A Collaborative Guide to Rapid Invisible Frequency Tagging (RIFT): Methods, Insights, and Recommendations

Kabir Arora[1,*], Cecília Hustá[2,*], Floortje Bouwkamp[3], Noor Seijdel[2], Songyun Bai[3], Qiu Han[3], J. Leon Kenemans[1], Stefan Van der Stigchel[1], Surya Gayet[1], Eelke Spaak[3,†], Samson Chota[1,†], Linda Drijvers[2,3,†]

[1] *Helmholtz Institute, Utrecht University, Heidelberglaan 1, 3584CS, Utrecht, The Netherlands.*

[2] *Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD, Nijmegen, The Netherlands.*

[3] *Donders Institute for Brain, Cognition and Behaviour, Radboud University, 6525 HR, Nijmegen, The Netherlands.*

Correspondence: k.arora@uu.nl

* These authors contributed equally
† These authors contributed equally

Abstract

Rapid Invisible Frequency Tagging (RIFT) is a recent advance in frequency tagging that exploits novel, high-frequency displays to modulate luminance at imperceptibly high frequencies. RIFT goes beyond low-frequency tagging by allowing researchers to track neural responses to rhythmic stimulation while avoiding perceptual confounds. RIFT is thus a promising method to address central questions in cognition, including attention, multimodal integration, and the neural mechanisms underlying oscillatory coordination in perception. However, setting up a RIFT study involves several technical and conceptual considerations. In an effort to make RIFT more accessible, we provide a comprehensive guide for implementing RIFT in cognitive neuroscience. On the basis of the joint experiences and data-driven insights of multiple labs, we provide practical recommendations derived from empirical datasets to improve reproducibility, covering hardware requirements, stimulus design, analysis approaches, and interpretation of results. We hope that this guide helps readers to both identify the conceptual areas where RIFT offers promising insights and navigate the technical caveats that come with the approach.

1 Introduction

Rhythms are omnipresent in our brain and environment, and shape how we attend to, perceive, and integrate information. In cognitive neuroscience, a fundamental challenge lies in the way we measure how rhythmic neural activity relates to ongoing cognitive processing without interfering with perception itself.

Frequency tagging has been a widely-used stimulus presentation method within the field of cognitive neuroscience (Norcia et al., 2015). In short, it involves varying a specific property of a stimulus (e.g., luminance) at a fixed frequency. This rhythmic modulation induces a neural response at the same frequency, called Steady-State Evoked Potentials (SSEPs). Because this response is frequency-specific, it provides a powerful marker of how the brain processes the tagged stimulus over time. Frequency tagging has therefore been widely used to study attention (e.g., spatial attention (Morgan et al., 1996; Müller & Hübner, 2002), feature-based attention (Müller et al., 2006; Pei et al., 2002), perceptual selection, and multimodal integration (Alp et al., 2018; Regan & Regan, 1988).

Recent advances in display technology have given rise to an innovative new branch of frequency tagging: Rapid Invisible Frequency Tagging (RIFT). Unlike low-frequency tagging, which typically flickers stimuli at frequencies below 30Hz, RIFT flickers stimuli at frequencies higher than 60Hz (Seijdel et al., 2023; Zhigalov et al., 2019), allowing the tagging to become virtually imperceptible **(Figure 1)**. RIFT therefore provides several advantages over existing SSEP applications. Firstly, at lower frequencies luminance changes are visible. This makes them easily identifiable, and they may automatically draw attention (Cass et al., 2011). RIFT, by flickering complex stimuli beyond the threshold of visibility, offers a tracker of visual processing without perceptual interference. This results in more naturalistic paradigms where the spatial spread of attention can be continually measured without any awareness of this probe. Another major benefit is that RIFT can measure the allocation of attention to locations which appear

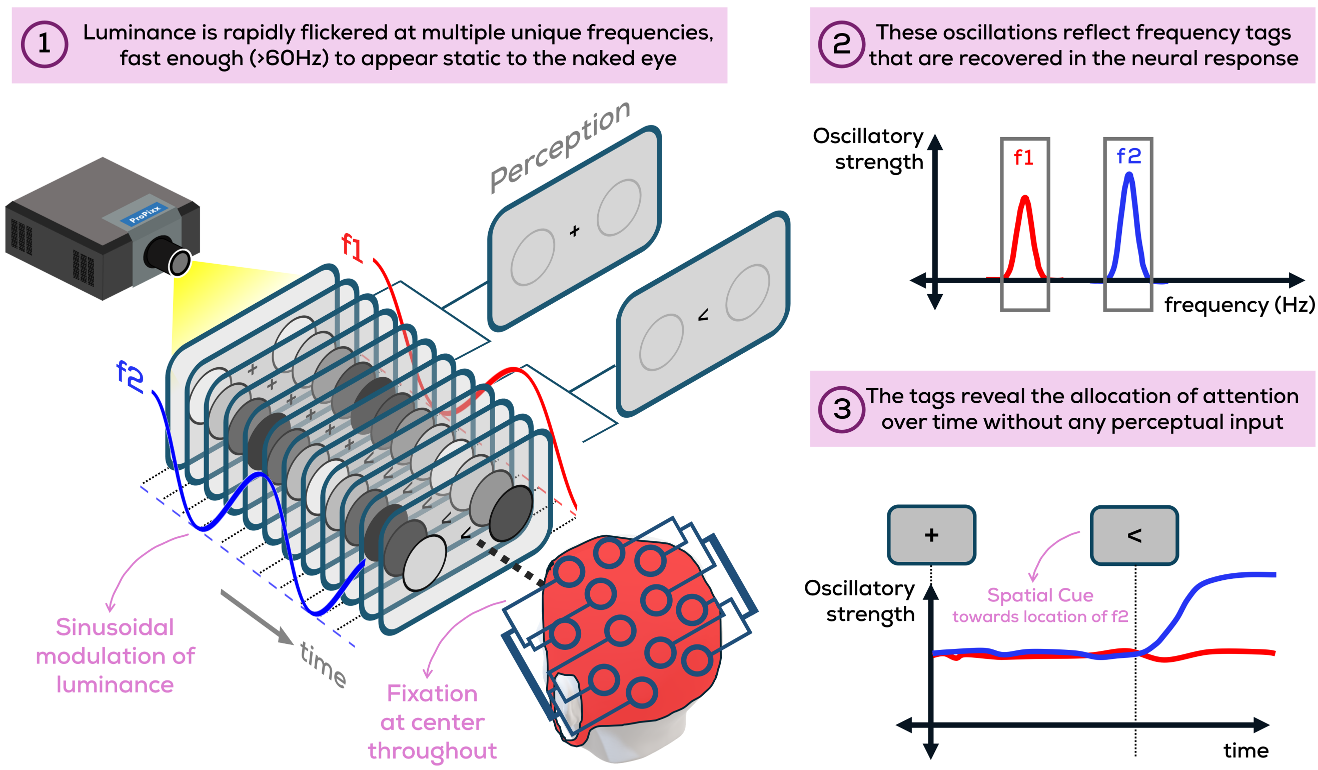

Figure 1: Rapid Invisible Frequency Tagging. The luminance of stimuli or areas on the screen is modulated sinusoidally at frequencies that exceed the threshold of perception. The response to these tags can then be uniquely recovered in the M/EEG response, and its amplitude over time can reveal location or feature based modulations in covert attention. Thus, RIFT forms a non-invasive, continuous tracker of attention in the absence of any visible probes.

indistinguishable from background space. In doing so, it offers a qualitatively different tool for studying relatively difficult-to-measure cognitive processes that occur in the absence of any stimuli. Lastly, in the low frequency range (<30Hz), SSEP responses may be difficult to disentangle from endogenous oscillations in similar frequency bands, or may even entrain or disrupt them (Notbohm et al., 2016; Spaak et al., 2014). The RIFT response does not interact with endogenous oscillations: its typical frequencies are far away from lower bands such as alpha (Arora et al., 2025; Gutteling et al., 2022; Zhigalov & Jensen, 2020), and even spectrally close oscillations such as gamma are not entrained (Duecker et al., 2021).

RIFT is essentially a tracker of visual processing, used most frequently as an invisible tracker of covert attention. Given the large range of cognitive phenomena that involve covert attention, RIFT has the potential to produce novel insights for a variety of cognitive fields, For example, RIFT has shown to be a powerful tool to capture spatial shifts of attention both to tagged stimuli (Zhigalov et al., 2019), as well as to tagged, but perceptually indistinguishable regions of visual space (i.e., tagging the background; Arora et al., 2025). Over the last few years RIFT has been applied within the domains of reading (Pan et al., 2021, 2024), distractor suppression (Ferrante et al., 2023), visual search (Bouwkamp et al., 2025; Duecker et al., 2025), visual working memory (Arora et al., 2025), brain-computer interfacing (Brickwedde et al., 2022), multimodal integration (Drijvers et al., 2021; Seijdel et al., 2024) and the interaction between speech planning and comprehension representations (Hustá et al., 2025). Aside from exploring novel cognitive contexts that may benefit from using the technique, current RIFT research is also focused on technical aspects such as optimizing display features (Minarik et al., 2023) and exploring alternate forms of tagging (Spaak et al., 2024).

In sum, RIFT has proven to be a sensitive method for tracking neural responses and their modulation by cognitive demands, and previous work shows clear and promising potential for RIFT within cognitive research. However, the method also introduces novel technical and analytical considerations that are not yet widely documented.

In this perspective we consolidate the joint experiences and data-driven insights of multiple independent labs to provide a comprehensive guide with resources and recommendations for running a RIFT study. We describe best practices for 1) hardware requirements, 2) experimental design and stimulus presentation, and 3) analysis of the neural response. This manual primarily compiles previously undocumented knowledge acquired through the setup and operation of new RIFT laboratories, as well as through in-depth exploration of technical aspects of previously collected RIFT datasets. Our goal is to make RIFT more accessible to researchers across various fields, and to highlight how RIFT can contribute to the study of rhythmic cognition across domains.

2 Hardware Considerations

2.1 Why does RIFT require devices with fast refresh rates?

An important feature of RIFT is that visual stimuli are tagged above the threshold of perceptibility (>60 Hz). Achieving this requires specialized display tools with high-speed refresh rates that go beyond the refresh rates available in standard monitors.

Conventional monitors, typically running at up to 120 Hz, are ill-suited for several reasons. Firstly, a 120Hz monitor can only produce a 60Hz RIFT tag, by alternating between black and white on each frame However, given that RIFT is mainly used as a tracker of spatial attention, most cognitive questions using RIFT depend on contrasting the attentional resources dedicated to different locations in the same visual environment. Thus, RIFT studies often use two or more high-frequency tags simultaneously, for example 60Hz and 64Hz (Arora et al., 2025), or 60Hz, 64Hz, and 68Hz (Bouwkamp et al., 2025), which a 120Hz monitor is not capable of doing. Secondly, standard monitors are not normally designed to be able to shift through their full luminance range (black - minimum, to white - maximum) during the short interval between two consecutive frames. Our own personal experiments demonstrated that standard monitors have trouble doing so, leading to visible tagging modulations on the screen. Though the tagging could also use a lower luminance range (e.g., flicker from grey to black) if a monitor is incapable of using its full luminance range quickly, this has been shown to weaken the tagging response (Spaak et al., 2024). Lastly, even if perfect control of luminance at 120Hz is achieved, there is a stark difference between the resulting step-function luminance trace achievable at a 120Hz refresh rate and the sine-approaching traces achievable at higher refresh rates. The latter avoid sudden and large jumps in luminance **(Figure 2)**, resulting in much reduced visibility of the luminance oscillation. Indeed, it has been shown that, when using 120Hz refresh rates, even 60Hz oscillations are visible (Waldin et al., 2017). However, with rapid refresh rates (i.e., 1440Hz) the 60Hz tag has been experimentally shown to be invisible (Spaak et al., 2024).

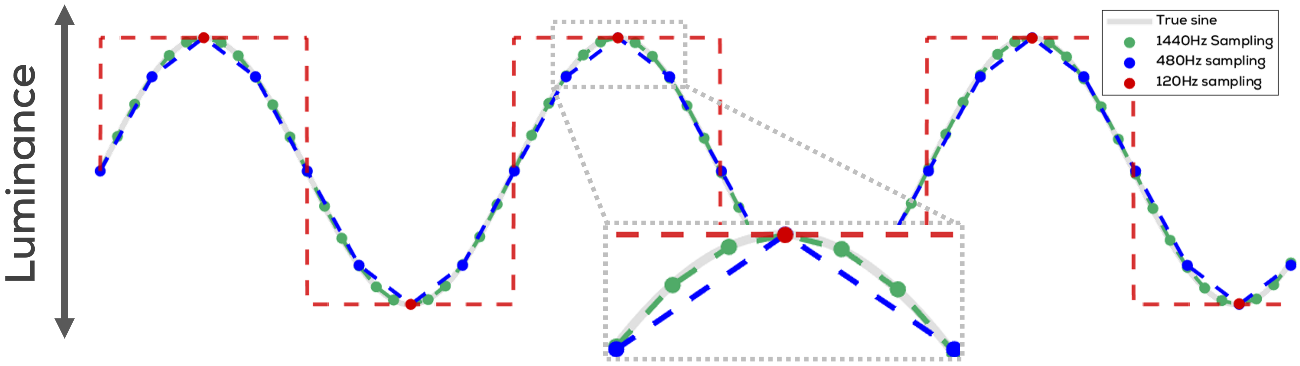

Figure 2: Higher refresh rates more closely reflect sinusoidal modulations of luminance.

2.2 Which devices can accomplish this?

Nearly all RIFT research so far has made use of a PROPixx DLP projector (VPixx technologies), which supports refresh rates up to 1440 Hz (if using greyscale) or 480Hz (if colour is used). PROPixx circumvents the bandwidth limitations of graphics cards by repackaging multiple frames into each input image (see Box 1), allowing for imperceptible rhythmic stimulation at a precision that is unmatched by standard hardware. However, given the expenses associated with this option, it is worth exploring viable alternatives.

More recently, newer commercially available OLED monitors operating at 480 Hz have emerged as a viable alternative, by using fast pixel response times, which increases accessibility (Dimigen et al., 2025). Early evidence exists that demonstrates that these consumer-grade monitors produce robust RIFT responses. We therefore believe that as the use of RIFT becomes more widespread, this monitor option will soon be more utilised. Thus, our discussions and suggestions below are general to any device and refresh rate (with the exception of **Box 1**).

Box 1. How does the PROPixx projector achieve higher-than-GPU refresh rates?

he core limitation overcome by the ProPixx projector is not merely high-speed display, since LED displays can also display rapidly changing visual luminance. In order to display complex stimuli, the display tool must be paired with a graphics card that can produce frame-by-frame images quickly enough to keep up with a fast display rate. The ProPixx takes away this requirement. That is, the graphics card being used can simply operate as if it was connected to a simple 120Hz refresh rate device, but the projector is able to convert this input to a much faster refresh rate (480Hz if using colour, 1440Hz if not using colour).

The ProPixx achieves this by allowing the user to "pre-package" up to 12 frames worth of information within a single image. For a normal monitor/PC setup, the graphics card sends out a frame which the monitor then displays. Thus, for a 120Hz monitor setup, 120 images are displayed per second **(Figure 3A)**. The projector is able to achieve a higher refresh rate with the same input speed (120 images per second) from the graphics card. First, it divides the image received from the graphics card into four equal quadrants. Each of these quadrants is then sequentially displayed **(Figure 3B)**. This quadruples how many images the projector can display per frame sent by the graphics card (each with half the pixel resolution), resulting in a 480Hz refresh rate. This means that any image which is intended to be drawn at one point on the screen, must actually be drawn at four different locations (the centers of each of the four "quadrants") which are offset in different directions from the true center of the image. Here we include code (`quadifier.m`) compatible with MATLAB-Psychtoolbox that carries out this operation by converting coordinate information from a screen-centered reference to a quadrant-centered reference.

This refresh rate can then further be tripled to achieve a 1440Hz refresh rate for grayscale images. For this, the three color channels (Red, Green, and Blue) that normally add up to produce one color image can be used to transmit three unique images **(Figure 3C)** which are then displayed in sequence. In this way, the projector can display three frames per quadrant, resulting in twelve total images per frame sent by the graphics card **(Figure 3D)**.

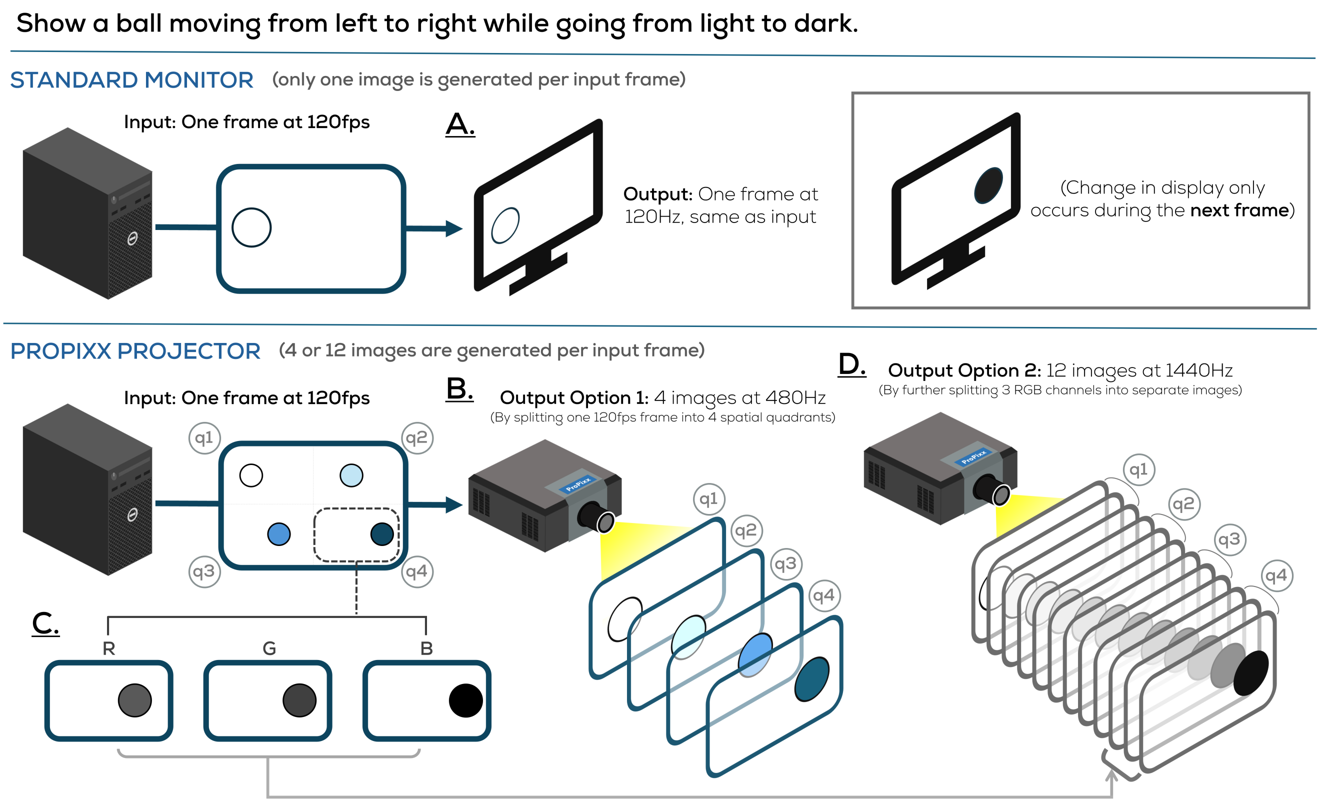

Figure 3: Operation of the PROPixx projector. When connected to a 120fps PC input, **A.** a standard monitor displays one image for each frame. In the same time, the PROPixx projector can display either **B.** 4 images by splitting up the frame into quadrants and displaying them sequentially, producing a 480Hz output, or, **C.** 12 images by further splitting the R, G, and B colour channels of the frame to produce three images per each of the four quadrants, **D.** resulting in a 1440Hz framerate. Note that the input from the PC has to be encoded specifically to produce the desired output after quadrant-splitting and RGB-splitting.

2.3 Components of a RIFT Setup

The main hardware component of a RIFT setup is the high-refresh display device (see *Section 2.2*). Because of the low demands the PROpixx projector places on the graphics card itself (Box 1), the projector setup has relatively lenient graphics card requirements. Our setups utilize relatively inexpensive graphics cards (e.g., NVIDIA Quadro K620 2GB, GeForce GTX960 2GB, and Radeon RX570 8GB). The GPU requirement is considerably higher when using a high refresh-rate monitor for RIFT instead of the PROpixx projector (Dimigen & Stein, 2024; Dimigen et al., 2025).

Irrespective of which device is used for RIFT, a photodiode (luminance sensor) is essential for identifying and verifying that the intended modulation is faithfully displayed. Even timing errors on a millisecond-level can shift the phase of the frequency-tagged signal, undermining experimental precision. For example, with a 60Hz tag (1 cycle = $\sim$16.6ms), an 8ms delay would put the tagging signal in anti-phase, so even a 1ms delay will have a notable impact, and thus accounting for these inaccuracies becomes critical.

3 Stimuli and Experimental Design Considerations

3.1 Which frequencies are best to use for tagging?

3.1.1 Theoretical Range

With RIFT, we measure neural responses to imperceptible periodic stimulation. This definition leads to two automatic (theoretical) thresholds for RIFT frequencies. The first is a maximum frequency above which the response to periodic stimulation is no longer observed in the M/EEG signal. The second is a minimum frequency below which the flicker is visible. Previous work has used LED-based displays to investigate the periodic neural response to flickers in the 1-100Hz range (Gulbinaite et al., 2019; Herrmann, 2001), from which a rough threshold of around 80Hz emerges as an upper limit (Minarik et al., 2023). On the other hand, mean estimates of the critical flicker-fusion threshold tend to be close to 50Hz but vary between 30-60Hz depending on numerous individual factors and cognitive states (Haarlem et al., 2024). Thus, for visual tagging,

Study	Topic	Tagged Stimuli	Refresh Rate (Hz)	Tag Type & Freq. (Hz)	Tag Duration (ms)	Stimuli (# / Size / Loc.)
Zhigalov et al., 2019	Attentional effect on cortical excitability	Photos of faces and houses	1440	63, 78	1500	2 / – / 8.3° eccentricity, below midline
Zhigalov et. al., 2020	Alpha oscillations and spatial attention	Photos of faces and houses	1440	60–70 (broad-band tag)	2000	2 / 5.7° / 3.8° eccentricity, below midline
Drijvers et al., 2021	Multimodal integration	Auditory verbs, gesture video	1440	A: 61, V: 68, IM: 7	∼2000 avg	2 / 10°×6.5° / Midline
Duecker et al., 2021	Entrainment of gamma	Background, circular moving gratings	1440	52–90 (2Hz steps)	2000	1 / 2.62° / Center
Gutteling et. al., 2021	Attentional gating via alpha-band oscillations	Clear or degraded photos of faces	1440	63, 70	2350 - 3350	2 / 8° / 7° eccentricity, below midline
Pan et al., 2021	Parafoveal processing	Background under the target word	1440	60	1000	1 / ∼2–3°×1° / Midline
Brickwedde et al., 2022	BCI control via covert attention	Section of background with grainy texture	1440	V: 56, 60	2000 (train); continuous segmented to 1000 (test)	2 / – / Bilateral below midline
Ferrante et al., 2023	Statistical learning of distractor location	Gabor patches	480	55–75 (broad-band tag)	1500 (filler) + 300 (search)	2 / 6°×6° / 4° eccentricity in all quadrants
Minarik et al., 2023	Optimal tag parameters examining: 1) tagging freq., 2) stim. size, 3) stim. position	Background section	1440	66 - freq. varied for 1) between 60–100 (4Hz steps)	1400 + tapers	1 / 10° / Center - size varied for 2) 2–12° in 1° steps - pos. varied for 3) midline, above, or below
Bouwkamp et al., 2025	Predictive visual search	T + L distractors	1440	60, 64, 68	≤2500	3 / 2.5°×2.5° / variable (within 11°×7.5°)
Pan et. al., 2024	Parafoveal semantic processing in reading	Background under the target word	1440	60	1000	1 / ∼2–3°×1° / Midline
Seijdel et al., 2024	Multimodal integration; attention	Auditory verbs, gestures	1440	A: 58, V-att: 65, V-unatt: 63, IM: 5, 7	∼2000	3 / – / Midline
Spaak et al., 2024	Optimal tag parameters examining: 1) imperceptibility, 2) type of tagging, 3) accounting for phase shifts	Circular gratings	1440	60, 66	1200 or 1500	1–2 / 2°–4° / Midline, above, or below
Arora et al., 2025	Internal vs. External Attention	Color gratings, Background	480	60, 64	Continuous	2 / 6° / 6° eccentricity, 2° below midline
Husta et al., 2025	Planning speech during comprehension	Auditory nouns, pictures	1440	A: 54, V: 68, IM: 14	A: ∼877, V: 1000	2 / 7.1°×8.4° / Center
Duecker et al., 2025	Guided visual search	T + L distractors	480	60, 67	≤4000	17 or 33 / 1°×1° / variable (within 10°×10°)
Dietz et al., 2025	Temporal expectations	Square-wave gratings	480	60	Whole trial (≤3800)	1 / 6° / Center
Wang et al., 2025	Attentional capture	Geometrical shapes (singleton targets + distractors)	480	60, 64	1300	6 / 4.3° × 4.3° / 6° eccentricity, in circular shape
Dimigen et al., 2025	RIFT on consumer monitor	disc stimulus matching background	480 (on OLED monitor)	60, 64	10000	1 / 2.55° + tapers / Center or Periphery at 12°
Duecker et al., 2025	Role of alpha in visual search	T + L distractors	480	60, 67	Variable	17 or 33 / – / –

Table 1: Overview of topics and tagging parameters in existing RIFT research. Notes: A = Auditory tagging; V = Visual tagging; IM = Intermodulation frequency. Where not specified, "–" indicates data not reported or not applicable. Tag location and stimulus size are reported in degrees of visual angle (dva).

the flicker appears visible to most people below 50Hz. This allows us to set a (conservative) theoretical RIFT tagging range of 60-80Hz in the visual domain. RIFT has also been implemented slightly below this range at 56Hz (Brickwedde et al., 2022), and similar tagging protocols have been applied at even lower frequencies (41-45Hz; (Marshall et al., 2024)), though here invisibility was not verified.

Beyond this range, there are a number of technical aspects to consider that produce a more limited set of frequency options in practice.

3.1.2 Higher frequencies evoke lower responses

It is worth noting that even though all frequencies within the theoretically plausible range (~60Hz-80Hz) have been shown to evoke a periodic response, this response decreases in strength as the frequency increases (Gulbinaite et al., 2019; Herrmann, 2001). An upper limit of 72Hz has been suggested for observing tagging in 90% of the participant pool, above which this percentage decreases quickly (Minarik et al., 2023). A summary of the stimuli and frequencies utilized in previous RIFT studies is summarised in **Table 1**.

3.1.3 Sampling

When sinusoidally modulating luminance to display a tagged stimulus, the choice of tagging frequency and phase can lead to the presentation of imperfect sine waves, which can in turn result in dynamic luminance range being underutilized by the tagging. A 'perfect' tag, i.e., one that only stimulates the intended frequency without introducing any harmonics or other low-frequency components, would always capture the same values on the sinusoid for every cycle including the high and low peaks (e.g., **Figure 4A**). This, however, is only possible when the tagging frequency is a factor of the refresh rate, in this example a 60Hz (480Hz / 8) tag. If the refresh rate is not an integer multiple of the tagging frequency, an artifact will arise because the peaks and troughs of the true sinusoid are not sampled regularly (e.g., **Figure 4B**), in this example a 61Hz (480Hz / ~7.86) tag. Similarly, if the sinusoid being sampled has a phase offset such that no frames sample its peak or troughs, even a 'perfect' tagging frequency can miss out on the full dynamic range of luminance (e.g., **Figure 4C**). Such sampling issues can be avoided entirely by simply using frequencies that are an integer factor of the refresh rate (480Hz/8 = 60Hz, 480Hz/7 = 68.57Hz, 480Hz/6 = 80Hz) at the appropriate phase offset to ensure that the

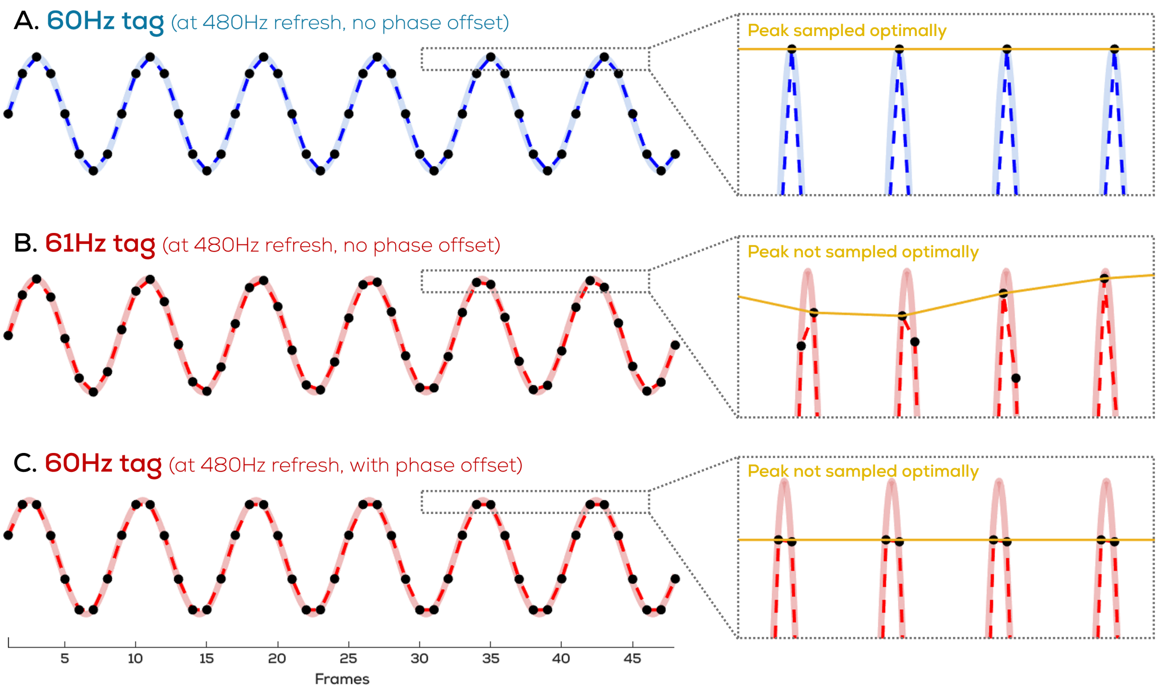

Figure 4: The choice of tagging frequency and phase can lead to presentation of imperfect sine waves. A. When drawing a 60Hz tag at a 480Hz refresh rate, 8 (480/60) luminance points are equally sampled on each cycle and perfect sampling is observed if the phase is aligned to the sinusoid's peaks. **B.** When drawing a 61Hz tag at the same refresh rate, every cycle does not have the same sampling of luminance points, and **C.** When drawing a 60Hz tag at the same refresh rate but a different phase offset, the peaks are not sampled uniformly thus under-utilizing the dynamic range of the tagging.

peaks of the sinusoid are sampled. This factor then equals how many points are sampled on each cycle, for example 8 points per cycle for a 60Hz tag. Though experimental constraints may result in cases where it is preferable to sacrifice a small amount of dynamic range in order to use a larger number of unique frequency tags, this route offers a starting point for frequencies to select in the absence of any other limitations. These concerns are less pressing in cases of very high refresh rates, such as the 1440 Hz capabilities of a PROpixx projector.

3.2 What needs to be considered when tagging multiple stimuli?

RIFT studies commonly implement multiple tags to track visual processing of multiple stimuli or locations simultaneously. This involves some extra considerations.

3.2.1 Spectral overlap

Any time series analysis involves a trade-off between frequency resolution and temporal resolution: improving precision for one reduces precision for the other.

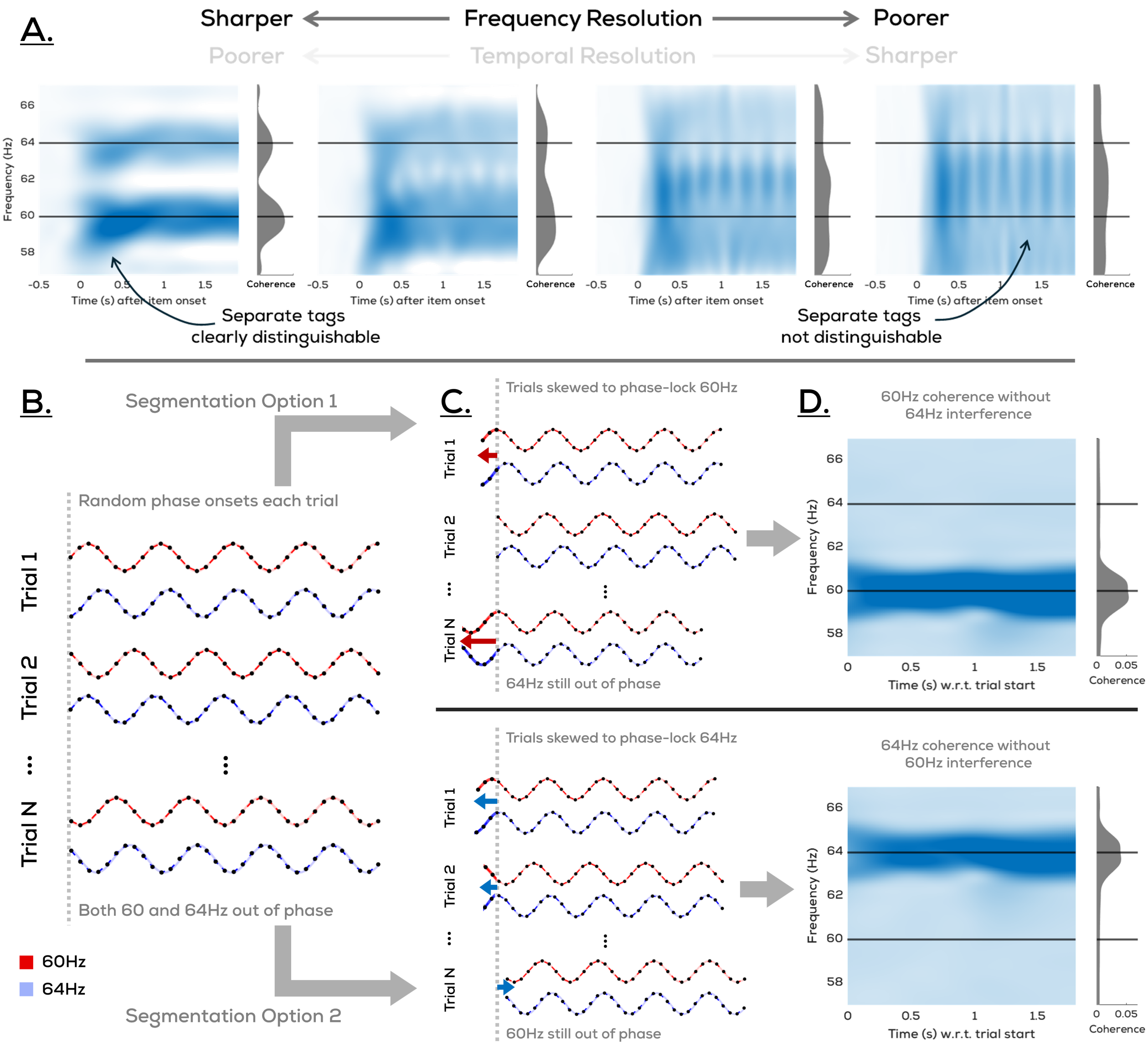

Figure 5: A. When using multiple tags, temporal resolution is limited due to overlap effects. To maintain high temporal resolution, frequency resolution must be sacrificed. However, when using multiple tags, this is not possible beyond a limit due to overlap between the responses from both tags. **B-D. Phase can be used to avoid overlap effects across multiple tags. B.** To avoid these overlap effects, trials can be presented with tagging envelopes initialized at random phases. **C.** Trials can then be separately analyzed by skewing to align the phases of one tag at a time. **D.** This results in independent isolation of each tagging signal without the contribution of the other using coherence. Data averaged across 23 participants from Dataset 2.

For studies looking at RIFT modulations over time, it is ideal to have high temporal resolution (and thus low frequency resolution). However, when tagging multiple stimuli at unique frequencies, frequency resolution cannot be too low because both tags must remain distinguishable. To illustrate this, we show how the mean discriminability of two frequency tags varies with parametrically modulated frequency resolution **(Figure 5A)**. Here, the periodic response is measured using filter-HT derived coherence (a technique to extract the amplitude of the oscillatory signal, explained further in *Section 4.1*).

It is possible to leverage tagging phase to increase the separability of different tags. Instead of maintaining a constant phase onset across trials that is also identical among different stimuli, the tagging signals can instead be offset by a random amount for each stimulus and each trial. Then, different tagging signals can later be separately analyzed from the same data without interference from one another **(Figure 5A)** by epoching the data separately such that the epochs are phase-locked with respect to only one particular signal at a time **(Figure 5B)**. All other tags will still remain scrambled in phase, and a measure such as coherence which relies on phase consistency will thus discount their contributions **(Figure 5C)**. It is worth noting that using this procedure prevents the use of phase as a tagging information channel (see *Section 3.5.1* for more about phase tagging).

3.2.2 Counterbalancing

Counterbalancing tagging frequencies across participants – or, if possible, across trials – is important to avoid confounding frequency-specific effects with experimental conditions. Since higher frequencies elicit weaker neural responses, unique frequency tags are treated as different “channels” of information. To make meaningful comparisons across conditions, frequency-condition assignments can be balanced across participants, or better yet, within participants. Here, ‘condition’ can refer to many feature dimensions such as stimulus type, location, shape, etc. Trial-level counterbalancing is especially powerful, as it controls for individual differences and increases sensitivity to condition-specific effects. However, it does reduce the available number of trials in each condition for measures such as coherence (see *Section 4.1*) which operate on a set of trials.

3.3 What needs to be considered when tagging grayscale vs. coloured stimuli?

When tagging a stimulus, its intensity is modulated sinusoidally; practically this involves multiplying some “luminance feature” of a stimulus with a time-varying sinusoidal envelope. Luminance can be tagged in various ways. For example, when tagging black and white gratings, it has been shown that modulating the white bands across the full luminance range (0-100%), or using contrast tagging (i.e., anti-phase flicker to white and black bands), produce a stronger response than other alternatives (Spaak et al., 2024). Furthermore, we have seen that applying a white luminance mask over the stimuli yields robust tagging responses, whereas modulating stimulus opacity to fade stimuli in and out against a grey or textured background results in markedly weaker tagging signals (unpublished data).

Most design tools for psychophysics experiments directly use RGB colour codes when drawing stimuli. If a grayscale object is tagged, it is most straightforward to use this RGB colour code as the luminance feature that is sinusoidally modulated over time. However, when tagging coloured objects, it is better to modulate a feature that explicitly represents luminance to ensure that modulation targets luminance directly without unintentionally altering color balance. This can easily be achieved by converting the RGB colour code to a perceptually uniform colour space, such as CIELAB, which includes a luminance dimension. This luminance component can then be multiplied with a sinusoidal envelope to produce the time-varying tagging amplitude. Finally, these scaled values can be converted back to RGB for drawing the stimulus.

Lastly, it should be noted that tagging colour affects the available dynamic tagging range of luminance. Tagging from 0% luminance (black) to 100% luminance (white) produces a grey tag. Since the luminance of any other colour falls somewhere within this range, tagging other colours reduces the fraction of the luminance range that is sinusoidally modulated. The more perceptually distinct colours (that are equivalently tagged) are required, the lower this possible range gets. Thus, tagging several colours equally comes at some cost to the luminance range, which then translates into a cost to the tagging amplitude in the neural response (Spaak et al., 2024).

3.4 Best practices to avoid visibility of the tagging

In addition to using a high tagging frequency (>60Hz), there are other parameters that heavily influence how “invisible” the tagging is. Consider a tagged patch of background that is beyond an individual’s critical flicker-fusion threshold, in that the luminance oscillation itself is invisible. Since this stimulus is presented against a static background, very abrupt shifts of luminance are produced at the border of the tagged region when the tagging cycle is at extreme values (grey background - white tagging area, grey background - black tagging area). Over time, this is not a concern since the luminance changes are too fast to be perceived. However, during eye movements, the invisibility of the tag is compromised. A faint boundary becomes momentarily visible around the tagged region, likely due to saccadic suppression (Krekelberg, 2010). One simple option to reduce this boundary detection is to place a visible outline around the tagged area, since this tagging boundary is then masked by an actual physical boundary. However, if the goal is to invisibly tag an area that is perceptually indistinguishable from a static background, then this is not an option. In that case, a transparency mask (Minarik et al., 2023) can be applied to the edges of the tagged region to produce a smooth edge that is not detected as a clear boundary between tagged and untagged regions. Similarly, to reduce visibility resulting from a sudden tagging onset or offset, the tagging may be ramped-up or ramped-down at the ends of the tagging interval (for e.g., 200ms ramp-up and ramp-down periods as used by (Minarik et al., 2023); but we have also used periods as low as 50ms). Regardless of such precautions, it is ideal to empirically confirm the luminance modulation being imperceptible. This can either be done beforehand by means of a pilot experiment, or alternatively by informing the participants in advance that they may see certain flickers or glitches during the experiment and conducting an appropriate questionnaire afterwards to see if they perceived any such effects (Arora et al., 2025; Drijvers et al., 2021; Pan et al., 2021; Seijdel et al., 2024).

3.5 Alternative forms of tagging

In this perspective, our suggestions and discussions of RIFT focus on periodic stimulation at specific frequencies, using frequency alone as a channel through which separate tags can be achieved. This approach is the standard taken in almost all existing RIFT work (see Table 1). However, there are also alternative tagging approaches, varying in how extensively they have already been explored in combination with RIFT.

3.5.1 Phase tagging

One alternative involves a different feature of oscillations: phase. Instead of creating only one channel of information for a given tagging frequency, by modulating phase it is possible to achieve multiple channels of information (multiple tags) per tagging frequency (Spaak et al., 2024). This is achieved by tagging separate stimuli or locations at the same frequency, but at some phase offset that is maintained consistently over time and/or trials. This consistency allows the neural responses at that frequency to later be disentangled and attributed to the uniquely phase-locked tags. Though unique frequency tags have separate response profiles and thus cannot be compared to each other and must be counterbalanced (see *Section 3.2.2*), two phase tags at the same frequency can be directly compared. Given that the frequency range within which RIFT is possible is rather small (See *Section 3.1*), this approach also offers a promising avenue for maximizing the number of RIFT tags that can be used simultaneously.

3.5.2 Broadband and noise tagging

In broadband tagging, as opposed to stimulating at one specific frequency, the stimulation used contains periodic components at a range of frequencies (for e.g., 55-75Hz broadband tagging as used in (Ferrante et al., 2023)). A similar alternative comes from the Brain-Computer Interface (BCI) literature in the form of Code Modulated Visual Evoked Potentials (c-VEPs), also referred to as noise tagging (Martínez-Cagigal et al., 2021). The central idea here is the modulation of luminance using a fully random or pseudorandom temporal sequence instead of a sinusoidal sequence or a sequence limited to a small frequency range. The strength of the neural response to this signal can then be recovered by correlating the M/EEG signal over time with this luminance sequence. c-VEPs, like SSVEPs, form an already established field of visual stimulation. However, the advent of RIFT offers a promising future direction: using devices with high-frequency refresh rates allows for these random luminance sequences to be high-pass filtered at 60Hz prior to display, resulting in random sequences that are compatible with c-VEP but also invisible.

3.6 Intermodulation frequencies and auditory tagging

As was briefly outlined in the introduction, RIFT can also be utilized to examine the interaction between two signals, such as audio-visual inputs, by examining intermodulation (IM) frequencies (Drijvers et al., 2021; Hustá et al., 2025; Seijdel et al., 2024). The IM frequency results from the nonlinear interaction of the base audio and visual tagging frequencies and peaks at the difference and sum of the two signals interacting ($f2 \pm f1$; e.g., (Regan et al., 1995)). The power at the IM frequency is thought to reflect the strength of nonlinear interaction between the representations of the two tagged stimuli. When tagging with frequencies above 60Hz, the IM component at f2+f1 exceeds 100Hz and is virtually undetectable in the M/EEG signal. Consequently, previous studies have focused on analyzing the IM frequency at f2−f1 (Drijvers et al., 2021; Hustá et al., 2025; Seijdel et al., 2024), which is more reliably measurable, but falls within the range of endogenous oscillations. This leads to additional considerations for analyzing IM frequencies. It is best when the oscillatory results for a paradigm are known (such as the established behaviour of alpha oscillations in visual attention paradigms), so the overlap between the endogenous effects and the (chosen) IM frequency can be avoided. This further constrains the selection of the main tagging frequencies. Considering that with certain paradigms, oscillatory effects are difficult to avoid, using a condition without tagging (i.e., tagging baseline) is always advised (see *Section 4.2*). Additionally, unpublished data shows that the negative relationship between frequency and tagging amplitude (higher frequencies generate lower responses) is not necessarily true for lower IM responses, which allows for selection of a wider range of tagging frequencies, especially when the main research question focuses on the IM frequency.

When tagging auditory stimuli to study the interaction between auditory and visually tagged stimuli, there are additional considerations beyond those of visual tagging alone. In particular, the duration of the auditory stimulus is critical: short audio segments may not provide sufficient time for reliable frequency tagging. As a result, auditory tagging has often been restricted to longer words (>700 ms). Moreover, the tagging should be applied offline to the audio stimuli before the experiment, so that tagged files can be presented directly during data acquisition and potential timing inaccuracies are avoided. Finally, the perceptual detectability of the modulation should be evaluated. For instance, Drijvers et al. (2021) showed in a pretest that amplitude modulation at 61 Hz did not impair speech intelligibility in clear speech conditions, indicating that high-frequency tagging can be applied without compromising stimulus clarity. We briefly note these caveats here for completeness, but a detailed discussion of auditory tagging lies outside the scope of this paper.

4 Analysis Considerations

4.1 Common Analysis Methods

This section outlines common techniques for quantifying the RIFT response. First, we describe two ways of computing the spectral content of the M/EEG signal on each trial. These include **power** as measured by the traditional **Fourier Transform**, as well as the **filter-Hilbert Transform** (filter-HT) approach. Then, we describe how phase-consistency across trials is commonly leveraged by computing (inter-trial) **coherence**. We subsequently provide an overview of the RIFT response as measured through each of these techniques **(Figure 6)**.

All techniques listed here are used to measure the magnitude of neural activity at specific frequencies induced by RIFT and to assess the stability and intensity of these responses across trials, conditions, or participants. They all operate on data that has been segmented into epochs after or during preprocessing, often around the onset of the tagging. Which technique to use for estimation of spectral coefficients (FFT versus filter-HT), and for subsequently quantifying the RIFT response (power, amplitude, or (inter-trial) coherence) depends on the exact nature of the experimental design. In general, coherence-like measures provide higher signal-to-noise ratio, so are preferred when aggregating across trials is not a problem. When single-trial estimates are necessary, power or amplitude can be used.

4.1.1 Fourier Transform

Power at the tagging frequencies can be computed by converting the time-domain data into the frequency-domain using a Fast Fourier Transform (FFT). The time–frequency trade-off must be considered here: e.g., if aiming to achieve 1Hz frequency resolution, a window length of 1000 ms is required.

Conversion to the frequency domain for regular M/EEG analysis is usually accompanied by tapering of the signal to reduce spectral leakage, i.e., the spreading of energy from one frequency to others. Hanning or Hamming windows provide strong attenuation of frequencies far away from the frequency-of-interest (i.e., low side lobes), which is desirable in standard M/EEG analyses. However, based on our findings, boxcar tapers (i.e., essentially not tapering at all) are more effective for capturing tagging responses, due to the narrow main lobe of the boxcar taper.

If multiple frequency tags are present in the signal, ensure that unique tagging frequencies are spaced at least N Hz apart for a 1/N second analysis window (e.g., at least 5Hz apart for a 0.2 s analysis window, or at least 2Hz apart for a 0.5 s analysis window) to avoid spectral overlap (or, see alternative suggestion using phase randomization in *Section 3.2.1*).

4.1.2 Filter-Hilbert Transform

An alternative spectral transformation to the FFT is the filter-Hilbert Transform (filter-HT). The Hilbert Transform is only practically interpretable when applied on monocomponent signals, i.e., those produced by a single periodic source. Thus, when computing the HT, an M/EEG signal must first be bandpass filtered at the frequency of interest. For example, filtering a 60Hz tagging response between 58Hz and 62Hz. This can then be referred to as the filter-HT approach. Bandpass filtering requires the selection of a bandpass width parameter. Narrower filter widths reduce the contribution of neighbouring frequencies and

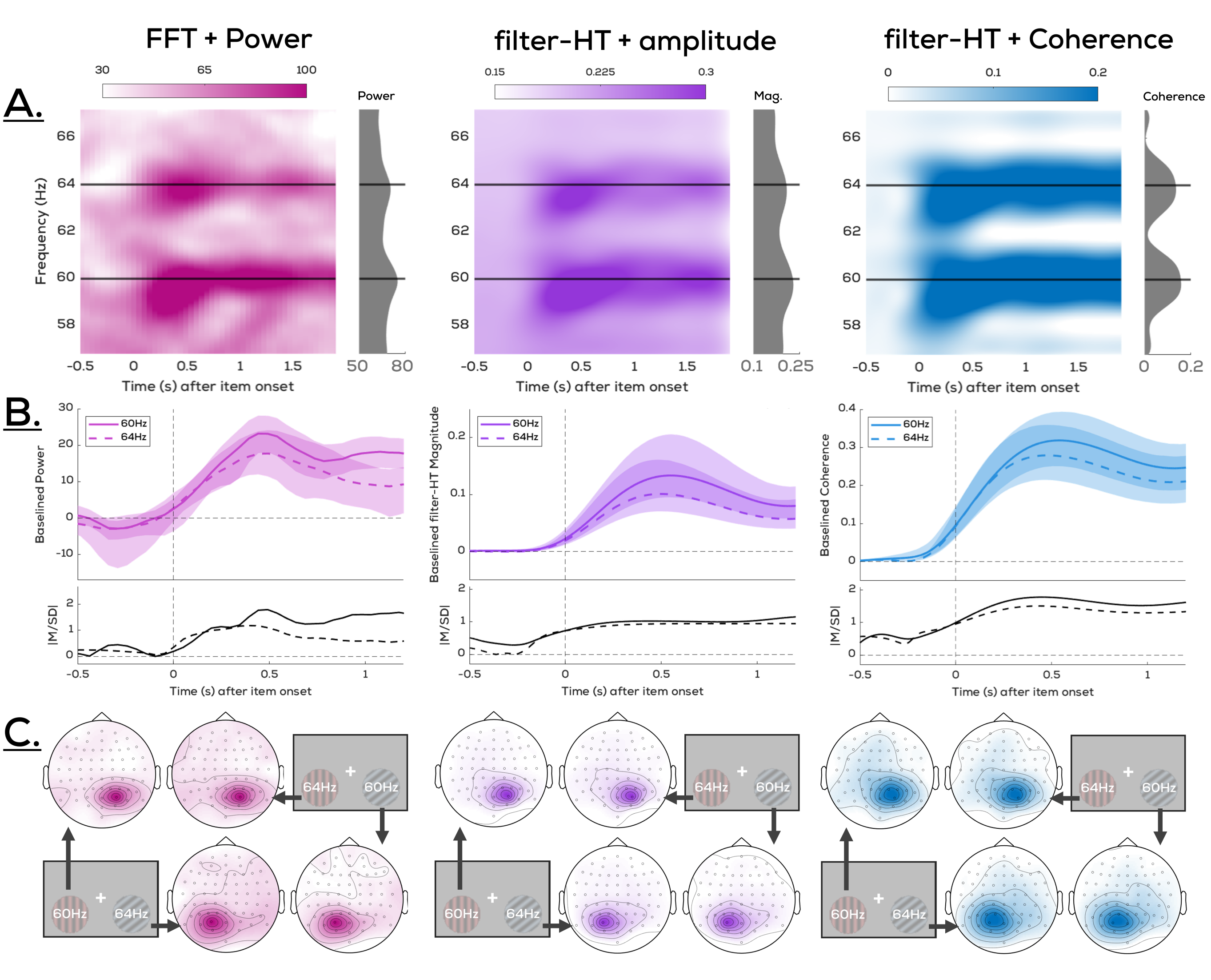

Figure 6: Overview of tagging response as measured by various techniques. A. Spectrograms, **B.** traces, and **C.** topoplots providing an overview of the RIFT response as measured by (left to right:) trial-averaged FFT-derived power, trial-averaged filter-HT amplitudes, and filter-HT derived coherence. Shaded regions in **B** represent 95% bootstrapped confidence intervals across participants. Data averaged across 24 participants from Dataset 1. For filter-HT based analysis (used here for filter-HT amplitudes and coherence) we used a bandpass width of 2Hz centered at panels A and C, and 3.8Hz for panel B. For FFT-based analysis (used here for Power), we computed FFTs at a resolution of 0.1Hz using sliding windows of 1 sec duration with 95% overlap.

noise but provide poorer temporal resolution. Ideally finite impulse response (FIR) filters are used (Widmann et al., 2015). See *Section 3.2.1* for an overview of how this parameter choice matters when multiple tags are used.

4.1.3 Coherence

Coherence measures the synchronization between M/EEG signals at specific frequencies and a reference signal, typically a pure sine wave at the tagging frequency or a photodiode recording. In addition to relying on oscillatory amplitude, coherence also relies on how phase-locked these oscillations are across a set of trials. To perform coherence analysis, a reference signal, such as a sine wave matching the tagging frequency, is generated to match the duration and sampling rate of the epochs. If the exact same reference signal (identical frequency and phase) is utilized across trials the resulting coherence is equivalent to intertrial coherence (ITC). Alternatively, some studies use the photodiode measurements from the on-screen tagging as a reference signal (Duecker et al., 2021; Pan et al., 2021), as this measure provides the ground truth even when encountering imprecisions in the tagging response (e.g., due to missed frames), which would be missed by the sine wave approach.

Equation 1 provides a measure of how consistently the brain's activity synchronizes with the tagged stimulus, resulting in a coherence estimate at time t.

$$coh(t) = \frac{|G_{xy}(t)|^2}{G_{xx}(t)\, G_{yy}(t)} \tag{1}$$

In Equation 1, G_{xy} is the cross-spectral density of the M/EEG signal and the reference signal, measuring how much the signals oscillate together. G_{xy} and G_{xy} are their auto-spectral densities, measuring the strength of oscillations in each signal individually. These spectral densities can be computed in different ways, for example through an FFT or through a filter-HT, as described above. High coherence implies a consistent response to the tagging frequency. Coherence is computed across time and/or trials. Previous RIFT literature describes this computation in more detail (Arora et al., 2025; Pan et al., 2021; Spaak et al., 2024).

Comparisons across experimental conditions can be made by computing coherence separately on groups of trials corresponding to specific conditions (rather than on individual trials), and contrasting the resulting coherence traces. Condition-wise coherence can in this manner be computed per channel, per tagging frequency, and per participant.

Coherence is particularly advantageous compared to power because it captures not only the amplitude of oscillatory activity but also the consistency of its phase alignment with the stimulus across trials (Figure 3 in Spaak et al., 2024). Consequently, coherence provides greater sensitivity for detecting reliable neural responses at the tagging frequency, even in the presence of amplitude variability or noise.

4.2 Baselines

For all experiments, we recommend a traditional baseline where no stimuli are presented, tagged or otherwise. This offers an additional SNR metric to demonstrate the strength of the tagged response by comparison to a no-tagging period. However, depending on the research question and its potential confounds, different experiments may require different additional baselines. In cases where cognitive (e.g., attentional) modulations of the RIFT response could be positive (e.g., enhancement) or negative (e.g., suppression), a 'tagging baseline' may also be of use. That is, a period of time where the tagged stimuli are presented but without the experimental manipulation of interest. This then provides a base amplitude of tagging compared to which modulations can be then observed as positive or negative. Without such a tagging baseline, even if it is possible to compare different tagging conditions to each other, it may not be possible to tell whether a particular condition reflects a suppression or enhancement of the RIFT response. Other experiments may warrant the use of a 'cognitive baseline', where experimental stimuli are presented without tagging. This can help identify whether and at which frequencies oscillatory effects arise independently of tagging. This baseline is especially critical when analyzing IM frequencies that overlap with lower-frequency bands.

4.3 Channel Selection

4.3.1 The RIFT topography differs based on tagging location

All analysis techniques described above in *Section 4.1* are computed independently for all channels. As is the case with any M/EEG analysis, channel selection is then an important step when analyzing the RIFT response. This is especially true here since the topography of the RIFT response varies based on the tagging location, i.e., the same tag may evoke different response patterns across channels in trials with different tagging locations. This is evident from the topographies in **Figure 6C**, where lateralized peaks are seen depending on which hemisphere the tagged stimulus is displayed in. Channels can either be selected separately for each location at which tags were displayed (but, importantly, still blind to the experimental conditions of interest; (Wang et al., 2025)), or, first the tagging amplitude across different location presentations can be averaged to produce a general tagging response topography from which channels can then be selected (Arora et al., 2025; Bouwkamp et al., 2025). The former would allow for a stronger response since channel selection can then reflect the unique topographies across different locations of presentation, however, the latter does not require splitting trials into various bins which can negatively affect statistical power and SNR.

Here, we show using Dataset 1 that there is very little difference between the two procedures, both in terms of the resulting RIFT response obtained, and the modulation (here by means of covert attentional shifts) to this response (**Figure 7** blue vs. pink). This conveys that it may not be necessary to split trials into conditions based on where the tagging was displayed, retaining higher SNR. Dataset 1 included two different tagging locations each at 6 degrees of visual angle (dva) horizontal eccentricity.

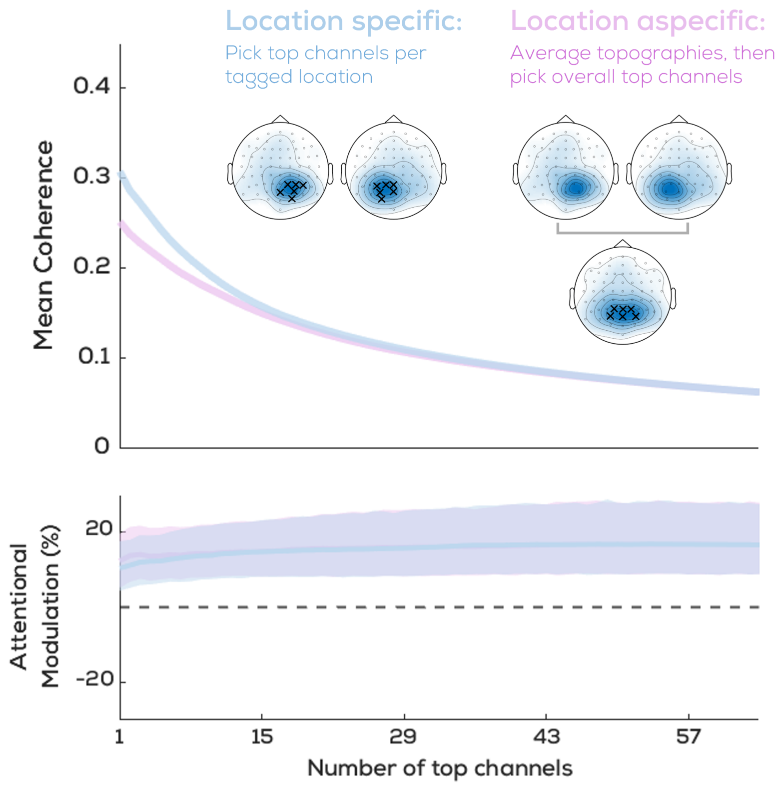

Figure 7: Top-down RIFT modulation is robust across channel selection choices. The attentional modulation of RIFT (as measured across 24 participants from Dataset 1) is independent of the sequence of channel selection and averaging, as well as the number of top channels selected. **Top:** Mean coherence amplitude across participants and **Bottom:** relative attentional modulation of this coherence amplitude as a function of the number of best (highest overall amplitude) channels selected.

4.3.2 Number of channels selected does not affect the top-down RIFT modulation

Since slight differences in the exact topographies can be expected across participants (see Arora et al., 2025 for an overview of topographies across 72 participants), most studies conduct a participant-wise selection of channels. This can be done either through a participant-wise selection of 'n' channels that show the strongest tagging response, or a selection of all channels that show a significant difference in tagging response from baseline. In case a channel-average referencing procedure is implemented during preprocessing, the former option is preferred. Making a selection of top channels should ideally be done using an independent dataset, for example, a tagging baseline (see *Section 4.2*) where the tagging is presented without the experimental manipulation of interest.

Here, we show that attentional modulations to RIFT are independent of the exact number of channels selected. We looked at the effect of how many channels are selected on coherence and its attentional modulation as measured from Dataset 1 **(Figure 7)**. Although average coherence decreases as more channels are selected, the relative top-down modulation of this coherence (% change with attention in/out) is unaffected regardless of how many channels are selected, meaning that the exact number of channels selected is not as relevant. Alternatively, spatial filters such as Rhythmic Entrainment Source Separation (Cohen & Gulbinaite, 2017) may be used, circumventing concerns with numbers of channels.

4.3.3 Additional considerations for MEG

With MEG there is an additional consideration, namely that of whether to use magnetometers or gradiometers. Previous work showed that the two measures are largely comparable, however magnetometers compared to planar gradiometers had stronger sensitivity to tagging (Minarik et al., 2023). A majority of the current studies use axial gradiometers, which some convert to planar gradiometers (Bouwkamp et al., 2025; Drijvers et al., 2021; Seijdel et al., 2024; Zhigalov et al., 2019).

A further concern with MEG, compared to EEG, is that the sensor topographies are likely more heterogeneous across participants (i.e., EEG is strongly spatially low-pass filtered). Therefore, more care is needed for channel selection in

MEG than in EEG. The case for (anatomical and/or signal-driven) spatial filtering may thus be stronger in MEG than in EEG.

4.4 Variation in the RIFT response across participants is not driven by standard collection noise.

With RIFT, we operate close to the threshold of stimulation that does not produce a measurable tagging response in the M/EEG signal (> 72Hz as discussed in *Section 3.1*). Even when operating at a feasible tagging frequency, making stimuli smaller, more peripheral, or using less of the dynamic luminance range may eliminate a measurable tagging response. Importantly, each participant's tagging amplitude drop-off may scale differently with any of these factors (Minarik et al., 2023). It is worth highlighting this because in the authors' experience a tagging response is not visible in every participant with every tagged stimulus design. In some designs, only two-thirds of the participants show viable peaks at the tagging frequencies compared to non-tagged frequencies (Pan et al., 2021), even when using more sensitive analysis methods such as coherence. Here, we visualize the spread of the tagging response across participants in two experiments **(Figure 8)**, one with a relatively large tagged area showing a response in almost all participants (Dataset 1) and one with a relatively smaller tagged area showing a response in roughly three-quarters of participants (Dataset 2).

Even without participants that can be qualitatively labelled as 'unresponsive' to the tagging, there is a lot of variability in the strength of the response of Dataset 1. We conducted further analyses in an attempt to identify the source of this variability. One possibility is that this variability only reflects standard collection noise, since any measure of frequency amplitude or phase would be negatively impacted in participants with noisier signals. Such noise would also be visible as variability in the raw voltage recordings. We correlated the participant-wise variability in the raw voltage (trial-wise standard error in event-locked ERPs) to coherence. Despite picking several points of time along the trial to use for the voltage variability measure, we observed no correlation between signal noise (ERP variability) and coherence **(Figure 9A-D)**. That is, participant variability in the tagging response is not driven by overall noise in the data.

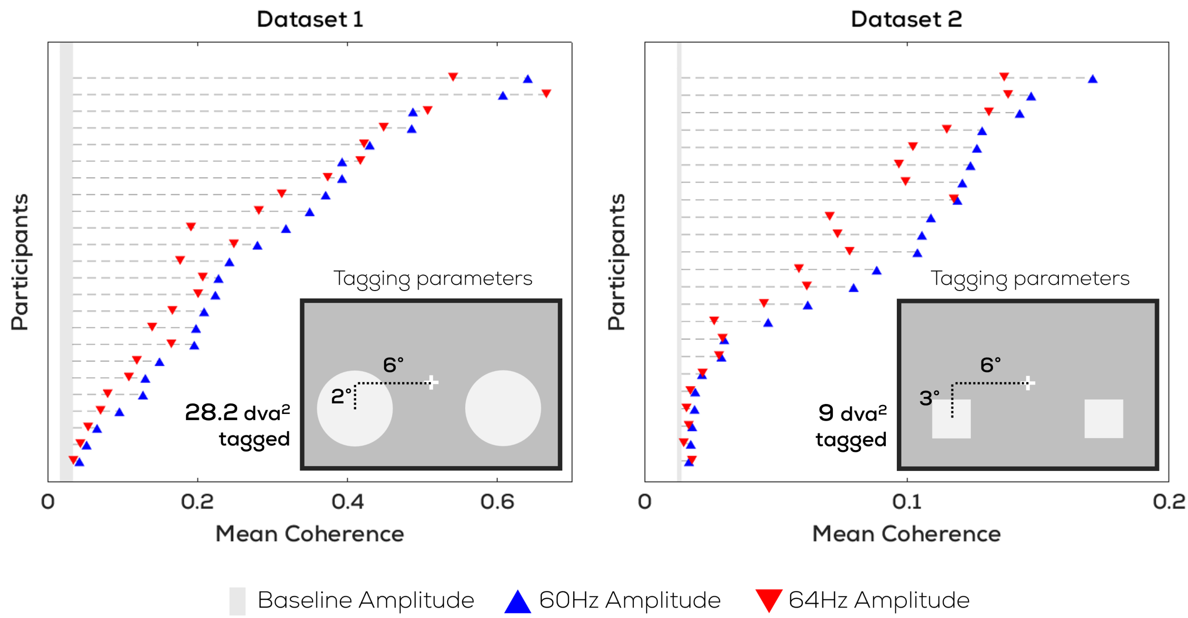

Figure 8: Spread of 60Hz and 64Hz coherence across participants in Datasets 1 and 2. Coherence amplitudes from two tagged frequencies compared to non-tagged frequency baseline for **A.** Dataset 1 **B.** Dataset 2. Insets show tagged stimulus size and eccentricity. Dataset 2 shows a larger proportion of participants with a tagging response minimally distinguishable from baseline, most likely due to smaller tagged area.

Next, we wanted to see whether this variability in the tagging amplitude affects the attentional modulations that can then be observed in the tag. With Dataset 1, for example, this effect of interest comes from increased covert attention at the tagging location to encode an item presented there. We compared the participant-wise coher-

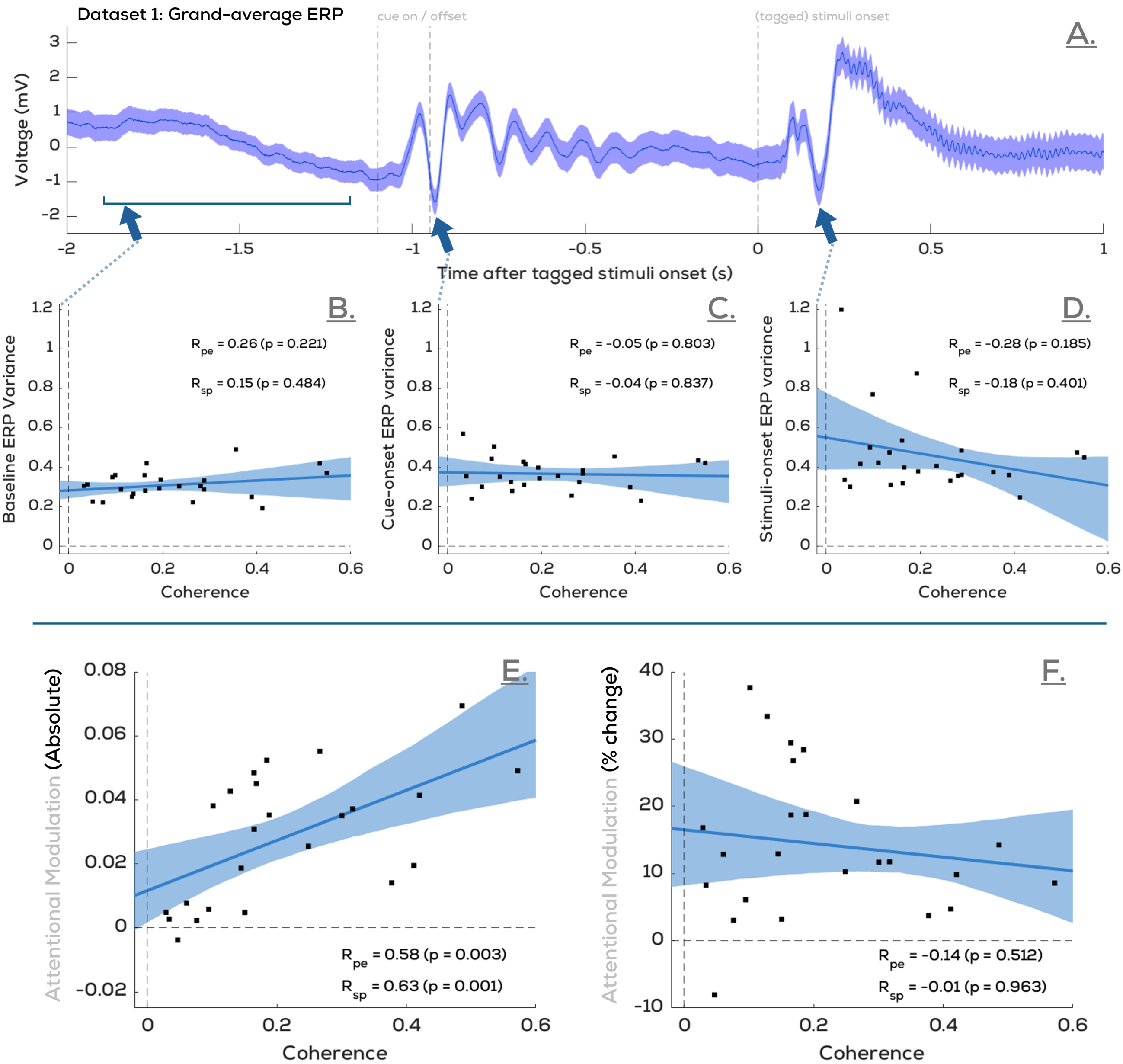

Figure 9: Coherence amplitude is not correlated with trial-wise ERP variation across participants. A. ERP averaged across participants and top channels similarly to coherence, shaded areas represent mean trial-wise SEM. Pearson (R_{pe}) and Spearman (R_{sp}) correlations between participant-wise coherence amplitude and trial-wise SEM of **B.** baseline (-1.8 to -1.1s), **C.** cue ERP (-0.97s), and **D.** stimuli ERP (0.17s). **Coherence amplitude is correlated with absolute, but not relative, attention modulation.** Pearson (R_{pe}) and Spearman (R_{sp}) correlations between coherence amplitude and **E.** absolute attentional modulation of coherence **F.** relative attentional modulation of coherence. Coherence and attentional modulation from Dataset 1 averaged in the interval with a significant attentional modulation at the group level (0.28-1.14s after tag onset). Data from 24 participants of Dataset 1.

ence to the attentional modulation of this signal (i.e., the boost in the RIFT response from covertly attending the tagged location) using Dataset 1. Naturally, having a higher overall coherence was strongly linked to a higher absolute attentional effect **(Figure 9E)**. However, interestingly, when looking at the relative attentional effect (i.e., the cognitive effect as a fraction of the overall coherence), there was no benefit afforded by higher overall coherence amplitudes **(Figure 9F)**. That is, the effect observed in **Figure 9E** is caused only by linear scaling of the amplitude modulation. Thus, variability in tagging strength across participants may not be of consequence to experimental manipulations, provided that a viable tagging peak is observed.

4.5 Spontaneous eye movements around fixation do not impact the RIFT response

Most studies on covert attention, or shifts of spatial attention, require participants to fixate the center of the screen while some stimuli of interest are presented in the periphery (Posner, 1980). Since in the visual domain RIFT acts as a tracker of spatial attention, this is also the design of many RIFT studies (Arora et al., 2025; Bouwkamp et al., 2025; Ferrante et al., 2023; Seijdel et al., 2024; Zhigalov et al., 2019). The fixation requirement is frequently controlled with an eye tracker, so that trials containing large saccades can be identified and excluded. But measures like RIFT are subject to an additional eye-movement related concern: visual stimuli close to the fovea are processed much more strongly than those in the periphery. The absence of large saccades towards a tagged stimulus or location does not eliminate the presence of small deviations in gaze position (<1 dva) that might nonetheless bring the tag closer or further away from the fovea. Does variability in gaze position at these small scales consistently drive the RIFT response?

This question has previously been addressed at the trial level when linked to consistent attentional modulations of gaze position through microsaccadic action (Arora et al., 2025), and we recommend researchers to conduct similar trial-level tests when using RIFT in designs where eye movements can be a potential confound. Here, we present a more sensitive analysis in which we conducted a timepoint-by-timepoint comparison of gaze position and the RIFT amplitude of two peripheral tags (one at 60Hz and

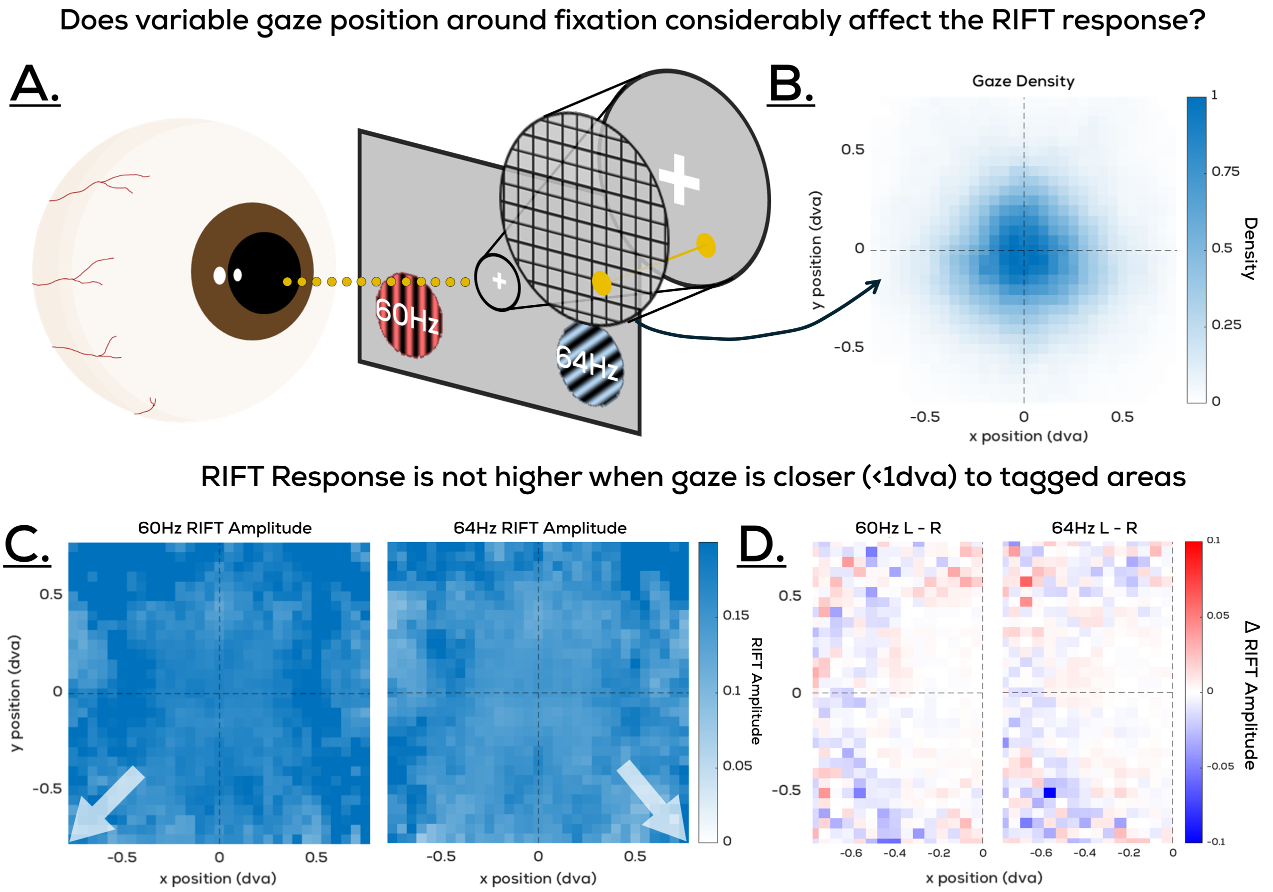

Figure 10: Minor gaze deviations (<0.75 dva) around fixation do not correlate with the RIFT response. A. Dataset 1 contained one 60Hz and one 64Hz tagged stimulus on each trial **B.** Density plot of gaze position averaged over participants and the 1s interval of item presentation. **C.** Mean 60Hz and 64Hz tagging signal when gaze position was in the corresponding bins. Arrows point towards the location of corresponding tagged stimuli. **D.** Lateralization in tagging signal; a cluster permutation test showed no significant clusters, i.e., small gaze deviations towards the left vs. the right do not enhance the tagged response of the stimulus on the left vs. the right.

the other at 64Hz) on the screen from Dataset 1. The trials from Dataset 1 consisted of two tagged stimuli slightly below left and right of fixation.

First, we computed participant-wise gaze densities within a square (side length 1.5 dva) centered at fixation per participant **(Figure 10B)**. Gaze was present within this square for 96.7% of the duration used. Then, we conducted the following analysis for each binned section of the gaze density plot (bin width 0.05 dva). We isolated all the time points during which gaze was within that particular bin. We then averaged the RIFT amplitude over all these timepoints. Thus, we obtained a measure of the mean RIFT amplitude from when participants' gaze position was localised within each bin **(Figure 10C)**. We next tested whether gaze deviation away from fixation and towards the tagged stimuli locations influenced the RIFT amplitude. The lateralization plots testing this (left minus right; averaged across participants; **Figure 10D**) show no consistently positive difference for 60Hz or negative difference for 64Hz. This was also statistically confirmed with a 2D cluster based permutation test (Maris & Oostenveld, 2007).

Thus, small fixational eye movements (<0.75 dva) do not meaningfully drive the RIFT response. This confirms that an enhancement of the RIFT response measured in such cognitive tasks can reflect a genuine top-down modulation of the responsiveness to the tagged location or stimulus, rather than simply being a correlate of changes in foveation.

5 Interpreting the RIFT Response

RIFT utilizes a relatively well understood property of the visual system: its responsiveness to changes in luminance. Unlike traditional SSVEP, it achieves this without being consciously perceived. Since conscious perception is associated with more downstream areas of the visual hierarchy, this implies that the response to RIFT stimulation is limited to upstream visual areas; i.e., early visual cortex. This selectivity is part of what makes RIFT attractive: conventional stimuli unavoidably elicit responses from the whole visual system, but this technique allows researchers to non-invasively measure a response that is selectively obtained from an early subset of this system and thus not a direct readout of high-level processing. The well-established attentional modulation of RIFT responses therefore likely reflects the influence of attentional processing in higher-order regions on the early visual cortex activity.

Where does the boundary lie between the neural activity captured by RIFT and the subsequent downstream processing? The few existing studies that have utilized RIFT do not convey a unanimous answer to this question. Terms such as "cortical excitability" (Zhigalov et al., 2019), or [responses from] the "early visual cortex" (Arora et al., 2025; Drijvers et al., 2021; Pan et al., 2021), "predominantly primary visual cortex" (Duecker et al., 2025), "occipital cortex" (Minarik et al., 2023), have been used to describe the RIFT response, both with and without the explicit mention of regions such as V1 or V2. There is no doubt that the peak RIFT response is in fact localized to relatively early processing areas; its retinotopy and existing MEG source localization work (Drijvers et al., 2021; Duecker et al., 2025; Ferrante et al., 2023; Minarik et al., 2023) is clear evidence of this. However, even within these studies, there are slight variations in regions to which the RIFT response is attributed (for e.g., "V1": Duecker et al., 2025 "V1/V2": Ferrante et al., 2023).

Applying a combination of visual and auditory tagging has given rise to IM frequencies that have been localized beyond regions typically responsive to sensory processing, such as (left) frontotemporal cortex (Drijvers et al., 2021; Hustá et al., 2025; Seijdel et al., 2024). This suggests that, intriguingly, RIFT signals may progress to downstream areas of the visual hierarchy, where consequences of their (nonlinear) processing, i.e., the IM peaks, are detectable.

6 Closing

RIFT provides a powerful new way to study rhythmic cognition. By embedding invisible, high-frequency tagging into stimuli, it offers a clean, and continuous marker of neural processing that avoids the perceptual confounds of traditional frequency tagging. As this manual outlines, careful attention to hardware, tagging parameters and analysis methods are imperative for the success of RIFT experiments. This manual therefore serves as a guide for implementing Rapid Invisible Frequency Tagging (RIFT) in cognitive neuroscience research. By outlining detailed protocols and recommendations for setup, experimental design, and analysis (based on both prior experience and data-derived evidence from multiple labs) it provides researchers with the necessary tools to leverage RIFT's capabilities effectively. With these best practices in place, RIFT opens opportunities to probe attention, perception, memory and multimodal integration, under naturalistic settings, thus bridging the gap between experimental research and everyday cognitive functions. In sum, we hope that this guide will help establish RIFT as a standard approach for studying the oscillatory dynamics that shape human cognition.

Looking ahead, there is a compelling need for ongoing research aimed at enhancing the accessibility of the required technology, refining the procedures to minimize variability, developing more robust methods for data analysis, and expanding the growing list of theoretical questions that RIFT is used to study. By continuing to develop and apply the RIFT technique in cognitive research, the neuroscience community can push the boundaries of our understanding of brain function and cognition.

Supplementary Material and Code

We make use of two previously collected datasets to demonstrate common RIFT design and analysis techniques.

In Dataset 1 (Arora et al., 2025), 24 participants performed 480 trials each of a working memory task (the 'pre-cue' experiment in Arora et al., 2025). Here, we only make use of the 1s period of this task when two gratings, tagged at 60Hz and 64Hz, were displayed in the lower visual field. Further details on the task, data preprocessing, and analysis can be found in Arora et al., 2025). In Dataset 2 (unpublished data), 23 participants viewed a grid display with three grid locations, tagged at 60Hz, 64Hz, and 68.5Hz respectively. Figure S1 provides an overview of the stimulus and tagging parameters for both datasets.

Example experimental files and useful functions for designing RIFT experiments in combination with the PROPixx projector can be found in the following repository: https://osf.io/9mv3e/

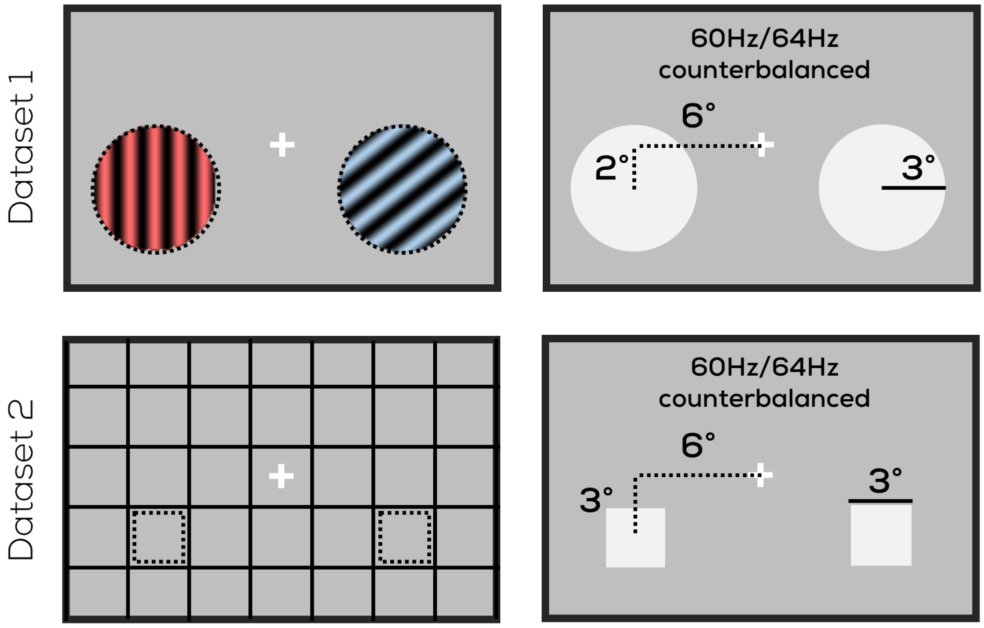

Figure S1: Screen display and tagging parameters for Datasets 1 and 2.

References

Alp, N., Kohler, P. J., Kogo, N., Wagemans, J., & Norcia, A. M. (2018). Measuring integration processes in visual symmetry with frequency-tagged EEG [Publisher: Nature Publishing Group UK London]. *Scientific Reports*, *8*(1), 6969.

Arora, K., Gayet, S., Kenemans, J. L., Van der Stigchel, S., & Chota, S. (2025). Dissociating external and internal attentional selection [Publisher: Elsevier]. *iScience*, *28*(4).

Bouwkamp, F. G., de Lange, F. P., & Spaak, E. (2025). Spatial predictive context speeds up visual search by biasing local attentional competition [Publisher: MIT Press 255 Main Street, 9th Floor, Cambridge, Massachusetts 02142, USA ...]. *Journal of Cognitive Neuroscience*, *37*(1), 28–42.

Brickwedde, M., Bezsudnova, Y., Kowalczyk, A., Jensen, O., & Zhigalov, A. (2022). Application of rapid invisible frequency tagging for brain computer interfaces. *Journal of neuroscience methods*, *382*, 109726.

Cass, J., Van der Burg, E., & Alais, D. (2011). Finding flicker: Critical differences in temporal frequency capture attention [Publisher: Frontiers Research Foundation]. *Frontiers in psychology*, *2*, 320.

Cohen, M. X., & Gulbinaite, R. (2017). Rhythmic entrainment source separation: Optimizing analyses of neural responses to rhythmic sensory stimulation [Publisher: Elsevier]. *Neuroimage*, *147*, 43–56.

Dimigen, O., Badea, I., Simon, I., & Span, M. M. (2025). Rapid invisible frequency tagging (RIFT) with a consumer monitor: A proof-of-concept [Publisher: Cold Spring Harbor Laboratory]. *bioRxiv*, 2025–08.

Dimigen, O., & Stein, A. (2024). A high-speed OLED monitor for precise stimulation in vision, eye-tracking, and EEG research [Publisher: Cold Spring Harbor Laboratory]. *bioRxiv*, 2024–09.

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 [Publisher: Society for Neuroscience]. *Journal of Neuroscience*, *41*(31), 6684–6698.

Duecker, K., Shapiro, K. L., Hanslmayr, S., Griffiths, B. J., Pan, Y., Wolfe, J. M., & Jensen, O. (2025). Guided visual search is associated with target boosting and distractor suppression in early visual cortex [Publisher: Nature Publishing Group]. *Communications Biology*, *8*(1), 1–11.

Ferrante, O., Zhigalov, A., Hickey, C., & Jensen, O. (2023). Statistical learning of distractor suppression downregulates prestimulus neural excitability in early visual cortex [Publisher: Society for Neuroscience]. *Journal of Neuroscience*, *43*(12), 2190–2198.

Gulbinaite, R., Roozendaal, D. H., & VanRullen, R. (2019). Attention differentially modulates the amplitude of resonance frequencies in the visual cortex [Publisher: Elsevier]. *NeuroImage*, *203*, 116146.

Gutteling, T. P., Sillekens, L., Lavie, N., & Jensen, O. (2022). Alpha oscillations reflect suppression of distractors with increased perceptual load [Publisher: Elsevier]. *Progress in Neurobiology*, *214*, 102285.

Haarlem, C. S., O'Connell, R. G., Mitchell, K. J., & Jackson, A. L. (2024). The speed of sight: Individual variation in critical flicker fusion thresholds [Publisher: Public Library of Science]. *Plos one*, *19*(4), e0298007.

Herrmann, C. S. (2001). Human EEG responses to 1-100hz 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

Hustá, C., Meyer, A., & Drijvers, L. (2025). Using rapid invisible frequency tagging (RIFT) to probe the neural interaction between representations of speech planning and comprehension [Publisher: MIT Press 255 Main Street, 9th Floor, Cambridge, Massachusetts 02142, USA ...]. *Neurobiology of Language*, 1–36.

Krekelberg, B. (2010). Saccadic suppression [Publisher: Elsevier]. *Current Biology*, *20*(5), R228–R229.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data [Publisher: Elsevier]. *Journal of neuroscience methods*, *164*(1), 177–190.

Marshall, T. R., Ruesseler, M., Hunt, L. T., & O'Reilly, J. X. (2024). The representation of priors and decisions in the human parietal cortex [Publisher: Public Library of Science San Francisco, CA USA]. *PLoS Biology*, *22*(1), e3002383.

Martínez-Cagigal, V., Thielen, J., Santamaria-Vazquez, E., Pérez-Velasco, S., Desain, P., & Hornero, R. (2021). Brain–computer interfaces based on code-modulated visual evoked potentials (c-VEP): A literature review [Publisher: IOP Publishing]. *Journal of Neural Engineering*, *18*(6), 061002.

Minarik, T., Berger, B., & Jensen, O. (2023). Optimal parameters for rapid (invisible) frequency tagging using MEG [Publisher: Elsevier]. *NeuroImage*, *281*, 120389.

Morgan, S. T., Hansen, J. C., & Hillyard, S. A. (1996). Selective attention to stimulus location modulates the steady-state visual evoked potential. *Proceedings of the National Academy of Sciences*, *93*(10), 4770–4774. https://doi.org/10.1073/pnas.93.10.4770

Müller, M. M., Andersen, S., Trujillo, N. J., Valdés-Sosa, P., Malinowski, P., & Hillyard, S. A. (2006). Feature-selective attention enhances color signals in early visual areas of the human brain. *Proceedings of the National Academy of Sciences*, *103*(38), 14250–14254. https://doi.org/10.1073/pnas.0606668103

Müller, M. M., & Hübner, R. (2002). Can the spotlight of attention be shaped like a doughnut? evidence from steady-state visual evoked potentials. *Psychological Science*, *13*(2), 119–124. https://doi.org/10.1111/1467-9280.00422

Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottereau, B. R., & Rossion, B. (2015). The steady-state visual evoked potential in vision research: A review [Publisher: The Association for Research in Vision and Ophthalmology]. *Journal of vision*, *15*(6), 4–4.

Notbohm, A., Kurths, J., & Herrmann, C. S. (2016). Modification of brain oscillations via rhythmic light stimulation provides evidence for entrainment but not for superposition of event-related responses [Publisher: Frontiers Media SA]. *Frontiers in human neuroscience*, *10*, 10.

Pan, Y., Frisson, S., Federmeier, K. D., & Jensen, O. (2024). Early parafoveal semantic integration in natural reading [Publisher: eLife Sciences Publications Limited]. *Elife*, *12*, RP91327.

Pan, Y., Frisson, S., & Jensen, O. (2021). Neural evidence for lexical parafoveal processing [Publisher: Nature Publishing Group UK London]. *Nature Communications*, *12*(1), 5234.

Pei, F., Pettet, M. W., & Norcia, A. M. (2002). Neural correlates of object-based attention [Publisher: The Association for Research in Vision and Ophthalmology]. *Journal of Vision*, *2*(9), 1–1.

Regan, M. P., & Regan, D. (1988). A frequency domain technique for characterizing nonlinearities in biological systems [Publisher: Elsevier]. *Journal of theoretical biology*, *133*(3), 293–317.

Regan, M., Regan, D., & He, P. (1995). An audio-visual convergence area in the human brain. *Experimental Brain Research*, *106*(3). https://doi.org/10.1007/BF00231071

Seijdel, N., Marshall, T. R., & Drijvers, L. (2023). Rapid invisible frequency tagging (RIFT): A promising technique to study neural and cognitive processing using naturalistic paradigms [Publisher: Oxford University Press]. *Cerebral Cortex*, *33*(5), 1626–1629.

Seijdel, N., Schoffelen, J.-M., Hagoort, P., & Drijvers, L. (2024). Attention drives visual processing and audiovisual integration during multimodal communication [Publisher: Society for Neuroscience]. *Journal of Neuroscience*, *44*(10).

Spaak, E., Bouwkamp, F. G., & de Lange, F. P. (2024). Perceptual foundation and extension to phase tagging for rapid invisible frequency tagging (RIFT) [Publisher: MIT Press 255 Main Street, 9th Floor, Cambridge, Massachusetts 02142, USA ...]. *Imaging Neuroscience*, *2*, 1–14.

Spaak, E., de Lange, F. P., & Jensen, O. (2014). Local entrainment of alpha oscillations by visual stimuli causes cyclic modulation of perception [Publisher: Society for Neuroscience]. *Journal of Neuroscience*, *34*(10), 3536–3544.

Waldin, N., Waldner, M., & Viola, I. (2017). Flicker observer effect: Guiding attention through high frequency flicker in images. *Computer Graphics Forum*, *36*(2), 467–476. https://doi.org/10.1111/cgf.13141

Wang, D., Arora, K., Theeuwes, J., Stigchel, S. V. d., Gayet, S., & Chota, S. (2025). Dynamic competition between bottom-up saliency and top-down goals in early visual cortex. *bioRxiv*, 2025–08.

Widmann, A., Schröger, E., & Maess, B. (2015). Digital filter design for electrophysiological data–a practical approach [Publisher: Elsevier]. *Journal of neuroscience methods*, *250*, 34–46.

Zhigalov, A., Herring, J. D., Herpers, J., Bergmann, T. O., & Jensen, O. (2019). Probing cortical excitability using rapid frequency tagging [Publisher: Elsevier]. *NeuroImage*, *195*, 59–66.

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