UvA-DARE (Digital Academic Repository)

Quantifying learning-dependent changes in the brain

Single trial multivoxel pattern analysis requires slow event related fMRI

Visser, R.M.; de Haan, M.I.C.; Beemsterboer, T.; Haver, P.; Kindt, M.; Scholte, H.S.

DOI

[10.1111/psyp.12665](https://doi.org/10.1111/psyp.12665)

Publication date

2016

Document Version

Final published version

Published in

Psychophysiology

License

Article 25fa Dutch Copyright Act

[Link to publication](#)

Citation for published version (APA):

Visser, R. M., de Haan, M. I. C., Beemsterboer, T., Haver, P., Kindt, M., & Scholte, H. S. (2016). Quantifying learning-dependent changes in the brain: Single trial multivoxel pattern analysis requires slow event related fMRI. *Psychophysiology*, 53(8), 1117-1127. <https://doi.org/10.1111/psyp.12665>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Quantifying learning-dependent changes in the brain: Single-trial multivoxel pattern analysis requires slow event-related fMRI

RENÉE M. VISSER,^{a,b,c} MICHELLE I. C. DE HAAN,^{b,d} TINKA BEEMSTERBOER,^e PIA HAVER,^e MEREL KINDT,^{a,b} AND H. STEVEN SCHOLTE^{b,e,f}

^aDepartment of Clinical Psychology, University of Amsterdam, Amsterdam, The Netherlands

^bAmsterdam Brain and Cognition (ABC), Amsterdam, The Netherlands

^cMedical Research Council—Cognition and Brain Sciences Unit, Cambridge, UK

^dDepartment of Psychiatry, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands

^eSpinoza Centre for Neuroimaging, University of Amsterdam, Amsterdam, The Netherlands

^fDepartment of Brain and Cognition, University of Amsterdam, Amsterdam, The Netherlands

Abstract

Single-trial analysis is particularly useful for assessing cognitive processes that are intrinsically dynamic, such as learning. Studying these processes with fMRI is problematic, as the low signal-to-noise ratio of fMRI requires the averaging over multiple trials, obscuring trial-by-trial changes in neural activation. The superior sensitivity of multivoxel pattern analysis over univariate analyses has opened up new possibilities for single-trial analysis, but this may require different fMRI designs. Here, we measured fMRI and pupil dilation responses during discriminant aversive conditioning, to assess associative learning in a trial-by-trial manner. The impact of design choices was examined by varying trial spacing and trial order in a series of five experiments (total $n = 66$), while keeping stimulus duration constant (4.5 s). Our outcome measure was the change in similarity between neural response patterns related to two consecutive presentations of the same stimulus (within-stimulus) and between patterns related to pairs of different stimuli (between-stimulus) that shared a specific outcome (electric stimulation vs. no consequence). This trial-by-trial similarity analysis revealed clear single-trial learning curves in conditions with intermediate (8.1–12.6 s) and long (16.5–18.4 s) intervals, with effects being strongest in designs with long intervals and counterbalanced stimulus presentation. No learning curves were observed in designs with shorter intervals (1.6–6.1 s), indicating that rapid event-related designs—at present, the most common designs in fMRI research—are not suited for single-trial pattern analysis. These findings emphasize the importance of deciding on the type of analysis prior to data collection.

Descriptors: Single-trial fMRI, Multivoxel pattern analysis, Representational similarity analysis, Associative learning, Aversive conditioning

Over the last two decades, much effort has been devoted to optimizing study designs for fMRI. The signal-to-noise ratio (SNR) of fMRI is low, given that events often do not evoke more than a 1% change in the blood-oxygen-level-dependent (BOLD) signal in an individual voxel (Huettel, Song, & McCarthy, 2004). A common method for improving SNR in event-related fMRI studies is to collect multiple observations for each experimental condition and to combine these as a single regressor in a general linear model (GLM). This method reduces noise and allows for a better estimation of the amplitude of the hemodynamic response. Rapid event-related designs (intervals of less than 10 s) generally produce the

strongest activation, as more trials can be presented in the same amount of time, increasing the total variance in the BOLD signal and thereby the experimental power (Dale & Buckner, 1997; Huettel & McCarthy, 2001).

However, many psychological constructs are intrinsically dynamic. The first time a picture is presented is not equivalent to the second time it is presented, as the picture may have become familiar. As a result, at least part of the brain will respond differently to consecutive trials. This change may even be of specific interest, as, for example, in classical (Pavlovian) conditioning or operant conditioning. In those learning paradigms, the information of interest is not available in the average responses and can only be obtained with single-trial analyses (Chadwick, Bonnici, & Maguire, 2012; Rey, Ahmadi, & Quiroga, 2015). Unfortunately, single-trial fMRI analyses are quite challenging due to the low SNR and the sluggishness of the hemodynamic response.

The advent of multivoxel pattern analysis (MVPA) opened up new avenues for single-trial analysis of BOLD-MRI data. Instead of average signal change, MVPA assesses distributed (multivoxel)

This study was funded by a Vici grant (MK, #453-07-006) from the Netherlands Organization for Scientific Research (NWO) and the research priority program on Brain and Cognition of the University of Amsterdam. The authors would like to thank Marcus Spaan for technical assistance

Address correspondence to: Renée M. Visser, Medical Research Council—Cognition and Brain Sciences Unit, 15 Chaucer Road, CB2 7EF Cambridge, UK. E-mail: renee.m.visser@gmail.com

patterns of BOLD signal to characterize the distinctive neural representation of a stimulus or condition. Over the last decade, numerous studies have underscored the superior sensitivity of MVPA compared to analysis of average activation for reading cognitive states from BOLD-MRI data (Haxby et al., 2001; Haynes & Rees, 2005; Kamitani & Tong, 2005; Li, Howard, Parrish, & Gottfried, 2008) and quantifying the relationships between patterns induced by different states or stimuli (Kriegeskorte, Mur, & Bandettini, 2008).

For many applications of MVPA, trials are modeled as single regressors in a GLM, instead of combined into one regressor per condition. The response patterns related to the different events are then used either for (binary) classification analysis, or (continuous) similarity analysis. In representational similarity analysis (RSA; Kriegeskorte et al., 2008), Pearson correlations are calculated between different response patterns, resulting in matrices that display the representational (dis)similarity between stimuli or trials. In previous work (Visser, Kunze, Westhoff, Scholte, & Kindt, 2015; Visser, Scholte, Beemsterboer, & Kindt, 2013; Visser, Scholte, & Kindt, 2011), we applied this technique to monitor trial-by-trial changes in neural response patterns as a function of aversive classical conditioning. In this paradigm, an initially neutral stimulus (conditioned stimulus [CS+]; e.g., a picture of a face) is repeatedly paired with an intrinsically aversive stimulus (unconditioned stimulus [UCS]; e.g., an electric shock), while another conditioned stimulus (CS-; e.g., a picture of another face) is never paired with the UCS. After sufficient CS+/UCS pairings, the CS+ acquires the same aversive qualities as the UCS and will elicit a conditioned defensive response on its own. Combining the conditioning paradigm with the trial-by-trial application of RSA enabled us to quantify changes in the representation of a stimulus as it acquires an aversive association (Visser et al., 2011, 2013, 2015), to assess the formation of a long-term aversive memory at the time of encoding (Visser et al., 2013, 2015), and to examine neuromodulatory factors associated with this memory formation (Visser et al., 2015).

Despite the potential of single-trial pattern analysis, relatively little is known about how to optimize study designs for this particular type of analysis. Rapid event-related designs, with many trial repetitions and jittered periods of prolonged stimulus intervals, are clearly the most efficient designs for estimating univariate signal changes. However, these designs presumably pose problems for the estimation of single-trial activation patterns, since overlapping BOLD signals cannot be decorrelated unless multiple trials are combined into a single regressor (Mumford, Davis, & Poldrack, 2014; Mumford, Turner, Ashby, & Poldrack, 2012).

The designs that we used for single-trial similarity analysis (Visser et al., 2011, 2013, 2015) differed in a number of ways from regular event-related fMRI designs. These design choices were based on theoretical assumptions, other studies (e.g., Formisano, De Martino, Bonte, & Goebel, 2008), and on pilot work. Crucial features included the fact that these designs were slow event-related and that the order of stimulus repetitions was not randomized, jittered, or optimized using standard algorithms (e.g., Kao, Mandal, Lazar, & Stufken, 2009), but designed in such a way that the time between consecutive presentations of a stimulus type was the same across stimulus types. We also utilized a partial reinforcement paradigm, in order to prevent shock-related confounds in the estimation of CS-related activation patterns.

While usually only the final protocol is published, the process of fine-tuning an experimental procedure can be very informative for researchers trying to replicate a published finding, especially when the attempted replication is conceptual in nature (i.e., not an

exact copy of the design), or when new analyses are performed on existing datasets. Since currently most fMRI datasets are derived from rapid event-related designs, a crucial question is whether single-trial pattern analysis is possible with these data sets. Based on our pilot data, we hypothesized that the key to successful single-trial analysis would depend on the length of the interstimulus intervals and the ordering of trials, and that standard rapid event-related designs are unsuited for this type of analysis. Here, we test this hypothesis by examining the effects of spacing and ordering of trials on the ability to quantify the dynamics of aversive learning with single-trial pattern analysis. In addition, we explore the effects of different ways of modeling single-trial responses in rapid event-related designs. Single-trial response patterns are usually obtained by modeling trials as separate regressors using a single GLM, that is, a least squares—all (LSA) approach. An alternative approach is to use a separate GLM for each trial, in which that specific trial is modeled as the regressor of interest and all other trials are combined into a single nuisance regressor per condition (least squares—single, LSS, Mumford et al., 2012). This technique has been shown to provide somewhat better parameter estimates in rapid event-related designs (Mumford et al., 2012), by reducing the correlation between regressors in the model (collinearity).

Method

Participants

Seventy-four participants were recruited by advertisements in the social media and the university website. For the analyses of BOLD-MRI data, participants were excluded because of sleep ($n = 1$), excessive head motion ($n = 3$), substantial signal drop-out ($n = 1$), or because they did not learn the associations ($n = 3$). In the remaining sample of 66 participants, three participants lacked eye-tracker data (see section on pupil dilation). Hence, BOLD-MRI data are reported for 66 participants (20 male, 4 left-handed, $M = 22.4$, $SD = 3.02$ years), and pupil data are reported for 63 participants. Participants earned a small amount of money or partial course credits for their participation. All participants gave their written informed consent before participation, declared no current or previous psychiatric illness or substance misuse, and had normal or corrected-to-normal vision. Procedures were executed in compliance with relevant laws and institutional guidelines, and were approved by the University of Amsterdam's ethics committee (2012-CP-2638).

Apparatus and Materials

Stimuli and conditioning procedure. The experiment consisted of one session of fMRI scanning during which we used a classical fear-conditioning paradigm, with delay conditioning and partial reinforcement (Figure 1a). Two pictures of neutral faces, derived from the Todorov database (Oosterhof & Todorov, 2008; generated with FaceGen Modeller 3.1), and two pictures of houses (Visser et al., 2013) were converted to grayscale and presented on a gray background. Each picture was presented 11 times for 4.5 s and served as a to-be conditioned stimulus (CS). One face and one house were followed by a mild electrical stimulus in five out of 11 presentations (CSs+). The electrical stimulation served as an unconditioned stimulus (UCS) and was delivered twice for 2 ms, with a delay of 300 ms (the second coterminating with the CS), by a Digitimer DS7A through MRI-compatible carbon electrodes attached to the right shinbone. The intensity of the electrical

Figure 1. a: Experimental design with five conditions, varying in trial spacing, order of presentation, and in use of jitter. Each of the four stimuli was repeatedly presented for 4.5 s, with two of the four stimuli (CSs+) coterminating with an electric stimulus on 45% of the trials. b: In Condition I, II, and III, the paradigm consisted of repeating sequences of target trials (black numbers), presenting the four stimuli in a fixed order such that the time between two consecutive target trials was equal over the four stimulus types. In between, the semirandom presentation of filler trials ensured the unpredictability of stimuli (gray numbers). Administration of a UCS was restricted to filler trials, and the first paired CS+ trials were presented between the second and third sequence of target trials. In Condition IV and V, the order of stimulus presentation was semirandom. c: In order to assess neural pattern similarity, correlations were calculated between patterns evoked by consecutive trials of the same stimulus (within-stimulus) and trials of stimuli that share (non)reinforcement. Average activation and within-stimulus correlations were averaged over face and house stimuli, to reduce the number of comparisons and because we were not interested in the difference between faces and houses with regard to the experimental manipulation. Images are not to scale.

stimulus was individually adapted at a level that the participant experienced as aversive but not painful (intensity range 4–60 mA, $M = 30.6$, $SD = 15.6$). The other two stimuli were never reinforced

(CSs-). Visual stimuli were backward-projected onto a screen that was viewed through a mirror attached to the head coil. Participants were told that two out of four stimuli could be followed by electrical stimulation while the other two would never be reinforced, and were instructed to learn the specific contingencies. We used these explicit instructions to minimize uncertainty and individual differences in learning rate, and thereby maximize differential learning effects. After scanning, knowledge of the contingencies was assessed using a paper-and-pencil questionnaire.

Pupil dilation. Pupil dilation responses were used as an independent measure of anticipatory autonomic arousal. Pupil size was recorded continuously throughout MRI scanning, using a remote nonferromagnetic infrared EyeLink 1000 Long Range Mount eye tracker (SR Research). Before task onset, a nine-point calibration procedure was performed. Participants were instructed not to move their eyes, but instead fixate on the center of the screen for as long as a stimulus was presented. Prior to stimulus onset, a white fixation cross turned pink for 1 s to enable the participant to focus in time.

Data were sampled at 500 Hz, and data preprocessing and analysis were identical to previous work (Visser et al., 2015). Samples around series of missing samples were regarded as unreliable and were removed (100 ms before and 100 ms after each series of 10 missing samples) and replaced by a linear trend at point, using the entire time series. Next, the interpolated pupil time series were low-pass filtered (third-order Butterworth, 4 Hz). The baseline pupil diameter was the average value during the 500 ms prior to each CS onset. The pupil response to the CS was calculated as the peak change from baseline in a window from 0 to 4 s after picture onset. Trials that suffered substantial signal loss affecting more than 50% of either the baseline samples or the 4 s after stimulus onset were eliminated and replaced entirely by estimating the linear trend at point over trials for each condition separately. Participants that ended up missing 25% or more of the trials (11 out of 44 trials) were excluded ($n = 3$), leaving 63 participants for the analysis of pupil responses (0–22.7% replaced trials per participant, median = 0%). Next, data were Z-transformed, to reduce between-subjects variability.

Image acquisition. Scanning was performed on a 3T Philips Achieva TX MRI scanner using a 32-channel head coil. Functional data were acquired using a gradient-echo, echo-planar pulse sequence (TR = 2,000 ms; TE = 27.63 ms; FA = 76.1°; 39 sagittal slices with interleaved acquisition; $3 \times 3 \times 3.3$ mm voxel size; 64×64 matrix; $192 \times 192 \times 141.24$ FoV) covering the whole brain. In Condition I and II, 500 volumes were recorded; in Condition III and IV, 304 volumes were recorded; in Condition V, 168 volumes were recorded. Foam pads minimized head motion. A high-resolution 3D T1-weighted image (TR = 8.11 ms, TE = 3.72 ms, FA = 8°; $1 \times 1 \times 1$ mm voxel size; $240 \times 220 \times 188$ FoV) was additionally collected for anatomical visualization.

Preprocessing. fMRI data processing was carried out using FEAT (fMRI Expert Analysis Tool) version 6.00, part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). Preprocessing included motion correction using MCFLIRT (Jenkinson, Bannister, Brady, & Smith, 2002); slice-timing correction using Fourier-space time-series phase-shifting; nonbrain removal using BET (Smith, 2002); high-pass temporal filtering ($\sigma = 50$ s), and voxelwise prewhitening (Woolrich, Ripley, Brady, & Smith, 2001). No spatial smoothing was applied. Functional images were coregistered to each individual's high resolution structural images using FLIRT (Jenkinson et al., 2002; Jenkinson & Smith, 2001) and 6 degrees of

freedom (rigid-body transformation). Registration from high resolution structural to standard space (MNI152 template, 2 mm) was then carried out using FNIIRT nonlinear registration (Andersson, Jenkinson, Smith, 2007). Initial voxelwise whole-brain analyses

(first-level) were performed in native space; trial-by-trial similarity analyses (higher-level) were performed in standard space.

Regions of interest selection. Regions of interest (ROI) were selected based on their role in fear learning and (extinction) memory and included the anterior cingulate cortex (ACC, 9,213 voxels), the insula (6,591 voxels), amygdala (2,967 voxels), hippocampus (5,837 voxels), and ventromedial prefrontal cortex (vmPFC, 4,160 voxels). We additionally included the superior frontal gyrus (SFG, 18,946 voxels), a region outside the salience network, for its large learning effects as revealed by previous similarity analysis, in the absence of differences in average activation (Visser et al., 2011, 2013, 2015). ROIs were obtained from the Harvard-Oxford cortical and subcortical structural atlases (Harvard Center for Morphometric Analysis). No probability thresholding was applied.

Trial spacing and order of presentation. In a series of five independent experiments (hereafter referred to as Condition I–V), we examined the effects of trial spacing and trial order on neural pattern similarity (Figure 1a). The stimulus duration of 4.5 s was constant over conditions, and is not included in what we refer to as interstimulus intervals (in our case, the time between stimulus offset and the onset of the next stimulus). In Condition I, interstimulus intervals were fixed (17.5 s), and the onset of each trial was triggered by the start of the acquisition of a BOLD-MRI volume. In Condition II, interstimulus intervals varied randomly between 16.5 and 18.5 s. In Condition III, IV, and V, interstimulus intervals were drawn from a truncated exponential distribution. In Condition III and IV, the mean interval was 9.2 s (range 8.1–12.5 s); in Condition V, the mean interval was 3.0 s (range 1.6–6.1 s).

A crucial aspect of the design was the distinction between *target* and *filler* trials (Figure 1b; Visser et al., 2011, 2013, 2015). Both target and filler trials entailed the presentation of one of the four stimuli (two faces, two houses), but only target trials were analyzed. The distinction between filler and target trials was important for two reasons. First, it enabled us to control for the effect of temporal proximity on neural pattern similarity: In Condition I, II, and III, the order of stimulus presentation was fixed (counterbalanced across participants), meaning that the time between two consecutive target trials (e.g., time between third and fourth presentation of CS+ house) was equal over the four stimulus types (i.e., was the same as the time between third and fourth presentation of CS+ face, CS- face, or CS- house) canceling out any effects of temporal proximity on correlation values. The presentation of filler trials (same stimuli, but discarded from the analyses) in between sequences of target trials ensured that the stimulus presentation remained unpredictable for the participant. In Condition IV and V, the order of stimulus presentation was semirandom (to examine what happens if we do not control for time), with the restriction that stimuli were roughly equally distributed across the experiment and that (as

Figure 2. Neural pattern similarity in the superior frontal gyrus, showing all trials (filler and target trials) in order of presentation (left). In Condition III, IV, and V, which have shorter intertrial intervals, high correlations are observed between adjacent trials. Still, with intermediate intervals and controlled stimulus presentation (Condition I, II, and III), higher pattern similarity is observed within and between CS+ stimuli compared to CS- stimuli (right, target trials only). This differential pattern similarity is weaker when the trial presentation is random (Condition IV), and absent when the trial presentation is both random and intervals are very short (Condition V). ISI = interstimulus interval (excludes the stimulus duration of 4.5 s).

Table 1. Summary of Statistics of the fMRI Data in Condition I, Repeated Measures ANOVA, in Six Anatomical ROIs

Region	Within-stimulus				Between-stimulus				Mean activation			
	Main effect of stimulus (2)		Interaction of Stimulus (2) × Trial (5)		Main effect of stimulus (2)		Interaction of Stimulus (2) × Trial (6)		Main effect of stimulus (2)		Interaction of Stimulus (2) × Trial (6)	
	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2
ACC	15.00 .003	.58	5.39 .001	0.33	10.80 .007	.50	1.23 .313	0.10	6.72 <i>.025</i>	.38	1.74 .142	.14
Amygdala	0.98 .344	.08	NT NT	NT	0.56 .469	.05	NT NT	NT	0.72 .416	.06	NT NT	NT
Hippocampus	1.91 .194	.15	NT NT	NT	2.32 .156	.17	NT NT	NT	3.20 .101	.23	NT NT	NT
Insula	12.99 .004	.54	0.63 .641	0.05	6.65 <i>.026</i>	.38	2.22 .065	.17	4.90 <i>.049</i>	.31	3.90 .004	.26
SFG	23.84 < .0005	.68	4.07 .007	0.27	16.78 .002	.60	1.69 .152	.13	0.04 .856	.00	NT NT	NT
vmPFC	1.36 .268	.11	NT NT	NT	3.62 .084	.25	NT NT	NT	2.33 .155	.18	NT NT	NT

$N = 12$.

Note. All significant values ($p < .05$) are in italics, and those that reach FDR-corrected significance are in bold. Areas without significant main effect of stimulus type are not tested for interaction effects. Main effects are calculated over all acquisition trials, 5 for within-stimulus pattern similarity, 6 for between-stimulus pattern similarity, 6 for average activation. FDR = false discovery rate; NT = not tested; ACC = anterior cingulate cortex; SFG = superior frontal gyrus; vmPFC = ventromedial prefrontal cortex.

in the other three conditions) the experiment started with two unreinforced presentations of each stimulus, to estimate a preconditioning baseline response to the pictures.

A second reason for the distinction between filler and target trials was so that we could ensure that UCS-related activity would not confound CS-related activity (all five conditions): Administration of a UCS was restricted to filler trials (Figure 1b), which were discarded from the analyses. In total, each condition consisted of 44 trials: 24 target trials (six per stimulus type) and 20 filler trials (five per stimulus type), including all CS+ trials that coterminated with electrical stimulation (*reinforced* trials). The first reinforced CS+ trials were presented between the second and third sequence of target trials (Figure 1b). As the relatively high temporal resolution of the pupil dilation response allowed for the analysis of reinforced trials (i.e., the response to the UCS could be easily distinguished from responses to CSs), the analysis of pupil data included both filler and target trials (Visser et al., 2015).

Trial-by-Trial Similarity Analysis

For the trial-by-trial representational similarity analysis, we modeled each trial as a separate regressor in a voxelwise whole-brain analysis using a single GLM (LSA approach) and including six motion parameters as regressors of no interest. For Conditions III–V, we additionally used a LSS approach (Mumford et al., 2012, 2014), where we ran a separate GLM for each trial, modeling that trial as the regressor of interest and combining all other trials into one nuisance regressor per condition (target and filler trials modeled in separate conditions). In both approaches, the resulting single-trial parameter estimates were transformed into t values to down-weight noisy voxels (Misaki, Kim, Bandettini, & Kriegeskorte, 2010). To this end, each voxel's parameter estimate was divided by the standard error of that voxel's residual error term after fitting the first-level GLM. In MATLAB (version 8.0; MathWorks), we created for each participant, for each ROI, a vector containing t values per voxel, for a particular trial (i.e., the spatial representation of that trial). Next we calculated pairwise Pearson

correlations (i.e., representational similarity) between all vectors of all single trials, resulting in a similarity matrix containing correlations among trials, for each participant, for each ROI (Figure 2, left panels). The strength of these correlations was used as a metric of similarity. From this matrix, two different types of correlations were selected (Figure 1c, Figure 2, right panels), discarding filler trials. First, we examined within-stimulus correlations on consecutive target trials of the same stimulus (e.g., similarity between Trial 4 and Trial 5 of CS+ face). Second, we examined between-stimulus correlations between target trials that shared (non)reinforcement (learned associations: Trial 4 CS+ face with Trial 4 CS+ house, and Trial 4 CS- face with Trial 4 CS- house). Note that the number of between-stimulus correlations is equal to the number of target trials, whereas the number of within-stimulus correlations is equal to the number of target trials minus one. Next, data were Z-transformed, to reduce between-subjects variability.

Trial-by-Trial Univariate Analysis

To visualize trial-by-trial changes in average activation, we analyzed data as described in the previous section, except that when we analyzed the normalized single-trial parameter estimates, we averaged across voxels in an ROI. Thus, instead of preserving the spatial information by creating a vector of voxels per ROI, we obtained one value per ROI (average response amplitude), which we then Z-transformed across trials to reduce between-subject variability.

Statistical Analyses

Z-transformed pupil dilation responses, average activation, and within-stimulus correlations were averaged over face and house stimuli. This was done to reduce the number of comparisons and because we were not interested in the difference between faces and houses with regard to the experimental manipulation.

Statistical comparisons of the learned associations were performed by within-subject analysis of variance (ANOVA), using Statistical Package for the Social Sciences (SPSS, version 21; SPSS Inc.). Statistical tests were equivalent for pupil dilation and

Table 2. Summary of Statistics of the fMRI Data in Condition II, Repeated Measures ANOVA, in Six Anatomical ROIs

Region	Within-stimulus				Between-stimulus				Mean activation			
	Main effect of stimulus (2)		Interaction of Stimulus (2) × Trial (5)		Main effect of stimulus (2)		Interaction of Stimulus (2) × Trial (6)		Main effect of stimulus (2)		Interaction of Stimulus (2) × Trial (6)	
	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2
	<i>p</i> value		<i>p</i> value		<i>p</i> value		<i>p</i> value		<i>p</i> value		<i>p</i> value	
ACC	16.50 .002	.62	2.17 0.090	.18	15.62 .003	.61	2.44 .094	.20	8.22 .017	.45	1.06 .393	.10
Amygdala	3.28 .100	.25	NT NT	NT	0.96 .351	.09	NT NT	NT	1.91 .197	.16	NT NT	NT
Hippocampus	9.21 .013	.48	2.26 .904	.02	6.30 .031	.39	0.61 .691	.06	1.32 .277	.12	NT NT	NT
Insula	7.07 .024	.41	1.63 .186	.14	16.45 .002	.62	1.85 0.121	.16	17.88 .002	.64	4.02 .004	.29
SFG	11.79 .006	.54	5.25 .002	.34	14.10 .004	.59	1.14 .344	.10	1.10 .320	.10	NT NT	NT
vmPFC	1.29 .283	.11	NT NT	NT	1.01 .339	.09	NT NT	NT	0.64 .443	.06	NT NT	NT

N = 11.

Note. All significant values ($p < .05$) are in italics, and those that reach FDR-corrected significance are in bold. Areas without significant main effect of stimulus type are not tested for interaction effects. Main effects are calculated over all acquisition trials, 5 for within-stimulus pattern similarity, 6 for between-stimulus pattern similarity, 6 for average activation. FDR = false discovery rate; NT = not tested; ACC = anterior cingulate cortex; SFG = superior frontal gyrus; vmPFC = ventromedial prefrontal cortex.

neural measures, the only difference being the number of trials that was included in the analysis. The term *trial* is somewhat ambiguous in this context as it might refer to the actual presentation of a stimulus (i.e., 11 trials per stimulus type for pupil dilation responses, including both target and filler trials, or six trials for average activation, only including target trials). Alternatively, it might refer to the correlation between a pair of stimuli. For within-stimulus pattern similarity, this is the correlation between target trial *n* and target trial *n* + 1 of the same stimulus type (i.e., five trials per stimulus type). For between-stimulus pattern similarity, this is the correlation between trial *n* from one stimulus type and trial *n* from another stimulus type (i.e., six trials for CS+ stimuli and six trials for CS- stimuli).

Differential aversive learning (a learning curve) was assessed by the interaction of Trial × Stimulus Type (2 levels [CS+ and CS-, averaged over faces and houses]), but was only tested when there was also a significant main effect of stimulus type. Likewise, the reported average effect sizes (see section on neural pattern similarity) represent the average over the tested main effects, and if significant also over the tested interaction effects (see also Table 1–5). Note that the sample sizes are too small to directly compare the learning effects (which include numerous parameters, i.e., different trials and stimuli) between groups. The aim of presenting a series of independent experiments was to provide researchers with effect size estimations that may guide choices for future study designs.

Table 3. Summary of Statistics of the fMRI Data in Condition III, Repeated Measures ANOVA, in Six Anatomical ROIs

Region	Within-stimulus				Between-stimulus				Mean activation			
	Main effect of stimulus (2)		Interaction of Stimulus (2) × Trial (5)		Main effect of stimulus (2)		Interaction of Stimulus (2) × Trial (6)		Main effect of stimulus (2)		Interaction of Stimulus (2) × Trial (6)	
	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2
	<i>p</i> value		<i>p</i> value		<i>p</i> value		<i>p</i> value		<i>p</i> value		<i>p</i> value	
ACC	5.32 .040	.31	2.79 .036	.19	12.81 .004	.52	1.31 .273	.10	0.90 .361	.07	NT NT	NT
Amygdala	1.79 .205	.13	NT NT	NT	2.40 .147	.17	NT NT	NT	0.15 .706	.01	NT NT	NT
Hippocampus	2.65 .129	.18	NT NT	NT	0.00 .949	.00	NT NT	NT	0.43 .522	.03	NT NT	NT
Insula	15.62 .002	.57	2.42 .061	.17	10.11 .008	.46	0.41 .843	.03	5.96 .031	.33	1.98 .095	.14
SFG	2.62 .132	.18	NT NT	NT	12.00 .005	.50	1.61 .172	.12	0.01 .921	.00	NT NT	NT
vmPFC	9.52 .009	.44	1.36 .263	.10	8.68 .012	.42		.15	31.26 < .0005a	.72	2.61 .031	.18

N = 13.

Note. All significant values ($p < .05$) are in italics, and those that reach FDR-corrected significance are in bold. Areas without significant main effect of stimulus type are not tested for interaction effects. Main effects are calculated over all acquisition trials, 5 for within-stimulus pattern similarity, 6 for between-stimulus pattern similarity and 6 for average activation. FDR = false discovery rate; NT = not tested; ACC = anterior cingulate cortex; SFG = superior frontal gyrus; vmPFC = ventromedial prefrontal cortex.

^aEffect caused by significantly higher values for CS- stimuli.

Table 4. Summary of Statistics of the fMRI Data in Condition IV, Repeated Measures ANOVA, in Six Anatomical ROIs

Region	Within-stimulus				Between-stimulus				Mean activation			
	Main effect of stimulus (2)		Interaction of Stimulus (2) × Trial (5)		Main effect of stimulus (2)		Interaction of Stimulus (2) × Trial (6)		Main effect of stimulus (2)		Interaction of Stimulus (2) × Trial (6)	
	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2
	<i>p</i> value		<i>p</i> value		<i>p</i> value		<i>p</i> value		<i>p</i> value		<i>p</i> value	
ACC	19.05 .001	.59	1.74 .155	.12	2.09 .172	.14	NT NT	NT	1.05 .324	.08	NT NT	NT
Amygdala	10.75 .006	.45	0.58 .681	.04	2.90 .112	.18	NT NT	NT	0.66 .432	.05	NT NT	NT
Hippocampus	6.09 .028	.32	0.51 .728	.04	1.42 .255	.10	NT NT	NT	0.00 .994	.00	NT NT	NT
Insula	6.83 .021	.34	2.53 .051	.16	0.72 .413	.05	NT NT	NT	11.06 .005	.46	2.19 .066	.14
SFG	17.76 .001	.58	3.00 .027	.19	1.18 .297	.08	NT NT	NT	0.91 .358	.07	NT NT	NT
vmPFC	2.92 .111	.18	NT NT	NT	1.70 .215	.12	NT NT	NT	2.24 .158	.15	NT NT	NT

N = 14.

Note. All significant values ($p < .05$) are in italics, and those that reach FDR-corrected significance are in bold. Areas without significant main effect of stimulus type are not tested for interaction effects. Main effects are calculated over all acquisition trials, 5 for within-stimulus pattern similarity, 6 for between-stimulus pattern similarity and 6 for average activation. FDR = false discovery rate; NT = not tested; ACC = anterior cingulate cortex; SFG = superior frontal gyrus; vmPFC = ventromedial prefrontal cortex.

If the assumption of sphericity was violated, a Greenhouse-Geisser correction was applied, and corrected p values, together with the uncorrected degrees of freedom, are reported (Jennings, 1987). All p values are reported two-sided, with the significance level set at $\alpha = .05$.

Results

Conditioned Pupil Dilation Responses

Pupil dilation responses were assessed as an independent measure of anticipatory arousal, to verify that aversive conditioning was

successful in each of the five conditions. Consistent with previous work (Visser et al., 2013, 2015), successful learning was evident from a trial-by-trial change in pupil dilation in response to the CS+, relative to the CS-, in all conditions (Figure 3a). Repeated measures ANOVA revealed significant interaction effects of trial (11) and stimulus type (2) in Condition II ($n = 11$) $F(10,100) = 4.59$, $p = .004$, $\eta_p^2 = .31$, Condition III ($n = 12$) $F(10,110) = 2.61$, $p = .007$, $\eta_p^2 = .19$, and Condition IV ($n = 14$) $F(10,130) = 2.77$, $p = .004$, $\eta_p^2 = .18$, but not in Condition I ($n = 11$) $F(10,100) = 1.45$, $p = .232$, $\eta_p^2 = .13$ and Condition V ($n = 15$) $F(10,140) = 0.91$, $p = .489$, $\eta_p^2 = .06$. Strong main effects of stimulus type were observed in Condition I, $F(1,10) = 15.04$,

Table 5. Summary of Statistics of the fMRI Data in Condition V, Repeated Measures ANOVA, in Six Anatomical ROIs

Region	Within-stimulus				Between-stimulus				Mean activation			
	Main effect of stimulus (2)		Interaction of Stimulus (2) × Trial (5)		Main effect of stimulus (2)		Interaction of Stimulus (2) × Trial (6)		Main effect of stimulus (2)		Interaction of Stimulus (2) × Trial (6)	
	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2
	<i>p</i> value		<i>p</i> value		<i>p</i> value		<i>p</i> value		<i>p</i> value		<i>p</i> value	
ACC	1.11 .310	.07	NT NT	NT	0.04 .849	.00	NT NT	NT	1.04 .323	.07	NT NT	NT
Amygdala	3.19 .094	.18	NT NT	NT	0.08 .781	.01	NT NT	NT	0.13 .722	.01	NT NT	NT
Hippocampus	5.35 .035a	.26	1.01 .389	.07	0.00 .983	.00	NT NT	NT	1.26 .280	.08	NT NT	NT
Insula	0.90 .357	.06	NT NT	NT	0.20 .662	.01	NT NT	NT	0.12 .731	.01	NT NT	NT
SFG	0.53 .478	.03	NT NT	NT	0.024 .878	.00	NT NT	NT	1.43 .250	.09	NT NT	NT
vmPFC	0.09 .766	.01	NT NT	NT	0.36 .559	.02	NT NT	NT	19.92 < .0005a	.57	1.35 .271	.08

N = 16.

Note. All significant values ($p < .05$) are in italics, and those that reach FDR-corrected significance are in bold. Areas without significant main effect of stimulus type are not tested for interaction effects. Main effects are calculated over all acquisition trials, 5 for within-stimulus pattern similarity, 6 for between-stimulus pattern similarity and 6 for average activation. FDR = false discovery rate; NT = not tested; ACC = anterior cingulate cortex; SFG = superior frontal gyrus; vmPFC = ventromedial prefrontal cortex.

^aEffect caused by significantly higher values for CS⁺ stimuli.

Figure 3. a: Pupil dilation responses, and (b) neural pattern similarity in the superior frontal gyrus—both within-stimulus—show clear learning over the course of conditioning in conditions with longer intervals, with the exception of between-stimulus pattern similarity in Condition IV and within- and between-stimulus pattern similarity in Condition V. c: For average activation in the superior frontal gyrus, no learning-dependent changes are observed. ISI = interstimulus interval (excludes the stimulus duration of 4.5 s). Error bars represent *SEM*. * $p < .05$ for main effect of stimulus type.

Figure 4. Neural pattern similarity in the superior frontal gyrus obtained using a least squares–single (LSS) approach, showing all trials (filler and target trials) in order of presentation (a, left) and sorted per condition (a, right). Compared to a least squares–all approach (LSA; Figure 2 and 3b), the effect of overlapping BOLD responses for adjacent trials is reduced in the conditions with intertrial intervals of medium length (Condition III and IV) and short intervals (Condition V). Still, only with intermediate intervals is higher pattern similarity observed within and between CS+ stimuli compared to CS- stimuli (a), showing typical trial-by-trial learning curves over the course of conditioning (b). ISI = interstimulus interval (excludes the stimulus duration of 4.5 s). Error bars represent *SEM*. * $p < .05$ for main effect of stimulus type.

$p = .003$, $\eta_p^2 = .60$; in Condition II, $F(1,10) = 28.06$, $p < .0005$, $\eta_p^2 = .74$; Condition III, $F(1,11) = 13.84$, $p = .003$, $\eta_p^2 = .56$; Condition IV, $F(1,13) = 85.15$, $p < .0005$, $\eta_p^2 = .87$; and Condition V, $F(1,14) = 7.74$, $p = .015$, $\eta_p^2 = .36$, indicating that pupil size increased when a CS+ was presented. These results confirm that the conditioning procedure was successful in all five conditions.

Neural Pattern Similarity

The left panels in Figure 2 present similarity matrices in the SFG, showing all trials (filler and target trials) in order of presentation,

for the different conditions ($n = 12$ in Condition I, $n = 11$ in Condition II; $n = 13$ in Condition III, $n = 14$ in Condition IV, and $n = 16$ in Condition V). With shorter intervals, correlations between adjacent trials became higher, illustrating the effects of collinearity and temporal autocorrelations on pattern similarity. Still, with intermediate intervals and controlled stimulus presentation (Condition I, II, and III), higher pattern similarity was observed within and between CS+ stimuli compared to CS- stimuli (Figure 2, right panels). This differential pattern similarity was weaker when the trial presentation was random (Condition IV) and absent when the trial presentation was both random and intervals were very short (Condition V).

From these similarity matrices, we selected the correlations between consecutive trials of the same stimulus (within-stimulus pattern similarity) and between stimuli that shared the same outcome (between-stimulus pattern similarity) to systematically quantify changes in trial-by-trial pattern similarity as a function of learning. A complete overview of the statistical tests for each of the six ROIs, per condition, can be found in Table (1–5); an example of trial-by-trial data is displayed in Figure 3b (SFG).

In short, within-stimulus pattern similarity (Figure 3b) revealed successful learning as evidenced by an increase in similarity for CS+ stimuli, relative to CS- stimuli, in designs with intermediate and long intervals (Table (1–4)). This is in line with our previous work (Visser et al., 2013, 2015). For between-stimulus pattern similarity, an increase in similarity for CS+ stimuli, relative to CS- stimuli, was observed when order of trial presentation was controlled (Condition I, II, and III), but not when presentation was semirandom (Condition IV). In general, main and interaction effects were stronger in ACC (mean $\eta_p^2 = .34$), insula (mean $\eta_p^2 = .28$), and SFG (mean $\eta_p^2 = .35$) than in amygdala (mean $\eta_p^2 = .16$), hippocampus (mean $\eta_p^2 = .17$), and vmPFC (mean $\eta_p^2 = .20$). Furthermore, effects were stronger in Condition I (mean $\eta_p^2 = .29$) and II (mean $\eta_p^2 = .30$) than in Condition III (mean $\eta_p^2 = .25$) and IV (mean $\eta_p^2 = .19$), indicating that long intervals had a benefit over intermediate intervals. In Condition V, the rapid event-related design, we did not observe any effects of conditioning on pattern similarity (Figure 3, Table 5).

To explore if suboptimal modeling could possibly explain the weaker effects in the conditions with intermediate intervals (Condition III and IV), and lack of effects in the condition with short intervals (Condition V), we reanalyzed these data using the LSS approach. Figure 4 shows that, although there seemed to be a general reduction in collinearity consistent with previous findings (Mumford et al., 2012), the rapid event-related design (Condition V) still did not show any learning-dependent changes in pattern similarity. Of note, the LSS approach is mainly recommended for classification analysis. Although classification analysis often relies on single-trial estimations as well (though not necessarily), the classifier is trained on multiple exemplars, thus this type of analysis is not really single trial.

Trial-by-Trial Univariate Activation

Results obtained with single-trial univariate analysis dissociated from results obtained with similarity analysis in some, but not all of the regions. Typical learning curves (i.e., a differential increase of CS+ stimuli compared to CS- stimuli over the course of conditioning) were observed in areas in the “salience network” (ACC and insula), but were absent in hippocampus, amygdala, vmPFC, and SFG (Figure 3c). For an overview of the statistics per ROI, see Table 1–5. The fact that effect sizes were substantially smaller for average activation (mean $\eta_p^2 = .18$) than for pattern similarity (see preceding section) is consistent with previous results (Visser et al., 2011, 2013, 2015). This again shows the high sensitivity of pattern analysis compared to analysis of average activation for quantifying the formation of aversive associations over time.

Discussion

Consistent with previous work (Visser et al., 2011, 2013, 2015), we show that the application of trial-by-trial similarity analysis produces clear learning curves that index the formation of aversive associations, even in the absence of differences in mean activation.

Importantly, and in line with our hypothesis, we only obtained these effects in designs with intermediate (8.1–12.6 s) to long (> 16 s) intervals, not in a rapid event-related design (1.6–6.1 s). The lack of effects in the rapid event-related condition cannot be explained by a failure to induce aversive learning, since an independent measure of anticipatory arousal (i.e., pupil dilation responses) confirmed that the aversive conditioning procedure was effective in all conditions.

A precise understanding of why single-trial responses are difficult to estimate with short intervals is beyond the scope of this article. One can think of a number of explanations for the problems that arise with short intervals, or combination thereof, including local stationary noise (causing temporal autocorrelations), collinearity in the model, or a nonlinear summation of BOLD responses for trials that are close in time. Although an alternative modeling approach (LSS) to estimate single-trial response patterns seemed to reduce the effects of collinearity and autocorrelations on similarity values, in line with previous findings (Mumford et al., 2012), the rapid event-related design still did not reveal any effects of learning.

Whereas rapid event-related designs are at present standard in fMRI research, Pavlovian conditioning paradigms typically use intermediate intervals of around 10 s (Fullana et al., 2015) for other reasons than single-trial estimation of BOLD responses. For example, the sensation of the electrical stimulus always requires a few seconds to fade. This poses problems if many trials were to be presented in a short amount of time, causing unwanted effects such as backward conditioning (Moscovitch & LoLordo, 1968). Furthermore, short intervals may possibly cause additional discomfort as the uncomfortable stimulation summates. Finally, the number of conditioning trials is usually limited, given that learning in these paradigms often reaches an asymptote within a few trials (a theoretical consideration), and given that peripheral indices of sympathetic activity tend to habituate over trials (a practical consideration; O’Gorman, 1977), rendering it more feasible to have longer stimulus intervals compared to designs with large numbers of trials. Learning paradigms thus seem ideally suited for the application of single-trial pattern analysis. However, although we were able to detect learning-dependent changes in designs with intermediate intervals—provided that the order in which stimuli were presented was fixed and counterbalanced—longer intertrial intervals (> 16 s) still yielded the strongest effects. With longer intervals, the trial-specific activation patterns have more power and are more independent from each other, as both the impact of temporal autocorrelation in the data and collinearity in the model is reduced (see also Mumford et al., 2012, 2014). However, given that longer scans are costly, boring, and increase the risk of head motion, intertrial intervals of medium length (8–12 s) may sometimes be preferred, as long as the experiment is designed in such a way that the comparisons of interest will not be biased by temporal proximity of the trials of interest (i.e., a fixed, counterbalanced stimulus presentation). The latter is not standard in conditioning paradigms and may also pose a problem for the reanalysis of existing data sets.

In conclusion, our data advocate for the use of slow event-related designs if single-trial pattern analysis is the main analysis of interest. Given that slow event-related designs are inefficient for studies in which multiple trials are combined into a single regressor (i.e., the most common univariate approach), an fMRI design will never be optimal for both types of analyses. This limits the possibilities for reanalyzing existing data sets and emphasizes the importance of deciding on the type of data analysis before collecting the data.

References

- Andersson, J. L., Jenkinson, M., & Smith, S. (2007). *Non-linear registration, aka spatial normalisation*. FMRIB technical report TR07JA2. FMRIB Analysis Group of the University of Oxford.
- Chadwick, M. J., Bonnici, H. M., & Maguire, E. A. (2012). Decoding information in the human hippocampus: A user's guide. *Neuropsychologia*, 50(13), 3107–3121. doi: 10.1016/j.neuropsychologia.2012.07.007
- Dale, A. M., & Buckner, R. L. (1997). Selective averaging of rapidly presented individual trials using fMRI. *Human Brain Mapping*, 5(5), 329–340. doi: 10.1002/(SICI)1097-0193(1997)5:5<329::
- Formisano, E., De Martino, F., Bonte, M., & Goebel, R. (2008). "Who" is saying "what"? Brain-based decoding of human voice and speech. *Science*, 322(5903), 970–973. doi: 10.1126/science.1164318
- Fullana, M. A., Harrison, B. J., Soriano-Mas, C., Vervliet, B., Cardoner, N., Ávila-Parcet, A., & Radua, J. (2015). Neural signatures of human fear conditioning: An updated and extended meta-analysis of fMRI studies. *Molecular Psychiatry*, 1–9. doi: 10.1038/mp.2015.88
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293(5539), 2425–2430. doi: 10.1126/science.1063736
- Haynes, J.-D., & Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience*, 8(5), 686–691. doi: 10.1038/nn1445
- Huettel, S. A., & McCarthy, G. (2001). The effects of single-trial averaging upon the spatial extent of fMRI activation. *NeuroReport*, 12(11), 2411–2416. doi: 10.1097/00001756-200108080-00025
- Huettel, S. A., Song, A. W., & McCarthy, G. (2004). *Functional magnetic resonance imaging* (Vol. 1). Sunderland, MA: Sinauer Associates.
- Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2), 825–841. doi: 10.1006/nimg.2002.1132
- Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, 5(2), 143–156. doi: 10.1016/S1361-8415(01)00036-6
- Jennings, J. R. (1987). Editorial policy on analyses of variance with repeated measures. *Psychophysiology*, 24(4), 474–475. doi: 10.1111/j.1469-8986.1987.tb00320.x
- Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, 8(5), 679–685. doi: 10.1038/nn1444
- Kao, M.-H., Mandal, A., Lazar, N., & Stufken, J. (2009). Multi-objective optimal experimental designs for event-related fMRI studies. *NeuroImage*, 44(3), 849–856. doi: 10.1016/j.neuroimage.2008.09.025
- Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, 2. doi: 10.3389/neuro.06.004.2008
- Li, W., Howard, J. D., Parrish, T. B., & Gottfried, J. A. (2008). Aversive learning enhances perceptual and cortical discrimination of indiscriminable odor cues. *Science*, 319(5871), 1842–1845. doi: 10.1126/science.1152837
- Misaki, M., Kim, Y., Bandettini, P. A., & Kriegeskorte, N. (2010). Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. *NeuroImage*, 53(1), 103–118. doi: 10.1016/j.neuroimage.2010.05.051
- Moscovitch, A., & LoLordo, V. M. (1968). Role of safety in the Pavlovian backward fear conditioning procedure. *Journal of Comparative and Physiological Psychology*, 66(3 Pt 1), 673–678. doi: 10.1037/h0026548
- Mumford, J. A., Davis, T., & Poldrack, R. A. (2014). The impact of study design on pattern estimation for single-trial multivariate pattern analysis. *NeuroImage*, 103, 130–138. doi: 10.1016/j.neuroimage.2014.09.026
- Mumford, J. A., Turner, B. O., Ashby, F. G., & Poldrack, R. A. (2012). Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *NeuroImage*, 59(3), 2636–2643. doi: 10.1016/j.neuroimage.2011.08.076
- O'Gorman, J. G. (1977). Individual differences in habituation of human physiological responses: A review of theory, method, and findings in the study of personality correlates in non-clinical populations. *Biological Psychology*, 5(4), 257–318. doi: 10.1016/0301-0511(77)90017-5
- Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, 105(32), 11087–11092. doi: 10.1073/pnas.0805664105
- Rey, H. G., Ahmadi, M., & Quiroga, R. Q. (2015). Single trial analysis of field potentials in perception, learning and memory. *Current Opinion in Neurobiology*, 31, 148–155. doi: 10.1016/j.conb.2014.10.009
- Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, 17(3), 143–155. doi: 10.1002/hbm.10062
- Visser, R. M., Kunze, A. E., Westhoff, B., Scholte, H. S., & Kindt, M. (2015). Representational similarity analysis offers a preview of the noradrenergic modulation of long-term fear memory at the time of encoding. *Psychoneuroendocrinology*, 55, 8–20. doi: 10.1016/j.psyneuen.2015.01.021
- Visser, R. M., Scholte, H. S., Beemsterboer, T., & Kindt, M. (2013). Neural pattern similarity predicts long-term fear memory. *Nature Neuroscience*, 16(4), 388–390. doi: 10.1038/nn.3345
- Visser, R. M., Scholte, H. S., & Kindt, M. (2011). Associative learning increases trial-by-trial similarity of BOLD-MRI patterns. *Journal of Neuroscience*, 31(33), 12021–12028. doi: 10.1523/JNEUROSCI.2178-11.2011
- Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. *NeuroImage*, 14(6), 1370–1386. doi: 10.1006/nimg.2001.0931

(RECEIVED November 4, 2015; ACCEPTED March 18, 2016)