Short-latency ocular-following responses: Weighted nonlinear summation predicts the outcome of a competition between two sine wave gratings moving in opposite directions

Boris M. Sheliga Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, MD, USA

Christian Quaia Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, MD, USA

Edmond J. FitzGibbon Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, MD, USA

Bruce G. Cumming Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, MD, USA

We recorded horizontal ocular-following responses to pairs of superimposed vertical sine wave gratings moving in opposite directions in human subjects. This configuration elicits a nonlinear interaction: when the relative contrast of the gratings is changed, the response transitions abruptly between the responses elicited by either grating alone. We explore this interaction in pairs of gratings that differ in spatial and temporal frequency and show that all cases can be described as a weighted sum of the responses to each grating presented alone, where the weights are a nonlinear function of stimulus contrast: a nonlinear weighed summation model. The weights depended on the spatial and temporal frequency of the component grating. In many cases the dominant component was not the one that produced the strongest response when presented alone, implying that the neuronal circuits assigning weights precede the stages at which motor responses to visual motion are generated. When the stimulus area was reduced, the relationship between spatial frequency and weight shifted to higher frequencies. This finding may reflect a contribution from surround suppression. The nonlinear interaction is strongest when the two components have similar spatial frequencies, suggesting that the nonlinearity may reflect interactions within single spatial frequency channels. This framework can be extended to stimuli composed of more than two components: our model was able to predict the responses to stimuli composed of three gratings. That this relatively simple model successfully captures the ocular-following responses over a wide range of spatial/temporal frequency and contrast parameters suggests that these interactions reflect a simple mechanism.

Introduction

Ocular-following responses (OFRs) are short-latency tracking eye movements evoked by the motion of a visual stimulus (Gellman, Carl, & Miles, 1990; Miles, Kawano, & Optican, 1986). The characteristics of the OFR (e.g., amplitude, latency) seem to be closely linked with many aspects of neuronal activity at early stages of visual processing (V1, MT, MST; for review, see Masson & Perrinet, 2012; Miles, 1998; Miles & Sheliga, 2010). This offers the hope that a relatively simple model might describe the visual processing that drives OFRs. However, even these reflexive movements cannot be explained as a linear summation over spatiotemporal channels: when two gratings move in opposite directions there is a "winner-take-all"–like behavior, in which small changes in contrast lead to rapid changes in response towards that produced by the dominant

Citation: Sheliga, B. M., Quaia, C., FitzGibbon, E. J., & Cumming, B. G. (2020). Short-latency ocular-following responses: Weighted nonlinear summation predicts the outcome of a competition between two sine wave gratings moving in opposite directions. *Journal of Vision*, *20*(1):1, 1–23, https://doi.org/10.1167/jov.20.1.1.

https://doi.org/10.1167/jov.20.1.1 Received July 10, 2019; published January 29, 2020 ISSN 1534-7362

component alone (Matsuura et al., 2008; Sheliga, FitzGibbon, & Miles, 2008; Sheliga, Kodaka, FitzGibbon, & Miles, 2006). This effect may then be related to similar effects that has been seen in cortical neurons (areas MT/MST; Kumbhani, Saber, Majaj, Tailby, & Movshon, 2008; Miura, Inaba, Aoki, & Kawano, 2014). Here we explore how this behavior depends on the properties of the stimulus in an effort to find a simple rule that describes such interactions over a broad range of spatiotemporal inputs. We hope it will eventually provide a route to predicting OFRs to arbitrary one-dimensional stimuli.

When these highly nonlinear interactions were first observed, it was assumed that the dominant spatial component was the one that produced the strongest response when presented alone. However, one study hinted at a possibility that certain spatial frequencies (SFs) had more weight than others: two equal-contrast gratings moving in opposite directions resulted in eye movements in the direction of motion of the higher SF grating (Sheliga et al., 2006).

Here we describe a series of experiments in which we used vertical sine wave gratings (hereafter "gratings") having a wide range of SFs and temporal frequencies (TFs). In Experiment 1, two gratings differed in contrast and SF, but not TF. Contrast, SF, and TF were manipulated in Experiment 2. Experiment 3 replicated the conditions of Experiment 1, except for a four-fold decrease in an overall stimulus size. Experiment 4 used a stimulus composed of three moving gratings, to test whether the rules revealed in Experiments 1–3 still held. In brief, we found the weight given to a single component depended on its SF and TF, but this dependence was different from that seen in the OFRs to single gratings (i.e., the weight did not simply reflect response amplitude). This finding suggests that the interactions preceded the stage at which motor responses to visual motion were generated.

Preliminary results of this study were presented in abstract form elsewhere (Sheliga, Quaia, FitzGibbon, & Cumming, 2018).

Material and methods

Many of the techniques are described only briefly, because they are similar to those used in this laboratory in the past (e.g., Sheliga, Chen, FitzGibbon, & Miles, 2005). Experimental protocols were approved by the Institutional Review Committee concerned with the use of human subjects. Our research was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki), and informed consent was obtained for experimentation with human subjects.

Subjects

Three subjects took part in this study: two were authors (BMS and EJF) and the third was a paid volunteer (TH) naïve as to the purpose of the experiments. All subjects had normal or corrected-to-normal vision. Viewing was binocular.

Eye-movement recording

The horizontal and vertical positions of the right eye were recorded with an electromagnetic induction technique (Robinson, 1963). A scleral search coil was embedded in a silastin ring (Collewijn, Van Der Mark, & Jansen, 1975), as described by Yang, FitzGibbon, & Miles (2003).

Visual display and stimuli

Dichoptic stimuli were presented using a Wheatstone mirror stereoscope. In a darkened room, each eye saw a computer monitor (HP p1230 21-in. CRT; HP, Palo Alto, CA) through a 45° mirror, creating a binocular image 521 mm straight ahead from the eyes' corneal vertices, which was also the optical distance to the images on the two monitor screens. Each monitor was driven by an independent PC (Dell Precision 490; Dell, Round Rock, TX), but the outputs of each computer's video card (PC NVIDIA Quadro FX 5600, NVIDIA, Santa Clara, CA) were frame-locked via NVIDIA Quadro G-Sync cards. The monitor screens were each 41.8° wide and 32.0° high, had 1024 × 768-pixel resolution (i.e., 23.4 pixels/° directly ahead of each eye), and the two were synchronously refreshed at a rate of 150 Hz. Each monitor was driven via an attenuator (Pelli, 1997) and a video signal splitter (Black Box Corp., AC085A-R2; Black Box Corp., Lawrence, PA), allowing presentation of black/white images with 11-bit grayscale resolution (mean luminance: 20.8 cd/m^2). Visual stimuli—single or two to three vertical sinusoidal luminance gratings (fully overlapping)—were seen through a ∼22° × ∼22° (512 × 512 pixels) rectangular aperture centered directly ahead of the eyes. The stimuli seen by the two eyes were always the same: we used the stereoscope at the outset of the project because we were not sure if we would need binocular manipulations to understand these responses.

Experiment 1: Different SF pairs

Horizontally moving gratings of eight different SFs (range: 0.08–1.25 cpd) composed the stimulus set of Experiment 1. Gratings shifted ⅛-wavelength each video frame, that is, with a TF of $18\frac{3}{4}$ Hz. Any given stimulus was either a single grating or a sum of two

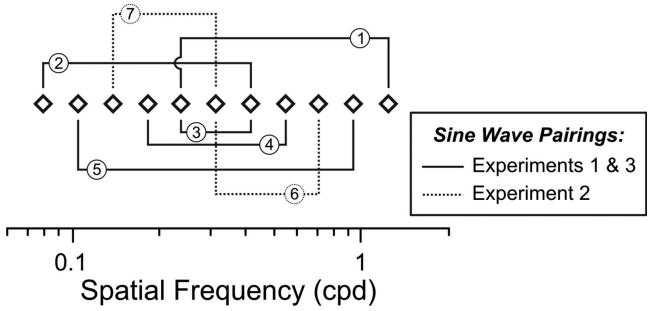

Figure 1. SFs of sine wave gratings used in Experiments 1, 2, and 3, shown as diamond symbols arranged along the SF axis. Solid thin lines (numbered from 1 to 5) connect SFs that form two-component stimuli of Experiments 1 and 3; dotted thin lines (numbered 6 and 7) connect SFs forming two-component stimuli of Experiment 2.

gratings—component gratings—moving in opposite directions. In a two-component stimulus, the Michelson contrast of one grating was always kept at 12%, whereas that of the other was 7%, 12%, or 21%. (Numbers from 1 to 5 in Figure 1 label the two-component stimuli used in Experiment 1.) The phases of the two components on the first video frame of a trial were randomized independently. A single block of trials had 62 randomly interleaved stimuli: 32 single-grating and 30 two-component conditions.

Before Experiment 1 was conducted in all subjects, we ran numerous pilot experiments in subject BMS. Those experiments employed gratings over a wide range of SFs (0.04–1.3 cpd) that formed 35 different SF pairings. This large dataset was used to select the most informative pairings to use in the remaining subjects.

Experiment 2: Variation in SF and TF

In Experiment 2, gratings of three different SFs (0.14, 0.32, or 0.72 cpd) were presented either in isolation or as a two-component stimulus (pairings 6 and 7 in Figure 1). In the two-component stimuli, one component (0.32 cpd) could be given one of three Michelson contrasts (∼7%, 12%, or ∼21%) but had constant TF of motion: $18\frac{3}{4}$ Hz in pairing 6 and 25 Hz in pairing 7. The contrast of the other component was always 12%, while its TF could be set to one of three values: $3\frac{1}{8}$, $18\frac{3}{4}$, or 30 Hz for 0.72-cpd grating in pairing 6 and $6\frac{1}{4}$, $12\frac{1}{2}$, or 25 Hz for 0.14-cpd grating in pairing 7. As before, the initial phases were randomized. A single block of trials had 60 randomly interleaved stimuli: 24 single-grating and 36 two-component conditions.

Pilot experiments in one subject (BMS) tested wider SF (0.05–0.72 cpd) and TF (from $3\frac{1}{8}$ to $37\frac{1}{2}$ Hz) ranges. Once again, this extended dataset allowed us to better formulate the conditions for Experiment 2 run in all subjects such that they would cover the most relevant portions of stimulus space.

Experiment 3

In Experiment 3, the total area covered by visual stimuli was reduced four-fold to ∼11° × ∼11° (256 × 256 pixels). Otherwise, the stimulus conditions were the same as in Experiment 1.

Experiment 4

Experiment 4 studied stimuli composed of three components, where the motion direction of one grating (G2) was opposite to that of the other two (G1 and G3). One component was 0.26 cpd in all subjects—close to the optimal SF for driving the OFR (G1; optimal). The second component had a much lower SF: 0.09 or 0.08 cpd (low). The third component had one of two much higher SFs, chosen (for each subject individually) such that it either had the same weight (as determined in Experiment 1) as the low SF component ($high^W$), or produced a response amplitude when presented alone similar to that of the low SF component ($high^A$). The TF was $18\frac{3}{4}$ Hz for all gratings. The optimal SF always moved in the same direction as one of the other two components, and these two gratings each had a Michelson contrast of 12%. The component moving in the direction opposite to the other two gratings could have Michelson contrast of ∼7%, 12%, or ∼21%. At the beginning of a trial, the phases of all components were randomized. A single block of trials had 68 randomly interleaved stimuli: 20 single-grating, 24 two-component, and 24 three-component conditions.

Procedures

Experimental paradigms were controlled by three PCs, which communicated via Ethernet (TCP/IP protocol). The first PC used a Real-time EXperimentation software (Hays, Richmond, & Optican, 1982), which provided the overall control of the experimental protocol, acquisition, display, and storage of the eye movement data. Two other PCs used the Psychophysics Toolbox extensions of MatLab (Brainard, 1997; Pelli, 1997) and generated the visual stimuli.

At the start of each trial a fixation target (dia. 0.25°) appeared at the screen center. After the subject's eye had been positioned within 2° of the fixation target and no saccades had been detected (using an eye velocity threshold of 18°/s) for a randomized period of 600–1,000 ms the fixation target disappeared and the first frame of the (randomly selected) stimulus appeared. Its first horizontal motion step commenced one video frame (6.7 ms) later. The motion lasted for 200 ms; the screen then turned to uniform gray (luminance, 20.8 cd/m^2) marking the end of the trial. A new fixation target appeared after a 500 ms intertrial interval, signaling a new trial. The subjects were asked to

refrain from blinking or shifting fixation except during the intertrial intervals, but were given no instructions relating to the motion stimuli. If no saccades were detected for the duration of the trial, then the data were stored; otherwise, the trial was aborted and repeated within the same block. Data collection usually occurred over several sessions until each condition had been repeated an adequate number of times to permit good resolution of the responses (through averaging).

Data analysis

The calibration procedure provided eye position data which were fitted with second-order polynomials and later used to linearize the horizontal eye position data recorded during the experiment. Eye-position signals were then smoothed with an acausal sixth-order Butterworth filter (3 dB at 30 Hz) and mean temporal profiles were computed for each stimulus condition. Trials with micro-saccadic intrusions (that had failed to reach the eye velocity cutoff of 18°/s used during the experiment) were deleted. We used position difference measures to minimize the impact of directional asymmetries and boost the signal-to-noise ratio. Position difference measures were calculated by subtracting the mean horizontal eye position for a given stimulus from the mean horizontal eye position to the same stimulus moving in the opposite direction. Because this work measures OFRs to pairs of gratings moving in opposite directions, we arbitrarily gave a positive sign to this value for one of the components, and a negative sign for the other.[1] The mean eye velocity was estimated by subtracting position difference measures 10 ms apart (central difference method) and evaluated every millisecond. Response latency was estimated by determining the time after stimulus motion onset when the mean eye velocity first exceeded 0.1°/s. The initial OFRs to a given stimulus were quantified by measuring the changes in the mean horizontal eye position signals—OFR amplitude—over the initial open loop period, that is, over the period up to twice the minimum response latency. This window always commenced at the same time after the stimulus motion onset ("stimulus-locked measures") and, for a given subject, was the same in all experiments reported in this article: 69–138, 74–148, and 66–132 ms for BMS, EJF, and TH, respectively. Bootstrapping procedures were used for statistical evaluation of the data and to construct 68% confidence intervals of the mean in the figures (these intervals were smaller than the symbol size in many cases and, therefore, not visible on most graphs).

Results

Data for one SF pairing in subject BMS is shown in Figure 2, as an example. Figure 2A displays averaged OFR velocity traces for different stimulus conditions. The black solid line is the OFR velocity trace in response to a 12%-contrast 0.22 cpd drifting grating when it was presented in isolation (labeled G1). Each dashed line is the OFR velocity trace induced by a single 0.36 cpd grating (labeled G2)—darker blue traces refer to higher contrast stimuli, which produced stronger OFRs. The OFR traces to two-component stimuli—G1

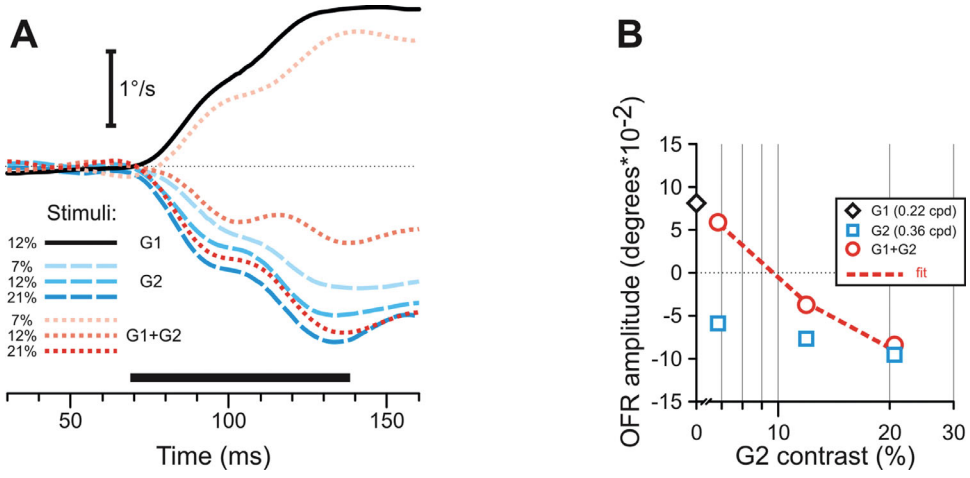

Figure 2. Experiment 1. Subject BMS. (A) Mean eye velocity profiles over time to pure sine wave gratings (G1 or G2) and two-component stimuli (note lines of different color and style; see the legend). Each trace is the mean response to 77–89 stimulus repetitions. The abscissa shows the time from stimulus onset; horizontal thin dotted black line represents zero velocity; the horizontal thick black line beneath the traces is the response measurement window. (B) Mean OFR amplitude (the same experiment as in A). Dependence on G2 contrast. G1 contrast was 12%. Symbols: data; dashed line: Equation 1 fit.

and G2 moving in opposite directions—are shown by red dotted lines. A combination of 12% contrast G1 and 7% contrast G2 produced the OFRs in the direction of G1 motion and only slightly smaller than that to 12% contrast G1 in isolation (cf. light-red dotted and solid black lines). This change in amplitude is much smaller than the one predicted by the linear summation of the two responses (i.e., by summing light blue dashed and solid black lines). A combination of 12% contrast G1 and 21% contrast G2 produced the OFRs in the direction of G2 motion and only marginally smaller than those obtained to motion of 21% contrast G2 in isolation (cf. dark red dotted and dark blue dashed lines), namely, the outcome of the competition between gratings differing in contrast is nonlinear, as described by Sheliga et al. (2006). Earlier studies described this pattern as winner-take-all behavior. Note that both gratings produce responses of similar magnitude when presented at 12% contrast, yet the response to the sum of these gratings (medium red dotted line) lies closer to the response to G2 (0.36 cpd). Thus, which component dominates cannot simply be predicted from the OFR responses to single gratings, indicating that the grating's SF per se plays a role.

Figure 2B summarizes the responses in Figure 2A with displacement measures—the change in eye position during the open-loop period. These OFR amplitudes to two-component stimuli (red circles) were fit ($r^2 = 0.995$; red dashed line) by the following equation:

$$R = \frac{R_1 * (WR * C_1)^n + R_2 * C_2{}^n}{(WR * C_1)^n + C_2{}^n}, \quad (1)$$

where R_1 and R_2 are the OFR amplitudes produced by each component in isolation, and C_1 and C_2 are the contrasts of each component. In this Experiment, R_2 is always negative, because the second grating moves in the opposite direction to the first one. The two free parameters, WR and n, summarize the interaction. n characterizes the steepness the transition between R_1 and R_2. WR stands for weight ratio; Equation 1 can be rewritten as

$$R = \frac{R_1 * (W_1 * C_1)^n + R_2 * (W_2 * C_2)^n}{(W_1 * C_1)^n + (W_2 * C_2)^n}, \text{ where}$$

$$WR = \frac{W_1}{W_2} \quad (2)$$

In this instance, the WR is equal to 0.81, that is, 0.36 cpd grating is given more weight than 0.22 cpd one and dominates the interaction (quantifying the observations made, while examining mean velocity traces in Figure 2A). Although it is not surprising that fits were very good (three modelled data points and two parameters), this function also provided an excellent account when more contrast levels were used (see Appendix B). It is for this reason that we were able to rely on just three contrast ratios to estimate WR and n for a given grating pair.

In our earlier report (Sheliga et al., 2006), we described these interactions using a different equation:

$$R = \frac{R_1 * C_1{}^{n_1} + R_2 * C_2{}^{n_2}}{C_1{}^{n_1} + C_2{}^{n_2}} \quad (3)$$

where R_1, R_2, C_1, and C_2 have the same meaning as in Equation 1, and n_1 and n_2 are two free parameters. Although Equations 1 and 3 are clearly different, there are situations in which both equations produce identical fits. One of these situations is when the contrast of one component is kept constant, as in the current study. In this case it can be shown that:

$$WR = C_1^{\frac{n_1}{n_2}-1} \text{ and } n = n_2 \quad (4)$$

where C_1 is the contrast used for the fixed contrast component, and n_2 is fit to the component that changes contrast. A second situation is where $WR \approx 1$, in which case both models give very similar description for the same patterns (i.e., they each provide a good fit to data generated by the other model), even when both contrasts are varied (hence, both models provide equally good descriptions of the data in Sheliga et al. (2006), see below). We ran a series of simulations to identify conditions where the two models were most distinguishable, which we found to be true when both components varied in contrast and the WR was far from 1. We ran one such condition (0.07/0.59 cpd pairing) in subject BMS and found that Equation 1 provided a better description of data than Equation 3 ($p = 0.0028$, see Figure A1 in Appendix A). Furthermore, the differences between the data and the Equation 1 fit were not statistically significant ($p = 0.14$). We proceeded, therefore, with using Equation 1 to describe interactions in two-component stimuli.

Experiment 1: Varying SF

We examined how two parameters of Equation 1—WR and n—depended on component SF. Figure 3A–C shows WR (color-coded small squares; see color bar to the right of panel C) as a function of SF_1 and SF_2 in a two-component stimulus[2] for 3 subjects. Each of five SF pairings contributed two symbols to the plot—symmetrically located across the main diagonal—and their WR values were the inverse of one another. This is because swapping the labels SF_1 and SF_2 and substituting $\frac{1}{WR}$ is an identity mapping. Extensive pilot experiments in one subject (Appendix

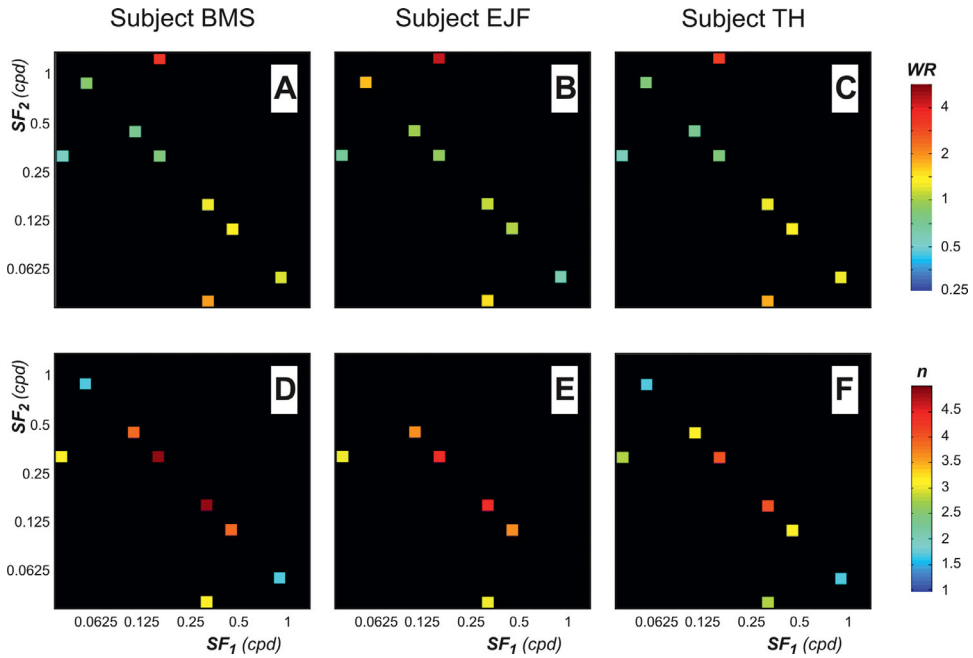

Figure 3. Experiment 1. Dependence of parameters *WR* (A–E) and *n* (D–F) (from Equation 1) on the SF of the components. (A–C) *WRs* as a function of SF of sine wave gratings in two-component stimulus (color-coded small squares; see the color bar to the right of C; the background color (black) is irrelevant and was chosen to make *WRs* of different colors clearly visible. (D–F) Values of *n* (color-coded small squares; see a color bar to the right of F) as a function of SF of sine wave gratings in two-component stimulus; the background color is arbitrary as in A–C.

B) showed that the value of WR was well-described by a separable function of SF, relating W_i to SF_i:

$$W_i = F_1(SF_i) * F_2(SF_i), i = 1, 2 \quad (5)$$

Equation 5 was constructed from the product of two functions: an exponential (Equation 6) and an inverted cumulative Gaussian function (Equation 7):

$$F_1(SF) = e^{\lambda_W * \log_2(SF)} \quad (6)$$

$$F_2(SF) = 1 - \frac{1}{2} * (1 + erf\left(\frac{\mu_{HC} - \log_2 SF}{sqrt(2) * \sigma_{HC}}\right) \quad (7)$$

where λ_W, μ_{HC}, and σ_{HC} are three free parameters. Note that the absolute scaling of Equation 6 is unconstrained, because W_i is only ever used to calculate the weight ratio, $WR = \frac{W_1}{W_2}$, for a given pair of SFs. *WRs* of three subjects, shown in Figure 3A–C, were very well fit by Equation 5: $r^2 = 0.998$, 0.995, and 0.996 for subjects BMS, EJF, and TH, respectively. Figure 3D–F plots *n* (color-coded small squares; see color bar to the right of Figure 3F) as a function of SF_1 and SF_2 in a two-component stimulus[3] for three subjects. Once again, each SF pairing contributed two symbols to the plot, with the same value in symmetrical locations around the main diagonal (the value of *n* does not depend on the order of SF_1 and SF_2). Pilot experiments, which tested much higher number of SF pairings (Appendix C), allowed to suggest a model, describing the dependence of *n* upon SF_1 and SF_2, which was then tested in 3 subjects. There are two notable features in the model. First, *n* is largest close to the identity line (when the SF ratio is close to 1) and falls off smoothly as the SF ratio increases. We described this with a Gaussian of the *log* SF ratio:

$$n = A_n * e^{-\frac{\left[\log_2\left(\frac{SF_1}{SF_2}\right)\right]^2}{2*\sigma_n^2}} + 1 \quad (8)$$

One is added because values of *n* less than one would lead to low contrast stimuli dominating. Second, *n* decreases more rapidly with the SF ratio at high and low frequency, with the slowest decline for pairs where the geometric mean SF is ~0.25 cpd, close to the optimal SF for driving the OFR with single gratings. We capture this feature by allowing the value of sigma in Equation 8 to be a Gaussian function of the log of the product of SFs:

$$\sigma_n = A_{\sigma n} * e^{-\frac{\left[\mu_{\sigma n} - \frac{\log_2(SF_1 * SF_2)}{2}\right]^2}{2*\sigma_{\sigma n}^2}} \quad (9)$$

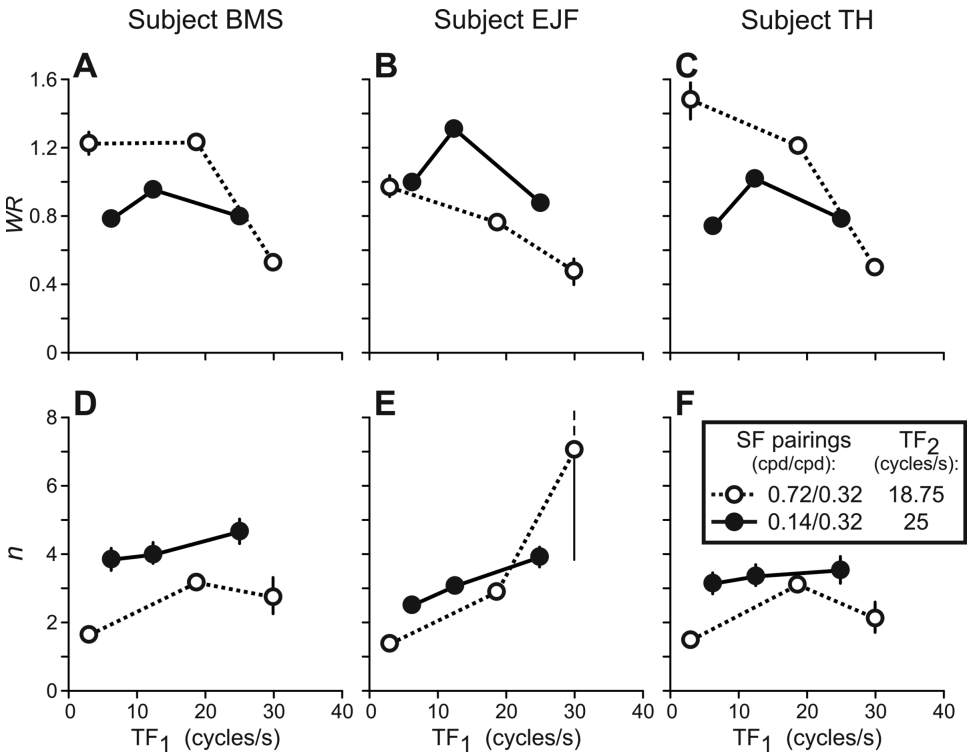

Figure 4. Experiment 2. Dependence of *WR* (A–C) and *n* (D–F) on the TF of the components. Thin colored vertical lines: 68% confidence intervals of the mean (bootstrapping). In many cases these intervals were smaller than a symbol size (especially, for *WR* data) and, therefore, not visible on the graphs.

Thus, n in Equation 1 seems to depend on both the ratio and the product of component SFs:

$$n = F_3\left(\frac{SF_1}{SF_2}; SF_1 * SF_2\right) + 1 \quad (10)$$

Equation 10 provided a good fit for the relationship between SF_1, SF_2 and n (shown in Figure 3D–F) using four free parameters: $r^2 = 0.999$, 0.976, and 0.993 for subjects BMS, EJF, and TH, respectively.

Experiment 2: Varying TF and SF

In Experiment 2, we manipulated the TF of components, in addition to SF and contrast. In a pairing, the TF of one component varied, whereas it was fixed for the other component. Data were collected from six SF_1TF_1, SF_2TF_2 pairings, and for each pairing we recorded the OFRs to three contrast combinations (see Methods). The fitted values of WR and n are shown in Figure 4A–C and 4D–F, respectively, and one can clearly see that, for a given SF pairing, the values of WR and n show further sizeable changes, depending on the TF of the components. Our pilot data suggested (Appendix D) that the changes in WR can be well described by a simple interaction between SF and TF, where WR is a Gaussian function of TF (Equation 11), but the TF at which WR peaks (i.e., optimal TF) depends on SF (Equation 12):

$$F_4(SF; TF) = e^{-\frac{[G(SF)-TF]^2}{2*\sigma_{TFw}^2}} \quad (11)$$

$$G(SF) = A_{TFw} * e^{-\lambda_{TFw}*\log_2(SF)} \quad (12)$$

We choose an exponential in Equation 12 because it is well-behaved as SF becomes large. We can now summarize all the factors influencing a component's weight by combining Equations 5 and 11 (adding three free parameters):

$$W_i = F_1(SF_i) * F_2(SF_i) * F_4(SF_i; TF_i), i = 1, 2 \quad (13)$$

Our pilot data showed (Appendix D) that n is a separable function of SF and TF, well-described with a Gaussian function of TF:

$$F_5(TF) = e^{-\frac{[\mu_{TFn}-TF]^2}{2*\sigma_{TFn}^2}} \quad (14)$$

Subject	λ_W	μ_{HC}	σ_{HC} (log2 units)	A_n	$A_{\sigma n}$	$\mu_{\sigma n}$	$\sigma_{\sigma n}$ (log2 units)	σ_{TFw}	A_{TFw}	λ_{TFw}	μ_{TFn}	σ_{TFn}
1.1 The single set of free parameters for Experiments 1 and 2												
BMS	0.30	1.00	0.35	4.78	1.94	0.20	1.03	13.8	10.1	0.16	22.8	17.1
EJF	0.27	0.81	0.57	4.13	2.33	0.08	2.00	12.7	8.60	0.18	20.7	10.9
TH	0.32	1.04	0.37	3.60	1.89	0.21	1.03	13.7	9.14	0.21	19.2	15.2
1.2 Free parameters for Experiment 3												
BMS	0.23	1.31	0.37	3.97	2.55	0.28	0.94					
EJF	0.16	1.08	0.42	3.65	2.48	0.28	1.16					
TH	0.22	1.78	0.76	2.87	1.75	0.29	0.94					
1.3 The single set of free parameters for Experiments 1, 2, and 3												
		Experiments 1 and 2 μ_{HC}	Experiment 3 μ_{HC}			Experiments 1 and 2 $\mu_{\sigma n}$	Experiment 3 $\mu_{\sigma n}$					
BMS	0.30	0.99	1.28	0.35	4.66	1.99	0.19	0.29	1.06	14.0	10.1	0.16	21.8	15.9
EJF	0.27	0.81	1.06	0.56	4.04	2.19	0.10	0.33	1.79	12.8	8.35	0.19	20.8	11.3
TH	0.32	1.04	1.51	0.38	3.65	1.81	0.17	0.77	1.35	13.7	9.07	0.21	19.7	15.5
1.4 The single set of free parameters for Experiments 1, 2, 3, and 4												
		Experiments 1, 2, and 4 μ_{HC}	Experiment 3 μ_{HC}			Experiments 1, 2, and 4 $\mu_{\sigma n}$	Experiment 3 $\mu_{\sigma n}$					
BMS	0.31	1.00	1.29	0.38	5.02	1.26	0.16	0.29	1.38	13.9	9.83	0.17	24.7	17.7
EJF	0.26	0.80	1.06	0.56	3.94	1.76	0.05	0.68	2.36	12.7	8.64	0.18	21.3	11.2
TH	0.30	1.07	1.41	0.36	7.05	1.51	0.09	1.69	2.11	14.0	9.43	0.20	47.1	32.8

Table 1. Best-fit values of free parameters in Experiments 1, 2, 3, and 4

Combined with Equation 10, this allowed us to summarize the factors influencing n for any spatiotemporal component:

$$n = F_3\left(\frac{SF_1}{SF_2}; SF_1 * SF_2\right) * \prod_{i=1}^{2} F_5(TF_i) + 1 \quad (15)$$

Experiments 1 and 2: Testing the model

Extensive pilot experiments allowed us to identify a good descriptive model with a modest number of free parameters (six describing WR, six describing n; see Table 1.1). We then used the model to test five SF pairs (0.08–1.25 cpd) with the same TF (Experiment 1) and six SF/TF combinations (Experiment 2, TFs from 3⅛ to 30 Hz). OFRs to these 11 pairings were measured in three subjects. The model was then fit to each subject's data. The results of Experiments 1 and 2 are shown in Figures 5 and 6, respectively, with responses to two-component gratings shown with red circles, and the model fit shown with dashed red lines. The fits were very good: $r^2 = 0.994$, 0.991, and 0.989 for subjects BMS, EJF, and TH, respectively. Figure 7 shows the fitted functions used for each subject, which are for the most part similar. The greatest intersubject variation seems to be in the effect of geometric mean SF on how rapidly n falls off with SF ratio (Figure 7F).

Experiment 3: The effect of stimulus size

In Experiment 3 we repeated Experiment 1, but the total area covered by visual stimuli was reduced four-fold: from 22° × 22° to 11° × 11°. Equations 5 and 10 were fit again to these data and provided an excellent description (Figure 8; red dashed lines): $r^2 = 0.993$, 0.993, and 0.984 for subjects BMS, EJF, and TH, respectively. Table 1.2 lists the best-fit values of free parameters. Because we did not vary TF in this experiment, Equations 13 and 15 were not used. Many of the fitted parameters were very similar to those obtained in Experiments 1 and 2. Indeed, we found that we could constrain five parameters to be the same as those used for Experiments 1 and 2, with almost no change in fit quality ($r^2 = 0.994$, 0.992, and 0.986 for subjects BMS, EJF, and TH, respectively), shown by grey dotted lines in Figures 5, 6, and 8. The two parameters that were substantially changed by stimulus size were μ_{HC} and $\mu_{\sigma n}$: both moved to higher SF for the smaller stimulus (see Table 1.3).

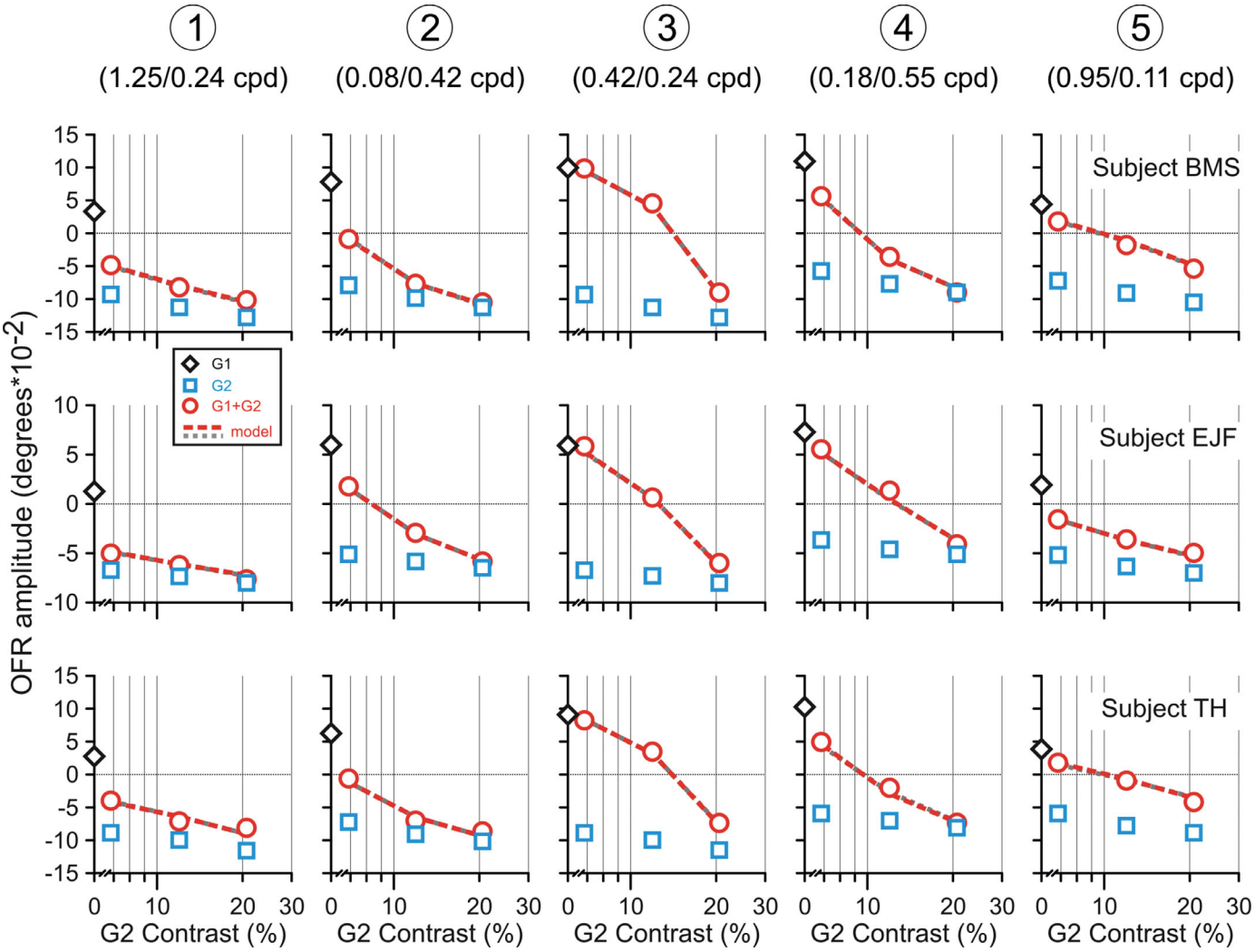

Figure 5. Experiment 1. Mean OFR amplitude. Dependence on contrast of grating 2 (G2). Grating 1 (G1) had a contrast of 12%. Symbols, data; red dashed lines, model fits in which one fit was used (per subject) to describe all the data in Experiments 1 and 2; gray dotted lines, fits of a single model applied to Experiments 1, 2, and 3 (almost indistinguishable from red lines). Each row shows data for one subject, each column shows data for one SF pairing (identified by a circled number on top of each column, see Figure 1). Subjects BMS (145–155 trials per condition), EJF (85–111 trials per condition), and TH (106–120 trials per condition).

The change in model fits with stimulus size could reflect two processes. First, this could reflect differences in the relative strength of interactions between components. Second, it could simply be that changes in stimulus size alter the relative effectiveness of different SFs, even when presented alone. To explore this, we examined the responses to single gratings in Experiments 1 and 3 (12% contrast), as a function of SF. Figure 9 shows that there is a clear shift in the SF tuning curve towards higher frequencies for the smaller stimulus. We fit the tuning with a Gaussian function of log SF (e.g., Sheliga et al., 2005; Sheliga, Quaia, Cumming, & Fitzgibbon, 2012). Table 2 lists the fitted parameters, where it is clear that there is a substantial change in the location of the peak SF, but little change in the standard deviation. Constraining the standard deviations to be the same in both Experiments, led to a marginal decrease in r^2 values, and those fits are shown by solid (Experiment 1) and dashed (Experiment 3) grey lines. For both stimulus sizes, the peak SF of these tuning curves is very similar to the $\mu_{\sigma n}$ parameter used to describe responses to two-component gratings in Equation 9. Thus, size seems to influence this parameter through its effect on single gratings. If, instead of allowing $\mu_{\sigma n}$ to be a free parameter in Equation 9, we constrain it to be the peak of the SF tuning to single gratings, we get almost no change in r^2s (0.993, 0.991, and 0.987 for subjects BMS, EJF, and TH, respectively). Overall then, this allows us to describe the interactions between components in 16 stimuli (32 values) with just 12 free parameters.

Furthermore, a closer inspection of Tables 1.3 and 2 reveals that in each subject the change in μ_{HC} (fit

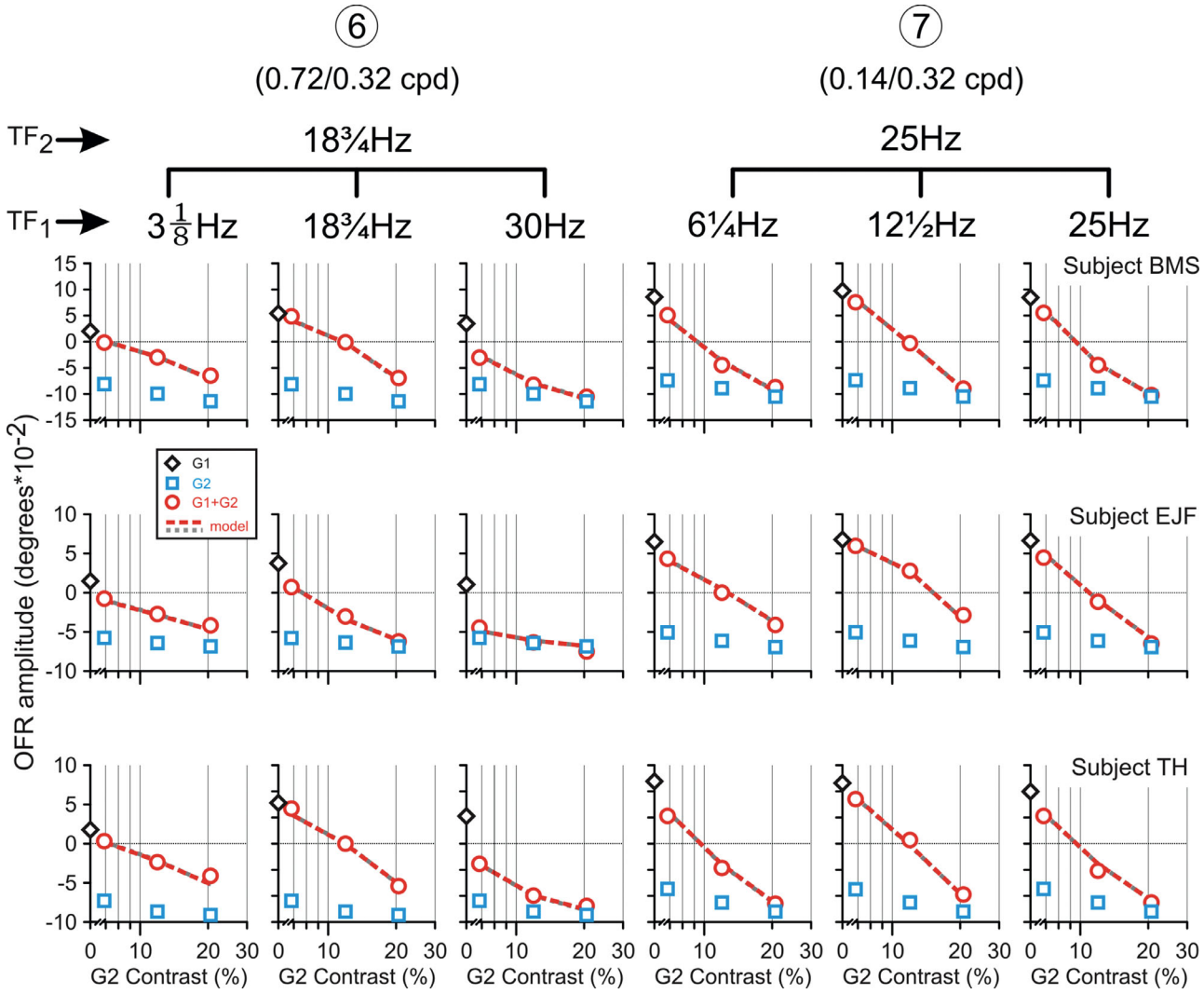

Figure 6. Experiment 2. Variation in SF and TF. Each panel plots the mean OFR amplitude as a function of the contrast of grating 2 (G2). Grating 1 (G1) had a contrast of 12%. Symbols, data; red dashed lines, fits of single model applied to Experiments 1 and 2; gray dotted lines, fits of model applied to Experiments 1, 2, and 3. Rows show data for one subject, columns show data for one SF/TF pairing (SF pair identified by a circled number on top, see Figure 1). Subjects BMS (65–72 trials per condition), EJF (89–113 trials per condition), and TH (107–125 trials per condition).

with Equation 7) owing to changes in stimulus size is very similar to the shift of the peak of the OFR SF tuning curve. If instead of allowing μ_{HC} to be a separate free parameter in each experiment we constrain it to be proportional to the peak of the SF tuning to single gratings, we get little change in r^2s (0.988, 0.984, and 0.987 for subjects BMS, EJF, and TH, respectively). The constant of proportionality is then just one parameter for both experiments, allowing us to describe all the data with 11 free parameters.

Experiment 4: Stimuli with three components

Here we consider whether the model that we developed above, that described interactions in stimuli with two components, can be extended to more complex stimuli. That is, would a more general equation

$$R = \frac{\sum_{i=1}^{K} R_i * (W_i * C_i)^n}{\sum_{i=1}^{K} (W_i * C_i)^n}, i = 1, 2, 3...K \quad (16)$$

be accurate? In Experiment 4, we recorded the OFRs to stimuli composed of three grating components, all moving horizontally, but such that the direction of motion of one of them was opposite to that of the other two. These data allow us to address two slightly different questions. First, does Equation 16 describe the interactions observed for each individual three-component stimulus? Second, do the values W_i and n depend on the SF and TF of the components in

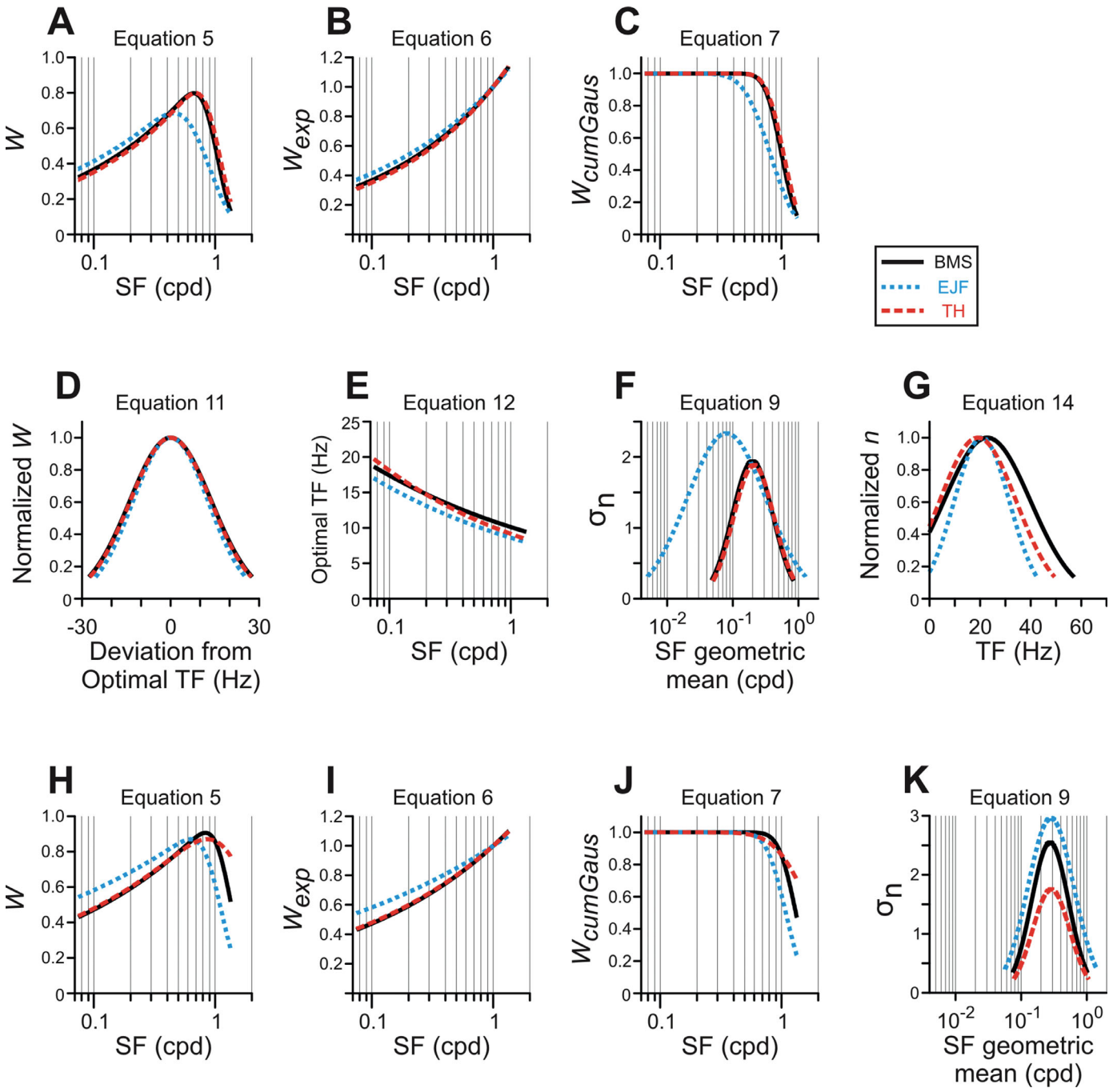

Figure 7. The relationships between stimulus parameters (SF, TF) and the model fits for each subject. (A–E) show the relationship between SF/TF and *W* for Experiments 1 and 2. (H–J) show this for Experiment 3 (smaller stimulus) where a rightward shift of peak SF can be seen. (F, G) The relationship between SF/TF and *n* for Experiments 1 and 2. (K) The relationship between SF/TF and *n* for Experiment 3 (smaller stimulus) where a rightward shift of peak SF geometric mean can be seen. (A) Equation 5: the function relating SF to *W*; product of Equations 6 and 7. (B) Equation 6: the high-pass component of the function relating SF to *W* is described with an exponential. (C) Equation 7: the lowpass component of the function relating SF to *W* is described with a Cumulative Gaussian. (D and E) A simple interaction between SF and TF: *W* is a Gaussian function of TF (D; Equation 11), but the TF at which *W* peaks (i.e., optimal TF) depends on SF (E; Equation 12). (F) Equation 9: the effect of geometric mean SF on how rapidly *n* falls off with SF ratio. (G) Equation 14: *n* is a Gaussian function of TF. Data of different subjects are shown using lines of different style and color; see figure legend.

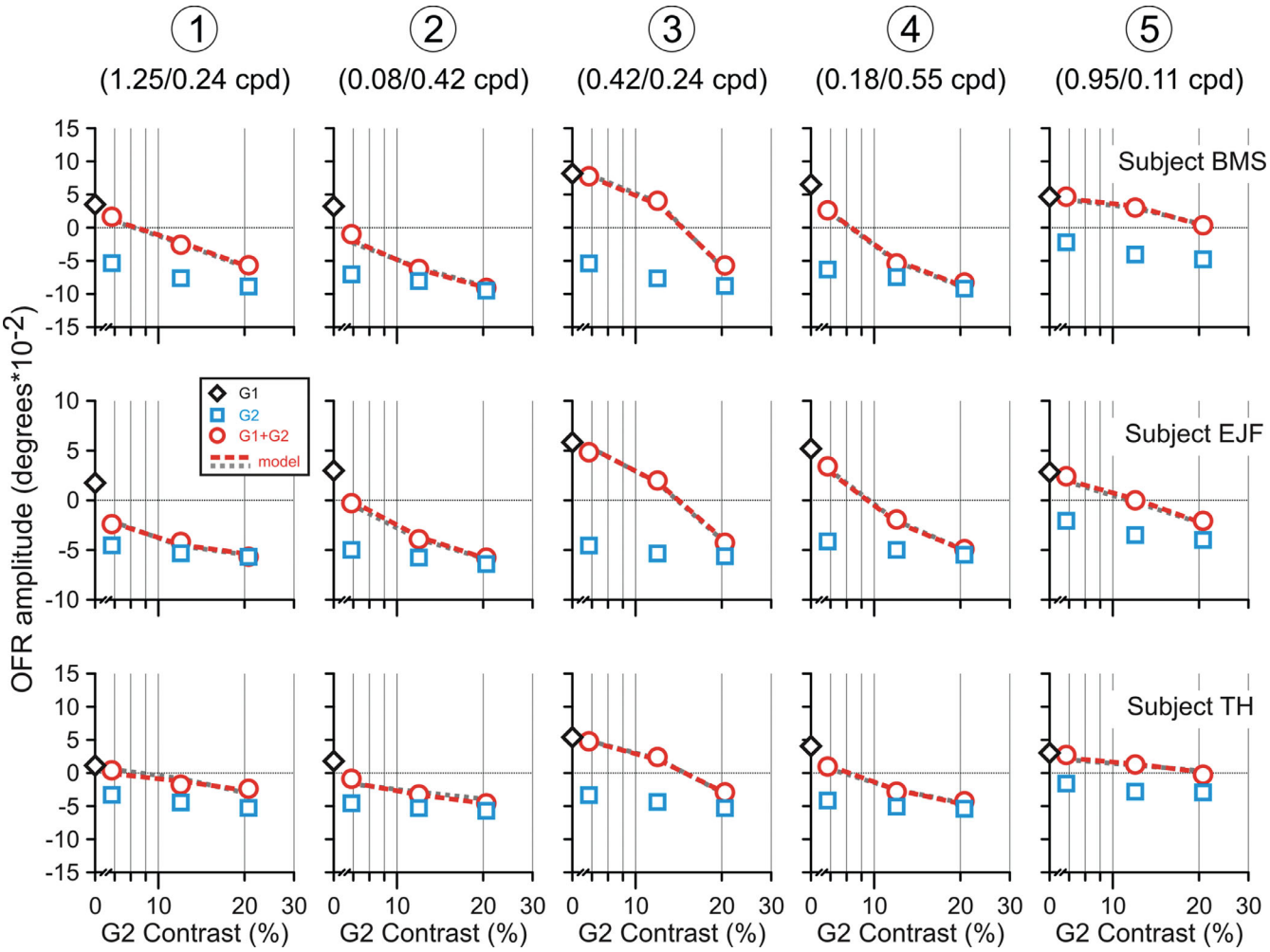

Figure 8. Experiment 3: as in Experiment 1, but with a smaller stimulus. Each panel plots the mean OFR amplitude as a function of the contrast of grating 2 (G2). The contrast of grating 1 (G1) was 12%. Symbols. data; red dashed lines, model fits; gray dotted lines, fits of single model applied to Experiments 1, 2, and 3. Rows show data for one subject, columns show one SF pairing (identified by a circled number on top of each column, see Figure 1). Subjects BMS (97–104 trials per condition), EJF (74–92 trials per condition), and TH (90–110 trials per condition).

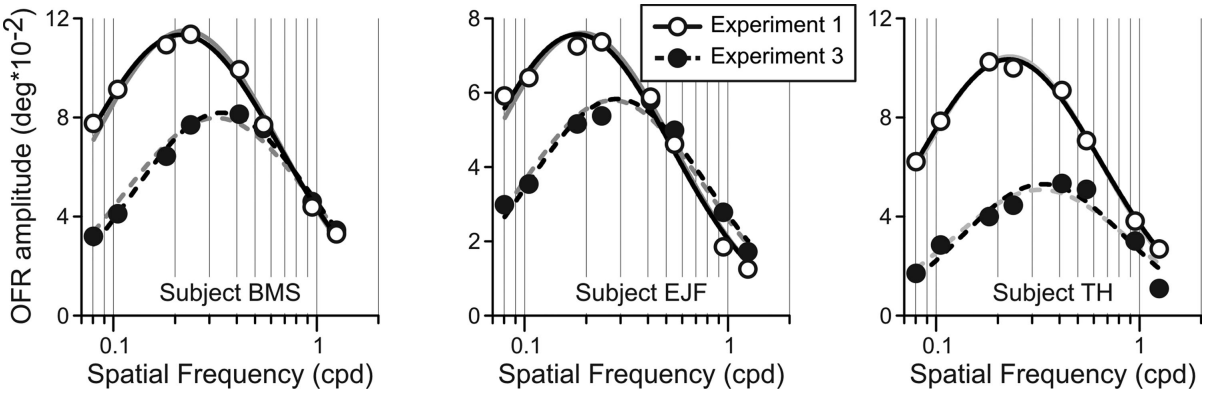

Figure 9. Mean OFR amplitude as a function of SF in single gratings. Symbols, data; lines, fits. Experiment 1 (stimulus size ∼22° × ∼22°): open symbols, solid lines. Experiment 3 (stimulus size ∼11° × ∼11°): filled symbols, dashed lines. Black lines, best-fit Gaussians; gray lines, Gaussian fits whose best-fit standard deviation free parameter was shared by both experiments. Note an almost complete overlap of black and gray fits for each subject.

	Experiment 1				Experiment 3			
Subject	Amplitude (°)	Offset (cpd)	Standard deviation (log2 cpd units)	r^2	Amplitude (°)	Offset (cpd)	Standard deviation (log2 cpd units)	r^2
BMS	0.12	0.22	1.59	0.993	0.08	0.34	1.44	0.989
EJF	0.08	0.18	1.52	0.985	0.06	0.29	1.46	0.964
TH	0.10	0.23	1.49	0.994	0.05	0.34	1.32	0.908

Table 2. Best-fit values of free parameters of SF tuning curves in Experiments 1 and 3

the same way? When considering the second question, several of the equations (Equations 6, 7, 11, 12, and 14) can be applied without modification. Equation 9 can simply be rewritten in a way that applies to the general case:

$$\sigma_n = A_{\sigma n} * e^{-\frac{\left[\mu_{\sigma n} - \frac{\log_2\left(\prod_{i=1}^{k} SF_i\right)}{k}\right]^2}{2*\sigma_{\sigma n}^2}}, k = 2 \text{ or } 3 \quad (17)$$

Equation 8, where n depends on the ratio of two SFs, does not readily apply to more than two components. However, a given SF ratio can also be described in terms of the variance of $\log_2(SF)$:

$$n = A_n * e^{-\frac{\left(\sqrt{\frac{\sum_{i=1}^{k}\left[\log_2(SF_i) - \overline{\log_2(SF)}\right]^2}{k-1}}\right)^2}{2*\sigma_n^2}}, k = 2 \text{ or } 3 \quad (18)$$

a form that can be applied to any number of components. This then provides us with a way to test our model with three component stimuli.

In our stimuli, the three components always consisted of the optimal SF (in the sense of producing the strongest response alone), one lower SF and one higher SF. The high SF took one of two values: high^A which produced OFRs of the same response amplitude as the low SF component (left column in Figure 10) or high^W, which had the same weight as the low SF component (according to the model), shown in the right column of Figure 10.

Figure 10 shows the OFRs to two- and three-component stimuli (red and green symbols, respectively) with the predictions of our model (dashed red and dotted green lines). These predictions were not fit to these data—we simply used the model and parameters derived from the two-component case to predict responses. Nonetheless, they accounted for most of the variance, with $r^2 = 0.965$, 0.975, and 0.908 for subjects BMS, EJF, and TH, respectively. We also fit our complete model to all of the data in Experiments 1–4, and this is shown by grey dashed and dotted lines: $r^2 = 0.990$ for all subjects (Figure 10). The fitted parameters are listed in Table 1.4. Clearly, the model provides an excellent account of responses to stimuli containing three components. One particularly telling case is the lower left quadrant for each subject. Here, the low and optimal frequencies moved in the same direction. The high frequency moved in the opposite direction and produced a smaller response when presented alone—it was chosen to have a high weight according to the model. Consequently, any competition that reflected the total driving signal in each direction would be dominated by the optimal + low pair. However, when the high frequency component is added (high^A), it has a powerful effect—compare the purple triangle with the green triangles. Similarly, comparing all three components (green triangles) with optimal + high^A (red circles) shows that adding the low frequency component (low) makes almost no difference. In this comparison, the low and high SF components here produced similar responses when presented alone, yet the effect of adding them to the other two components is very different.

Discussion

We explored the interaction between different grating components generating OFRs and discovered that a relatively simple model—a nonlinear weighed summation model—describes the interaction over a wide range of SFs and TFs. For any given pair, the interaction is described by two parameters (Equation 1): WR describes the relative weight given to the two components, and n describes how steeply the response changes with relative contrast. Large values of n correspond to the winner-take-all–like outcome described in Sheliga et al. (2006), which we replicate when the components have similar SF. For widely different SFs n approaches 1, which describes a linear summation.

The relative weight of two components

For any given pair of gratings, their relative weight is sufficient to describe their interaction in Equation 1 (parameter WR). Importantly, we found that this WR was a separable function of SF_1, SF_2, implying that the weight assigned to a component is determined by its SF, not a more complex pairwise interaction.

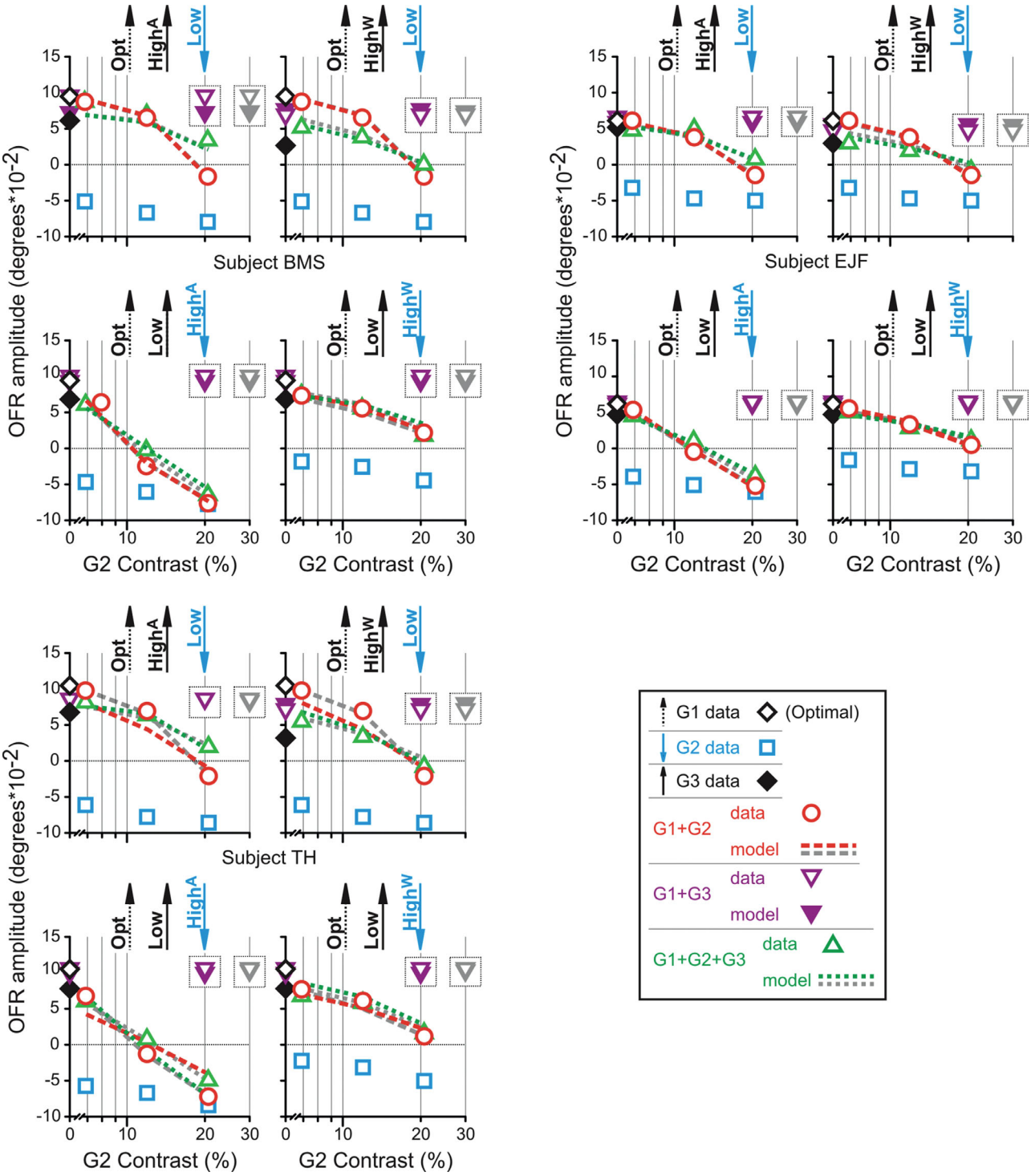

Figure 10. Experiment 4. Responses to the sum of three gratings (low, optimal [Opt], and high SF). Each panel plots the mean OFR amplitude as a function of grating 2 (G2) contrast, whose motion direction was opposite to that of the other two components (G1 and G3). Those components each had 12% contrast. Red dashed lines and maroon inverted filled triangles, predictions for two-component stimuli (data shown with red circles and maroon inverted open triangles, respectively); green dotted line, predictions for three-component stimuli. These predictions we made using the model fit to Experiments 1–3, with no additional fitting. All other symbols, data. Black vertical arrows on top of each panel show each component's direction of motion. Clusters of four neighboring panels show the data for one subject. In each cluster, the upper row graphs depict OFRs when the low-frequency component moved in the opposite direction to the other two; lower row graphs depict OFR when the high-frequency component ($High^A$ or $High^W$) moved in the opposite direction to the other two. In each cluster, left column graphs are where $High^A$ was used for the

→

←
high-frequency component, the right column is for High^W. See Methods for further details. Maroon open/filled inverted triangles are shown twice on each graph: in appropriate locations given the components' contrast and the OFR amplitude—where they are often, however, obscured by other symbols—and, for clarity, in a separate insert toward the right side of each graph. Gray open/filled inverted triangles, gray dashed and dotted lines: model fits whose best fit values of free parameters were shared by Experiments 1–4. Subjects BMS (90–96 trials per condition), EJF (63–81 trials per condition), and TH (85–97 trials per condition).

That is, the *WR* for a pair composed of SF_i and SF_j is simply determined by $(\frac{W_i}{W_j})$. Surprisingly, we found that weight is not simply proportional to the response elicited by that component alone. Rather, as SF is increased beyond the optimal value for driving the OFR, the weight continued to increase for ∼1 octave. This finding in turn implies that the neuronal circuits determining weights must precede the stage at which motor responses to visual motion are generated.

Halving the stimulus linear dimensions resulted in a substantial rightward shift of the function relating SF to weight (compare Figures 7A and 7H). This also produced quantitatively similar shifts in the OFR SF-tuning curves (Figure 9), suggesting that similar circuitry may be responsible for both changes. Surround inhibition may play an important role here—as stimulus size is increased, recruitment of neurons may be counterbalanced by increased surround inhibition. The surround inhibition may be relatively stronger for neurons preferring a high SF, because the same physical size corresponds with a greater multiple of RF width. If the ability of one grating component to suppress responses to a second component (high weight) uses the same pathway as surround suppression, this could explain why both functions change in the same way with stimulus size.

We found that the weight of a grating component also depended on the TF of its motion. For a component of any given SF, the dependence was well captured by a Gaussian function of TF (Figure 7D). The peak of this function changed somewhat with SF (Figure 7E), whereas the standard deviation was fixed. This spatiotemporal inseparability is not seen for OFR amplitudes elicited by single gratings (Sheliga, Quaia, FitzGibbon, & Cumming, 2016), so this is another way in which the factors determining the weight are different from those that drive the OFR. This inseparability could reflect a correlation between SF and TF preferences in the underlying neurons. The association of higher TF with lower SF is reminiscent of the difference between M and P ganglion cells (Derrington & Lennie, 1984), although this correlation is not clear in cortical neurons (V1 or MT; Hawken, Shapley, & Grosof, 1996; Lui, Bourne, & Rosa, 2007). However, it is difficult to reconcile a simple explanation like this with the fact that OFR amplitudes for single gratings are a separable function of SF and TF.

The nonlinear summation of contrast

Parameter *n* determines how steeply the OFR amplitude changes with the contrasts of the components, with values greater than 1, indicating a nonlinear interaction. This nonlinear interaction is greatest for similar SF. At first sight this might suggest that the competition is taking place within (rather than between) spatiotemporal channels (e.g., contrast gain control). However, this interaction also depends on absolute SF—dissimilar SFs interact more strongly when they are near the peak suggesting a more complex mechanism.

Previous work

In experiments very similar to those reported in this paper, Sheliga et al. (2006) recorded the OFRs to two-component grating stimuli moving in opposite directions, in which they varied the contrasts of both components. In principle, these data could pose a significant challenge to our model (see Methods). We show here that the OFRs across all conditions in that study are well-described by a single value of *n* and *WR*: $r^2 = 0.985$, 0.990, and 0.987 for subjects BMS, FAM, and JKM, respectively[4] (Figure 11). Equation 3 also provides good fits for all the data. Both models succeed in this case, because the individual components have similar weights.

Sheliga et al. (2016) recorded the OFRs to white noise stimuli in human subjects. Responses were shaped by two factors: (1) an excitatory drive that reflected the OFRs to individual Fourier components and (2) a suppression by higher SF channels where the temporal sampling of the display led to flicker. The explanation we offered there depended on the idea that flickering components were suppressive. However, the model we have developed here might explain the same result without treating flicker as a special case—it may simply reflect the summed contribution of the flickering SFs to the denominator in Equation 1. We therefore applied our model to one of the experiments reported by Sheliga et al. (2016) in which a moving bandpass stimulus was added to a flickering bandpass stimulus. The data from the two subjects who were also the participants of the current study are reproduced in Figure 12.[5] The central SF of the moving pattern was

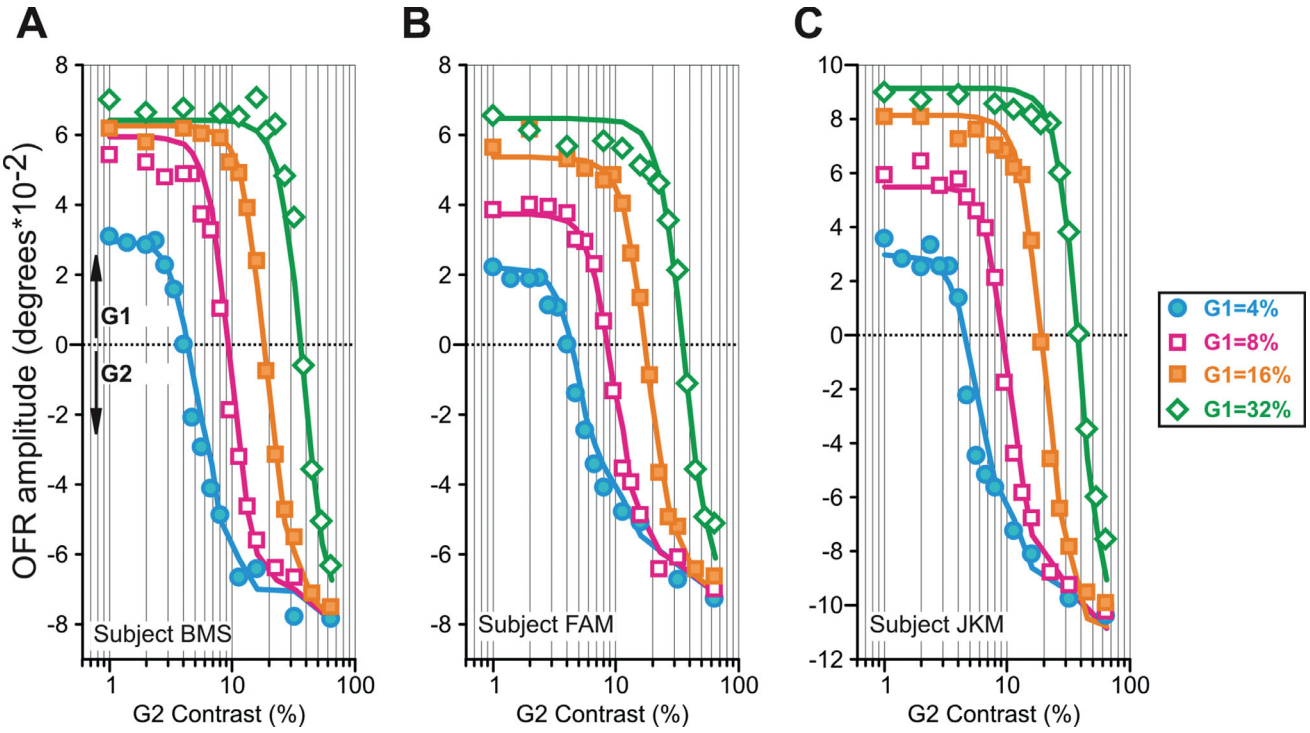

Figure 11. Data from Experiment 2 of Sheliga et al. (2006), who recorded the OFRs to two-component sine wave grating stimuli moving in opposite directions. The OFR amplitudes for four sets of experimental conditions are shown by different symbols and colors. In each set, the contrast of one grating was fixed (G1; see Figure legend), whereas the contrast of the other was varied (G2). Solid lines: Equation 1 fits. For each subject a single pair of parameters (*WR* and *n*) was used to describe all of the data. (A) Subject BMS (153–171 trials per condition). (B) Subject FAM (133–150 trials per condition). (C) Subject JKM (150–177 trials per condition).

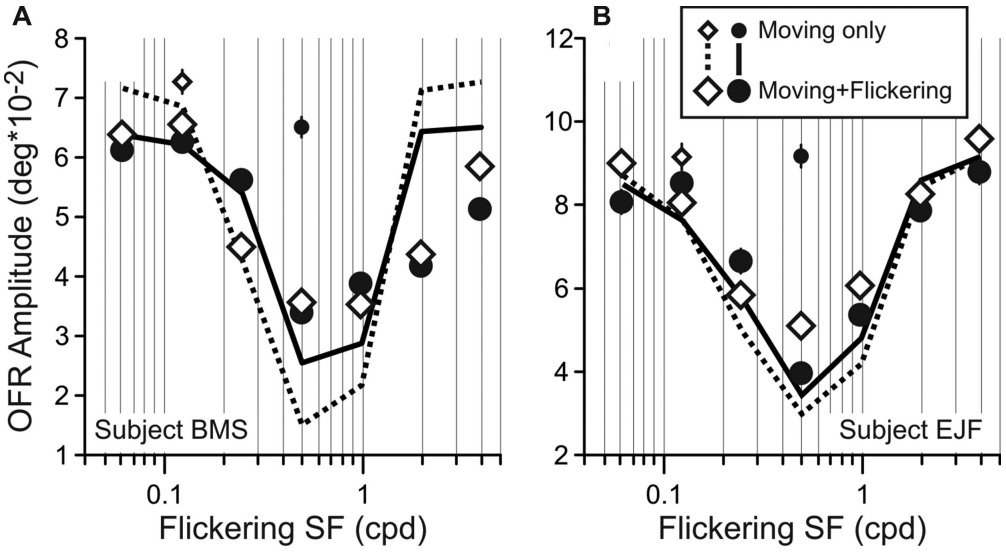

Figure 12. Data from Experiment 1D of Sheliga et al. (2016), showing the dependence of mean OFR amplitude on flicker SF. Symbols, data; lines, fits; small symbols, moving band-pass filtered noise only; large symbols, combined moving and flickering band-pass filtered noise; lines, predictions from our model fit to Experiments 1–3 of this work. The moving noise stimulus was filtered with one of two central SFs: 0.125 (filled circles, solid line fits) or 0.5 cpd (open diamonds, dotted line fits). (A) Subject BMS (174–187 trials per condition). (B) Subject EJF (76–99 trials per condition).

either 0.125 (open diamonds) or 0.5 (filled circles) cpd, and it was paired with flickering samples, whose central SFs varied from 0.0625 to 4 cpd in octave increments: in effect, a SF tuning for flicker. Moving stimuli were also presented in isolation (smaller in size diamond and circle symbols), for comparison. Figure 12 shows that the flickering stimuli with central SFs in the range of 0.5 to 1.0 cpd were the most detrimental for both subjects. At the time, we were puzzled that these SFs were the most powerful, since they are higher than the optimum SF for driving the OFR. Since we have now demonstrated that the weight given to a component in Equation 1 is also biased towards higher SF, we tested whether Equation 1 offers a good description of those results:

$$\vec{R} = \frac{\vec{R}_M * (WR * C_M)^n + \vec{R}_F * {C_F}^n}{(WR * C_M)^n + {C_F}^n} = \frac{\vec{R}_M * (WR * C_M)^n}{(WR * C_M)^n + {C_F}^n} \quad (19)$$

because flickering samples in isolation produce no response (i.e., $\vec{R}_F = 0$). The band-pass filtered noise stimuli of Sheliga et al. (2016) were modelled as pure sine waves, whose contrast and SF were set equal to the RMS contrast and central SF of the band-pass filtered noise stimuli used in the study. Parameters WR and n were calculated from Equations 13 and 15 using best fit values from Table 1.1. The predictions (no parameters were fit to the noise data) are shown by black solid (0.125 cpd data) and dotted (0.5 cpd data) lines in Figure 12. Thus, the relationship between SF and weight that we report here provides a good description of the suppressive effects attributed to flicker in our previous study (Sheliga et al., 2016).

A notable feature of the interaction we propose in Equation 1 is that there is no pooling of directional signals before competition. That is, the model predicts the same competition between gratings moving in the same direction, or even between moving and stationary gratings. In these cases, the model proposed here could capture phenomena some of us have previously described with contrast normalization (Quaia, Optican, & Cumming, 2017), and indeed contrast normalization could account for the terms in the denominator of Equation 1. The models have different structure because they explain complementary features of the data. The model formulation of Quaia et al. (2017) predicts responses as a function of contrast, fitting parameters that capture the weight of grating components to each stimulus. For the model we present here, the contrast response function is not represented explicitly because the measured response to single gratings at different contrasts are used as inputs to the model. Our model predicts the weights assigned to different spatiotemporal components. To predict responses to arbitrary one-dimensional patterns successfully, it will be necessary to combine both models, but additional empirical data will be required to constrain that more complete model. Nonetheless, our success in describing responses to certain simple broadband stimuli suggests that we have moved one step closer to the goal of a model that describes OFRs to any broadband stimulus. Motion cloud stimuli, used by others (Gekas, Meso, Masson, & Mamassian, 2017; Simoncini, Perrinet, Montagnini, Mamassian, & Masson, 2012), could be a very useful tool for this purpose, because they provide a principled way to control spatial and temporal bandwidths of the components.

Keywords: visual motion, spatial and temporal frequency tuning, nonlinear weighed summation model, surround inhibition

Acknowledgments

Supported by the Intramural Program of the National Eye Institute at the NIH.

Commercial relationships: none.
Corresponding author: B.M. Sheliga.
Email: bms@lsr.nei.nih.gov.
Address: Laboratory of Sensorimotor Research, National Institutes of Health, Bethesda, MD, USA.

Footnotes

[1]Experiment 4 studied 3-component cases, but the same principle applies – the pair moving together had one sign, the third component was given the opposite sign.
[2]The background color of Figure 2A-C is irrelevant and was chosen black in order to make WR values of different color clearly visible.
[3]The background color of Figure 3D-F is irrelevant and was chosen black in order to make n values of different color clearly visible.
[4]WR and n best-fit values: 1.16 and 4.4, 1.10 and 4.2, 1.19 and 4.5 for subjects BMS, FAM, JKM, respectively.
[5]The overall stimulus size in Sheliga et al. (2016) was very similar to that in Experiments 1, 2, and 4 of the current study (25°x25° vs. 22°x22°).
[6]Best-fit values: λ_W=0.31, μ_{HC}=−0.037, and σ_{HC}=0.50.
[7]Best-fit values: A_n=4.7, $A_{\sigma n}$=2.7, $\mu_{\sigma n}$=−2.5, and $\sigma_{\sigma n}$=1.4.
[8]Best-fit values: λ_W=0.33, μ_{HC}=-0.30, and σ_{HC}=0.12, σ_{TFw}= 13.4, A_{TFw}=11.4, and λ_{TFw}=0.12.
[9]Best-fit values: A_n=5.2, $A_{\sigma n}$=2.2, $\mu_{\sigma n}$=−2.6, and $\sigma_{\sigma n}$=1.4, μ_{TFn}= 20.5, and σ_{TFn}=16.7.

References

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433–436.

Collewijn, H., Van Der Mark, F., & Jansen, T. C. (1975). Precise recording of human eye movements. *Vision Research, 15*, 447–450.

Derrington, A. M., & Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. *Journal of Physiology, 357*, 219–240.

Gekas, N., Meso, A. I., Masson, G. S., & Mamassian, P. (2017). A normalization mechanism for estimating visual motion across speeds and scales. *Current Biology, 27*(10), 1514–1520 e1513.

Gellman, R. S., Carl, J. R., & Miles, F. A. (1990). Short latency ocular-following responses in man. *Visual Neuroscience, 5*(2), 107–122.

Hawken, M. J., Shapley, R. M., & Grosof, D. H. (1996). Temporal-frequency selectivity in monkey visual cortex. *Visual Neuroscience, 13*(3), 477–492.

Hays, A. V., Richmond, B. J., & Optican, L. M. (1982). A UNIX-based multiple process system for real-time data acquisition and control. *WESCON Conference Proceedings, 2*(1), 1–10.

Kumbhani, R. D., Saber, G. T., Majaj, N. J., Tailby, C., & Movshon, J. A. (2008). Contrast affects pattern direction selectivity in macaque MT neurons. *Program No. 460.26, 2008 Neuroscience Meeting Planner*. Washington, DC: Society for Neuroscience. Online.

Lui, L. L., Bourne, J. A., & Rosa, M. G. (2007). Spatial and temporal frequency selectivity of neurons in the middle temporal visual area of new world monkeys (Callithrix jacchus). *The European Journal of Neuroscience, 25*(6), 1780–1792.

Masson, G. S., & Perrinet, L. U. (2012). The behavioral receptive field underlying motion integration for primate tracking eye movements. *Neuroscience and Biobehavioral Reviews, 36*, 1–25.

Matsuura, K., Miura, K., Taki, M., Tabata, H., Inaba, N., & Kawano, K. et al. (2008). Ocular following responses of monkeys to the competing motions of two sinusoidal gratings. *Neuroscience Research, 61*(1), 56–69.

Miles, F. A. (1998). The neural processing of 3-D visual information: Evidence from eye movements. *The European Journal of Neuroscience, 10*(3), 811–822.

Miles, F. A., Kawano, K., & Optican, L. M. (1986). Short-latency ocular following responses of monkey. I. Dependence on temporospatial properties of visual input. *Journal of Neurophysiology, 56*(5), 1321–1354.

Miles, F. A., & Sheliga, B. M. (2010). Motion detection for reflexive tracking. In U. Ilg, & G. S. Masson (Eds.), *Dynamics of Visual Motion Processing: Neuronal, Behavioral and Computational Approaches* (pp. 141–160). New York: Springer-Verlag.

Miura, K., Inaba, N., Aoki, Y., & Kawano, K. (2014). Responses of MT/MST neurons elicited by dual-grating stimulus: Differences between areas MT and MST. *Program No. 726.709. 2014 Neuroscience Meeting Planner*. Washington, DC: Society for Neuroscience. Online.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442.

Quaia, C., Optican, L. M., & Cumming, B. G. (2017). Suppression and contrast normalization in motion processing. *Journal of Neuroscience, 37*(45), 11051–11066.

Robinson, D. A. (1963). A method of measuring eye movement using a scleral search coil in a magnetic field. *Institute of Electronic and Electrical Engineers: Transactions in Biomedical Engineering, BME-10,* 137–145.

Sheliga, B. M., Chen, K. J., FitzGibbon, E. J., & Miles, F. A. (2005). Initial ocular following in humans: A response to first-order motion energy. *Vision Research, 45*, 3307–3321.

Sheliga, B. M., FitzGibbon, E. J., & Miles, F. A. (2008). Human ocular following: Evidence that responses to large-field stimuli are limited by local and global inhibitory influences. *Progress in Brain Research, 171*, 237–243.

Sheliga, B. M., Kodaka, Y., FitzGibbon, E. J., & Miles, F. A. (2006). Human ocular following initiated by competing image motions: Evidence for a winner-take-all mechanism. *Vision Research, 46*(13), 2041–2060.

Sheliga, B. M., Quaia, C., Cumming, B. G., & Fitzgibbon, E. J. (2012). Spatial summation properties of the human ocular following response (OFR): Dependence upon the spatial frequency of the stimulus. *Vision Research, 68*, 1–13.

Sheliga, B. M., Quaia, C., FitzGibbon, E. J., & Cumming, B. G. (2016). Ocular-following responses to white noise stimuli in humans reveal a novel nonlinearity that results from temporal sampling. *Journal of Vision, 16*(1), 8.

Sheliga, B. M., Quaia, C., FitzGibbon, E. J., & Cumming, B. G. (2018). Short-latency ocular-following responses: Weighted average mechanism predicts the outcome of a competition between two sine wave gratings moving in opposite directions. *Program No. 220.05, 2018 Neuroscience Meeting Planner*. San Diego, CA: Society for Neuroscience. Online.

Simoncini, C., Perrinet, L. U., Montagnini, A., Mamassian, P., & Masson, G. S. (2012). More

is not always better: Adaptive gain control explains dissociation between perception and action. *Nature Neuroscience, 15*(11), 1596–1603.

Yang, D. S., FitzGibbon, E. J., & Miles, F. A. (2003). Short-latency disparity-vergence eye movements in humans: Sensitivity to simulated orthogonal tropias. *Vision Research, 43*(4), 431–443.

Appendix A

Figure A1 shows the results for one SF pairing (0.07/0.59 cpd), in which we changed contrasts of both components, and the *WR* of two components was far from 1. Figure A1 also shows the fits provided by the two models. Figure A1B zooms in on the relevant part of this plot (two-component responses).

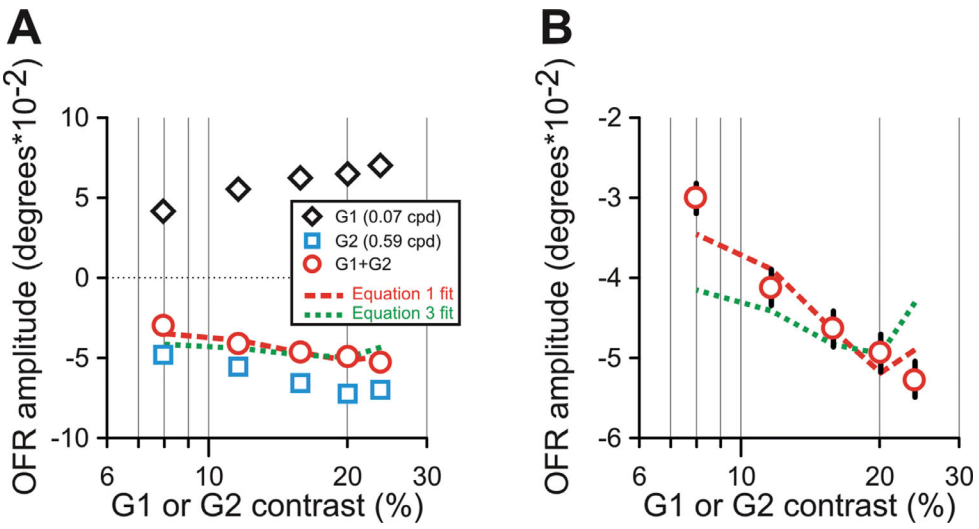

Figure A1. Experiment 1. Pilot experiments. Subject BMS. (A) Mean OFR amplitude in a control experiment designed to test predictions of Equations 1 and 3. Dependence on contrast of sine wave gratings. Symbols, data; dashed line, Equation 1 fit; dotted line, Equation 3 fit. (B) The same experiment as in A, but with an expanded ordinate showing two-component stimuli only. Black vertical lines are 68% confidence intervals.

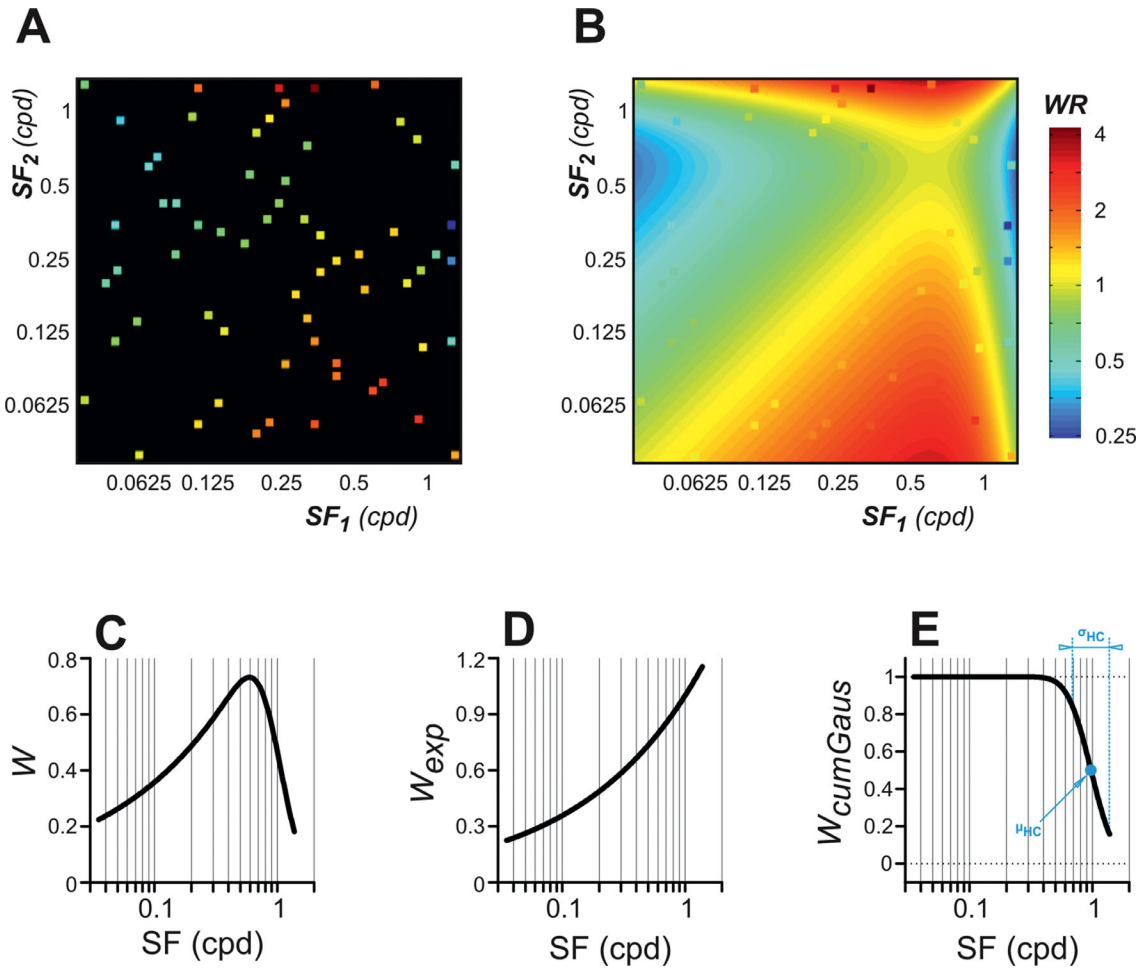

Figure B1. Experiment 1. Pilot experiments. Subject BMS. Dependence of parameter *WR* (from Equation 1) upon SF of the components. (A) *WRs* as a function of SF of sine wave gratings in two-component stimulus (color-coded small squares; see the color bar to the right of B; the background color (black) is irrelevant and was chosen to make *WRs* of different colors clearly visible. (B) As in A, but the background is replaced by the *WR* surface ($WR = \frac{W_1}{W_2}$) predicted by Equation 5 (plotted in C). The SF tuning in Equation 5 is the product of two functions: an exponential (D; W_{exp}; Equation 6) and an inverted cumulative Gaussian function (E; $W_{cumGaus}$; Equation 7).

Appendix B

Figure B1A shows WR (color-coded small squares; see color bar in Figure B1B) as a function of SF_1 and SF_2 in a two-component stimulus. We found that the value of WR was well-described by a separable function of SF ($r^2 = 0.917$). Figure B1C shows the function relating W_i to SF_i, and Figure B1B shows the resulting surface $WR = \frac{W_1}{W_2}$, with the data squares superimposed. The function relating SF to W was constructed from the product of two functions: an exponential (Equation 6; Figure B1D) and an inverted Cumulative Gaussian function (Equation 7; Figure B1E), where λ_W, μ_{HC}, and σ_{HC} are three free parameters.[6] μ_{HC} is the high cutoff frequency, where the function in Figure B1E is 0.5 (mean of the Gaussian function). σ_{HC} describes the steepness of this cutoff.

Appendix C

Figure C1 plots n (color-coded small squares; see color bar in Figure C1B) as a function of SF_1 and SF_2 in a two-component stimulus. There are two notable features in Figure C1A. First, n is largest close to the identity line (when the SF ratio is close to 1) and falls off smoothly as the SF ratio increases. We described this with a Gaussian of the *log* SF ratio

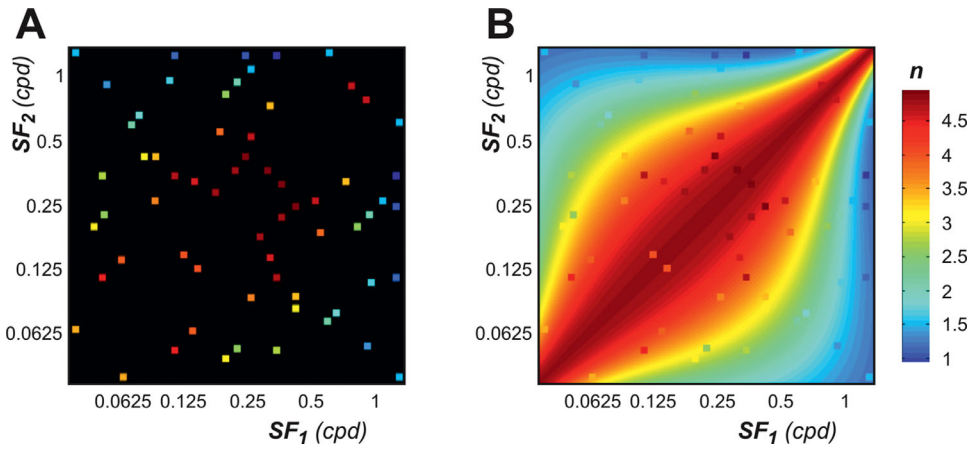

Figure C1. Experiment 1. Pilot experiments. Subject BMS. Dependence of parameter *n* (from Equation 1) upon SF of the components. (A) Values of *n* (color-coded small squares; see a color bar to the right of B) as a function of SF of sine wave gratings in two-component stimulus; the background color (black) is irrelevant and was chosen to make *n* of different colors clearly visible. (B) As in A, but the background is replaced by the surface predicted by Equation 10.

(Equation 8). Second, *n* decreases more rapidly with the SF ratio at high and low frequency, with the slowest decline for pairs where the geometric mean SF is ~0.25 cpd, close to the optimal SF for driving the OFR with single gratings. We were able to describe this feature successfully by allowing the value of sigma in Equation 8 to be a Gaussian function of the log of the product of SFs (Equation 9). Equation 10 provides a good fit ($r^2 = 0.909$) for the relationship between SF_1, SF_2 and *n*, using four free parameters (see the resulting surface in Figure C1B).[7]

Appendix D

Pilot experiments in subject BMS included seven different SF pairings (see legend of Figure D1). In a pairing, the TF of one component varied, whereas the TF of the other component was fixed (6.25, $18\frac{3}{4}$, or 25 Hz in different SF pairings). In total, data were collected from 22 SF_1TF_1, SF_2TF_2 pairings, and for each pairing we recorded the OFRs to three to five contrast combinations (from 4 to 36%). The values of *WR* and *n* fit with Equation 1 are shown by symbols in Figures D1A and D1D, respectively.

The effect of TF on the WR

Figure D1A plots *WR* as a function of TF for seven different pairs of SF (each pair shown with a different color). It is clear that the weight assigned to a component declines as the TF moves away from the optimal. Changes in the absolute value of *WR* make it difficult to determine if this is a separable function of SF and TF. To clarify this, we exploit the fact that for 5 out of 7 SF pairings there was a condition in which $TF_1 = TF_2 = 18\frac{3}{4}$ Hz. Figure D1B replots the data from Figure D1A normalized by the value of *WR* in this condition. That is, for the case where $TF_1 = TF_2 = 18\frac{3}{4}$ Hz, any differences in the *WR* can be attributed to the values of SF alone. If the *WR* is a separable function of SF and TF, then after this normalization a single function should describe the effect of TF. Although the dependence on TF is similar in these normalized curves, there are significant differences, especially at low TF. The error bars (68% confidence intervals) in Figure D1B are in most cases smaller than the data points. We find that these differences can be well described by a simple interaction between SF and TF (Equation 11), where *WR* is a Gaussian function of TF, but the TF at which *WR* peaks (i.e., optimal TF) depends on SF (Figure D1C; Equation 12). We choose an exponential in Equation 12 because it is well-behaved as the SF becomes large. The data over the range we observe could equally well be described by a straight line. We can now summarize all the factors influencing a component's weight by combining Equations 5 and 11; see Equation 13. Colored solid lines in Figure D1A show how well ($r^2 = 0.951$) this model fits the pilot dataset.[8] Note that the fits are also very good for pairings in which TF of neither component equaled $18\frac{3}{4}$ Hz (black dotted and grey dashed lines).

The effect of TF on *n*

Figure D1D plots *n* as a function of TF, for the same seven SF pairings as described. For the five pairs that included $TF_1 = TF_2 = 18\frac{3}{4}$ Hz, Figure D1E plots *n* normalized by the value for that condition.

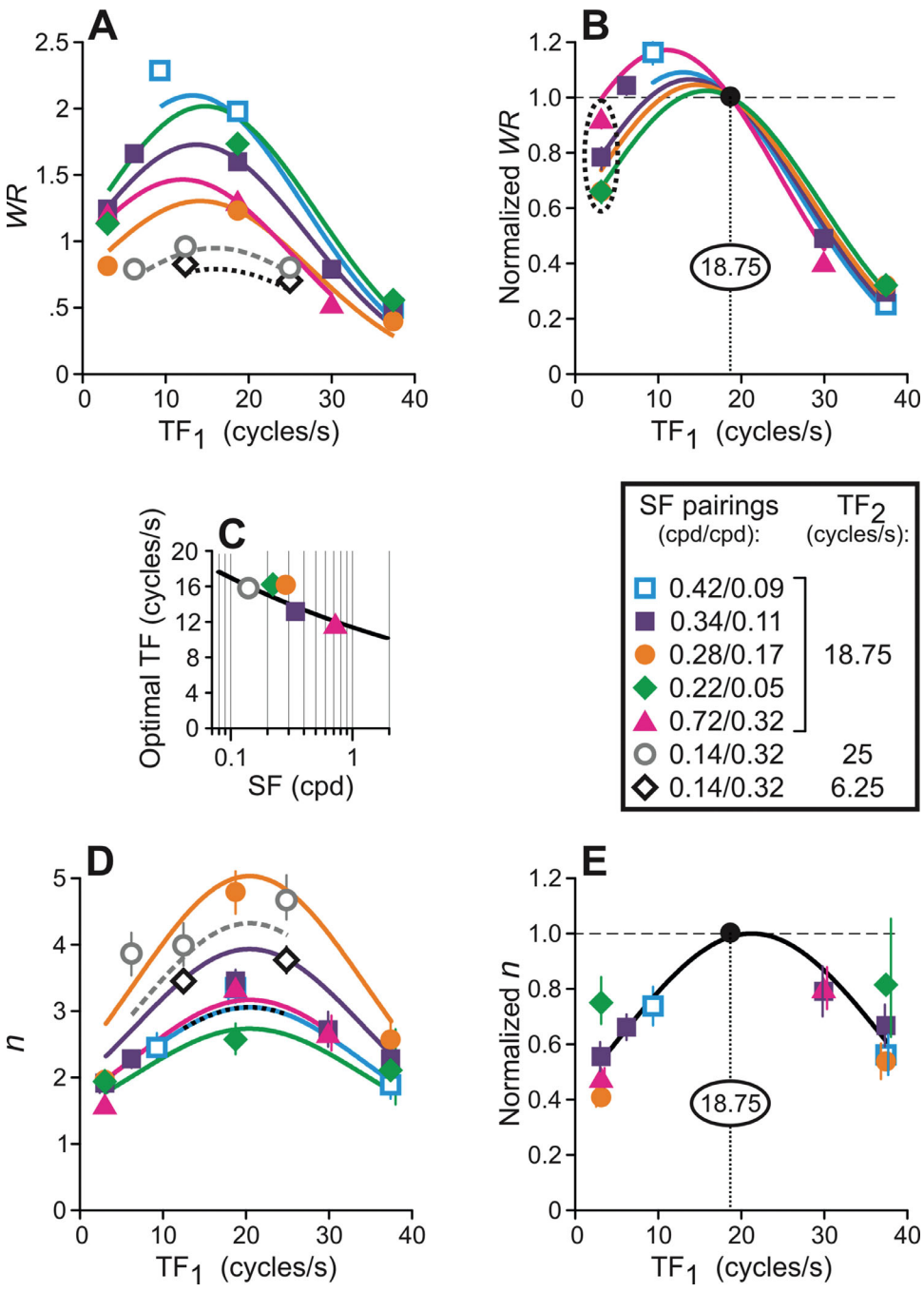

Figure D1. Experiment 2. Pilot experiments. Subject BMS. Dependence of *WR* (A–C) and *n* (D, E) on TF of the components. (A) *WRs* as a function of TF_1. Symbols, data; lines, Equation 13 fits. (B) Same as (A), but *WRs* were normalized with respect to $TF_1 = TF_2 = 18\frac{3}{4}$ Hz condition, black filled circle. Symbols, data; lines, Equations 11 and 12 fits. Black dotted oval highlights substantial differences in WR for the same TF, suggesting some dependence on SF values. (C) Optimal TF (i.e., at which *WR* peaks) as a function of SF: Equation 12 (black solid line). Symbols, optimal TFs for 5 (out of 7) SF pairs when *WRs* for each SF pair were fitted by a separate Gaussian. Data for two remaining SF pairs (blue squares and black diamonds) were poorly constrained and, therefore, not fitted with a Gaussian. (D) Values of *n* as a function of TF_1. Symbols, data; lines, Equation 15 fits. (E) Normalized values of *n* as a function of TF_1. The *n* values were normalized with respect to $TF_1 = TF_2 = 18\frac{3}{4}$ Hz condition, black filled circle; symbols, data; lines, Equation 14 fit. Thin colored vertical lines: 68% confidence intervals (bootstrapping). In many cases, these intervals were smaller than a symbol size (especially, for *WR* data) and, therefore, not visible on the graphs.

As with the *WR*, the value of *n* falls as TF deviates from the optimal. This effect is similar for all SF pairings except one (0.22/0.05 cpd), shown in green. This pair was also associated with a larger confidence interval than the other cases (because the response amplitudes were smaller), so we felt that a separable function of SF and TF would suffice to describe these data. We fit this with a Gaussian function of TF (Equation 14). The black solid line in Figure D1E shows this fit. Combined with Equation 10, this allows us to summarize the factors influencing *n* for any spatiotemporal component: Equation 15. Colored solid lines in Figure D1D show how well ($r^2 = 0.749$) this complete model fits the whole pilot dataset.[9]