Received: 17 February 2020 | Revised: 29 November 2020 | Accepted: 26 December 2020
DOI: 10.1002/hbm.25339

RESEARCH ARTICLE

# The dual nature of the BOLD signal: Responses in visual area hMT+ reflect both input properties and perceptual decision

Teresa Sousa[1,2,3] | João V. Duarte[1,2] | Gabriel N. Costa[1,2] | Valentin G. Kemper[3] | Ricardo Martins[1,2] | Rainer Goebel[3,4] | Miguel Castelo-Branco[1,2,3,5]

[1]Coimbra Institute for Biomedical Imaging and Translational Research (CIBIT), University of Coimbra, Coimbra, Portugal

[2]Institute of Nuclear Sciences Applied to Health (ICNAS), University of Coimbra, Coimbra, Portugal

[3]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, University of Maastricht, Maastricht, Netherlands

[4]Department of Vision and Cognition, Netherlands Institute for Neuroscience, Royal Netherlands Academy of Arts and Sciences (KNAW), Amsterdam, Netherlands

[5]Faculty of Medicine, University of Coimbra, Coimbra, Portugal

**Correspondence**
Miguel Castelo-Branco, ICNAS – Instituto de Ciências Nucleares Aplicadas à Saúde, Pólo das Ciências da Saúde Universidade de Coimbra, Azinhaga de Santa Comba, 3000-548 Coimbra, Portugal.
Email: mcbranco@fmed.uc.pt

**Funding information**
Fundação Bial, Grant/Award Numbers: 207/16, 252/18; Fundação para a Ciência e a Tecnologia, Grant/Award Numbers: SFRH/BD/80735/2011, BIGDATIMAGE CENTRO-01-0145-FEDER-000016, COMPETE UID/04950/2017, PTDC/PSI-GER/1326/2020

## Abstract

Neuroimaging studies have suggested that hMT+ encodes global motion interpretation, but this contradicts the notion that BOLD activity mainly reflects neuronal input. While measuring fMRI responses at 7 Tesla, we used an ambiguous moving stimulus, yielding the perception of two incoherently moving surfaces—component motion—or only one coherently moving surface—pattern motion, to induce perceptual fluctuations and identify perceptual organization size-matched domains in hMT+. Then, moving gratings, exactly matching either the direction of component or pattern motion percepts of the ambiguous stimulus, were shown to the participants to investigate whether response properties reflect the input or decision. If hMT+ responses reflect the input, component motion domains (selective to incoherent percept) should show grating direction stimulus-dependent changes, unlike pattern motion domains (selective to the coherent percept). This hypothesis is based on the known direction-selective nature of inputs in component motion perceptual domains versus non-selectivity in pattern motion perceptual domains. The response amplitude of pattern motion domains did not change with grating direction (consistently with their non-selective input), in contrast to what happened for the component motion domains (consistently with their selective input). However, when we analyzed relative ratio measures they mirrored perceptual interpretation. These findings are consistent with the notion that patterns of BOLD responses reflect both sensory input and perceptual read-out.

KEYWORDS
ambiguous visual motion, BOLD mechanisms, direction-selective input, perceptual bistability

## 1 | INTRODUCTION

During continuous observation of a physically unchanging ambiguous stimulus, visual perception may alternate between competing interpretations, depending on the integration or segregation of visual cues. This phenomenon, known as multistable visual perception (Attneave, 1971; Leopold & Logothetis, 1999; Sterzer, Kleinschmidt, & Rees, 2009), allows us to experimentally dissect neuronal mechanisms related to perceptual interpretation from the ones related to the nature of visual input. Its study offers powerful insights into mechanisms of visual awareness, perceptual organization, and decision (Blake, Brascamp, & Heeger, 2014; Blake & Logothetis, 2002; Crick & Koch, 1998; Pomerantz & Kubovy, 1981).

Brain activity measured with functional magnetic resonance imaging (fMRI) includes response patterns that may be sensory and/or perceptually driven (Bartels, Logothetis, & Moutoussis, 2008; Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001). Accordingly, understanding the contribution of each one to the cortical response to visual stimuli, as measured with fMRI, is of great interest. It has been suggested that neuronal populations within the extrastriate cortex may alter their activity according to the perceptual interpretation of an observed stimulus, rather than its physical properties (Leopold & Logothetis, 1999; Sterzer et al., 2009). In the visual motion perception domain, it was initially suggested, based on animal studies, that there are both neurons in early visual processing streams reflecting the perceptual experience and neurons reflecting the nature of the retinal input (Logothetis & Schall, 1989). However, more recently, it was shown that the same cell can carry both types of signals depending upon the exact type of stimulus conflict (Maier, Logothetis, & Leopold, 2007). These questions are here explored taking advantage of high-field fMRI, which enhances the possibility of detecting functionally homogeneous neuronal ensembles that constitute fundamental computational units in the brain. It is particularly suited to investigate motion domains within the human visual motion complex (hMT+) (Emmerling et al., 2016; Schneider et al., 2019; Zimmermann et al., 2011). Importantly, this also allows us to investigate the nature of the blood-oxygen-level-dependent (BOLD) signal response, and to which extent it reflects regional input. The hMT+ region contains component neurons, which receive direction-selective input, and pattern-neuron populations, in which input comprises multiple directions (Movshon & Newsome, 1996; Orban, Kennedy, & Bullier, 1986; Rust, Mante, Simoncelli, & Movshon, 2006). Here, we hypothesized that if hMT+ responses reflect the neuronal input, component-neuron domains should show stimulus-dependent changes, unlike pattern-neuron populations.

hMT+ activity has been shown to reflect global motion perceptual content (Castelo-Branco et al., 2002). Furthermore, the conscious experience of a specific motion axis was recently demonstrated to be related to the level of activity in direction-selective clusters of this brain region (Schneider et al., 2019). Here, we investigated to what extent the hMT+ response, as measured by fMRI, reflects the sensory input and/or the perceptual output of visual motion. This remains an outstanding question given the two following observations: (a) fMRI signals reflect mainly synaptic activity and are strongly correlated with local field potential (LFP) signals (Logothetis et al., 2001); (b) it has been shown that the tuning of LFP signals in primate MT and MST (which in humans jointly correspond to hMT+) mainly reflects the component selectivity of the inputs of these regions (in particular from V1), instead of global motion direction (Khawaja, Tsui, & Pack, 2009).

In order to investigate to which extent fMRI signals reflect input processing versus perceptual output, a paradigm of bistable visual perception of ambiguous motion was used here. Ambiguous moving stimuli can be constructed from superimposed gratings, such as the moving plaids (Adelson & Movshon, 1982; Castelo-Branco et al., 2002; Rust et al., 2006; Sousa et al., 2018). Observers' perception spontaneously switches between an incoherent percept (component motion), segregating the visual cues in two moving surfaces, and a coherent percept (pattern motion), integrating the visual cues as one single moving surface. However, the highly overlapped local motion contours of this type of stimuli may artefactually bias the influence of local mechanisms on neuronal responses and perception (Adelson & Bergen, 1985; Adelson & Movshon, 1982; Alais, Wenderoth, & Burke, 1994). Thus, here we have used a roof-shaped ambiguous moving stimulus in a non-overlapping configuration whereby 1D components are presented to each visual hemifield (Wallach, 1935; Wuerger, Shapley, & Rubin, 1996), thereby requiring long-range integration. Accordingly, pattern or component motion requires interhemispheric binding or segregation, respectively (Sousa et al., 2019).

Animal and human-based fMRI studies have suggested the existence of functional domains selective to pattern or component motion perception within hMT+ (Purushothaman & Bradley, 2005; Schneider et al., 2019). We hypothesized that the neuronal response of these perceptual sub-domains' to different directions of motion depends on their neuronal input profile. This prediction is based on the notion that hMT+ cells that signal pattern motion are distinguished by having convergent excitatory input from component-selective cells with a wide range of preferred directions (Rust et al., 2006). Figure 1 depicts the variable tuning curves of strongly direction-selective neurons that converge to pattern-selective neurons preferring the vertical motion. This broad range of input directions for pattern selective cells contrasts with component cells and their inputs, which are strongly direction-selective given the strong direction-selective input from V1 (Movshon & Newsome, 1996; Orban et al., 1986). We have also hypothesized that the relative response ratio profiles of such domains convey relevant information to what extent hMT+ response to local grating stimuli predicts global percepts.

In this work, we mapped functional sub-domains within hMT+ related to coherent and incoherent motion interpretations. We found that when tested with different axes of motion, the response amplitude of each one is in agreement with the hMT+ input contribution for each percept, whereas the ratio of activity between both revealed the content of perception. Thus, we suggest that, based on different metrics, both the sensory input and the perceptual output can be decoded from these functional domains.

## 2 | METHODS

### 2.1 | Participants

Ten healthy participants (six males; 28.7 ± 7.6 years; 9 right-handed) with normal or corrected-to-normal vision participated in this study. Participants gave informed consent and were paid for their participation.

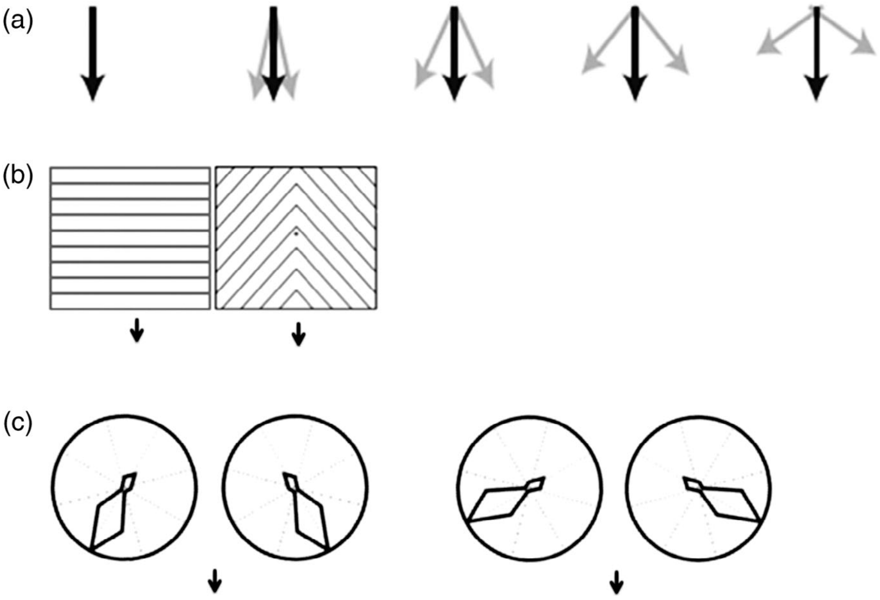

**FIGURE 1** Pattern-selective cells receive highly variable motion direction-selective inputs. (a) The same vertical pattern direction vector is consistent with the activation of inputs (gray vectors) with variable directional tuning. (b) Two distinct stimuli yielding the same type of pattern direction irrespective of component motion. (c) Pairs of distinct tuning curves from component-selective input neurons leading to similar pattern direction (vertical)

All experimental procedures followed the principles expressed in the Declaration of Helsinki and were conducted with approval from the Ethical Committees of Coimbra University and the Faculty of Psychology and Neuroscience of Maastricht University.

## 2.2 | Experimental design

All participants performed an ambiguous visual motion perception task during which behavioral and fMRI data were acquired. The scanning session included the acquisition of a standard structural MRI sequence and fMRI data (hMT+ functional localizer and ambiguous stimulus presentation). Prior to the scanning, a familiarization session was performed offline. It helped participants to get used to the ambiguous stimulus and allowed us to verify whether robust perceptual bistability was present. The minimum average duration of 6 s per each type of motion percept was set as eligibility criteria for the imaging study.

Stimuli were created with MATLAB (version 2016a; The Mathworks, Inc.), using the Psychophysics Toolbox version 3.0 extensions (Brainard, 1997), and displayed on a screen located 99 cm away from the participant (screen size: 17.2° × 10.4° [horizontal × vertical]; stimuli size: 11° × 10°; projection display: resolution of 1920 × 1080 and refresh rate of 60 Hz). Imaging data were acquired with a Siemens MAGNETOM 7 T scanner (Siemens; Erlangen, Germany) and a 32-channel head-coil (Nova Medical Inc.; Wilmington, MA) accessible at the Faculty of Psychology and Neuroscience of Maastricht University facilities. A response box device compatible with the magnetic resonance system was used to collect the perceptual reports from participants (Current Designs, 4-button response device, Philadelphia). Data processing and analyses were performed using BrainVoyager QX (version 2.8.4; Brain Innovation, Maastricht, The Netherlands) and MATLAB. Statistical analyses were performed with the IBM (Armank, 169 NY) SPSS Statistics 22.0 software package.

## 2.3 | Stimuli

### 2.3.1 | hMT+ functional localizer

Eight visual motion conditions (both directions of motion along axes of 0°, 45°, 90°, and 135°) were randomly interleaved with a no-motion condition and repeated nine times across three stimulation runs with 196 volumes each. Motion conditions consisted of black moving lines on a white background (contrast: 100%; motion speed: 3°/s; duty cycle: 6%; spatial frequency: 0.6 cycle/°; stimulus visual angle: 11° × 10° – horizontal × vertical). A central blue cross (visual angle: 0.2°) at the center of the image was used as a fixation target. During the no-motion condition, only the fixation target was shown. Each condition lasted 8 s. The design of the localizer stimulus is based on a previous experiment where it was shown that brain responses to different axes of motion are mainly confined to the MT portion of the hMT+ complex (Zimmermann et al., 2011).

### 2.3.2 | Ambiguous visual motion

A roof-shaped ambiguous moving stimulus in a non-overlapping configuration whereby 1D components are presented to each visual hemifield was used (Figure 2a). The perceived pattern of motion (Figure 2b) alternated between coherent (the stimulus was perceived as a single roof-shaped surface covering both hemifields and moving downward) and incoherent (the stimulus was perceived as two separate surfaces, one in each visual hemifield, moving inward).

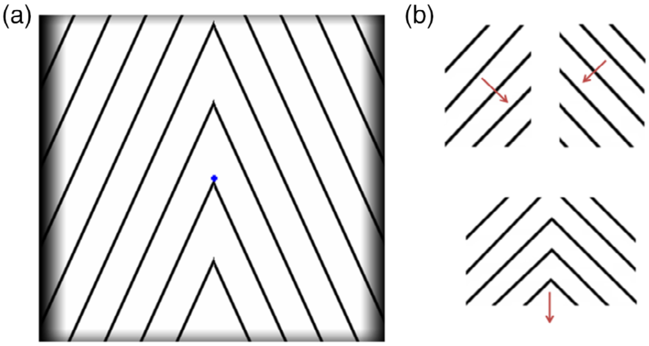

**FIGURE 2** Components of the ambiguous stimulus. A roof-shaped ambiguous moving stimulus in a non-overlapping configuration whereby 1D components are presented to each visual hemifield was used in this study (a). The 1D moving components could be perceived as two separate surfaces, one in each visual hemifield (incoherent percept—component motion) or integrated as a single roof-shaped surface covering both hemifields (coherent percept—pattern motion) (b). Arrows depict the direction of the ambiguous components (top) and of the overall integrated pattern motion (bottom). The amplified difference in phase in (b) represents only a perceptual feature of the illusion, as the physical stimulus was constant

The stimulus design was based on the original description of Wallach (Wallach, 1935; Wuerger et al., 1996). It consisted of oblique black lines equally separated forming an inverted V-shape on a white background continuously moving (orientation: ±25° relative to *x*-axis; contrast: 100%; motion speed: 3°/s; duty cycle: 6%; spatial frequency: 0.6 cycle/°; stimulus visual angle: 11° × 10° – horizontal × vertical) with a central blue fixation cross (visual angle: 0.2°). In order to optimize the presence of more stable percepts, the stimulus was prepared with a slight offset of 0.06° between the two 1D components, and the line terminations on the stimulus border were smoothed using a mask with a central aperture (9.9° × 10°), superimposed on the roof-shaped surface (Lorenceau & Shiffrar, 1992). The mask was prepared using a bi-dimensional squared Gaussian kernel (width and height: 0.6° ± 0.3°). Participants fixated the central cross, leading to relatively independent stimulus processing in each visual hemifield.

Importantly, our stimulus was physically constant, with unchanging luminance, and without perceptual bias related to motion energy. It was designed to reduce the contrast of intersections/terminations. In fact, in our case, the contrast is zero, and there are no intersections but just terminations. We used the strategy explained by (Stoner & Albright, 1996) to avoid luminance artifacts and unambiguous motion information. Luminance confounds might lead to a “blob tracking mechanism” that was minimized in our experiment, given the null contrast at intersections. Although the stimulus might contain rather small Fourier components away from component motion energy, they are not determining the percept, which depends instead on surface segregation cues, as previously shown (Stoner & Albright, 1996).

Perceptual tests consisted of four runs with 180 volumes each, with blocks of ambiguous motion for 44 to 90 s (4 repetitions per run), preceded by periods of 10–16 s with the static roof-shaped figure.

## 2.4 | Imaging data acquisition and processing

A T1-weighted magnetization prepared rapid acquisition gradient echo (3D-MPRAGE) sequence was applied to acquire structural brain images for anatomical reference (256 sagittal slices; isotropic resolution of 0.6 mm; repetition time (TR) = 3100 ms; echo time (TE) = 2.52 ms; flip angle = 5°; matrix = 384 × 384). Additionally, gradient echo proton-density images (same parameters as 3D-MPRAGE, except TR = 1440 ms) were acquired for intensity inhomogeneities correction. High-resolution functional images were then acquired using T2-weighted gradient echo echo-planar imaging (2D GE-EPI) (28 slices; isotropic resolution of 0.8 mm; TR = 2000 ms; TE = 25.6 ms; flip angle = 69°; matrix = 186 × 186). Five functional volumes with reverse encoding direction were also acquired to correct for EPI distortions. Furthermore, an initial slice positioning run was required to ensure the functional coverage of hMT+ (39 coronal slices; isotropic resolution of 1.6 mm; TR = 2000 ms; TE = 17.2 ms; flip angle = 70°; matrix = 88 × 88). The stimulation protocol of this run was as described in previous studies (Emmerling et al., 2016; Huk & Heeger, 2002; Zimmermann et al., 2011).

Anatomical data were corrected for spatial intensity inhomogeneities based on proton density information (Van de Moortele et al., 2009) and interpolated to an isotropic resolution of 0.8 mm using *sinc*-weighted interpolation to match the resolution of the functional images. Then, data were normalized to AC-PC space (Talairach & Tournoux, 1988). 3D body motion correction, aligning all subsequent functional runs to the closest one to the anatomical scans, and temporal high-pass filtering (GLM-Fourier with two cycles sine/cosine per run, including linear trend removal) were applied to the functional data. Possible distortions were also corrected, based on information of recorded functional volumes with opposite phase encoding (Andersson, Skare, & Ashburner, 2003). The functional images were then co-registered to the 3D anatomical data.

## 2.5 | Data analysis

### 2.5.1 | Behavioral data analysis

Participants' reports were analyzed to study bistable perceptual dynamics. The mean duration of each percept and the perceptual switch probability across time were estimated for each individual. The probability density function of percept duration was applied using a time-window from zero to the longest percept duration. Gamma and lognormal distributions were fitted to the data using the maximum likelihood method (Borsellino, Marco, Allazetta, Rinesi, & Bartolini, 1972; Zhou, Gao, White, Yao, & Merk, 2004). The goodness of fit was assessed using the Kolmogorov–Smirnov test.

### 2.5.2 | Imaging data analysis

Neuronal responses to the localizer and ambiguous stimuli were assessed by applying a general linear model (GLM) analysis (Kutner et al., 2005). Predictors for each condition were created by convolving the stimulation blocks with a standard hemodynamic response function (Friston, Josephs, Rees, & Turner, 1998). For the localizer runs, four predictors were created according to each axis of motion of the stimulation conditions. In the ambiguous runs, two distinct predictors were created, representing the two perceptual states reported by the participants. All runs of each stimulation protocol were analyzed together, at the single-subject level, using fixed-effects GLM and including the motion correction parameters as confound predictors. Data were normalized with *z*-transformation and corrected for serial correlations with a second-order autoregressive method (Lenoski et al., 2008). Data from one participant were excluded from imaging analysis due to highly unstable percepts.

First, the hMT+ complex was defined per participant as the voxels in the middle temporal region responding significantly, at *q* (FDR) = 0.05, to the contrast between motion and static conditions in the localizer. Then, hMT+ sub-domains selective for coherent and incoherent motion perception were mapped based on the ambiguous runs. The *t* map resulting from significant response to each motion percept was created using the static condition as the baseline. hMT+ voxels with a preference for coherent and incoherent percepts were then determined based on the winner map between both, restricted to the hMT+ ROIs. Afterward, voxels with the strongest preference for each percept were selected to define the coherent and incoherent perceptual sub-domains. Perceptual sub-domains were defined to be size-matched, which was determined as the minimum number of voxels responding significantly ($p \leq .05$) to the contrast between percepts.

Finally, we tested the differential responses of each perceptual domain to the exact direction of the ambiguous components, and vertical pattern motion, resulting from the integration of the two component-directions (Figure 2b). Gratings matching either the direction of the pattern or the component motion percepts of the ambiguous stimulus were used. Importantly, the definition of the perceptual sub-domains and directionally varying stimuli conditions used to test perceptual domains' responses are independent, as in one case we used plaids and in the other gratings, respectively. The hMT+ perceptual sub-domains response properties were studied to test to what extent the most selective voxels reflected pattern or component neuronal inputs.

### 2.5.3 | Statistical analysis

Group analyses were performed using a related-sample Wilcoxon signed-rank test. We tested the null hypothesis of no differences in the response of each hMT+ perceptual sub-domain (coherent and incoherent) to grating motion directions corresponding to pattern or component motion of the ambiguous stimulus. The significance of the differences in the ratio between incoherent and coherent perceptual sub-domains responses to each axis of motion was also tested. The null hypotheses were rejected to $p < .05$.

## 3 | RESULTS

### 3.1 | Behavioral results

Participants perceived the ambiguous stimulus as pattern motion (coherent percept) for periods of 8.35 ± 0.96 s (mean ± *SEM*) alternating with component motion (incoherent percept) during 7.25 ± 0.70 s on average.

The gamma and lognormal distributions fit well each percept duration histogram, as revealed by the Kolmogorov–Smirnov test results. The test demonstrated no significant deviation, at *p-value* = .05, between the fitted distributions and each percept duration histogram. Additionally, percept duration histograms are similarly fitted by both distributions (see Figure 3 for an example).

### 3.2 | Definition of perceptual sub-domains in hMT+

The hMT+ complex was localized for each participant. The contrast between responses to the motion and static conditions of the localizer stimulus allowed us to define hMT+ at the right and left hemispheres for all participants. The resulting statistical maps revealed clear bilateral activations in the expected spatial location of hMT+ (see Figure 4 for an example). The defined right and left hMT+ ROIs included on average 512.89 ± 84.29 and 811.33 ± 102.45 voxels (isotropic resolution 0.8 mm), respectively. Then, the responses to the ambiguous stimulus were used to map hMT+ sub-domains selective for each motion percept (coherent or incoherent). hMT+ division into voxels subsets with higher activation during incoherent or coherent percept was the basis for defining such perceptual domains (see Figure S1 for an example).

Contrasting each type of motion percept with the static control we found that the hMT+ response to coherent and incoherent percepts was on average 1.41 ± 0.21 and 1.75 ± 0.23 (*beta* values), respectively. Furthermore, the contrast between both percepts revealed a higher number of hMT+ voxels with a stronger response during the incoherent than the coherent percept (414.20 ± 46.16 vs. 212.20 ± 47.15), as expected from the known distribution of component and pattern populations in hMT+. However, not all of them contributed significantly ($p \leq .05$) to differentiate both percepts, which is consistent with the notion that these domains are quasi-columnar. To isolate such quasi-columnar domains, ensuring specificity, and to prevent unbalanced analyses between large and small coherent/incoherent motion domains we identified across all subjects, a set with the most significant five voxels per hemisphere with larger responses for coherent than incoherent percepts. A matched approach was followed for the domains selective for incoherent percepts.

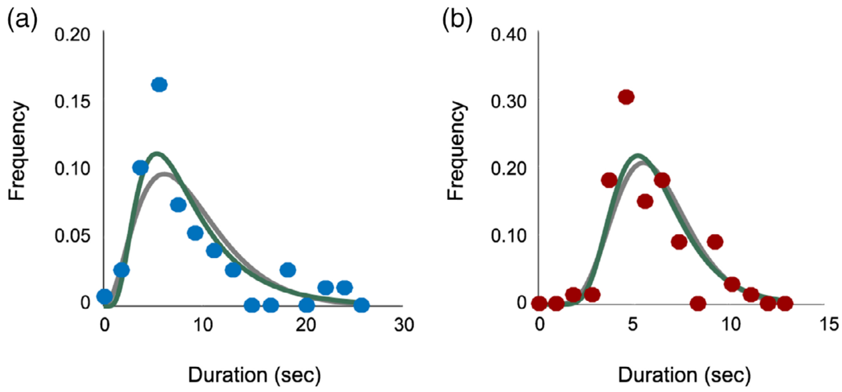

**FIGURE 3** Distribution of coherent and incoherent percepts durations during ambiguous stimulation. The dots represent the coherent (a) and incoherent (b) percept duration histograms from one participant, as an example. Gray and green lines illustrate gamma and lognormal distributions fitted to the data, respectively

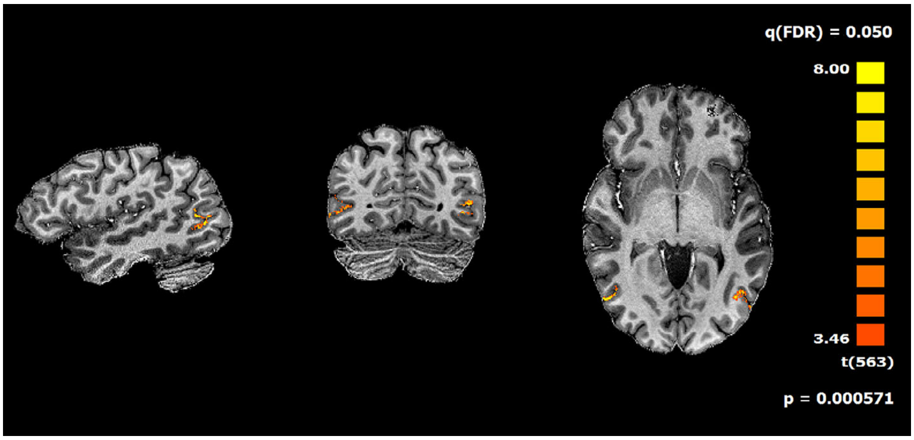

**FIGURE 4** Example of bilateral hMT+ localization. Activation map resulting from the contrast between motion and static conditions during the localizer experiment. Left and right hMT+ ROIs are shown at the same statistical threshold level (*q(FDR)* = 0.05)

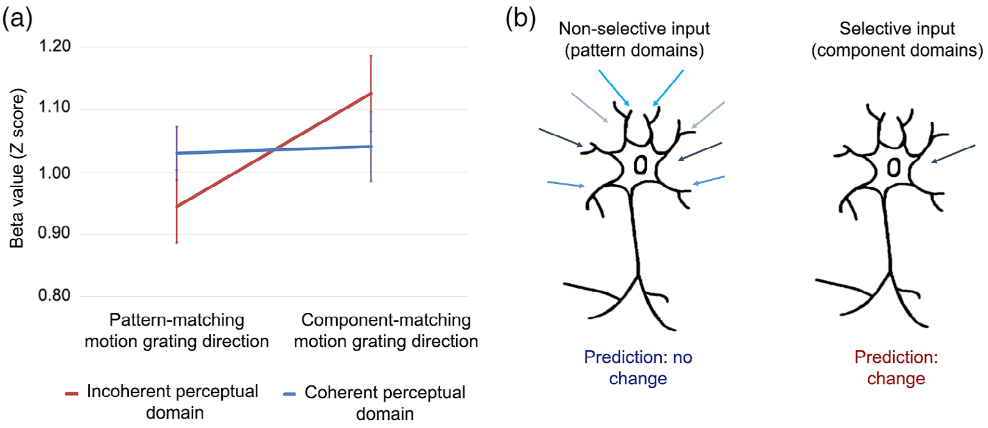

**FIGURE 5** Motion direction selectivity of the hMT+ perceptual domains. Responses in the coherent perceptual domain do not change with grating direction (consistently with its non-selective input), in contrast with the incoherent perceptual domain (consistently with the selective nature of inputs in component versus pattern domains). (a) Neuronal responses of the hMT+ perceptual sub-domains to moving gratings matching pattern or component motion direction (GLM beta values—group average). The perceptual ROIs with a preference for the incoherent percept responded significantly more to gratings matching the axes of motion of the components of the ambiguous stimulus than to the ones matching its global pattern of motion (*p* = .025). Error bars represent ± SEM. (b) hMT+ region contains component neurons, which receive direction-selective input and pattern-neuron populations, in which input comprises multiple directions. We hypothesized that if hMT+ responses reflect the input, component-neuron domains should show stimulus-dependent changes, unlike pattern-neuron populations

The incoherent perceptual subdomain exhibited an average *beta* estimation of 1.61 ± 0.07 and 0.71 ± 0.05 for the incoherent and coherent percepts, respectively. On the other hand, the coherent perceptual subdomain exhibited an average beta value of 0.44 ± 0.07 and 1.04 ± 0.09 for the incoherent and coherent percepts, respectively.

### 3.3 | Perceptual sub-domains selectivity for different axes of motion

The neuronal responses of the coherent and incoherent perceptual sub-domains were tested for the axes of motion of the ambiguous components and of the two components integration (as in pattern motion).

The response profiles for axes of motion matching either pattern and component motion varied according to each perceptual domain (Figure 5). Note that only the incoherent perceptual domains show modulation of responses (Figure 5a) with gratings of different directions, as expected from the input-based prediction that only these should be sensitive to a variation of component direction (Figure 5b). The domains selective to incoherent perception showed stronger responses for the gratings matching component than pattern motion direction (average *beta* estimation of 1.13 ± 0.05 and 0.95 ± 0.04, respectively; *p* = .025). On the other hand, coherent perceptual domains did not show any differential response to direction, and responded similarly to the different axes of motion conditions. In line with the notion that pattern neurons are not selective to the specific motion of their inputs, the average *beta* estimation for the pattern motion direction was 1.03 ± 0.06, and 1.04 ± 0.07 for the component motion direction (no significant differences, *p* = .59).

### 3.4 | Perceptual sub-domains' response ratio

We then investigated relative response profiles. When comparing tuning properties of both perceptual domains, as depicted in Figure 5, we found that the highest ratio response to the pattern motion direction was from the coherent domain (1.03 ± 0.06 vs. 0.95 ± 0.04, average *beta* estimations), whereas the highest ratio response to the component motion direction was from the incoherent domain (1.13 ± 0.05 vs. 1.04 ± 0.07, average *beta* estimations). The analysis of response ratios between incoherent and coherent perceptual sub-domains to each axis of motion revealed that this effect is significant at the group level (*p* = .015). The response ratios were on average 0.93 ± 0.04 and 1.10 ± 0.04 for the axes of motion matching pattern and component motion, respectively (Figure 6). Ratio values below and above one depict whether the coherent or the incoherent perceptual domain is more responsive, respectively.

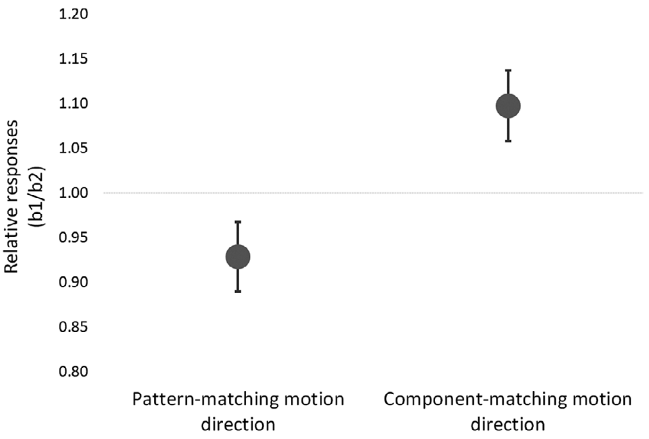

**FIGURE 6** Perceptual sub-domains response ratios. Ratio between incoherent and coherent perceptual domains responses (*beta* values *b1* and *b2*) to moving gratings matching either component or pattern motion directions. The group values, lower and higher than one, reveal that coherent domains presented the highest response ratio for moving gratings matching the pattern motion direction, whereas incoherent domains presented the highest response ratio for moving gratings matching the component motion direction (significant differences with *p* = .015). Error bars represent ± SEM

## 4 | DISCUSSION

In this study, we investigated whether size-matched sub-domains related to perceptual content can be found within hMT+ and whether they mostly reflect the nature of the input or decision. To answer these questions, we used a bistable perception paradigm and assessed neuronal responses during ambiguous visual motion stimulation, as well as for moving grating stimuli matching either pattern or component motion direction of the ambiguous stimulus.

During ambiguous stimulation, the perceptual dynamics were consistent with typical gamma and lognormal distributions (Borsellino et al., 1972; Duarte, Costa, Martins, & Castelo-Branco, 2017; Kaneoke, Urakawa, Hirai, Kakigi, & Murakami, 2009; Leopold, Wilke, Maier, & Logothetis, 2002), as expected in the presence of competing neuronal representations (Brascamp, van Ee, Pestman, & van den Berg, 2005; Logothetis & Schall, 1989; Zhou et al., 2004). Our perceptual data shows that there is no perceptual bias, as we found balanced percepts between inward and downward interpretation (no significant difference between both durations) (see also Costa, Duarte, Martins, Wibral, & Castelo-Branco, 2017).

In line with previous reports, we found that the incoherent percept induced higher hMT+ activity than the coherent percept (Castelo-Branco et al., 2002; Duarte et al., 2017; Sousa et al., 2018). This possibly reflects a larger number or size of functional units responding to the component motion. Indeed, the perception of component motion involves two moving surfaces, whereas the perception of pattern motion involves only one. Perceiving incoherent motion may be associated with activation of a larger pool of neurons, which could account for a higher BOLD signal, if one assumes a linear relationship between neuronal activity and the BOLD signal (Boynton et al., 1996; Logothetis et al., 2001). It is also consistent with the lower number of voxels that we found to prefer the coherent percept, compared to the incoherent one. Therefore, it was important to assess size-matched functional domains to test our hypothesis.

By contrasting brain responses to both motion percepts, we found functional sub-domains, within hMT+, differentially recruited according to the participants' perceptual state during ambiguous stimulation. It has been shown that overall activity in the hMT+ reflects global motion interpretation of perceptual bistability (Castelo-Branco et al., 2002). This 7 Tesla study goes a step further by investigating perceptual domain responses instead of looking at the overall activity. This allows us to disentangle between theories postulating that the BOLD signal mainly reflects input or instead perceptual readout. We first identified different hMT+ functional sub-domains modulated according to the visual motion perceptual content. Coherent perceptual domains encompass voxels highly activated by motion in the coherent percept and incoherent perceptual domains encompass voxels highly activated by motion in the incoherent percept. We then investigated whether the identified functional sub-domains could reflect hMT+ input, as suggested by the observations that LFP signals are a mirror of BOLD signal (Logothetis et al., 2001) and mainly reflect the nature of the neuronal input (Khawaja et al., 2009). Logothetis and Schall (1989) have suggested that some neurons might be recruited to respond to the early sensory input but not to the perceptual experience. On the other hand, Maier et al. (2007) have suggested that although the majority of feature-responsive neurons in the visual cortex can contribute to the resolution of sensory conflicts, this depends upon the exact type of stimulus conflict.

When tested for different axes of motion using grating stimuli matching either pattern or component motion directions, hMT+ perceptual sub-domains revealed different stimulus dependence profiles. The incoherent perceptual domain responses were revealed to be direction-selective, in contrast to the coherent perceptual domain. This asymmetry is consistent with the input hypothesis. Component-selective domains should respond in a stimulus-dependent fashion, because their input is direction-selective, unlike pattern motion domains, which receive input that is not direction-selective. There is substantial evidence for a two-stage mechanism of global motion computation. It involves a local component stage in highly direction-selective neurons in V1 (Movshon & Newsome, 1996) followed by a pattern stage in hMT+, where vectors comprising multiple directions can be combined due to the input non-selectivity of pattern responding neurons (Adelson & Movshon, 1982; Albright & Stoner, 1995; Movshon & Newsome, 1996; Rust et al., 2006). Unlike V1, hMT+ contains not only local motion-selective component neurons (around 40%) but also pattern neurons (around 25%) which can combine the component cells' outputs as moving surfaces (Movshon, Adelson, Gizzi, & Newsome, 1985). Therefore, the lack of selectivity for axes of motion in the coherent perceptual domain might be related to the varying tuning of input neurons (with selectivity for different axes of motion) involved in the coherent percept. The fact that a lower (pattern) neurons percentage is activated was controlled by analyzing domains with equally sized volumes. Our input-hypothesis data are consistent with the previous demonstration that BOLD signal reflects mainly LFP signals (Logothetis et al., 2001), which in turn has been shown to reflect afferent inputs, specifically for primate MT/MST (Khawaja et al., 2009). In light of this, the tuning properties of recorded perceptual sub-domains might reflect mainly the component-selective input within hMT+ and from V1.

We also investigated whether the ratio of activity between perceptual domains reflects the content of motion perception. The highest response to the pattern-matching motion direction was found in the coherent domain, whereas the highest response to the component-matching motion direction was found in the incoherent domain. This suggests that the ratio of the responses of hMT+ sub-domains to pattern and component-matching motion directions reflects either coherent or incoherent perception. Therefore, information about the perceptual content, and not solely of the nature of the input, is also present in population responses. The notion that one can retrieve in hMT+ both perceptual content and response matching axes of motion input support the reports of previous studies that suggested quasi columnar-level neuronal correlates of perceived direction in the hMT+ region (Schneider et al., 2019).

In sum, our study shows that bistable perception is reflected not only at a global level in the hMT+ region but also at the level of perception-related functional sub-domains. Such perceptual sub-domains exhibited response properties possibly reflecting both the nature of the hMT+ input, as well as perceptual content, as indexed by responses ratio. We suggest that, based on different metrics, both the sensory input and the perceptual output can be decoded from perceptual sub-domains of hMT+. Future studies investigating the inherent limitations of BOLD fMRI to discriminate between input signals from local computations will be further required.

## 5 | CONCLUSION

It remains controversial whether fMRI BOLD signals truly reflect the local regional activity. This debate stems from monkey studies showing that they might in fact mostly reflect the input. Here, we took advantage of the unique organization of visual hMT+ to directly address this question in humans. This region contains *component* neurons, which receive direction-selective input, and pattern -neuron populations, which input comprises multiple directions. We hypothesized that if responses reflect input, component -neuron domains should show stimulus direction -dependent changes, unlike pattern -neuron populations. The response amplitude of the perceptual domains is in agreement with the input hypothesis. However, when we analyzed relative ratio response ratios they mirrored perceptual interpretation. These findings unveil a dual nature for the BOLD signal and show that largely conveys information about the functional properties of regional inputs.

### ACKNOWLEDGMENTS

This research work was funded by the Portuguese Foundation for Science and Technology (FCT) (grants: COMPETE UID/NEU/04539/2013, COMPETE POCI-01-0145-FEDER-007440, BIGDATIMAGE CENTRO-01- 0145-FEDER-000016, Centro 2020 FEDER, COMPETE, PTDC/PSI-GER/1326/2020 and SFRH/BD/80735/2011) and by the BIAL projects 207/16 and 252/18.

### CONFLICT OF INTEREST

The authors have no conflict of interests.

## DATA AVAILABILITY STATEMENT

Data used in the study are available upon direct request. Conditions for its sharing involve the formalization of a research agreement. The data and code sharing adopted by the authors comply with the requirements of the funding body or institute, and with the institutional ethics approval. Parts of the data are confidential and additional ethical approval may be needed for reuse.

## ORCID

*Teresa Sousa* https://orcid.org/0000-0003-2652-3152
*João V. Duarte* https://orcid.org/0000-0001-8586-9554
*Gabriel N. Costa* https://orcid.org/0000-0001-9630-1155
*Valentin G. Kemper* https://orcid.org/0000-0003-0412-3086
*Ricardo Martins* https://orcid.org/0000-0001-7184-185X
*Rainer Goebel* https://orcid.org/0000-0003-1780-2467
*Miguel Castelo-Branco* https://orcid.org/0000-0003-4364-6373

## REFERENCES

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, *2*(2), 284–299. https://doi.org/10.1364/JOSAA.2.000284

Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, *300*(5892), 523–525. https://doi.org/10.1038/300523a0

Alais, D., Wenderoth, P., & Burke, D. (1994). The contribution of one-dimensional motion mechanisms to the perceived direction of drifting plaids and their aftereffects. *Vision Research*, *34*(14), 1823–1834. https://doi.org/10.1016/0042-6989(94)90307-7

Albright, T. D., & Stoner, G. R. (1995). Visual motion perception. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(7), 2433–2440. https://doi.org/10.1073/pnas.92.7.2433

Andersson, J. L. R., Skare, S., & Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. *NeuroImage*, *20*(2), 870–888.

Attneave, F. (1971). Multistability in perception. *Scientific American*, *225*(6), 62–71. https://doi.org/10.1038/scientificamerican1271-62

Bartels, A., Logothetis, N. K., & Moutoussis, K. (2008). fMRI and its interpretations: An illustration on directional selectivity in area V5/MT. *Trends in Neurosciences*, *31*(9), 444–453. https://doi.org/10.1016/j.tins.2008.06.004

Blake, R., Brascamp, J., & Heeger, D. J. (2014). Can binocular rivalry reveal neural correlates of consciousness? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1641), 20130211. https://doi.org/10.1098/rstb.2013.0211

Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience*, *3*, 13–21. https://doi.org/10.1038/nrn701

Borsellino, A., Marco, A., Allazetta, A., Rinesi, S., & Bartolini, B. (1972). Reversal time distribution in the perception of visual ambiguous stimuli. *Kybernetik*, *10*(3), 139–144.

Boynton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *The Journal of Neuroscience*, *16*(13), 4207–4221. https://doi.org/10.1523/JNEUROSCI.16-13-04207.1996.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Brascamp, J. W., van Ee, R., Pestman, W. R., & van den Berg, A. V. (2005). Distributions of alternation rates in various forms of bistable perception. *Journal of Vision*, *5*, 287–298. https://doi.org/10.1167/5.4.1

Castelo-Branco, M., Formisano, E., Backes, W., Zanella, F., Neuenschwander, S., Singer, W., & Goebel, R. (2002). Activity patterns in human motion-sensitive areas depend on the interpretation of global motion. *Proceedings of the National Academy of Sciences*, *99*(21), 13914–13919. https://doi.org/10.1073/pnas.202049999

Costa, G. N., Duarte, J. V., Martins, R., Wibral, M., & Castelo-Branco, M. (2017). Interhemispheric binding of ambiguous visual motion is associated with changes in Beta oscillatory activity but not with gamma range synchrony. *Journal of Cognitive Neuroscience*, *29*(11), 1829–1844.

Crick, F., & Koch, C. (1998). Consciousness and neuroscience. *Cerebral Cortex*, *8*(2), 97–107. https://doi.org/10.1093/cercor/8.2.97

Duarte, J. V., Costa, G. N., Martins, R., & Castelo-Branco, M. (2017). Pivotal role of hMT+ in long-range disambiguation of interhemispheric bistable surface motion. *Human Brain Mapping*, *38*(10), 4882–4897. https://doi.org/10.1002/hbm.23701

Emmerling T. C., Zimmermann J., Sorger B., Frost M. A., & Goebel R. (2016). Decoding the direction of imagined visual motion using 7 T ultra-high field fMRI. *NeuroImage*, *125*, 61–73. http://dx.doi.org/10.1016/j.neuroimage.2015.10.022.

Friston, K. J., Josephs, O., Rees, G., & Turner, R. (1998). Nonlinear event-related responses in fMRI. *Magnetic Resonance in Medicine*, *39*(1), 41–52.

Huk, A., & Heeger, D. (2002). Pattern-motion responses in human visual cortex. *Nature Neuroscience*, *5*(1), 72–75. https://doi.org/10.1038/nn774

Kaneoke, Y., Urakawa, T., Hirai, M., Kakigi, R., & Murakami, I. (2009). Neural basis of stable perception of an ambiguous apparent motion stimulus. *Neuroscience*, *159*, 150–160. https://doi.org/10.1016/j.neuroscience.2008.12.014

Khawaja, F. A., Tsui, J. M. G., & Pack, C. C. (2009). Pattern motion selectivity of spiking outputs and local field potentials in macaque visual cortex. *Journal of Neuroscience*, *29*(43), 13702–13709. https://doi.org/10.1523/JNEUROSCI.2844-09.2009

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models*. (5th Edition). New York: McGraw-Hill, Irwin.

Lenoski B., Baxter L. C., Karam L. J., Maisog J., Debbins J. (2008). On the Performance of Autocorrelation Estimation Algorithms for fMRI Analysis. *IEEE Journal of Selected Topics in Signal Processing*, *2*(6), 828–838. http://dx.doi.org/10.1109/jstsp.2008.2007819.

Leopold, D. A., & Logothetis, N. K. (1999). Multistable phenomena: Changing views in perception. *Trends in Cognitive Sciences*, *3*(7), 254–264. https://doi.org/10.1016/S1364-6613(99)01332-7

Leopold, D. A., Wilke, M., Maier, A., & Logothetis, N. K. (2002). Stable perception of visually ambiguous patterns. *Nature Neuroscience*, *5*, 605–609. https://doi.org/10.1038/nn0602-851

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(12), 150–157. https://doi.org/10.1038/35084005

Logothetis, N. K., & Schall, J. D. (1989). Neuronal correlates of subjective visual perception. *Science*, *245*, 761–763. https://doi.org/10.1126/science.2772635

Lorenceau, J., & Shiffrar, M. (1992). The influence of terminators on motion integration across space. *Vision Research*, *32*(2), 263–273. https://doi.org/10.1016/0042-6989(92)90137-8

Maier, A., Logothetis, N. K., & Leopold, D. A. (2007). Context-dependent perceptual modulation of single neurons in primate visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(13), 5620–5625. https://doi.org/10.1073/pnas.0608489104

Movshon, J. A., Adelson, E. H., Gizzi, M. S., & Newsome, W. T. (1985). The analysis of moving visual patterns. In C. Chagas, R. Gattass, & C. Gross (Eds.), *Pattern recognition mechanisms* (54th ed., pp. 117–151). Rome: Vatican Press.

Movshon, J. A., & Newsome, W. T. (1996). Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *Journal of Neuroscience*, *16*(23), 7733–7741. https://doi.org/10.1167/9.8.752

Orban, G. A., Kennedy, H., & Bullier, J. (1986). Velocity sensitivity and direction selectivity of neurons in areas V1 and V2 of the monkey:

Influence of eccentricity. *Journal of Neurophysiology*, *56*(2), 462–480. https://doi.org/10.1152/jn.1986.56.2.462

Pomerantz, J. R., & Kubovy, M. (1981). Perceptual organization: An overview. In *Perceptual Organization* (1st ed., pp. 423–456). Hillsdale: Lawrence Erlbaum Associates, Inc.

Purushothaman, G., & Bradley, D. C. (2005). Neural population code for fine perceptual decisions in area MT. *Nature Neuroscience*, *8*, 99–106. https://doi.org/10.1038/nn1373

Rust, N. C., Mante, V., Simoncelli, E. P., & Movshon, J. A. (2006). How MT cells analyze the motion of visual patterns. *Nature Neuroscience*, *9*(11), 1421–1431. https://doi.org/10.1038/nn1786

Schneider M., Kemper V. G., Emmerling T. C., De Martino F., Goebel R. (2019). Columnar clusters in the human motion complex reflect consciously perceived motion axis. *Proceedings of the National Academy of Sciences*, *116*(11), 5096–5101. http://dx.doi.org/10.1073/pnas.1814504116.

Sousa, T., Duarte, J. V., Costa, G. N., Kemper, V. G., Martins, R., Goebel, R., & Castelo-Branco, M. (2019). Tracking perceptual decision mechanisms through changes in interhemispheric functional connectivity in human visual cortex. *Scientific Reports*, *9*(1), 1242. https://doi.org/10.1038/s41598-018-37822-x

Sousa, T., Sayal, A., Duarte, J. V., Costa, G. N., Martins, R., & Castelo-Branco, M. (2018). Evidence for distinct levels of neural adaptation to both coherent and incoherently moving visual surfaces in visual area hMT+. *NeuroImage*, *179*, 540–547. https://doi.org/10.1016/j.neuroimage.2018.06.075

Sterzer, P., Kleinschmidt, A., & Rees, G. (2009). The neural bases of multistable perception. *Trends in Cognitive Sciences.*, *13*, 310–318. https://doi.org/10.1016/j.tics.2009.04.006

Stoner, G. R., & Albright, T. D. (1996). The interpretation of visual motion: Evidence for surface segmentation mechanisms. *Vision Research*, *36*(9), 1291–1310. https://doi.org/10.1016/0042-6989(95)00195-6

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: 3-D proportional system*, New York: Thieme Medical Publishers, Inc.

Van de Moortele P-F., Auerbach E. J., Olman C., Yacoub E., Uğurbil K., & Moeller S. (2009). T1 weighted brain images at 7 Tesla unbiased for Proton Density, T2* contrast and RF coil receive B1 sensitivity with simultaneous vessel visualization. *NeuroImage*, *46*(2), 432–446. http://dx.doi.org/10.1016/j.neuroimage.2009.02.009.

Wallach, H. (1935). Über visuell wahrgenommene Bewegungsrichtung. *Psychologische Forschnung*, *20*, 325–380.

Wuerger, S., Shapley, R., & Rubin, N. (1996). "On the visually perceived direction of motion" by Hans Wallach: 60 years later. *Perception*, *25* (11), 1317–1367.

Zhou, Y. H., Gao, J. B., White, K. D., Yao, K., & Merk, I. (2004). Perceptual dominance time distributions in multistable visual perception. *Biological Cybernetics*, *90*(4), 256–263. https://doi.org/10.1007/s00422-004-0472-8

Zimmermann, J., Goebel, R., de Martino, F., van de Moortele, P. F., Feinberg, D., Adriany, G., … Yacoub, E. (2011). Mapping the organization of axis of motion selective features in human area MT using high-field fMRI. *PLoS One*, *6*(12), e28716. https://doi.org/10.1371/journal.pone.0028716

## SUPPORTING INFORMATION

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

**How to cite this article:** Sousa T, Duarte JV, Costa GN, et al. The dual nature of the BOLD signal: Responses in visual area hMT+ reflect both input properties and perceptual decision. *Hum Brain Mapp*. 2021;1–10. https://doi.org/10.1002/hbm.25339