

Neural responses to binocular in-phase and anti-phase stimuli

Bruno Richard^{a,*}, Daniel H. Baker^b

^a*Department of Math and Computer Sciences, Rutgers University, Newark, New Jersey, USA*

^b*Department of Psychology, University of York, York, UK*

Abstract

1 Binocular vision fuses compatible inputs from the two eyes into a single percept, whereas
2 incompatible inputs can produce rivalry, lustre, or diplopia. We measured neural responses
3 to binocular stimuli with different phase relationships to test predictions from contemporary
4 binocular combination models. Steady-State Visually Evoked Potentials (SSVEPs) were
5 recorded from 15 observers in response to monocular and binocular stimulation at 3 Hz, us-
6 ing either On/Off or counterphase flicker with varied spatial and temporal phase relationships.
7 On/Off flicker elicited responses at the fundamental frequency (3 Hz), and its integer har-
8 monics, while counterphase flicker generated responses at the even integer harmonics (6Hz,
9 12Hz, 18Hz). Manipulating phase relationships modulated these response patterns, includ-
10 ing a reduction in the fundamental amplitude for On/Off flicker. The data were modeled
11 with a series of binocular combination algorithms, ranging in complexity from a simple lin-
12 ear sum to a two-stage binocular gain-control model with parallel monocular and binocular
13 phase-selective channels. The model required parallel monocular channels to account for our
14 data, whereas phase selectivity was not essential. Overall, the two-stage contrast gain-control
15 model remains a powerful and flexible framework for describing binocular combinations across
16 various experimental conditions and modalities.

Keywords: Binocular Combination, SSVEP, Contrast Gain Control, Phase Selectivity,
Monocular Channels

*Corresponding author

Email address: br379@newark.rutgers.edu (Bruno Richard)

17 **Introduction**

18 The human visual system integrates input from both eyes to form a unified binocular rep-
19 resentation of the world. Combining monocular inputs enhances sensitivity to the presented
20 stimuli, particularly when the contrast is low or near detection threshold (Baker et al., 2018;
21 Campbell and Green, 1965; Meese et al., 2006). Contrast sensitivity can improve by a factor
22 of $\sqrt{2}$ or more when stimuli are presented binocularly versus monocularly (Baker et al., 2018;
23 Blake and Wilson, 2011; Campbell and Green, 1965; Richard et al., 2018). Notably, the
24 visual system also attempts to combine inputs even when the stimuli presented to each eye
25 are markedly different (i.e., incompatible). In such cases, observers may perceive binocular ri-
26 valy (Blake, 1989; Wilson, 2003), diplopia, or visual lustre (Wendt and Faul, 2022). Despite
27 differences in perceptual outcomes, computationally, the underlying processes of binocular
28 combination for compatible and incompatible inputs appear similar (Baker et al., 2007a;
29 Legge, 1984a). Psychophysical responses to compatible and incompatible stimuli can be ef-
30 fectively explained by a single psychophysical model that involves nonlinear transduction,
31 followed by summation across monocular and binocular phase-selective channels (Baker et
32 al., 2007a; Baker and Meese, 2007). Here, we explore if the integrative processes defined over
33 multiple behavioural tasks are also reflected at the neural level.

34 It has long been known that stimuli presented binocularly are summed across the eyes.
35 In contrast detection tasks, where stimulus contrast is low, binocular presentation increases
36 sensitivity by approximately $\sqrt{2}$ (Campbell and Green, 1965). This implies that observers
37 require roughly 1.4 times more contrast to detect a monocular stimulus than a binocular one.
38 The binocular improvement in sensitivity is consistent with a non-linearity operating before
39 the signals from the two eyes (c_L and c_R) are combined (Legge, 1984b):

$$R_B = c_L^m + c_R^m. \quad (1)$$

40 Here, the exponent m determines the degree of summation (R_B). When $m = 1$, summation is
41 linear, yielding a doubling of sensitivity. When $m = 2$, summation is reduced to $\sqrt{2}$. Several
42 studies have reported summation ratios over $\sqrt{2}$ with some approaching 1.8 (Meese et al.,
43 2006; Simmons, 2005; Simmons and Kingdom, 1998). A recent meta-analysis of 65 studies
44 ($N = 716$) found an average binocular summation ratio of 1.5 (Baker et al., 2018). This work

45 highlighted the challenges in accurately describing the binocular summation process (e.g.,
46 m) as individual variability and methodological differences can greatly impact binocular
47 summation.

48 The contrast of stimuli is crucial in measuring binocular summation. Binocular summa-
49 tion can be very large when stimulus contrast is low; however, the binocular advantage is
50 seldom observed in tasks that involve higher contrasts. In contrast discrimination tasks,
51 where observers judge the contrast difference between otherwise identical stimuli, binocular
52 presentation no longer confers a benefit in sensitivity: discrimination thresholds are identi-
53 cal whether stimuli are shown to one eye or both (Legge, 1984a; Maehara and Goryo, 2005;
54 Meese et al., 2006). However, this does not indicate that binocular summation does not
55 occur at higher contrasts; the monocular signals are still summed (Meese et al., 2006; Meese
56 and Baker, 2011). Instead, the advantage of summation is counteracted by normalization
57 mechanisms that maintain consistency across viewing conditions (referred to as ‘ocularity
58 invariance’). In gain control models of early vision, normalization can stem from interocular
59 and self-suppressive signals (Meese et al., 2006):

$$R_B = \frac{c_L^m}{S + c_L + c_R} + \frac{c_R^m}{S + c_R + c_L}. \quad (2)$$

60 When both eyes are stimulated, the suppressive terms on the denominators offset the en-
61 hanced excitatory signals, nullifying the binocular advantage. **The new parameter, S, is**
62 **included to prevent division by zero and sets overall sensitivity of the monocular response.**

63 A similar pattern of results is observed in neural measurements of binocular summation. In
64 functional Magnetic Resonance Imaging (fMRI) recordings, neural responses are significantly
65 larger for binocular than monocular responses when stimulus contrast is low (Moradi and
66 Heeger, 2009). At higher contrasts, observers no longer show the binocular advantage. As
67 with behavioural results, these findings are well-explained computationally by interocular
68 suppression and binocular contrast normalization. The equivalency in binocular summation
69 between psychophysics and neuroimaging findings means that both data types can constrain
70 binocular summation models. We, for example, have demonstrated that a popular model
71 of binocular summation could be easily adapted to capture Steady-State Visually Evoked
72 Potentials (SSVEPs) to monocular and binocular stimuli when one eye was occluded by

73 a neutral density filter (Richard et al., 2018). Placing a neutral density filter in front of
74 one eye darkens its input, reducing the amplitude and altering the phase of SSVEPs to
75 stimuli presented to the filtered eye. Steady-State response amplitudes and phases to stimuli
76 presented through different neutral density filter strengths were well described by a model
77 that first used a biophysically plausible temporal filter on the input stimuli, followed by self
78 and interocular suppression, and finally, binocular contrast normalization. This model also
79 explained psychophysically measured binocular summation in the same group of observers.
80 A comprehensive description of the processes involved in binocular summation should be able
81 to explain behavioural and neuroimaging findings under various experimental conditions of
82 binocular summation.

83 Many studies have worked towards the development of a comprehensive description of
84 the process of binocular summation in human vision using a variety of psychophysical and
85 neuroimaging approaches (Baker et al., 2008, 2007b; Ding et al., 2013; Ding and Sperling,
86 2006; Legge, 1984a; Lygo et al., 2021; Maehara and Goryo, 2005; May and Zhaoping, 2016;
87 Richard et al., 2018). This is not an insignificant challenge; such a description must account
88 for multiple components of early vision, including identifying relevant signals, defining how
89 these signals might interact, what non-linearities are present, and most importantly, how
90 these signals are summed and/or differenced. One model that has proven very informative
91 and capable of describing binocular summation under various experimental conditions is
92 the two-stage contrast gain control model developed by Meese et al. (2006). In a two-
93 stage process, the model captured detection and discrimination thresholds for monocular
94 and binocular presentation, in addition to dichoptic masking (when the stimuli presented
95 to both eyes are not identical). First, the input is rectified by an excitatory non-linearity
96 ($m \approx 1.3$) and normalized by self and interocular suppression (see Equation 2). The binocular
97 (e.g., combined) signal then undergoes a second contrast normalization before the decision
98 stage,

$$R = \frac{R_B^p}{Z + R_B^q}, \quad (3)$$

99 where the exponents p and q typically have quite large values, and determine the shape of
100 the contrast discrimination ('dipper') functions. As with the parameter S in Equation 2,

101 the parameter Z is included to prevent division by 0 and sets the overall sensitivity of the
102 binocular response.

103 Subsequent iterations of the two-stage contrast gain control model added channels for
104 opposite contrast polarities (Baker and Meese, 2007), and monocular channels parallel to
105 the binocular summing channel (Georgeson et al., 2016). Polarity-specific channels were
106 included to explain masking effects when the stimuli presented to each eye had opposite
107 phase polarities (i.e., dichoptic presentation). Antiphase stimuli do not cancel each other
108 out (as would be expected if their luminances were linearly summed): the stimuli remain
109 detectable and the two inputs can sum (Bacon, 1976; Baker and Meese, 2007; Simmons,
110 2005), at least in a probabilistic sense. Parallel monocular channels are consistent with
111 adaptation after-effects that suggest monocular signals may be preserved and available for
112 perception following binocular summation (Blake et al., 1981; Moulden, 1980).

113 While the possibility of monocular channels had been considered (Legge, 1984a), many
114 assumed that only the binocularly summed signal contributed to perception. Georgeson et al.
115 (2016) developed a specific experimental condition to assess the involvement of monocular
116 channels. They devised a discrimination task where the target interval presented a contrast
117 increment to one eye (e.g., 10% pedestal + 2%) and a contrast decrement to the other (e.g.,
118 10% pedestal - 2%). If the only available signal is a binocularly summed one, the task would
119 be nearly impossible to complete; the target interval would be perceptually identical to the
120 pedestal-only interval. However, observers were able to complete the task. The two-stage
121 contrast gain control model with parallel monocular channels was the only model able to
122 capture observer thresholds from all experimental conditions, including the binocular incre-
123 ment and decrement tasks. Evidence of an additional differencing channel accompanying
124 the summing channel commonly described in work on binocular combination (Chen and Li,
125 1998; Li and Atick, 1994; May et al., 2012; May and Zhaoping, 2016) has also accumulated.
126 As the name implies, this channel encodes the difference between stimuli presented to the
127 left and right eye; it is involved in early contributions to disparity processing and stereovi-
128 sion. While computational models of binocular vision do not explicitly include differencing
129 channels, encoding the difference between monocular signals is, indirectly, included in the

¹³⁰ two-stage contrast gain control model defined by Georgeson et al. (2016).

¹³¹ Steady-State Visually Evoked Potentials offer a unique opportunity to bridge the gap
¹³² between models developed from psychophysical data and neuroimaging, as responses to stim-
¹³³ ulus contrast (i.e., SSVEP amplitude) are directly associated with behavioral sensitivity to
¹³⁴ contrast (Norcia et al., 2015; Wade and Baker, 2025). Two common stimulus presentation
¹³⁵ protocols are used to generate SSVEPs: a sinusoidal On/Off flicker, where the stimulus al-
¹³⁶ ternates between a blank background (0% contrast) and the peak contrast, and sinusoidal
¹³⁷ counterphase flicker, where the stimulus alternates in phase (i.e., the black regions become
¹³⁸ white, and the white regions become black). On/Off flicker activates different populations
¹³⁹ of on and off-cells independently once per cycle, generating frequency-following responses at
¹⁴⁰ the frequency of the sinusoidal modulation (1F). On/Off flicker can also generate SSVEPs
¹⁴¹ at the odd and even integer harmonics (2F, 3F, 4F, etc.), which reflect nonlinear process-
¹⁴² ing in the visual system (Regan and Regan, 1988). Counterphase flicker will generate two
¹⁴³ transients per cycle, resulting in SSVEPs at the even harmonics of the flicker frequency (2F,
¹⁴⁴ 4F, 6F, etc.). These frequency-doubling responses are argued to provide a cleaner measure
¹⁴⁵ of nonlinear visual responses (Kim et al., 2011; Skottun and Skoyles, 2007). Utilizing both
¹⁴⁶ stimulation protocols enables a comprehensive description of linear and nonlinear visual pro-
¹⁴⁷ cesses, thereby generating the data necessary to evaluate psychophysical models of visual
¹⁴⁸ perception.

¹⁴⁹ The current architecture of the two-stage contrast gain control model has been rigorously
¹⁵⁰ evaluated on psychophysical data, providing a solid foundation for our understanding of
¹⁵¹ binocular combination (Baker and Meese, 2007; Georgeson et al., 2016; Meese et al., 2006).
¹⁵² While previous studies have applied the two-stage contrast gain control model to neuroimag-
¹⁵³ ing data (Lygo et al., 2021; Richard et al., 2018), they only included experimental conditions
¹⁵⁴ where the phases of the sinusoidal gratings presented to each eye were identical. To accurately
¹⁵⁵ assess the current architecture of the two-stage contrast gain control model, neuroimaging
¹⁵⁶ data for binocular presentation of stimuli in opposite phase polarity are required. Here, we
¹⁵⁷ recorded SSVEPs to monocular and binocular stimuli with different spatial and temporal
¹⁵⁸ phase relationships and stimulation protocols (On/Off and counterphase flicker) to obtain

159 the data needed to evaluate the two-stage contrast gain control model. By progressively
160 increasing the complexity of the model, we demonstrate that many components of binocular
161 combination, such as monocular non-linearities, interocular interactions, and parallel mono-
162 ocular channels, are required to explain neural responses to our set of experimental conditions.
163 The two-stage contrast gain control model remains a powerful and flexible descriptor of the
164 architecture of binocular combination for data collected across many experimental conditions
165 and modalities.

166 Methods

167 Participants

168 Fifteen observers ($N_{male} = 4$; age range [19 34]), including authors BR and DHB, par-
169 ticipated in this study. All observers had normal or corrected-to-normal visual acuity, as
170 verified by a Snellen chart, and normal binocular vision, as verified by a Titmus test. Writ-
171 ten informed consent was obtained from all participants, and experimental procedures were
172 approved by the ethics committee of the Department of Psychology at the University of York.

173 Apparatus

174 All stimuli were presented using a gamma-corrected ViewPixx 3D display (VPixx Tech-
175 nologies, Canada) driven by a Mac Pro. Binocular separation with minimal crosstalk was
176 achieved by synchronizing the display's refresh rate with the toggling of a pair of Nvidia
177 stereo shutter goggles using an infrared signal. The monitor refresh rate was set to 120 Hz;
178 each eye updated at 60 Hz (every 16.67 msec). The display resolution was set to 1920×1080
179 pixels. A single pixel subtended 0.027° of visual angle (1.63 arc min) when viewed from 57 cm.
180 The mean luminance of the display viewed through the shutter goggles was 26 cd/m^2 . EEG
181 signals were recorded from 64 electrodes distributed across the scalp according to the 10/20
182 EEG system (Chatrian et al., 1985) in a WaveGuard cap (ANT Neuro, Netherlands). We
183 monitored eye blinks with an electrooculogram consisting of bipolar electrodes placed above
184 the eyebrow and the cheek on the left side of the participant's face. Stimulus-contingent trig-
185 gers were sent from the ViewPixx display to the amplifier using a parallel cable. Signals were
186 amplified and digitized using a PC with the ASA lab software (ANT Neuro, Netherlands).

187 All EEG data were imported into MATLAB (Mathworks, MA, USA) using components of
188 the *EEGlab* toolbox (Delorme and Makeig, 2004) and then exported for subsequent offline
189 analysis using R.

190 *Stimulus Creation*

191 Observer SSVEPs were measured with a single horizontal sinusoidal grating that sub-
192 tended 15° of visual angle on the retina with a spatial frequency of 3 cycles/ $^\circ$ of visual angle
193 (Figure 1A). All stimuli were generated using *MATLAB* (MathWorks, Natick, MA) and pre-
194 sented to observers using *Psychtoolbox* (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997).
195 Our experimental conditions modulated the interocular spatial phase of stimuli (Figure 1A).
196 Under binocular viewing, the sinusoidal gratings could be presented in spatial phase or spa-
197 tial anti-phase. When stimuli were presented in spatial phase, the aligned sinusoidal gratings
198 were identical in both eyes ($\Delta\phi = 0$). The spatial anti-phase condition phase-shifted one of
199 the sinusoidal gratings by 180° ($\Delta\phi = \pi$). Stimuli were also modulated in their oscillatory
200 pattern, which could be On/Off or counterphase flicker at a frequency of 3Hz (Figure 1B).
201 Under On/Off contrast flicker, the relative contrast of the gratings began at 0%, increased
202 smoothly to 100% of the nominal maximum (100% Michelson contrast), and then returned
203 to 0% over 333 ms (i.e., one cycle). On/Off flicker will generate SSVEPs at the fundamental
204 frequency (3Hz) and its integer harmonics (2F, 3F, 4F, see Figure 1C). Counterphase flicker
205 reversed the phase of the gratings at a frequency of 3Hz. The contrast of the grating began
206 at the relative maximum (100%), gradually decreased to 0% of the relative maximum, and
207 then increased again to 100% but in the opposite phase polarity. Unlike On/Off flicker, coun-
208 terphase flicker generates two nearly identical transients per cycle and thus does not produce
209 SSVEPs at the fundamental frequency (3Hz) but only its even harmonics (Wade and Baker,
210 2025).

211 To aid with binocular fusion, stimuli were surrounded by a static binocular texture pre-
212 sented beyond the central 19° stimulus aperture. These textures were constructed by first
213 low-pass filtering a white ($\text{amplitude} \propto 1/f^0$) noise pattern, dichotomizing its output into
214 a binary image, and taking its phase spectrum. A second flat ($\text{amplitude} \propto 1/f^0$) was ad-
215 justed by multiplying each spatial frequency's amplitude coefficient by f^{-1} to generate a

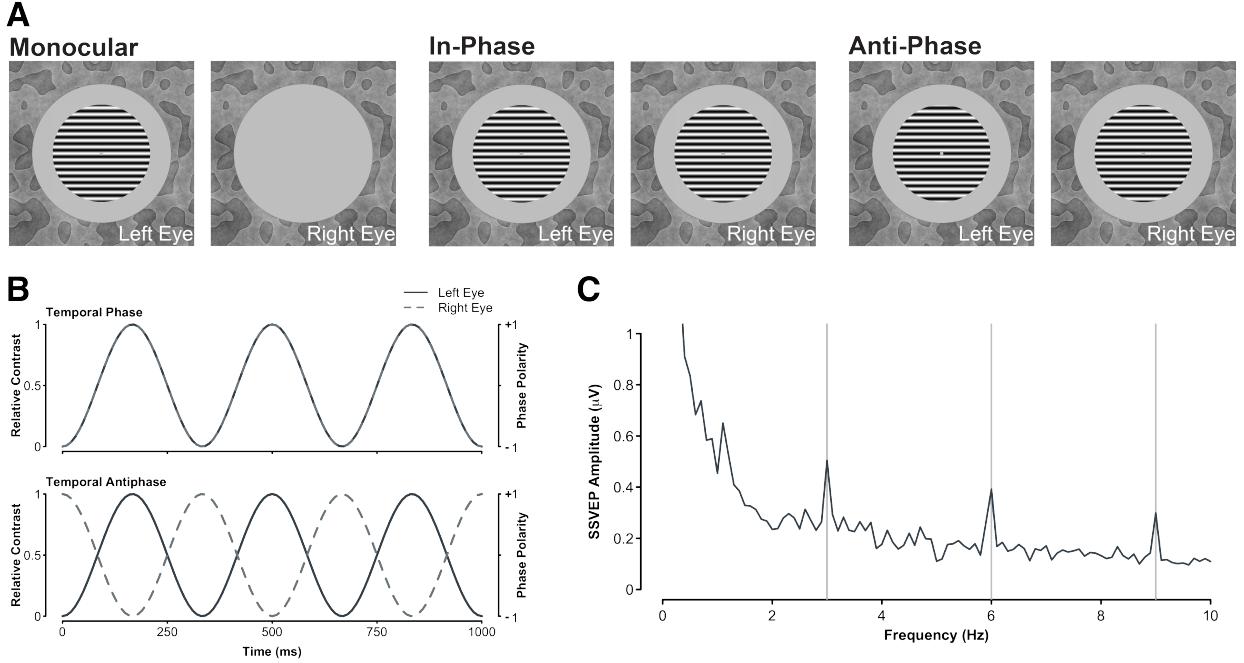


Figure 1: **A.** The spatial configuration of stimuli presented to observers in our experiment. Monocular conditions presented the sinusoidal grating to the left or right eye of observers (counterbalanced) while the other eye was presented with a gray screen set to mean luminance. Binocular conditions could be shown with stimuli in spatial phase, whereby the phase of both sinusoidal gratings was identical, or in spatial anti-phase, where the phase of the sinusoidal gratings presented to each eye was opposite. The background texture did not change throughout the trial to aid with binocular fusion. **B.** The temporal configuration of our stimuli. To generate SSVEPs, stimuli were contrast modulated in two ways: on/off (left Y axis) or counterphase (right Y axis). The oscillatory pattern could also be in phase (upper plot), where both stimuli were modulated in the same manner, or in counterphase (lower plot), where as one stimulus increased in contrast, the other decreased in contrast (or increased in opposite polarity). **C.** Example SSVEPs generated under binocular spatial and temporal in-phase viewing of stimuli for one observer, averaged across four electrodes (Oz , POz , $O1$, $O2$) and 12 repetitions.

216 pink amplitude spectrum (Hansen and Hess, 2006; Tadmor and Tolhurst, 1994). The pink
217 amplitude spectrum and the phase spectrum of the binary image were rendered in the spatial
218 domain by taking the inverse Fourier transform, resulting in the pattern shown in Figure 1A.
219 The surrounding texture was fixed during a trial, but its phase structure was changed across
220 trials.

221 *Procedures*

222 Steady-State Visually Evoked Potentials (SSVEPs) were recorded with monocular and
223 binocular stimulation using either on-off or counterphase flicker at 3Hz. Across the eyes,
224 binocular stimuli could be in spatial and temporal phases, temporal phase but spatial anti-
225 phase, spatial phase but temporal anti-phase, or spatial and temporal anti-phase (on-off
226 flicker only). Stimuli presented in spatial and temporal anti-phase under counterphase flicker
227 are identical to stimuli presented in spatial and temporal phase (and so we did not duplicate
228 this condition). Thus, this experiment was comprised of nine conditions - two monocular and
229 seven binocular - each repeated 12 times for a total of 108 trials. Stimulus presentation was
230 separated into four experimental blocks, each containing 27 trials. A trial lasted 15 seconds;
231 a grating stimulus flickered onscreen for 11 seconds, followed by a screen with its central 19°
232 set to mean luminance for 4 seconds. Participants completed all 27 trials of an experimental
233 block in a single sequence (6.75 minutes) and were given breaks between experimental blocks.
234 The trial order was pseudo-randomized on each block. Participants did not receive explicit
235 task instructions other than to fixate the marker in the center of the display and blink only
236 during the blank period between stimulus presentations.

237 *SSVEP Analysis*

238 We used whole-head average referencing to normalize each electrode to the mean signal of
239 all 64 electrodes (for each sample point). The EEG waveforms were Fourier transformed at
240 each electrode for a 10-second window, beginning one second after the stimulus onset to avoid
241 onset transients. The Fourier spectra were coherently averaged (i.e., retaining the phase
242 information) across four occipital electrodes (*Oz*, *POz*, *O1*, and *O2*) and trial repetitions
243 (see Figure 1C). We then calculated signal-to-noise ratios (SNRs) by dividing the absolute
244 amplitude in the signal bin (e.g., 3 Hz) by the mean of the absolute value of the ten adjacent

245 bins (± 0.5 Hz in steps of 0.1 Hz). There were no differences in the amplitude of the adjacent
 246 bins across experimental conditions. Given that distributions of ratios (including SNRs) are
 247 inherently skewed, the median SNR was taken across all participants; the median is a more
 248 robust descriptor of central tendency in skewed distributions.

249 **Results**

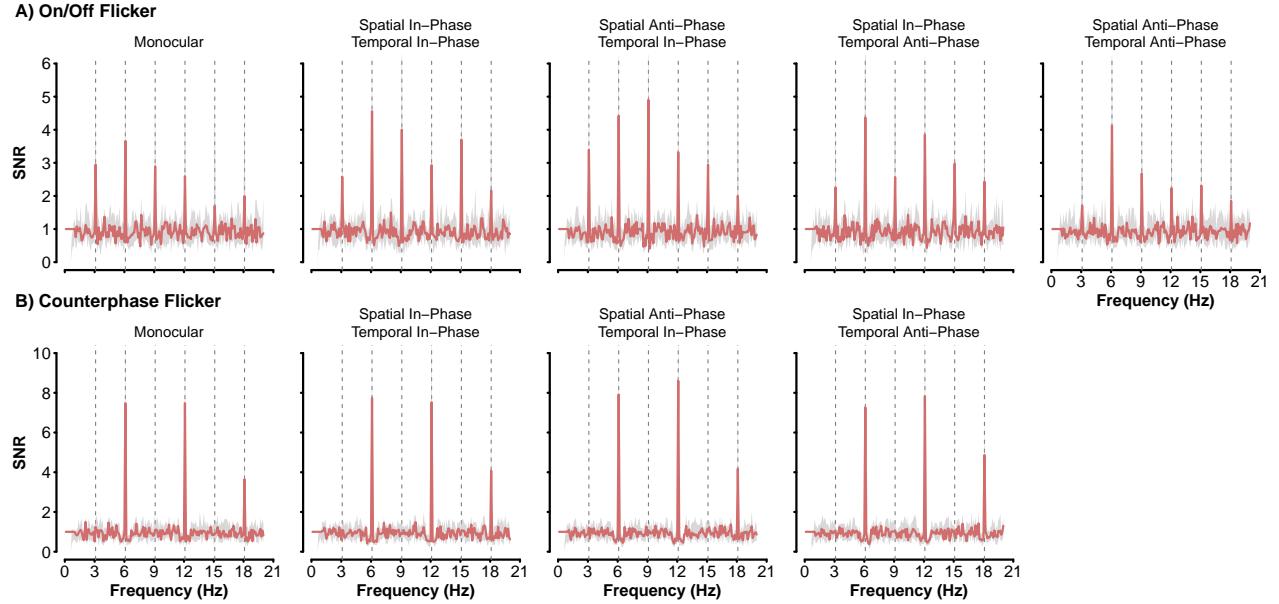


Figure 2: Cross-participant median SNRs for frequencies up to 20Hz. SNRs generated by On/Off flicker are shown in the top row (A), while those generated by counterphase flicker are shown in the bottom row (B). The light gray area represents bootstrapped 95% confidence intervals that were calculated by resampling (with replacement) participant SNRs 2000 times.

250 Figure 2 shows the cross-participant median SNR spectra for all experimental conditions.
 251 Responses for all On/Off flicker experimental conditions generated peaks at the fundamental
 252 frequency (3Hz) and its harmonics (integer multiples of 3Hz). Similarly, counterphase flicker
 253 produced responses at twice the flicker frequency (6 Hz) and its harmonics. We assessed
 254 differences in SNR magnitude across experimental conditions via a permutation test, allowing
 255 a non-parametric comparison of a statistic between two conditions. We first take the median
 256 difference between two experimental conditions (i.e., the observed difference) to conduct the
 257 permutation tests. A null distribution is then constructed by combining SNR values from
 258 both experimental conditions, and randomly sampling them without replacement to create

two groups of sizes identical to their original, but with values not associated with a particular experimental condition. The median difference of the randomly sampled SNRs is then taken. This process is repeated multiple times (e.g., $N = 2000$ iterations) to build a distribution of median differences with no association of experimental condition (i.e., a null-hypothesis distribution). The observed median difference is then compared to this distribution. The proportion of scores greater than the observed difference represents the p value associated with the test. When comparing SNRs at the fundamental frequency (3Hz) for On/Off flicker, we find no statistically significant difference in median SNR magnitude between experimental conditions where stimuli were presented in temporal phase (see Figure 3). Monocular and binocular presentation for stimuli presented in temporal phase resulted in a similar response pattern under both On/Off and counterphase flicker modulations. This is consistent with ocularity invariance; binocular and monocular stimuli appear equal in magnitude at high contrast (Baker et al., 2007a; Meese et al., 2006).

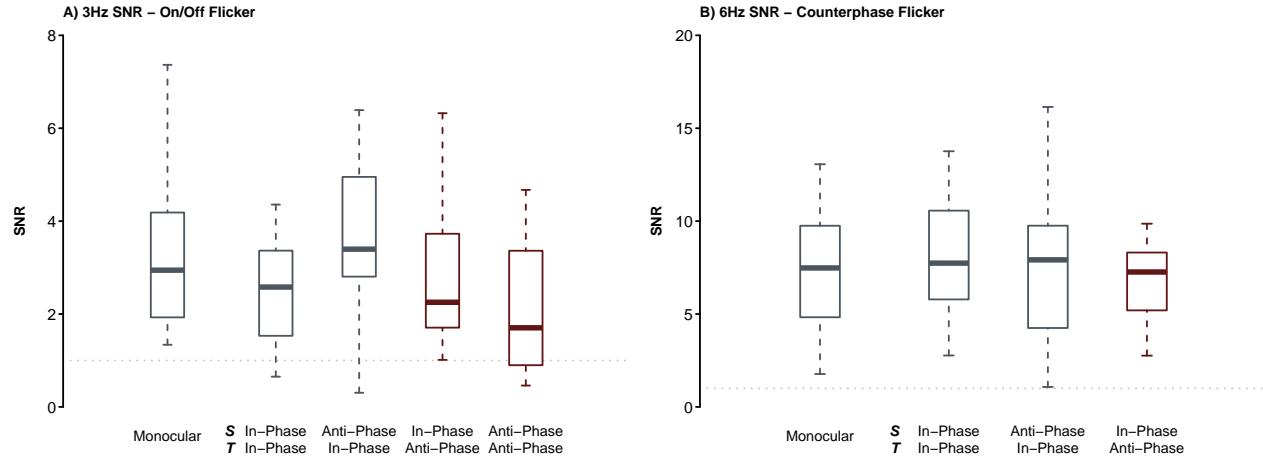


Figure 3: Boxplots of participant SNRs at the fundamental frequency (3Hz) under On/Off stimulation and its second harmonic for counterphase stimulation. A) Boxplots represent the distribution of participant SNRs at 3Hz. The median SNR is shown by the thick line within the box, with the lower and upper border of the box representing the first (25%) and third (75%) quartile of the distribution. The lower and upper whisker limits represent 1.5 times the interquartile range (distance between the third and first quartile). Boxplots for the binocular conditions have labels for their spatial (S) and temporal (T) phase relationships. Experimental conditions where stimuli are presented in temporal anti-phase are shown in maroon. B) As in A, boxplots show the distribution of participant SNRs at 6Hz. In both graphs, the dashed line represents an SNR of 1.0.

272 Changing the phase relationships of stimuli under On/Off flicker had some interesting
273 impacts on the fundamental frequency (Figure 3A). Stimuli presented in spatial phase and
274 temporal anti-phase generated smaller SNRs (median_{SNR} = 2.25) than stimuli presented in
275 spatial anti-phase and temporal phase (median_{SNR} = 3.4, $p = .047$). The reduction in the
276 amplitude relative to stimuli in spatial anti-phase and temporal phase was also observed
277 for stimuli in temporal and spatial anti-phase (median_{SNR} = 1.7; $p = .023$). No other
278 statistically significant difference in median SNRs were observed for all counterphase flicker
279 conditions or the other harmonics for On/Off flicker (all p values were greater than .05).
280 The median SNRs of temporal anti-phase stimuli under On/Off flicker, while smaller than
281 those of other conditions, were nevertheless statistically significantly greater than 1.0 (spatial
282 in-phase: $p < .001$ and spatial anti-phase: $p = .004$). The presence of a 3Hz response for
283 binocular stimuli presented in temporal anti-phase is a strong indication that monocular
284 responses remain and contribute to the SSVEP, as these conditions generate two transients
285 per cycle. A purely binocular signal would only generate responses at 6Hz (Blake et al., 1981;
286 Georgeson et al., 2016; Moulden, 1980).

287 *Modelling*

288 The perception of binocular stimulus contrast is well-explained by psychophysical mod-
289 els that process input contrast in two sequential contrast gain control stages interposed by
290 binocular summation (Baker et al., 2008, 2007a, 2007b; Baker and Meese, 2007; Meese et al.,
291 2006). This simple, yet powerful, family of models not only captures behavioural data well,
292 but can also explain neural responses to binocular and dichoptic stimuli (Baker and Wade,
293 2017; Lygo et al., 2021; Richard et al., 2018). Our SSVEP results show the expected pattern
294 of binocular combination for stimuli presented at high contrast (i.e., ocularity invariance)
295 but also intriguing effects that are likely explainable by the most recent extension of the
296 two-stage contrast gain control model, as defined by Georgeson et al. (2016). To explore
297 the architecture required to describe our effects adequately, we progressively increase the
298 complexity of binocular combination, beginning with a deliberately wrong model (i.e., lin-
299 ear combination) and building up to a multi-channel model with monocular, binocular, and
300 phase-selective pathways (Figure 4).

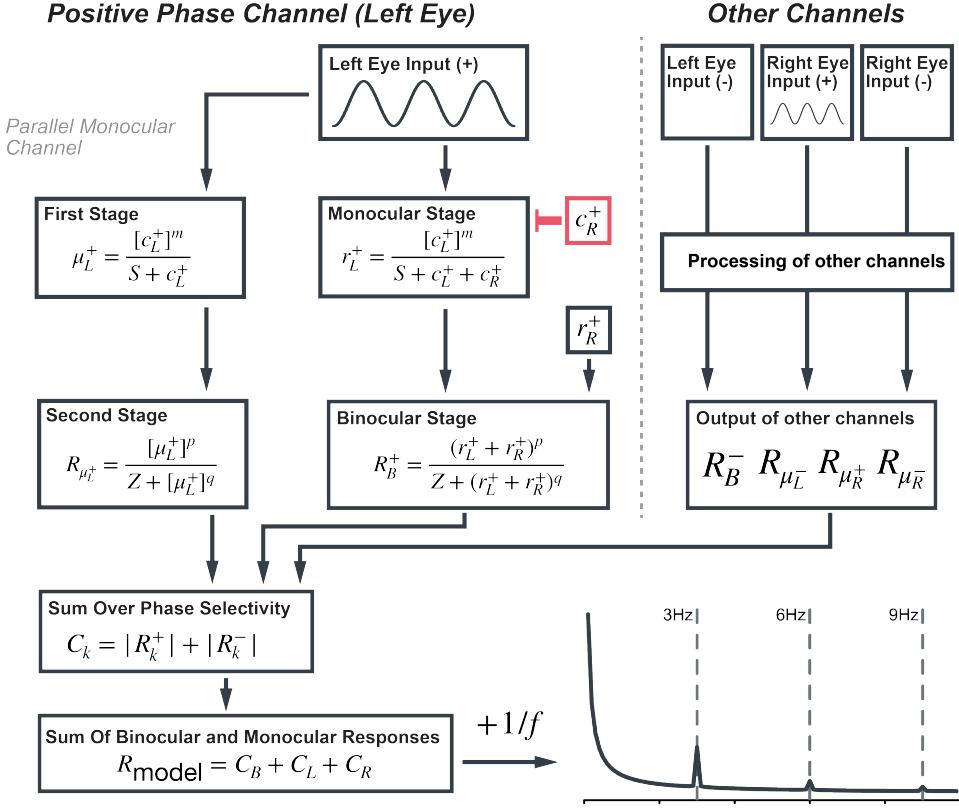


Figure 4: A simplified diagram of the two-stage contrast gain control model with parallel monocular and phase-selective channels. This diagram represents the processing stages of the positive phase channel in the left eye. These operations are identical for the other channels (negative phase left eye, positive and negative phase right eye) of the model. The sinusoidal input to the left eye (c_L^+) is fed through the first stage of the parallel monocular channel and the monocular stage of the binocular channel, which both apply a non-linearity to the input (m) and divisive inhibition (i.e., contrast gain control). The monocular stage of the binocular channel receives suppression from itself and the positive phase channel from the other eye. The output of these stages is fed into a second contrast gain control stage. It is at this stage that the monocular inputs are combined in the binocular channel. Finally, all responses are fast Fourier Transformed and their absolute values are summed over phase selectivity, followed by a sum over ocularity (binocular and monocular responses). This approach to defining the model's output differs from that of Georgeson (2016) to account for methodological differences. A pink noise spectrum was added to facilitate a comparable calculation of model SNRs, as is done with human data.

301 The architecture of the models explored differs, but all received the same input and had
 302 their final outputs processed identically. The input to all models was a 3 Hz sine wave, ad-
 303 justed to accurately represent the various experimental conditions of this study (see Figure 1).
 304 For example, stimuli presented with an On/Off flicker in temporal anti-phase had the left
 305 eye input generated by the following equation,

$$c_L = A * (\cos(2\pi ft) + 1)/2, \quad (4)$$

306 while the right eye input is defined as,

$$c_R = A * (-\cos(2\pi ft) + 1)/2. \quad (5)$$

307 A represents stimulus contrast (amplitude), f the temporal flicker frequency (i.e., 3 Hz), and
 308 t time in milliseconds. The input to the right eye (c_R) is phase shifted by 180° , which is
 309 accomplished using the negative cosine function ($-\cos$). Finally, sine waves are rectified to
 310 the range between 0 and 1 to represent the relative contrast presented to observers. The
 311 same experimental condition with counterphase flicker has the following sinusoidal profile for
 312 the left eye,

$$c_L = A * [\cos(2\pi ft)]_+ \quad (6)$$

313 and for the right eye,

$$c_R = A * [-\cos(2\pi ft)]_+. \quad (7)$$

314 These profiles are identical to the On/Off flicker, but the sine waves are half-wave rectified
 315 to represent the counterphase oscillation. To fit model outputs (rectified sine waves) to
 316 observer data, the final response of the models was Fast Fourier Transformed, and a pink
 317 noise spectrum was added to the Fourier amplitudes, $|FFT(R_{model})| + 1/f$ to simulate neural
 318 noise (see e.g. Donoghue et al., 2020), before calculating model SNRs. All models developed
 319 in this study were fit by minimizing the sum of squared errors between the model output and
 320 the observer median SNRs for the first 6 SSVEP components (3Hz, 6Hz, 9Hz, 12Hz, 15Hz,
 321 and 18Hz).

322 *Evidently wrong models*

323 As a first step in defining the necessary architecture to capture our results, we built baseline
 324 models with no monocular stage or phase selectivity that we do not expect to explain all of

325 our effects. The first is a purely linear summation model of binocular combination;

$$R_B = c_L + c_R, \quad (8)$$

326 the binocular response (R_B) is the sum of the inputs c_L and c_R . The fits of the linear
327 summation model are shown in Figure 5 and its performance metrics in Table 1. For On/Off
328 flicker, the linear summation model only generates responses at the fundamental frequency
329 (3Hz) that grossly overestimate observer SNRs. This is expected as this model lacks the
330 rectification and non-linearities required to generate responses at the harmonics (Regan and
331 Regan, 1988; Wade and Baker, 2025). In a linear sum, stimuli presented under On/Off flicker
332 in temporal anti-phase cancel each other, and thus the model generates no response. The
333 model does generate responses at the fundamental and harmonics of the counterphase flicker
334 condition (Figure 5B), but this is attributable to the input's rectification (the half-wave
335 rectification applied to the input; Equation 6, Equation 7) and not the model architecture.

Model	R^2	RMSE	AIC
Linear sum	-	3.234	281.99
Linear sum, with Rectification	0.575	1.405	197.95
Two-Stage, no interocular interactions	0.822	0.908	154.86
Two-Stage, with interocular interactions	0.81	0.94	158.62
Two-Stage with parallel monocular channels	0.877	0.756	135.1
Two-Stage with phase-selective channels & monocular channels	0.876	0.759	135.39

Table 1: Goodness-of-fit metrics for all models compared in this study. Errors in predictions for the linear sum model were too large to calculate R^2 . RMSE is the Root Mean Square error and AIC is the Akaike Information Criterion.

336 Responses of neurons to contrast in the visual system are well-modeled by a saturating
337 non-linearity: as contrast increases, the magnitude of responses saturates (the rise in response

per unit contrast decreases at higher contrast values (Heeger, 1992)). The saturating non-linearity can be modeled in different ways, but generally contains a divisive suppression and an exponentiation of the excitatory and inhibitory inputs. Including suppression can aid the model in better capturing the magnitude of responses in our observers, while exponents introduce the non-linearities required to generate responses at the harmonic frequencies. Thus, the next increment of model complexity defines the binocular response as the contrast gain control equation (after Legge, 1984b):

$$R_B = \frac{(c_L + c_R)^p}{Z + (c_L + c_R)^q}, \quad (9)$$

where the binocular response of the model (R_B) is defined as the sum of monocular inputs (c_L and c_R) raised to the power p , normalized by the sum of monocular inputs raised to q and where ($p > q$). The parameter Z prevents division by zero and sets overall sensitivity. The model can now generate responses at the harmonic frequencies for stimuli presented in temporal phase under On/Off flicker (see Figure 5 and Table 1). While this model iteration improves on the fits, it nevertheless struggles to fit SNR values at the fundamental frequency (3Hz) and is, as with the linear summation model, incapable of generating responses to stimuli presented with On/Off flicker in temporal anti-phase; the linear sum of stimuli presented in temporal anti-phase will always return zero. Therefore, the model still lacks the necessary architecture to define neural responses to our stimuli adequately.

355 *The Two-Stage Contrast Gain Control Model*

The simple models described above could not accurately represent the observer SSVEPs we recorded. They overestimated SNRs at the fundamental frequency and failed to generate responses for stimuli presented in temporal anti-phase with On/Off flicker. A potential model refinement is adding a monocular transducer before binocular combination (Meese et al., 2006). The architecture of this model now begins with a monocular stage

$$r_L = \frac{c_L^m}{S + c_L}, \quad r_R = \frac{c_R^m}{S + c_R} \quad (10)$$

where an exponent (m) is applied to the monocular inputs (c_L and c_R) in addition to self-suppression. The parameter S , as the parameter Z above, prevents devision by zero. The

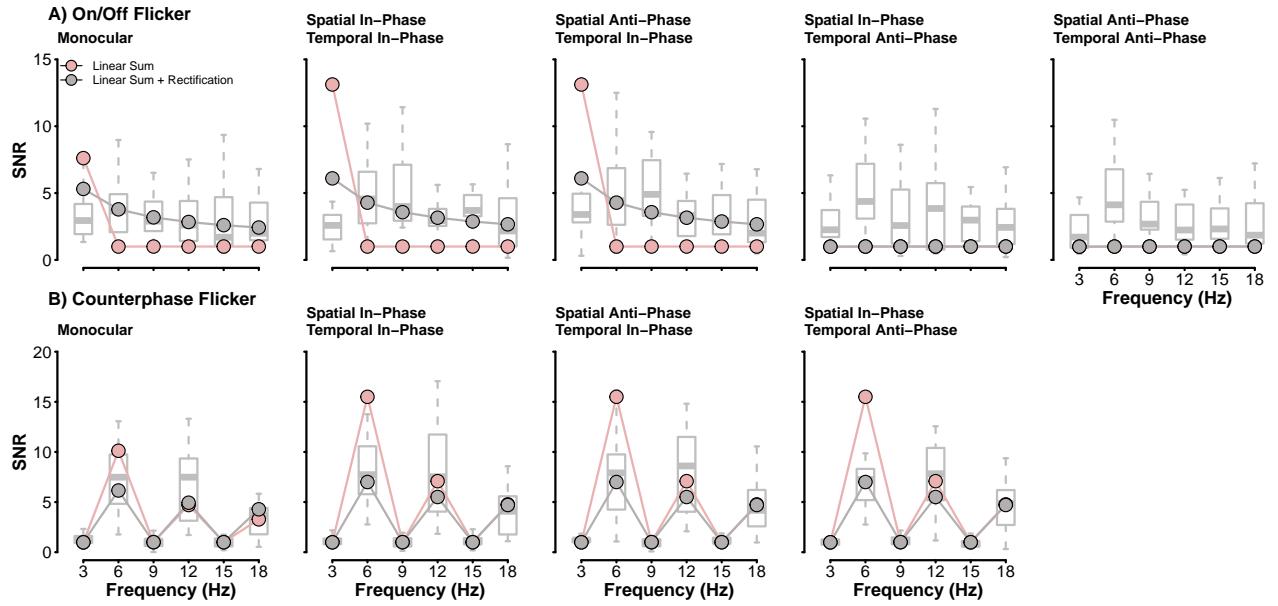


Figure 5: Fits of the linear sum (green) and the rectified linear sum (brown) models. Boxplots behind model responses show the distribution of observer SNRs. Model SNRs were fit to the median SNR of observers, which is represented by the thicker line within each box.

363 outputs of the monocular stage (r_L and r_R) are then fed into the binocular stage as defined
 364 in Equation 9,

$$R_B = \frac{(r_L + r_R)^p}{Z + (r_L + r_R)^q}. \quad (11)$$

365 In this model variant, m is the monocular exponent that determines the extent of summa-
 366 tion at detection threshold, with moderation by the suppressive term. In the second stage,
 367 $p > q$ as with Equation 9, which is necessary to capture the mild facilitatory effects of di-
 368 choptic masking (Meese et al., 2006), and determines the shape of contrast discrimination
 369 functions. We can strengthen the normalization of the monocular input by adding interocular
 370 suppression and replacing Equation 10 (the first stage) with:

$$r_L = \frac{c_L^m}{S + c_L + c_R}, \quad r_R = \frac{c_R^m}{S + c_R + c_L}. \quad (12)$$

371 This model iteration is identical to the two-stage contrast gain control model defined by
 372 Meese et al. (2006).

373 The fits of both model variants, with and without interocular suppression, are shown in
 374 Figure 6. The difference in their quality is negligible (see Table 1) as both describe most

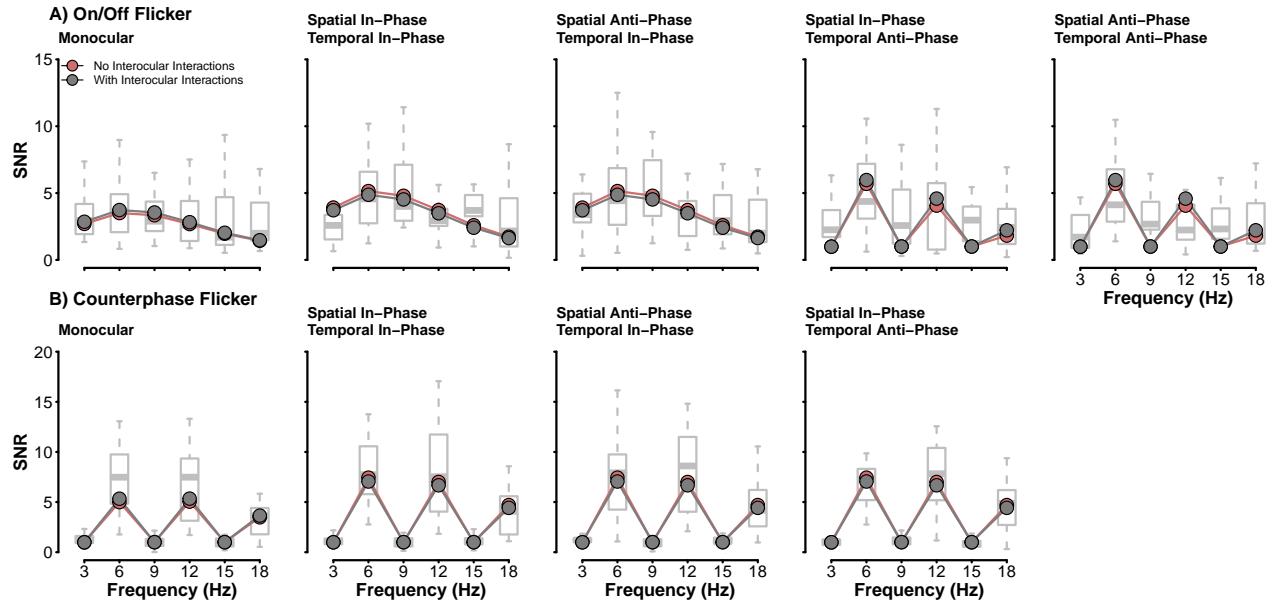


Figure 6: Fits of the two-stage contrast gain control model without (green) and with (red) interocular suppression. Boxplots behind model responses show the distribution of observer SNRs. Model SNRs were fit to the median SNR of observers, which is represented by the thicker line within each box.

experimental conditions well. The addition of the monocular transducer enables the model to fit observer SNRs at the fundamental frequency for stimuli presented in temporal phase under On/Off flicker, and importantly, it now generates responses for stimuli presented in temporal anti-phase. The transduced monocular inputs no longer cancel each other at the binocular stage. While the model can create responses to temporal anti-phase stimuli, it only does so at the even harmonics (2F-6Hz, 4F-12Hz, and 6F-18Hz) of the SSVEP spectrum. This is expected as the model can only generate binocular responses; it does not preserve monocular signals beyond the first stage. As the two rectified sine waves are in anti-phase, their sum will generate a new waveform with frequencies twice the original (6Hz) frequency and its integer harmonics (2F-12Hz and 3F-18Hz). The responses we recorded at the fundamental frequency (3Hz) and its odd integer harmonics (3F - 9Hz, 5F - 15Hz) cannot be explained by an architecture with a purely binocular output. Next, we explore methods of preserving the monocular signal to explain observer responses to stimuli presented in temporal anti-phase.

388 *Parallel Monocular and Phase-Selective Channels*

389 Based on the models described above, observer SSVEPs for stimuli presented in temporal
 390 anti-phase imply the presence of both a binocular and monocular response in the population
 391 response. To preserve the monocular response in the modelling, we add parallel monocular
 392 channels to the two-stage contrast gain control model, similar to Georgeson et al. (2016).
 393 These channels operate in an identical manner to the traditional two-stage contrast gain
 394 control model channel but are fully monocular (i.e., they do not include interocular
 395 suppression),

$$\mu_L = \frac{c_L^m}{S + c_L}, \quad \mu_R = \frac{c_R^m}{S + c_R}. \quad (13)$$

396 These equations, identical to those presented in Equation 10, return μ_L , the output of the
 397 left eye's first stage of the monocular channel, and μ_R is that of the right eye. The output
 398 of the monocular channels undergoes a second contrast gain control stage similar to that of
 399 the binocular channel,

$$R_{\mu_L} = \frac{\mu_L^p}{Z + \mu_L^q}, \quad R_{\mu_R} = \frac{\mu_R^p}{Z + \mu_R^q}. \quad (14)$$

400 R_{μ_L} and R_{μ_R} represent the final responses of the left and right monocular channels. No
 401 additional free parameters are required to define the parallel monocular channels; m , p , q , S ,
 402 and Z are used to define the monocular channel and binocular channel responses.

403 The inclusion of parallel monocular channels poses an interesting problem in our modelling
 404 as we now contend with three visual cues from which to generate model SSVEPs: monocular
 405 left (R_L), monocular right (R_R), and binocular (R_B). Psychophysically, cue selection has
 406 been implemented as a Minkowski sum with a large (≈ 30) exponent (Georgeson et al., 2016),
 407 approximating a MAX rule. This method is inappropriate when modelling neural data, as
 408 the recordings from the scalp represent an amalgamation of all responses generated by the
 409 stimuli (Wade and Baker, 2025). Our model output must therefore represent the combination
 410 of signals instead of the selection of the strongest signal. Here, it is implemented as the
 411 sum of the Fourier amplitude spectra of all three channel outputs. This sum preserves the
 412 amplitude responses required to generate SSVEPs and prevents the nullifying of responses
 413 from summing signals in anti-phase.

414 Preserving monocular responses until the output stage with parallel monocular channels

improved the fit to our data (see Table 1 and Figure 7). The model now captures responses at the fundamental (3Hz) and odd-integer harmonics (9Hz and 15Hz) of the temporal anti-phase conditions under On/Off flicker. Monocular signals, while weaker than binocular signals, are preserved in the neural responses of observers and thus must be accounted for in their computational description. As proposed by Georgeson et al. (2016), parallel monocular channels appear to be an adequate descriptor of monocular signals.

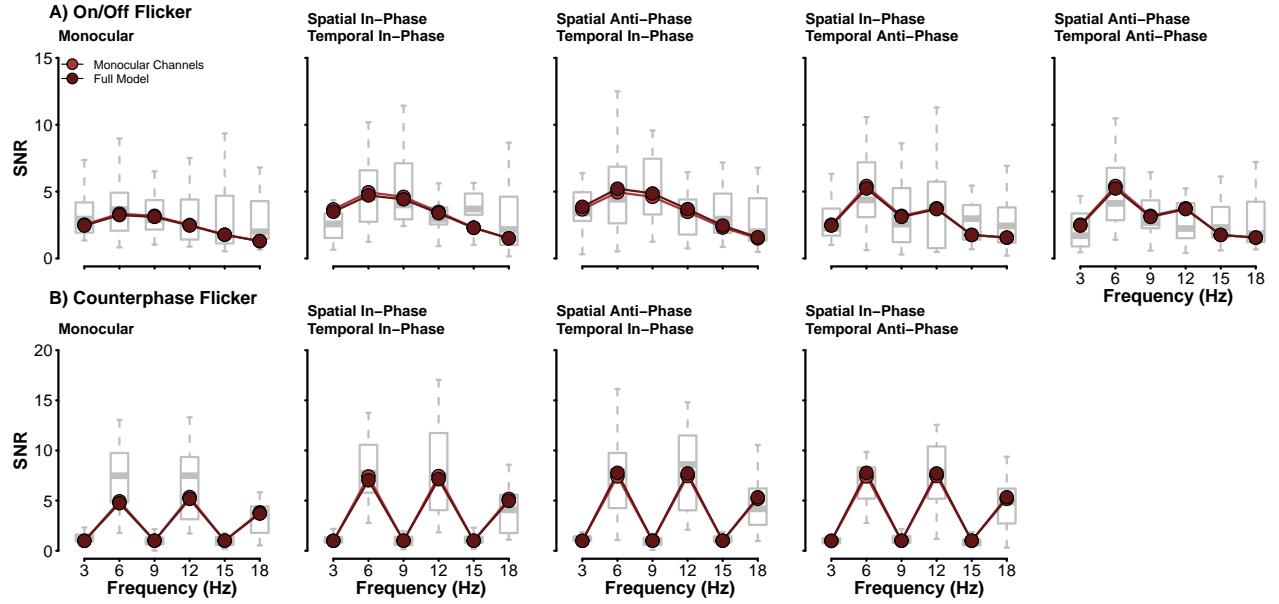


Figure 7: Fits of the two best performing models to our observer data. Both models, that including parallel monocular channels and the full model with the added phase-selective channels perform quite similarly.

Psychophysically, spatial phase has a meaningful impact on binocular combination (Bacon, 1976; Baker and Meese, 2007; Simmons, 2005) and the two-stage contrast gain control often includes phase-selective channels to account for these effects. Our study found little evidence of any influence of spatial phase on observer SSVEPs, as no statistically significant difference in signal-to-noise ratios across spatial phase was found. Still, we felt it prudent to verify if adding phase-selective channels improves the fit of the model to our data. Phase-selectivity was added to the binocular and monocular channels by replicating the equations of the first and second stage, for six channels (see Figure 4). As with the previous model, no additional model parameters were required to include phase-selectivity; the parameters m , p , q , S , and

430 Z were used to define the responses of all six channels in this model.

431 Model responses were generated from sine waves (in temporal phase or anti-phase) fed
432 into their respective positive or negative phase channels. This simulated the spatial phase of
433 stimuli presented to observers. We refer to this model iteration as the full model, as it includes
434 a binocular channel with a monocular stage (with interocular interactions), a binocular stage,
435 parallel monocular channels (with their respective stages), and phase-selectivity. This is, in
436 essence, the 6-channel 2-cue model proposed by Georgeson et al. (2016) to explain binocular
437 combination under multiple experimental conditions. As with the addition of monocular
438 channels, including phase-selectivity to the model means we now have six final signals to
439 contend with: a binocular and two monocular channels selective for positive spatial phase,
440 and a binocular and two monocular channels selective for negative spatial phase. We used the
441 same method of combining signals as above: we first summed the Fourier amplitude spectra
442 of observers across the positive and negative phase-selective channels and then combined
443 the binocular and monocular channels. The inclusion of phase-selectivity did not improve
444 the model's fit. The R^2 value calculated across the six frequencies and nine experimental
445 conditions was no different than that of the previous model (see Table 1), indicating that
446 the spatial phase of stimuli does not have a meaningful impact on the SSVEP amplitudes of
447 observers.

448 Discussion

449 The computational architecture of binocular combination, under the two-stage contrast
450 gain control model framework, has been carefully evaluated on psychophysical data (Baker
451 and Meese, 2007; Georgeson et al., 2016; Meese et al., 2006), yet its ability to explain neural
452 data has only been explored on a limited set of stimulus conditions (Baker and Wade, 2017;
453 Lygo et al., 2021; Richard et al., 2018). This study investigated the effects of stimulus spatial
454 and temporal phase on observer SSVEPs. We explored the ability of the two-stage contrast
455 gain control model - and its variants - to capture our neural data. On/Off flicker generated
456 responses at the fundamental frequency (3Hz) and its harmonics (6Hz, 9Hz, 12Hz, and 15Hz)
457 while counterphase flicker generated responses at twice the fundamental frequency (6Hz) and
458 its even integer harmonics (12Hz and 18Hz). No statistically significant difference was found

459 between the monocular and binocular conditions presented with On/Off or counterphase
460 flicker. This is consistent with ocularity invariance; binocular and monocular stimuli appear
461 equal in magnitude at high contrast (Baker et al., 2007a; Meese et al., 2006). We additionally
462 found no statistically significant differences in SSVEPs for all stimulus conditions under
463 counterphase flicker.

464 Our data proved useful in selecting the best-performing model, albeit in an unexpected way.
465 The stimulus presentation protocols (On/Off and counterphase) individually were unable to
466 discern the best model, as all models, except the evidently wrong ones, were able to capture
467 neural responses to counterphase stimuli. Instead, it was the spatial and temporal phase of
468 our stimuli that proved most useful in model selection. Presenting stimuli in temporal anti-
469 phase reduced the magnitude of responses at the fundamental frequency (3Hz) compared to
470 its temporal phase counterpart, but they were not abolished. This finding indicates that,
471 even under binocular viewing, monocular signals can be measured at the scalp with SSVEPs.
472 This was crucial in the selection of the best fitting model. Only the two-stage contrast
473 gain control model with parallel monocular channels, which preserves monocular signals
474 throughout the model architecture, could accurately capture the data in all experimental
475 conditions. Simpler models failed to generate the necessary responses at the fundamental
476 and odd-integer harmonics for stimuli presented in temporal anti-phase. In contrast, a more
477 complex model that included phase selectivity did not improve the quality of fits. Overall, we
478 find that the same general framework used to explain psychophysical stimulus combination
479 can be successfully applied to neural data collected under various experimental conditions.

480 *Monocular channels.*

481 It has long been assumed by most computational descriptions of binocular combination
482 that only the binocular response was available to later stages of perception and decision
483 (Baker and Meese, 2007; Ding et al., 2013; Ding and Sperling, 2006; Legge, 1984a; Meese et
484 al., 2006). The monocular pathways, which represent the early stages of the visual processing
485 stream, only serve as the input to combine. However, psychophysical evidence from adapta-
486 tion and discrimination experiments has demonstrated that monocular signals do contribute
487 to perception (Blake et al., 1981; Georgeson et al., 2016; Moulden, 1980) and that monocular

488 channels parallel to the binocular channel should be included in models of binocular combi-
489 nation. This study found that monocular responses to stimuli can be recorded with SSVEPs
490 under binocular stimulation. Presenting stimuli in temporal anti-phase under On/Off flicker
491 generates two transients per cycle, and a binocular system, which combines the monocular
492 inputs of each eye, would only create responses at 6Hz. While the 6Hz component was sig-
493 nificant, we still found SSVEPs at 3Hz that exceeded the noise level in our data. These 3Hz
494 components represent the individual oscillations for each eye and are therefore likely represen-
495 tative of a monocular response. The reduction in response magnitude of the 3Hz component
496 we observed from the temporal in-phase to the temporal anti-phase conditions can thus be
497 explained as the transition from a binocular to an inadvertently weaker monocular response.

498 Computationally, we demonstrated that many mechanisms of binocular combination, in-
499 cluding monocular non-linearities, interocular interactions, and parallel monocular channels,
500 were required to explain the SNR spectra of our observers. Parallel monocular channels were
501 critical in capturing observer data in the two temporal anti-phase conditions presented with
502 On/Off flicker. Without adding these channels, the models could not generate a response at
503 the fundamental (3Hz) or its odd integer harmonics. While an essential inclusion to explain
504 our data, the addition of monocular channels did not meaningfully alter the ability of the
505 two-stage contrast gain control model to capture observer SNRs in the other experimental
506 conditions. Monocular contributions to the model output are most significant for percep-
507 tion when the stimulus contrast between the two eyes differs (Georgeson et al., 2016), an
508 experimental scenario we did not explore here.

509 *SSVEP responses to spatial phase.*

510 Our results did not suggest an impact of stimulus spatial phase on observer SSVEPs. It is
511 well-known that spatial phase affects binocular combination when measured psychophysically
512 (Bacon, 1976; Baker and Meese, 2007; Simmons, 2005). The binocular combination of two
513 opposite-polarity stimuli does not cancel; it sums, thus influencing sensitivity to the stimuli.
514 It is therefore odd to find little to no influence of spatial phase on the SSVEPs of observers,
515 particularly as SSVEP amplitude is associated with perceptual sensitivity (Bosse et al., 2018;
516 Campbell and Kulikowski, 1972; Campbell and Maffei, 1970; Norcia et al., 2015). However,

517 we note that at high contrasts (as used here), perceived contrast is equivalent between in-
518 phase and antiphase stimuli (Baker et al., 2012). The absence of an apparent spatial phase
519 effect also complicates our ability to investigate any contributions of a differencing channel
520 in our study (Chen and Li, 1998; May et al., 2012; May and Zhaoping, 2016).

521 Spatial phase-dependent effects in SSVEPs have been recorded with motion stimuli (Cot-
522 tereau et al., 2014; Kohler et al., 2018). The dichoptic presentation of moving bars, whereby
523 they either move in-phase (lateral motion) or in anti-phase (motion-in-depth), impacts the
524 amplitude of SSVEPs. In-phase motion generates SSVEP amplitudes that are twice those of
525 motion in anti-phase (Cottetrau et al., 2014). A different response can also be recorded for
526 dichoptic stimuli when frequency tagged (Katyal et al., 2018; Sutoyo and Srinivasan, 2009),
527 where the “conflict” response is represented at the intermodulation frequencies. All stimuli
528 in this study were presented at 3Hz; we did not frequency tag stimuli presented to the left
529 and right eyes. Thus, the responses we recorded at the scalp are a spatial aggregation of the
530 stimuli presented to observers (Wade and Baker, 2025). Spatially aggregated responses will
531 be identical regardless of stimulus phase and thus generate phase-insensitive responses.

532 Conclusion

533 We investigated the effects of stimulus spatial and temporal phase on observer SSVEPs
534 and explored the necessary computational components to explain our data. We worked un-
535 der the framework of the two-stage contrast gain control model of binocular vision (Meese et
536 al., 2006). We examined the impact of interocular interactions, parallel monocular channels,
537 and phase-selective channels on the model’s fit with our data. The most significant effect
538 to capture was the presence of odd-integer harmonics in the temporal anti-phase conditions,
539 representing the monocular response to stimuli. This was well-explained by adding parallel
540 monocular channels to the model. The two-stage contrast gain control model of binocu-
541 lar combination remains a robust descriptor of binocular vision as it can explain various
542 experimental conditions and modalities (psychophysics and neuroimaging).

543 **Data and Code Availability**

544 We provide the raw (.cnt) and processed (.RData) observer SSVEPs, in addition to all
545 code used in this study, on the Open Science Framework, which can be accessed using the
546 following project link: <https://osf.io/cn894/>. A computationally reproducible version of this
547 manuscript is available at the linked GitHub repository: <https://github.com/brunoRichard/>
548 [SSVEP_Phase_AntiPhase](#).

549 **References**

- 550 Bacon, J.H., 1976. The interaction of dichoptically presented spatial gratings. *Vision Res*
551 16, 337–44. [https://doi.org/10.1016/0042-6989\(76\)90193-0](https://doi.org/10.1016/0042-6989(76)90193-0)
- 552 Baker, D.H., Lygo, F.A., Meese, T.S., Georgeson, M.A., 2018. Binocular summation revisited:
553 Beyond $\sqrt{2}$. *Psychol Bull* 144, 1186–1199. <https://doi.org/10.1037/bul0000163>
- 554 Baker, D.H., Meese, T.S., 2007. Binocular contrast interactions: Dichoptic masking is not a
555 single process. *Vision Res* 47, 3096–107. <https://doi.org/10.1016/j.visres.2007.08.013>
- 556 Baker, D.H., Meese, T.S., Georgeson, M.A., 2007a. Binocular interaction: Contrast matching
557 and contrast discrimination are predicted by the same model. *Spat Vis* 20, 397–413.
558 <https://doi.org/10.1163/156856807781503622>
- 559 Baker, D.H., Meese, T.S., Hess, R.F., 2008. Contrast masking in strabismic amblyopia:
560 Attenuation, noise, interocular suppression and binocular summation. *Vision Res* 48,
561 1625–40. <https://doi.org/10.1016/j.visres.2008.04.017>
- 562 Baker, D.H., Meese, T.S., Summers, R.J., 2007b. Psychophysical evidence for two routes to
563 suppression before binocular summation of signals in human vision. *Neuroscience* 146,
564 435–448. <https://doi.org/10.1016/j.neuroscience.2007.01.030>
- 565 Baker, D.H., Wade, A.R., 2017. Evidence for an optimal algorithm underlying signal combi-
566 nation in human visual cortex. *Cereb Cortex* 27, 254–264. <https://doi.org/10.1093/cercor/bhw395>
- 568 Baker, D.H., Wallis, S.A., Georgeson, M.A., Meese, T.S., 2012. The effect of interocular
569 phase difference on perceived contrast. *PLoS ONE* 7, e34696. <https://doi.org/10.1371/journal.pone.0034696>

- 571 Blake, R., 1989. A neural theory of binocular rivalry. *Psychol Rev* 96, 145–67. <https://doi.org/10.1037/0033-295x.96.1.145>
- 572
- 573 Blake, R., Overton, R., Lema-Stern, S., 1981. Interocular transfer of visual aftereffects. *J Exp Psychol Hum Percept Perform* 7, 367–81. <https://doi.org/10.1037/0096-1523.7.2.367>
- 574
- 575 Blake, R., Wilson, H., 2011. Binocular vision. *Vision Research* 51, 754–770. <https://doi.org/10.1016/j.visres.2010.10.009>
- 576
- 577 Bosse, S., Acqualagna, L., Samek, W., Porbadnigk, A.K., Curio, G., Blankertz, B., Müller, K.-R., Wiegand, T., 2018. Assessing perceived image quality using steady-state visual evoked potentials and spatio-spectral decomposition. *IEEE Transactions on Circuits and Systems for Video Technology* 28, 1694–1706. <https://doi.org/10.1109/TCSVT.2017.2694807>
- 578
- 579
- 580
- 581 Brainard, D.H., 1997. The psychophysics toolbox. *Spatial Vision* 10, 433–436. <https://doi.org/10.1163/156856897X00357>
- 582
- 583 Campbell, F.W., Green, D.G., 1965. Monocular versus binocular visual acuity. *Nature* 208, 191–2. <https://doi.org/10.1038/208191a0>
- 584
- 585 Campbell, F.W., Kulikowski, J.J., 1972. The visual evoked potential as a function of contrast of a grating pattern. *J Physiol* 222, 345–56. <https://doi.org/10.1113/jphysiol.1972.sp009801>
- 586
- 587
- 588 Campbell, F.W., Maffei, L., 1970. Electrophysiological evidence for the existence of orientation and size detectors in the human visual system. *J Physiol* 207, 635–52. <https://doi.org/10.1113/jphysiol.1970.sp009085>
- 589
- 590
- 591 Chatrian, G.E., Lettich, E., Nelson, P.L., 1985. Ten percent electrode system for topographic studies of spontaneous and evoked EEG activities. *American Journal of EEG Technology* 25, 83–92. <https://doi.org/10.1080/00029238.1985.11080163>
- 592
- 593
- 594 Chen, D., Li, Z., 1998. A psychophysical experiment to test the efficient stereo coding theory, in: Wong, K.-Y.M., King, I., Yeung, D.Y. (Eds.), *Theoretical Aspects of Neural Computation: A Multidisciplinary Perspective*. Springer Verlag, New York, pp. 225–235.
- 595
- 596
- 597 Cottereau, B.R., McKee, S.P., Norcia, A.M., 2014. Dynamics and cortical distribution of neural responses to 2D and 3D motion in human. *Journal of Neurophysiology* 111, 533–543. <https://doi.org/10.1152/jn.00549.2013>
- 598
- 599
- 600 Delorme, A., Makeig, S., 2004. EEGLAB: An open source toolbox for analysis of single-

- 601 trial EEG dynamics including independent component analysis. Journal of Neuroscience
602 Methods 134, 9–21. <https://doi.org/10.1016/J.JNEUMETH.2003.10.009>
- 603 Ding, J., Klein, S.A., Levi, D.M., 2013. Binocular combination of phase and contrast ex-
604 plained by a gain-control and gain-enhancement model. J Vis 13, 13. <https://doi.org/10.1167/13.2.13>
- 605 Ding, J., Sperling, G., 2006. A gain-control theory of binocular combination. Proc Natl Acad
606 Sci U S A 103, 1141–6. <https://doi.org/10.1073/pnas.0509629103>
- 607 Donoghue, T., Haller, M., Peterson, E.J., Varma, P., Sebastian, P., Gao, R., Noto, T., Lara,
608 A.H., Wallis, J.D., Knight, R.T., Shestyuk, A., Voytek, B., 2020. Parameterizing neural
609 power spectra into periodic and aperiodic components. Nat Neurosci 23, 1655–1665.
610 <https://doi.org/10.1038/s41593-020-00744-x>
- 611 Georgeson, M.A., Wallis, S.A., Meese, T.S., Baker, D.H., 2016. Contrast and lustre: A
612 model that accounts for eleven different forms of contrast discrimination in binocular
613 vision. Vision Research 129, 98–118. <https://doi.org/10.1016/j.visres.2016.08.001>
- 614 Hansen, B.C., Hess, R.F., 2006. Discrimination of amplitude spectrum slope in the fovea
615 and parafovea and the local amplitude distributions of natural scene imagery. J Vis 6,
616 696–711. <https://doi.org/10.1167/6.7.3>
- 617 Heeger, D.J., 1992. Normalization of cell responses in cat striate cortex. Vis Neurosci 9,
618 181–97. <https://doi.org/10.1017/s0952523800009640>
- 619 Katyal, S., Vergeer, M., He, S., He, B., Engel, S.A., 2018. Conflict-sensitive neurons gate
620 interocular suppression in human visual cortex. Scientific Reports 8, 1239. <https://doi.org/10.1038/s41598-018-19809-w>
- 621 Kim, Y.J., Grabowecky, M., Paller, K.A., Suzuki, S., 2011. Differential roles of frequency-
622 following and frequency-doubling visual responses revealed by evoked neural harmonics.
623 J Cogn Neurosci 23, 1875–86. <https://doi.org/10.1162/jocn.2010.21536>
- 624 Kleiner, M., Brainard, D.H., Pelli, D.G., 2007. What's new in psychtoolbox-3?, in: Percep-
625 tion.
- 626 Kohler, P.J., Meredith, W.J., Norcia, A.M., 2018. Revisiting the functional significance of
627 binocular cues for perceiving motion-in-depth. Nature Communications 9, 3511. <https://doi.org/10.1038/s41467-018-05918-7>
- 628
- 629
- 630

- 631 Legge, G.E., 1984b. Binocular contrast summation-II. Quadratic summation. Vision Res 24,
632 385–94. [https://doi.org/10.1016/0042-6989\(84\)90064-6](https://doi.org/10.1016/0042-6989(84)90064-6)
- 633 Legge, G.E., 1984a. Binocular contrast summation-i. Detection and discrimination. Vision
634 Res 24, 373–83. [https://doi.org/10.1016/0042-6989\(84\)90063-4](https://doi.org/10.1016/0042-6989(84)90063-4)
- 635 Li, Z., Atick, J.J., 1994. Efficient stereo coding in the multiscale representation. Network:
636 Computation in Neural Systems 5, 157â 174. https://doi.org/10.1088/0954-898x_5_2_003
- 638 Lygo, F.A., Richard, B., Wade, A.R., Morland, A.B., Baker, D.H., 2021. Neural markers of
639 suppression in impaired binocular vision. Neuroimage 230, 117780. <https://doi.org/10.1016/j.neuroimage.2021.117780>
- 641 Maehara, G., Goryo, K., 2005. Binocular, monocular and dichoptic pattern masking. Optical
642 Review 12, 76–82. <https://doi.org/DOI 10.1007/s10043-004-0076-5>
- 643 May, K.A., Zhaoping, L., 2016. Efficient coding theory predicts a tilt aftereffect from viewing
644 untilted patterns. Curr Biol 26, 1571–1576. <https://doi.org/10.1016/j.cub.2016.04.037>
- 645 May, K.A., Zhaoping, L., Hibbard, P.B., 2012. Perceived direction of motion determined by
646 adaptation to static binocular images. Curr Biol 22, 28–32. <https://doi.org/10.1016/j.cu b.2011.11.025>
- 648 Meese, T.S., Baker, D.H., 2011. Contrast summation across eyes and space is revealed
649 along the entire dipper function by a “swiss cheese” stimulus. J Vis 11, 1–23. <https://doi.org/10.1167/11.1.23>
- 651 Meese, T.S., Georgeson, M.A., Baker, D.H., 2006. Binocular contrast vision at and above
652 threshold. Journal of vision 6, 1224–1243. <https://doi.org/10.1167/6.11.7>
- 653 Moradi, F., Heeger, D.J., 2009. Inter-ocular contrast normalization in human visual cortex.
654 J Vis 9, 13 1–22. <https://doi.org/10.1167/9.3.13>
- 655 Moulden, B., 1980. After-effects and the integration of patterns of neural activity within a
656 channel. Philos Trans R Soc Lond B Biol Sci 290, 39–55. <https://doi.org/10.1098/rstb.1 980.0081>
- 658 Norcia, A.M., Appelbaum, L.G., Ales, J.M., Cottetereau, B.R., Rossion, B., 2015. The steady-
659 state visual evoked potential in vision research: A review. Journal of Vision 15, 4. <https://doi.org/10.1167/15.6.4>

- 661 Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: Transforming num-
662 bers into movies. *Spatial Vision* 10, 437–442. <https://doi.org/10.1163/156856897X00366>
- 663 Regan, M., Regan, D., 1988. A frequency domain technique for characterizing nonlinearities
664 in biological systems. *Journal of theoretical biology* 133, 293–317.
- 665 Richard, B., Chadnova, E., Baker, D.H., 2018. Binocular vision adaptively suppresses delayed
666 monocular signals. *Neuroimage* 172, 753–765. <https://doi.org/10.1016/j.neuroimage.201>
667 8.02.021
- 668 Simmons, D.R., 2005. The binocular combination of chromatic contrast. *Perception* 34,
669 1035–1042. <https://doi.org/10.1088/p5279>
- 670 Simmons, D.R., Kingdom, F.A.A., 1998. On the binocular summation of chromatic contrast.
671 *Vision Research* 38, 1063–1071. [https://doi.org/10.1016/S0042-6989\(97\)00272-1](https://doi.org/10.1016/S0042-6989(97)00272-1)
- 672 Skottun, B.C., Skoyles, J.R., 2007. Some remarks on the use of visually evoked poten-
673 tials to measure magnocellular activity. *Clinical Neurophysiology* 118, 1903–1905.
674 <https://doi.org/https://doi.org/10.1016/j.clinph.2007.06.007>
- 675 Sutoyo, D., Srinivasan, R., 2009. Nonlinear SSVEP responses are sensitive to the perceptual
676 binding of visual hemifields during conventional “eye” rivalry and interocular “percept”
677 rivalry. *Brain Research* 1251, 245–255. <https://doi.org/https://doi.org/10.1016/j.brainr>
678 es.2008.09.086
- 679 Tadmor, Y., Tolhurst, D.J., 1994. Discrimination of changes in the second-order statistics of
680 natural and synthetic images. *Vision Research* 34, 541–554. <https://doi.org/10.1016/00>
681 42-6989(94)90167-8
- 682 Wade, A.R., Baker, D.H., 2025. Measuring contrast processing in the visual system using
683 the steady state visually evoked potential (SSVEP). *Vision Research* 231, 108614. <https://doi.org/10.1016/j.visres.2025.108614>
- 685 Wendt, G., Faul, F., 2022. A simple model of binocular luster. *J Vis* 22, 6. <https://doi.org/10.1167/jov.22.10.6>
- 687 Wilson, H.R., 2003. Computational evidence for a rivalry hierarchy in vision. *Proceedings*
688 of the National Academy of Sciences of the United States of America 100, 14499–14503.
689 <https://doi.org/10.1073/pnas.2333622100>