

## Extensive training leads to temporal and spatial shifts of cortical activity underlying visual category selectivity



Tim C. Kietzmann<sup>a,\*</sup>, Benedikt V. Ehinger<sup>a</sup>, Danja Porada<sup>a</sup>, Andreas K. Engel<sup>b</sup>, Peter König<sup>a,b</sup>

<sup>a</sup> Institute of Cognitive Science, University of Osnabrück, 49076 Osnabrück, Germany

<sup>b</sup> Dept. of Neurophysiology and Pathophysiology, University Medical Center Hamburg Eppendorf, 20246 Hamburg, Germany

### ARTICLE INFO

#### Article history:

Received 29 April 2015

Revised 24 March 2016

Accepted 26 March 2016

Available online 06 April 2016

#### Keywords:

Category learning

MEG

Occipitotemporal cortex

Plasticity

Prefrontal cortex

### ABSTRACT

The human visual system is able to distinguish naturally occurring categories with exceptional speed and accuracy. At the same time, it exhibits substantial plasticity, permitting the seamless and fast learning of entirely novel categories. Here we investigate the interplay of these two processes by asking how category selectivity emerges and develops from initial to extended category learning. For this purpose, we combine a rapid event-related MEG adaptation paradigm, an extension of fMRI adaptation to high temporal resolution, a novel spatiotemporal analysis approach to separate adaptation effects from other effect origins, and source localization. The results demonstrate a spatiotemporal shift of cortical activity underlying category selectivity: after initial category acquisition, the onset of category selectivity was observed starting at 275 ms together with stronger activity in prefrontal cortex. Following extensive training over 22 sessions, adding up to more than 16,600 trials, the earliest category effects occurred at a markedly shorter latency of 113 ms and were accompanied by stronger occipitotemporal activity. Our results suggest that the brain balances plasticity and efficiency by relying on different mechanisms to recognize new and re-occurring categories.

© 2016 Elsevier Inc. All rights reserved.

### Introduction

One of the most essential tasks of our visual system is to make sense of the complex signals it receives from the world around us. A central aspect of this is the ability to group objects into various categories, allowing for considerable simplification, generalization and supporting higher cognitive function. To advance our understanding of the underlying cortical mechanisms, a large body of experimental work focuses on temporal aspects of category selectivity, asking for the earliest point in time at which category information is extracted. As a result, we now have ample psychophysical and electrophysiological evidence that naturally occurring categories can be extracted in only little more than 100 ms of processing (Carlson et al., 2013; Antzoulatos and Miller, 2011; Cichy et al., 2014; Hung et al., 2005; Kirchner and Thorpe, 2006; Liu et al., 2009; 2002; Sugase et al., 1999). However, apart from the necessity for fast and robust categorization of re-occurring categories, our ever-changing environment poses the additional challenge to retain considerable plasticity in order to support the rapid learning of entirely novel categories. Here, the study of naturally occurring categories provides only limited possibilities, as it focuses on categories with which we already have extended experience (for instance, all of us can be considered face- and house-experts, as these categories play a vital role in our everyday life). It therefore remains an open question, how cortical

representations and the temporal dynamics of category selectivity develop from the initial learning of a category towards category expertise.

To elucidate this issue, we performed a longitudinal study in which we investigated the impact of extended category training of two artificial visual categories in a parametric feature space on the visually evoked responses using a rapid event-related magnetoencephalography (MEG) adaptation paradigm. Adaptation paradigms, also known as repetition-suppression and repetition-enhancement (Krekelberg et al., 2006; Segaert et al., 2013), are widely applied in the field of functional magnetic resonance imaging (fMRI; see Grill-Spector and Malach, 2001 for an adaptation review) and offer the advantage to bypass the limited spatial resolution of any imaging method by focusing on response-changes in neuronal subpopulations which are measurable in the average response of a pre-defined region of interest (ROI). Adaptation paradigms therefore have the potential to reveal differences in neuronal selectivity that would remain unnoticed in more traditional univariate designs. While the limits of spatial resolution are even more drastic in case of MEG/EEG, these methods offer the possibility to investigate cortical processes with high temporal resolution. A combination of a rapid, event-related MEG adaptation paradigm with perceptual category training is therefore a promising candidate to resolve changes in the temporal aspects of category processing, indicative of changes in the underlying cortical mechanism. An additional advantage of our longitudinal paradigm is a control for effects of low-level stimulus properties. Data recorded during a baseline session allowed us to exclude the possibility that potentially found category effects are an

\* Corresponding author at: Albrechtstr. 28, 49080 Osnabrück, Germany.  
E-mail address: [tkietzma@uos.de](mailto:tkietzma@uos.de) (T.C. Kietzmann).

inherent low-level property of the utilized feature space. Such differences in low-level statistics have previously lead to considerable challenges in the interpretations of category effects in studies using naturally occurring categories (Crouzet and Thorpe, 2011; Rossion and Jacques, 2008; Thierry et al., 2007; VanRullen, 2011).

We investigated the emergence and development of category selectivity by recording MEG data in a baseline session, prior to any category training, a second time after five training sessions, and a third time after extensive category training in 22 training sessions. Category selectivity was estimated by comparing the visually evoked responses to stimuli that were either preceded by a different adaptor stimulus from the same category (category-internal), or by an adaptor stimulus of a different category (category-external), while holding low-level stimulus differences constant. To analyze the high-dimensional MEG data, a novel spatiotemporal analysis approach was employed. Building on the observation that true adaptation effects should occur in the same cortical regions activated previously by the adaptor stimulus, the analysis exploits the linear additivity of MEG sources in order to explicitly separate experimental effects, i.e. differences between category-internal and -external conditions, into adaptation- and other, non adaptation-related effects. In the adaptation phase of the experiment, each trial consisted of two stimuli: an adaptor and a repetition stimulus. Interpreting each MEG topography as a high-dimensional vector, the cortical response to the second stimulus can be understood as a linear combination of (a) a re-activation of the regions previously responding to the adaptor stimulus, as required for adaptation, and (b) other cortical regions, activated uniquely during the processing of the second stimulus. This lead to the insight that the response to the first, adaptor stimulus can be used to decompose experimental effects, observed in the second response, into adaptation-based and non-adaptation effects. Importantly, the involved vector projection maps the 271 dimensional sensor space onto a single, yet highly informative subspace and thereby avoids problems of multiple comparison (see Methods and Materials for details).

Using this approach to focus on effects driven by adaptation we observed a temporal shift in category selectivity from a latency of 275 ms after initial category acquisition to only 113 ms following extensive training. This speedup suggests a marked change in the cortical network mediating the categorization of visual input. Indeed, source analysis revealed an anterior-to-posterior shift of cortical activity from initial to extensive category training. While the time-window of category selectivity found after five training sessions exhibited stronger activation in the prefrontal cortex (PFC), the early category effects found after 22 training sessions showed increased activation in occipitotemporal regions. Previous theories on visual categorization viewed either PFC or regions in the ventral stream as the origin of category selectivity. Our findings now reconcile these contrasting views by suggesting that both processes are likely to contribute to categorization at different stages of category learning. While PFC is involved in the categorization of rather novel and dynamic categories, extensively used categories seem to obtain a privileged status and are resolved faster relying more heavily on cortical resources in occipitotemporal cortex.

## Materials and methods

### Participants

Nine healthy, right-handed subjects (five female, aged 19–30) participated in the study. All subjects had normal or corrected-to-normal visual acuity, were naïve to the purpose of the study and gave written informed consent to participate. The experimental procedures were approved by the ethics committees of the University of Osnabrück and the Ärztekammer Hamburg. Each subject participated in a total of 23 experimental sessions (one baseline session and 22 training sessions). MEG data were recorded during the first baseline session, as well as after training sessions five and 22. The MEG recording from subject nine in

session 22 was excluded from the analyses due to excessive noise in the data.

### Stimulus space

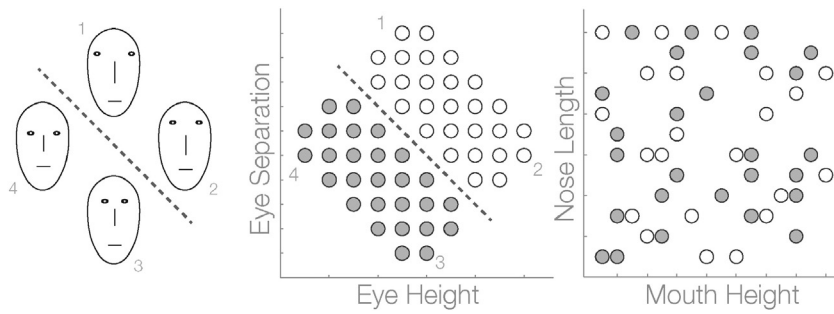
Similar to previous work with macaques (Sigala et al., 2002) and humans (Kietzmann and König, 2010; Nosofsky, 1991; Reed and Friedman, 1973; Sigala et al., 2002), category training was based on two artificial categories of Brunswik faces (Brunswik and Reiter, 1938), defined in a four-dimensional, parameterized stimulus space (Fig. 1), also known as a factorial morphspace (Folstein et al., 2012a; Goldstone et al., 1996; Gureckis and Goldstone, 2008). Two of the dimensions were category-relevant (eye height and eye separation), while the two others (mouth height and nose length) were assigned pseudo-randomly, ensuring that no stimulus clusters of the same category existed that could potentially render these task-irrelevant dimensions informative. A linear category boundary split the category space of the two relevant dimensions in half, such that no single stimulus property was decisive for the category membership. This design is optimized to search for effects of category selectivity, since no linear re-weighting of singular input dimensions will lead to optimal training performance. The overall stimulus space consisted of 60 stimuli, six of which defined the respective category boundary and were not included in the training and testing. The final two categories comprised 27 stimuli each. Moreover, the category boundary was rotated by 90° for every other subject. The subjects were at no point in time instructed about the design of the stimulus space or the relevant category dimensions. Post training, no participant was able to verbally describe the relevant category rule. All stimuli shown during training and the MEG adaptation sessions were presented using the Psychophysics Toolbox 3 (Brainard, 1997; Kleiner et al., 2007) running under Matlab 2010a.

### Category training

In order to allow subjects to learn the two artificial categories of faces, they received category training in a total of 22 sessions with 756 trials each. Here, we largely followed our previous procedure (Kietzmann and König, 2010). In each training trial, subjects were presented with a single stimulus and were then asked to categorize it as either category A or B with their index- or middle-finger. Auditory feedback was provided as training signal. A high-pitch tone indicated a correct response, whereas a low-pitch tone and a forced break of two seconds indicated an incorrect response. To prevent a fixed association between the category membership and motor response, the finger used to indicate the category decision was switched three times across each training session. The subjects were notified of the switches.

### Rapid event-related adaptation paradigm

To estimate the time-course of electrophysiological category effects, we used a rapid event-related MEG adaptation paradigm. This approach is similar to the more common fMRI adaptation (Grill-Spector and Malach, 2001) or repetition suppression/enhancement, and has only recently been introduced to the field of EEG and MEG (Caharel et al., 2009; Harris and Nakayama, 2007; Huberle and Lutzenberger, 2013; Marinkovic et al., 2003; Scholl et al., 2014; Vizioli et al., 2010; Zimmer and Kovács, 2011). While fMRI adaptation paradigms are traditionally associated with effects of repetition suppression, repetition enhancement is now commonly observed across a wide variety of cortical regions (Krekelberg et al., 2006; Segaert et al., 2013). Especially for experiments investigating adaptation effects across time, a temporal sequence of early enhancement and late suppression has been reported (Marinkovic et al., 2003; Petit et al., 2006), in line with the prediction of the more recent accumulation model of adaptation (James and Gauthier, 2006). Taken together, the repeated activation of neuronal



**Fig. 1.** Stimulus space. Subjects were trained to distinguish two artificial categories of faces, defined in a four-dimensional parametric feature space. Two dimensions were category-relevant (eye height and eye separation), and two were irrelevant (nose length and mouth height). No single feature was decisive for the category of a given face, only the combination of features allowed for correct categorization. The category boundary was rotated by 90° for every other subject.

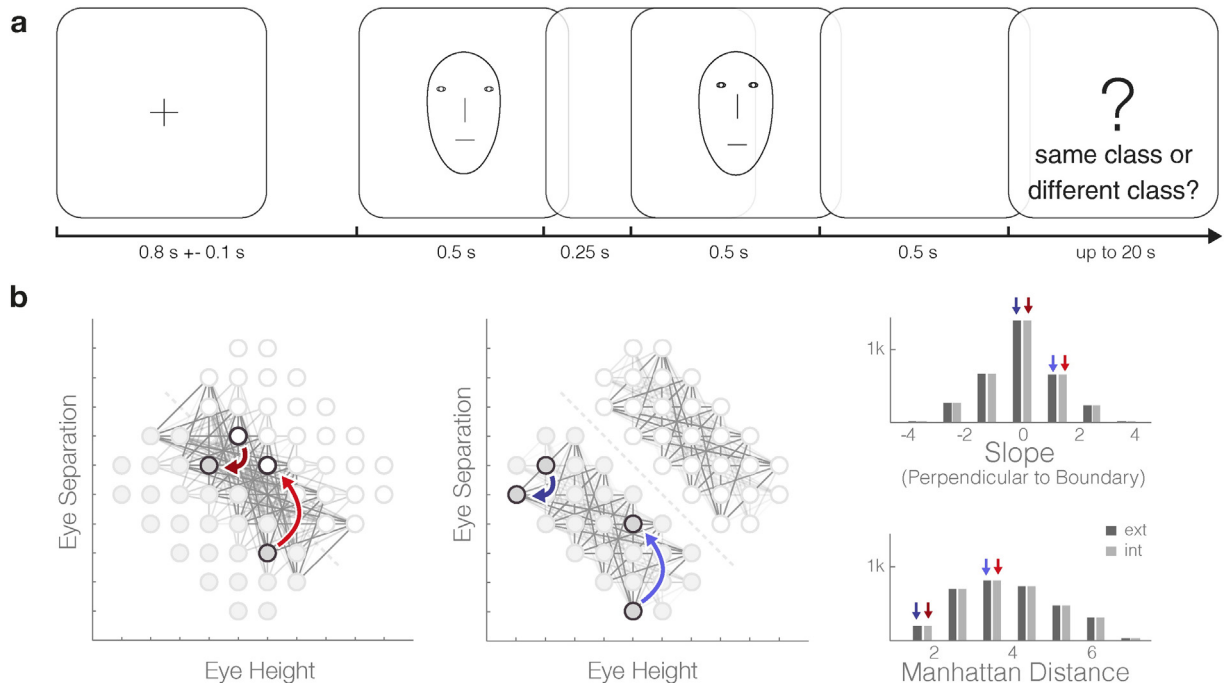
populations can either lead to a suppressed or an enhanced response amplitude in fMRI and MEG/EEG.

During the rapid MEG adaptation paradigm applied here, participants performed a delayed match-to-category (DMC) task, indicating whether two subsequently presented stimuli belonged to the same or a different category (Fig. 2a). To test for effects of neuronal adaptation, i.e. repetition enhancement and repetition suppression, we analyzed the magnetic fields evoked by the second stimulus, when either preceded by a different adaptor stimulus from the same category or an adaptor from a different category. In total, 432 trials were recorded per session and subject, including 216 trials with a category-internal and 216 trials with a category-external adaptor stimulus. The sequence of category-internal and category-external conditions was randomized across trials. To control for low-level feature differences in the two adaptation conditions, all category-internal and category-external trials were matched in distance and direction in the two relevant dimensions of feature space (Fig. 2b). This has the additional advantage that no linear re-weighting of the category-relevant dimensions can account for category selectivity (Folstein et al., 2012a; Goldstone et al., 2001), because all category internal and external adaptation trials will be affected likewise. Our setup is

therefore tuned for observing effects of category selectivity, rather than attentional re-weighting of (single) features, required for later category extraction. During the randomized adaptation trials, a fixed mapping of experimental condition to motor response was prevented by switching the target keys for the two answers after half of the experiment. The structure of the adaptation trials was as follows. First, a fixation cross was presented for 800 ms with an SOA of  $\pm 100$  ms. Then, a first stimulus was presented for 500 ms, followed by an inter-stimulus-interval of 250 ms. The second stimulus was again shown for 500 ms. Finally, after an additional delay of 500 ms a question mark was displayed on the screen, indicating to the subject that a response can be given (Fig. 2a). This timed response was introduced to keep presentation of the second stimulus free of cortical activity related to the motor-execution.

#### MEG acquisition

MEG data were recorded in a baseline session, prior to category training, as well as after five and 22 training sessions. The selection of five and 22 training sessions was based on previous work using a similar feature space, in which subjects were able to perform at >90% accuracy



**Fig. 2.** Adaptation paradigm. To test for electrophysiological correlates of category information, an adaptation paradigm was used. Each trial either crossed the category boundary (category-external) or stayed within a category (category-internal). (a) Temporal sequence of an adaptation trial. (b) Low-level properties of the category-internal and category-external adaptation trials were controlled by matching the distance and slopes of the corresponding stimulus-pairs. Exemplary trials are highlighted in color (category-external in red, category-internal in blue).

after only five training sessions (Kietzmann and König, 2010), while exhibiting high-level category effects only after prolonged training. MEG data were acquired at 1200 Hz using a CTF whole-head system with 271 axial gradiometers (CTF275, VSM MedTech). The position of the participants' head was continuously recorded using three head localization coils (NAS/LPA/RPA). Moreover, a bipolar electrocardiogram (ECG) and an electrooculogram (EOG) with three channels were recorded. The EOG electrodes were placed below the eyes and on the forehead. The reference was positioned on the tip of the nose. The experimental stimuli were back-projected on a screen with a LCD projector (Sanyo XP51) at 60 Hz refresh rate. The presentation distance was 60 cm, leading to a display size of  $2^\circ \times 3.3^\circ$  of visual angle.

#### Data analyses

All analyses were performed using custom code in Matlab R2010a (Mathworks, Natick, MA, USA), fieldtrip (Oostenveld et al., 2011) and Brainstorm (Tadel et al., 2011).

#### Preprocessing

After downsampling the data to 600 Hz, artifacts due to muscle activity, sensor-jumps and extreme noise were first detected automatically using fieldtrip, followed by manual cleaning of the data. To account for sensor drifts, the data were high-pass filtered at 1 Hz. Moreover, frequencies above 100 Hz and the artifactual frequency bands around 50 Hz and 60 Hz were excluded using a zero-phase Butterworth IIR filter. To remove eye-related and cardiac artifacts from the data, we used an automated procedure based on an independent component analysis. The underlying algorithm relies on a correlation-based and a weight-based artifact-metric computed for each independent component. Components surpassing a selected threshold were labeled as artifacts and removed from the data. The optimal thresholds were determined automatically based on a receiver operator characteristic (ROC) analysis applied to a subset of the data for which two experts had classified components as artifacts. The resulting algorithm was able to detect 98.1% of the components tagged by the experts, with only 0.3% false positives (see Material S1 and Fig. S1 for more details). Without making use of additional eye-tracking data, our approach reaches performance levels of a state of the art algorithm for automatic artifact removal that require ground-truth eye-movement data (Plöchl et al., 2012). Finally, although our fully randomized design prevents systematic effects of head-position, we removed any residual effects from our data. We first extracted a six dimensional description of the head position and direction from the simultaneously recorded localization coils (NAS/LPA/RPA) and used this to regress out the effects of head-position (Stolk et al., 2013). All evoked potentials were baseline-corrected with respect to the 700 ms fixation period prior to the presentation of the first stimulus.

#### Spatiotemporal projection approach: separating effect sources

The current experiment makes use of a rapid event-related MEG adaptation paradigm. In each trial, two stimuli are presented: one adaptor and subsequently a second, repetition stimulus. Two experimental conditions are compared. The second stimulus can either be of the same category as the adaptor, or of a different category. Differences between these two conditions thus indicate category selectivity. Common to every adaptation paradigm, experimental effects can either originate from true adaptation, i.e. the differential re-activation of category selective regions (Fig. 3a), or from other sources that are uniquely activated during the presentation of the second stimulus. As an example of the latter, if category-internal and category-external conditions were reported unbalanced, via different hands, condition-dependent differences in the activations of the two motor-cortices would be expected. Such effects, despite originating from category-related signals, would not be due to adaptation (please note that this example is for illustration purposes only, as the current design balanced different motor responses across experimental conditions).

In order to separate these different adaptation and non-adaptation effect sources, we here employ a spatiotemporal analysis approach that decomposes the MEG signals of the second stimulus into parts that are due to the re-activation of regions previously involved in processing the adaptor, and parts that cannot be accounted for by re-activation (Fig. 3b).

*Effects due to adaptation.* To focus on adaptation effects at a given point in time, we project the high-dimensional MEG response vector of the second stimulus ( $\vec{r}$ ) onto the normalized adaptor response vector ( $\hat{a}$ ):

$$\vec{r}_{\hat{a}} = (\vec{r} \cdot \hat{a})\hat{a}.$$

The adaptation-based category effect (*ace*) is then computed as the difference in amplitude between projected category external ( $\vec{e}$ ) and internal ( $\vec{i}$ ) response vectors (Fig. 3c), yielding a scalar estimate of the adaptation-based category effect:

$$ace = |\vec{e}_{\hat{a}}| - |\vec{i}_{\hat{a}}| = \vec{e} \cdot \hat{a} - \vec{i} \cdot \hat{a}.$$

Applied for every point in time, this projection yields one-dimensional “effect traces” for every participant and session, which are subsequently subject to statistical analyses (Fig. 3e).

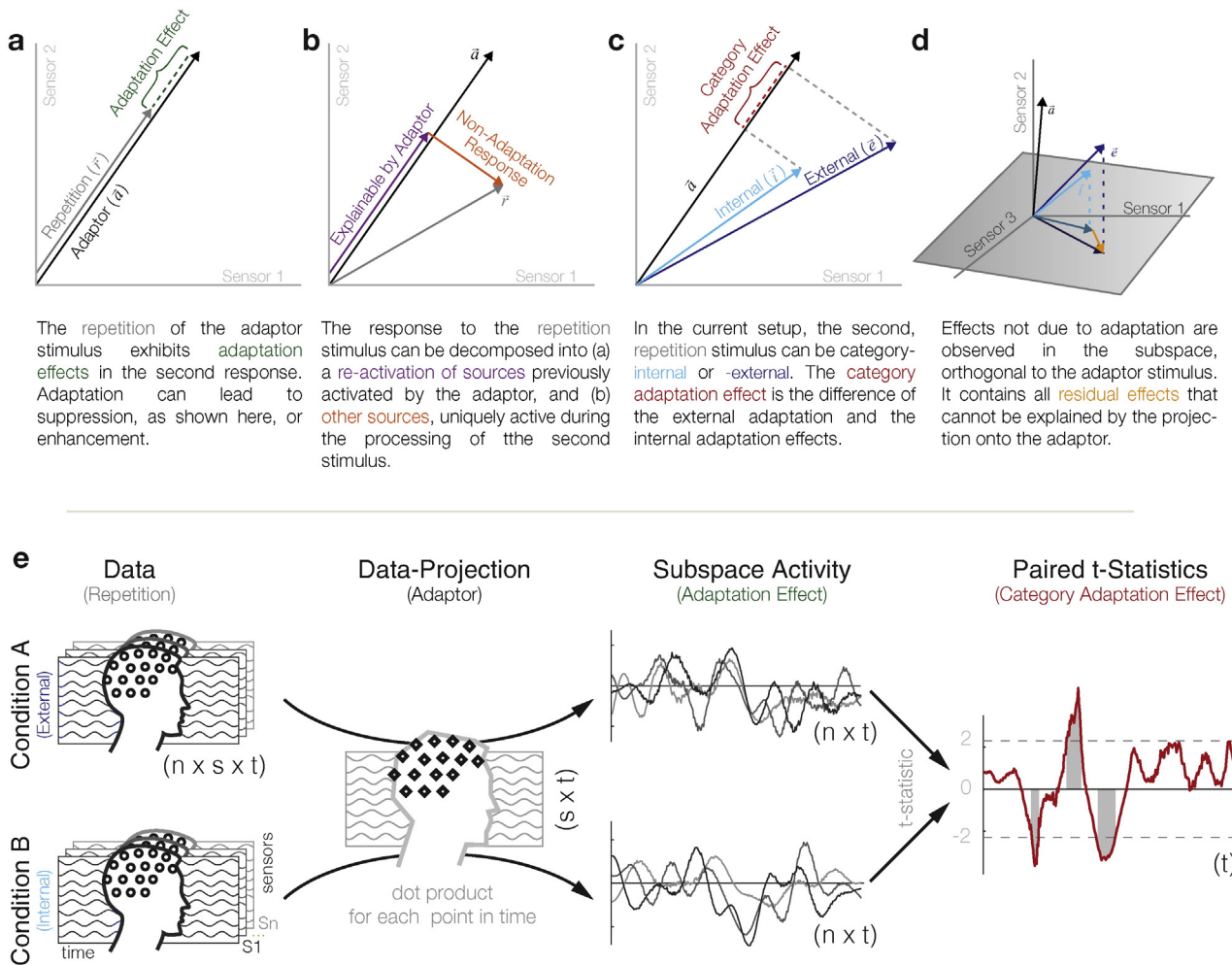
To define the projection vector,  $\hat{a}$ , we here chose to use the group average response, evoked by the first, adaptor stimulus, as recorded in the baseline session (time window between 0 and 300 ms after stimulus onset, low-pass filtered at 35 Hz using a zero-phase Butterworth IIR filter, Fig. 4). Using the same projection vector as basis for effect estimation has the advantage that it allows for comparisons of effect sizes across experimental sessions. If different projection vectors were used for every session, it would introduce unnecessary ambiguity as possible changes in effect-sizes might be merely due to changes in the underlying projection vector. To ensure that the current approach is appropriate, we performed a non-parametric cluster test based on an F-statistic in which we compared the responses to the first stimulus across all three sessions (baseline, session five, session 22) within the first 300 ms of processing (the cluster-threshold was set to  $p < 0.05$ , cluster inclusion was at  $\alpha < 0.05$ ). No significant differences were found, indicating that the same projection vector can be used across sessions. Finally, to be able to compare effects across time, the projection vector ( $\vec{a}$ ) was normalized to unit length ( $\hat{a}$ ) at each sample.

#### Non-adaptation effects

While the projection approach is straight-forwardly applicable to adaptation effects, as the adaptor stimulus uniquely defines the projection target, non-adaptation based effects can occur independently, rendering a projection approach inapplicable: given adaptation-based effects have been accounted for via projection, such residual effects can occur anywhere in the space, orthogonal to the adaptation vector. Thus, to test for such residual, non-adaptation effects, we statistically compared the previously unexplained parts of the external ( $\vec{e}$ ) and internal ( $\vec{i}$ ) vectors (shown as a projection onto the orthogonal plane in Fig. 3d), which reside in the original 271-dimensional sensor space.

In summary, we use the response to the first, adaptor stimulus to decompose the response to the second, category-external or category-internal, stimulus into two possible sources: adaptation-based effects, and other, non-adaptation effect sources. In contrast to this, more traditional analyses of adaptation paradigms leave the origin of observed effects ambiguous, and thereby severely complicate their interpretation. In addition to this important advantage, the current projection method has further statistical benefits because adaptation-based effects can be investigated using a one-dimensional projection-signal, created from the original 271 dimensional data. This circumvents the problems of





The repetition of the adaptor stimulus exhibits adaptation effects in the second response. Adaptation can lead to suppression, as shown here, or enhancement.

The response to the repetition stimulus can be decomposed into (a) a re-activation of sources previously activated by the adaptor, and (b) other sources, uniquely active during the processing of the second stimulus.

In the current setup, the second, repetition stimulus can be category-internal or -external. The category adaptation effect is the difference of the external adaptation and the internal adaptation effects.

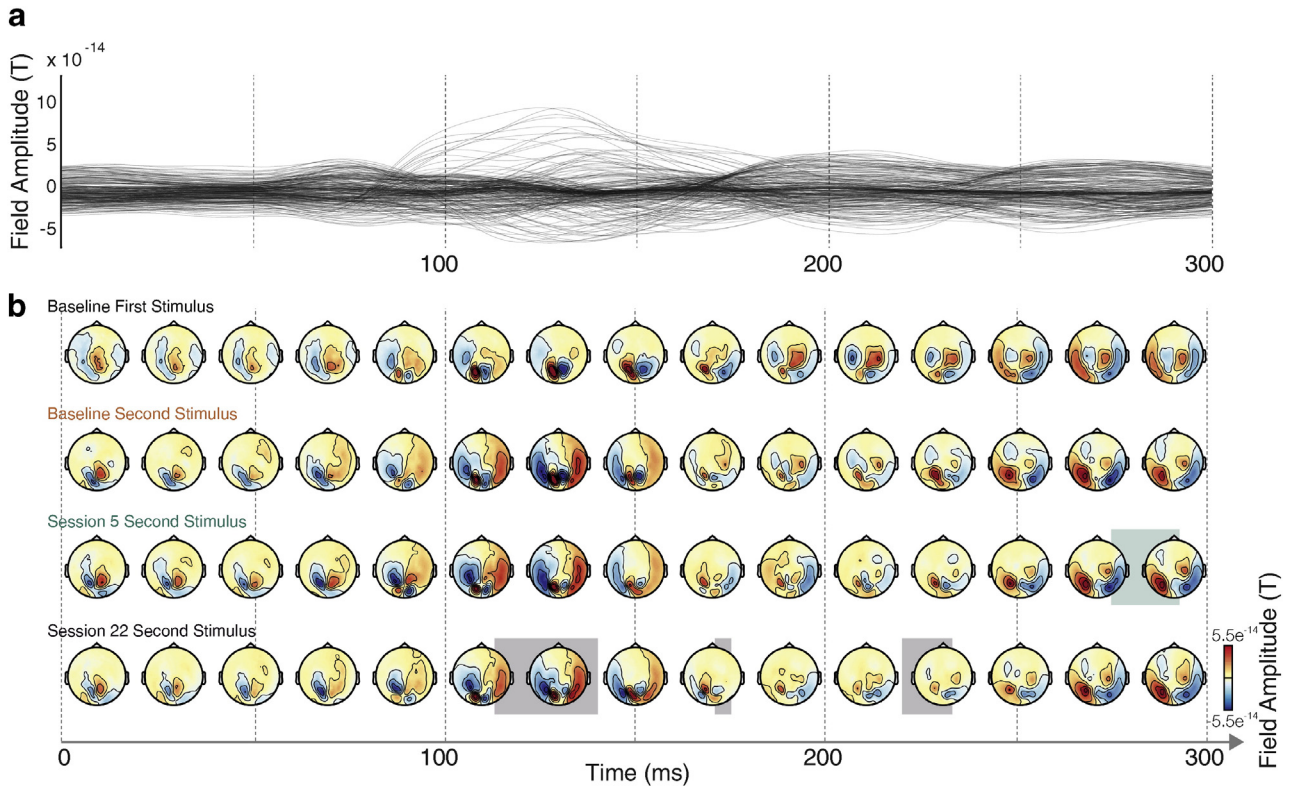
Effects not due to adaptation are observed in the subspace, orthogonal to the adaptor stimulus. It contains all residual effects that cannot be explained by the projection onto the adaptor.

**Fig. 3.** Spatiotemporal projection approach — separating effect sources. A spatiotemporal projection approach was used to decompose the MEG response to the second stimulus, and thereby category effects, into parts due to a re-activation of the regions previously responding to the adaptor stimulus and activity in other cortical regions. (a) At any given point in time, the MEG activity pattern can be interpreted as a vector in high-dimensional space (for clarity, we here show a 2-dimensional example, without loss of generality). Adaptation effects are typically expected to lead to an altered amplitude in the second, repetition response, compared to the first, adaptor response. (b) The response to the second stimulus, however, is a linear combination of a repeated activation of the previously activated sources (as required by the adaptation paradigm), and other sources (e.g. response preparation, etc.). (c) Experimental effects can be decomposed in a similar fashion. To detect effects based on adaptation, the experimental conditions are projected onto the adaptor vector. (d) Non-adaptation based effects are found in the residual activity of both conditions, which reside in a subspace orthogonal to  $\hat{a}$ . If no sensible projection target can be defined a priori, a test can be performed for each sensor and time point, adjusting accordingly for multiple comparisons. (e) Adaptation-based effects are investigated across time, based on the projection-approach yielding activity traces for each subject, condition and session. These traces are subject to statistical analyses, highlighting temporal candidate clusters that show significant differences between category-internal and category-external trials. These candidates are then subject to subsequent analyses. Abbreviations: n, trial number; s, sensor number; t, time point.

multiple comparisons occurring when all sensors are considered individually. Traditionally, this problem is solved by either selecting sensors of interest a priori (Rossion and Jacques, 2008), or alternatively by testing all sensors individually and afterwards controlling for multiple comparisons, for instance by applying non-parametric cluster-based correction methods (Ehinger et al., 2015; Maris and Oostenveld, 2007). Unfortunately, both options are not without problems. An a priori, fixed selection of sensors is problematic, if it is unclear which sensors should be selected or, even more so, if sensors of interest change over time. The second solution, testing all sensors and time-points individually while applying a cluster-based correction, provides the liberty of observing effect clusters anywhere in space and time, but at the cost of much decreased statistical sensitivity. These limitations are overcome by the current projection method. Another statistical consideration worth noting is that the data used for the projection vector are independent of the experimental data in question (comparing category-internal and category-external responses in the second stimulus). This avoids the (spatial) dangers of double-dipping in neuroimaging (Kriegeskorte et al., 2009). Notably, the current projection method is directly applicable

to fMRI adaptation paradigms. In fMRI, the traditional use of ROIs, too, exhibits the problem that observed effects are intermixed and therefore cannot be unambiguously attributed to mechanisms of adaptation. That is, effects observed can either originate from adaptation, or other effect sources.

A further benefit of our approach, exploited later in this paper, is that the effect decomposition allows for efficient source localization of MEG data. Effects due to adaptation can be interpreted as a differential activation of the regions contributing to the first, adaptor stimulus. Therefore, to source localize the effect, the data from the adaptor stimulus can be used. Put differently, the adaptor functions as a spatiotemporal localizer that explicitly focuses the analysis on stimulus-repetition effects. Thus, if an effect is found via projection, the same localizer determines the underlying sources. This approach resembles the standard methodology used for analyses based on independent components, for which effects are first investigated based on component activations and localized based on the component topography (Ehinger et al., 2014; Makeig et al., 2002; Pockett et al., 2007; Tsai et al., 2014). The assumption of our, and in fact any localizer approach, is that the same cortical



**Fig. 4.** Grand average responses to adaptor and repetition stimulus. (a) Butterfly plot of the first, adaptor stimulus. (b) Grand average response topographies of the first, adaptor stimulus during the baseline session (as later used for data projection), and for the second, repetition stimulus, shown separately for each session.

processes are active during the trials used to define the localizer and the experimental trials of interest. While many experimental settings meet this assumption, adaptation paradigms are particularly suited for this approach. This is due to the fact that they already presuppose the same cortical regions to be active during the processing of the first and second stimulus. The use of the first-stimulus response as a spatio-temporal filter for the evoked response to the second is therefore simply a consequent translation of the experimental paradigm to the analysis methodology.

It should be noted that the projection approach, described so far, relies on the assumption that similar cortical regions are active at comparable latencies, because  $\hat{a}$  and  $\vec{r}$  are taken from the same point in time. If responses exhibit significant temporal shifts effects can potentially be missed. To partially counter this effect, we here low-pass filtered the adaptor stimulus, resulting in a temporally more robust fit. As a theoretical alternative, one could use the full response matrices (sensor x time) of adaptor and repetition to explicitly test for effects at all possible delays. Instead of an effect trace, this yields an effect matrix ( $M$ ):

$$M = R \cdot A'$$

The diagonal of this matrix corresponds to a zero-delay and therefore to the effect traces used here. While able to detect effects at different latencies, this approach introduces a quadratic increase in the number of tests. The required corrections for multiple comparisons thereby decrease overall sensitivity.

#### Statistical analyses

For statistical analyses of the adaptation effects, yielded through our spatiotemporal projection method, we computed the one-dimensional response traces for every participant, condition, and session (Fig. 3e). Based on these signals, we then tested for training-induced category effects, following a two-staged approach. First, we temporally localize time-windows of interest, i.e. time-windows exhibiting significant

category effects ( $p < 0.05$ ) by performing a paired t-test at every point in time, contrasting category-external and category-internal trials. Following this, we investigated, whether the observed category selectivity was indeed the result of category training. To accomplish this, we tested, for each candidate time-window, whether the respective effect is significantly larger after training compared to the difference observed in the baseline session (training interaction). This was accomplished by estimating the corresponding interaction effect size and its 95% confidence intervals. Corrections for multiple comparisons across time were performed at this final stage by applying a Bonferroni correction at the cluster-level, i.e. by enlarging the 95% intervals according to the number of clusters tested in each session. As a result of this statistical procedure, any cluster reported in the following will not only have shown significant category effects, but also a significant training interaction, verifying that the found effects are indeed caused by category training. Testing for a training interaction is an important additional prerequisite in investigations of developing category selectivity, as observed differences between category-internal and category-external conditions could also be an inherent property of the selected feature space and not the result of category training. This possibility is ruled out by the statistical procedure described.

Summing up, we focused on adaptation-based effects by combining a spatiotemporal projection method with rigorous statistical analyses. This allowed us to overcome the need to use heuristics in selecting sensors and time-windows of interest, while controlling for multiple comparisons in space and time. The only free parameter of the overall approach for finding adaptation-based effects is the p-value for the selection of temporal candidate windows, which was selected to be  $p < 0.05$ .

To test for non-adaptation effects, we performed a two-sided t-test for every point in time and space, and corrected the family wise error rate using a nonparametric, cluster-based permutation test (cluster inclusion threshold  $\alpha < 0.05$ , left- and right-sided cluster  $p < 0.025$ , respectively). This approach can find “unpredicted” effects, but at the cost

of decreased statistical sensitivity. All analyses performed, adaptation and non-adaptation, focus on the first 300 ms of processing after stimulus onset, as this time-window approximately resembles typical fixation durations during free-viewing of natural scenes (Underwood et al., 1998).

#### Behavioral relevance

To estimate the behavioral relevance of the observed category effects, we contrasted the effect size of adaptation trials in which the response of the subject was correct and trials in which an incorrect response was given. The reasoning of this approach was that if the found effects are behaviorally relevant, larger effects should be expected upon correct performance in the delayed match-to-category task. Similar to the statistical analyses of the training-interactions, we focused on clusters that exhibited significant category effects and training interactions, estimated the effect size and bootstrapped the respective upper- and lower bounds of the 95% confidence intervals (with replacement) while applying a Bonferroni correction for multiple comparisons at the cluster level. Matching the behavioral accuracy in the DMC task, on average 136 incorrect trials were compared to 277 correct trials in session five, whereas 94 incorrect trials were compared to 277 correct ones in session 22.

#### Source analysis

To compare source activity on the cortical surface, we used the sLORETA algorithm (Pascual-Marqui, 2002), as implemented in the Brainstorm software (Tadel et al., 2011), on the adaptor stimulus data, which was used as projection target, to localize the adaptation-specific category effects. For every subject, we first segmented the individual MRI into white and gray matter using Freesurfer (Dale et al., 1999; Fischl et al., 1999). We then performed the source reconstruction based on each individual anatomy and aligned the results to MNI space (Colin27) using spherical averaging of the cortical surfaces. For statistical analyses, we contrasted the average source activity (L2-Norm) during the earliest time-window of category selectivity in session five (275–293 ms) and session 22 (113–140 ms) at every surface vertex and applied a clusterwise correction for multiple comparisons based on a nonparametric permutation test (Maris and Oostenveld, 2007). Only vertices showing  $p < 0.05$  were included in the cluster estimates.

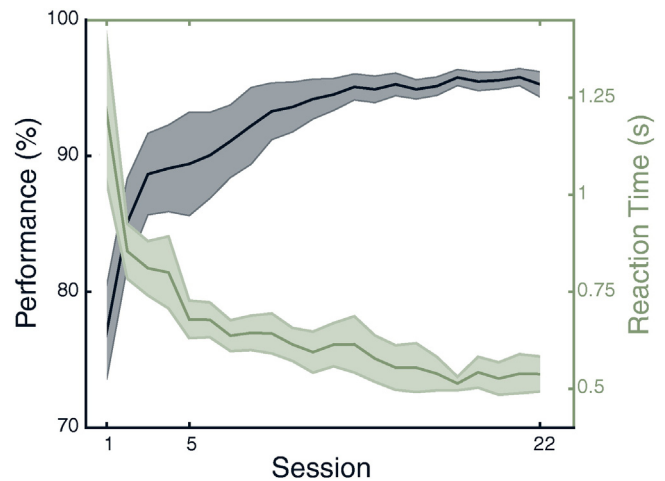
## Results

#### Behavioral data on category training

Subjects were trained to categorize two artificial categories of faces defined in a parametric feature space (Fig. 1). Training lasted for a total of 22 sessions consisting of 756 training trials each. In each trial, participants were required to make a category judgment for a given stimulus and received auditory feedback as training signal. Classification accuracy reached 89.4% after five training sessions, and 95.3% after training was completed in session 22 (Fig. 5). At the same time, reaction times continued to decrease with training (from 679 ms in session five to 538 ms in session 22,  $p < 0.01$  paired t-test). Thus, although high classification performance was reached already after five training sessions, the behavioral data indicate continued improvements over the whole training period.

#### Behavior during the delayed match-to-category task

The electrophysiological correlates of category effects were estimated using an adaptation paradigm in which subjects performed a delayed match-to-category task (Fig. 2a). During the baseline session, and therefore prior to category training, the DMC performance of our subjects did not differ significantly from chance (49% accuracy,  $p = 0.128$ , t-test against a chance-level of 50%). This demonstrates that our artificial



**Fig. 5.** Effects of training on performance and reaction times. Subjects received auditory feedback as a training signal, but no explicit information about the underlying feature space or category structure. Recognition performance and reaction times improved until session 22, illustrating continued training effects. Error bars depict SEM.

category structure is not an inherent property of the stimulus space. With training, DMC performance increased to 66.2% in session five and 76.0% in session 22 (Fig. S2). A repeated measures ANOVA (with session (baseline, five, 22) and category membership (internal, external) as factors) revealed a significant main effect of session ( $p < 0.01$ , all pairwise comparisons are significant at  $p < 0.01$ , t-test, Bonferroni corrected), but no main effect of category membership ( $p > 0.05$ ) and no significant interaction ( $p > 0.05$ ). Thus, although there was an overall increase in task performance with training, there was no significant difference in the performance of the category-internal and category-external trials indicating that the task was equally demanding in trials of both conditions. Effects of condition difficulty can therefore not explain the categorical effects observed.

The accuracy in the DMC task was lower than expected from the high training performance (95% training accuracy predicts around 90% accuracy for two consecutive decisions). This is in line with observations by Helie and Ashby (2012), who observed sub-optimal DMC performance even for comparably simple one-dimensional category boundaries. Multiple reasons can account for this difference. First, a successful DMC trial requires, in addition to the correct classification of both stimuli, successful working memory encoding and retrieval, a successful category comparison, and a successful match to the correct motor response. Moreover, the electromagnetic shielding required for the MEG measurements required the use of a back-projected display with decreased contrast compared to the training monitor. Most importantly, our participants had considerably less experience with the structure of the DMC task, compared to the excessive amount of trials in the training paradigm.

#### Training-induced category effects in visual responses (MEG data)

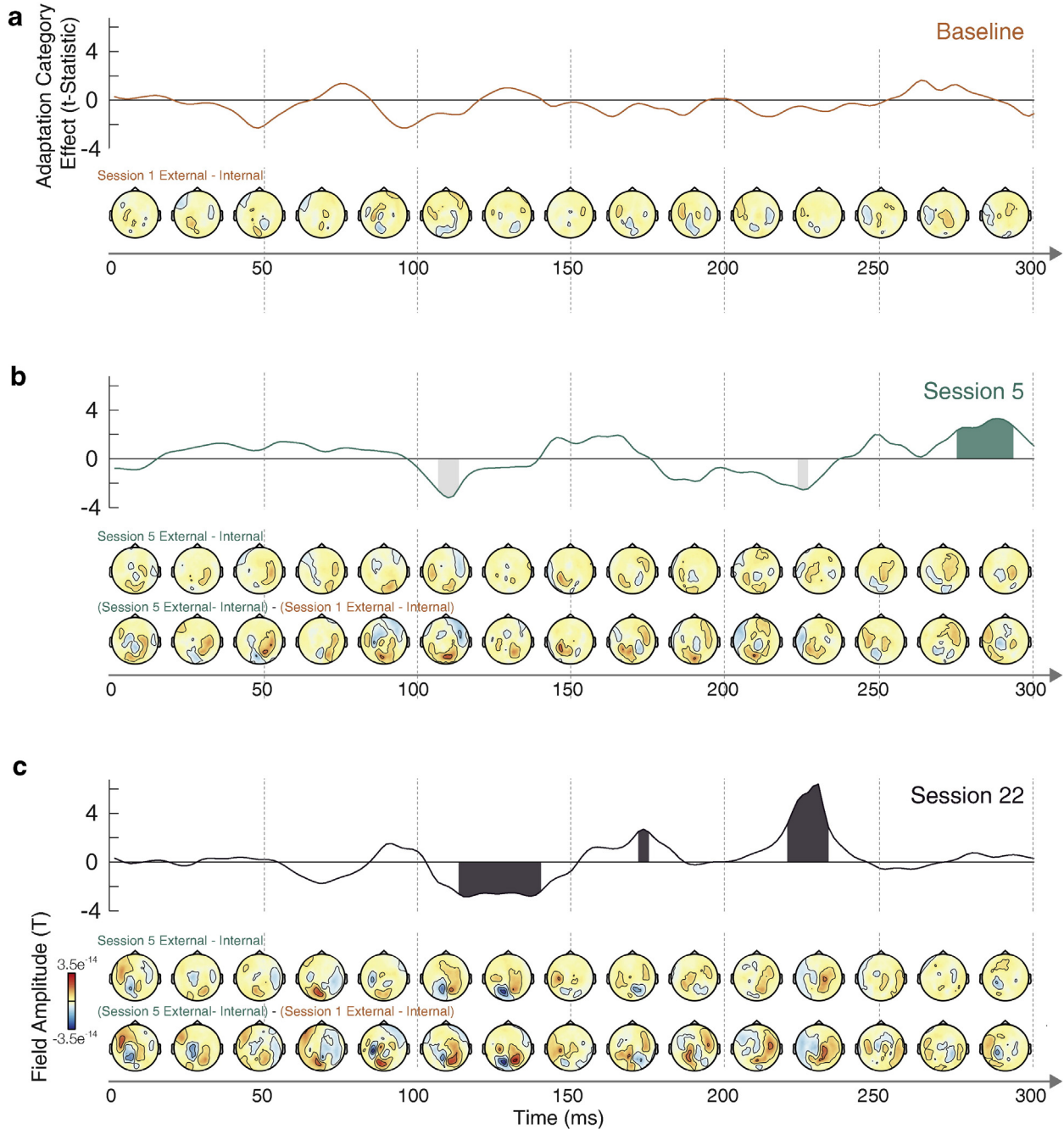
To test for category effects in the visually evoked responses, we compared the magnetic fields evoked by the second stimulus in the category-internal and category-external adaptation trials in the MEG adaptation paradigm, while controlling the low-level stimulus properties of the two conditions (Fig. 2b). This indirectly tests for category selectivity, as differences between these two conditions will only be detectable if category-information is encoded in the underlying cortical activity. Importantly, adaptation paradigms were previously shown to result in both, effects of repetition suppression as well as repetition enhancement (Krekelberg et al., 2006; Segaert et al., 2013), depending on stimulus timing (James and Gauthier, 2006), effect latency (Marinkovic et al., 2003; Petit et al., 2006), and region of interest (Zago et al., 2005).



For analyses of visually evoked responses, we employed a spatio-temporal projection approach that allowed us to focus our analyses on adaptation-based effects, and an unconstrained cluster-based analysis for effects that are not due to adaptation.

Focusing on adaptation-based effects first, we projected the evoked fields in response to the second stimulus onto the adaptor response, and thereby created activity traces for each session, subject and condition (category-external and category-internal). We then performed a paired t-test at every time-point to test for differences between category-external and category-internal conditions (positive t-values

indicate a larger response for category-external trials). This provided us with temporal candidate clusters that exhibit significant category effects for every session. To ensure that the observed category effects were indeed the result of category training, it had to be shown that category effects were significantly larger post-training as compared to the baseline session. As a final step, we therefore estimated the effect sizes and confidence intervals of the training interaction for each temporal candidate cluster (Bonferroni corrected at the cluster level, thereby controlling for multiple comparisons). Only temporal clusters surviving this rigorous control will be reported in the following. The clusters reported



**Fig. 6.** Training-induced MEG category effects across time. The traces in each panel represent the adaptation-based t-statistics across time, as obtained from the spatiotemporal projection method, comparing category-internal and category-external trials. Candidate temporal windows during which visually evoked responses showed significant category effects and a significant training interaction are shaded in dark colors. Candidate windows exhibiting no significant training effects are marked in light gray. Together with traces of t-statistics, each panel shows the underlying effect topographies and training interactions (i.e. effect difference between baseline and post-training) where applicable. (a) During the baseline session, no significant category effects could be found. (b) After five training sessions, the first significant training-induced window of category selectivity is present from 275 to 293 ms. (c) After extended category training in 22 sessions, the earliest training-induced category effects are present from 113 to 140 ms. Additional clusters of significant training-induced category effects were found between 171 and 175 ms and 220–233 ms.



will not only exhibit significant category effects, but also show significantly stronger category effects compared to baseline, indicating that the seen category effects are the result of category training.

We first analyzed the data from the baseline session. Here, we found no significant category effects (Fig. 6a), confirming that the category structures used for training were not an inherent property of the used stimulus space. We then analyzed the data of the two post-training sessions five and 22, testing for category effects and training interactions. After five training sessions, the earliest significant, training-induced category effects were evident between 275 and 293 ms (Fig. 6b, Fig. S3a). With developing category expertise, however, a temporal shift in category effects was observed. After 22 training sessions, the earliest cluster exhibiting significant training effects occurred already after 113 ms (lasting from 113 to 140 ms). Additional time-windows of significant category effects were found between 171 and 175 ms and between 220 and 233 ms (Fig. 6c, Fig. S3b). The corresponding template and effect topographies are shown in Figs. 4 and 6, respectively (individual sensor-traces are provided in Fig. S4). Importantly, the earliest cluster found in session 22 (113 to 140 ms) is not only significantly different to the baseline session, indicating an overall training effect, but also significantly larger compared to session five ( $CI_{95} = [-8.47 * 10^{-15}, -1.55 * 10^{-13}]$ ). The latter indicates that the extensive training between sessions five and 22 lead to the temporal shift in category effects. The observed speed-up of more than 160 ms from session five to 22 is remarkable, as our subjects already categorized the stimuli at about 90% accuracy during training session five. Moreover, it is comparable to the observed decrease in reaction times of around 140 ms from training session five to 22.

As a necessary result of the close control of low-level feature differences (same stimulus-space distance and direction for category-external and category-internal trials), stimuli close to the category boundary were shown more frequently in the DMC task. Based on this, it could be argued that effects of long-term adaptation might specifically affect category external trials, thereby contributing to the early category-effects observed in session 22. Speaking against this possibility we found no significant differences during the baseline session. Furthermore, we report only clusters exhibiting both, significant category effects and training interactions to ensure that the effects reported are indeed the results of category training.

Following adaptation-based effects, we tested whether category-specific effects exist that are not due to adaptation (Fig. 3). Based on the residual data, i.e. the parts of the evoked fields that cannot be explained based on the adaptor response, we compared category-internal and category-external conditions, while correcting for multiple comparisons using a nonparametric cluster-based permutation test. This analysis revealed no significant effects of category selectivity (all cluster  $p > 0.3$ , residual effect topographies shown in Fig. S5). This highlights the successful balance of motor-response mapping across conditions and indicates that, in the current setup, category-selectivity was only observed adaptation-based.

#### *Relation of physiological category effects to behavior*

To test whether the observed adaptation-based category effects were behaviorally relevant, we compared the category effect sizes for successful and erroneous trials during the delayed match-to-category task. Again, we estimated the effect sizes and confidence intervals, while Bonferroni correcting for multiple comparisons at the cluster level. This analysis revealed significant differences for the earliest cluster in session 22, indicating the behavioral relevance of the effect. No other cluster in session 22 and five exhibited significant behavioral effects. Considering the absence of significant differences for session five, it should be noted that behavioral errors in the delayed match-to-category task can have various origins. Apart from the variability in the category signal, which is of interest here, they include subjects' inattentiveness, errors in working memory and an incorrect mapping of the

perceptual decision to the appropriate behavioral response. These additional sources of error significantly complicate the search for behavioral relevance, as they all do not predict differences in category-selectivity. Moreover, it is possible that effects of behavioral relevance occurred at an even later point, extending beyond the 300 ms analyzed here.

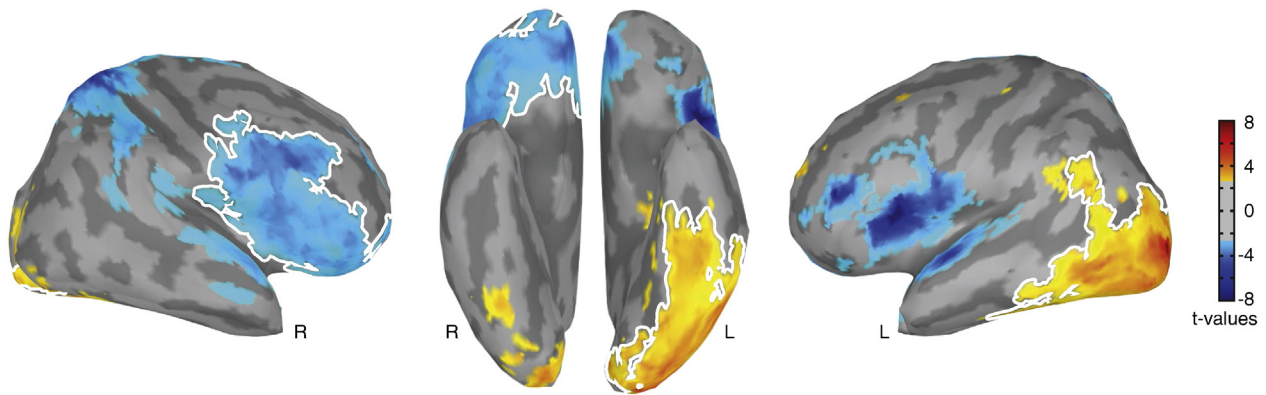
#### *Source analyses*

Following the analyses in sensor space, we tested whether the temporal shift in category selectivity observed between session five (275–293 ms) and session 22 (113–140 ms) is due to altered neuronal processing in the same cortical areas, or whether different sets cortical areas are activated during these two time windows of interest. To this end, we computed a standardized low resolution brain electromagnetic tomography (sLORETA) (Pascual-Marqui, 2002) on the data of the adaptor stimulus, which underlies the observed category effects (see Materials and Methods for details). We estimated the average source activations during the two time-windows of interest and tested for significant differences based on a t-statistic, while controlling for multiple comparisons using a nonparametric cluster-based permutation test (Maris and Oostenveld, 2007) on the cortical surface. This analysis revealed that the previously shown temporal shift in category selectivity was accompanied by an anterior-to-posterior shift of cortical activation (Fig. 7, positive t-values indicate a stronger activation in the early time-window observed in session 22). Although the source distributions exhibited considerable overlap, among others in parietal regions (see Fig. S6 for source activations of all clusters compared to baseline), the time-window of training-induced category effects in session five showed a significantly stronger activation in the ventrolateral and ventromedial parts of the PFC. In contrast, the cortical activation during the earlier time-window of category selectivity in session 22 exhibited significantly stronger activity in more posterior regions, including the occipitotemporal cortex. All results were reproduced in a separate source localization analysis based on dynamical statistical parametric mapping, dSPM (Dale et al., 2000) instead of sLORETA (Fig. S7).

As stated above, the source analysis approach taken here specifically focuses on time-points at which adaptation-effects were observed. Using the adaptor response as basis, it highlights sources that are differentially activated at the two earliest significant timepoints in sessions 5 and 22. This assumes that the respective effects observed are the result of altered activity in the strongest sources. This is not necessarily true, as the projection approach can yield significant results that are driven by weaker sources and sensor-patterns. In the current case, however, this concern is not warranted, as the effect topographies nicely match the topographies of the adaptor stimulus (Fig. S4). Moreover, the fact that no category-effects were observed in the residual activity indicates that the localized sources are the sole contributor to category selectivity in the current data.

#### **Discussion**

Previous work on naturally occurring categories has demonstrated that category information can be rapidly extracted from visually presented objects. It remained unclear, however, how the visual system copes with the challenge to reach such rapid recognition speeds while at the same time allowing for sufficient plasticity to encompass the fast learning of entirely new categories. Are the same neuronal mechanisms and structures involved in recognizing re-occurring and newly learned categories, or are they different? And, if they are different, are novel categories implemented differently with prolonged experience? Here we investigated these issues by extensively training nine subjects to categorize two artificial visual categories. During training, we recorded MEG data in a rapid event-related adaptation paradigm to investigate the emergence of category selectivity in visually evoked responses. Additionally, MEG data were recorded prior to category training to serve as a baseline. Using a novel data projection approach, which



**Fig. 7.** Source localization results. Source activations of the earliest clusters of category selectivity in sessions five and 22 were contrasted. Shown are uncorrected t-values with a cutoff at  $p < 0.05$ . Blue regions show larger activity during the category-selective time-window in session five, red regions show larger activity in the early category-cluster in session 22. A white border highlights clusters after controlling for multiple comparisons (cluster-based permutation test).

allowed us to separate adaptation-based and non-adaptation effects, we demonstrate the emergence and, following this, a temporal shift in category selectivity. The data recorded in the baseline session did not exhibit any category effects, indicating successful control for low-level stimulus properties. After five training sessions, the earliest training-induced category effects were found around 280 ms of processing. With extensive training in 22 sessions, we observed a temporal shift in category selectivity. The first significant differences were now found about 160 ms earlier, between 113 and 140 ms. We then investigated whether the temporal shift in category selectivity was accompanied by changes in the spatial pattern of the underlying cortical activity. We compared the source activations during the two earliest temporal clusters of sessions five and 22 and found a significant anterior-to-posterior shift. While the cortical activity during the late category effects in session five showed stronger signals in PFC, the early time-window of category selectivity in session 22 exhibited an increased activation in occipitotemporal regions.

An interesting aspect of the results is that temporally late effects in session five and 22 exhibit positive t-values, suggesting a decreased response for category-internal compared to category-external trials and therefore repetition suppression. The earliest effect observed, cluster one in session 22, however, exhibits a reverse effect, indicating effects of repetition enhancement. This finding is in line with previous EEG adaptation experiments that demonstrated early enhancement, but late suppression effects (Marinkovic et al., 2003; Petit et al., 2006), and contributes to an ongoing debate about the mechanisms underlying differential repetition effects in electrophysiology and neuroimaging (James and Gauthier, 2006; Krekelberg et al., 2006; Segaert et al., 2013).

Our finding of an early cluster of category selectivity, starting at 113 ms and lasting until 140 ms, is fully compatible with previous studies of natural categories in macaque and human. In the macaque, Sugase et al. (1999) recorded from inferotemporal cortex (IT) and observed a peak in category information after only 117 ms of processing. In line with this, Hung et al. (2005) demonstrated that relatively small numbers of randomly selected neurons in IT allow for reliable category decoding, peaking 125 ms after stimulus onset. Interestingly, the authors also show decoding of low-level properties such as size and position of an object, arguing for residual retinotopic information in the neuronal response. This emphasizes the necessity to control for low-level stimulus properties and underlines the benefits of baseline measurements in category training. Finally, Freedman et al. (2003) applied a receiver operator characteristic approach to recordings from macaque IT and PFC. They showed that IT cells exhibited category selectivity after 127 ms. In humans, electrocorticographic recordings provided direct evidence that natural categories can successfully be decoded at a mean latency of 115 ms (Liu et al., 2009). Remarkably, decoding was possible

based on single trials, allowing for generalization across rotation and changes in scale. In line with this, MEG recordings of human subjects provided evidence that visually evoked responses of houses and faces can be separated already at the time of the M100 component (Liu et al., 2002). In the same study, a positive correlation of response amplitude and categorization performance was shown, indicating the behavioral relevance of the early category signals. Using a multivariate decoding approach, Carlson et al. (2011) showed that it is possible to differentiate two visual categories (faces and cars) after 135 ms of processing, even if the retinal locations of trained and tested stimuli were different. Similar results were later obtained using a wider range of categories (Carlson et al., 2013). Extending this approach, Cichy et al. (2014) performed a temporally fine-grained representational similarity analysis based on 92 object images and demonstrated successful decoding of different types of category selectivity at approximately the same latency. The authors furthermore showed a correlation between the brain responses in macaque and human, providing further evidence for a common representational space (Kriegeskorte et al., 2008). These results of early category selectivity were extended to a more natural, cluttered stimulus set contrasting faces to other stimulus categories (Cauchoix et al., 2014). Finally, electrooculography (EOG) data provided by Kirchner and Thorpe (2006) suggest that category information is present and behaviorally relevant after only 120 ms of processing. However, it should be noted that all of the studies mentioned above either investigated neuronal responses to naturally occurring categories or did not include a pre-training baseline. Apart from the inherent challenges to differentiate category selectivity from systematic differences in the low-level statistics (Crouzet and Thorpe, 2011; Rossion and Caharel, 2011; Thierry et al., 2007; VanRullen, 2011; Wichmann et al., 2010), these setups do not allow for an investigation of emerging category selectivity with increasing category experience, which is the focus of the current study.

Overall, the neuronal mechanisms underlying the categorization of visual input have been in the focus of a lively debate over the recent years. A prominent view centers around the idea that category information is extracted by PFC (Antzoulatos and Miller, 2011; Cromer et al., 2010; Roy et al., 2010; Serre et al., 2007). Accordingly, neuronal selectivity in temporal regions is seen as merely providing a sufficiently complex vocabulary from which the category information can be flexibly read out. This view is consistent with the predictions of the two-stage model of perceptual category learning (Riesenhuber and Poggio, 2002), which hypothesized that neurons in IT obtain sharper tuning to re-occurring stimulus features, while regions in frontal cortex learn to associate these features with the corresponding category membership. In humans, experimental evidence supporting such division of labor was provided by Jiang et al. (2007). They showed that category training can lead to an increased shape selectivity in ventral areas whereas

category selectivity was found only in the lateral PFC (but see Minamimoto et al., 2010). Moreover, there is evidence for enhanced shape selectivity in ventral areas in human and macaque (Freedman et al., 2006; Sigala and Logothetis, 2002; van der Linden et al., 2013; van der Linden et al., 2010). Nevertheless, the large body of evidence for rapid category selectivity in IT, as reviewed above, supports a contrasting view according to which category information might already be extracted at the level of the temporal lobe (DiCarlo et al., 2012; Liu et al., 2013; Mur et al., 2012). Closely mirroring this controversy, different labs have studied the cortical representations of spatial and motion-related categories in the parietal and prefrontal cortex, arriving at opposite conclusions. Whereas some observed stronger and earlier category signals in prefrontal compared to parietal cortex (Crowe et al., 2013; Goodwin and Blackman, 2012), others reported the reverse: earlier category-selective signals in parietal cortex preceding prefrontal category selectivity (Fitzgerald et al., 2012; Swaminathan and Freedman, 2012). Providing a potentially unifying solution to these controversies, we have demonstrated here that prolonged category training can lead to a temporal shift in category selectivity, which is accompanied by an anterior-to-posterior shift in cortical activity. These data provide a clue as to how the brain could balance the need for robust and fast recognition of re-occurring categories while still allowing for considerable flexibility and rapid plasticity. Selectivity for novel categories relies more heavily on PFC and, as indicated by the long latency of the observed effect, potentially recurrent processing. Sufficient expertise with the categories, however, leads to changes in the cortical implementation of the trained categories, thereby allowing for a substantial speedup in processing times and emphasizing cortical processes in occipitotemporal regions.

A comparable view was recently described by Seger and Miller (2010) who proposed that the brain might simultaneously implement fast and slow learning processes. Fast learning provides multiple advantages, such as increased flexibility and rapid adjustments, but at the cost of an increased risk of erroneous classification. Slow learning, on the other hand, is less error-prone but at the cost of extended training requirements. In line with this suggestion, Helie et al. (2010) trained participants in a rule-based categorization task and demonstrated an initial transition from subcortical to cortical areas, including PFC, and a second transition towards the premotor cortex with emerging automaticity. The current setup, using MEG, is not particularly suited to resolve subcortical activity. However, it is possible that a similar transition from subcortical to cortical areas also occurred in our participants during initial category training, potentially even earlier than our first post-training MEG recording. The question of the respective contribution of subcortical and cortical regions in category learning was recently addressed in the macaque (Antzoulatos and Miller, 2014; Antzoulatos and Miller, 2011; Muhammad et al., 2006), suggesting that the striatum is indeed involved during initial category learning, potentially entraining prefrontal circuitry.

In line with this suggested learning transition, our results provide a potential explanation as to why some previous studies did not see (early) category selectivity in temporal areas after category training (Gillebert et al., 2009; Jiang et al., 2007; Li et al., 2007; Scholl et al., 2014). Apart from many differences between these experiments and our study, our data suggest that the extent of training is a decisive factor. Comparably short training times might only reveal rather late category selectivity in frontal regions, as observed in session five here, whereas prolonged training is required for early occipitotemporal effects. Another important difference is given by the type of category space used during training. Using psychophysical measurements Folstein et al. (2012a) demonstrated that factorial, but not blended morphspaces, lead to an enhanced discriminability of category-relevant feature dimensions, implying that studies using the latter (Gillebert et al., 2009; Jiang et al., 2007; Scholl et al., 2014; van der Linden et al., 2010), were less likely to observe category-selectivity in visual areas. This

suggestion was corroborated by a follow-up fMRI study in which the authors demonstrated robust category signals in visual areas (see Folstein et al., 2015 for a review; Folstein et al., 2012b). In line with such evidence for category-selectivity in visual areas, effects of expertise have been demonstrated in the FFA (Gauthier et al., 2000), and the N170 ERP component (Tanaka and Curran, 2001).

By contrasting correct and incorrect responses, we demonstrated significant behavioral relevance of the early category effects starting at 113 ms in session 22. It has to be noted, however, that the time-points of category selectivity observed in sessions five and 22 do not necessarily mark the end point of the perceptual decision process. Successful performance in the DMC task requires the successful completion of additional processing steps, such as the successful comparison of the two shown categories and the mapping of the perceptual decision to the appropriate motor response. Moreover, effects of perceptual certainty (Philiastides and Sajda, 2006) and ongoing evidence accumulation (Donner et al., 2009) can be expected to play a vital role in the perceptual decision process.

While further experiments are required to fully disentangle the contribution of these different factors, we have shown here that the brain is capable of extracting visual categories based on two different modes. Novel categories are recognized late, involving recurrent processing and increased activity in PFC. This pattern of results is consistent with a re-labeling of existing visual features, which would allow the system to flexibly learn new categories and to quickly adjust to changing task-demands (Mckee et al., 2014). Extended category experience, however, leads to a significant speed-up in category selectivity, accompanied by increased activity in occipitotemporal cortex. This suggests that re-occurring categories are processed differently to allow for quick and reliable recognition. Taken together, our results suggest that the brain balances plasticity for acquisition of new and efficiency in processing of known categories by relying on different networks.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2016.03.066>.

## Acknowledgements

We would like to thank José P. Ossandón for valuable discussions and input during the data analyses, Guido Nolte for his advice on MEG source localization and Maïke Gaca for her help during data acquisition. This work was supported by the EU (grants ERC-2010-AdG-269716 and H2020-641321) and by the DFG (SFB 936 and SPP1665).

## References

- Antzoulatos, E.G., Miller, E.K., 2011. Differences between neural activity in prefrontal cortex and striatum during learning of novel abstract categories. *Neuron* 71, 243–249. <http://dx.doi.org/10.1016/j.neuron.2011.05.040>.
- Antzoulatos, E.G., Miller, E.K., 2014. Increases in functional connectivity between prefrontal cortex and striatum during category learning. *Neuron* 83, 216–225. <http://dx.doi.org/10.1016/j.neuron.2014.05.005>.
- Brainard, D.H., 1997. The psychophysics toolbox. *Spat. Vis.* 10, 433–436.
- Brunswick, E., Reiter, L., 1938. Eindrucks-charaktere schematisierter Gesichter. *Z. Psychol.* 142, 67–134.
- Caharel, S., d'Arripe, O., Ramon, M., Jacques, C., Rossion, B., 2009. Early adaptation to repeated unfamiliar faces across viewpoint changes in the right hemisphere: evidence from the N170 ERP component. *Neuropsychologia* 47, 639–643. <http://dx.doi.org/10.1016/j.neuropsychologia.2008.11.016>.
- Carlson, T., Hogendoorn, H., Kanai, R., 2011. High temporal resolution decoding of object position and category. *J. Vis.* 11, 1–17. <http://dx.doi.org/10.1167/11.10.9.Introduction>.
- Carlson, T., Tovar, D., Alink, A., Kriegeskorte, N., 2013. Representational dynamics of object vision: the first 1000 ms. *J. Vis.* 13, 1–19. <http://dx.doi.org/10.1167/13.10.1.doi>.
- Cauchoix, M., Barragan-Jason, G., Serre, T., Barbeau, E.J., 2014. The neural dynamics of face detection in the wild revealed by MVPA. *J. Neurosci.* 34, 846–854. <http://dx.doi.org/10.1523/JNEUROSCI.3030-13.2014>.
- Cichy, R.M., Pantazis, D., Oliva, A., 2014. Resolving human object recognition in space and time. *Nat. Neurosci.* 17, 455–462. <http://dx.doi.org/10.1038/nn.3635>.
- Cromer, J.A., Roy, J.E., Miller, E.K., 2010. Representation of multiple, independent categories in the primate prefrontal cortex. *Neuron* 66, 796–807. <http://dx.doi.org/10.1016/j.neuron.2010.05.005>.
- Crouzet, S.M., Thorpe, S.J., 2011. Low-level cues and ultra-fast face detection. *Front. Psychol.* 2, 342. <http://dx.doi.org/10.3389/fpsyg.2011.00342>.



- Crowe, D.A., Goodwin, S.J., Blackman, R.K., Sakellaridi, S., Sponheim, S.R., MacDonald, A.W., Chafee, M.V., 2013. Prefrontal neurons transmit signals to parietal neurons that reflect executive control of cognition. *Nat. Neurosci.* 16, 1484–1491. <http://dx.doi.org/10.1038/nn.3509>.
- Dale, A.M., Fischl, B.R., Sereno, M.I., 1999. Cortical surface-based analysis. I. Segmentation and Surface Reconstruction. *Neuroimage* 9, 179–194. <http://dx.doi.org/10.1006/nimg.1998.0396>.
- Dale, A.M., Liu, A.K., Fischl, B.R., Buckner, R.L., Belliveau, J.W., Lewine, J.D., Halgren, E., 2000. Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron* 26, 55–67.
- DiCarlo, J.J., Zoccolan, D., Rust, N.C., 2012. How does the brain solve visual object recognition? *Neuron* 73, 415–434. <http://dx.doi.org/10.1016/j.neuron.2012.01.010>.
- Donner, T.H., Siegel, M., Fries, P., Engel, A.K., 2009. Buildup of choice-predictive activity in human motor cortex during perceptual decision making. *Curr. Biol.* 19, 1581–1585. <http://dx.doi.org/10.1016/j.cub.2009.07.066>.
- Ehinger, B.V., Fischer, P., Gert, A.L., Kaufhold, L., Weber, F., Pipa, G., König, P., 2014. Kinesthetic and vestibular information modulate alpha activity during spatial navigation: a mobile EEG study. *Front. Hum. Neurosci.* 8, 1–12. <http://dx.doi.org/10.3389/fnhum.2014.00071>.
- Ehinger, B.V., König, P., Ossandon, J.P., 2015. Predictions of visual content across eye movements and their modulation by inferred information. *J. Neurosci.* 35, 7403–7413. <http://dx.doi.org/10.1523/JNEUROSCI.5114-14.2015>.
- Fischl, B.R., Sereno, M.I., Tootell, R.B., Dale, A.M., 1999. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum. Brain Mapp.* 8, 272–284.
- Fitzgerald, J.K., Swaminathan, S.K., Freedman, D.J., 2012. Visual categorization and the parietal cortex. *Front. Integr. Neurosci.* 6, 18. <http://dx.doi.org/10.3389/fnint.2012.00018>.
- Folstein, J.R., Gauthier, I., Palmeri, T.J., 2012a. How category learning affects object discrimination: not all morphspaces stretch alike. *J. Exp. Psychol. Mem. Cogn.* 38, 4.
- Folstein, J.R., Palmeri, T.J., Gauthier, I., 2012b. Category learning increases discriminability of relevant object dimensions in visual cortex. *Cereb. Cortex* <http://dx.doi.org/10.1093/cercor/bhs067>.
- Folstein, J.R., Palmeri, T.J., Van Gulick, A.E., Gauthier, I., 2015. Category learning stretches neural representations in visual cortex. *Curr. Dir. Psychol. Sci.* 24, 17–23. <http://dx.doi.org/10.1177/0963721414550707>.
- Freedman, D.J., Riesenhuber, M., Poggio, T., Miller, E.K., 2003. A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *J. Neurosci.* 23, 5235–5246.
- Freedman, D.J., Riesenhuber, M., Poggio, T., Miller, E.K., 2006. Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. *Cereb. Cortex* 16, 1631–1644. <http://dx.doi.org/10.1093/cercor/bhj100>.
- Gauthier, I., Skudlarski, P., Gore, J.C., Anderson, A.W., 2000. Expertise for cars and birds recruits brain areas involved in face recognition. *Nat. Neurosci.* 3, 191–197. <http://dx.doi.org/10.1038/72140>.
- Gillebert, C.R., Op De Beeck, H.P., Panis, S., Wagemans, J., 2009. Subordinate categorization enhances the neural selectivity in human object-selective cortex for fine shape differences. *J. Cogn. Neurosci.* 21, 1054–1064. <http://dx.doi.org/10.1162/jocn.2009.21089>.
- Goldstone, R.L., Steyvers, M., Larimer, K., 1996. Categorical perception of novel dimensions. Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society, pp. 243–248. <http://dx.doi.org/10.1080/713756735>.
- Goldstone, R.L., Lippa, Y., Shiffrin, R.M., 2001. Altering object representations through category learning. *Cognition* 78, 27–43.
- Goodwin, S., Blackman, R., 2012. Executive control over cognition: stronger and earlier rule-based modulation of spatial category signals in prefrontal cortex relative to parietal cortex. *J. Neurosci.* 32, 3499–3515. <http://dx.doi.org/10.1523/JNEUROSCI.3585-11.2012>.
- Grill-Spector, K., Malach, R., 2001. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychol. (Amst)* 107, 293–321.
- Gureckis, T.M., Goldstone, R.L., 2008. The effect of the internal structure of categories on perception. *Proc. Thirtieth Annu. Conf. Cogn. Sci. Soc.*, pp. 1876–1881.
- Harris, A., Nakayama, K., 2007. Rapid face-selective adaptation of an early Extrastriate component in MEG. *Cereb. Cortex* 17, 63–70. <http://dx.doi.org/10.1093/cercor/bhj124>.
- Helie, S., Ashby, F.G., 2012. Learning and transfer of category knowledge in an indirect categorization task. *Psychol. Res.* 76, 292–303. <http://dx.doi.org/10.1007/s00426-011-0348-1>.
- Helie, S., Roeder, J.L., Ashby, F.G., 2010. Evidence for cortical automaticity in rule-based categorization. *J. Neurosci.* 30, 14225–14234. <http://dx.doi.org/10.1523/JNEUROSCI.2393-10.2010>.
- Huberle, E., Lutzenberger, W., 2013. Temporal properties of shape processing by event-related MEG adaptation. *Neuroimage* 67, 119–126. <http://dx.doi.org/10.1016/j.neuroimage.2012.10.070>.
- Hung, C.P., Kreiman, G., Poggio, T., DiCarlo, J.J., 2005. Fast readout of object identity from macaque inferior temporal cortex. *Science* 310 (80-), 863–866. <http://dx.doi.org/10.1126/science.1117593>.
- James, T.W., Gauthier, I., 2006. Repetition-induced changes in BOLD response reflect accumulation of neural activity. *Hum. Brain Mapp.* 27, 37–46. <http://dx.doi.org/10.1002/hbm.20165>.
- Jiang, X., Bradley, E., Rini, R.A., Zeffiro, T., Vanmeter, J., Riesenhuber, M., 2007. Categorization training results in shape- and category-selective human neural plasticity. *Neuron* 53, 891–903. <http://dx.doi.org/10.1016/j.neuron.2007.02.015>.
- Kietzmann, T.C., König, P., 2010. Perceptual learning of parametric face categories leads to the integration of high-level class-based information but not to high-level pop-out. *J. Vis.* 10, 1–14.
- Kirchner, H., Thorpe, S.J., 2006. Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited. *Vision Res.* 46, 1762–1776. <http://dx.doi.org/10.1016/j.visres.2005.10.002>.
- Kleiner, M., Brainard, D., Pelli, D., 2007. What's new in Psychtoolbox-3? *Perception* 36, 14. [http://dx.doi.org/10.1016/0028-3932\(83\)90075-1](http://dx.doi.org/10.1016/0028-3932(83)90075-1).
- Krekelberg, B., Boynton, G.M., van Wezel, R.J., 2006. Adaptation: from single cells to BOLD signals. *Trends Neurosci.* 29, 250–256. <http://dx.doi.org/10.1016/j.tins.2006.02.008>.
- Kriegeskorte, N., Mur, M., Ruff, D.A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., Bandettini, P.A., 2008. Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* 60, 1–16.
- Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S.F., Baker, C.I., 2009. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* 12, 535–540. <http://dx.doi.org/10.1167/8.6.88>.
- Li, S., Ostwald, D., Giese, M., Kourtzi, Z., 2007. Flexible coding for categorical decisions in the human brain. *J. Neurosci.* 27, 12321–12330. <http://dx.doi.org/10.1523/JNEUROSCI.3795-07.2007>.
- Liu, J., Harris, A., Kanwisher, N., 2002. Stages of processing in face perception: an MEG study. *Nat. Neurosci.* 5, 910–916. <http://dx.doi.org/10.1038/nn909>.
- Liu, H., Agam, Y., Madsen, J.R., Kreiman, G., 2009. Timing, timing, timing: fast decoding of object information from intracranial field potentials in human visual cortex. *Neuron* 62, 281–290. <http://dx.doi.org/10.1016/j.neuron.2009.02.025>.
- Liu, N., Kriegeskorte, N., Mur, M., Hadj-Bouziane, F., Luh, W.-M., Tootell, R.B.H., Ungerleider, L.G., 2013. Intrinsic structure of visual exemplar and category representations in macaque brain. *J. Neurosci.* 33, 11346–11360. <http://dx.doi.org/10.1523/JNEUROSCI.4180-12.2013>.
- Makeig, S., Westerfield, M., Jung, T.P., Enghoff, S., Townsend, J., Courchesne, E., Sejnowski, T.J., 2002. Dynamic brain sources of visual evoked responses. *Science* 295 (80-), 690–694. <http://dx.doi.org/10.1126/science.1066168>.
- Marinkovic, K., Dhond, R.P., Dale, A.M., Glessner, M., Carr, V., Halgren, E., 2003. Spatiotemporal dynamics of modality-specific and supramodal word processing. *Neuron* 38, 487–497.
- Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164, 177–190. <http://dx.doi.org/10.1016/j.jneumeth.2007.03.024>.
- McKee, J.L., Riesenhuber, M., Miller, E.K., Freedman, D.J., 2014. Task dependence of visual and category representations in prefrontal and inferior temporal cortices. *J. Neurosci.* 34, 16065–16075. <http://dx.doi.org/10.1523/JNEUROSCI.1660-14.2014>.
- Minamimoto, T., Saunders, R.C., Richmond, B.J., 2010. Monkeys quickly learn and generalize visual categories without lateral prefrontal cortex. *Neuron* 66, 501–507. <http://dx.doi.org/10.1016/j.neuron.2010.04.010>.
- Muhammad, R., Wallis, J.D., Miller, E.K., 2006. A comparison of abstract rules in the prefrontal cortex, premotor cortex, inferior temporal cortex, and striatum. *J. Cogn. Neurosci.* 18, 974–989. <http://dx.doi.org/10.1162/jocn.2006.18.6.974>.
- Mur, M., Ruff, D.A., Bodurka, J., De Weerd, P., Bandettini, P.A., Kriegeskorte, N., 2012. Categorical, yet graded - single-image activation profiles of human category-selective cortical regions. *J. Neurosci.* 32, 8649–8662. <http://dx.doi.org/10.1523/JNEUROSCI.2334-11.2012>.
- Nosofsky, R.M., 1991. Tests of an exemplar model for relating perceptual classification and recognition memory. *J. Exp. Psychol. Hum. Percept. Perform.* 17, 3–27. <http://dx.doi.org/10.1037/0096-1523.17.1.3>.
- Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011, 1–10. <http://dx.doi.org/10.1155/2011/156869>.
- Pascual-Marqui, R., 2002. Standardized low resolution brain electromagnetic tomography (sLORETA): technical details. *Methods Find. Exp. Clin. Pharmacol.* 24, 5–12.
- Petit, J.-P., Midgley, K.J., Holcomb, P.J., Grainger, J., 2006. On the time course of letter perception: a masked priming ERP investigation. *Psychon. Bull. Rev.* 13, 674–681.
- Philiastides, M.G., Sajda, P., 2006. Temporal characterization of the neural correlates of perceptual decision making in the human brain. *Cereb. Cortex* 16, 509–518. <http://dx.doi.org/10.1093/cercor/bhi130>.
- Plöchl, M., Ossandon, J.P., König, P., 2012. Combining EEG and eye tracking: identification, characterization, and correction of eye movement artifacts in electroencephalographic data. *Front. Hum. Neurosci.* 6, 278. <http://dx.doi.org/10.3389/fnhum.2012.00278>.
- Pockett, S., Whalen, S., McPhail, A.V.H., Freeman, W.J., 2007. Topography, independent component analysis and dipole source analysis of movement related potentials. *Cogn. Neurodyn.* 1, 327–340. <http://dx.doi.org/10.1007/s11571-007-9024-y>.
- Reed, S.K., Friedman, M.P., 1973. Perceptual vs conceptual categorization. *Mem. Cognit.* 1, 157–163. <http://dx.doi.org/10.3758/BF03198087>.
- Riesenhuber, M., Poggio, T., 2002. How Visual Cortex Recognizes Objects: The Tale of the Standard Model. In: Chapula, L.M., J.S.W. (Eds.), *The Visual Neurosciences*. MIT Press, pp. Cambridge, MA, pp. 1640–1653.
- Rossion, B., Caharel, S., 2011. ERP evidence for the speed of face categorization in the human brain: disentangling the contribution of low-level visual cues from face perception. *Vision Res.* 51, 1297–1311. <http://dx.doi.org/10.1016/j.visres.2011.04.003>.
- Rossion, B., Jacques, C., 2008. Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *Neuroimage* 39, 1959–1979. <http://dx.doi.org/10.1016/j.neuroimage.2007.10.011>.
- Roy, J.E., Riesenhuber, M., Poggio, T., Miller, E.K., 2010. Prefrontal cortex activity during flexible categorization. *J. Neurosci.* 30, 8519–8528. <http://dx.doi.org/10.1523/JNEUROSCI.4837-09.2010>.
- Scholl, C., Jiang, X., Martin, J., Riesenhuber, M., 2014. Time course of shape and category selectivity revealed by EEG rapid adaptation. *J. Cogn. Neurosci.* 26, 408–421. <http://dx.doi.org/10.1162/jocn>.
- Segaert, K., Weber, K., de Lange, F.P., Petersson, K.M., Hagoort, P., 2013. The suppression of repetition enhancement: a review of fMRI studies. *Neuropsychologia* <http://dx.doi.org/10.1016/j.neuropsychologia.2012.11.006>.
- Seger, C.A., Miller, E.K., 2010. Category learning in the brain. *Annu. Rev. Neurosci.* 33, 203–219. <http://dx.doi.org/10.1146/annurev.neuro.051508.135546>.

- Serre, T., Oliva, A., Poggio, T., 2007. A feedforward architecture accounts for rapid categorization. *Proc. Natl. Acad. Sci. U. S. A.* 104, 6424–6429. <http://dx.doi.org/10.1073/pnas.0700622104>.
- Sigala, N., Logothetis, N.K., 2002. Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature* 415, 318–320. <http://dx.doi.org/10.1038/415318a>.
- Sigala, N., Gabbiani, F., Logothetis, N.K., 2002. Visual categorization and object representation in monkeys and humans. *J. Cogn. Neurosci.* 14, 187–198. <http://dx.doi.org/10.1162/089892902317236830>.
- Stolk, A., Todorovic, A., Schoffelen, J.-M., Oostenveld, R., 2013. Online and offline tools for head movement compensation in MEG. *Neuroimage* 68, 39–48. <http://dx.doi.org/10.1016/j.neuroimage.2012.11.047>.
- Sugase, Y., Yamane, S., Ueno, S., Kawano, K., 1999. Global and fine information coded by single neurons in the temporal visual cortex. *Nature* 400, 869–873. <http://dx.doi.org/10.1038/23703>.
- Swaminathan, S.K., Freedman, D.J., 2012. Preferential encoding of visual categories in parietal cortex compared with prefrontal cortex. *Nat. Neurosci.* 15, 315–320. <http://dx.doi.org/10.1038/nn.3016>.
- Tadel, F., Baillet, S., Mosher, J.C., Pantazis, D., Leahy, R.M., 2011. Brainstorm: a user-friendly application for MEG/EEG analysis. *Comput. Intell. Neurosci.* 2011, 879716. <http://dx.doi.org/10.1155/2011/879716>.
- Tanaka, J.W., Curran, T., 2001. A neural basis for expert object recognition. *Psychol. Sci. A J. Am. Psychol. Soc./APS* 12, 43–47. <http://dx.doi.org/10.1111/1467-9280.00308>.
- Thierry, G., Martin, C.D., Downing, P., Pegna, A.J., 2007. Controlling for interstimulus perceptual variance abolishes N170 face selectivity. *Nat. Neurosci.* 10, 505–511. <http://dx.doi.org/10.1038/nn1864>.
- Tsai, A.C., Jung, T.-P., Chien, V.S.C., Savostyanov, A.N., Makeig, S., 2014. Cortical surface alignment in multi-subject spatiotemporal independent EEG source imaging. *Neuroimage* 87, 297–310. <http://dx.doi.org/10.1016/j.neuroimage.2013.09.045>.
- Underwood, G., Henderson, J.M., Hollingworth, A., 1998. Eye guidance in reading and scene perception. In: Underwood, G. (Ed.), *Eye Movements during Scene Viewing: An Overview*. Elsevier, pp. 269–293.
- van der Linden, M., van Turenout, M., Indefrey, P., 2010. Formation of category representations in Superior temporal sulcus. *J. Cogn. Neurosci.* 22, 1270–1282. <http://dx.doi.org/10.1162/jocn.2009.21270>.
- van der Linden, M., Wegman, J., Fernández, G., 2013. Task- and experience-dependent cortical selectivity to features informative for categorization. *J. Cogn. Neurosci.* 1–15. <http://dx.doi.org/10.1162/jocn>.
- VanRullen, R., 2011. Four common conceptual fallacies in mapping the time course of recognition. *Front. Psychol.* 2, 365. <http://dx.doi.org/10.3389/fpsyg.2011.00365>.
- Vizioli, L., Rousselet, G.A., Caldara, R., 2010. Neural repetition suppression to identity is abolished by other-race faces. *Proc. Natl. Acad. Sci. U. S. A.* 107, 20081–20086. <http://dx.doi.org/10.1073/pnas.1005751107>.
- Wichmann, F., Drewes, J., Rosas, P., Gegenfurtner, K., 2010. Animal detection in natural scenes: critical features revisited. *J. Vis.* 10, 1–27. <http://dx.doi.org/10.1167/10.4.6.Introduction>.
- Zago, L., Fenske, M.J., Aminoff, E., Bar, M., 2005. The rise and fall of priming: how visual exposure shapes cortical representations of objects. *Cereb. Cortex* 15, 1655–1665. <http://dx.doi.org/10.1093/cercor/bhi060>.
- Zimmer, M., Kovács, G., 2011. Position specificity of adaptation-related face aftereffects. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 586–595. <http://dx.doi.org/10.1098/rstb.2010.0265>.