Supplementary Methods

Descriptives on phase trigger to stimulus onset delays

The experimental setup in the current studies (described in ten Oever et al., 2016) involved sub-millisecond precise and extensively validated timing, when it comes to the sending of 'stimulus triggers' relative to tACS phase angle. In other words, at the level of these triggers, the phase conditions are near-perfect. In experiments 1 and 2 of the current series of studies, we could not take full advantage of this precision. Upon receiving the stimulus triggers, the stimulus computer had to process these triggers, prepare trigger value based stimuli (i.e. left or right hemifield, stimulus parameters based on predetermined condition matrix, etc.), and send a command to the monitor. Importantly, this monitor could then only actually present the stimulus with a delay tied to the monitor refresh rate. Taking all this into account, we here provide extensive descriptives of the delays caused by this processing pipeline, across participants and conditions. As expected, there is a variance in delays in line with the monitor refresh rate. Also as expected, there is no clear difference in these descriptives across conditions or participants.

To determine the delay between stimulus trigger (i.e. tACS phase angle) and actual stimulus presentation, we used timestamps provided by our experiment code in Matlab and Psychtoolbox. There is a caveat here; the true onset of visual stimuli is still not captured by these timestamps. Once the monitor 'presents' the stimulus, there is a further delay depending for instance on visual field location. Only the use of a light diode can timestamp the actual onset perfectly. However, these further delays should be extremely consistent (as assessed in our lab repeatedly across different situations) and therefore do not affect comparisons between conditions. After all, most important in evaluating these delays is *not* the absolute delay from tACS phase to stimulus onset (on average per experiment: 13.5ms exp1, 17.5 ms exp2). In fact the absolute delay is completely irrelevant for our research questions and analyses; it should only be acknowledged that there *is* an absolute delay between tACS phase angles and the visual stimulus onsets. From this follows only that, in the case of potential phase-condition specific effects, no conclusions should be drawn about the relevance of specific oscillatory phases. What this setup instead allows, is inferences about patterns across tACS phases; are there differences between phases or more specifically is there an oscillatory pattern in outcome measures across tACS phases.

However, to address this question it *is* relevant to know how *consistent* the delays are 1) from trial to trial, and 2) from phase condition to phase condition. To estimate these delay consistencies, we calculated the 95% range of delays, per participant and condition. These values, in milliseconds, reflect the difference between the 2.5^{th} and 97.5^{th} percentile of recorded trigger-stimulus delays, and thus an indication of consistency trial-by-trial. For example, with a hypothetical range value of 15 ms for a given participant/condition, 95% of trials had delays within 15 milliseconds from each other. This scale and approach seemed to us a useful and intuitive marker of delay consistency, because 1) for 10 Hz tACS, the six tACS phase angle conditions were separated by approximately 16 milliseconds, and 2) a single frame of the monitor refresh rate was 16.7 milliseconds. As can be seen below, generally the recorded delays indeed ranged within 16 milliseconds, and rather consistently across experimental (phase and hemifield) conditions.

	left hemifield						right hemifield					
participant	phase 1	phase 2	phase 3	phase 4	phase 5	phase 6	phase 1	phase 2	phase 3	phase 4	phase 5	phase 6
1	14.78	14.60	23.53	17.84	14.77	18.40	16.15	15.61	19.38	14.93	15.92	21.34
2	15.57	16.18	15.86	15.87	15.59	16.99	16.73	16.82	16.30	16.28	15.91	16.37
3	15.47	15.50	14.67	14.27	15.73	15.65	15.40	15.22	15.44	15.08	14.85	15.12
4	14.57	14.78	14.77	14.20	15.09	16.01	13.23	15.96	15.60	13.63	14.31	15.12
5	16.21	28.24	14.55	12.88	14.55	15.79	14.61	14.55	15.21	14.73	16.04	15.36
6	15.59	15.43	15.46	15.33	15.82	13.28	14.97	14.95	14.88	15.56	15.34	15.00
7	15.78	16.27	14.54	15.86	15.52	13.65	17.55	15.73	15.30	15.36	15.93	14.36
8	16.03	16.09	16.10	16.48	15.36	16.10	16.28	16.06	15.68	17.42	16.20	16.55
9	14.63	17.85	15.81	15.58	15.64	17.16	15.36	15.82	14.70	17.40	15.06	15.71
10	15.38	13.72	14.60	15.67	17.04	14.01	14.90	14.47	14.79	14.51	14.98	15.55
11	15.94	16.85	15.68	16.06	16.27	14.36	14.39	14.03	15.82	15.74	15.79	15.45
12	18.76	15.82	16.43	19.11	14.90	15.93	22.01	15.39	15.79	19.21	20.61	14.55
13	15.69	14.80	15.27	16.10	15.92	15.53	15.25	15.16	13.93	14.57	16.10	15.14
14	15.91	15.18	14.64	15.49	15.25	16.81	15.38	16.11	14.72	14.53	16.21	14.90
median	**15.64**	**15.66**	**15.36**	**15.76**	**15.56**	**15.86**	**15.37**	**15.50**	**15.37**	**15.22**	**15.91**	**15.25**
variance	**1.02**	**12.44**	**5.30**	**2.27**	**0.41**	**2.11**	**4.24**	**0.57**	**1.59**	**2.22**	**2.15**	**2.95**

Supplementary Table 1: experiment 1 95% range values of recorded stimulus trigger to stimulus delays, in milliseconds, per experimental condition (hemifield and phase condition) and participant.

	left hemifield						right hemifield					
participant	phase 1	phase 2	phase 3	phase 4	phase 5	phase 6	phase 1	phase 2	phase 3	phase 4	phase 5	phase 6
1	15.81	18.36	14.85	16.57	15.88	15.56	15.88	17.30	16.43	16.21	14.70	15.31
2	15.42	17.59	15.29	16.88	15.11	13.68	15.46	15.98	14.83	15.20	15.55	15.10
3	15.26	18.28	15.55	16.05	15.97	17.61	16.17	17.99	15.71	15.00	15.16	16.02
4	16.29	15.61	15.53	14.98	16.03	17.48	15.94	19.25	15.13	16.26	16.15	15.71
5	15.35	18.92	14.99	15.59	16.29	14.76	14.48	15.70	15.91	13.91	15.85	15.37
6	16.13	15.57	15.47	15.55	15.07	15.22	14.95	19.66	15.90	15.96	15.31	15.39
7	15.81	15.42	15.03	16.17	18.62	15.41	14.76	14.42	15.60	16.12	15.02	14.91
8	14.08	15.69	14.95	15.41	15.08	15.17	15.73	14.67	17.46	15.60	15.20	15.47
9	15.09	15.91	15.91	16.88	16.51	15.36	16.55	14.85	19.68	18.78	15.73	15.57
10	15.29	15.38	15.45	16.37	14.94	15.92	16.05	14.66	15.20	15.26	15.90	14.79
median	**15.38**	**15.80**	**15.37**	**16.11**	**15.93**	**15.39**	**15.81**	**15.84**	**15.80**	**15.78**	**15.43**	**15.38**
variance	**0.39**	**2.06**	**0.11**	**0.42**	**1.20**	**1.39**	**0.45**	**3.90**	**2.06**	**1.59**	**0.21**	**0.13**

Supplementary Table 2: experiment 2 95% range values of recorded stimulus trigger to stimulus delays, in milliseconds, per experimental condition (hemifield and phase condition) and participant.

		left hemifield					right hemifield					
	estimator	phase 1	phase 2	phase 3	phase 4	phase 5	phase 6	phase 1	phase 2	phase 3	phase 4	phase 5	phase 6
experiment 1	median of range value	15.64	15.66	15.36	15.76	15.56	15.86	15.37	15.50	15.37	15.22	15.91	15.25
	mean of median delays	13.13	13.29	13.27	13.22	13.89	13.53	13.58	12.89	13.22	13.69	13.30	12.81
	st.dev. of median delays	4.71	4.60	4.32	4.67	4.61	4.43	4.96	4.33	4.38	4.62	5.11	4.87
experiment 2	median of range value	15.38	15.80	15.37	16.11	15.93	15.39	15.81	15.84	15.80	15.78	15.43	15.38
	mean of median delays	18.24	17.81	17.27	16.57	16.83	17.36	17.51	17.83	17.66	17.20	17.03	17.71
	st.dev. of median delays	2.82	2.76	3.00	3.72	2.11	2.53	1.64	2.71	2.81	2.75	2.52	3.05

Supplementary Table 3: not separated for participants, but still for experimental conditions, summary statistics for both experiments in milliseconds. Median of range value *is the same as the medians in SM tables 1 and 2.* Mean of median delays *is the average of the individual medians of* absolute *recorded delays, which we propose is irrelevant but is here reported for transparency. St. dev. of median delays is the standard deviation of the individual medians, thus corresponding to the mean of median delays but offering a summary statistics of consistency across participants.*

Supplementary Results

Since the current report obtained null findings on its primary outcome variables, with the planned analyses, any findings presented below must be considered post-hoc and interpreted cautiously. They make inspire future studies or analysis strategies. Presented here are:

- High- versus low-intensity tACS sub-samples, for exp 1.
- Results for all experiments in alternative analysis: no Z-scoring prior to curve fitting.
- Results for all experiments in alternative analysis: fast fourier transform approach.

1. *Experiment 1: high- and low-intensity tACS sub-samples*

In experiment 1, stimulation intensity varied across participants, prompting the question whether group-level results might differ for participants that were stimulated with lower intensity tACS as compared to participants stimulated with higher intensity tACS. The group analyses as previously described were repeated for two halves of the (previously included) participant sample based on median split of stimulation intensity: low intensities (n = 6, range 0.8-1.3 mA peak-to-peak, mean 1.1 mA) and high intensities (n = 6, range 1.6-2.0 mA peak-to-peak, mean 1.9 mA). This split matched the distribution of stimulation intensities reasonably well (figure S1 below).

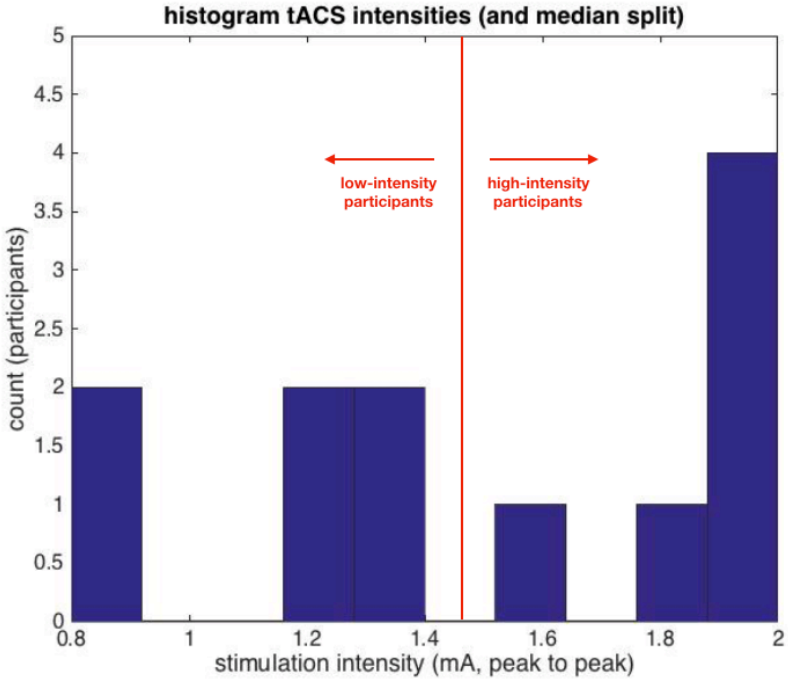


Figure S1: stimulation intensities in experiment 1 and median split.

The non-parametric fixed-effects group-level analysis phase-shifted individual outcome measures (contrast threshold, reaction time) based on peak performance across phase conditions, then averaged these shifted data across participants, and then fitted a single-cycle sinusoid to the group average after removing the data point used for phase-alignment. Relevance values quantified the goodness of fit as well as magnitude of modulation. Identical analysis on 2000 permutations of trial-level phase labels provided null distributions to evaluate the observed relevance values statistically.

In the low-intensity sample, no effects were found for either contrast thresholds (left hemifield: p = 0.87, right hemifield: p = 0.41) or reaction times (left: p = 0.23, right: p = 0.90). In the high-intensity sample, there was now a peri-threshold phase modulation of contrast thresholds in the right hemifield (p = 0.05), not left (p = 0.93), and a peri-threshold phase modulation of reaction times in the left hemifield (p = 0.07), not right (p = 0.23).

The second group-level analysis performed analogous analysis on individual participant data, providing hemifield and participant specific p-values based on observed relevance value versus a permutation-based null distribution. For group-level analysis, these individual p-values were transformed into Z-scores, which were tested against 0 in a one-sided t-test. Note that, here, these t-tests were based only one 6 participants per sub-sample.

In the low-intensity sample, no effects were found for contrast threshold (left: p = 0.96, right: p = 0.11), but there was peri-threshold evidence of phase modulation of reaction times in left (p = 0.05) and also right (0.08) hemifields. In the high-intensity sample, no effects were found for contrast threshold (left: p = 0.88, right: p = 0.33). There was a significant phase modulation of reaction times in left hemifield (p = 0.007), but not right (p = 0.32).

The final group-level analysis phase-aligned individual behavioral outcomes (as in the first described analysis), and then compared the average outcome for shifted phases 1 and 3 (leaving out 2, since it was used for the phase-shifting) versus phases 4, 5, 6, both averaged over participants. This comparison was achieved by subtracting these two values, and then testing this difference against a null distribution based on identical computations across 2000 permutations of trial phase-labels.

For the low-intensity sample, this analysis yielded no effects for either thresholds (left: p = 0.98, right: p = 0.29) or reaction times (left: p = 0.33, right: p = 0.83). For the high-intensity sample, mirroring results of the full sample described in main test, there were no effects for left hemifield (threshold: p = 0.85, reaction times: p = 0.13). There was a peri-threshold modulation of right hemifield thresholds (p = 0.06) and no effect on reaction times (p = 0.38).

Overall, these group-level analyses on the two intensity-based participant sub samples seem coherent with the overall group analyses reported in the main text. It appears that the post-hoc effects most consistently observed across all experiments and analyses, a left-hemifield reaction time modulation, was in experiment 1 predominantly driven by the participants with high stimulation intensity. Also the trend modulations of right hemifield thresholds, observed across the full sample, seem more apparent in the high-intensity sub sample. Future work might take these as preliminary indications that right hemifield performance should be considered, alongside left hemifield performance. But since we did not hypothesize an effect in right hemifield this should be interpreted with caution.

2. *No Z-scoring prior to curve fitting*

In the main analysis, we Z-scored individual performance patterns per hemifield, prior to group analyses. These concern the analysis of phase-shifting (aligning) individual performance patterns, where performance patterns 1) were Z-scored, then 2) removing the alignment-source data point, and 3) averaging across participants. Then we fitted the best-fitting sinusoid in one analysis, or compared the mean score between the 'up-phase' and 'down-phase' half cycles of the group average performance pattern. P-values were always determined based on a permutation test; shuffling phase labels at the trial level within subject and hemifield and recalculating the outcome measure, 2000 times to create a null distribution. P-values reported reflect the proportion of this null distribution with outcomes equal to or greater than the actually observed outcome. We here report the analogous P-

values for the same analyses but foregoing the step of Z-scoring. The rationale for Z-scoring was that individual participants with huge modulations should have less impact on the group average outcomes, but the downside is that – also in permutated data – the 'extent of modulation' is somewhat lost. This was a post-hoc consideration, so these results are presented here in supplementary material for transparency and completeness.

no Z scoring		left hemifield		right hemifield	
		thresh./HR	RT	thresh./HR	RT
exp 1	group curve fitting	0.75	**0.01**	0.09	0.38
	up- vs down-phase	0.94	**0.02**	**0.04**	0.37
exp 2	group curve fitting	0.72	**0.05**	*0.07*	0.41
	up- vs down-phase	0.56	*0.06*	0.12	0.33
exp 3	group curve fitting	0.92			
	up- vs down-phase	0.66			

Supplementary Table 4: P-values from permutation tests analogous to main text analyses but without the processing step of Z-scoring. Uncorrected. Bold values are <0.05, Italics <0.01

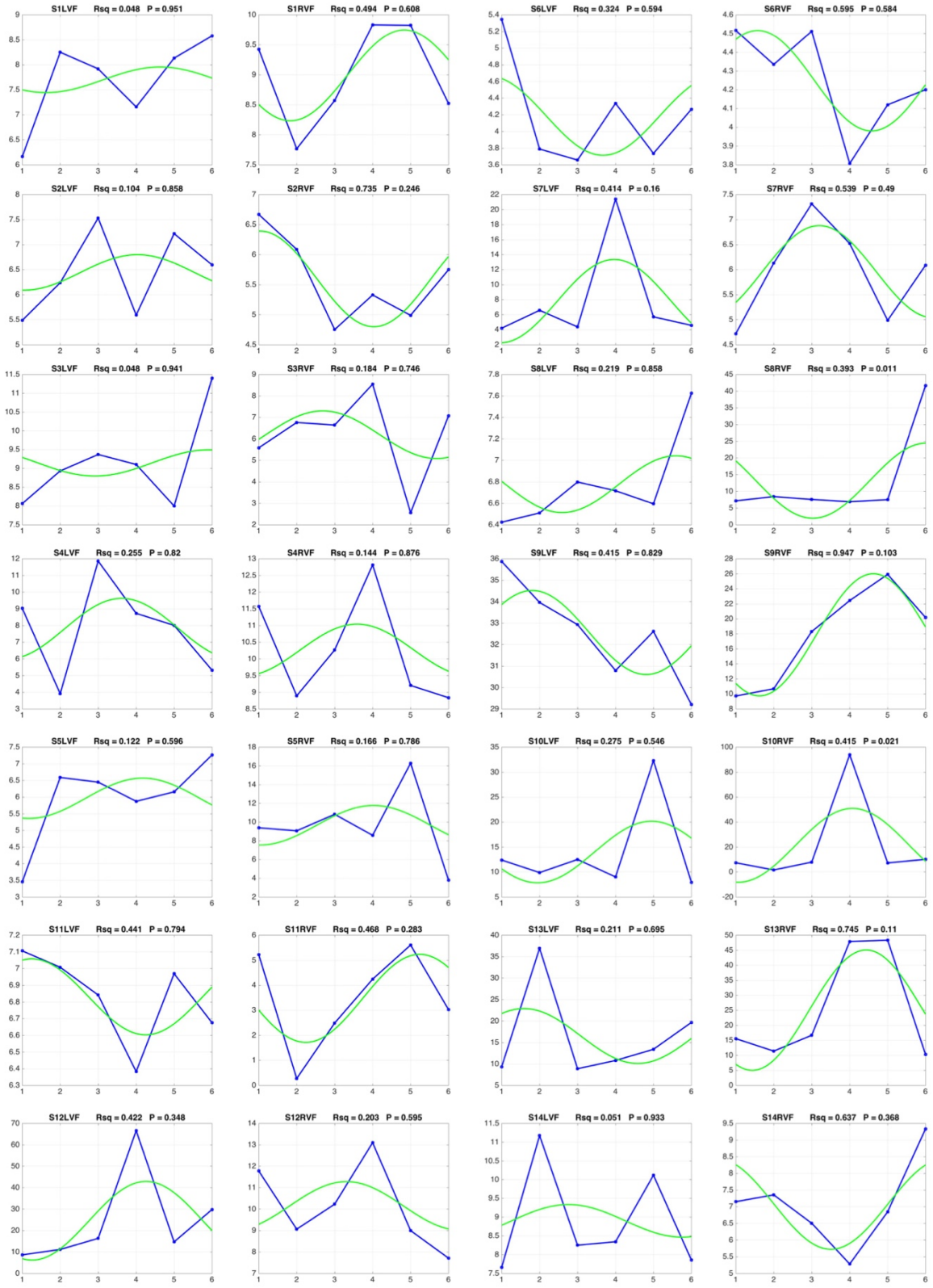
3. Alternative analysis: FFT

We had a priori devised an analysis approach that revolves around curve fitting: the fitting of sinusoids to behavioral patterns and evaluating the goodness of fit and extent of modulation. Prompted by an anonymous reviewer we performed a final alternative analysis, in which observed, and then to create a null distribution also permuted, performance patterns were subjected to fast fourier transform (FFT). We took FFT of the Z-scored performance pattern, squared its absolute output, and took the second peak of the result as outcome variable (see Fiebelkorn et al. 2013). We could apply this approach to two of our three group analyses: the second-level t-test and the group average performance analysis. As reminder, in the second-level t-test analysis we first performed statistical analysis on each individual participant (and hemifield and dependent variable), to obtain individual P-values. Here, these were based on a permutation test of this FFT approach. Then, we converted these P values to Z values, and performed a one-sided t-test against zero. In the group average performance analysis, or 'group curve fitting' to maintain nomenclature, we now made one change to accommodate our implementation of the FFT: we did not remove the data points used or the phase-alignment across participants (i.e. we did not remove the 'peak' in the group average curve). Note that this procedure was the same in both observed and permuted data. We do not here report the up- vs down-phase comparison since there was no curve fitting (or FFT) involved in that analysis.

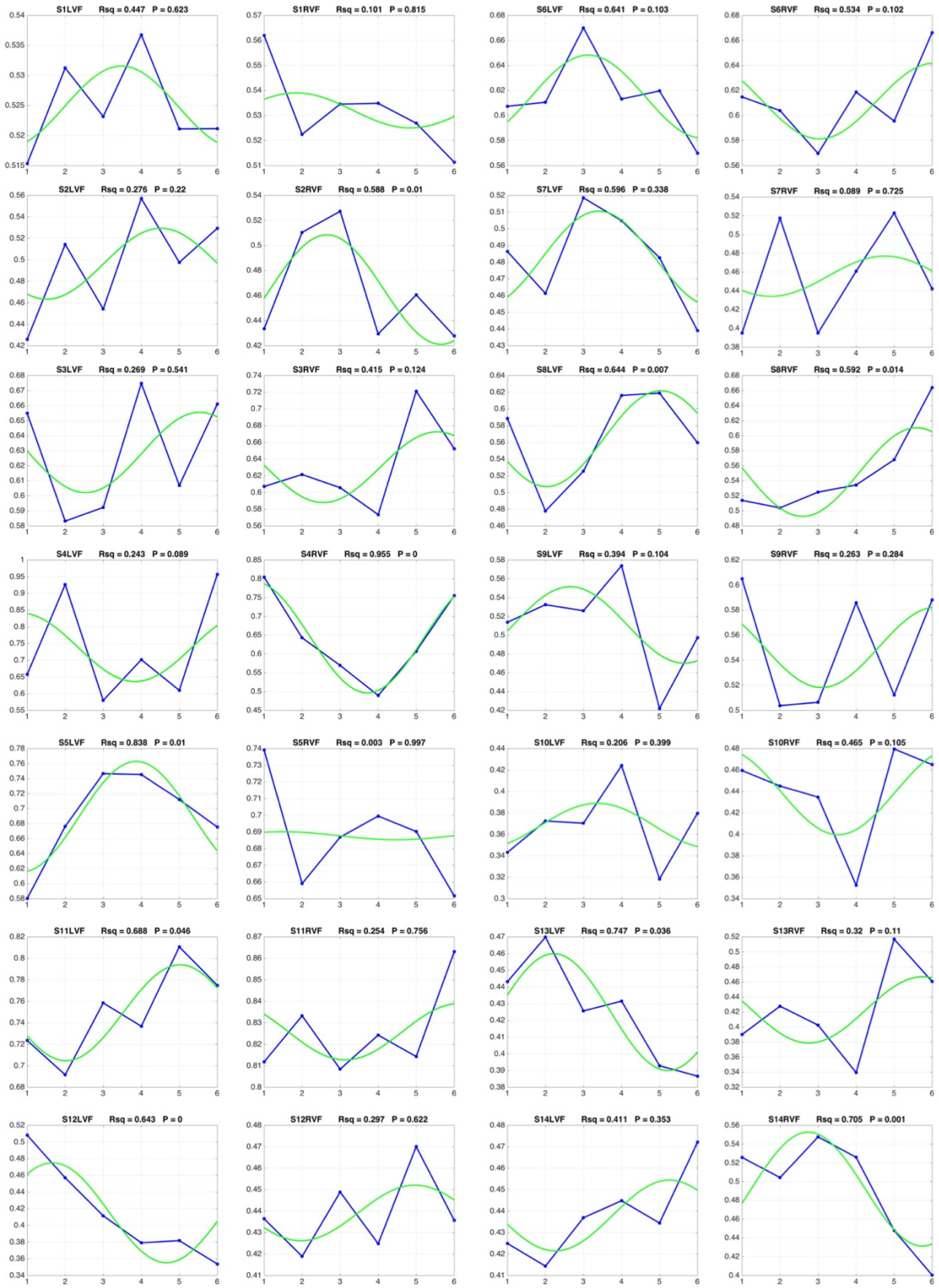
FFT approach		left hemifield		right hemifield	
		thresh./HR	RT	thresh./HR	RT
exp 1	group curve fitting	0.95	**0.02**	0.11	0.42
	2nd-level t-test	0.98	**0.0009**	0.12	**0.05**
exp 2	group curve fitting	0.44	0.14	0.20	0.28
	2nd-level t-test	0.60	0.16	0.67	0.48
exp 3	group curve fitting	0.72			
	2nd-level t-test	0.40			

Supplementary Table 5: P-values from permutation tests of FFT analysis approach instead of curve fitting and relevance values. Uncorrected. Z-scoring applied (without Z-scoring the significant outcomes become more significant). Bold values are <0.05, Italics <0.01

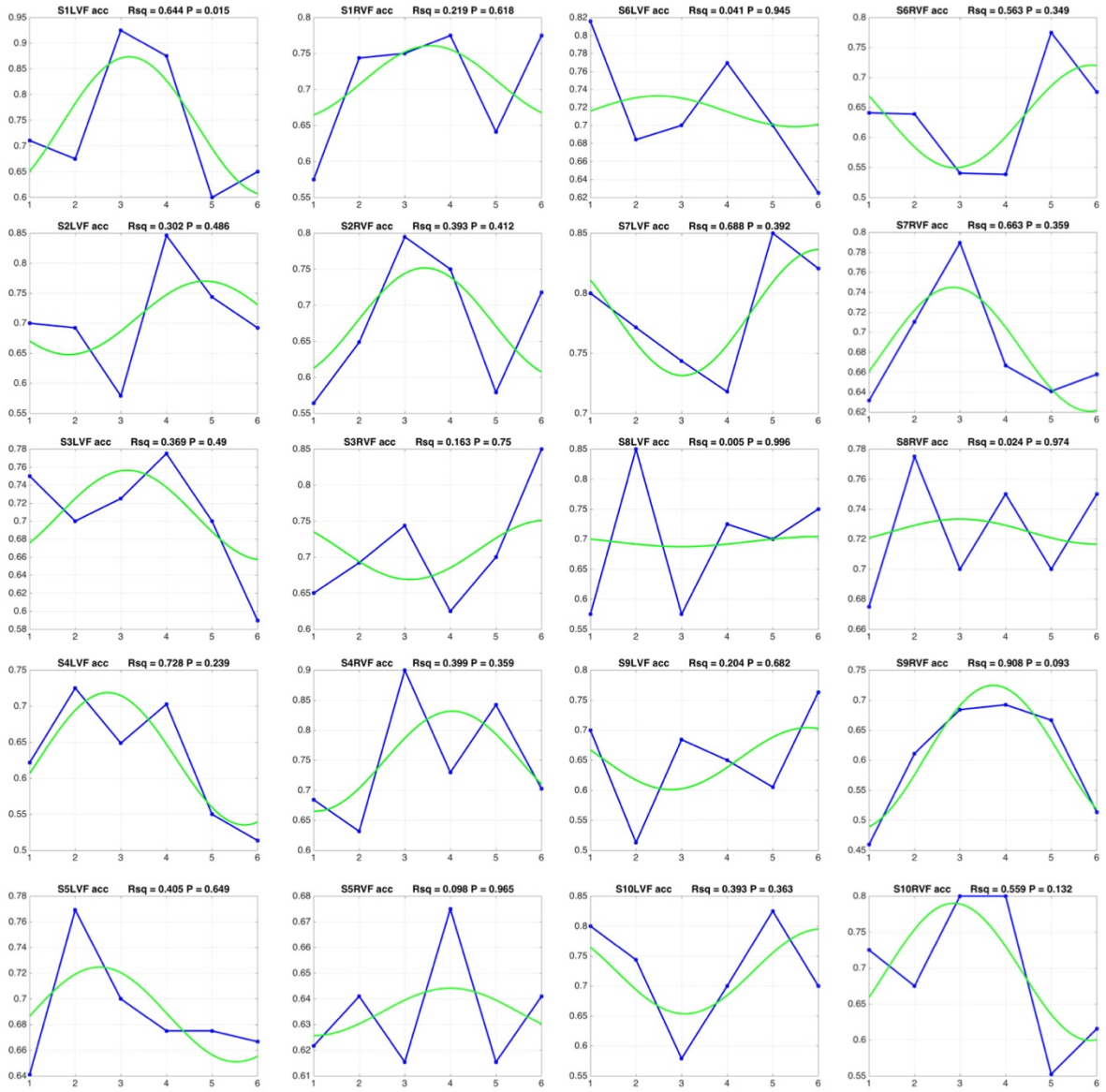
4. Individual patterns of task performance



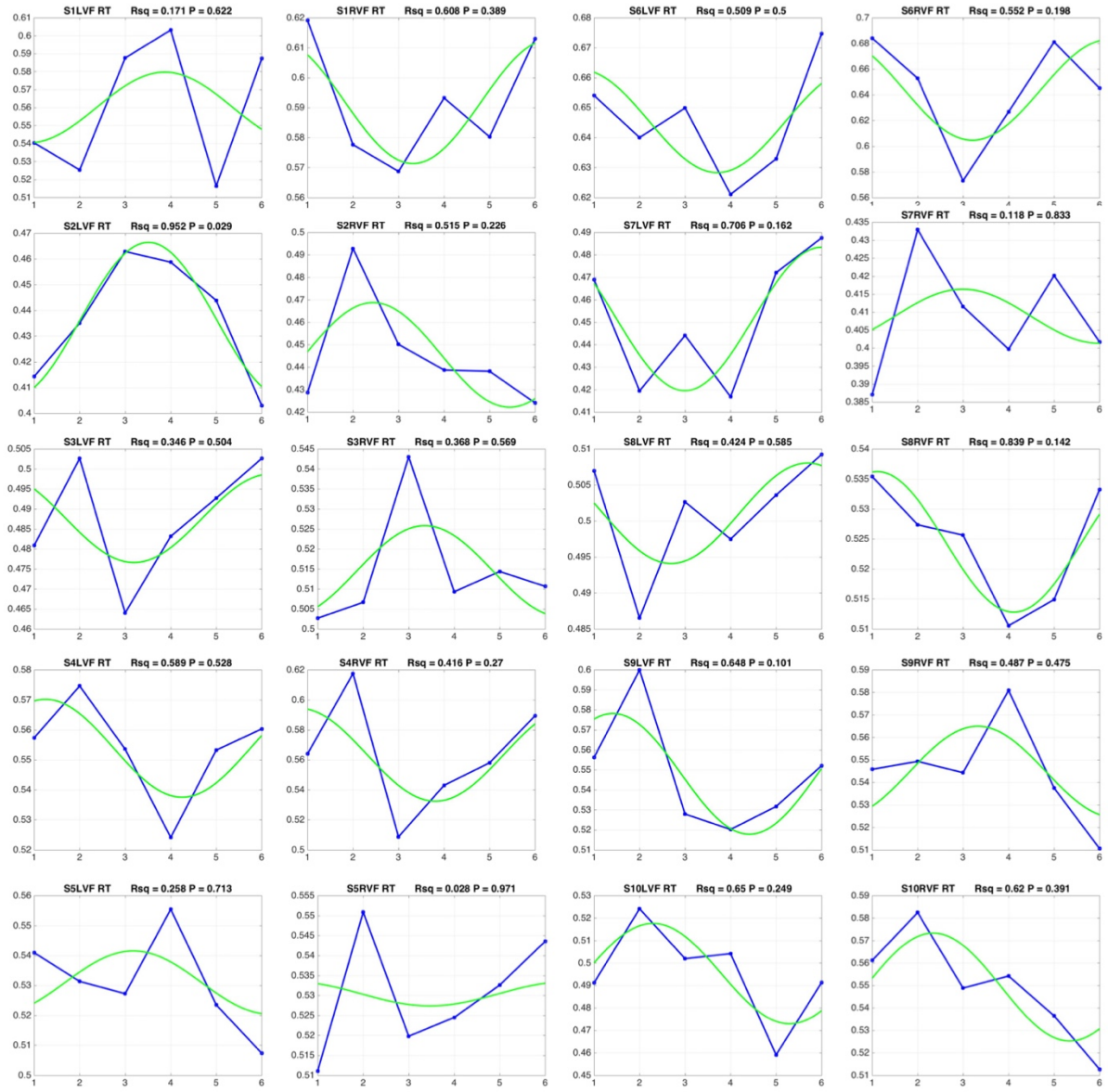
SF2: individual results experiment 1: thresholds



SF3: individual results experiment 1: reaction times



SF4: individual results experiment 2: accuracy



SF5: individual results experiment 2: reaction times

SF6: individual results experiment 3: detection rates