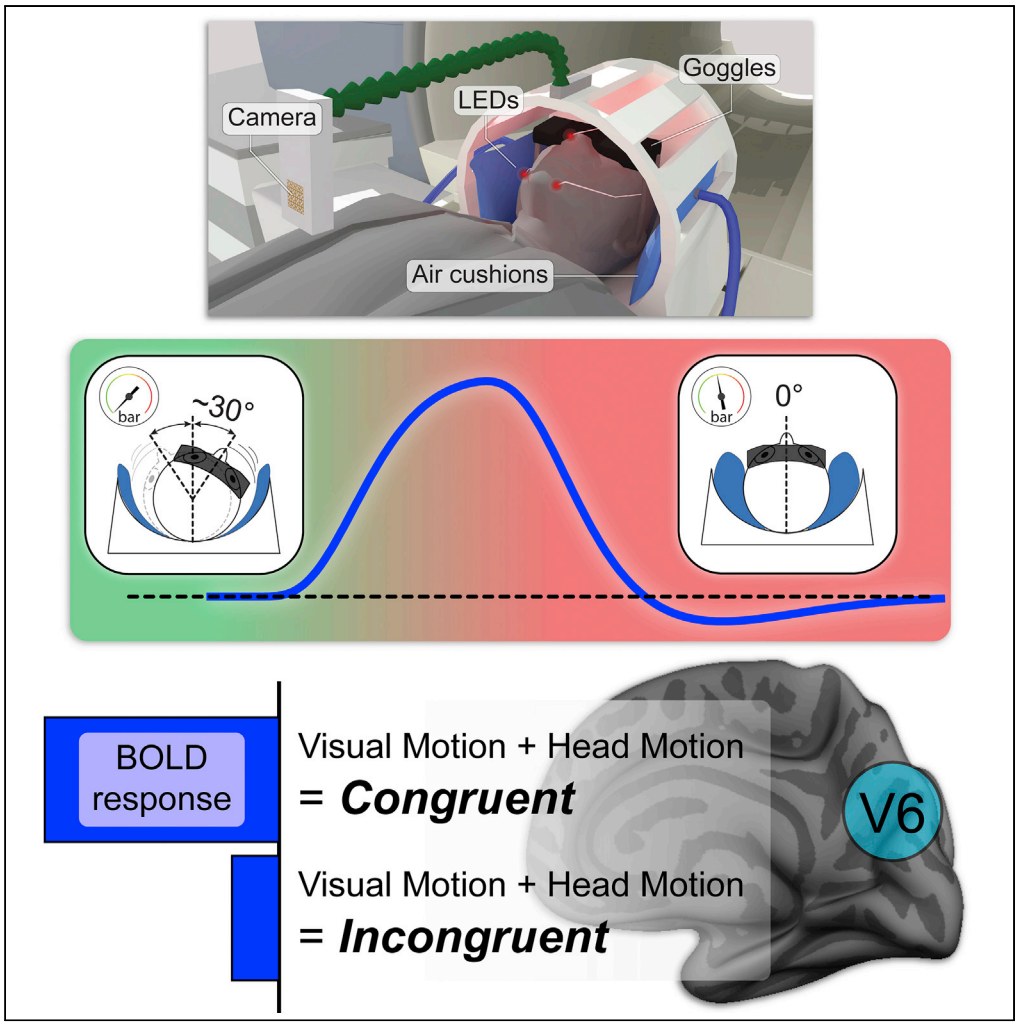
iScience

CellPress

Article

Human V6 Integrates Visual and Extra-Retinal Cues during Head-Induced Gaze Shifts



Andreas Schindler, Andreas Bartels

andreas.schindler@tuebingen.mpg.de (A.S.)
andreas.bartels@tuebingen.mpg.de (A.B.)

HIGHLIGHTS

Setup with head-mounted goggles and head movement during fMRI

Simulation of forward flow in stable or unstable world during head rotation

Human V6 integrates visual self-motion with head motion signals

Likely mediated by efference copy or proprioception as V6 lacks vestibular input

Schindler & Bartels, iScience 7, 191–197
September 28, 2018
https://doi.org/10.1016/j.isci.2018.09.004

iScience

CellPress

Article

Human V6 Integrates Visual and Extra-Retinal Cues during Head-Induced Gaze Shifts

Andreas Schindler[1,2,3,4,*] and Andreas Bartels[1,2,3,5,*]

SUMMARY

A key question in vision research concerns how the brain compensates for self-induced eye and head movements to form the world-centered, spatiotopic representations we perceive. Although human V3A and V6 integrate eye movements with vision, it is unclear which areas integrate head motion signals with visual retinotopic representations, as fMRI typically prevents head movement executions. Here we examined whether human early visual cortex V3A and V6 integrate these signals. A previously introduced paradigm allowed participant head movement during trials, but stabilized the head during data acquisition utilizing the delay between blood-oxygen-level-dependent (BOLD) and neural signals. Visual stimuli simulated either a stable environment or one with arbitrary head-coupled visual motion. Importantly, both conditions were matched in retinal and head motion. Contrasts revealed differential responses in human V6. Given the lack of vestibular responses in primate V6, these results suggest multi-modal integration of visual with neck efference copy signals or proprioception in V6.

INTRODUCTION

A remarkable property of the visual system is its feat to provide us with stable vision despite continuously changing retinal input induced by our movements of eyes, head, and body. This feat appears especially intriguing as the majority of visual areas are organized retinotopically, yet stability requires integration of visual input with cues from other modalities. Although the integration of eye movements with retinal signal has been studied extensively in both monkeys (Galletti et al., 1984, 1988, 1990; Erickson and Thier, 1991; Ilg et al., 2004; Dicke et al., 2008) and humans (Goossens et al., 2006; Arnoldussen et al., 2011; Fischer et al., 2012; Nau et al., 2018), the integration of visual signal with voluntary head movements remains barely studied at the level of neocortex (see Carriot et al. [2013]; Cullen and Taube [2017] for subcortical function).

In macaques and humans, previous studies examining cortical function focused on passive head motion or artificial vestibular stimulation to examine visual-vestibular integration (Gu et al., 2008; Chen et al., 2011; Smith et al., 2012; Frank et al., 2014, 2016; Billington and Smith, 2015). However, active gaze shifts beyond eye movements also involve head rotation (Land, 1992). In fact, gaze change commands reach eye and head effector muscles at the same time (Bizzi et al., 1971), and human observers compensate for eye- and head-induced self-motion with equal precision (Crowell et al., 1998). Notably, however, despite this prominent role for head motion in visual stability almost nothing is known about which visual processing stages integrate head motion signals with retinotopic representations as technical limitations have hindered human neuroimaging to study the neural underpinnings of voluntary head movements.

We recently circumvented these limitations and introduced an approach that allows participants to move their heads during fMRI scanning by exploiting the delay of several seconds between neural processing and blood-oxygen-level-dependent (BOLD) signal (see Figures 1A–1C) (Schindler and Bartels, 2018). We constructed a custom-built air pressure-based head stabilization system that permitted head rotation during trials, but stabilized head position during data acquisition. Observers wore head-mounted magnetic resonance-compatible goggles while head movement was tracked online. This allowed generation of visual stimuli that could be modulated by head motion in real time (Schindler and Bartels, 2018). In two conditions, observers viewed approaching visual flow that was modulated by head motion. A congruent condition simulated a scenario of constant forward motion where head rotation resulted in looking around while being driven along a straight road. In the incongruent condition, observers performed identical head

[1]Vision and Cognition Lab, Centre for Integrative Neuroscience, University of Tübingen, Otfried-Müller-Str. 25, Tübingen 72076, Germany

[2]Department of Psychology, University of Tübingen, Tübingen 72076, Germany

[3]Max Planck Institute for Biological Cybernetics, Tübingen 72076, Germany

[4]Centre for Integrative Neuroscience & MEG Center, University of Tübingen, Tübingen 72076, Germany

[5]Lead Contact

*Correspondence: andreas.schindler@tuebingen.mpg.de (A.S.), andreas.bartels@tuebingen.mpg.de (A.B.)

https://doi.org/10.1016/j.isci.2018.09.004

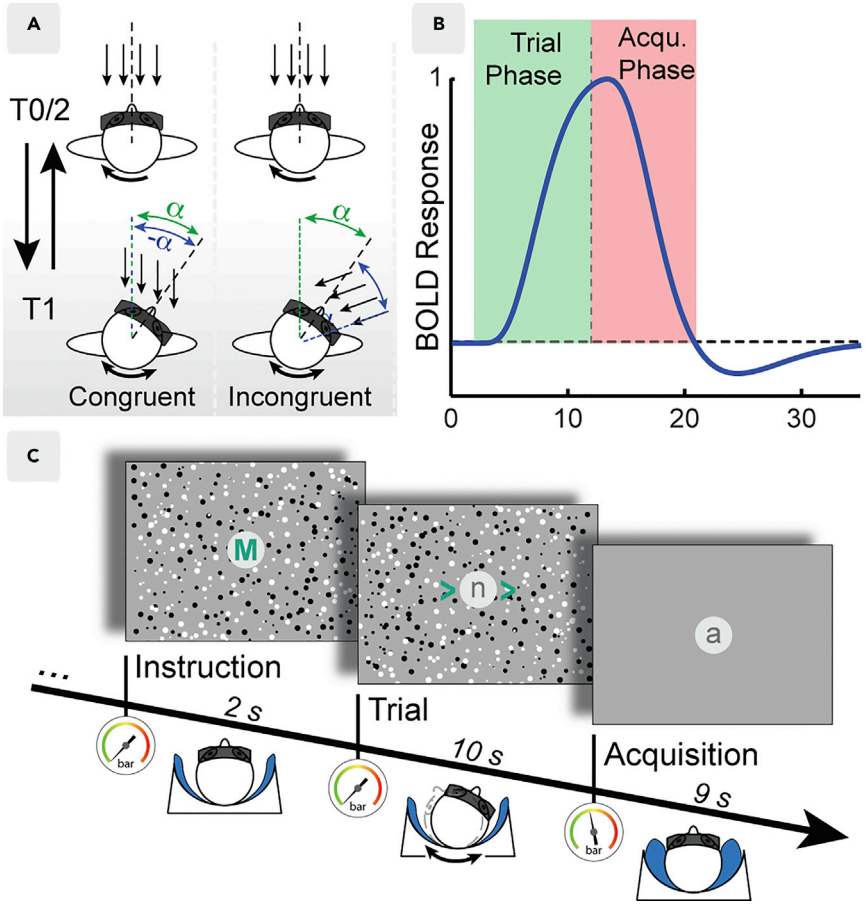


Figure 1. Illustration of Visual Stimuli and Head-Rotation Task, BOLD Signal Acquisition during a Trial, and Experimental Paradigm

(A) Observers performed voluntary head rotations while being approached by a simulated 3D dot cloud in both congruent and incongruent conditions. Head rotations in the congruent condition lead to cloud rotation in opposite direction (-α) to the observer's head (α), as would be experienced when moving forward in a stable environment and looking around. In the incongruent condition, the cloud and head rotated in the same direction (α), resulting in perceptually arbitrary motion of the environment. Note that retinal flow as well as head motion were matched in both conditions.

(B) Model of the evoked BOLD time course as predicted by the paradigm in (C). Stimulus presentation and active head movements induced BOLD signals during the trial phase (green shade) while the slow dynamics of the BOLD signal allowed acquisition of these responses even after stimulus offset, at a time when the observer's head was stabilized (acquisition phase, red shade).

(C) Each trial started with an instruction phase when air cushions were emptied. In the trial phase, green arrowheads guided the observer's head rotation. In the acquisition phase, air cushions were inflated again to record BOLD responses. Observers performed a demanding fixation task across all phases and conditions, except the instruction phase (see Methods).

rotations, but visual consequences of head rotation were inversed such that visual and extra-retinal cues did not combine in any meaningful way but retinal motion was matched to the congruent condition. In both conditions, a demanding letter detection task assured fixation. Based on this paradigm we previously examined the integration of head movements and visual signals in a network of areas with established vestibular input. Particularly, a contrast between congruent and incongruent conditions revealed evidence consistent with the multi-modal integration of visual cues with head motion into a coherent "stable world" percept in the parietal operculum and in the anterior part of the parieto-insular cortex. This also applied for a subset of visual motion-responsive areas such as human medial superior temporal area (MST) (at uncorrected level), the dorsal part of the ventral intraparietal area (VIP), the cingulate sulcus visual area (CSv), and a region in the precuneus (Pc) (Schindler and Bartels, 2018). However, the important question whether retinotopic cortex and especially areas V3A and V6 play a role in visual stability during voluntary head movement remained open.

RESULTS

We here fill this gap by combining the previously collected dataset with newly acquired retinotopic localizer data in the same participants for early visual areas V1, V2, V3, and V4 as well as higher visual areas V3A and V6 (see Methods). We probed each area using the BOLD contrast between congruent and incongruent conditions. Notably, these conditions were matched in both, head motion and retinal visual flow. However, only the congruent condition reflected a stable world and hence a match between visual and extra-retinal signals, whereas the incongruent condition constituted an arbitrary combination of both features. We thus operationalized "visual stability"—i.e., to what extent a given visual area might integrate its retinotopic input with extra-retinal signals—as differential BOLD activity to this contrast. Although we used this measure to probe a given area for integration, we had no a priori hypothesis regarding the sign of this difference. In fact, given the lack of the underlying integration mechanism as well as the complex relation between neural activity and BOLD signals, many scenarios are conceivable (Logothetis et al., 2001; Logothetis and Wandell, 2004; Goense and Logothetis, 2008; Logothetis, 2008).

Neither BOLD responses in early visual cortex (V1: $T(18) = -1.32$; $p = 0.771$; V2: $T(18) = -1.35$; $p = 0.771$; V3: $T(18) = -1.10$; $p = 0.771$; V4: $T(18) = -1.24$; $p = 0.771$; corrected for family-wise error (FWE) nor activation in V3A ($T(18) = -0.60$; $p = 0.559$; FWE-corrected) distinguished between congruent and incongruent conditions (Figure 2A). We found, however, significantly increased responses to the congruent combination of visual and extra-retinal signals in area V6 ($T(18) = 2.52$; $p = 0.043$; FWE-corrected) (Figure 2A). To reexamine these results, we performed a multivariate pattern analysis training classifiers to distinguish congruent and incongruent conditions based on multi-voxel region of interest (ROI) patterns. Using an odd/even cross-validation approach we trained support vector machines to distinguish both conditions on one-half of the data and tested them on the other half. Notably, this analysis confirmed our positive result in V6 (mean accuracy = 0.64 ± 0.06; $p = 0.004$; FWE-corrected) and the lack of differential activity in the early visual cortex (V1: mean accuracy = 0.46 ± 0.05, $p = 0.987$; V2: mean accuracy = 0.54 ± 0.06, $p = 0.548$; V3: mean accuracy = 0.58+-0.04, $p = 0.126$; V4: mean accuracy = 0.57 ± 0.06, $p = 0.217$; FWE-corrected). In addition, it revealed a small effect in V3A (mean accuracy = 0.58 ± 0.05, $p = 0.024$, uncorrected; $p = 0.101$, FWE-corrected) that did not survive correction for multiple comparisons.

We took great care to rule out alternative accounts of our results. There were no differences in fixation task performance between congruent and incongruent conditions, and quantitative analysis of head movements as well as visual motion energy showed that both were matched between conditions (Schindler and Bartels, 2018).

DISCUSSION

Despite substantial progress in vision research, the mechanisms of how transient retinotopic representations may give rise to stable vision are far from being understood (Burr and Morrone, 2011). In particular, evidence for the compensation of visual self-motion induced by voluntary head motion is scarce. Based on a recently established paradigm (Schindler and Bartels, 2018) we here provide first evidence for the integration of retinotopic representations and voluntary head movement signals in the human homolog of area V6.

Already early studies in macaques found strong visual motion selectivity in V6 (Galletti et al., 1996, 1999). These findings are paralleled by a growing body of human neuroimaging studies that point to human V6 as a prominent visual motion area specialized in the processing of self-motion flow fields (Cardin and Smith, 2010; Pitzalis et al., 2010; Arnoldussen et al., 2011; Cardin et al., 2012b; Fischer et al., 2012; Schindler and Bartels, 2017). However, the role that self-motion signals in V6 may play has been an open question. Monkey electrophysiology found visual heading selectivity in V6 neurons (Fan et al., 2015), whereas human V6 shows no adaptation effects when presented with successive optic flow patterns (Cardin et al., 2012a). In fact, the lack of BOLD and single cell responses to vestibular signals in V6 of man (Smith et al., 2012) and monkey (Fan et al., 2015) even further deemphasize a role for V6 in heading perception. Rather, V6 has been suggested to play a major role in the compensation of self-induced visual motion to compute object motion. In line with this, "real motion" neurons in V6 receive efference copies of eye movements and attenuate or stop firing when retinal motion is induced by active pursuit over a static target (Galletti and Fattori, 2003). In fact, human V6 even shows near-complete multi-modal integration of planar visual and non-visual motion signals during pursuit (Fischer et al., 2012). Corroborating the self-motion compensation view, our results provide evidence for the integration of self-motion-compatible visual flow signals with non-retinal

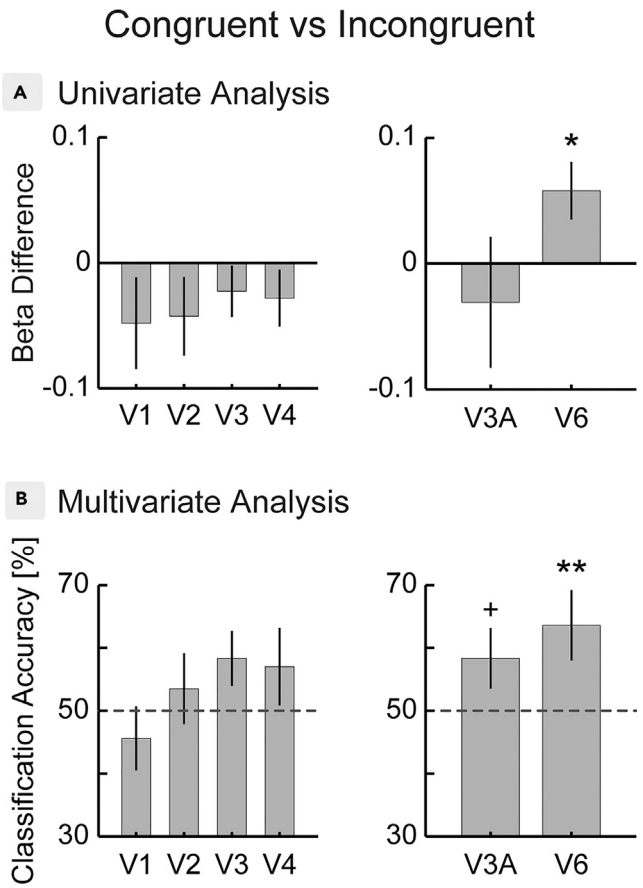


Figure 2. Univariate and Multivariate Differences between Congruent and Incongruent Combinations of Visual and Extra-Retinal Signals in Visual Cortex

(A) A contrast between both conditions revealed significant responses to the congruent combination of both cues in area V6, whereas V3A and early visual areas showed no differential activation.

(B) A subsequent multivariate pattern analysis found robust classification of both conditions in area V6, whereas effects in V3A did not survive multiple comparison correction. Patterns in early visual cortex did not distinguish between congruent and incongruent conditions. Dashed lines indicate chance level.

$^{+}p < 0.05$, uncorrected; *$p < 0.05$, FWE-corrected; **$p < 0.005$; FWE-corrected. Error bars indicate SEM. See Figure S1 for responses to an additional "head only" condition and Figure S2 for retinotopic area definitions of example observers.

voluntary head motion signals in human V6. Although the integration of retinotopic information and eye movements is predominantly mediated by corollary discharge (Sommer and Wurtz, 2002; Cavanaugh et al., 2016), less is known about integration signals of voluntary head movements. The lack of passive inertial head motion signals in monkey V6 (Fan et al., 2015), together with the absence of responses to galvanic vestibular stimuli in human V6, render it likely that the integration of visual and extra-retinal head motion signals observed here may have been driven by efference copy signals or proprioception. Given the self-motion hypothesis of V6, this idea is further corroborated by the fact that proprioceptive information in combination with neck efference copy signals are sufficient to allow human observers accurate compensation for head-induced visual self-motion, whereas vestibular cues presented in isolation are not (Crowell et al., 1998). Following these arguments, we appreciate the sources of extra-retinal signals as an important difference between the present results in V6 and our previous findings concerning the integration of visual and head motion signals in areas with prominent vestibular responses. Although these included the posterior insula and human VIP, integration responses in human MST reached significance only at the uncorrected level, whereas area MT/V5 showed no differential responses to either condition (Schindler and Bartels, 2018). Nevertheless, monkey V6 as well as its human homolog form major connections to both (h)MST and (h)VIP (Galletti et al., 2001; Tosoni et al., 2015). In fact, all these areas have been found to contain neurons that code visual space in craniotopic coordinates (Galletti et al., 1993; Duhamel

et al., 1997; Colby and Goldberg, 1999; Ilg et al., 2004). Notably, however, in these experiments the monkey's head was typically restrained such that head-centered and body-centered coordinates coincided. Against this background, the present results render it likely that, at least to some extent, craniotopic responses in human V6 (in addition to hMST and hVIP [Schindler and Bartels, 2018]) may actually generalize to a body-centered reference system, as it has been proposed for other parietal regions (Brotchie et al., 1995). A body-centered reference framework would also well be in line with the notion that V6 plays a major role in mapping vision onto action (Galletti et al., 1999, 2001; Pitzalis et al., 2012; Tosoni et al., 2015). This seems particularly evident for grasping as, directly adjacent to V6, cells in V6A are tuned to arm movements (Fattori et al., 2001) as well as to hand orientation (Fattori et al., 2009). In light of the present results, interactions of V6 and V6A could thus mediate visually guided arm movements during gaze shifts regardless of whether they are induced by eye or head movements. In addition, a body-centered reference frame is also consistent with a possible role for V6 in flow parsing for navigation as it has been suggested based on functional properties and prominent connections between V6 and para-hippocampal cortex in humans and monkeys (Galletti and Fattori, 2003, 2018; Pitzalis et al., 2010; Tosoni et al., 2015).

In contrast to V6, evidence for the integration of head movement and retinotopic signals in human V3A was only evident in a more sensitive multivariate pattern analysis (MVPA) analysis and did not survive corrections for multiple comparisons (see Figure 2B). Given the prominent role of V3A in the integration of eye movements with visual signals, this finding seems counter-intuitive at first sight. However, the additive forward flow used in the present study had abolished the ability of V3A to compensate planar motion for eye movements previously, whereas V6 still compensated (Fischer et al., 2012). This suggests a more advanced flow parsing mechanism in V6 able to isolate planar motion from expansion flow. Additional studies are hence needed to carefully examine the possible role of V3A in the compensation for head-induced self-motion during planar flow alone. We also found no evidence for the integration of retinotopic signals and head movements in early visual cortex. As for V3A this does not mean that early visual areas might not integrate these signals in a different setting. However, the fact that responses there did not differentiate between congruent and incongruent conditions corroborates that both conditions were well matched in low-level features.

Considering the substantial body of literature on visuo-vestibular integration in monkeys (Gu et al., 2008; Chen et al., 2011), caloric and galvanic stimulation in human neuroimaging (Brandt et al., 1998; Smith et al., 2012; Frank et al., 2014, 2016; Billington and Smith, 2015), as well as the results of our previous study (Schindler and Bartels, 2018), evidence converges toward a network of cortical brain areas that may directly or indirectly be involved in maintaining our stable experience of the visual world while we navigate in it. This network includes not only regions of the posterior insula as central hubs of vestibular processing but also areas with vestibular input that were initially associated with visual (self-)motion processing such as MST, VIP, and CSv. Although this list is far from being complete, the present results point to V6 as an important node in this network concerned with the integration of visual and extra-retinal signals during head-induced gaze shifts. Given the apparent lack of vestibular input to this area, future studies are required to examine possible sources of head-driven signals that may enable such integration in V6.

Limitations of the Study

As shown in our previous study, voluntary head motion and artificial vestibular stimulation both led to similar BOLD responses in insular cortex and areas of the self-motion network (Schindler and Bartels, 2018). However, a major advantage of voluntary head movements over artificial vestibular stimulation is the ecological validity, and foremost, the possibility to precisely match high-frequency temporal aspects and intensity between visual and non-visual signals and therefore provide an immersive, realistic experience. A disadvantage of head movements is that contributions of non-retinal signals, i.e., efference copy, and proprioceptive and vestibular signals coincide and thus cannot be disentangled. Although no previous evidence for vestibular signals in primate V6 exists (Smith et al., 2012; Fan et al., 2015), the present study thus cannot conclusively answer the question what source head signals for multi-modal integration in V6 may have.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Transparent Methods and two figures and can be found with this article online at https://doi.org/10.1016/j.isci.2018.09.004.

ACKNOWLEDGMENTS

We thank Mirsat Memaj and Alexander Baum of the MPI workshop for building the air-pressure-based head stabilizing system, Joachim Werner and Winfried Engler for interfacing it to the stimulus computer, and Mario Kleiner for integrating the head tracker with Psychtoolbox. This work was funded by the Centre for Integrative Neuroscience, Tübingen; by the German Excellence Initiative of the German Research Foundation (DFG) grant number EXC307; by DFG grants BA 4914/3-1 and SFB 1233 Robust Vision TP09; by University Clinics Tübingen Fortüne grant number 2251-0-0; and by the Max Planck Society, Germany.

AUTHOR CONTRIBUTIONS

A.S. and A.B. conceived the study. A.S. programmed stimuli and acquired and analyzed data. A.S. and A.B. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: June 8, 2018
Revised: July 13, 2018
Accepted: September 4, 2018
Published: September 28, 2018

REFERENCES

Arnoldussen, D.M., Goossens, J., and van den Berg, A.V. (2011). Adjacent visual representations of self-motion in different reference frames. Proc. Natl. Acad. Sci. USA *108*, 11668–11673.

Billington, J., and Smith, A.T. (2015). Neural mechanisms for discounting head-roll-induced retinal motion. J. Neurosci. *35*, 4851–4856.

Bizzi, E., Kalil, R.E., and Tagliasco, V. (1971). Eye-head coordination in monkeys: evidence for centrally patterned organization. Science *173*, 452–454.

Brandt, T., Bartenstein, P., Janek, A., and Dieterich, M. (1998). Reciprocal inhibitory visual-vestibular interaction. visual motion stimulation deactivates the parieto-insular vestibular cortex. Brain *121* (Pt 9), 1749–1758.

Brotchie, P.R., Andersen, R.A., Snyder, L.H., and Goodman, S.J. (1995). Head position signals used by parietal neurons to encode locations of visual stimuli. Nature *375*, 232–235.

Burr, D.C., and Morrone, M.C. (2011). Spatiotopic coding and remapping in humans. Philos. Trans. R. Soc. Lond. B Biol. Sci. *366*, 504–515.

Cardin, V., and Smith, A.T. (2010). Sensitivity of human visual and vestibular cortical regions to egomotion-compatible visual stimulation. Cereb. Cortex *20*, 1964–1973.

Cardin, V., Hemsworth, L., and Smith, A.T. (2012a). Adaptation to heading direction dissociates the roles of human MST and V6 in the processing of optic flow. J. Neurophysiol. *108*, 794–801.

Cardin, V., Sherrington, R., Hemsworth, L., and Smith, A.T. (2012b). Human V6: functional characterisation and localisation. PLoS One *7*, e47685.

Carriot, J., Brooks, J.X., and Cullen, K.E. (2013). Multimodal integration of self-motion cues in the vestibular system: active versus passive translations. J. Neurosci. *33*, 19555–19566.

Cavanaugh, J., Berman, R.A., Joiner, W.M., and Wurtz, R.H. (2016). Saccadic corollary discharge underlies stable visual perception. J. Neurosci. *36*, 31–42.

Chen, A., DeAngelis, G.C., and Angelaki, D.E. (2011). Convergence of vestibular and visual self-motion signals in an area of the posterior Sylvian fissure. J. Neurosci. *31*, 11617–11627.

Colby, C.L., and Goldberg, M.E. (1999). Space and attention in parietal cortex. Annu. Rev. Neurosci. *22*, 319–349.

Crowell, J.A., Banks, M.S., Shenoy, K.V., and Andersen, R.A. (1998). Visual self-motion perception during head turns. Nat. Neurosci. *1*, 732–737.

Cullen, K.E., and Taube, J.S. (2017). Our sense of direction: progress, controversies and challenges. Nat. Neurosci. *20*, 1465–1473.

Dicke, P.W., Chakraborty, S., and Thier, P. (2008). Neuronal correlates of perceptual stability during eye movements. Eur. J. Neurosci. *27*, 991–1002.

Duhamel, J.R., Bremmer, F., Ben Hamed, S., and Graf, W. (1997). Spatial invariance of visual receptive fields in parietal cortex neurons. Nature *389*, 845–848.

Erickson, R.G., and Thier, P. (1991). A neuronal correlate of spatial stability during periods of self-induced visual motion. Exp. Brain Res. *86*, 608–616.

Fan, R.H., Liu, S., DeAngelis, G.C., and Angelaki, D.E. (2015). Heading tuning in macaque area V6. J. Neurosci. *35*, 16303–16314.

Fattori, P., Gamberini, M., Kutz, D.F., and Galletti, C. (2001). 'Arm-reaching' neurons in the parietal area V6A of the macaque monkey. Eur. J. Neurosci. *13*, 2309–2313.

Fattori, P., Breveglieri, R., Marzocchi, N., Filippini, D., Bosco, A., and Galletti, C. (2009). Hand orientation during reach-to-grasp movements modulates neuronal activity in the medial posterior parietal area V6A. J. Neurosci. *29*, 1928–1936.

Fischer, E., Bulthoff, H.H., Logothetis, N.K., and Bartels, A. (2012). Human areas V3A and V6 compensate for self-induced planar visual motion. Neuron *73*, 1228–1240.

Frank, S.M., Wirth, A.M., and Greenlee, M.W. (2016). Visual-vestibular processing in the human Sylvian fissure. J. Neurophysiol. *116*, 263–271.

Frank, S.M., Baumann, O., Mattingley, J.B., and Greenlee, M.W. (2014). Vestibular and visual responses in human posterior insular cortex. J. Neurophysiol. *112*, 2481–2491.

Galletti, C., and Fattori, P. (2003). Neuronal mechanisms for detection of motion in the field of view. Neuropsychologia *41*, 1717–1727.

Galletti, C., and Fattori, P. (2018). The dorsal visual stream revisited: stable circuits or dynamic pathways? Cortex *98*, 203–217.

Galletti, C., Battaglini, P.P., and Aicardi, G. (1988). 'Real-motion' cells in visual area V2 of behaving macaque monkeys. Exp. Brain Res. *69*, 279–288.

Galletti, C., Battaglini, P.P., and Fattori, P. (1990). 'Real-motion' cells in area V3A of macaque visual cortex. Exp. Brain Res. *82*, 67–76.

Galletti, C., Battaglini, P.P., and Fattori, P. (1993). Parietal neurons encoding spatial locations in craniotopic coordinates. Exp. Brain Res. *96*, 221–229.

Galletti, C., Squatrito, S., Battaglini, P.P., and Grazia Maioli, M. (1984). 'Real-motion' cells in the primary visual cortex of macaque monkeys. Brain Res. *301*, 95–110.

Galletti, C., Fattori, P., Gamberini, M., and Kutz, D.F. (1999). The cortical visual area V6: brain location and visual topography. Eur. J. Neurosci. *11*, 3922–3936.

Galletti, C., Fattori, P., Battaglini, P.P., Shipp, S., and Zeki, S. (1996). Functional demarcation of a border between areas V6 and V6A in the superior parietal gyrus of the macaque monkey. Eur. J. Neurosci. *8*, 30–52.

Galletti, C., Gamberini, M., Kutz, D.F., Fattori, P., Luppino, G., and Matelli, M. (2001). The cortical connections of area V6: an occipito-parietal network processing visual information. Eur. J. Neurosci. *13*, 1572–1588.

Goense, J.B., and Logothetis, N.K. (2008). Neurophysiology of the BOLD fMRI signal in awake monkeys. Curr. Biol. *18*, 631–640.

Goossens, J., Dukelow, S.P., Menon, R.S., Vilis, T., and van den Berg, A.V. (2006). Representation of head-centric flow in the human motion complex. J. Neurosci. *26*, 5616–5627.

Gu, Y., Angelaki, D.E., and Deangelis, G.C. (2008). Neural correlates of multisensory cue integration in macaque MSTd. Nat. Neurosci. *11*, 1201–1210.

Ilg, U.J., Schumann, S., and Thier, P. (2004). Posterior parietal cortex neurons encode target motion in world-centered coordinates. Neuron *43*, 145–151.

Land, M.F. (1992). Predictable eye-head coordination during driving. Nature *359*, 318–320.

Logothetis, N.K. (2008). What we can do and what we cannot do with fMRI. Nature *453*, 869–878.

Logothetis, N.K., and Wandell, B.A. (2004). Interpreting the BOLD signal. Annu. Rev. Physiol. *66*, 735–769.

Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. Nature *412*, 150–157.

Nau, M., Schindler, A., and Bartels, A. (2018). Real-motion signals in human early visual cortex. Neuroimage *175*, 379–387.

Pitzalis, S., Fattori, P., and Galletti, C. (2012). The functional role of the medial motion area V6. Front. Behav. Neurosci. *6*, 91.

Pitzalis, S., Sereno, M.I., Committeri, G., Fattori, P., Galati, G., Patria, F., and Galletti, C. (2010). Human v6: the medial motion area. Cereb. Cortex *20*, 411–424.

Schindler, A., and Bartels, A. (2017). Connectivity reveals sources of predictive coding signals in early visual cortex during processing of visual optic flow. Cereb. Cortex *27*, 2885–2893.

Schindler, A., and Bartels, A. (2018). Integration of visual and non-visual self-motion cues during voluntary head movements in the human brain. Neuroimage *172*, 597–607.

Smith, A.T., Wall, M.B., and Thilo, K.V. (2012). Vestibular inputs to human motion-sensitive visual cortex. Cereb. Cortex *22*, 1068–1077.

Sommer, M.A., and Wurtz, R.H. (2002). A pathway in primate brain for internal monitoring of movements. Science *296*, 1480–1482.

Tosoni, A., Pitzalis, S., Committeri, G., Fattori, P., Galletti, C., and Galati, G. (2015). Resting-state connectivity and functional specialization in human medial parieto-occipital cortex. Brain Struct. Funct. *220*, 3307–3321.

ISCI, Volume *7*

Supplemental Information

Human V6 Integrates Visual and Extra-Retinal Cues during Head-Induced Gaze Shifts

Andreas Schindler and Andreas Bartels

Supplemental Figures

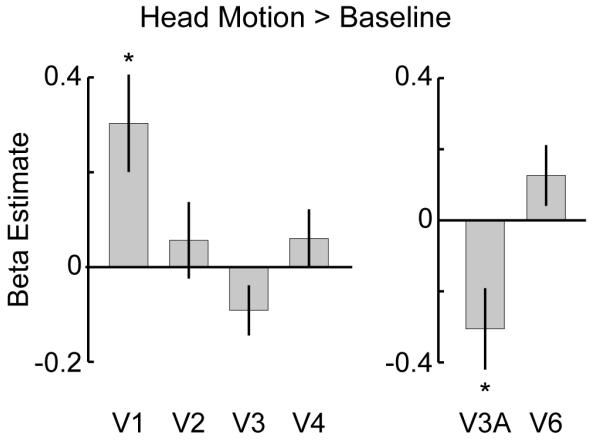


Figure S1 (relates to main figure 2): Head motion responses in visual areas. A contrast between the "head only" condition and a baseline condition involving static dots found differential BOLD responses in V1 and V3A while the others areas showed no differential activation. *= $p < 0.05$. FWE corrected. In an additional explorative analysis we compared responses to a "head only" condition that involved observer head motion while fixating a static dot display to a baseline condition involving "static" dots in the absence of head motion. Based on this analysis, we found differential responses in V1 while the remaining areas of early visual cortex showed no effects (V1: $T(18) = 2.95$; $p = 0.034$; V2: $T(18) = 0.70$; $p = 0.679$; V3: $T(18) = -1.73$; $p = 0.301$; V4: $T(18) = 0.98$; $p = 0.679$; FWE corrected). We additionally found BOLD modulation to "head only" in V3A, while V6 did not respond to head motion in the absence of additional visual motion (V3A: $T(18) = -2.67$; $p = 0.031$; V6: $T(18) = 1.47$; $p = 0.158$; FWE corrected). As described in the previous study (Schindler and Bartels, 2018), the "head only" condition was consistent with an additional interpretation. As the dot field was static with respect to the head (no dot motion in the head-mounted goggles) this condition was also consistent with world-centered object motion. Responses to this condition could thus have reflected object motion selectivity or rather head motion driven efference copy, proprioceptive or vestibular (even though less likely cf. (Smith et al., 2012)) signals in these areas. Scenarios for both, or even interactions between both stimulus interpretations are conceivable. While head motion was perfectly matched between congruent and incongruent conditions, there was a small (mean differences in translation: 0.01 mm ± 0.01 mm and rotation: 0.02° ± 0.00°) but significant ($p < 0.05$, FWE-corrected) difference in head motion between the "head only" and baseline conditions (see figure 6 in (Schindler and Bartels, 2018)). While we deem this influence as marginal given the fact that the affected contrast "head only" vs. "static" evoked opposing responses even in closely adjacent areas (i.e. V5/MT and MST; PIVC and pPIC, aPIC) previously (Schindler and Bartels, 2018), we cannot entirely exclude that this effect might have influenced results of this particular contrast in one of the areas examined here.

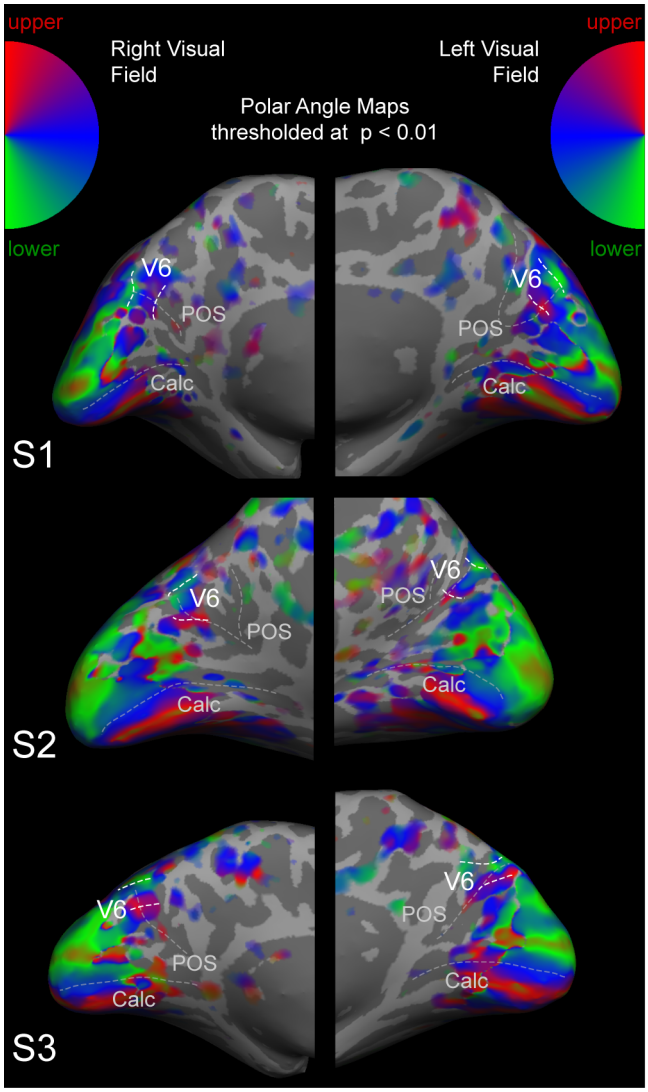


Figure S2 (relates to main figure 2): Retinotopic data of three example observers. Polar angle maps are thresholded at $p < 0.01$. Based on previous evidence (Pitzalis et al., 2006; Pitzalis et al., 2010; Pitzalis et al., 2012, 2015), V6 was defined as the contralateral hemi-field representation in the dorsal-most part of the parieto-occipital sulcus (POS).

Transparent Methods

Stimulus paradigm

The data used in the present study were acquired within a counterbalanced design involving a total of five conditions (Schindler and Bartels, 2018). For the purpose of the present study, two conditions were used for analysis. While a summary is given below, a full and comprehensive description of the paradigm, the customized hardware as well as all details of head motion and data analyses have been published previously (Schindler and Bartels, 2018).

In each condition, a trial comprised three phases. In the instruction phase (2 s), air cushions were emptied and a green “M” indicated that the next phase involved head rotation. The onset of the trial phase was indicated by the appearance of two green arrows to the left and right of the fixation dot. During the trial phase (10 s) the observer then moved the head as directed by the arrows, while the condition specific stimulus was presented. In the acquisition phase (9 s), air cushions were filled to record stimulus induced BOLD signals. Observers were instructed to continuously and slowly rotate their heads two times from center position (0°) to either ~+30° or ~ -30° and back. A given trial either involved rotations between 0° and ~-30° or between 0° and ~+30°, and trials of both directions were counterbalanced. Two green arrows guided observers' head rotations and switched orientation direction when the head reached a given angle threshold (i.e. 0°, 30°, −30°) based on online head tracking data. At the end of each head rotation cycle, both arrows disappeared and the fixation disc turned red until the observer's head reached center position again. In case the observer failed to complete the current rotation cycle, the fixation disc turned red after 10 s (i.e. after the entire trial phase) elapsed. Such trials were removed from further analysis. Subsequently, the acquisition phase started with filling of the air cushions.

Observers viewed stimuli via MR proof video goggles (resolution: 800 × 600; refresh rate: 60 Hz; VisuaStim, Resonance Technology Inc.) providing a visual field of 30 × 23 visual degrees. The paradigm was programmed using Psychtoolbox 3 (http://psychtoolbox.org/) and ran on MATLAB 2010 (Mathworks) on a Windows PC.

Trial phases of both conditions were based on random dot patterns that contained on average 400 black and white dots, presented on a gray (90 cd/m2) background at 100% contrast. Both conditions simulated a 3D dot cloud in front of the observer. The visibility of the cloud began at 360 cm and ended at 120 cm. Dots changed size with distance, subtending a maximal diameter of 2.2°. No disparity cues were provided. To study self-motion integration during simulated forward motion we added a translational component to dot motion such that dots approached the observer at a constant speed of 120 cm/s. Dot motion was modulated dependent on observers’ head rotation using real-time feedback from the head-tracking camera. In the “congruent” condition the 3D dot cloud rotated in opposite direction relative to the observer's head. Notably, this condition allowed the observer to integrate visual self-motion motion cues and head movement into a coherent stable world percept, i.e. as experienced in the real world when a navigator rotates his/her head while traveling along a straight linear path. It hence allowed for integration of multimodal cues just as experienced in real-life scenarios. In contrast, head rotation in the “incongruent” condition involved rotation of the 3D dot cloud in the same direction as the observer's head. Thus, this condition provided identical visual motion energy (in retinal coordinates) and head motion as the “congruent” condition, but only provided arbitrary coupling of both cues in ways that cannot be encountered in real-world scenarios. In the “static” and “head motion“ conditions each trial involved a single static projection of the 3D dot cloud, i.e. devoid of any dot motion. Note that since in the “head motion” condition the

dot field was static with respect to the head (no dot motion in the head-mounted goggles), this condition also involved world-centered object motion.
To ensure fixation, vigilance and balanced attention across all conditions, observers performed a demanding letter detection task during all conditions in trial and acquisition phases (but not during the instruction phase). The task was presented within the fixation disc (1.3° diameter) that was presented at the center of the head mounted display. Observers were asked to press a button whenever an alphabetical letter (a–z, size: 0.9°) with same identity appeared twice in a row. Letters appeared every 830 ms with repetitions occurring on average every 6 s.

Observers

A sub sample of 19 of the originally acquired 21 observers (Schindler and Bartels, 2018) took part in this study (8 males, 11 females, age 20–36 years; mean 27 years). The remaining two observers were no longer available at the time we acquired retinotopic data. All participating observers gave written informed consent prior to participation. The study was approved by the ethics committee of the University Clinic Tübingen.

Region of interest definition

Retinotopic maps of areas V1, V2, V3 and V4 as well as areas V3A and V6 were delineated in 19 observers using standard retinotopic mapping techniques (Sereno et al., 1995; Swisher et al., 2007). In contrast to the main experiment that used MR-compatible head-mounted goggles, retinotopic stimuli were back-projected using a projector onto a screen positioned behind the observers' head and viewed via a front-surfaced mirror mounted on the head coil, with 1280 × 1024 pixels resolution at 60 Hz. The visual display subtended 24 × 18 visual degree and was viewed at 80 cm distance. The stimulus consisted of a wedge shaped checkerboard with an arc angle of 90 degrees and a radius of 12 visual degrees (100% contrast, 6 Hz contrast inversion flicker, check sizes increased logarithmically with eccentricity) rotating around the central fixation dot (55.7 s period, matched clockwise and counterclockwise rotation) in front of a gray background. An attention task was superimposed onto the checkerboard to facilitate parietal responses (Swisher et al., 2007). For preprocessing, images were slice time corrected and smoothed by a Gaussian kernel of 4 mm full width half maximum. Polar angle maps were calculated as described in previous studies (Sereno et al., 1995), and projected on inflated brain surfaces of individual participants. This way, polar maps of occipital cortex were used to define areas V1, V2, V3 and V4. V3A was delineated as the intersection of the retinotopically defined V3AB-complex and a dedicated functional motion localizer. This localizer defined V3A as the strong response on the superior occipital lobe to pursuit with co-moving objective motion versus pursuit on a static background, corresponding to contrasting objective with retinal motion, which replicably leads to selective activation of voxels overlapping with retinotopically defined V3A (Fischer et al., 2012).
Based on previous evidence, we defined human area V6 as the contralateral hemi-field representation in the dorsal-most part of the parieto-occipital sulcus (POS) (Pitzalis et al., 2006; Pitzalis et al., 2010; Pitzalis et al., 2012, 2015) (see also example data in supplemental figure 2). We used this approach instead of a functional motion localizer for the delineation of V6 as the latter may have increased the risk to

incorrectly label voxels of the adjacent V6A as V6 (Pitzalis et al., 2015). In fact, V6 as well as V6A have both shown to respond to visual motion while only V6 has a prominent retinotopic representation including the center of gaze (Pitzalis et al., 2015). While definition of area V6 is facilitated using wide-field retinotopic stimuli (Pitzalis et al., 2006; Pitzalis et al., 2010), V6 has previously also been delineated using standard retinotopic procedures (Pitzalis et al., 2010; Cardin et al., 2012b; Fischer et al., 2012). We here used retinotopic stimuli comparable in size with the main flow stimuli to assure that the selected V6 voxels were also stimulated by the main optic flow stimuli. Surface reconstruction (Dale et al., 1999), analysis of polar angle maps as well as retinotopic ROI delineation was done in FreeSurfer (http://surfer.nmr.mgh.harvard.edu/).

MRI acquisition

Functional gradient-echo echoplanar T2-weighted images (EPI) were acquired on a Siemens 3T Prisma scanner with a 20-channel phased-array head coil (Siemens, Erlangen, Germany), with the following parameters: TR 754 ms, TE 30 ms, flip angle 55°, field of view 192 × 192 mm, multiband factor 3. Images consisted of 36 slices with 64 × 64 pixels (3 mm thick, no gap), resulting in 3 × 3 × 3 mm voxels. Slices were acquired in an interleaved fashion. We recorded six runs of the main experiment each consisting of 1415 images (including all 5 conditions initially acquired). This involved images for all three experimental phases of all trials. Motion localizer and retinotopy data were acquired with a higher resolution of 2 × 2 × 2 mm in 56 slices (TR 870 ms, TE 30 ms, flip angle 56°, GRAPPA factor 2 and multiband factor 4) using a 64 channel head coil. One run of the motion localizer consisted of 710 images. One run of polar wedge retinotopy involved 660 images. We acquired one or two runs motion localizer and four runs of retinotopy per observer, respectively. For each run, the initial eight images were discarded to allow for equilibration of T1 signal. A high-resolution anatomical scan was also obtained for each observer with a T1-weighted MP-RAGE sequence of 1 × 1 × 1 mm resolution.

Data Analysis

The central steps of the fMRI analysis are provided in the following, as the full description has been provided elsewhere (Schindler and Bartels, 2018). Only images that were recorded during the acquisition phase (see figure 1B) were considered for further analysis. Preprocessing involved rejection of head-motion artifacts, followed by preprocessing using SPM12 (http://www.fil.ion.ucl.ac.uk/spm/) including head-motion correction and spatial normalization to MNI space. Images were spatially smoothed using a Gaussian Kernel of 6 mm full-width at half maximum.

Each observer was analyzed separately using a GLM. We initially built the full design matrix based on the entire time series including time points that were affected by motion artifacts. Thus, each of the five initially acquired conditions was modeled separately and was convolved with the hemodynamic response function as implemented by SPM12. To account for potentially remaining head movement related variance and for drifts, a set of nuisance regressors was included. For model estimation we removed all time points affected by motion artifacts from design matrix and time series. This procedure maintained the temporal structure of data and model despite the removal of artifact-polluted images.

For ROI analyses, condition betas for congruent and incongruent conditions were extracted in each observer individually after ROIs had been projected to standard space. ROIs of both hemispheres in a given observer were averaged for each visual area. Significant differences between conditions were assessed using one-sample, two-tailed t-tests (i.e. the difference scores of two given conditions were tested against zero), and corrected for multiple comparisons.

For multivariate pattern analyses, we trained linear support vector machines (Fan et al., 2008) to distinguish voxel response patterns of congruent and incongruent conditions. Beta estimates from voxels of a given ROI were combined to form vectors of brain responses. For every run and every voxel, the time series of beta estimates was quadratically detrended to filter out low-frequency noise and subsequently *z*-scored. To make our analysis more robust against outliers, we set all values with a difference of >2 SD from the mean to −2 and 2, respectively. Classification was performed using cross-validation across runs in an odd/even fashion. Statistical significance of classification accuracies was evaluated using permutation tests. For every time we trained a classifier to discriminate between fMRI patterns, we refit new classifiers after randomly permuting the labels in the training set 10^4 times. We used the 10^4 classification accuracies from each observer to obtain a null distribution of mean accuracies at the group level expected under the null hypothesis (including the accuracy that was actually observed using the un-permuted dataset). From these null distributions, p values for a one-tailed test can be calculated as the number of values in the distribution that exceeded the observed accuracies divided by the number of permutations. We controlled the family-wise error (FWE) by constructing a common null distribution by taking the maximum value across ROI group means in each permutation step while making sure that the same label permutations were used in every ROI (Nichols and Holmes, 2002). The resulting null distribution was then used to calculate FWE-corrected p values. MVPA-Analysis were implemented using in-house code and the Princeton-MVPA toolbox (https://github.com/PrincetonUniversity/princeton-mvpa-toolbox).

Supplemental References

Fischer E, Bulthoff HH, Logothetis NK, Bartels A (2012) Human areas V3A and V6 compensate for self-induced planar visual motion. Neuron 73:1228-1240.

Nau M, Schindler A, Bartels A (2018) Real-motion signals in human early visual cortex. Neuroimage.

Petro LS, Vizioli L, Muckli L (2014) Contributions of cortical feedback to sensory processing in primary visual cortex. Front Psychol 5:1223.

Pitzalis S, Fattori P, Galletti C (2012) The functional role of the medial motion area V6. Front Behav Neurosci 6:91.

Pitzalis S, Fattori P, Galletti C (2015) The human cortical areas V6 and V6A. Vis Neurosci 32:E007.

Pitzalis S, Sereno MI, Committeri G, Fattori P, Galati G, Patria F, Galletti C (2010) Human v6: the medial motion area. Cereb Cortex 20:411-424.

Pitzalis S, Galletti C, Huang RS, Patria F, Committeri G, Galati G, Fattori P, Sereno MI (2006) Wide-field retinotopy defines human cortical visual area v6. J Neurosci 26:7962-7973.

Schindler A, Bartels A (2018) Integration of visual and non-visual self-motion cues during voluntary head movements in the human brain. Neuroimage 172:597-607.

Smith AT, Wall MB, Thilo KV (2012) Vestibular inputs to human motion-sensitive visual cortex. Cereb Cortex 22:1068-1077.