in V1 itself. Such a high-SNR V1 source may obscure sources in other (nearby) cortical areas. Therefore, we consider the evidence for a limited spread beyond early visual cortex based on MEG and beamforming inconclusive, at present.

More direct evidence for such a limited spread was recently put forward by Schneider et al. (2023), who recorded intracranially from mouse V1, lateral geniculate nucleus (LGN), and hippocampal CA1, while presenting the animals with high-frequency visual flicker. They found that the flicker response was strong in LGN, weaker in V1, and absent from CA1. In V1 specifically, the response was strongest in the granular layer, and much reduced in the other (deep/superficial) layers (from which V1 output projections originate). This is again taken as evidence that the flicker response likely does not propagate beyond V1. Although this is stronger evidence than that based on beamforming, two issues remain before it should be considered conclusive. First, the MEG signal is thought to rely primarily on the magnetic field caused by primary electric currents along the apical dendrites of layer five pyramidal cells (Baillet, 2017). Since we clearly measure a tagging

response in MEG, at least these dendritic currents in layer five cells are likely entrained as well. Second, although to date no studies have conclusively demonstrated a clear high-frequency RIFT response beyond early visual cortex, a so-called 'intermodulation peak' as a consequence of visual RIFT has been identified in fronto-temporal regions (Drijvers et al., 2021; Seijdel et al., 2024), suggesting at least some transmission of the flicker response beyond (early) visual cortex (perhaps depending on the use of naturalistic vs artificial stimuli). Taken together, we consider the spatial extent of cortical entrainment due to RIFT an important, as yet unsettled, issue.

(It is finally worth noting that, even if the RIFT response itself turns out to be exclusively generated in early visual cortex, that feedforward neural response is still subject to the consequences of activity in higher-order regions, such as attentional modulation or other feedback processes. Therefore, even in the case of strictly anatomically limited RIFT responses, the technique likely has value in cognitive neuroscience.)

4.7. Outlook

Previous authors have emphasized the advantages of RIFT, like its leaving endogenous neural oscillations unperturbed, and of course its (hitherto alleged) invisibility (Seijdel et al., 2022). Invisible flicker by itself is already much more naturalistic than classical SSVEP/F paradigms, with highly apparent luminance modulation. We speculate that RIFT may offer a unique way forward in naturalistic visual cognitive neuroscience. The complex stimuli inherent in the natural visual world, and therefore those used in naturalistic experiments, do not lend themselves as easily to the rapid 'readout' of neural responses in M/EEG (whether via univariate or multivariate analyses) as well-controlled experimental stimuli (Sonkusare et al., 2019). There is often simply too much going on to be able to reliably say that this particular neural response was a consequence of that visual stimulus. With RIFT, one can tag only certain elements of the visual scene (or movie), without reducing the level of naturalism, and still be able to reliably identify the neural response corresponding to only specific spatial locations or objects. This is a considerable and unique advantage. Of course, further methodological work is needed before RIFT's potential can be leveraged in naturalistic settings (e.g., the above considerations on the spatial extent of neural entrainment, and the unknown strength of the tagging response for objects within natural scenes). Nonetheless, we expect to see a happy marriage between RIFT and naturalistic visual cognitive neuroscience. With such a promising future ahead, RIFT does need a solid empirical foundation, and we hope to have provided one in this paper.

DATA AND CODE AVAILABILITY

All recorded data (with the exception of privacy-sensitive headshape data) are available from the Donders/Radboud Data Repository under a CC-BY license at <https://doi.org/10.34973/8dx5-7e51>. All experiment and analysis code are available on GitHub under a BSD license at <https://github.com/Spaak/rift-phase-and-visibility>.

AUTHOR CONTRIBUTIONS

E.S: Conceptualization, Methodology, Software, Investigation, Writing – Original Draft, Writing – Review & Editing, Visualization, Supervision. F.G.B.: Conceptualization, Investigation. F.P. dL.: Conceptualization, Writing – Review & Editing, Funding Acquisition.

FUNDING

This work was partially funded by the EC Horizon 2020 program (ERC starting grant 678286 awarded to F.P.dL).

DECLARATION OF COMPETING INTEREST

We declare no competing interests.

ACKNOWLEDGEMENTS

We would like to thank Virág Fodor for help in data acquisition.

REFERENCES

- Adrian, E. D., & Matthews, B. H. C. (1934). The Berger rhythm: Potential changes from the occipital lobes in man. *Brain*, 57(4), 355–385. <https://doi.org/10.1093/brain/57.4.355>
- Alais, D., Locke, S. M., Leung, J., & Van der Burg, E. (2016). No attentional capture from invisible flicker. *Scientific Reports*, 6(1), 29296. <https://doi.org/10.1038/srep29296>
- Baillet, S. (2017). Magnetoencephalography for brain electrophysiology and imaging. *Nature Neuroscience*, 20(3), Article 3. <https://doi.org/10.1038/nn.4504>
- Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7(6), 1129–1159. <https://doi.org/10.1162/neco.1995.7.6.1129>
- Bouwkamp, F. G., Lange, F. P. de, & Spaak, E. (2023). Spatial predictive context speeds up visual search by biasing local attentional competition. *bioRxiv*. <https://doi.org/10.1101/2023.07.14.548976>
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436. <https://doi.org/10.1163/156856897x00357>
- de Graaf, T. A., Gross, J., Paterson, G., Rusch, T., Sack, A. T., & Thut, G. (2013). Alpha-band rhythms in visual task performance: Phase-locking by rhythmic sensory stimulation. *PLoS One*, 8(3), e60035. <https://doi.org/10.1371/journal.pone.0060035>
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- Drijvers, L., Jensen, O., & Spaak, E. (2021). Rapid invisible frequency tagging reveals nonlinear integration of auditory and visual information. *Human Brain Mapping*, 42(4), 1138–1152. <https://doi.org/10.1002/hbm.25282>
- Duecker, K., Gutteling, T. P., Herrmann, C. S., & Jensen, O. (2021). No evidence for entrainment: Endogenous gamma oscillations and rhythmic flicker responses coexist in visual cortex. *Journal of Neuroscience*, 41(31), 6684–6698. <https://doi.org/10.1523/JNEUROSCI.3134-20.2021>
- Gibson, E. J. (1969). *Principles of perceptual learning and development*. Appleton-Century-Crofts. <https://psycnet.apa.org/record/1969-35014-000>
- Greenaway, R. E. (2017). Chapter 13—ABX discrimination task. In L. Rogers (Ed.), *Discrimination testing in sensory science* (pp. 267–288). Woodhead Publishing. <https://doi.org/10.1016/B978-0-08-101009-9.00013-7>
- Gross, J., Kujala, J., Hämäläinen, M., Timmermann, L., Schnitzler, A., & Salmelin, R. (2001). Dynamic imaging of coherent sources: Studying neural interactions in the human brain. *Proceedings of the National Academy of Sciences*, 98(2), 694–699. <https://doi.org/10.1073/pnas.98.2.694>
- Haegens, S. (2020). Entrainment revisited: A commentary on Meyer, Sun, and Martin (2020). *Language, Cognition and Neuroscience*, 35(9), 1119–1123. <https://doi.org/10.1080/23273798.2020.1758335>
- Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography—Theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, 65(2), 413–497. <https://doi.org/10.1103/RevModPhys.65.413>
- Herrmann, C. S. (2001). Human EEG responses to 1–100 Hz flicker: Resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Experimental Brain Research*, 137(3), 346–353. <https://doi.org/10.1007/s002210100682>
- Keitel, C., Keitel, A., Benwell, C. S. Y., Daube, C., Thut, G., & Gross, J. (2019). Stimulus-driven brain rhythms within the alpha band: The attentional-modulation conundrum. *Journal of Neuroscience*, 39(16), 3119–3129. <https://doi.org/10.1523/JNEUROSCI.1633-18.2019>
- Kumar, R., Carroll, C., Hartikainen, A., & Martin, O. A. (2019). ArviZ a unified library for exploratory analysis of Bayesian models in Python. *The Journal of Open Source Software*, 4(33), 1143. <https://doi.org/10.21105/joss.01143>
- Marshall, T. R., Ruesseler, M., Hunt, L. T., & O'Reilly, J. X. (2024). The representation of priors and decisions in the human parietal cortex. *PLOS Biology*, 22(1), e3002383. <https://doi.org/10.1371/journal.pbio.3002383>
- Mathewson, K. E., Prudhomme, C., Fabiani, M., Beck, D. M., Lleras, A., & Gratton, G. (2012). Making waves in the stream of consciousness: Entraining oscillations in EEG alpha and fluctuations in visual awareness with rhythmic visual stimulation. *Journal of Cognitive Neuroscience*, 24(12), 2321–2333. <https://doi.org/10.1162/jocn.a.00288>
- Minarik, T., Berger, B., & Jensen, O. (2023). Optimal parameters for rapid (invisible) frequency tagging using MEG. *NeuroImage*, 281, 120389. <https://doi.org/10.1016/j.neuroimage.2023.120389>

- Mundy, M. E., Honey, R. C., & Dwyer, D. M. (2009). Superior discrimination between similar stimuli after simultaneous exposure. *Quarterly Journal of Experimental Psychology*, 62(1), 18–25. <https://doi.org/10.1080/17470210802240614>
- Nolte, G. (2003). The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Physics in Medicine and Biology*, 48(22), 3637–3652. <https://doi.org/10.1088/0031-9155/48/22/002>
- Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottureau, B. R., & Rossion, B. (2015). The steady-state visual evoked potential in vision research: A review. *Journal of Vision*, 15(6), 4. <https://doi.org/10.1167/15.6.4>
- Obleser, J., & Kayser, C. (2019). Neural entrainment and attentional selection in the listening brain. *Trends in Cognitive Sciences*, 23(11), 913–926. <https://doi.org/10.1016/j.tics.2019.08.004>
- Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011, 1–9. <https://doi.org/10.1155/2011/156869>
- Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. (2016). Probabilistic programming in Python using PyMC3. *PeerJ Computer Science*, 2, e55. <https://doi.org/10.7717/peerj-cs.55>
- Schmidt, T. (2015). Invisible stimuli, implicit thresholds: Why invisibility judgments cannot be interpreted in isolation. *Advances in Cognitive Psychology*, 11(2), 31–41. <https://doi.org/10.5709/acp-0169-3>
- Schneider, M., Tzanou, A., Uran, C., & Vinck, M. (2023). Cell-type-specific propagation of visual flicker. *Cell Reports*, 42(5), 112492. <https://doi.org/10.1016/j.celrep.2023.112492>
- Seijdel, N., Marshall, T. R., & Drijvers, L. (2022). Rapid invisible frequency tagging (RIFT): A promising technique to study neural and cognitive processing using naturalistic paradigms. *Cerebral Cortex*, 33(5), 1626–1629. <https://doi.org/10.1093/cercor/bhac160>
- Seijdel, N., Schoffelen, J.-M., Hagoort, P., & Drijvers, L. (2024). Attention drives visual processing and audiovisual integration during multimodal communication. *Journal of Neuroscience*, 44(10), e0870232023. <https://doi.org/10.1523/JNEUROSCI.0870-23.2023>
- Sonkusare, S., Breakspear, M., & Guo, C. (2019). Naturalistic stimuli in neuroscience: Critically acclaimed. *Trends in Cognitive Sciences*, 23(8), 699–714. <https://doi.org/10.1016/j.tics.2019.05.004>
- Spaak, E., Lange, F. P. de, & Jensen, O. (2014). Local entrainment of alpha oscillations by visual stimuli causes cyclic modulation of perception. *The Journal of Neuroscience*, 34(10), 3536–3544. <https://doi.org/10.1523/JNEUROSCI.4385-13.2014>
- Stolk, A., Todorovic, A., Schoffelen, J.-M., & Oostenveld, R. (2013). Online and offline tools for head movement compensation in MEG. *NeuroImage*, 68, 39–48. <https://doi.org/10.1016/j.neuroimage.2012.11.047>
- Vadillo, M. A., Malejka, S., Lee, D. Y. H., Dienes, Z., & Shanks, D. R. (2022). Raising awareness about measurement error in research on unconscious mental processes. *Psychonomic Bulletin & Review*, 29(1), 21–43. <https://doi.org/10.3758/s13423-021-01923-y>
- Vialatte, F.-B., Maurice, M., Dauwels, J., & Cichocki, A. (2010). Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives. *Progress in Neurobiology*, 90(4), 418–438. <https://doi.org/10.1016/j.pneurobio.2009.11.005>
- Yarkoni, T., & Westfall, J. (2016). Bambi: A simple interface for fitting Bayesian mixed effects models. *Journal of Statistical Software*, 103(15), 1–29. <https://doi.org/10.31219/osf.io/rv7sn>
- Zhigalov, A., Herring, J. D., Herpers, J., Bergmann, T. O., & Jensen, O. (2019). Probing cortical excitability using rapid frequency tagging. *NeuroImage*, 195, 59–66. <https://doi.org/10.1016/j.neuroimage.2019.03.056>
- Zhigalov, A., & Jensen, O. (2020). Alpha oscillations do not implement gain control in early visual cortex but rather gating in parieto-occipital regions. *Human Brain Mapping*, 41(18), 5176–5186. <https://doi.org/10.1002/hbm.25183>