Cerebral Cortex, 2017; 1–12

doi: 10.1093/cercor/bhx039
Original Article

ORIGINAL ARTICLE

Long Time No See: Enduring Behavioral and Neuronal Changes in Perceptual Learning of Motion Trajectories 3 Years After Training

Sebastian M. Frank[1], Mark W. Greenlee[2], and Peter U. Tse[1]

[1]Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH 03755, USA, and
[2]Institute for Experimental Psychology, University of Regensburg, Regensburg 93053, Germany

Address correspondence to Sebastian M. Frank. Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH 03755, USA.
Email: sebastian.m.frank@dartmouth.edu

Abstract

Here, we report on the long-term stability of changes in behavior and brain activity following perceptual learning of conjunctions of simple motion features. Participants were trained for 3 weeks on a visual search task involving the detection of a dot moving in a "v"-shaped target trajectory among inverted "v"-shaped distractor trajectories. The first and last training sessions were carried out during functional magnetic resonance imaging (fMRI). Learning stability was again examined behaviorally and using fMRI 3 years after the end of training. Results show that acquired behavioral improvements were remarkably stable over time and that these changes were specific to trained target and distractor trajectories. A similar pattern was observed on the neuronal level, when the representation of target and distractor stimuli was examined in early retinotopic visual cortex (V1–V3): training enhanced activity for the target relative to the surrounding distractors in the search array and this enhancement persisted after 3 years. However, exchanging target and distractor trajectories abolished both neuronal and behavioral effects, suggesting that training-induced changes in stimulus representation are specific to trained stimulus identities.

Key words: area MT+ (V5), long-term stability, perceptual learning, early visual cortex, visual search

Introduction

Successful learning (e.g., learning symbols in an orthographic system, learning a new computer program, or learning how to drive a car) should establish long-term representations that outlast the period of training, and that would ideally last a lifetime. The acquired information should be readily accessible long after acquisition without any need for further practice. For perceptual learning (Gibson 1963; Sasaki et al. 2010), it is not uncommon to find such long-lasting learning effects. This has been shown for learning of low-level visual features, such as textures (Karni and Sagi 1993) or motion directions (Ball and Sekuler 1982; Watanabe et al. 2002), as well as for learning of more complex stimuli, such as feature conjunctions in visual search (Sireteanu and Rettenbach 1995; Frank et al. 2014b).

Perceptual improvements are associated with brain activity changes (e.g., Furmanski et al. 2004; Kourtzi et al. 2005; Sigman et al. 2005; Yotsumoto et al. 2008; Shibata et al. 2012; Frank et al. 2014b, 2016; Chen et al. 2015, 2016). If perceptual learning is highly stable over time, there should be neuronal signatures of these enduring learning effects. In particular, one might expect to find enduring changes in those brain areas or circuits where learned representations are maintained or processed more efficiently. However, only a few studies have addressed this question (Yotsumoto et al. 2008; Bi et al. 2014; Chen et al. 2015). In these studies, retesting during neuronal recording was performed a few weeks after the final training session. Results have been mixed: activity in task-relevant areas either reverted back to pre-training baseline levels (Yotsumoto et al. 2008), or

persisted in the trained neuronal activity patterns (Bi et al. 2014; Chen et al. 2015).

In our previous work, we have examined perceptual learning of complex motion trajectories in the context of a visual search task (see Frank et al. 2016; Reavis et al. 2016). In this task, participants had to detect the presence or absence of a dot cycling through a "v"-shaped motion trajectory, which served as the target, among dots at other locations cycling through inverted "v"-shaped trajectories, which served as distractors. The first and last training sessions were performed during functional magnetic resonance imaging (fMRI). Behavioral results showed that participants improved dramatically on this task over the course of 3 weeks of training. Functional imaging results suggested that learning was associated with an increase of activity in human area MT+ (V5) and retinotopic visual cortex (V1–V3, V3ab), whereas no learning effects were observed in parietal cortex. A second result of this study was that learning changed the representation of target and distractor stimuli in early visual cortex, such that more activity was observed for the target relative to the surrounding distractors after training (see Reavis et al. 2016).

In this study, we aimed to address the long-term stability of behavioral and neuronal learning effects in the visual search paradigm for motion trajectories described above. Three years after the end of our study, we were able to rescan 7 of our original trained participants. Thus, in addition to the original first and last training sessions, these participants took part in a retest session, while being scanned with fMRI. Furthermore, retest participants completed an additional fMRI scan with exchanged target and distractor trajectories, in order to measure whether behavioral and neuronal learning effects were specific to trained target and distractor identities.

Based on previous studies (e.g., Ball and Sekuler 1982; Karni and Sagi 1993; Sireteanu and Rettenbach 1995; Watanabe et al. 2002; Frank et al. 2014b), we predicted that we would find stable behavioral learning effects, even 3 years after the end of original training. Using univariate analysis, we hypothesized that training-induced activity increases during task performance (target present and absent conditions) would revert back to pre-training baseline levels during retest, as reported by Yotsumoto et al. (2008). Furthermore, we examined whether the training-induced target over distractor enhancement in target-present trials would also revert back to baseline levels or outlast the period of training.

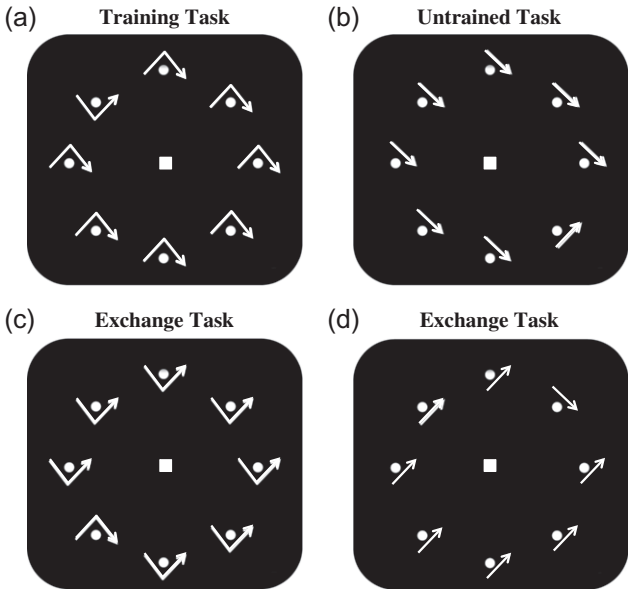

Figure 1. Search stimuli used for training, retest, exchange, and control conditions. (*a*) Training task. The search array consisted of 8 small white dots, arranged circularly in the periphery (10° from central fixation). Participants covertly searched the array and indicated with a buttonpress the presence or absence of a target moving in a "v"-shaped trajectory. Distractor dots moved in inverted "v"-shaped trajectories. White arrows, showing motion trajectories, are inserted for illustrative purposes only and were not present in the actual stimulus display. Participants received training on this task and performed the first and last training sessions during fMRI. Three years after the end of original training, participants performed the task once again during fMRI. (*b*) Untrained control task. For half of the participants the target moved rightwardly upward and distractors moved rightwardly downward (vice versa in the other half). This task was designed to elicit target pop-out without training and participants performed the task only during scanning sessions. (*c*) Exchange task. Same as (*a*) but with exchanged target and distractor identities. The target moved in an inverted "v"-shaped trajectory and distractors moved in "v"-shaped trajectories. This task was performed in a separate fMRI session after retest and was designed to measure the specificity of learning to trained target and distractor trajectories. (*d*) Exchange task. Same as (*b*) but with exchanged target and distractor identities. For half of the participants the target moved rightwardly downward, whereas distractors moved rightwardly upward (vice versa in the other half). This task was performed in the same fMRI session as the exchange training task. Since the untrained task elicits target pop-out without training, we predicted performance transfer to the exchange condition.

Materials and Methods

General Design

Participants performed 12 training sessions on separate days in a visual search task for complex motion trajectories (Fig. 1*a*). The first and last training sessions were carried out during fMRI. All other training sessions were performed in a psychophysics testing room. On average, participants completed the training sessions in 3 weeks. Participants carried out the original visual search task in a third fMRI "retest" session 3 years after the post-training fMRI scan. After the retest, participants performed a fourth fMRI session with exchanged target and distractor trajectories (Fig. 1*c*). During each fMRI session an untrained pop-out control task (Fig. 1*b*,*d*) was performed as well.

Participants

Participants were recruited from a pool of 19 subjects from the University of Regensburg community who had completed training in the learning task for a previous study (see Frank et al. 2016; Reavis et al. 2016). Three years after the end of their original training, 7 participants (4 females, mean age at training start = 28 ± 7 years) could be located and recruited for a retest. Participants were recruited based on availability and willingness to participate. Participants gave informed written consent. The study was approved by the local ethics committee at the University of Regensburg.

Training Task

Participants were trained on a visual search task involving covert search for the presence or absence of a target among an array of distractors. Details of the task have been described elsewhere (Frank et al. 2016). In brief, there were 8 white dots arranged radially at a distance of 10° from central fixation on a black background. The target dot cycled through a "v"-shaped motion trajectory. Distractor dots cycled through inverted "v"-shaped trajectories (Fig. 1*a*). Each dot had a different

starting position within the trajectory in order to avoid any perception of global motion patterns. No dot ever overlapped with, or crossed the path of, another dot in the array. Also, the spatial position of each dot was slightly jittered on each trial to prevent the formation of after-images. Dots moved at a speed of 7.5°/s. After completion of a motion trajectory (duration of full trajectory cycle = 600 ms, followed by blank interval of 217 ms) dots again cycled through the trajectory. Dots continued to cycle until the trial ended. Each trial was 4 s long. Participants were requested to respond as quickly and as accurately as possible after trial onset. Feedback (green or red fixation spot for correct or incorrect responses, respectively) was provided at the end of each trial.

Overall, 160 trials per session were performed (half of which were target-present trials). In the fMRI sessions, trials were split across 5 runs with 32 trials each. We used a fast event-related fMRI-design with a jittered interstimulus interval between 4, 6, and 8 s (counterbalanced for each run). Half of the trials in each fMRI run were target-present trials. Trial order was random. Each behavioral training session took 10–15 min. The fMRI session took 30 min due to longer intertrial intervals. In each session, the target was presented equally often (10×) at each stimulus location (2× each location in each fMRI run).

Untrained Task

The untrained task (Fig. 1*b*) was designed to serve as a target pop-out control condition where the target could be easily detected without any training. In this case, each dot moved only in a single direction and half of the participants searched for a target moving up and to the right among distractors moving down and to the right (vice versa for the other half). Other parameters were identical with the training task. The untrained task was only performed during fMRI.

Exchange Tasks

The exchange tasks (Fig. 1*c*,*d*) were identical with the training and untrained tasks, except that target and distractor trajectories were interchanged. Exchange tasks were performed in a separate fMRI session after retesting.

Functional Localizers for ROIs

Separate localizer scans were performed for motion-sensitive area MT+, retinotopic visual cortex (V1–V3, V3ab), and the representation of stimulus locations in retinotopic visual cortex. In each of these localizer scans, participants were requested to maintain central fixation and to perform a speeded dimming detection task at the central fixation spot. Regions of interest (ROIs) were defined on each participant's inflated left and right cortical hemispheres at a threshold of $P < 0.001$ (false-discovery rate corrected, or uncorrected, if activity could not be identified at the expected location using a corrected threshold).

Human Area MT+ (V5)

A standard visual motion localizer was performed in order to define human area MT+ (see Frank et al. 2014a, 2014b, 2016). In this localizer, 200 white dots were presented on a black background. During visual motion blocks, dots moved with a speed of 15°/s in one of 12 translational directions for 1 s each. Dots did not spatially overlap and had a limited random lifetime between 167 and 333 ms. Blocks with visual motion alternated with baseline blocks during which all dots remained static. Each block lasted 12 s. One run was performed (9.6 min). Area MT+ was defined by contrasting the motion and no-motion conditions (see Fig. 3, for the location of MT+ in an example participant).

For the purpose of previous projects in our laboratory (e.g., Frank et al. 2014a), participants in the retest group performed localizer scans in order to differentiate area MST from area MT within the MT+ complex, using the approach suggested by Huk et al. (2002). Specifically, MST was defined by means of more pronounced activation during ipsilateral visual motion stimulation and was located in the anterior part of the MT+ complex, whereas MT was defined as the posterior section of MT+ that did not respond to ipsilateral stimulation (see inset on the right in Supplementary Fig. 5, and Frank et al. 2014a, for details). The subregion MT of the combined MT+ complex was used for the analysis of target relative to distractor activity (see below), because MT, in contrast to MST, has a fairly clearcut contralateral hemispheric representation (Huk et al. 2002), which is necessary for this analysis.

Retinotopic Visual Cortex

For the definition of retinotopic visual cortex, phase-encoded retinotopic mapping was performed (see DeYoe et al. 1996; Engel et al. 1997). Therefore, a bow-tie shaped double-wedge, flickering in different colors, rotated in clockwise (run 1) and counterclockwise directions (run 2) across 18 locations on the screen (3 s each location) for 12 cycles. Each run was 10.8 min long. This localizer revealed the borders between visual areas V1, V2, V3, and V3ab; V3a and V3b could not be reliably separated in each participant using the current localizer data set and therefore both regions were combined into a single V3ab ROI. As in our previous work (see Frank et al. 2016), we combined visual areas V1, V2, and V3 into a single mask for the definition of stimulus representations in early retinotopic visual cortex (see below). V3ab was treated as a separate ROI (see Supplementary Fig. 1, for the location of V3ab in an example participant), following recent reports of perceptual learning effects at this site (see Shibata et al. 2012; Chen et al. 2015, 2016).

Stimulus Representation in Retinotopic Visual Cortex

In order to define the representations of the stimuli in retinotopic visual cortex, circular checkerboard patterns flickering in different colors were presented at each of the 8 dot locations, in separate scan blocks. Each stimulation block was followed by a blank baseline and all blocks were 12 s long. Two runs (9.6 min each) were performed. ROIs were defined by contrasting activity during stimulation at a retinotopic location with the activity combined across stimulation at all other locations. Representations of stimulus locations spanned across different subregions of retinotopic visual cortex (V1–V3, see Fig. 4 for ROI representations in left visual cortex of an example participant). Since it was not possible, using the current localizer data set, to define the representation of each stimulus location separately in visual areas V1, V2, and V3, we combined V1–V3 and used it as an inclusion mask for the definition of retinotopic stimulus representations. Therefore, activation in the stimulus localizer experiment was only used to define the representation of a stimulus location, if the activation was located within the V1–V3 mask. A downside of this approach is that results cannot be interpreted separately for different subregions of early visual

cortex. In one participant, 2 stimulus locations could not be defined inside the mask and therefore, only the remaining 6 locations were analyzed for this subject.

Anatomical Localizer for PPC

Following our previous publication (Frank et al. 2016), we also included posterior parietal cortex (PPC) in the analysis. PPC was defined by using the automated parcellation of each participant's high-resolution anatomical scan and consisted of the conjunction between parietal gyrus and sulcus (Desikan et al. 2006) (see Supplementary Fig. 1, for the location in an example participant).

Stimulus Generation and Presentation

Stimuli were generated using Psychtoolbox (Brainard 1997; Pelli 1997) running in MATLAB (Mathworks). In the scanner, stimuli were projected onto a circular screen, located at the end of the scanner bore. Participants viewed the screen (viewing distance = 63 cm, screen diameter: 30°) with a head-coil mounted mirror. For behavioral training sessions, stimuli were presented on a computer screen (34° × 26°), located 63 cm in front of the participants. Stimulus appearance in the visual search task was adjusted to match between behavioral and scanner settings by matching the stimulus size (see above) and by using similar luminance levels of stimuli and background (Michelson contrast: Psychophysics = 0.99, MRI = 0.98; luminance of black background: Psychophysics = 0.16 cd/m^2, MRI = 1.7 cd/m^2; luminance of white stimuli: Psychophysics = 185.44 cd/m^2, MRI = 193 cd/m^2).

Scanning Parameters

Imaging data were collected on a 3-Tesla Allegra scanner (Siemens) using a one-channel head coil. A high-resolution anatomical scan of each participant's brain was acquired in the first scanning session with a magnetization prepared rapid gradient echo sequence (time-to-repeat [TR] = 2.25 s, time-to-echo [TE] = 2.6 ms, flip-angle [FA] = 9°, voxel-size = 1 × 1 × 1 mm, 160 sagittal slices, no interslice gap, field of view [FOV] = 240 × 256 mm). All functional MRI data were collected with a standard T2*-weighted echoplanar imaging sequence (TR = 2 s, TE = 30 ms, FA = 90°, voxel-size = 3 × 3 × 3 mm, 34 transverse slices, interslice gap = 0.5 mm, FOV = 192 × 192 mm).

Eye-Tracking

Participants were requested to maintain central fixation and to perform the search task using their covert attention. However, it is possible that fixation quality changed across training sessions (e.g., better fixation in later training sessions). In order to exclude the possibility that changes in fixation quality occurred together with behavioral improvements, participants performed behavioral training sessions during eye-tracking. A video-based eye-tracking system was used (Cambridge Research Systems) that sampled the horizontal and vertical position of the right eye with a frequency of 250 Hz.

Data Analysis

Behavior

The search task required participants to perform as rapidly and as accurately as possible. Therefore, training-induced behavioral improvements in the learning task are to be expected for both reaction time and accuracy. For statistical assessment, following our previous work (Frank et al. 2016), we combined reaction time and accuracy into a learning score (= "weighted reaction time" or "inverse efficiency score"; see Townsend and Ashby 1978), which we will refer to as the "learning index" henceforth. Therefore, for each session and task, median reaction time (across target present and absent conditions, in s) was divided by accuracy (hits + correct rejections, in proportions of one). For the training task, the index was also log-transformed in order to correct for distortions of the learning index by low accuracy values in the pre-training and exchange sessions (see Bruyer and Brysbaert 2011). The lower the learning index, the faster and better participants performed.

Previous studies reported correlations between behavioral and neuronal learning effects, such that better learning was associated with more pronounced changes in brain activity (e.g., Furmanski et al. 2004; Kourtzi et al. 2005; Sigman et al. 2005; Yotsumoto et al. 2008; Shibata et al. 2012; Bi et al. 2014; Frank et al. 2014b, 2016; Chen et al. 2015, 2016). Therefore, we computed each participant's behavioral learning rate on the trajectory training task by performing a linear fit across learning indices in the 12 training sessions for each participant (including the first and last sessions, which were performed during fMRI). The slopes of these linear regressions, quantifying the rate of improvement by training, were correlated with brain activity changes between the pre-training and post-training fMRI sessions.

The slope represents a quantification of learning rate, which includes the amount of improvement from pre-training to post-training as well as the variation of improvement across behavioral training sessions. Therefore, it provides a more specific estimate of participants' improvement across training compared with the difference between post-training and pre-training sessions only.

Eye-Tracking

Eye-tracking data were analyzed as in our previous work (see Frank et al. 2016). For each participant, training session, trial, and time-point within a trial, the Eucledian distance of the position of the eye from central fixation was computed and averaged across all time-points of a trial until trial end, or, if participants responded during the trial, until buttonpress. This trial-wise deviation from fixation was then averaged across all trials for each session. Following the computation of perceptual learning rate (see above), a linear fit was performed to eye deviation scores across all behavioral training sessions for each participant. The slopes of these fits, across participants, were compared with zero (corresponding to no change in eye deviation from fixation across training sessions), using a Wilcoxon signed-rank test as implemented in MATLAB.

Imaging

Imaging data were analyzed using Freesurfer and the FSFAST toolbox (Martinos Center for Biomedical Imaging). High-resolution anatomical scans were reconstructed and inflated (Dale et al. 1999; Fischl et al. 1999). Functional images were motion-corrected, coregistered to the high-resolution anatomical scan that was collected on the pre-training session, smoothed with a 3D Gaussian kernel (full-width-at-half-maximum = 5 mm), and intensity-normalized. The coregistration of functional imaging data to the anatomical scan was carefully inspected and manually corrected, if necessary. Preprocessed fMRI data were analyzed

with a general linear model (GLM) approach. The blood oxygenation level dependent (BOLD) response was modeled using the SPM hemodynamical response function. Each GLM-model contained a linear scanner drift predictor and motion-correction parameters as regressors of no interest.

In order to quantify the neuronal effects of learning and the stability of training-induced brain activity changes over time, 2 fMRI analyses were performed, as described previously (Frank et al. 2016; Reavis et al. 2016). These analyses were performed for the training task and the untrained pop-out control task and for each scanning session (that is, pre-training, post-training, retest, and target/distractor exchange sessions).

In the first analysis, brain activation (target-present and target-absent conditions) was compared between the different fMRI sessions. The GLM-model contained 2 regressors of interest for target present and absent trials. Only correct trials were included in these regressors (that is, only hits and correct rejections were modeled). Incorrect target present and absent trials (misses and false alarms) were modeled with a separate regressor of no interest. For each trial, the BOLD response was only modeled until participants responded. Since the analysis is restricted to correct trials and the period of search, changes in the BOLD response due to differences in participants' performance between sessions are minimized. Moreover, additional control analyses with matched number of trials across sessions were performed (see below) to exclude the possibility that differences in the number of correct trials between sessions influenced measures of brain activity. Therefore, BOLD-signal activity for each session can be interpreted as corresponding to stimulus processing. For the untrained pop-out control task, 2 regressors of interest for target present and absent trials were constructed. Again, the BOLD-signal response was only modeled until participants responded. Since performance was, as predicted, close to ceiling in the control task (see Fig. 2*d*), the use of a separate regressor for incorrect trials was unnecessary. As in our previous study (Frank et al. 2016), 4 ROIs were analyzed: MT+, the representation of the stimuli in early retinotopic visual cortex (all stimulus locations combined in V1–V3, which we will refer to in the following as TRL for "trained retinotopic locations" in early visual cortex), area V3ab, and PPC. For each ROI, BOLD % signal change for each regressor in the trained and untrained tasks was computed relative to implicit baseline (consisting of blanks).

In a second fMRI analysis, the representation of the target among the distractors was examined in early retinotopic visual cortex (V1–V3). To this end, the analysis was carried out on correct target-present trials only. The GLM-model contained 8 regressors of interest for target presentation at each of 8 locations in the search array (see Fig. 1). A ninth regressor of no interest covered all other trials (that is, target absent and incorrect target-present trials). Each trial was modeled until participants responded. BOLD % signal change from implicit baseline (consisting of blanks) was computed for each of the 8 retinotopic stimulus locations. For each stimulus location, activation for a distractor present in that retinotopic representation was subtracted from activation for the target present in the same retinotopic representation. The resulting difference in BOLD-signal represents the activation gain when the target is shown within a given retinotopic location versus when a distractor is presented within the same retinotopic location (implying that the target is presented at a different location). Stimulus locations on the upper and lower vertical meridian are represented in both left and right hemispheres and target relative to distractor

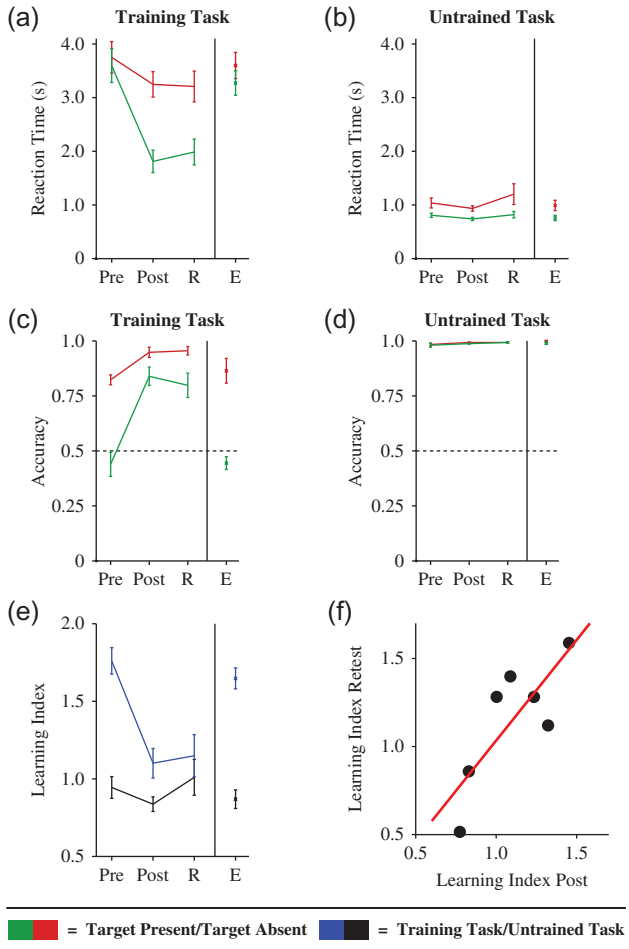

Figure 2. Behavioral performance in pre-training (Pre), post-training (Post), retest (R), and target/distractor exchange (E) sessions. Shown are average data (with standard error of the mean, SE) for 7 participants who completed all scanning sessions. Green line = target-present condition, red line = target-absent condition. (*a*) Median reaction time (in s) in training task across sessions. (*b*) Same as (*a*) but for untrained pop-out control task. (*c*) Accuracy (in proportion of one) in training task across sessions. The dashed line indicates chance level. (*d*) Same as (*c*) but for untrained pop-out control task. (*e*) Performance, quantified as learning index (= reaction time weighted by accuracy, see Materials and Methods for details), in the training task (blue) and the untrained pop-out control task (black). Please note that training task learning-indices were log-transformed and should not be directly compared with learning indices in the pop-out control task. (*f*) Correlation between performance, measured as learning index, on the final training session (*x*-axis) and the retesting session (*y*-axis). Participants with faster and better performance on the final training session (= lower values on *x*-axis) transferred this benefit to the retesting session (= lower values on *y*-axis). Each dot represents the result from a different participant.

signals in these ROIs were merged between hemispheres, before results were averaged across all stimulus location ROIs.

A similar analysis was performed for the posterior part of the MT+ complex (corresponding to human area MT). To this end, only correct trials during which the target was presented in the left and right hemifields were used. That is, trials with target presentation at 12 and 6 o'clock were excluded. The 2 conditions (target in left vs. right visual hemifield) were analyzed in left and right MT. Activity during target present in the contralateral hemifield was compared with activity during target presentation in the ipsilateral hemifield (that is, with activity when only distractors were shown in the represented contralateral site of the visual field).

Activations were statistically analyzed as follows: For each fMRI analysis and ROI, a 2 × 3 repeated measures ANOVA with factors task (trained vs. untrained) and time-point (pre vs. post vs. retest) was performed. Of particular interest was a significant interaction between task and time-point because it indicated that activity between training and untrained task changed differently over time. Specifically, we predicted that activity in the training task would change over the course of the different testing sessions, as a result of learning and long-term consolidation, whereas no such change was predicted for the untrained pop-out control task. In order to determine whether activity in the training task changed over the course of the 3 scanning time-points (pre, post, retest), post hoc Wilcoxon signed-rank tests were performed between (1) post-training versus pre-training, (2) retest versus post-training, and (3) retest versus pre-training. Activity in the target and distractor exchange condition was compared with post-training and pre-training activity, again using Wilcoxon signed-rank tests. BOLD-signal activations in the training task were statistically evaluated relative to the untrained pop-out control task that was performed in the same session, in order to account for any BOLD-signal changes between sessions due to factors unrelated to learning (e.g., different levels of noise). Therefore, for each session, activation in the training task was subtracted from the pop-out control task and this activity difference score was submitted to Wilcoxon signed-rank tests as described above.

Since participants improve with training, there are more correct trials available for post-training compared with pre-training scans in the training task. On average, across participants and conditions, with standard deviation (SD): 89 ± 8% of all trials in post-training versus 63 ± 9% of all trials in pre-training were correct. Therefore, a control analysis was performed for which the same number of correct target present and absent trials as in the pre-training session was randomly selected from the pool of correct trials in the post-training session for each participant. Differences in the number of available trials did not appear to influence retest results in area MT+ (for the analysis of activation combined across target present and absent conditions, see Fig. 3), because, even though more trials were available for retest (88 ± 9%) compared with pre-training (see above) and exchange conditions (65 ± 7% of all trials), the BOLD signal in MT+ during retest reverted back to pre-training baseline and exchange condition levels. However, the training-induced enhancement of target relative to surrounding distractor stimuli in retinotopic visual cortex (Fig. 4) appeared to be stable over time, but could be influenced by differences in the number of target-present trials between sessions (pre-training: 44 ± 14%; post-training: 84 ± 11%; retest: 80 ± 15%; exchange: 44 ± 8%). Therefore, in a control analysis, the same number of correct target-present trials in the training task was randomly selected for each fMRI session (including target and distractor exchange), based on the minimum number of correct target-present trials in any of the scanning sessions, for each participant.

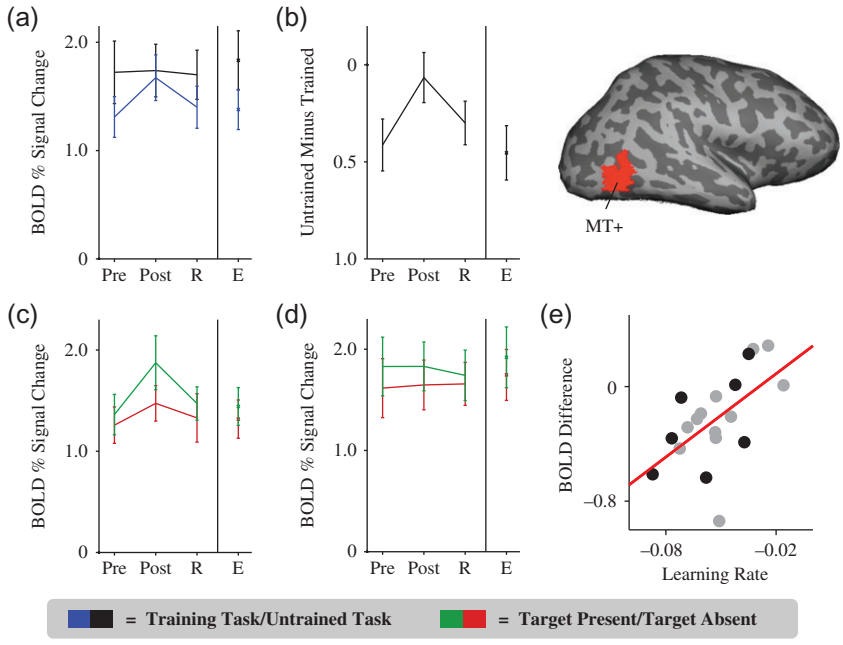

Figure 3. Activity changes (combined across target present and absent conditions) in human area MT+ (V5) across pre-training (Pre), post-training (Post), retest (R), and target/distractor exchange (E) sessions. Shown are average data (with SE) for 7 participants who performed all scanning sessions. The inset on the right depicts the location of MT+ on the right inflated hemisphere of an example participant. (*a*) Activations (in BOLD % signal change from zero = blank baseline) in the training task (blue line) and the untrained pop-out control task (black line). (*b*) Difference-score between training task and untrained pop-out control task, computed by subtracting activity in the trained task from activity in the untrained task, for each session. The closer this score was to zero, the smaller was the activity difference between pop-out and training task after learning. Note that the polarity of the y-axis has been reversed, such that positive values are below zero and negative values are above zero, in order to preserve the general direction of the activity changes over time, as shown in (*a*). (*c*) Activity in the training task separately for target present (green) and absent conditions (red). (*d*) Same as (*c*) but for untrained pop-out control task. (*e*) Correlation between learning rate (x-axis, computed as slope across learning indices in different training session, see Materials and Methods for details) and activity change from pre to post (y-axis, computed by subtracting the difference score between pop-out and training task [see *b*] in pre-training from that in post-training). The lower the values on x- and y-axes the faster participants learned and the greater was the reduction of activity difference between pop-out and training task from pre- to post-training. Each dot represents the result from a different participant. Participants in the current study (black dots) were a subset of a larger group of subjects ($n = 19$ in total) who performed the same behavioral training sessions as well as pre- and post-fMRI and localizer scans for area MT+ for the purpose of a previous study (Frank et al. 2016). Other participants from this original group are included in the correlational analysis and their results are shown as gray dots.

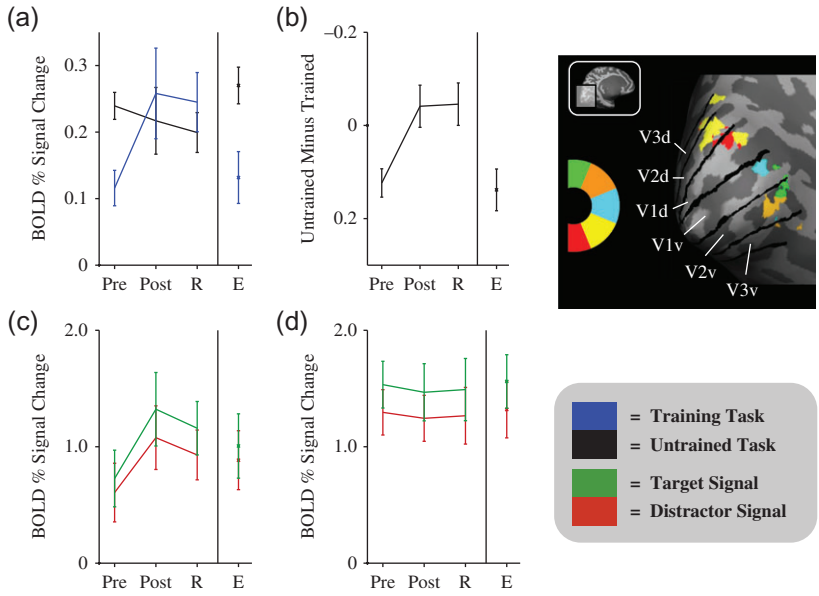

Figure 4. Same as Figure 3, but for activity of the target relative to the surrounding distractors in the retinotopic representations of the stimuli in early visual cortex (areas V1–V3). Only target-present trials were used for this analysis. The inset on the right shows stimulus representations in the left visual cortex of an example participant (posterior-medial view of the inflated occipital lobe). Different colors correspond to the retinotopic representations of different stimulus locations (green = 12 o'clock, brown = 1:30, cyan = 3, yellow = 4:30, red = 6). Stimulus representations spanned across visual areas V1, V2, and V3, which were defined by means of phase-encoded retinotopic mapping. (*a*) Target relative to distractor enhancement (in BOLD % signal change, see Materials and Methods for details). The larger the value, the more activation for the target among the distractors in the retinotopic representations of the stimuli in visual areas V1–V3 (zero on *y*-axis = no signal difference between target and distractor). Blue line = training task, black line = untrained pop-out control task. (*b*) Difference score, computed by subtracting the target relative to distractor signal in the training task (blue line in [*a*]) from that recorded for the untrained pop-out control task (black line in [*a*]), for each session. The closer this score was to zero, the smaller was the difference in target relative to distractor enhancement in the training task compared with that recorded for the untrained pop-out control task. Note that the polarity of the *y*-axis has been reversed, such that positive values are below zero and negative values are above zero, in order to preserve the general direction of the activity changes over time, as shown in (*a*). (*c*) Activity (in BOLD % signal change from zero = blank baseline) for the target (green line) and the surrounding distractors in the search array (red line) in target-present trials of the training task. The difference between activations for target and distractors was used in order to compute the target relative to distractor enhancement that is shown as a blue line in (*a*). (*d*) Same as (*c*) but for the untrained pop-out control task.

Results

Behavior

Training improved participants' performance on the learning task (see Fig. 2). Learning indices, quantifying improvements in reaction time relative to accuracy across target present and absent conditions (Fig. 2*e*), were significantly lower in post-training compared with pre-training ($W = 0$, $P = 0.02$), revealing the beneficial effect of training. This improvement was preserved over the course of 3 years: performance during retest was not significantly different from post-training performance ($W = 19$, $P = 0.47$), and significantly better than pre-training performance ($W = 0$, $P = 0.02$). However, training improvements were specific to learned target and distractor trajectories, because an exchange of target and distractors (see Fig. 1*c*) decreased performance significantly compared with post-training ($W = 28$, $P = 0.02$). Performance during exchange was not significantly different from pre-training ($W = 8$, $P = 0.38$). In the untrained pop-out control task (Fig. 2*b*,*d*,*e*), performance remained, as expected, unchanged across sessions, including target and distractor exchange conditions (repeated measures ANOVA on learning indices across testing sessions: $F(3,18) = 1.99$, $P = 0.15$). Individual differences in performance on the training task during post-training were preserved over time, such that higher performance during post-training was correlated with higher performance during retest ($r = 0.80$, $P = 0.03$) (Fig. 2*f*). Since the number of participants in the retesting group was low ($n = 7$), the effect size of this correlation should be interpreted with caution.

Similar results were obtained when only accuracy (across target present and absent conditions) was analyzed (Fig. 2*c*,*d*). Post-training performance in the learning task was significantly better than pre-training performance ($W = 28$, $P = 0.02$) and this performance gain was stable over the course of 3 years (retest vs. post: $W = 8.5$, $P = 0.38$; retest vs. pre: $W = 28$, $P = 0.02$). However, an exchange of target and distractors decreased performance significantly (exchange vs. post: $W = 0$, $P = 0.02$; exchange vs. pre: $W = 17.5$, $P = 0.61$). Accuracy in the pop-out control task did not change over time ($F(3,18) = 1.58$, $P = 0.23$).

The analysis of median reaction time in the training task suggested a similar trend as for accuracy, but weaker effects when target present and absent conditions were combined (Fig. 2*a*). Compared with pre-training,, participants tended to be faster in both post-training and retest, but both effects were not significant (post-training vs. pre-training: $W = 3$, $P = 0.08$; retest vs. pre-training: $W = 4$, $P = 0.11$). An exchange of target and distractors prolonged reaction times significantly (exchange vs. post-training: $W = 27$, $P = 0.03$; exchange vs. pre: $W = 11$, $P = 0.69$). In the untrained control task (Fig. 2*b*), reaction times remained unchanged across sessions ($F(3,18) = 1.97$, $P = 0.15$). The non-significant results in the post-training versus pre-training and retest versus pre-training comparisons in the training task were primarily caused by less pronounced changes from pre to post in the target-absent condition (Fig. 2*a*, red line), which might be expected if participants continued searching in the absence of finding a target. However, effects were strong in the target-present case (Fig. 2*a*, green line; post-training vs.

pre-training: W = 0, P = 0.02; retest vs. pre-training: W = 0, P = 0.02), indicating that the target, when present, was found more quickly with training.

Eye-Tracking

Similar to our previous report, using a larger sample of participants (Frank et al. 2016), we found that the mean deviation of the eye from central fixation in the current sample of subjects did not change across training sessions (slopes, quantifying changes in eye position in target present and absent trials over time, did not differ significantly from zero: W = 9, P = 0.47). Similar results were also obtained when slopes in target present and absent conditions were analyzed separately (target present: W = 9, P = 0.47; target absent: W = 14, P = 1.0). Overall, the deviation of the eye was small (average deviation from central fixation across all sessions and participants with SD: 1.7 ± 0.7°). Therefore, improvements in performance or changes in brain activity as a result of training are unlikely to be influenced by changes in fixation quality over time.

Brain Activity During Task Performance

In a first analysis, brain activity during task performance (target present and absent conditions) was analyzed over time (Fig. 3 and Supplementary Fig. 2). A 2 × 3 repeated measures ANOVA with factors task (trained vs. untrained) and time-point (pre vs. post vs. retest) revealed significant interactions between task and time in area MT+ ($F(2,12) = 7.46$, $P = 0.008$), TRL ($F(2,12) = 7.65$, $P = 0.007$), and V3ab ($F(2,12) = 4.47$, $P = 0.04$), but not in PPC ($F(2,12) = 0.53$, $P = 0.60$), suggesting that activity in MT+, TRL, and V3ab changed over time in the training task and remained stable in the pop-out control task (see Fig. 3*a* and Supplementary Fig. 2*a*,*e*). In PPC, activity in both training and control tasks remained stable over time (see Supplementary Fig. 2*i*). There were also significant main effects of task in TRL ($F(1,6) = 12.3$, $P = 0.01$) and V3ab ($F(1,6) = 8.22$, $P = 0.03$), indicating that activity in the pop-out control task was stronger than in the training task across all time-points (Supplementary Fig. 2*a*,*e*). No other significant main effects were found in any ROI (all P > 0.05).

Post hoc comparisons, based on the difference score "untrained pop-out minus training task" for each session, suggested that the activity difference between pop-out and training task was significantly smaller in post-training compared with pre-training (MT+: W = 1, P = 0.03; TRL: W = 0, P = 0.02; V3ab: W = 2, P = 0.047; Fig. 3*b* and Supplementary Fig. 2*b*,*f*), indicating that activation in the training task became more similar to the pop-out control task, after training. This is comparable to results from a larger group of participants, including subjects from the current study (see Frank et al. 2016). This change in activity from pre to post in the training task was not influenced by differences in the number of correct trials in each session; even when the same number of correct trials was used for post-training as in pre-training, effects remained similar to those found in the primary analysis (MT+: W = 2, P = 0.047; TRL: W = 0, P = 0.02; V3ab: W = 2, P = 0.047).

Finally, there was a significant correlation between the activity increase in MT+ (again, referenced to the untrained pop-out control task) from pre to post and participants' learning rates across the 12 training sessions: higher learning rate was associated with greater reduction in activity difference between pop-out and training task after learning (using the original set of 19 participants: $r = 0.57$, $P = 0.01$; limited to participants in the current sample: $r = 0.53$, $P = 0.23$; Fig. 3*e*). No such correlations were observed for TRL (current sample: $r = 0.31$, $P = 0.50$; original sample: $r = 0.004$, $P = 0.98$) or V3ab (current sample: $r = 0.03$, $P = 0.95$; original sample: $r = –0.01$, $P = 0.98$).

Results on the long-term stability of training-induced activity changes in TRL and V3ab remained inconclusive, because activity difference scores during retest were not significantly different from either pre-training (TRL: W = 3, P = 0.08; V3ab: W = 4, P = 0.11; Supplementary Fig. 2*b*,*f*) or post-training (TRL: W = 19, P = 0.47; V3ab: W = 22, P = 0.22; Supplementary Fig. 2*b*,*f*). Therefore, we focused the following analyses on area MT+.

During retest, the activity difference between pop-out and training task in MT+ was significantly larger than during post-training (W = 26, P = 0.047) and did not differ significantly from pre-training (W = 8, P = 0.38) (Fig. 3*b*). This reversal of activity in the training task relative to the untrained control task (Fig. 3*a*) is similar to results reported by Yotsumoto et al. (2008). In target and distractor exchange the activity difference between pop-out and training task tended to be larger compared with post-training (W = 25, P = 0.08) and was not significantly different from pre-training (W = 15, P = 0.94). Supplementary Figure 3 shows the activity difference score in MT+ for each session and participant.

Representation of Target Among Distractors

In a second analysis, the relative enhancement of BOLD signal for target versus distractors during target-present trials was examined for the retinotopic representations of the stimuli in early visual cortex (V1–V3) (Fig. 4 for group and Supplementary Fig. 4 for individual participant results): a 2 × 3 repeated measures ANOVA with factors task (trained vs. untrained) and time-point (pre vs. post vs. retest) revealed no main effects of task ($F(1,6) = 0.18$, $P = 0.69$) or time ($F(2,12) = 1.10$, $P = 0.36$), but a significant interaction between task and time ($F(2,12) = 7.31$, $P = 0.008$), indicating that the signal in the trained task changed with training and remained stable in the untrained pop-out control task (Fig. 4*a*). Post hoc tests confirmed this observation: the activity difference between pop-out and training task was significantly lower post-training versus pre-training (W = 0, P = 0.02), showing that the BOLD signal for target relative to distractors in the training task was enhanced, as a result of learning (Fig. 4*a*,*b*). Moreover, this change in representation from pre-training to post-training was significantly correlated with learning rate ($r = 0.79$, $P = 0.04$), such that faster learning was associated with greater reduction in signal difference between pop-out and training task after learning. Since the number of participants in the retesting group was low ($n = 7$), the effect size of this correlation should be interpreted with caution.

The induced target over distractor enhancement was stable over the course of 3 years: activity difference scores (pop-out task minus training task) were not significantly different between retest and post-training (W = 11, P = 0.69) and were significantly smaller in retest versus pre-training (W = 0, P = 0.02) (Fig. 4*b*). However, the enhancement was specific to learned target and distractor trajectories; in the exchange condition, training-induced enhancements of the target relative to the distractors reverted back to pre-training baseline levels (Fig. 4*a*, right side), such that the activity difference between untrained and training task (Fig. 4*b*, right side) was significantly larger in exchange compared with post-training (W = 26, P = 0.047) and did not differ significantly from pre-training (W = 13, P = 0.94).

A control analysis, for which the number of included correct target-present trials in the training task was matched across fMRI sessions for each participant (see Materials and Methods

for details), yielded similar results: a significantly smaller signal difference between untrained pop-out and training task in the post- versus pre-fMRI scan ($W = 0$, $P = 0.02$) was evident; no significant signal difference between retest and post-fMRI ($W = 14$, $P = 1.0$) could be determined; and a significantly smaller signal difference in retest versus pre ($W = 1$, $P = 0.03$) was found. In target and distractor exchange, there were trends for a larger signal difference between pop-out and training task compared with post-training ($W = 25$, $P = 0.08$) and the activity difference score in exchange did not differ significantly from that measured for pre-training ($W = 12$, $P = 0.81$). These results are comparable to the primary analysis and demonstrate that the observed activity changes are unlikely to be related to differences in the number of correct trials between sessions.

An analysis of target relative to distractor activity in the posterior part of the MT+ complex (corresponding to human area MT) suggested a similar pattern of activity changes over time (see Supplementary Fig. 5). However, results remained nonsignificant; a 2 × 3 repeated measures ANOVA with factors task (trained vs. untrained) and time-point (pre vs. post vs. retest) revealed no significant interaction between task and time-point ($F(2,12) = 1.70$, $P = 0.22$). There was a significant main effect of task ($F(1,6) = 10.9$, $P = 0.02$), indicating stronger target relative to distractor activity in the pop-out control compared with the training task across time. There was no significant main effect of time ($F(2,12) = 2.47$, $P = 0.13$).

Discussion

In this study, the long-term stability of perceptual learning of conjunctive motion trajectories was examined with a visual search task. The results show that learning is remarkably stable over time. Even 3 years after the end of original training, participants performed as if no time had passed between the final training session and retest. Learning was specific to trained stimulus identities and did not transfer when target and distractors were exchanged. On the neuronal level, learning was associated with an increase of activity during task performance in sensory cortex, including human area MT+ and retinotopic visual cortex (V1–V3, V3ab), whereas activity in parietal cortex did not change with training. Moreover, learning increased the activity for the target relative to the surrounding distractors in target-present trials in early retinotopic visual cortex (V1–V3). During retest, activity in MT+ during task performance reverted back to pre-training baseline levels, corroborating previous reports (Yotsumoto et al. 2008). Results on the long-term stability of activity changes in retinotopic visual cortex (V1–V3, V3ab) remained inconclusive, when the overall level of activation within an ROI during task performance was examined. However, when the analysis was limited to the representation of the target relative to the surrounding distractors in V1–V3, long-lasting learning effects were revealed. In particular, the training-induced enhancement of neuronal responses to the target relative to those to distractors outlasted the period of training and remained stable over the course of 3 years. Consistent with behavioral evidence for learning, this neuronal enhancement was specific to trained target and distractor identities and reverted back to pre-training baseline levels when target and distractors were exchanged. A similar, though nonsignificant, pattern of neuronal changes over time was observed for the posterior part of MT+, when target relative to distractor activity was analyzed. Overall, our results suggest that behavioral improvements in perceptual learning are enduring, whereas the neuronal correlates of learning differ in their respective time-courses, depending on the analysis performed (see below).

Long-Term Stability of Behavioral Learning

Similar to previous studies, we find that perceptual learning in the visual domain is stable over time (e.g., Ball and Sekuler 1982; Karni and Sagi 1993; Watanabe et al. 2002; Hussain et al. 2011; Frank et al. 2014b; Yashar et al. 2015). This suggests that enduring changes took place in the way the brain processes the learned stimuli. However, learning does not transfer when target and distractor trajectories are exchanged (see also Frank et al. 2014b, 2016), even though the distractor stimulus was seen more frequently than the target during training. Therefore, specificity to the trained target and distractor configurations (i.e., learned target and distractor identities) is preserved over time. Moreover, the individual degree of learning influenced retest performance. Participants who learned the task well and achieved higher post-training performance transferred this benefit to retest. This is comparable to results by Hussain et al. (2011) who reported that participants with better learning on a face/texture identification task also performed better during retest.

Long-Term Stability of Neuronal Changes

In our original study, using a larger sample of participants (Frank et al. 2016), we reported that perceptual learning in the trajectory search task is associated with activity increases in MT+ and retinotopic visual cortex (V1–V3, area V3ab). These effects were replicated in the smaller participant sample that was available for retest after 3 years, suggesting that the retest subjects were a representative subsample of our original participant group. Activity increased in both target present and absent conditions, even though changes appeared to be more pronounced for the target-present case (see Fig. 3 and Supplementary Fig. 2), as was found for behavioral improvements (Fig. 2). The only region where activity changes between pre-training and post-training correlated with perceptual learning rate was area MT+. Together with differences in pre-training cortical thickness of this region, which predict subsequent perceptual learning speed on the trajectory search task (see Frank et al. 2016), area MT+ appears to be critical to learning in this task. However, here we find that during retest the activity of MT+ reverted back to pre-training baseline levels, when the overall level of activation during target present and absent conditions was examined. Even though there were individual differences in how strongly MT+ was activated during retest (Supplementary Fig. 3), the majority of participants showed a pattern following the trend on the group-level.

These results are comparable to Yotsumoto et al. (2008) who reported that perceptual improvements in a texture discrimination learning task were associated with an initial increase of activity in early retinotopic visual cortex, which reverted back to pre-training baseline levels during a retest carried out a few weeks after the end of original learning. Importantly, behavioral improvements in the training task remained stable over time. Yotsumoto et al. (2008) interpreted this difference in the long-term stability of behavioral and neuronal effects of perceptual learning to result from long-term consolidation, leading to a more efficient representation of the trained visual information. On the cellular level, they argued that this could be caused by the downscaling of activity of synapses that are initially recruited but not critial to the long-term retention of the learned information, which would lead to a decrease of BOLD

signal. The learned information would presumably be represented over time in a smaller proportion of critical synapses that survive the process of downscaling (see Yotsumoto et al. 2008). Given the similarity between their results and our results for area MT+ (in the analysis of overall activation during task performance), the interpretation of Yotsumoto et al. might apply here as well. In that case, our results add to their original findings, by showing a similar neuronal pattern for a different perceptual learning task (visual search vs. texture discrimination), a different brain area (human MT+), and a much longer time-interval until retest (3 years compared with a few weeks). However, some of the neuronal correlates of learning in our task outlasted the period of training and persisted over time, as discussed below.

Recent results suggest that area V3a, a region likely sensitive to trackable motion features (Caplovitz and Tse 2007), is involved in perceptual learning of motion discrimination tasks (Shibata et al. 2012; Chen et al. 2015, 2016). Our findings are consistent with these reports in showing activity increases in V3ab after training. Unfortunately, the long-term stability of this effect remained inconclusive in our data set because retest activity in V3ab did not differ significantly from either pre-training or post-training. Similarly, the long-term stability of increased activity during task performance in early retinotopic visual cortex (trained stimulus locations in V1–V3) after training remained inconclusive for the same reason.

Representation of Target Among Distractors

A more conclusive pattern of results emerged when the representation of the target among the surrounding distractors was analyzed in the trained stimulus locations in V1–V3. We recently reported a neuronal marker of perceptual learning in visual search, namely, an enhancement of target relative to distractor BOLD signal in retinotopic visual cortex (Frank et al. 2014b; Reavis et al. 2016). Importantly, this measure is based on target among distractor trials only (that is, target-present trials), therefore providing a neuronal measure for the amount of activity evoked by the target relative to the simultaneously presented distractors. We found training-induced increases in target over distractor activity in various search tasks (see Frank et al. 2014b; Reavis et al. 2016), and even under conditions where participants do not perform any search on the stimuli and focus instead on an orthogonal task at screen center, suggesting that the effect is primarily stimulus-driven (Reavis et al. 2016).

Results of the current study show that this neuronal signature is long-lasting and specific. Even 3 years after the end of original training the learning-induced target versus distractor enhancement was significantly larger compared with pre-training baseline (see Fig. 4 and Supplementary Fig. 4). Both the BOLD-signal level for the target as well as the BOLD-signal levels for the surrounding distractors increased with learning and remained above pre-training levels during retest (Fig. 4c). However, the induced enhancement was specific to trained stimulus identities; when target and distractors were exchanged, acquired activity changes reverted back to pre-training baseline, similar to results found for behavioral performance. This specificity to trained target and distractor identities, make it unlikely that the target over distractor enhancement is a nonrelevant side-effect of learning.

A similar time-course of target relative to distractor activity was observed for the posterior part of area MT+ (corresponding to human MT). The nonsignificant results in MT might be caused by the lower sensitivity of the analysis in this area, because left and right MT each represented 3 contralateral stimulus locations, therefore reducing the target signal by the activation evoked by 2 distractors in the same representation.

The analysis of target relative to distractor activity differed from the analysis of overall activation because it included only target-present trials. The greatest learning effects, in terms of both reaction time and accuracy, occurred in this condition (see Fig. 2*a*,*c*) and therefore, the most pronounced neuronal learning effects are to be expected when the target was present. However, when the activity for target-present was analyzed for all stimulus locations together, there was still a decline of activation toward baseline during retest, at least in MT+ (see green line in Fig. 3c). Only when stimulus representations were separated into clearcut target and distractor retinotopic locations and when the difference between these 2 representations was analyzed, were long-lasting neuronal learning effects observed. This was the case for early retinotopic visual cortex and there were trends for a similar effect in MT. These long-lasting changes in target and distractor activations were potentially diluted by the decrease of overall activity in the combined target present and absent analysis across all stimulus locations (in particular in the MT+ complex). In the framework of Yotsumoto et al. (2008), the cellular mechanism might be that critical synapses that became tuned to target and distractor trajectories over the course of training survived the process of synaptic downscaling and continued to produce a stronger signal during retest when a target was presented in their receptive field compared with when a distractor was presented. This explanation is post hoc since we did not have a hypothesis for such a pattern of results. Clearly, more research will be required to clarify the conditions under which enhanced neural activation persists and when it does not.

Perceptual learning in this type of visual search task is context dependent, because neurons become tuned to specific stimulus identities ("v"-trajectory for the target and inverted "v"-trajectory for the distractor). Therefore, an exchange of target and distractor stimuli in the training task (Fig. 1c) leads to a decrease of performance to pre-training baseline levels (Fig. 2) and eliminates the target relative to distractor enhancement in the trained stimulus locations of early retinotopic visual cortex (Fig. 4). The pronounced specificity to trained target and distractor trajectories also rules out the possibility that behavioral improvements and changes in neuronal activity reflect learning of the visual search task per se. Rather, learning in this task reflects the specific association between stimuli and target and distractor identities.

For each analysis and ROI, activity in the untrained pop-out control task appeared to remain stable across sessions. This was expected for stimuli that elicit target pop-out and that require no learning, making the untrained control task a suitable reference condition for the comparison with activity in the training task. Even though the activity difference between untrained pop-out control and training task was very small in the post-training scan for both target among distractor enhancement (Fig. 4*b*) and overall activation in MT+ (Fig. 3*b*), participants were still faster and better in the pop-out control task (Fig. 2*b*,*d*) compared with the training task (Fig. 2*a*,*c*). However, even after learning, reaction times in particular are expected to be longer in the trajectory task, because information about the complete trajectory cycle (duration of one cycle = 600 ms) has to be collected before a decision about target presence or absence can be made. Since each dot moves in only one direction in the control task, information does not have to be integrated over as long a time-interval as in the

motion conjunction case, and therefore responses, even after extensive training on the trajectory task, are expected to be longer than in the control task. Additional training sessions might have further reduced differences in reaction time and accuracy between training and control task.

Limitations

A major limitation of our study is the small sample size of trained participants who were available for retesting 3 years after the end of their original training sessions. This was unavoidable given the difficulty of locating willing subjects after such a long-time interval. This small sample limited the statistical power and rendered some effects inconclusive (e.g., concerning activity in area V3ab during retest). Future studies might overcome this problem by performing the retest on a larger sample of participants at an earlier time-point.

Our analysis is further limited by the design of the study, which required participants to perform the search task during each scanning session. On the one hand, this has the advantage that the same task is performed during imaging and behavioral training sessions; on the other hand, it renders the usage of more advanced multivariate or connectivity analyses difficult, because of the sensitivity of these analyses to imbalances in the number of correct trials in each session. Therefore, changes and long-term stability in multivariate activity or functional connectivity patterns might exist (as indicated, for instance, in recent studies by Bi et al. 2014 and Chen et al. 2015, 2016, as well as by results in nonhuman primates, see Law and Gold 2008), but could not be assessed given the design of the current experiment.

One might also argue that improvements in performance (accuracy and reaction time) by means of training could affect fMRI activations in the training task. However, we minimized this problem by analyzing, for each session, only trials with correct participant responses and by modeling the BOLD response for each trial only until participants responded. This approach did not artificially lead to changes in activity across sessions, as the absence of any such effects in PPC shows. Moreover, a control analysis for target over distractor enhancement, with matched number of trials across sessions, revealed similar effects as the primary analysis for which all correct trials in each session were used. Therefore, we deem it unlikely that our fMRI results are affected by differences in perfomance between sessions.

Even though our analysis of target and distractor representations in retinotopic visual cortex is a sensitive analysis, the results had to be averaged across different stimulus locations and different subregions of early retinotopic visual cortex (V1–V3 combined) in order to increase the signal to noise ratio through averaging, and thus the power of our analysis. This was primarily caused by the small number of trials in which the target was presented at each stimulus location (10 trials per location in each session, which could be lower because of incorrect subject responses) and the difficulty of defining each stimulus representation separately in visual areas V1, V2, and V3 (for an example, see inset on the right in Fig. 4). Future studies might therefore include more target-present trials and use a smaller number of stimulus locations in order to increase experimental power.

Finally, it should be emphasized that the long-lasting neuronal learning effects in this study were found for trajectory stimuli presented in the context of a visual search paradigm (e.g., Frank et al. 2014b, 2016; Reavis et al. 2015, 2016). Therefore, it remains a goal for future studies to demonstrate similar effects in tasks other than visual search.

Conclusion

In this study, we examined the long-term stability of perceptual learning using a visual search task for complex motion trajectories. Participants were trained on the task for 3 weeks and performed a retest in the training task 3 years after the end of their original training sessions. Brain activity during task performance was measured at different time-points using fMRI. Results show that perceptual improvements were stable over time. The analysis of overall activation during task performance suggested that initial activity increases in area MT+, evident over the course of training, reverted back to pre-training baseline levels during retest. An analysis of target relative to distractor activity revealed that a training-induced enhancement of target relative to distractor representations in early retinotopic visual cortex (V1–V3) persisted for years beyond the period of training. This neuronal signature was specific to trained stimulus trajectories and reverted back to pre-training baseline levels when target and distractors were exchanged. A similar, though nonsignificant, pattern of target relative to distractor activity over time was observed for the posterior part of MT+ (human MT). Overall, these results suggest that perceptual learning of motion feature conjunctions in the context of a visual search task induces enduring neuronal changes in the way the learned stimuli are represented.

Supplementary Material

Supplementary material are available at *Cerebral Cortex* online.

Funding

The study was supported by the Alexander von Humboldt Foundation (to authors P.U.T. and S.M.F.) and by the Deutsche Forschungsgemeinschaft (GR988 22-1 to author M.W.G.).

Notes

We wish to thank Lisa Forster for help with data collection and our participants for their time. We also wish to thank 4 anonymous reviewers for helpful comments on a previous version of this manuscript. *Conflict of Interest*: None declared.

References

Ball K, Sekuler R. 1982. A specific and enduring improvement in visual motion discrimination. Science. 218:697–698.

Bi T, Chen J, Zhou T, He Y, Fang F. 2014. Function and structure of human left fusiform cortex are closely associated with perceptual learning of faces. Curr Biol. 24:222–227.

Brainard DH. 1997. The psychophysics toolbox. Spat Vis. 10: 433–436.

Bruyer R, Brysbaert M. 2011. Combining speed and accuracy in cognitive psychology: is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)? Psychol Belg. 51:5–13.

Caplovitz GP, Tse PU. 2007. V3A processes contour curvature as a trackable feature for the perception of rotational motion. Cereb Cortex. 17:1179–1189.

Chen N, Bi T, Zhou T, Li S, Liu Z, Fang F. 2015. Sharpened cortical tuning and enhanced cortico-cortical communication

contribute to the long-term neural mechanisms of visual motion perceptual learning. Neuroimage. 115:17–29.
Chen N, Cai P, Zhou T, Thompson B, Fang F. 2016. Perceptual learning modifies the functional specializations of visual cortical areas. Proc Natl Acad Sci USA. 113:5724–5729.
Dale AM, Fischl B, Sereno MI. 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage. 9: 179–194.
Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, et al. 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage. 31:968–980.
DeYoe EA, Carman GJ, Bandettini P, Glickman S, Wieser J, Cox R, Miller D, Neitz J. 1996. Mapping striate and extrastriate visual areas in human cerebral cortex. Proc Natl Acad Sci USA. 93: 2382–2386.
Engel SA, Glover GH, Wandell BA. 1997. Retinotopic organization in human visual cortex and the spatial precision of functional MRI. Cereb Cortex. 7:181–192.
Fischl B, Sereno MI, Dale AM. 1999. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. Neuroimage. 9:195–207.
Frank SM, Baumann O, Mattingley JB, Greenlee MW. 2014a. Vestibular and visual responses in human posterior insular cortex. J Neurophysiol. 112:2481–2491.
Frank SM, Reavis EA, Greenlee MW, Tse PU. 2016. Pretraining cortical thickness predicts subsequent perceptual learning rate in a visual search task. Cereb Cortex. 26:1211–1220.
Frank SM, Reavis EA, Tse PU, Greenlee MW. 2014b. Neural mechanisms of feature conjunction learning: enduring changes in occipital cortex after a week of training. Hum Brain Mapp. 35:1201–1211.
Furmanski CS, Schluppeck D, Engel SA. 2004. Learning strengthens the response of primary visual cortex to simple patterns. Curr Biol. 14:573–578.
Gibson EJ. 1963. Perceptual learning. Annu Rev Psychol. 14:29–56.
Huk AC, Dougherty RF, Heeger DJ. 2002. Retinotopy and functional subdivision of human areas MT and MST. J Neurosci. 22:7195–7205.
Hussain Z, Sekuler AB, Bennett PJ. 2011. Superior identification of familiar visual patterns a year after learning. Psychol Sci. 22:724–730.
Karni A, Sagi D. 1993. The time course of learning a visual skill. Nature. 365:250–252.
Kourtzi Z, Betts LR, Sarkheil P, Welchman AE. 2005. Distributed neural plasticity for shape learning in the human visual cortex. PLoS Biol. 3:e204.
Law CT, Gold JI. 2008. Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. Nat Neurosci. 11:505–513.
Pelli DG. 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis. 10: 437–442.
Reavis EA, Frank SM, Greenlee MW, Tse PU. 2016. Neural correlates of context-dependent feature conjunction learning in visual search tasks. Hum Brain Mapp. 37:2319–2330.
Reavis EA, Frank SM, Tse PU. 2015. Caudate nucleus reactivity predicts perceptual learning rate for visual feature conjunctions. Neuroimage. 110:171–181.
Sasaki Y, Náñez JE, Watanabe T. 2010. Advances in visual perceptual learning and plasticity. Nat Rev Neurosci. 11: 53–60.
Shibata K, Chang LH, Kim D, Náñez JE, Kamitani Y, Watanabe T, Sasaki Y. 2012. Decoding reveals plasticity in V3A as a result of motion perceptual learning. PLoS One. 7:e44003.
Sigman M, Pan H, Yang Y, Stern E, Silbersweig D, Gilbert CD. 2005. Top-down reorganization of activity in the visual pathway after learning a shape identification task. Neuron. 46: 823–835.
Sireteanu R, Rettenbach R. 1995. Perceptual learning in visual search: fast, enduring, but non-specific. Vision Res. 35: 2037–2043.
Townsend JT, Ashby FG. 1978. Methods of modeling capacity in simple processing systems. In: Castellan J, Restle F, editors. Cognitive theory. Vol. 3. Hillsdale, NJ, USA: Erlbaum. p. 200–239.
Watanabe T, Náñez JE, Koyama S, Mukai I, Liederman J, Sasaki Y. 2002. Greater plasticity in lower-level than higher-level visual motion processing in a passive perceptual learning task. Nat Neurosci. 5:1003–1009.
Yashar A, Chen J, Carrasco M. 2015. Rapid and long-lasting reduction of crowding through training. J Vis. 15:1–15.
Yotsumoto Y, Watanabe T, Sasaki Y. 2008. Different dynamics of performance and brain activation in the time course of perceptual learning. Neuron. 57:827–833.