Quantitative assessment of visual cortex function with fMRI at 7 Tesla—test–retest variability

Aini Ismafairus Abd Hamid [1, 2*], Oliver Speck [1, 3, 4, 5] and Michael B. Hoffmann [3, 6]

[1] Department of Biomedical Magnetic Resonance, Institute for Experimental Physics, Otto-von-Guericke University Magdeburg, Magdeburg, Germany, [2] Department of Neurosciences, School of Medical Sciences Health Campus, Universiti Sains Malaysia, Kubang Kerian, Malaysia, [3] Center for Behavioral Brain Sciences, Magdeburg, Germany, [4] Leibniz Institute for Neurobiology, Magdeburg, Germany, [5] German Center for Neurodegenerative Disease, Magdeburg, Germany, [6] Visual Processing Laboratory, Department of Ophthalmology, Otto-von-Guericke University Magdeburg, Magdeburg, Germany

fMRI-based retinotopic mapping was used to assess systematic variations in activated cortical surface area, amplitude, and coherence across sessions. Seven healthy subjects were scanned at 7 T in three separate sessions with intervals of 51.4 ± 5.4 days (Sessions 1 and 2) and 167.9 ± 24.4 days (Sessions 2 and 3). We found a reduction between Sessions 1 and 2 for activated cortical surface area, between Sessions 1 and 3 for amplitude, and between Sessions 1 and 2/3 for coherence. The results do not support head motion as a major cause of the observed effect seen in Session 1, suggesting that cognitive effects were the underlying cause of change. The phase correlations for both eccentricity and polar angle mapping were highly correlated between sessions, demonstrating the stability of the maps. Furthermore, the sensitivity in determining inter-session changes of cortical surface area, response amplitude, and coherence were, at a 5% significance level, estimated to be 1.5, 6, and 5%, respectively. Any future longitudinal fMRI study should carefully evaluate activation across sessions to determine the eligibility of inclusion of all time points. This experimental design provides guidance in methodological issues of clinical longitudinal fMRI-studies, specifically regarding effects of subject experience.

Keywords: test-retest, longitudinal, reliability, reproducibility, retinotopic maps, fMRI, visual areas

Introduction

Functional magnetic resonance imaging (fMRI) is a widely used method for investigating human brain functions. Retinotopic mapping is the gold standard to define visual field locations in the occipital cortex. Recent advances in fMRI using retinotopic mapping stimuli enable the identification of individual cortical visual areas and the functional specialization of the visual system. Besides fMRI, positron emission tomography (PET), and event-related potentials (ERPs) can also be used to identify visual topography or retinotopic organization in the human visual cortex (Woldorff et al., 1997). There are a number of factors that influence fMRI results; these factors can stem from the image signal-to-noise ratio, motion artifacts, magnetic field inhomogeneities, cognitive processes, cognitive strategies over time, or other causes (Krüger and Glover, 2001; Huettel et al., 2008; Bennett and Miller, 2010).

Test-retest studies have been used in the field of neuroscience to understand how response variabilities occur over time and influence detection sensitivity. In clinical studies of neurological disorders, for example, repeated measures have the potential to identify changes in brain activation, structure, and connectivity associated with disease progression, treatment or rehabilitation. To understand treatment effects, knowledge of the underlying test–retest variability for repeated measures of normal subjects' brain activation patterns is essential. Test-retest studies of normal subjects can provide guidance for the consideration of experience-driven effects and other methodological issues (Kelly et al., 2006).

Previous studies have investigated fMRI test–retest results from a wide range of tasks relevant to auditory, visual, motor, and cognitive stimuli (Specht et al., 2003; Johnstone et al., 2005; Kong et al., 2007; Raemaekers et al., 2007; Friedman et al., 2008; Caceres et al., 2009; Bressler and Silver, 2010; Seibert et al., 2012). For the visual areas and retinotopic mapping itself, however, there is only little knowledge. Bressler and Silver (2010) used a rotating checkerboard wedges stimulus, and the subjects were instructed to direct their attention to either the wedge (attention-to-wedges task) or the central fixation point (attention-to-fixation task). They reported that spatial attention tasks improve the reliability of fMRI retinotopic mapping signals in the occipital and parietal cortex.

Test-retest reliability of fMRI results has been assessed in many ways, for example using the Pearson product-moment correlation coefficient (Pearsons's r; Fernández et al., 2003), maximum likelihood (Maitra et al., 2002), or intraclass correlation coefficient (Raemaekers et al., 2007; Friedman et al., 2008; Caceres et al., 2009; Bressler and Silver, 2010; Seibert et al., 2012). To our knowledge no gold standard exists. In this study we aim to assess systematic variations in the fMRI responses between sessions. Therefore, we have applied a number of previously used measures to compare the activation between sessions. In order to make our work comparable to previous studies, we have included activation extent, response amplitude and coherence. Additionally, to evaluate the similarity of the obtained maps across sessions, we examined the reproducibility of the phase maps in the early visual areas.

Although retinotopic mapping has been a common and widely used method for visual studies, there is only limited knowledge on within-subject changes of brain activity, specifically in test–retest paradigms. In this study, conventional retinotopic mapping, i.e., eccentricity and polar angle mapping, and full field tasks were used to assess systematic variations in activated cortical surface area, fMRI signal amplitude, and coherence across sessions.

Materials and Methods

Subjects
Seven subjects (3 female; age 29 ± 4 years), with no history of ophthalmological or neurological disease, participated in the study. The procedures followed the tenets of the declaration of Helsinki (World Medical Association, 2000) and the protocol was approved by the ethics committee of the University of Magdeburg, Germany. The subjects gave written and informed consent. The test–retest experiment took place at the same time of day with intervals of 51 ± 5 days (Sessions 1 and 2) and 168 ± 24 days (Sessions 2 and 3). Two of the subjects involved were unfamiliar with the procedure (subjects 1 and 7), the other five were experienced fMRI subjects including one who had participated in a different retinotopic mapping study (subject 6) 10 months prior to session 1 of this study (see Supplementary information for the individual result patterns).

Visual Stimulation for fMRI
In order to avoid differential order effects of visual stimulation conditions across subjects, each of the three fMRI sessions comprised the same sequence of experiments: eccentricity mapping with (a) expanding and (b) contracting rings (scan 1 and 3); polar-angle mapping with rotating wedges (a) counter-clockwise and (b) clockwise (scan 2 and 4); and block design full-field stimulation starting with (a) stimulation on and then (b) stimulation off (scan 5 and 6). The visual stimuli were back projected with a video projector (DLA-G150CL, JVC Ltd.) onto a screen (horizontal viewing angle ~14.8° and vertical viewing angle ~8.3°) behind the subject and viewed via a surface mirror mounted onto the head coil. The stimulus patterns were approximately m-scaled to compensate for the cortical magnification of the visual stimulus (width of the most central and most peripheral ring 0.1° and 1.0°) and comprised a contrast inverting (with 8 reversals per second) circular checkerboard of 7.4° radius consisting of black and white checks with 24 segments and 26 rings (mean luminance 62 cd/m^2, contrast 99%). Identical presentation of the visual stimulation is necessary to allow comparability between the three-time points. The average stimulus luminance of 62 cd/m^2 and 99% contrast were ensured at all three points based on photometer calibration (CS-100A Konica Minolta). To produce the desired luminance and contrast of the stimuli, the gamma function, which relates numeric input and visual output for each projector color channel, has been considered in the computed RGB-values.

The subjects were instructed to fixate on a central-fixation marker (0.34° diameter) and to report any color changes (200 ms duration at intervals between 5 and 10 s) by pressing a button. Each experimental scan started with a 12 s scan dummy stimulation period and was followed by seven 36 s cycles of visual stimulation for rings and rotating wedges and ten 24 s cycles of visual stimulation for full field. Phase-encoded paradigms in retinotopic mapping stimulated different parts of the visual field during different stimulation epochs, i.e., increasing or decreasing polar angle for (i) polar-angle mapping and increasing or decreasing eccentricities for (ii) eccentricity mapping. Thus, only a section of the contrast-reversing checkerboard was presented at a time: (i) Polar angle mapping: Polar angle stimuli stepping through clockwise and counterclockwise direction were presented. The wedge was six segments (90°) wide and stepped 54 times by 6.7° in each of the 36 s cycles. (ii) Eccentricity mapping: The stimulus moved through the eccentricities as an expanding or contracting ring. The ring was 0.97° wide and stepped 54 times by 0.32° in each of the cycles. It expanded uniformly beyond the maximum extent of the screen for 5 s and

then wrapped around to the fovea. This stimulation gap helped to distinguish peripheral from fovea responses in the eccentricity mapping data (Hoffmann et al., 2009). For the third type of visual stimulation, i.e., full field stimulation, subjects viewed 12 s checkerboard reversal (8 reversals per second) followed by 12 s gray (same mean luminance in both epochs) in a block design. Altogether, the fMRI measurements took around 30 min.

fMRI Acquisition

For functional imaging, T2*-weighted echo-planar images (EPI) were acquired with a 7 Tesla whole body MRI scanner (Siemens Healthcare, Erlangen, Germany) using a 32 channel head coil (Nova Medical, Wilmington, MA). The gradient-echo EPI pulse sequence had the following parameters: repetition time (TR) = 3000 ms, echo time (TE) = 21 ms, field of view (FOV) = 140 mm, flip angle (α) = 90, and voxel size = $1 \times 1 \times 1$ mm. Forty-five slices were acquired perpendicular to the calcarine sulcus for a duration 264 s with 88 volumes (dummy stimulation period and 7 cycles) for eccentricity and polar angle mapping, and 252 s with 84 volumes (dummy stimulation period and 10 cycles) for the full-field stimulation. In addition, high-resolution anatomical images of the occipital region were obtained using T1-weighted 3D-MPRAGE with the following parameters: acquisition time = 296 s, TR = 2300 ms, TE = 2.66 ms, inversion time = 1050 ms, FOV = 256 mm, isotropic resolution = 1 mm. 3D-MPRAGE inhomogeneity was corrected by division with 3D gradient echo (GE) reference data (Van de Moortele et al., 2009) without inversion and otherwise identical parameters but: acquisition time = 170 s, TR = 1340 ms. The functional images were corrected online for motion and spatial distortions (In and Speck, 2012).

Over the course of all the experiments, no MR system upgrade or modification took place and no auxiliary equipment had been installed or removed. The MR system performance for fMRI experiments has been tested at least once per week as part of the regular quality assurance procedure and no relevant parameters (SNR, signal stability, signal homogeneity, RF-parameters, EPI ghost level) showed any deviation from the established reference values.

Cortical Flattening

In order to provide homogenous white and gray matter contrast for segmentation, T1-weighted inhomogeneity corrected MPRAGE images (Van de Moortele et al., 2009) were used to segment the cortex and to create a flattened representation of the cortical gray matter (Teo et al., 1997; Wandell et al., 2000). Gray and white matter were segmented using Freesurfer, ITK Gray, 1.6.0.1., and mrGray (VISTA) and were manually edited to minimize segmentation errors. The anatomical segmentation and flattening was done once per subject.

Preprocessing of Functional Images

In the three scanning sessions, identical methods were used for fMRI data analysis. The functional images were analyzed using the Stanford VISTA-Tools (VISTA; http://white.stanford.edu/software/). The T2*-weighted images were aligned to the T1-weighted images' coordinate frame to project the fMRI time series onto the flattened cortex representation. Each voxel's time series (TS) underwent the following analysis steps: (1) dummy stimulation period (12 s or 4 volumes) of each functional run was discarded to avoid transient-onset artifact; (2) final single cycle (36 s, i.e., 12 volumes) at the end of each functional run was also discarded for eccentricity and polar angle due to a reconstruction failure and an incorrect volume setting for two subjects in one of the sessions. These two scans were not excluded because only a few volumes at the end of the scans were involved. Thus, six stimulation cycles were retained for eccentricity and polar angle mapping and 10 stimulation cycles for full-field analysis for each subject; (3) TS were divided by their mean intensity; (4) for one of the opposing directions in the three types of visual stimulation (contracting vs. expanding rings, clockwise vs. counter clockwise rotating wedges, and on vs. off start of full field stimulation) the TS were flipped in time. Then the hemodynamic delay of the flipped TS was corrected by shifting the TS by 6 s (i.e., two TR). (5) TS of opposing directions were averaged to remove or cancel the residual phase lags caused by residual hemodynamic delays of the two stimulus directions. This technique allows examining the true phase of the eccentricity and polar angle (Sereno et al., 1995; Smith et al., 1998; Silver et al., 2005; Larsson and Heeger, 2006); (6) data were Fourier transformed to calculate the amplitude and the response phase for the stimulus frequency; (7) the correlation, technically coherence, was measured (Engel, 1997). The coherence is calculated from the time series with a sinusoid with the frequency having the same value as the fundamental frequency (i.e., 1/36 Hz for 36 s stimulus described above) of the visual stimulation (Engel, 1997). The coherence value was determined independently for each voxel in the functional scan. The coherence value represents the strength of the response of that voxels to each visual stimulation condition.

For qualitative assessment and presentation of the response maps in the flattened representation, the coherence and phase values were smoothed by convolving a Gaussian Kernel (full width at half maximum: 4 mm) with the complex-vector representation of the BOLD response. The statistical significance levels, or p-values, for associated coherence values were estimated according to Silver et al. (2005). For all repeated measures of individual results, a threshold of $p = 0.05$ was applied for multiple comparison (Zandbelt et al., 2008; Plichta et al., 2012): only voxels with coherence values exceeding 0.18 were considered for further analysis. No spatial or temporal filtering was applied for any quantitative analyses (amplitude, coherence, or any responses in the EPI planes) to avoid loss of resolution. Data quality evaluation was performed through direct visual inspection and did not show signal dropouts due to local susceptibility artifact.

The identification of the visual areas V1, V2, and V3 was performed on the data averaged across all three sessions in order to avoid the intrusion of systematic inter-session effects. The boundaries between the visual areas were delineated manually on the cortical flat maps based on the response phase reflected in the false-color maps for polar angle and eccentricity (Engel et al., 1994; Engel, 1997; Wandell et al., 2007). The borders of the visual areas were representations of the vertical and horizontal meridians in the visual field according to established procedures

(Wandell et al., 2007). Subsequently, visual area definitions were used to examine the correlation of phase maps and fMRI signal difference between sessions (Swisher et al., 2007).

Head Motion Assessment

Head motion during the fMRI acquisitions was quantified using the uncorrected functional images, i.e., without the online motion and distortion correction. For the relevant volumes (72 volumes for eccentricity and polar angle, 80 volumes for full field), head motion was quantified for each scan with McFlirt version 5.0 [motion correction using FMRIB's Linear Image Registration Tool (Jenkinson et al., 2002)] and FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). The head motion metrics were evaluated from three translation (x, y, and z in mm) and three rotation (x, y, and z in rad) parameters. Then, a head motion index (in mm) was calculated by the root mean square (RMS) of the six parameters (three rotations and three translations) (Silver et al., 2005).
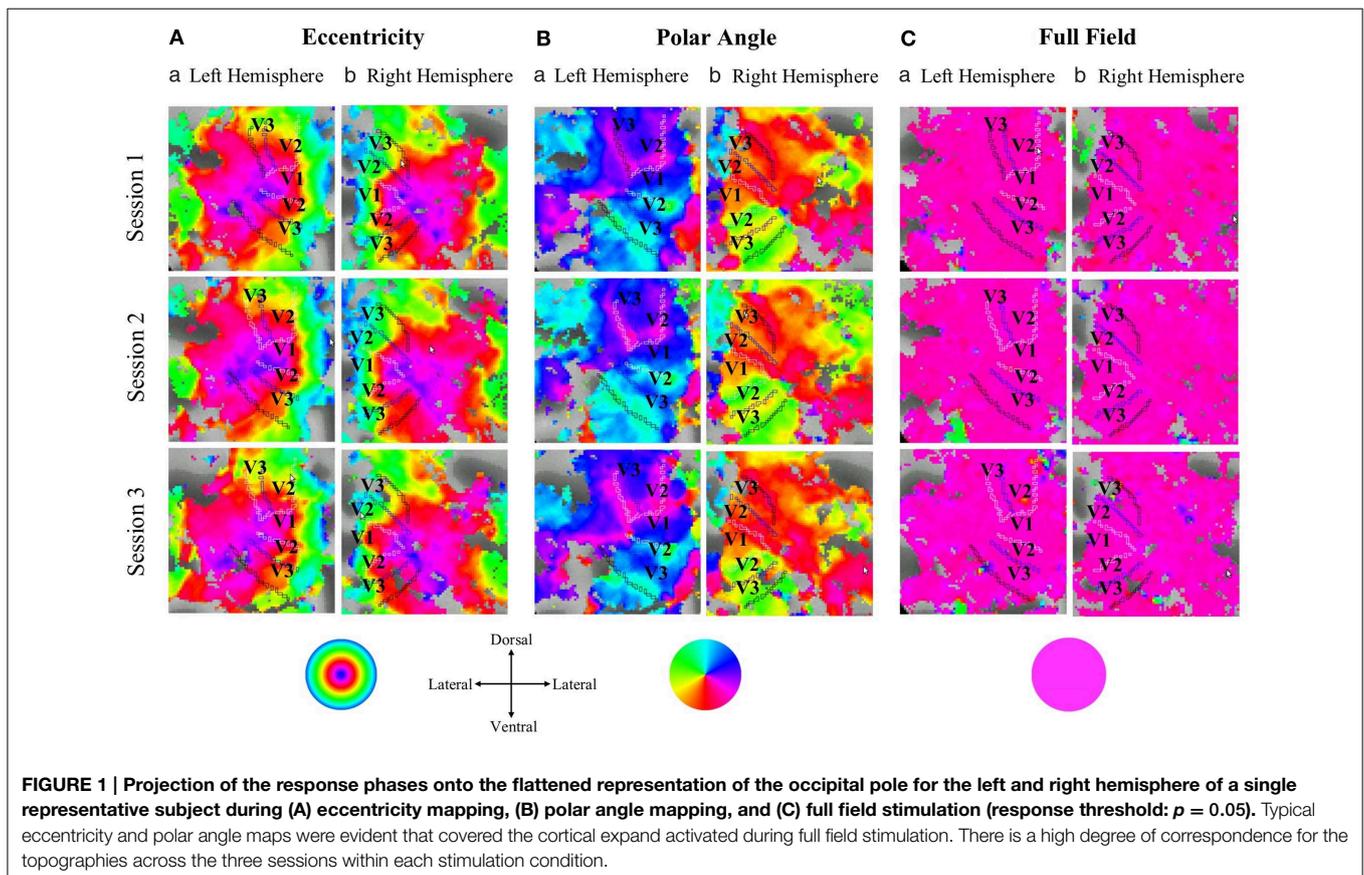
Statistical Analysis

First, in order to explain the data variance and assign it to the different possible sources (e.g., brain areas and condition), we investigated the variability or change in activated cortical surface area, amplitude, and coherence using repeated measures analysis of variance (ANOVA) by session as a within-subject factor, and visual area and stimulation condition as the between-subject factors. Because the coherence values were not normally distributed, they were converted into normally distributed values by Fisher transformation for further statistical analysis. In this study, *post-hoc* test Sidak correction was used to account for multiple testing to detail the factor effect. Second, we assessed the minimal effect sizes to be detected with the current design for 95% confidence intervals. Third, in order to evaluate the similarity of the obtained maps across sessions, we examined the reproducibility of the phase maps in the early visual areas using circular correlation [circular Statistic toolbox in Matlab as suggested by Philipp (2009)]. The resulting correlations were converted by Fisher's z-transformation for further statistical analysis. Quantitative comparison of phase correlation was determined separately for each visual area (V1, V2, and V3) and visual stimulation (eccentricity and polar angle) using independent *t*-tests. Additionally, we evaluated the change in head motion across sessions by repeated measures ANOVA.

Results

Bold Signals

In **Figure 1** flat maps of the occipital pole are depicted for one representative individual to indicate the typical response topography obtained for eccentricity mapping, polar angle mapping, and full field stimulation [see Supplementary information (Figure S5), for other subjects]. A high degree of

FIGURE 1 | Projection of the response phases onto the flattened representation of the occipital pole for the left and right hemisphere of a single representative subject during (A) eccentricity mapping, (B) polar angle mapping, and (C) full field stimulation (response threshold: *p* = 0.05). Typical eccentricity and polar angle maps were evident that covered the cortical expand activated during full field stimulation. There is a high degree of correspondence for the topographies across the three sessions within each stimulation condition.

consistency of the obtained response patterns across sessions 1–3 is evident. For a quantitative account of the effects we determined for each visual area the activated cortical surface area, i.e., the voxels with significant fMRI-responses ($p < 0.05$; **Figure 2**). Subsequently we determined for these super-threshold regions the average amplitude and coherence separately for each visual area and session (**Figures 3**, **4**). The significance of the effects was determined with separate Three-Way repeated measures ANOVAs [factors: session, stimulation condition, visual area; see **Table 1** and Supplementary information (Figures S6–S8)]. Initial Three-Way repeated measures ANOVAs comparing the dorsal and ventral parts of the early visual areas did not indicate any significant effect (individual and interaction factors) on the activated cortical surface area, response amplitude, and coherence ($p > 0.05$). Therefore, the dorsal and ventral parts of the early visual areas V1, V2, and V3 were combined for further analysis.

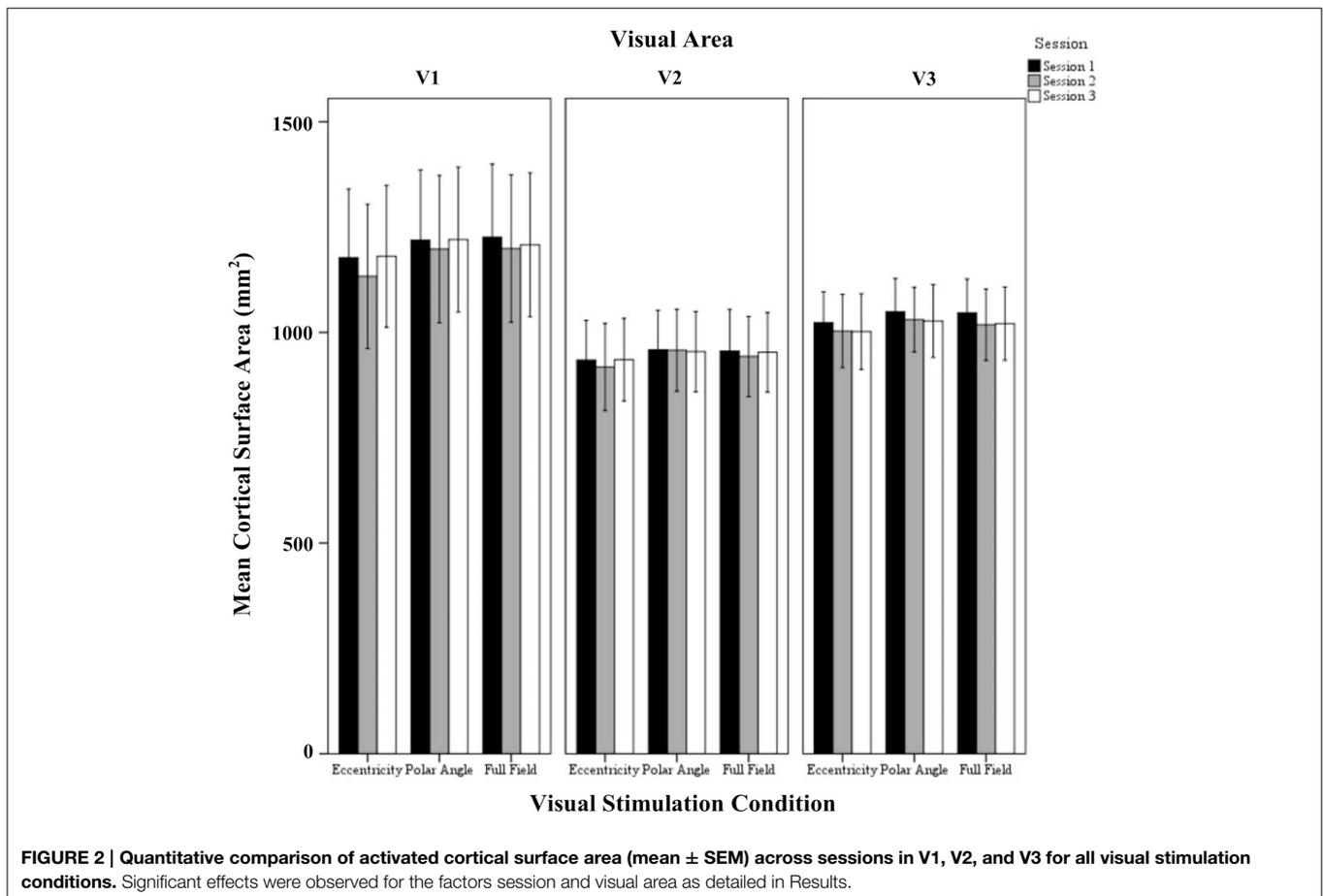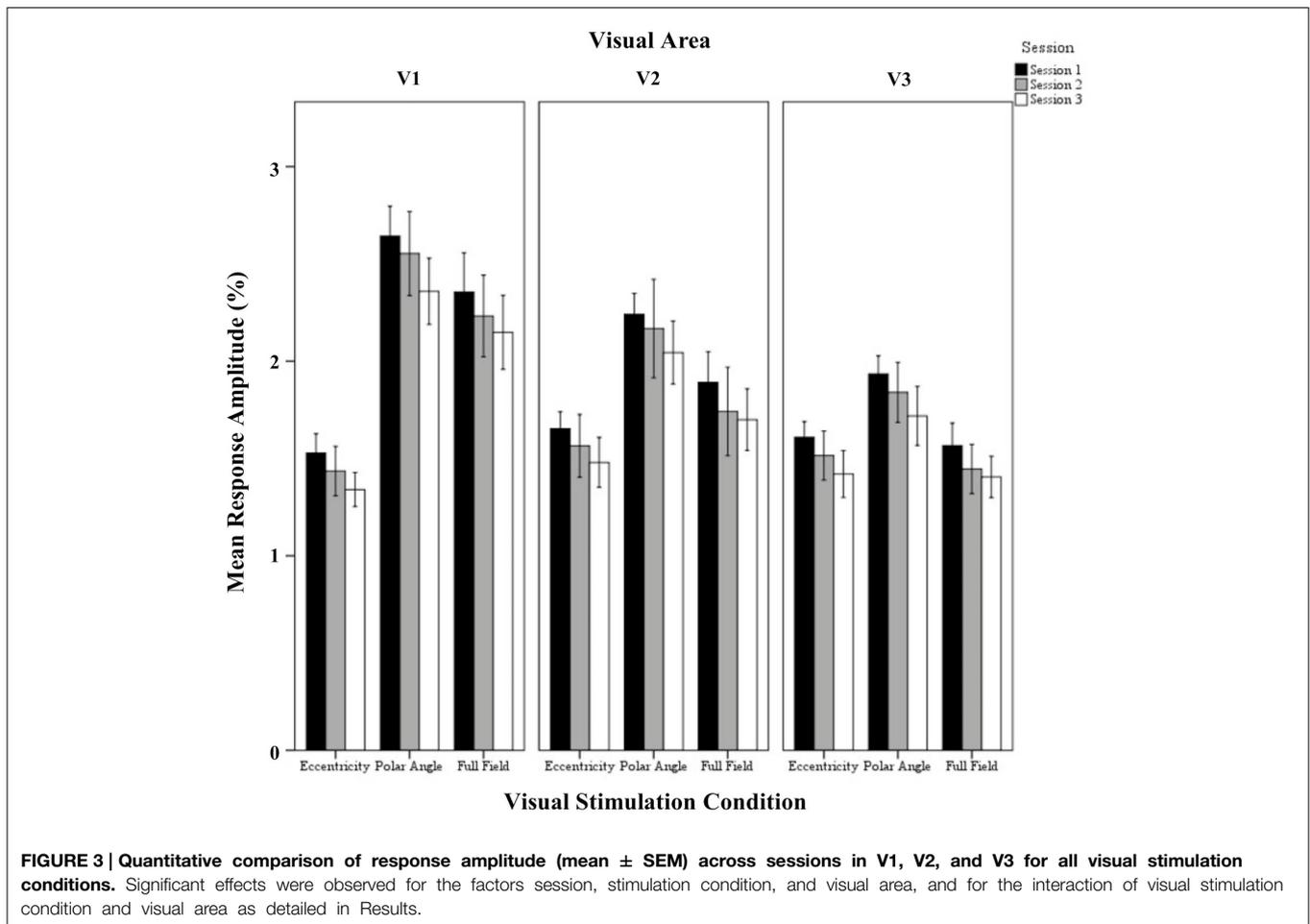### Activated Cortical Surface Area

The repeated measures ANOVA indicated significant effects of the factors session ($p = 0.005$) and visual area ($p = 0.046$), but not visual stimulation condition. No significant interactions of the factors were evident. *Post-hoc* tests were performed to detail the effect of session and visual areas. Importantly, for the factor session, there was a small, but significant reduction of

activated area from sessions 1 to 2 ($1066 \pm 40\,\text{mm}^2$ vs. $1044 \pm 42\,\text{mm}^2$; $p = 0.006$). Furthermore, for the factor visual areas, V1 exceeded V2 significantly ($1196 \pm 71\,\text{mm}^2$ vs. $946 \pm 71\,\text{mm}^2$; $p = 0.045$) which is in accordance with Dougherty et al. (2003).

Response Amplitude

The repeated measures ANOVA indicated highly significant effects of the factors session ($p < 0.001$), visual stimulation condition ($p < 0.001$), and visual area ($p < 0.001$) on response amplitude and a significant interaction only for visual stimulation condition × visual area ($p = 0.019$). Importantly, *post-hoc* tests showed a significant amplitude reduction from sessions 1 to 3 ($1.94 \pm 0.04\%$ vs. $1.74 \pm 0.05\%$; $p < 0.001$). Furthermore, for the factor visual stimulation, *post-hoc* tests showed increased amplitudes for (1) polar angle compared to eccentricity ($2.17 \pm 0.08\%$ vs. $1.51 \pm 0.08\%$; $p < 0.001$), (2) full field compared to eccentricity ($1.83 \pm 0.08\%$ vs. $1.51 \pm 0.08\%$; $p = 0.012$), and (3) polar angle compared to full field ($2.17 \pm 0.08\%$ vs. $1.83 \pm 0.08\%$; $p = 0.010$). Significant differences for the response amplitudes of visual stimulation conditions were expected as they have different stimulation duty cycles, which influences the Fourier-analysis-based quantification of the response amplitudes and coherences (see Materials and Methods). For the factor visual area *post-hoc* tests showed greater amplitudes in V1 than V3 ($2.07 \pm 0.08\%$ vs.

FIGURE 2 | Quantitative comparison of activated cortical surface area (mean ± SEM) across sessions in V1, V2, and V3 for all visual stimulation conditions. Significant effects were observed for the factors session and visual area as detailed in Results.

FIGURE 3 | Quantitative comparison of response amplitude (mean ± SEM) across sessions in V1, V2, and V3 for all visual stimulation conditions. Significant effects were observed for the factors session, stimulation condition, and visual area, and for the interaction of visual stimulation condition and visual area as detailed in Results.

$1.61 \pm 0.08\%$; $p < 0.001$), which is in accordance with Hoffmann et al. (2009).
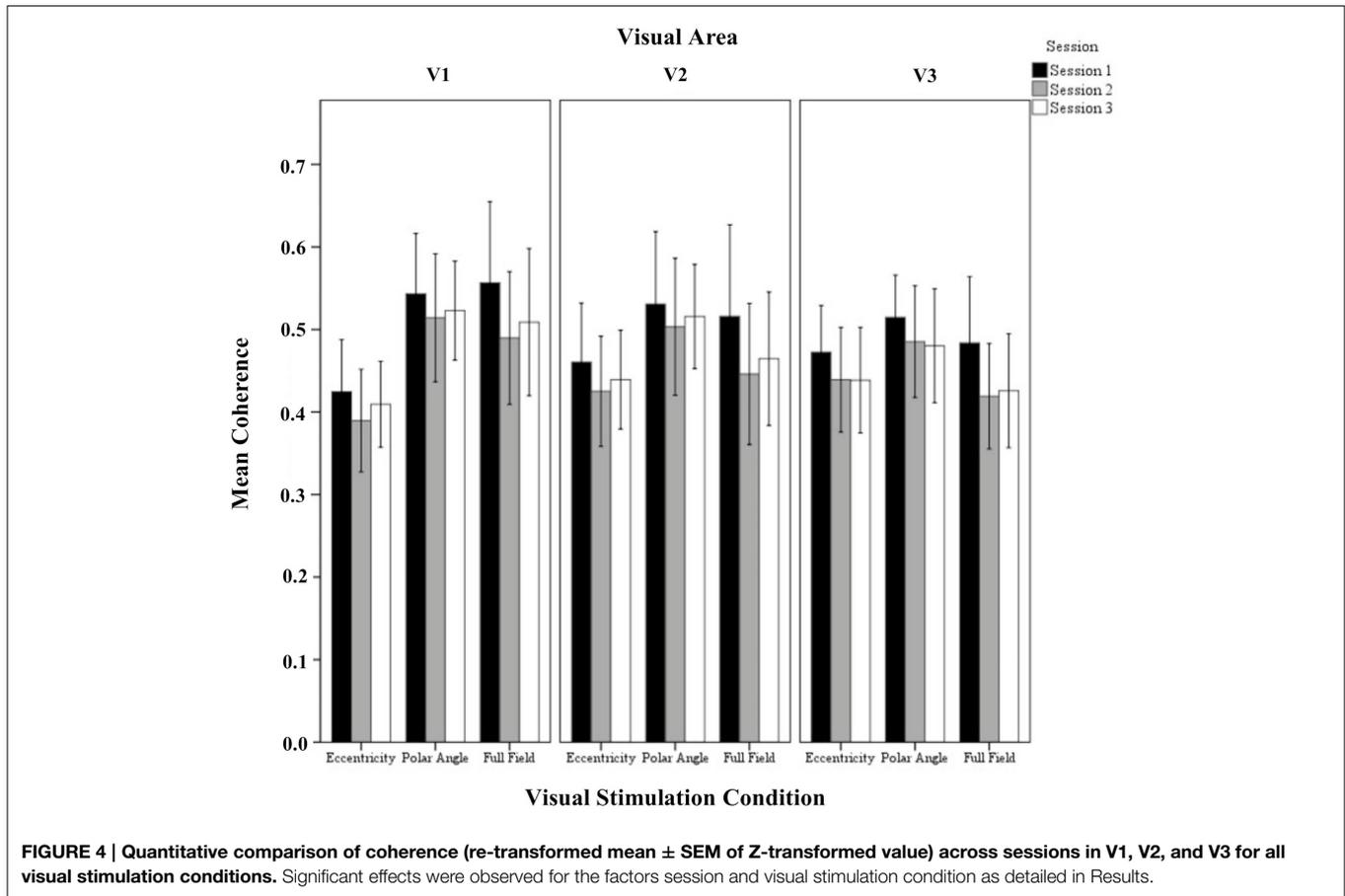
Coherence

The repeated measures ANOVA indicated significant differences for session and visual stimulation condition. No significant interactions of the factors were evident. Importantly, for the factor session, *post-hoc* tests showed a significant reduction from (1) sessions 1 to 2 (0.55 ± 0.01 vs. 0.50 ± 0.01 $p < 0.001$) and (2) sessions 1 to 3(0.55 ± 0.01 vs. 0.51 ± 0.01 $p < 0.001$). Furthermore, for the factor visual stimulation, *post-hoc* tests showed that coherences were greater for polar angle compared to eccentricity (0.57 ± 0.02 vs. 0.46 ± 0.02; $p = 0.001$). Similar to the effect of stimulation on the amplitude of the BOLD signal, it is assumed that this is due to the effects of the different stimulation duty cycles for the different conditions on the Fourier-based analysis of the BOLD signals (see Materials and Methods and above).

Taken together, effects of the factor session, i.e., a reduction of the measures in the later sessions compared to session 1, were evident for each measure tested, independent of the visual area analyzed. Specifically, activated cortical surface area decreased from sessions 1 to 2 by 2% ($p < 0.006$), response amplitudes

decreased from sessions 1 to 3 by 11.5% ($p < 0.001$), and coherence decreased from sessions 1 to 2 by 10% ($p < 0.001$) and session 3 by 8% ($p < 0.001$). Importantly, no session effects between sessions 2 and 3 were evident. This suggests that in session 1 a "novelty effect" contributed to the obtained measures as discussed below. From the scatter of the data, specifically the 95% confidence interval of session 2, the minimal effect sizes can be inferred that can be detected with a 5% significance threshold ($p < 0.05$) for a design similar to that of the present study, i.e., 7 subjects, repeated measures design, and an initial scan to eliminate novelty effects. The 95% confidence interval ranges for cortical surface area ±1.5% of the mean cortical surface area, for response amplitude ±6% for the mean percentage of the amplitude modulation, and for coherence ±5% of the Z-transformed coherence.

Reproducibility of Phase Maps

The phase map similarity or reproducibility within defined visual areas for eccentricity and polar angle were determined (1) between sessions 1 and 2 (S1S2) and (2) between sessions 2 and 3 (S2S3). We applied a threshold of $p = 0.05$ for all visual areas in both sessions for these analyses. **Figure 5** shows the mean correlation coefficient of S1S2 and S2S3 within the visual areas

FIGURE 4 | Quantitative comparison of coherence (re-transformed mean ± SEM of Z-transformed value) across sessions in V1, V2, and V3 for all visual stimulation conditions. Significant effects were observed for the factors session and visual stimulation condition as detailed in Results.

TABLE 1 | Results from repeated measures analyses for (a) activated cortical surface area, (b) responses amplitude, and (c) coherence.

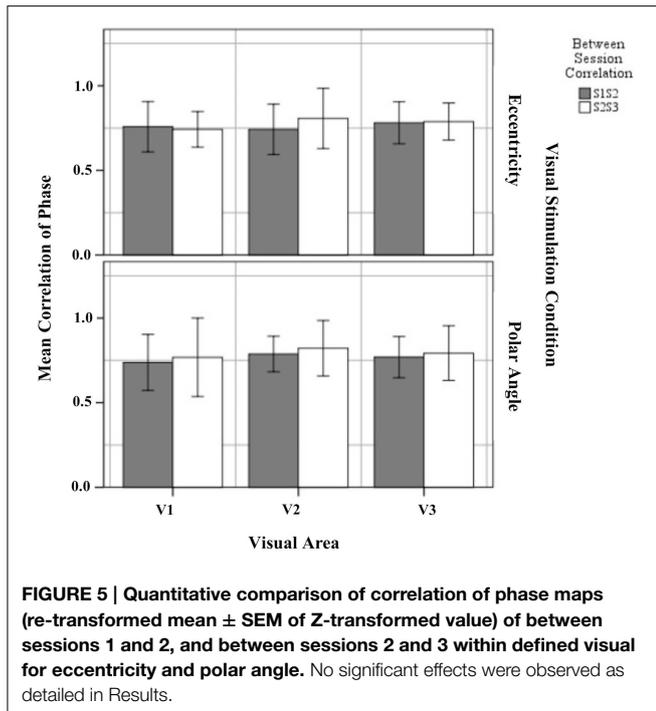| Factor	(a) Activated cortical surface area	(b) Amplitude	(c) Coherence
	F-ratio (df1, df2), p-value	F-ratio (df1, df2), p-value	F-ratio (df1, df2), p-value
Session	$F_{(2, 108)} = 5.590, p = 0.005^*$	$F_{(2, 108)} = 10.184, p < 0.001^*$	[a]$F_{(1.785, 96.417)} = 19.607, p < 0.001^*$
Visual stimulation condition	$F_{(2, 54)} = 0.068, p = 0.934$	$F_{(2, 54)} = 18.467, p < 0.001^*$	$F_{(2, 54)} = 7.242, p = 0.002^*$
Visual area	$F_{(2, 54)} = 3.272, p = 0.046^*$	$F_{(2, 54)} = 8.933, p < 0.001^*$	$F_{(2, 54)} = 0.673, p = 0.514$
Session × Visual area	$F_{(4, 108)} = 1.198, p = 0.316$	$F_{(4, 108)} = 0.057, p = 0.994$	[a]$F_{(3.571, 96.417)} = 0.266, p = 0.881$
Session × Visual stimulation condition	$F_{(4, 108)} = 0.405, p = 0.804$	$F_{(4, 108)} = 0.192, p = 0.942$	[a]$F_{(3.571, 96.417)} = 1.609, p = 0.184$
Visual stimulation condition × Visual area	$F_{(4, 54)} = 0.005, p = 1.000$	$F_{(4, 54)} = 3.247, p = 0.019^*$	$F_{(4, 54)} = 1.381, p = 0.253$
Session × Visual stimulation condition × Visual area	$F_{(8, 108)} = 0.210, p = 0.989$	$F_{(8, 108)} = 0.023, p = 1.000$	[a]$F_{(7.142, 96.417)} = 0.027, p = 1.000$

[a]*Greenhouse-Geisser correction.*
The effect factor is significant at the 0.05 level ($p < 0.05$).

for eccentricity and polar angle. The results showed that phase maps in V1, V2, and V3, for both eccentricity and polar angle, were highly correlated, with a mean correlation coefficient >0.70. This is in agreement with a previous studies on intra-session variability (Hoffmann et al., 2012) and exceeds the value of a study on inter-session variability (Swisher et al., 2007).

Next, the significance of the variability on the Z-transform correlation between session group (S1S2 and S2S3) was determined separately for each visual area (V1, V2, and V3) and visual stimulation (eccentricity and polar angle). The results showed that the difference between the mean correlation coefficient in S1S2 and S2S3 was not statistically significant for eccentricity mapping within visual area V1 [$t_{(12)} = 0.388, p = 0.705$], V2 [$t_{(12)} = -1.404, p = 0.186$], and V3 [$t_{(12)} = -0.211, p = 0.386$] and for polar angle mapping within visual areas V1 [$t_{(12)} = -0.484, p = 0.637$], V2 [$t_{(12)} = -0.938, p = 0.345$], and V3 [$t_{(12)} = -0.600, p = 0.560$]. It is concluded that the phase maps for both eccentricity and polar angle mapping correlate highly across sessions and that there is no systematic intersession effect on the obtained mapping.
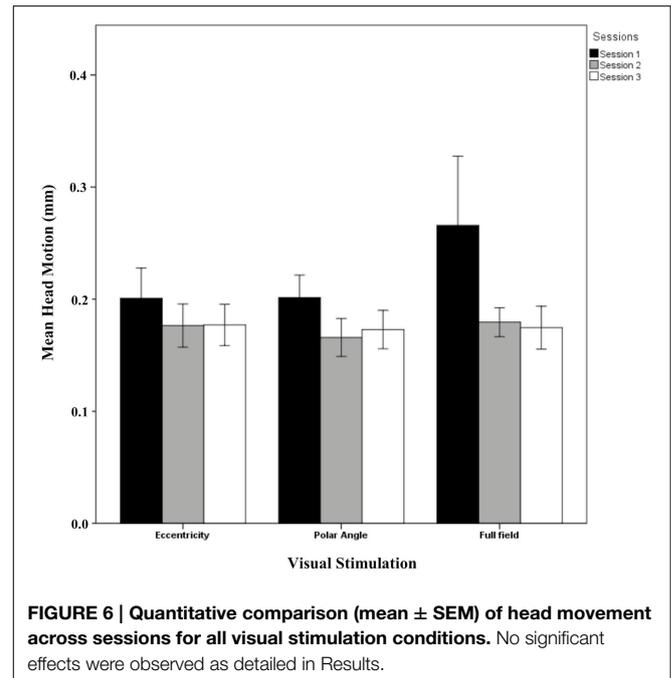
FIGURE 5 | Quantitative comparison of correlation of phase maps (re-transformed mean ± SEM of Z-transformed value) of between sessions 1 and 2, and between sessions 2 and 3 within defined visual for eccentricity and polar angle. No significant effects were observed as detailed in Results.

FIGURE 6 | Quantitative comparison (mean ± SEM) of head movement across sessions for all visual stimulation conditions. No significant effects were observed as detailed in Results.

Head Motion

The results showed that the phase maps were stable across sessions. However, the activation between sessions indicated significant differences from sessions 1 to 2 and 3 whereas no session effects from sessions 2 to 3 were evident. A potential cause of different BOLD responses in session 1 compared to the others might be a different extent of head motion in session 1. **Figure 6** shows the mean head motion across sessions. Overall, the mean head motion was <0.3 mm over time in all stimulation conditions. A repeated measures One-Way ANOVA (factors: session and stimulation condition; see **Table 2**) did not indicate any significant effects on head movement ($p > 0.05$), neither for the individual factors nor for their interaction. A non-significant trend for the factor session ($p = 0.074$), mainly a reduction of head movement by 0.049 mm from sessions 1 to 2, was observed. As the scatter was particularly high for the full field stimulation condition, we excluded this condition in a subsequent analysis. Again, a similar non-significant trend for the factor time ($p = 0.160$) was observed (a reduction of head movement by 0.030 mm). In conclusion, these results do not support head motion as a major cause of the observed novelty effect in the BOLD responses for session 1. This is further underlined by the individual response patterns given in Supplementary information (Additional results and Figures S1–S4).

Discussion

We assessed the inter-session variability of fMRI, for retinotopic mapping and full field stimulation and report sizable, but reduced BOLD responses after the initial session as reflected by decreased activated cortical surface area from sessions 1 to 2, decreased response amplitude from sessions 1 to 3, and decreased coherence from sessions 1 to 2 and 3. Importantly, the findings show no session effect from sessions 2 to 3 and do not support head motion as a major cause of the higher BOLD responses in session 1. Additionally, we examined the correlation coefficient of phase maps between sessions 1 and 2, and between sessions 2 and 3. The results showed that phase maps in early visual areas were highly correlated for eccentricity and polar angle and the correlation of session 2 with sessions 1 and 3 did not differ significantly, demonstrating that the phase maps and the cortical representations were stable.

In the present study, the BOLD responses are reduced after the first session, which might be related to a "novelty effect" associated with this initial scan. This corresponds to several previous reports. Raemaekers et al. (2007) observed a reduction in activation from sessions 1 to 2 (1 week intervals between sessions) in an anti-and prosaccade study. The observed amplitude reduction may reflect lower subject alertness in the follow-up sessions due to familiarity with the procedure (Raemaekers et al., 2007). In a motor task with a stop signal paradigm Zandbelt et al. (2008) did not find inter-session effects in the BOLD signal changes but they observed a reduction of cortical surface area from sessions 1 to 2 (1 week intervals between sessions) and no session effect from sessions 2 to 3 (1 week intervals between sessions). Another potential neuronal mechanism proposed for response decreases is increased neural efficiency with experience (Poldrack, 2000; Kelly et al., 2006; Zandbelt et al., 2008). It is suggested that fewer neurons are needed to fire strongly in response to a particular task (Poldrack, 2000) or a more precise functional circuit is recruited (Garavan et al., 2000). An alternative cause for a reduction in the brain activation pattern might be associated with inter-session differences in head motion during the MRI

TABLE 2 | Results from repeated measures analyses of the head movement for (a) eccentricity, polar angle and full field and (b) eccentricity and polar angle.

Factor	(a)	(b)
	F-ratio (df1, df2), p-value	F-ratio (df1, df2), p-value
Session	[a]$F_{(1.185, 21.322)} = 3.377$, $p = 0.074$	[a]$F_{(1.128, 13.537)} = 2.204$, $p = 0.160$
Visual stimulation condition	$F_{(2, 18)} = 0.276$, $p = 0.762$	$F_{(1, 12)} = 0.024$, $p = 0.879$
Sessions × Visual stimulation condition	[a]$F_{(2.369, 21.322)} = 0.615$, $p = 0.576$	[a]$F_{(1.128, 13.537)} = 0.066$, $p = 0.830$

[a]*Greenhouse-Geisser correction.*

scans. Head movement can lead to erroneous conclusions in fMRI even for normal subjects (Savoy, 2005). In our analysis, however, there was an absence of session-dependence of the head movements, which favors the "novelty effect" hypothesis. It should be noted that several sources of noise influence fMRI data, i.e., physiological (e.g., respiratory, cardiovascular, motion artifacts) and technical noise (e.g., differences in viewing distance, susceptibility). While these noise intrusions enhance the variance of the obtained fMRI data, they are not expected to introduce systematic inter-session effects as observed in the present study. Such noise effects are independent of the session sequence, as has here been tested explicitly for head motion.

The presented data suggest that longitudinal studies on visually driven cortical activity can be confounded by sequential effects, particularly in the initial scan. This is of particular relevance in the face of current therapeutic initiatives to improve or even restore visual function in patients. Visually driven fMRI responses appear to be powerful biomarkers for the efficacy of these treatments (e.g., Ashtari et al., 2011, 2014; Baseler et al., 2011). The present study underscores that a simple baseline vs. post-treatment comparison of fMRI responses may be error-prone. Any future longitudinal fMRI study should therefore carefully evaluate whether the applied task induces stable activation quantified by the respective outcome measure of the study. Only if this is established can all time points be included in the evaluation. In this study we demonstrate, for the visual stimulation paradigm applied, that the initial MRI session may not be representative and could lead to erroneous conclusions in longitudinal fMRI studies. It appears that this might be resolved by the inclusion of a control group, which will show sequential effects that resemble those seen in the patients. It should be noted, however, that in this case, the patient and control group, in addition to the standard age- and sex-matched, should also be matched according to their experience with the scanning paradigm. Furthermore, it is not clear whether novelty effects scale with visual impairment. Therefore, an alternative approach should be taken into consideration. Implementation of initial scanning sessions can contribute to the stability of the BOLD-fMRI signal. This is possible because no inter-session effects between sessions 2 and 3 were evident in this study, both in cortical surface area, coherence, and also in amplitude, even though we can see a linear drop of amplitude (**Figure 3**) from sessions 1 to 3. This absence of significant differences in the mean amplitude between sessions 2 and 3 does not proof the equality of the data, however, it does not show relevant response instabilities either. Importantly, this effect may last for several months, which was the inter-session interval that we used. Consequently, an initial MRI session (indicated as session 1 in this study) before a baseline scan (indicated as session 2 in this study) should be considered in longitudinal study designs. In practice, this initial session could be used to acquire additional MRI scans that are required only once (e.g., anatomical scans, functional localisers). This way, all scans entering the actual longitudinal observation can follow the exact same sequence and protocol. Our experimental design was to scan a sample of 7 subjects at 7 T, leading to a sensitivity that allowed determination of inter-session changes of cortical surface area, response amplitude, and coherence, at 1.5, 6, and 5%, respectively.

Taken together, reduced cortical surface areas, amplitudes, and coherences of visually driven BOLD responses were found for follow-up sessions in comparison to an initial scanning session, while the phase maps were reproducible across sessions. Cognitive effects such as novelty, learning, and adaptation are plausible reasons for the observed inter-session effects. As a consequence, data of an initial session in a longitudinal study might not be representative. This should be taken into consideration when interpreting fMRI results from longitudinal studies.

Author Contributions

Acknowledgments

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fnhum.2015.00477

References

Ashtari, M., Cyckowski, L. L., Monroe, J. F., Marshall, K. A., Chung, D. C., Auricchio, A., et al. (2011). The human visual cortex responds to gene therapy-mediated recovery of retinal function. *J. Clin. Invest.* 121, 2160–2168. doi: 10.1172/JCI57377

Ashtari, M., Cyckowski, L., Yazdi, A., Viands, A., Marshall, K., Bókkon, I., et al. (2014). fMRI of retina-originated phosphenes experienced by patients with leber congenital amaurosis. *PLoS ONE* 9:e86068. doi: 10.1371/journal.pone.0086068

Baseler, H. A., Gouws, A., Crossland, M. D., Leung, C., Tufail, A., Rubin, G. S., et al. (2011). Objective visual assessment of antiangiogenic treatment for wet age-related macular degeneration. *Optom. Vis. Sci.* 88, 1255–1261. doi: 10.1097/OPX.0b013e3182282f13

Bennett, C. M., and Miller, M. B. (2010). How reliable are the results from functional magnetic resonance imaging? *Ann. N.Y. Acad. Sci.* 1191, 133–155. doi: 10.1111/j.1749-6632.2010.05446.x

Bressler, D. W., and Silver, M. A. (2010). Spatial attention improves reliability of fMRI retinotopic mapping signals in occipital and parietal cortex. *Neuroimage* 53, 526–533. doi: 10.1016/j.neuroimage.2010.06.063

Caceres, A., Hall, D. L., Zelaya, F. O., Williams, S. C. R., and Mehta, M. A. (2009). Measuring fMRI reliability with the intra-class correlation coefficient. *Neuroimage* 45, 758–768. doi: 10.1016/j.neuroimage.2008.12.035

Dougherty, R. F., Koch, V. M., Brewer, A. A., Modersitzki, J., and Wandell, B. A. (2003). Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *J. Vis.* 3, 586–598. doi: 10.1167/3.10.1

Engel, S. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cereb. Cortex* 7, 181–192. doi: 10.1093/cercor/7.2.181

Engel, S. A., Rumelhart, D. E., Wandell, B. A., Lee, A. T., Glover, G. H., Chichilnisky, E. J., et al. (1994). fMRI of human visual cortex. *Nature* 369, 525. doi: 10.1038/369525a0

Fernández, G., Specht, K., Weis, S., Tendolkar, I., Reuber, M., Fell, J., et al. (2003). Intrasubject reproducibility of presurgical language lateralization and mapping using fMRI. *Neurology* 60, 969–975. doi: 10.1212/01.WNL.0000049934.34209.2E

Friedman, L., Stern, H., Brown, G. G., Mathalon, D. H., Turner, J., Glover, G. H., et al. (2008). Test-retest and between-site reliability in a multicenter fMRI study. *Hum. Brain Mapp.* 29, 958–972. doi: 10.1002/hbm.20440

Garavan, H., Kelley, D., Rosen, A., Rao, S. M., and Stein, E. A. (2000). Practice-related functional activation changes in a working memory task. *Microsc. Res. Tech.* 51, 54–63. doi: 10.1002/1097-0029(20001001)51:1<54::AID-JEMT6>3.0.CO;2-J

Hoffmann, M. B., Kaule, F. R., Levin, N., Masuda, Y., Kumar, A., Gottlob, I., et al. (2012). Plasticity and stability of the visual system in human achiasma. *Neuron* 75, 393–401. doi: 10.1016/j.neuron.2012.05.026

Hoffmann, M. B., Stadler, J., Kanowski, M., and Speck, O. (2009). Retinotopic mapping of the human visual cortex at a magnetic field strength of 7T. *Clin. Neurophysiol.* 120, 108–116. doi: 10.1016/j.clinph.2008.10.153

Huettel, S. A., Song, A. W., and McCarthy, G. (2008). *Functional Magnetic Resonance Imaging: Data Acquisition and Analysis |Health Sciences and Technology |MIT OpenCourseWare. 2nd Edn.* Sunderland, MA: Sinauer Assoc. Sinauer Associates. Available online at: http://ocw.mit.edu/courses/health-sciences-and-technology/hst-583-functional-magnetic-resonance-imaging-data-acquisition-and-analysis-fall-2008/readings/

In, M.-H., and Speck, O. (2012). Highly accelerated PSF-mapping for EPI distortion correction with improved fidelity. *MAGMA.* 25, 183–192. doi: 10.1007/s10334-011-0275-6

Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17, 825–841. doi: 10.1006/nimg.2002.1132

Johnstone, T., Somerville, L. H., Alexander, A. L., Oakes, T. R., Davidson, R. J., Kalin, N. H., et al. (2005). Stability of amygdala BOLD response to fearful faces over multiple scan sessions. *Neuroimage* 25, 1112–1123. doi: 10.1016/j.neuroimage.2004.12.016

Kelly, C., Foxe, J. J., and Garavan, H. (2006). Patterns of normal human brain plasticity after practice and their implications for neurorehabilitation. *Arch. Phys. Med. Rehabil.* 87(12 Suppl. 2), S20–S29. doi: 10.1016/j.apmr.2006.08.333

Kong, J., Gollub, R. L., Webb, J. M., Kong, J.-T., Vangel, M. G., and Kwong, K. (2007). Test-retest study of fMRI signal change evoked by electroacupuncture stimulation. *Neuroimage* 34, 1171–1181. doi: 10.1016/j.neuroimage.2006.10.019

Krüger, G., and Glover, G. H. (2001). Physiological noise in oxygenation-sensitive magnetic resonance imaging. *Magn. Reson. Med.* 46, 631–637. doi: 10.1002/mrm.1240

Larsson, J., and Heeger, D. J. (2006). Two retinotopic visual areas in human lateral occipital cortex. *J. Neurosci.* 26, 13128–13142. doi: 10.1523/JNEUROSCI.1657-06.2006

Maitra, R., Roys, S. R., and Gullapalli, R. P. (2002). Test-retest reliability estimation of functional MRI data. *Magn. Reson. Med.* 48, 62–70. doi: 10.1002/mrm.10191

Philipp, B. (2009). CircStat: A MATLAB toolbox for circular statistics. *J. Stat. Softw.* 31, 1–22. Available online at: http://www.jstatsoft.org/v31/i10

Plichta, M. M., Schwarz, A. J., Grimm, O., Morgen, K., Mier, D., Haddad, L., et al. (2012). Test-retest reliability of evoked BOLD signals from a cognitive-emotive fMRI test battery. *Neuroimage* 60, 1746–1758. doi: 10.1016/j.neuroimage.2012.01.129

Poldrack, R. A. (2000). Imaging brain plasticity: conceptual and methodological issues–a theoretical review. *Neuroimage* 12, 1–13. doi: 10.1006/nimg.2000.0596

Raemaekers, M., Vink, M., Zandbelt, B., van Wezel, R. J. A., Kahn, R. S., and Ramsey, N. F. (2007). Test-retest reliability of fMRI activation during prosaccades and antisaccades. *Neuroimage* 36, 532–542. doi: 10.1016/j.neuroimage.2007.03.061

Savoy, R. L. (2005). Experimental design in brain activation MRI: cautionary tales. *Brain Res. Bull.* 67, 361–367. doi: 10.1016/j.brainresbull.2005.06.008

Seibert, T. M., Majid, D. S. A., Aron, A. R., Corey-Bloom, J., and Brewer, J. B. (2012). Stability of resting fMRI interregional correlations analyzed in subject-native space: a one-year longitudinal study in healthy adults and premanifest Huntington's disease. *Neuroimage* 59, 2452–2463. doi: 10.1016/j.neuroimage.2011.08.105

Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268, 889–893. doi: 10.1126/science.7754376

Silver, M. A., Ress, D., and Heeger, D. J. (2005). Topographic maps of visual spatial attention in human parietal cortex. *J. Neurophysiol.* 94, 1358–1371. doi: 10.1152/jn.01316.2004

Smith, A. T., Greenlee, M. W., Singh, K. D., Kraemer, F. M., and Hennig, J. (1998). The processing of first- and second-order motion in human visual cortex assessed by functional magnetic resonance imaging (fMRI). *J. Neurosci.* 18, 3816–3830.

Specht, K., Willmes, K., Shah, N. J., and Jäncke, L. (2003). Assessment of reliability in functional imaging studies. *J. Magn. Reson. Imaging* 17, 463–471. doi: 10.1002/jmri.10277

Swisher, J. D., Halko, M. A., Merabet, L. B., McMains, S. A., and Somers, D. C. (2007). Visual topography of human intraparietal sulcus. *J. Neurosci.* 27, 5326–5337. doi: 10.1523/JNEUROSCI.0991-07.2007

Teo, P. C., Sapiro, G., and Wandell, B. A. (1997). Creating connected representations of cortical gray matter for functional MRI visualization. *IEEE Trans. Med. Imaging* 16, 852–863. doi: 10.1109/42.650881

Van de Moortele, P.-F., Auerbach, E. J., Olman, C., Yacoub, E., Ugurbil, K., and 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, 432–446. doi: 10.1016/j.neuroimage.2009.02.009

Wandell, B. A., Chial, S., and Backus, B. T. (2000). Visualization and measurement of the cortical surface. *J. Cogn. Neurosci.* 12, 739–752. doi: 10.1162/089892900562561

Wandell, B. A., Dumoulin, S. O., and Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron* 56, 366–383. doi: 10.1016/j.neuron.2007.10.012

Woldorff, M. G., Fox, P. T., Matzke, M., Lancaster, J. L., Veeraswamy, S., Zamarripa, F., et al. (1997). Retinotopic organization of early visual spatial attention effects as revealed by PET and ERPs. *Hum. Brain Mapp.* 5, 280–286.

World Medical Association. (2000). Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* 284, 3043–3045. doi: 10.1001/jama.284.23.3043

Zandbelt, B. B., Gladwin, T. E., Raemaekers, M., van Buuren, M., Neggers, S. F., Kahn, R. S., et al. (2008). Within-subject variation in BOLD-fMRI signal changes across repeated measurements: quantification and implications for sample size. *Neuroimage* 42, 196–206. doi: 10.1016/j.neuroimage.2008.04.183