NeuroImage 73 (2013) 225–236

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Comments and Controversies

Teaching an adult brain new tricks: A critical review of evidence for training-dependent structural plasticity in humans☆

Cibu Thomas*, Chris I. Baker

Laboratory of Brain and Cognition, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892, USA
Center for Neuroscience and Regenerative Medicine at the Uniformed Services University of the Health Sciences, Bethesda, MD, USA

ARTICLE INFO

Article history:
Accepted 22 March 2012
Available online 30 March 2012

Keywords:
VBM
DTI
Structural plasticity
Training
MRI

ABSTRACT

A growing number of structural neuroimaging studies have reported significant changes in gray matter density or volume and white matter microstructure in the adult human brain following training. Such reports appear consistent with animal studies of training-dependent structural plasticity showing changes in, for example, dendritic spines. However, given the microscopic nature of these changes in animals and the relatively low spatial resolution of MRI, it is unclear that such changes can be reliably detected in humans. Here, we critically evaluate the robustness of the current evidence in humans, focusing on the specificity, replicability, and the relationship of the reported changes with behavior. We find that limitations of experimental design, statistical methods, and methodological artifacts may underlie many of the reported effects, seriously undermining the evidence for training-dependent structural changes in adult humans. The most robust evidence, showing specificity of structural changes to training, task and brain region, shows changes in anterior hippocampal volume with exercise in elderly participants. We conclude that more compelling evidence and converging data from animal studies is required to substantiate structural changes in the adult human brain with training, especially in the neocortex.

Introduction

The relationship between brain structure and function has gained recent prominence in human neuroimaging. Studies have reported correlations between behavioral performance and localized brain structure (for a recent review, see Kanai and Rees, 2011) and have also identified possible training-dependent changes in structure (e.g. changes in measures of gray matter density or white matter integrity). Evidence for effects of training comes from both cross-sectional studies, comparing different groups of subjects with different experiences (e.g. musicians *versus* non-musicians (Bengtsson et al., 2005) or taxi *versus* bus drivers (Maguire et al., 2006)), as well as longitudinal studies, examining the effect of training over time in individuals (for a review see, Draganski and May, 2008; May and Gaser, 2006). However, with cross-sectional studies it is impossible to determine which came first, the structural differences or the experience (May, 2011). Longitudinal training studies provide the strongest evidence for training-dependent changes in brain structure since experience is directly manipulated and the changes are measured within a participant.

Such MRI evidence for adult structural plasticity seems consistent with animal studies of experience-dependent plasticity (Draganski and May, 2008) and based on this apparent convergence, it has been proposed that changes in the MRI signal may reflect changes in axonal myelination, neurogenesis, angiogenesis, dendritic spine motility, glial cell proliferation, and synaptogenesis (Draganski and May, 2008; Scholz et al., 2009). However, while animal studies do suggest that experience-dependent structural plasticity in the adult brain persists throughout the life span (Fu and Zuo, 2011), it is highly constrained. For example, longitudinal *in vivo* studies suggest that experience does not cause any change in large scale axons and dendrites (Mizrahi and Katz, 2003; Trachtenberg et al., 2002), although some cross-sectional studies have reported changes in glial cells, unmyelinated axons or dendritic length in adult animals exposed to an enriched environment (Juraska et al., 1980; Markham et al., 2009), altered visual input (McBride et al., 2008), or motor tasks (Black et al., 1990; Kleim et al., 2007). There is also no direct evidence for experience-driven increase in axonal myelination in the adult brain (Demerens et al., 1996; Markham et al., 2009). Likewise, experience-dependent angiogenesis in adults has been shown to be specific to exercise (Black et al., 1990; Kleim et al., 2002) (but see Isaacs et al., 1992), and the only undisputed claim regarding neurogenesis in the adult brain is that it is primarily observed in the dentate gyrus of the hippocampal complex and the olfactory bulb (Rakic, 2002).

☆ No conflicts of interest.

* Corresponding author at: Laboratory of Brain and Cognition, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892, USA. Fax: +1 301 443 7111.
E-mail address: cibu.thomas@nih.gov (C. Thomas).

1053-8119/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2012.03.069

Overall, the bulk of the evidence from animal studies suggests that experience-dependent structural plasticity is mediated by remodeling of neuronal processes (Lerch et al., 2011; McBride et al., 2008), synaptogenesis (Black et al., 1990; Briones et al., 2004; Knott et al., 2006) or transient changes in dendritic spines (Holtmaat et al., 2006; Trachtenberg et al., 2002; Xu et al., 2009b) and axonal boutons (Stettler et al., 2006; Yamahachi et al., 2009). Importantly, the experience-related increase in structures like dendritic spines is also accompanied by spine elimination resulting in similar total spine densities between the trained and untrained animals after training (Trachtenberg et al., 2002; Xu et al., 2009b). At the systems level, such subtle changes are considered sufficient to remodel patterns of activity in neuronal circuits (Chen and Nedivi, 2010), without inducing large-scale structural alterations in cortical networks. Thus, the evidence from animal studies suggests that the large-scale organization of axons and dendrites is very stable and experience-dependent structural plasticity in the adult brain occurs locally and is transient (for a review see, Holtmaat and Svoboda, 2009).

One of the big advantages of MRI is the capacity to image the whole brain, rather than individual cellular structures as in the case of, for example, 2-photon microscopy. However, given the large-scale stability of structures described above, it is not clear that human MRI, with typically 1 mm^3 voxels, can detect the type of microscopic structural changes reported in animal studies. In addition, it is important to note that much of the evidence from animal studies comes from highly invasive or demanding experimental manipulations such as trimming whisker barrels or rearing animals in enriched *versus* isolated environment, and the animals are motivated by requiring performance to receive food. In comparison, human studies use less intensive and demanding training tasks and some have suggested that a controlled training protocol is not even necessary for inducing structural changes in the adult brain (Bezzola et al., 2011). Finally, we note that the effect size reported in some human studies is very small relative to the size of the voxels. For example, memory training was reported to increase cortical thickness by ~0.05 mm (Engvig et al., 2010). Similarly, aerobic exercise was reported to increase hippocampal volume by ~0.10 mm^3 (Erickson et al., 2011a). Given that such effect sizes are many times smaller than the sampling frequency of the method, these results need to be carefully evaluated and interpreted with caution.

Taking into account these considerations, we conducted a critical review of the evidence from all longitudinal studies of training-dependent structural plasticity in adult humans. There are two key questions. First, how reliable is the evidence for training-dependent changes in MRI measures of brain structure? Second, if there are reliable changes, what do these changes in MRI measures reflect in terms of the biological substrate? Here, we focus primarily on the first question, but will discuss the second issue towards the end of the review. Specifically, in contrast to previous reviews of this literature (Draganski and May, 2008; May, 2011), we focus on the robustness of the experimental design and statistical methods, as well as the limitations of MRI-based structural imaging techniques.

In the first part of the review, we will discuss the different methods used to measure human adult structural plasticity and briefly survey the extant literature. In the central part of the review, we will evaluate the reported findings in terms of specificity, replicability and correlation with behavior. Finally, we will consider in more detail the inherent limitations of MRI measures of structure and the relationship between MRI measures and the biological substrate.

Measuring training-dependent structural changes

In total, we identified 20 research articles (see Table 1) that satisfy the following inclusion criteria: (a) the studies involved healthy adults (mean age > 18). (b) A longitudinal design was employed and participants were scanned before and after training in a specific task. (c) MRI-based techniques were used to measure structural changes (Fig. 1). The training tasks employed in these studies range from visuomotor tasks such as, juggling (for e.g., Draganski et al., 2004), golf (Bezzola et al., 2011) or balancing (Taubert et al., 2010) to cognitive tasks such as deciphering Morse code (Schmidt-Wilcke et al., 2010), working memory (for e.g., Takeuchi et al., 2010), and learning for an exam (Draganski et al., 2006). Most studies involved young adults (<30 years of age) although some focused on older adults (>60 years of age). The duration of training in all these studies varied from 3 days (Kwok et al., 2011) to 1 year (Erickson et al., 2011a). Differences in any of these factors (age, duration of training, etc.) could influence the nature of any underlying structural change. Here, however, we will restrict our focus simply to the strength of the evidence presented in individual studies and the techniques employed.

The most common techniques used in these studies to measure longitudinal structural changes are voxel-based morphometry (VBM)

Table 1
Details of the twenty longitudinal studies that tested for training-related structural plasticity in the adult human brain using MRI methods. The studies are arranged in chronological order (# approximate mean age of the training group, * Young Group, ^ Elderly Group).

Study	Method	Control condition	Sample		Mean age # (~years)	Task		Training duration (~days)
			Training	Control		Training	Control	
Draganski et al., 2004	VBM	Between subjects	12	12	22	Juggling	None	90
Colcombe et al., 2006	VBM	Between subjects	30	29	66	Aerobics	Stretching	120
Draganski et al., 2006	VBM	Between subjects	38	12	24	Learning abstract information	None	90
Boyke et al., 2008	VBM	Between subjects	25	25	60	Juggling	None	90
Ilg et al., 2008	VBM	Between subjects	18	18	24	Reading Mirrored words	None	14
Driemeyer et al., 2008	VBM	None	20	None	27	Juggling	None	7
Thomas et al., 2009	VBM	Within subjects	12	12	33	Visuo-motor		14
Scholz et al., 2009	VBM and DTI	Between subjects	24	24	25	Juggling	None	45
Tang et al., 2010	VBM and DTI	Between subjects	22	23	21	Integrative body mind therapy	Relaxation therapy	30
Taubert et al., 2010	VBM and DTI	Between subjects	14	14	26	Whole body dynamic balancing	None	30
Schmidt-Wilcke et al., 2010	VBM	Between subjects	16	15	26	Decipher Morse code	None	120
Takeuchi et al., 2010	VBM and DTI	None	11	None	22	Working memory	None	60
Lövdén et al., 2010	DTI and volume analysis	Between subjects	32	23	25*, 69^	Working memory	None	101
Engvig et al., 2010	Cortical thickness	Between subjects	22	20	62	Memory	None	56
Erickson et al., 2011a	Volume analysis	Between subjects	60	60	67	Aerobics	Stretching	365
Kwok et al., 2011	VBM	None	19	None	20	Learning color names	None	3
Landi et al., 2011	VBM and DTI	None	12	None	26	Visuo-motor tracking	None	7
Bezzola et al., 2011	VBM	Between subjects	12	12	51	Golf	None	40
Engvig et al., 2011	DTI	Between subjects	21	20	62	Memory	None	56
Woollett and Maguire, 2011	VBM	Between subjects	39	31	38	Spatial memory	None	1095

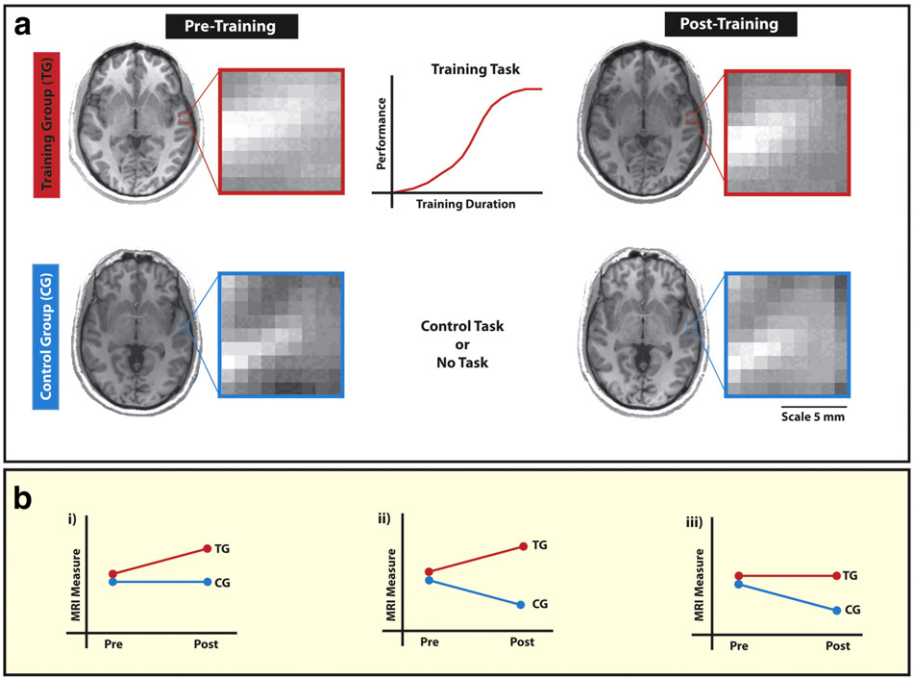

Fig. 1. Studying training-dependent structural plasticity. (a) Typical experimental design employed by many of the studies reporting training-related structural changes in the adult human brain. A group of adults are randomly assigned to a training group (TG) or a control group (CG) and structural brain data is collected before and after the training period. The example illustrated here, shows a set of T_1-weighted MPRAGE scans (voxel size 1 mm isotropic) of two young adult volunteers collected within the same scan session without repositioning. The raw anatomical data are aligned to the anterior–posterior commisure (AC–PC) plane. An arbitrary region of the brain is enlarged in both subjects, across the two scans, to show the variation in pixel intensities that is used to segment gray and white matter voxels in techniques like voxel based morphometry. Over and beyond the natural variability observed between scans of the same volunteer, the prediction with regard to training-dependent structural plasticity is that TG individuals will evince significant changes in measures of gray and/or white matter in a specific brain region relative to the CG. (b) Training-dependent structural plasticity predicts an interaction between time-point and group in structural measures. Specifically, the difference between TG and CG should be greater post- compared with pre-training. Three possible types of interaction are illustrated. (i) TG shows a significant increase in the MRI measure post-training with no change in CG. (ii) CG shows a longitudinal decline in the MRI measure whereas the TG shows an increase. Such a decrease in CG might be expected in elderly subjects. (iii) CG shows a longitudinal decline, which is arrested in the training group due to the training.

(Ashburner and Friston, 2000) and diffusion tensor imaging (DTI) (Basser et al., 1994). VBM is a whole-brain, automatic technique that enables voxel-wise statistical comparison of local gray matter volume (GMV) or gray matter density (GMD) between groups or time-points. In this technique, T_1-weighted structural brain images are provided as input to the standard VBM processing stream (Ashburner and Friston, 2000), and statistical parametric maps (SPM) indicating differences in the density or volume of brain tissue are provided as the output. While VBM is predominantly used for examining differences in gray matter (although see, Colcombe et al., 2006), DTI is used specifically to detect differences in white matter microstructure. In this technique, the diffusion properties of water molecules in each voxel are mathematically characterized by a tensor (Basser and Pierpaoli, 1996) and after standard processing routines (Pierpaoli, 2011), statistical tests are performed to identify changes in measures of white matter microstructural integrity such as, fractional anisotropy (FA). In addition to these two approaches, the same T_1-weighted images used in VBM can also be used to create maps of cortical thickness (Fischl and Dale, 2000), which can then be compared before and after training (Engvig et al., 2010). Further, structures such as the hippocampus can be readily segmented and volume measures over time estimated (Erickson et al., 2011a). Alternatively, rather than employ techniques that rely on segmentation algorithms to measure gray and white matter changes, the T_1-weighted images of a subject from multiple time points can be non-linearly aligned and deformation measures that reflect an increase or decrease in volume over time can be computed using a technique called deformation based morphometry (DBM) (Chung et al., 2001; Hyde et al., 2009). Finally, MRI measures of longitudinal (T_1) and transverse (T_2) relaxation times (Deoni et al., 2008; Tofts, 2003) of brain tissue that offer quantitative measures of the properties of white and gray matter could also be used for detecting microstructural changes following training.

As with any MRI measure, the raw data from structural MRI scans reflect both the true underlying signal and noise. Further, the raw data are subjected to several stages of processing to remove potential artifacts and enable valid statistical inference (Ashburner and Friston, 2000; Pierpaoli, 2011), and the final outcome is far removed from the original data. At each stage, small biases can be inadvertently introduced in the processing pipeline, which can give rise to false positives (Jones et al., 2005; Thomas et al., 2009) and make it difficult to detect small changes in brain structure (Klauschen et al., 2009).

In light of these concerns, the reliability of the evidence for training-dependent structural plasticity critically depends on the rigor of the experimental design and analysis methods adopted. Here, we review the robustness of the evidence provided by published studies focusing on three criteria, which apply to any study investigating the impact of training on neural processing: a) specificity, both to the training regimen and to particular brain regions; b) replicability; and c) correlation with behavioral measures of training. While these criteria are not pre-requisites for demonstrating training-dependent changes, fulfillment of all three would provide the most compelling evidence. In the following sections, we discuss each of these criteria in turn and evaluate the current evidence in the context of each.

Specificity

The level of specificity of any apparent training effect is critical for its interpretation. At a basic level, it is important to demonstrate that any changes are specific to training and are not observed in untrained subjects. However, trained and untrained subjects are not well

matched in terms of their overall experience and therefore, demonstration of specificity to a particular task provides more compelling evidence. Further, finding specificity to a particular task can give important insights into what aspects of training contribute to any structural change. Finally, it is important to understand how specific any training effects are to particular brain regions, especially in terms of elucidating the underlying mechanisms.

Training

Given the measurement error inherent even in structural imaging scans and the extensive pre-processing typically involved in assessing changes in brain structure, it is important to take into account the reliability of the structural measures used to assess the impact of training. Typically, this is done by collecting control data in the absence of the specific training protocol, often in a separate group of subjects (Fig. 1). Such a control group is particularly important when studying older adults where there may be changes in structural properties over time in the absence of any training (Fig. 1b). Training-dependent structural plasticity is then evidenced by any MRI measurement that shows an interaction between group (training *versus* control) and time-point (pre- *versus* post-training), that is a larger difference between the two groups post- compared with pre-training. Despite the importance of collecting control data, four of the 20 studies we identified did not include any control condition (Driemeyer et al., 2008; Kwok et al., 2011; Landi et al., 2011; Takeuchi et al., 2010), making it impossible to assess the training-dependence of any reported effects.

Of the remaining 16 studies, the vast majority (15/16) performed a between-subjects design with separate groups of subjects either assigned to a specific training protocol (training group) or not (control group). In most cases the control group was not assigned any task between the pre- and post-training scans and the comparison between these scans can be considered a measure of the reliability of the structural measures. Despite the universal claim of training-dependent structural changes in these studies, it is hard to find clear and compelling evidence for a robust statistical interaction between the factors group and time-point.

In particular, in a whole-brain analysis as employed in most studies, it would be ideal to run a direct test, with appropriate correction for multiple comparisons, for voxels showing a significant group by time-point interaction. However, many studies did not appear to run such a test and in some, the precise details of the statistical analysis are hard to establish (Boyke et al., 2008; Draganski et al., 2004, 2006). Overall, we identified two major concerns in the studies we reviewed. First, in some studies, the differences in structural measures between time-points were first examined in separate longitudinal analyses within each group (Boyke et al., 2008; Draganski et al., 2004, 2006; Woollett and Maguire, 2011), with the presence of training-dependent structural changes being inferred from significant changes in a set of voxels in the training but not the control group. The group by time point interaction was either not performed (Woollett and Maguire, 2011), not corrected for multiple comparisons (Boyke et al., 2008), or the results of statistical tests were not explicitly reported (Draganski et al., 2004, 2006). As highlighted by Nieuwenhuis et al. (2011) without clear evidence of a significant interaction, the presence of a significant effect in one group but not the other is only weak evidence for a specific effect of training.

Second, in several studies that we reviewed, the interaction analysis was restricted to voxels that were first selected based on a longitudinal analysis in the training group only, and the reported effect was based on performing the interaction test using the same rather than independent data (Bezzola et al., 2011; Ilg et al., 2008; Scholz et al., 2009; Tang et al., 2010; Taubert et al., 2010). This approach of using the same data to select voxels and further analyze effects produces a selection bias that has been demonstrated to produce potentially spurious effects (Baker et al., 2007; Kriegeskorte et al., 2009, 2010) and precludes reliable estimation of the true effect size of training on brain structure. This problem of circularity often extends to the plotting of the data, with magnitude estimates and error bars displayed for pre-selected voxels rather than using data independent of the selection criteria (for e.g., Takeuchi et al., 2010; Woollett and Maguire, 2011). While such plots may be useful for illustrating the trends and direction of effects, the results will be biased and distorted toward the effects that were selected for (see supplementary materials in Kriegeskorte et al., 2009). To avoid circularity, voxel selection should be conducted independent of further analysis such as in a split-half analysis (Kriegeskorte et al., 2009) or in pre-defined regions-of-interest (ROIs) (see also, 'Conclusions').

In addition to these two concerns, it is worth noting that although the repeated scans in untrained control subjects enable assessment of reliability of the structural measures, there is rarely a clear estimation or presentation of the measurement variability. Sometimes the data are normalized to the control data, which obscures the variability within these scans (Draganski et al., 2004, 2006).

While a between-subjects design has most commonly been employed, the remaining study containing a control condition used a within-subject design (Thomas et al., 2009). By avoiding the effect of between-subjects variance, such a design is potentially much more powerful (Poldrack, 2000). In particular, as shown in Fig. 2, subjects were scanned three times at two-week intervals, the first two scans serving as a control condition with no training in-between, and the second two scans occurring before and after training (Fig. 2a). Although significant changes in behavioral performance and functional cortical activity were observed after training, the magnitude and the location of concomitant changes in GMD varied as a function of the scan used for alignment and spatial normalization as well as the software tools employed (Figs. 2b,c). Differences in results due to choice of alignment scan likely arise from the effect of interpolation during the spatial normalization (see also, 'challenges in MRI based structural imaging'). Importantly, when optimized analysis methods were employed, the authors did not find any evidence for training-dependent structural changes (Thomas et al., 2009).

Overall, the bulk of the evidence reporting training-dependent structural changes in adults does not appear to contain rigorous statistical evidence for specificity to training and the results should be interpreted with caution. In only five studies (Colcombe et al., 2006; Engvig et al., 2010, 2011; Erickson et al., 2011a; Schmidt-Wilcke et al., 2010) could we find any indication that the appropriate statistical tests to identify training-related structural changes were conducted. It is also worth noting that the study that employed a powerful within-subject control found no evidence for training-related structural plasticity (Thomas et al., 2009).

Task

While comparison of a trained group with an untrained group can provide evidence for specificity to training, the strongest test is to compare two groups who have been trained on different tasks and show that the changes are specific to a given task and not a general effect of any training. Such a comparison also makes it possible to determine what aspects of any given task are critical for inducing structural changes. Absence of a training task for the control group also raises an additional concern about the matching of the two groups of subjects. Specifically, the lack of training in a control group may introduce confounding factors. For example, a group of subjects that has been trained on a specific task over a period of weeks is much more invested in the study and may exhibit less head motion in the second scan or even feel less anxious about the scan, compared to a control group without a task. This is important because the signal-to-

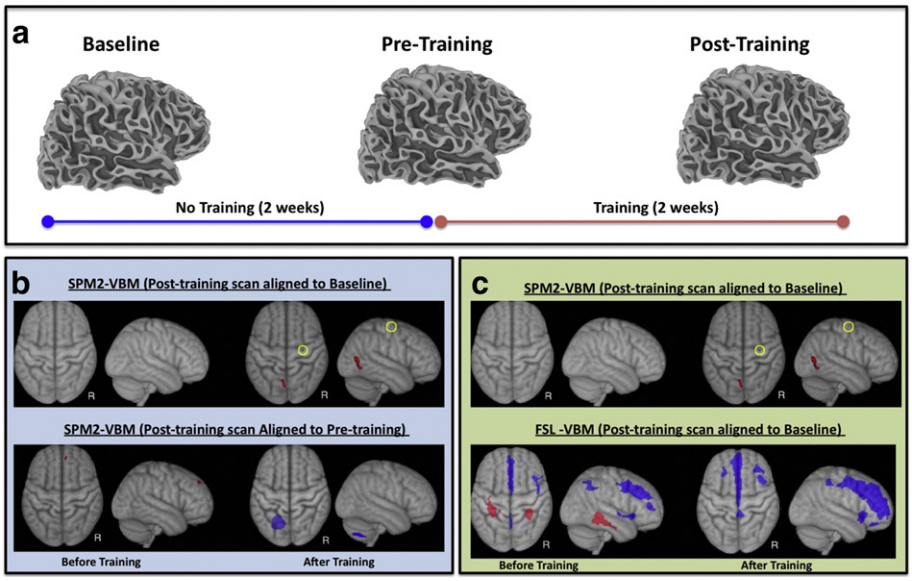

Fig. 2. Impact of analysis pipeline on measures of structural plasticity. (a) Schematic illustration of the within-subjects design used by Thomas et al. (2009). In addition to a pre-training scan, a baseline scan was acquired separated by the same interval as the training period and this served as the control scan. Statistical parametric maps from whole brain voxel-wise analyses indicated that the foci of the training-related gray matter changes vary based on: (b) the scan used as the target for spatial normalization and, (c) the type of VBM implementation used to analyze the data. Note that the bottom panel of (b) and (c) suggests significant gray matter changes between the baseline and pre-training scans even though participants were not given a task during this two-week control period.

noise ratio (SNR) in MRI techniques is generally sensitive to subject-related noise artifacts like head motion, and SNR in diffusion MRI sequences is especially sensitive to physiological effects (for a review see, Pierpaoli, 2011; Pierpaoli et al., 2003). These artifacts could manifest as significant differences between the training and control groups (Walker et al., 2011).

Among the 15 studies that included a control group/condition, only 3 employed a separate training task in the control condition. In two studies (Colcombe et al., 2006; Erickson et al., 2011a), the effect of aerobics exercise was investigated by comparing an aerobics group with an age-matched control group that practiced stretching exercises for the same training duration. Likewise, Tang et al. (2010) assigned different meditation techniques to the training and the control group and tested whether a specific type of meditation technique can induce significant structural changes.

Thus, only three studies include a trained control group and potentially provide any evidence for task-specific training-dependent structural changes rather than general effects of any form of training. However, it is worth noting that Tang and colleagues performed the group by time point interaction on voxels pre-selected for a training effect in one group of subjects and therefore the robustness of the results is unclear (see above). In principle, data could be collected both to determine the reliability of the structural measure in the absence of any intervention (two scans prior to training as in, Thomas et al., 2009) and to determine the task-specificity of any effects (two groups trained on separate tasks), although we found no study that adopted such an approach.

Brain region

All of the studies reporting training-dependent structural changes describe effects that are found in some brain regions and not others. Understanding the distribution and specificity of any structural changes is critical for interpreting the results. However, in the same way that a direct comparison between trained and untrained groups is necessary to show specificity to training, direct comparison between brain regions is also necessary to show anatomical specificity. Without any direct test between regions, it is unclear whether any highlighted regions show a change over time that is *significantly* different from other regions. With few exceptions (Erickson et al., 2011a; Lövdén et al., 2010), the vast majority of the studies we reviewed simply present statistical parametric maps of the differences between the pre- and post-training scans. However, the mass univariate approach cannot address regional specificity (Chumbley et al., 2010). In this type of analysis, specificity is evidenced as a significant effect in some voxels compared with non-significant effects in other voxels. This is equivalent to testing for the main effects in each set of voxels independently without ever directly testing for the interaction, and as discussed elsewhere (Nieuwenhuis et al., 2011), is weak evidence for anatomical specificity. Further, some studies have reported differences in the temporal profile of training effects in different brain regions (Draganski et al., 2006; Taubert et al., 2010). While the plots of signal change over time are qualitatively different between the regions, no direct test of the difference was performed.

Importantly, any test of anatomical specificity needs to avoid the statistical bias introduced by pre-selecting voxels for a specific effect and then testing the anatomical specificity using the same data (Kriegeskorte et al., 2009). This requires independent data for selection of voxels and tests of anatomical specificity. This could be achieved by pre-defining ROIs based on anatomy or prior studies. Alternatively, multiple data sets could be collected in each subject and part of the data used to determine regions showing training effects and the remaining data used to test anatomical specificity. Finally, in existing or published data sets this could be achieved by performing a split-half analysis, splitting the data over subjects (Kriegeskorte et al., 2009; Poldrack and Mumford, 2009).

The problem of comparing across brain regions is particularly challenging when there are no hypothesis-driven predictions. Having such predictions allows *a priori* selection of test and control regions. For example, in a recent study (Landi et al., 2011), subjects were trained to perform a complex visuo-motor task using their right hand only, with the prediction that any training-dependent structural change should be observed primarily in the left motor cortex. Consistent with this prediction, training was reported to evoke significant functional and structural changes in the left motor cortex. While this confirmation of the original prediction is indeed suggestive of specificity to left motor cortex, right motor cortex is an obvious control region, and a direct test between left and right motor cortex

would confirm such specificity. In this case, however, no direct test of anatomical specificity was reported. In principle, the strongest evidence would have been to demonstrate that a second group trained to perform the same task with their left hand showed the opposite pattern of results with structural changes in the right but not left hemisphere. Such hemisphere-specific training is often used in animal model studies, yielding evidence that is highly compelling (Chang and Greenough, 1982; Xu et al., 2009b).

In some cases, strong hypotheses have been generated from the animal literature. For example, taking into account the evidence supporting exercise induced neurogenesis and angiogenesis in the dentate gyrus of the adult hippocampus (Cotman and Berchtold, 2002; Pereira et al., 2007; Van Praag et al., 1999, 2005), Erickson et al. (2011a) tested whether training in aerobics compared with stretching exercises could induce structural changes in the hippocampus of elderly human subjects. They used a semi-automatic segmentation procedure to compute volume changes in both anterior (including the dentate gyrus) and posterior hippocampus. Consistent with their predictions, a relatively small (~2%), but significant increase in the volume of the bilateral anterior hippocampi was observed only in the aerobics group (Erickson et al., 2011a). More importantly, a direct statistical comparison (i.e. interaction between group, time point and ROI) revealed that the aerobics-related changes in volume were stronger in the anterior hippocampus than in the posterior hippocampus.

In sum, very few studies have rigorously tested the regional specificity of any structural changes they report. However, demonstrating such specificity provides compelling evidence for structural changes and is critical for understanding the effects of training and the underlying mechanisms.

Replicability

Three-ball juggling has been used as the training task in four separate studies (see Table 1) allowing us to evaluate the replicability, of the structural changes reported (Fig. 3). In the first of these studies, Draganski et al. (2004) reported changes in GMD bilaterally in the middle temporal region (visual motion area hMT/V5 — although hMT was not functionally localized in this or later studies) and near the left posterior intra-parietal sulcus (IPS). Further, the increases observed with training over 3 months declined over the following 3 months without practice. However, as is evident from Fig. 3, subsequent studies from the same research group showed limited replication of the location of the initially reported training-related peak changes (Thomas and Baker, 2012) and additionally revealed several regions not directly associated with processing visual motion including nucleus accumbens, hippocampus, as well as frontal and cingulate cortex. A recent conjunction analysis of data from three of these studies (Boyke et al., 2008; Draganski et al., 2004; Driemeyer et al., 2008), using an image based rather than co-ordinate based approach (Salimi-Khorshidi et al., 2009), suggests that the most consistent effect of training in juggling was a change in GMD close to the expected location of right hMT/V5 (May and Gaser, 2012; Thomas and Baker, 2012).

In contrast to these reports, the most recent juggling study (Scholz et al., 2009) reported a significant training-related increase (~4%) in GMD in the medial occipital and parietal lobe as well as a significant increase (~5%) in white matter integrity near the right posterior IPS. Importantly, Scholz and colleagues found no evidence for structural changes in or around hMT or the occipito-temporal cortex, even when the statistical criteria for significance were relaxed. In addition, the time course of the structural change was strikingly different. While the earlier juggling studies (Boyke et al., 2008; Draganski et al., 2004) reported effects that decreased after the training ceased, Scholz and colleagues reported increases in GMD and FA even after practice was suspended. Thus, while four separate studies have used a similar training paradigm and analyses, there is limited consistency across studies with regard to the location (Fig. 3) or the temporal profile of training-dependent structural changes in the adult human brain.

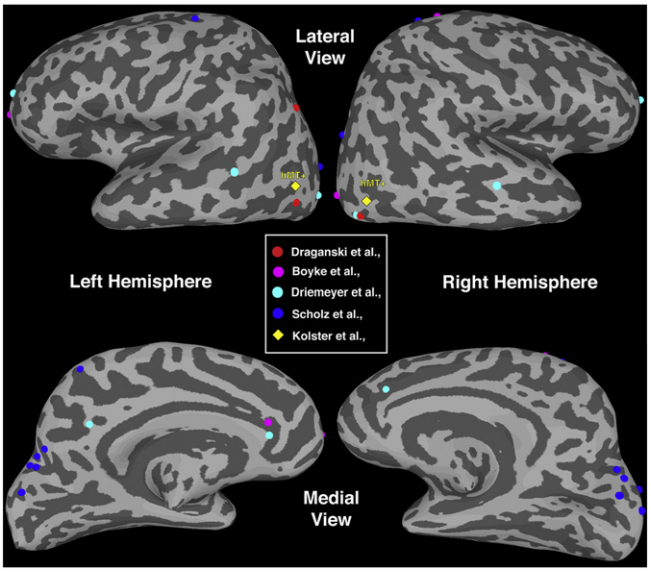

Fig. 3. Replicability of structural changes. Variability in the location of cortical regions showing significant structural changes from four studies reporting changes in gray matter following training in a three-ball juggling task. The colored circles indicate the peak voxel co-ordinates as per the Montreal Neurological Institute (MNI) atlas reported in the respective studies and are plotted on an inflated cortical surface map created by averaging 27 scans of a single subject (Saad et al., 2004). Apart from the two peaks in the right occipitotemporal cortex, there is little consistency in the reported locations of structural plasticity across studies. The yellow diamond represents the average peak MNI coordinate of area hMT functionally defined in a group of 11 healthy volunteers (Kolster et al., 2010). The regions in occipitotemporal cortex reported as corresponding to hMT (e.g. Draganski et al., 2004), tend to be posterior and largely inferior to the functionally-defined hMT. Note: The subcortical co-ordinates reported in Boyke et al. (2008) and Driemeyer et al. (2008) have not been plotted. All MNI coordinates have been remapped to correspond to the closest vertex on the gray-white matter boundary (average error<1 mm).

Besides juggling, training in working memory (Lövdén et al., 2010; Takeuchi et al., 2010) and aerobic exercise (Colcombe et al., 2006; Erickson et al., 2011a) are the only other paradigms that have been tested in multiple studies. However, due to differences in the training protocols, morphometric techniques used to measure structure, and the power of the analytical approaches, it is difficult to assess replicability in these studies.

Overall, across all the studies that have reported training-dependent structural changes, there is limited evidence for replicability, the most fundamental criterion for the robustness of an effect. While the weak replicability across the juggling studies could in principle reflect real differences, it is important to take into account the concerns about the robustness of the statistical analyses in these studies described above.

Correlation with behavior

Individual variations in GMD and white matter integrity have been reported to account for variance in behavioral measures (for a review, see Kanai and Rees, 2011). If the reported structural changes are the direct result of training, it seems reasonable to suppose that the change in structure should correlate with some measure of training behavior. Such a correlation is not necessary to conclude training-dependent structural changes but would significantly bolster support for this conclusion, and has been reported in some animal studies. For example, dendritic spine formation was found to correlate with the number of successful reaches in mice trained to reach for a food reward (Xu et al., 2009b). However, of the 14 studies

that tested for correlations between behavior and structural changes, only half reported evidence for significant effects.

While the original juggling study (Draganski et al., 2004) reported a close relationship between structural changes and juggling performance, no data or statistics were provided to support this assertion. Further, none of the later juggling studies found any significant correlation between change in brain structure and improvement in juggling performance (Boyke et al., 2008; Driemeyer et al., 2008; Scholz et al., 2009). A similar lack of correlation between behavioral and structural change was reported in three other studies using a range of different training tasks (Draganski et al., 2006; Lövdén et al., 2010; Schmidt-Wilcke et al., 2010). However, such negative results could simply reflect the fact that it is unclear which behavioral measure should correlate with the structural change (e.g. absolute performance at the end of training, relative task improvement, amount of practice).

Significant correlations between behavioral measures and apparent structural changes have been reported in seven studies. However, in evaluating the strength of these results, there are three issues worth considering. First, a general concern addressed by some (Landi et al., 2011) but not others (for example, Bezzola et al., 2011; Engvig et al., 2010, 2011) is that when multiple behavioral measures, multiple structural measures (e.g. FA, radial diffusivity, axial diffusivity) and/or multiple brain regions are tested, the multiple comparisons need to be taken into account to make sure that any effects are not spurious.

Second, in cases where behavioral measures were collected in both training and control groups (Engvig et al., 2010, 2011; Erickson et al., 2011a), the specificity of the correlation should be tested i.e., whether there is significantly stronger correlation between structural changes and behavior in the training compared with the control group.

Finally, it is unclear how to interpret differences in regions that show a significant effect of training on structural measures and those that show a correlation between behavior and structural changes. For example, following training on a dynamic balance beam task, whole brain analyses revealed one set of regions showing significant changes in VBM and diffusivity measures and a separate set of regions showing correlation between changes in GMV or diffusivity and behavior, with limited overlap between the regions (Taubert et al., 2010). Similarly, it is worth considering what it means if only a subset of the regions showing structural changes show a correlation with behavior (Bezzola et al., 2011; Takeuchi et al., 2010).

Overall, only half the studies that tested for correlations between behavior and structural changes reported any significant effects. While those studies reporting significant correlations appear to provide strong support for training-dependent structural changes, these results should be considered carefully in light of the concerns we have raised above.

Robustness of evidence

So far our review of the strength of the existing evidence for training-dependent structural plasticity in adult humans reveals a number of limitations. In many cases, the statistical evidence does not appear to be robust and the evidence for replication is weak. Out of 20 studies, only one (Erickson et al., 2011a) demonstrates effects that are specific to training, task and brain region, with a significant correlation with behavioral performance (but see, Coen et al., 2011; Erickson et al., 2011b). Specifically, anterior, but not posterior, hippocampal volume was found to increase in elderly subjects trained to perform aerobic exercise compared with subjects performing stretching exercises. Further, the changes in hippocampal volume correlated with improvements in a spatial memory task. However, replication of this result would provide the strongest support for structural changes. While this study provides strong evidence for structural plasticity, the pre-defined ROI approach is limited by the focus on specific regions (Friston et al., 2006) (but see, Saxe et al., 2006) and does not capitalize on the ability to image the whole brain with MRI (but see, Colcombe et al., 2006).

We do not mean to suggest that the other studies reporting structural changes are invalid or provide no evidence for training-dependent structural plasticity, just that the strength of the evidence is limited and alternative interpretations of the apparent effects are possible. In principle, the concerns we raise could be addressed by re-analysis of already published data, providing stronger statistical support for the conclusions.

Besides the issues associated with experimental design and statistical analysis discussed above, it is also important to be mindful of the limitations and challenges inherent to MRI-based imaging techniques. In the following section, we examine in more detail whether the results concerning training-dependent structural changes could have been influenced by noise in the data and by biases introduced while processing the MRI structural data.

Challenges in MRI based structural imaging

As noted earlier, unlike the structural imaging techniques used in animal studies, MRI-based techniques have relatively poor spatial resolution. Moreover, in both techniques the raw data undergoes several stages of processing to provide a measure related to the underlying biological structure. While these procedures are employed in order to reduce the effect of various sources of noise and to improve statistical inference, they make many assumptions and may introduce specific biases into the data. Details of the limitations of techniques like VBM and DTI have been discussed elsewhere (Ashburner and Friston, 2001; Bookstein, 2001; Crum et al., 2003; Davatzikos, 2004; Pierpaoli, 2011), but here we will provide a brief overview of some of the major concerns.

Signal-to-noise

The MR signal, which forms the basis for the structural images of the brain, is corrupted by various sources of noise that originate from the scanner (e.g. signal dropouts, eddy current distortions, susceptibility artifacts) or the subject (e.g. head motion, cardiac pulsation, respiration). These artifacts have a greater impact on DTI-based analysis (Basser and Jones, 2002; Pierpaoli, 2011) compared with VBM-based analysis because a reduction in SNR in diffusion MRI scan results in an artifactual increase in tensor-derived measures like FA (Farrell et al., 2007; Pierpaoli and Basser, 1996), the most widely used measure of microstructural integrity of white matter.

In the seven DTI studies we reviewed here, nearly all corrected for both head motion and eddy current distortions, although one study appeared to have corrected for head motion only (Taubert et al., 2010), and one did not report performing any corrections (Takeuchi et al., 2010). Importantly, none of the studies employed the correction techniques necessary for reducing the impact of physiological artifacts such as cardiac pulsation (Pierpaoli et al., 2003), which has been reported to cause significant artifactual changes in FA in several brain regions (Fig. 4). Such outlier data points can be identified and removed from further analysis (Chang et al., 2005; Mangin et al., 2002). If not, they can manifest as significant group differences (Walker et al., 2011) that are unrelated to the experimental manipulation. For example, the changes in FA reported in participants who took part in integrative body-mind training rather than relaxation therapy (Tang et al., 2010) could in principle arise from physiological changes (e.g. cardiac or respiration) induced in one of the groups following the training, rather than the specific meditation technique.

In addition to applying appropriate corrections to the data, it may be prudent to use measures of diffusivity such as Trace or Mean Diffusivity (MD) rather than just FA, since such raw measures of diffusivity are more tolerant to changes in SNR (Farrell et al., 2007; Marenco et al., 2006; Pierpaoli and Basser, 1996) and could potentially provide

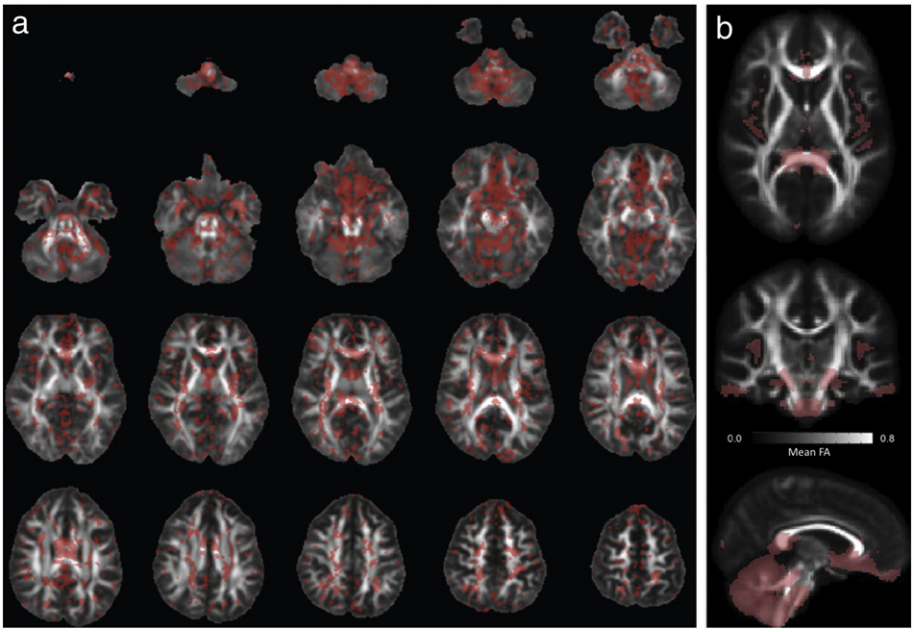

Fig. 4. Impact of physiology on diffusivity measures. (a) Axial images of fractional anisotropy (FA) from a single healthy volunteer superimposed with red areas indicating regions where FA is most likely to be affected by cardiac pulsation (data from Pierpaoli et al., 2003). (b) From top to bottom: axial, coronal, and mid-sagittal slices of the mean FA map based on 40 healthy adult volunteers (data from Walker et al., 2011). The pink spots indicate brain region where the FA outlier rejection probability is significantly high.

different assays of microstructural changes in white matter (Beaulieu, 2011). It is worth noting that only three of the seven diffusion MRI studies we reviewed here (Engvig et al., 2011; Lövdén et al., 2010; Taubert et al., 2010) tested for consistency in the training-related structural changes across different diffusivity measures.

Spatial normalization

A fundamental requirement for longitudinal studies of structural plasticity is that a voxel in a specific location in the participant's brain at time-point 1 should be aligned with the same location imaged at time-point 2. This alignment is complicated by differences in factors such as, the subject's head position and the scanner temperature across the two time-points. Further, in group analyses, each subjects' brain needs to be aligned with each other. This alignment across sessions and subjects is typically achieved by spatially normalizing to a standardized template allowing voxel-wise statistical analyses to be conducted. The assumptions made in these procedures have been a matter of debate (see for example, Ashburner and Friston, 2001; Bookstein, 2001) and although many of the concerns have been addressed in recent years (Klein et al., 2009), it is important to note that the choices made regarding spatial normalization can have a significant impact on the results. For example, with very few exceptions (Engvig et al., 2011; Landi et al., 2011; Scholz et al., 2009), most of the studies reviewed here normalized the structural scans from time-point 2 to the scans from time-point 1. This procedure introduces different levels of interpolation-related changes and other biases in the two data sets (Smith et al., 2002), which can manifest as longitudinal differences. Indeed, as is evident in Fig. 2b, the results for longitudinal structural differences can depend on which of two pre-training scans were used as the source for spatial normalization. Importantly, when the scans in this study were aligned to the halfway point between the two pre-training scans, no significant differences due to training were observed (Thomas et al., 2009).

Smoothing

MRI images of brain structure show considerable non-normality or non-stationarity (Salmond et al., 2002) and therefore parametric cluster-size tests cannot be performed without violating assumptions regarding normality of the data and the variance of the residuals (Ashburner and Friston, 2000; Hayasaka and Nichols, 2003). This is partly remedied by smoothing the data with a Gaussian filter of some standard size. Although a filter with a kernel size approximately 2–3 times the voxel size of the data has been recommended (Worsley et al., 1992), some have pointed out (Jones et al., 2005) that this estimate was derived empirically for analyzing fMRI and PET data and there is very little justification for the same rule to be used for VBM or voxel-wise analyses of DTI data. In the studies we reviewed, some did not report the exact parameters of smoothing (Draganski et al., 2004, 2006) making it difficult to compare results across studies. The remaining studies used filter sizes ranging from 3 to 12 mm. Such an arbitrary use of filter sizes can be counter-productive for a proper estimation of the effect size. For example, in a test of structural differences between a patient group and a control group, the foci of structural differences in FA between the two groups varied as a function of the filter size (Jones et al., 2005). In addition, studies that have examined the rate of false positives in VBM analysis, as a function of smoothing size concluded that large smoothing kernels (i.e. filter size >12 mm) and stringent cluster thresholds ($p<0.001$) are necessary to reduce the rate of false positives to an acceptable level (Silver et al., 2011). However, given the microscopic scale of structural changes observed in the animal studies, the use of such large filters may come with the cost of smoothing away the variance due to the experimental manipulation. In the context of training-dependent structural changes then, a legitimate question is to what extent the magnitude and location of the structural change reported in many of the studies reviewed here are resistant to changes in smoothing parameters.

Assumptions underlying statistical analyses

As noted above, smoothing the data does not ensure non-stationarity since strict assumptions need to be satisfied (Hayasaka and Nichols, 2003). Accordingly, procedures for non-stationarity correction have been included in the VBM pipeline, and at low degrees of freedom (<30), nonparametric permutation based tests, an assumption-free approach, have been recommended as a conservative alternative (Hayasaka et al.,

2004; Nichols and Holmes, 2002; Silver et al., 2011). Some of the studies reviewed here (Engvig et al., 2011; Landi et al., 2011; Scholz et al., 2009; Tang et al., 2010; Thomas et al., 2009) adopted nonparametric tests to identify training related structural changes. However, among the studies that relied on parametric tests, only few corrected for non-stationarity (Ilg et al., 2008; Schmidt-Wilcke et al., 2010) and in one study (Taubert et al., 2010) the parametric test may have been anti-conservative due to the small sample size. Given that failure to correct for non-stationarity can produce false positives and negatives (Hayasaka et al., 2004; Silver et al., 2011) the evidence reported from such studies should be interpreted with caution.

In summary, at each stage of the data processing stream, from preprocessing to statistical analyses, significant biases can be introduced inadvertently, and these can give rise to spurious changes in brain structure. Specifically, if the necessary corrections are not adopted, and rigorous statistical methods are not employed, as is the case in many of the studies we reviewed, it is difficult to rule out the possibility that imaging artifacts could have influenced the results concerning training-dependent structural changes.

Interpreting the evidence from MRI-based structural imaging

So far we have considered the reliability of evidence for training-dependent changes in MRI measures of brain structure. However, a second critical question is what any changes in human MRI measures, such as cortical thickness, gray matter density or fractional anisotropy, might reflect in terms of the biological substrate and how they relate to the animal literature. Specifically, the T_1-weighted structural images that are used as the raw data for morphometric analysis such as cortical thickness, VBM, DBM, etc., do not offer quantitative information regarding the nature of the change in brain tissue. Although cortical thickness translates to a realistic measure of cortical morphometry that is precise (Lerch and Evans, 2005) and consistent with post mortem data (Fischl and Dale, 2000), the gray matter probability values used in VBM do not correlate with quantitative histological measures of neuronal density (Eriksson et al., 2009). Likewise, diffusion measures of white matter microstructure such as FA, could reflect any number of the properties of axonal fibers including, myelination, packing density, and diameter (Beaulieu, 2002, 2011). However, the data from two recent animal studies that used rigorous experimental design and a combination of high spatial resolution MRI and immunohistochemistry provide some insight into what the MRI-based measures might reflect.

In a DTI-based study, (Blumenfeld-Katzir et al., 2011) the structural correlates of spatial learning were examined in a group of adult rats relative to two control groups matched for different aspects of training. The DTI analyses revealed training-related structural changes in the dentate gyrus, piriform cortex, the posterior cingulate cortex, and the corpus callosum. The immunohistochemistry analyses revealed that changes in synapses and astrocytes may have contributed to the change in the diffusion measures in the dentate gyrus, whereas an increase in myelin was identified as the contributing factor for the increase in FA in the corpus callosum. Interestingly, only the change in the dentate gyrus was found to be consistent across all the age groups. In contrast, a significant change in FA in the corpus callosum was observed most prominently in the developing group.

Similarly, Lerch et al. (2011) used a converging approach to test whether specific types of learning (spatial *versus* non-spatial) could evoke structural changes in specific brain regions of adult mice. *Ex vivo* analysis of the MRI structural data indicated that spatial learning was associated with structural changes in the hippocampus, while non-spatial learning was associated with changes in the striatum. In addition, compared to an untrained control group, they also reported non-specific structural changes in the training groups in several regions across the neocortex. Finally, immunohistochemical analysis suggested that the structural change in the hippocampus and striatum could be attributed to remodeling of neuronal processes.

In summary, these two studies suggest that: (a) with sufficient spatial resolution and rigorous experimental design MRI-based techniques can be used to detect training-related structural changes in the adult brain; (b) the hippocampal complex may have a special status in propensity for training-dependent structural plasticity, but changes may extend to neocortical regions; (c) the primary structural change may be mediated by changes in synapses and astrocytes, remodeling of neuronal processes, with changes in myelination appearing to be age-dependent.

There are obvious limitations in employing similar methods to explicate the mechanisms underlying structural plasticity in humans. Nevertheless, MRI-based imaging methods are currently the only tools available to measure changes in brain structure *in vivo*. A major challenge is that structural changes in the adult brain may be subserved by subtle changes in microscopic structures like dendritic spines rather than large-scale structure remodeling of axonal tracts, for example. Further, learning a novel skill may be mediated by a multi-stage process (Dayan and Cohen, 2011) such that rapid skill learning is facilitated by an increase in spine density within an hour and stabilizes to pre-training levels within 2 weeks (Xu et al., 2009b), but further augmentation of learning may be supported by a consolidation phase and a slow learning phase (Karni et al., 1998) that may range from days to months contributing to improvements in performance (for a recent review see, Dayan and Cohen, 2011). Such slow changes over long periods of training can be mediated by changes in other cellular processes such as angiogenesis, myelination or axonal remodeling. However, the exact nature of the structural change may be constrained by the type of training task and the neuroanatomy. For example, cognitive tasks, such as learning for an exam, may involve changes at the synaptic level, whereas aerobics training may involve a combination of neurogenesis, synaptogenesis and angiogenesis in the hippocampus, but only synaptogenesis in the cortex.

Animal studies that combine immunohistochemistry analysis and different MRI techniques (Blumenfeld-Katzir et al., 2011; Lerch and Evans, 2005) will be critical to explicate the spatial and temporal profile of such a multi-stage process of structural plasticity and how such changes are reflected in the MRI signal. Moreover, by collecting quantitative T_1 and T_2 measures (Tofts, 2003) in addition to Diffusion MRI data, the biological basis of structural changes detected with one type of MRI-method can be understood better (for e.g. Draganski et al., 2011). Thus, by acquiring both diffusion MRI and multiple components T_1/T_2 relaxometry data across time-points (Deoni et al., 2008) one can test quantitatively whether the learning related changes in FA for example is related to changes in myelination or due to partial voluming. Such converging data will not only provide a compelling demonstration of structural plasticity, but will also offer insight into the underlying mechanisms.

Conclusion

Based on our review of the literature and the limitations of MRI-based measures of structure, we conclude that the current literature on training-dependent plasticity in adult humans does not provide unequivocal evidence for training-dependent structural changes and more rigorous experimentation and statistical testing is required. Of the 20 studies we reviewed here, only one (Erickson et al., 2011a) provides strong evidence for effects that are specific both to the training task and to particular brain regions. Furthermore, these results are consistent with prior animal studies using related training paradigms (for a review see, Van Praag, 2008). Given the sometimes-weak experimental design, anti-conservative statistical methods and the potential for methodological artifacts, the remaining studies do not provide compelling evidence. Many of the studies we reviewed

are suggestive of training-dependent structural changes but in the absence of rigorous statistical testing, and a clear description of the methods used to analyze the data (Ridgway et al., 2008), the results remain ambiguous, and so far have not been shown to replicate convincingly. One of the difficulties for assessing replication of spatial location is the paucity of reliable methods to assess the variability of spatial estimates (for reviews see, Lazar et al., 2002; Wager et al., 2007). Meta-analytic approaches can be used but only if the full data are made available (Salimi-Khorshidi et al., 2009) (but see, Kang et al., 2011; Xu et al., 2009a).

Our review of this literature also highlights how two prominent statistical concerns, namely circular or non-independent analyses (Kriegeskorte et al., 2009; Vul et al., 2009) and failure to test appropriately for interactions (Nieuwenhuis et al., 2011), can impact a specific domain of research. However, there are two well-established procedures to avoid circular analyses that are worth considering. First, if there are *a priori* hypotheses regarding the brain regions expected to change with training, the ROIs can be defined independently of the experimental data and statistical tests can be restricted to the independently defined set of voxels. Alternatively, in an exploratory study where there may not be any strong hypotheses regarding the location of structural changes, unbiased estimates of magnitude and effect size can be obtained by performing a split-half analysis (Kriegeskorte et al., 2009; Poldrack and Mumford, 2009). In this type of analysis, the data from part of the sample can be used to define the ROIs and subsequent analysis and characterization of the effects can be performed in these independently defined ROIs using the remaining data.

Despite the numerous limitations of the studies reviewed here, we do not mean to suggest that the training protocols used in the human studies do not induce structural changes in the adult brain, nor do we mean to imply that MRI cannot be used to detect training-dependent structural changes in the adult brain. On the contrary, the two animal studies (Blumenfeld-Katzir et al., 2011; Lerch et al., 2011) described earlier provided compelling evidence for the feasibility of MRI-based techniques to detect training-related structural changes in adult animals. Further, *in vivo* imaging of structural plasticity in adults is of critical importance to characterize the mechanisms governing neural plasticity and also to develop rehabilitation strategies that facilitate structural plasticity in patients with brain injury.

The challenge moving forward is to address the limitations of the existing studies and present the strongest possible evidence for training-dependent structural plasticity in the adult human brain. In terms of design, we would argue that there should always be a control group (matched for factors such as age, gender and IQ) and this group should be engaged in a separate trained task to equate participants' overall experience as much as possible. In terms of analysis, direct comparisons must be made between the experimental and control groups. Above all, a rigorous statistical approach should be adopted — this may require larger data sets than previously collected to allow for independent estimates of effect size and testing of anatomical specificity. Only once robust data are presented, will it be possible to tease apart the contributions of different factors such as duration of training, age, gender, or intensity, complexity or novelty of the training task, which are all likely to impact the magnitude, nature or time course of any plasticity.

Structural MRI is a powerful tool to investigate brain plasticity and with the same sort of advances that revolutionized functional MRI (in hardware, analytical techniques, and understanding of the neurobiological correlates), it offers the potential for exciting and important insights.

Acknowledgments

This work was supported by the Intramural Research Program of NIMH. Support for this work also included funding from Department of Defense in the Center for Neuroscience and Regenerative Medicine. We thank Alan Koretsky, Sean Marrett, Alex Martin, Carlo Pierpaoli, Adam Thomas, Marta Ceko, and members of the Laboratory of Brain and Cognition, National Institute of Mental Health, for their helpful comments and discussion. Special thanks to Ziad Saad for the help in producing Fig. 3.

References

Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry — the methods. NeuroImage 11, 805–821.
Ashburner, J., Friston, K.J., 2001. Why voxel-based morphometry should be used. NeuroImage 14, 1238–1243.
Baker, C.I., Hutchison, T.L., Kanwisher, N., 2007. Does the fusiform face area contain subregions highly selective for nonfaces? Nat. Neurosci. 10, 3–4.
Basser, P.J., Jones, D.K., 2002. Diffusion tensor MRI: theory, experimental design and data analysis — a technical review. NMR Biomed. 15, 456–467.
Basser, P.J., Pierpaoli, C., 1996. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. J Magn Reson B 111, 209–219.
Basser, P.J., Mattiello, J., LeBihan, D., 1994. MR diffusion tensor spectroscopy and imaging. Biophys. J. 66, 259–267.
Beaulieu, C., 2002. The basis of anisotropic water diffusion in the nervous system — a technical review. NMR Biomed. 15, 435–455.
Beaulieu, C., 2011. What Makes Diffusion Anisotropic in the Nervous System? Diffusion MRI Oxford Publications, pp. 92–109.
Bengtsson, S., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., Ullén, F., 2005. Extensive piano practicing has regionally specific effects on white matter development. Nat. Neurosci. 8, 1148–1150.
Bezzola, L., Merillat, S., Gaser, C., Jäncke, L., 2011. Training-induced neural plasticity in golf novices. J. Neurosci. 31, 12444–12448.
Black, J.E., Isaacs, K.R., Anderson, B.J., Alcantara, A.A., Greenough, W.T., 1990. Learning causes synaptogenesis, whereas motor activity causes angiogenesis, in cerebellar cortex of adult rats. Proc. Natl. Acad. Sci. U. S. A. 87, 5568–5572.
Blumenfeld-Katzir, T., Pasternak, O., Dagan, M., Assaf, Y., 2011. Diffusion MRI of structural brain plasticity induced by a learning and memory task. PLoS One 6, e20678.
Bookstein, F.L., 2001. "Voxel-based morphometry" should not be used with imperfectly registered images. NeuroImage 14, 1454–1462.
Boyke, J., Driemeyer, J., Gaser, C., Buchel, C., May, A., 2008. Training-induced brain structure changes in the elderly. J. Neurosci. 28, 7031–7035.
Briones, T.L., Klintsova, A.Y., Greenough, W.T., 2004. Stability of synaptic plasticity in the adult rat visual cortex induced by complex environment exposure. Brain Res. 1018, 130–135.
Chang, F.L.F., Greenough, W.T., 1982. Lateralized effects of monocular training on dendritic branching in adult split-brain rats. Brain Res. 232, 283–292.
Chang, L.C., Jones, D.K., Pierpaoli, C., 2005. RESTORE: robust estimation of tensors by outlier rejection. Magn. Reson. Med. 53, 1088–1095.
Chen, J.L., Nedivi, E., 2010. Neuronal structural remodeling: is it all about access? Curr. Opin. Neurobiol. 20, 557–562.
Chumbley, J., Flandin, G., Seghier, M., Friston, K., 2010. Multinomial inference on distributed responses in SPM. NeuroImage 53, 161–170.
Chung, M., Worsley, K., Paus, T., Cherif, C., Collins, D., Giedd, J., Rapoport, J., Evans, A., 2001. A unified statistical approach to deformation-based morphometry. NeuroImage 14, 595–606.
Coen, R.F., Lawlor, B.A., Kenny, R.A., 2011. Failure to demonstrate that memory improvement is due either to aerobic exercise or increased hippocampal volume. Proc. Natl. Acad. Sci. 108, E89.
Colcombe, S.J., Erickson, K.I., Scalf, P.E., Kim, J.S., Prakash, R., McAuley, E., Elavsky, S., Marquez, D.X., Hu, L., Kramer, A.F., 2006. Aerobic exercise training increases brain volume in aging humans. J. Gerontol. A Biol. Sci. Med. Sci. 61, 1166–1170.
Cotman, C.W., Berchtold, N.C., 2002. Exercise: a behavioral intervention to enhance brain health and plasticity. Trends Neurosci. 25, 295–301.
Crum, W., Griffin, L., Hill, D., Hawkes, D., 2003. Zen and the art of medical image registration: correspondence, homology, and quality. NeuroImage 20, 1425–1437.
Davatzikos, C., 2004. Why voxel-based morphometric analysis should be used with great caution when characterizing group differences. NeuroImage 23, 17–20.
Dayan, E., Cohen, L.G., 2011. Neuroplasticity subserving motor skill learning. Neuron 72, 443–454.
Demerens, C., Stankoff, B., Logak, M., Anglade, P., Allinquant, B., Couraud, F., Zalc, B., Lubetzki, C., 1996. Induction of myelination in the central nervous system by electrical activity. Proc. Natl. Acad. Sci. U. S. A. 93, 9887–9892.
Deoni, S.C.L., Rutt, B.K., Arun, T., Pierpaoli, C., Jones, D.K., 2008. Gleaning multicomponent T1 and T2 information from steady-state imaging data. Magn. Reson. Med. 60, 1372–1387.
Draganski, B., May, A., 2008. Training-induced structural changes in the adult human brain. Behav. Brain Res. 192, 137–142.
Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., May, A., 2004. Neuroplasticity: changes in grey matter induced by training. Nature 427, 311–312.
Draganski, B., Gaser, C., Kempermann, G., Kuhn, H., Winkler, J., Buchel, C., May, A., 2006. Temporal and spatial dynamics of brain structure changes during extensive learning. J. Neurosci. 26, 6314–6317.
Draganski, B., Ashburner, J., Hutton, C., Kherif, F., Frackowiak, R., Helms, G., Weiskopf, N., 2011. Regional specificity of MRI contrast parameter changes in normal ageing revealed by voxel-based quantification (VBQ). NeuroImage 55, 1423–1434.

Driemeyer, J., Boyke, J., Gaser, C., Büchel, C., May, A., 2008. Changes in gray matter induced by learning — revisited. PLoS One 3, e2669.
Engvig, A., Fjell, A.M., Westlye, L.T., Moberget, T., Sundseth, O., Larsen, V.A., Walhovd, K.B., 2010. Effects of memory training on cortical thickness in the elderly. NeuroImage 52, 1667–1676.
Engvig, A., Fjell, A.M., Westlye, L.T., Moberget, T., Sundseth, Y., Larsen, V.A., Walhovd, K.B., 2011. Memory training impacts short term changes in aging white matter: a longitudinal diffusion tensor imaging study. Hum. Brain Mapp. 32, 11–12.
Erickson, K.I., Voss, M.W., Prakash, R.S., Basak, C., Szabo, A., Chaddock, L., Kim, J.S., Heo, S., Alves, H., White, S.M., 2011a. Exercise training increases size of hippocampus and improves memory. Proc. Natl. Acad. Sci. U. S. A. 108, 3017–3022.
Erickson, K., Voss, M., Prakash, R., Basak, C., Szabo, A., Chaddock, L., White, S., Wojcicki, T., Mailey, E., McAuley, E., 2011b. Reply to Coen et al.: exercise, hippocampal volume, and memory. Proc. Natl. Acad. Sci. 108, E90.
Eriksson, S., Free, S., Thom, M., Symms, M., Martinian, L., Duncan, J., Sisodiya, S., 2009. Quantitative grey matter histological measures do not correlate with grey matter probability values from *in vivo* MRI in the temporal lobe. J. Neurosci. Methods 181, 111–118.
Farrell, J.A.D., Landman, B.A., Jones, C.K., Smith, S.A., Prince, J.L., van Zijl, P., Mori, S., 2007. Effects of signal-to-noise ratio on the accuracy and reproducibility of diffusion tensor imaging-derived fractional anisotropy, mean diffusivity, and principal eigenvector measurements at 1.5 T. J. Magn. Reson. Imaging 26, 756–767.
Fischl, B., Dale, A.M., 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc. Natl. Acad. Sci. U. S. A. 97, 11050–11055.
Friston, K., Rotshtein, P., Geng, J., Sterzer, P., Henson, R., 2006. A critique of functional localisers. NeuroImage 30, 1077–1087.
Fu, M., Zuo, Y., 2011. Experience-dependent structural plasticity in the cortex. Trends Neurosci. 34, 177–187.
Hayasaka, S., Nichols, T.E., 2003. Validating cluster size inference: random field and permutation methods. NeuroImage 20, 2343–2356.
Hayasaka, S., Phan, K.L., Liberzon, I., Worsley, K.J., Nichols, T.E., 2004. Nonstationary cluster-size inference with random field and permutation methods. NeuroImage 22, 676–687.
Holtmaat, A., Svoboda, K., 2009. Experience-dependent structural synaptic plasticity in the mammalian brain. Nat. Rev. Neurosci. 10, 647–658.
Holtmaat, A., Wilbrecht, L., Knott, G.W., Welker, E., Svoboda, K., 2006. Experience-dependent and cell-type-specific spine growth in the neocortex. Nature 441, 979–983.
Hyde, K.L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A.C., Schlaug, G., 2009. Musical training shapes structural brain development. J. Neurosci. 29, 3019–3025.
Ilg, R., Wohlschlager, A.M., Gaser, C., Liebau, Y., Dauner, R., Woller, A., Zimmer, C., Zihl, J., Muhlau, M., 2008. Gray matter increase induced by practice correlates with task-specific activation: a combined functional and morphometric magnetic resonance imaging study. J. Neurosci. 28, 4210–4215.
Isaacs, K.R., Anderson, B.J., Alcantara, A.A., Black, J.E., Greenough, W.T., 1992. Exercise and the brain: angiogenesis in the adult rat cerebellum after vigorous physical activity and motor skill learning. J. Cereb. Blood Flow Metab. 12, 110–119.
Jones, D.K., Symms, M.R., Cercignani, M., Howard, R.J., 2005. The effect of filter size on VBM analyses of DT-MRI data. NeuroImage 26, 546–554.
Juraska, J.M., Greenough, W.T., Elliott, C., Mack, K.J., Berkowitz, R., 1980. Plasticity in adult rat visual cortex: an examination of several cell populations after differential rearing. Behav. Neural. Biol. 29, 157–167.
Kanai, R., Rees, G., 2011. The structural basis of inter-individual differences in human behaviour and cognition. Nat. Rev. Neurosci. 12, 231–242.
Kang, J., Johnson, T.D., Nichols, T.E., Wager, T.D., 2011. Meta analysis of functional neuroimaging data via Bayesian spatial point processes. J. Am. Stat. Assoc. 106, 124–134.
Karni, A., Meyer, G., Rey-Hipolito, C., Jezzard, P., Adams, M.M., Turner, R., Ungerleider, L.G., 1998. The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. Proc. Natl. Acad. Sci. U. S. A. 95, 861–868.
Klauschen, F., Goldman, A., Barra, V., Meyer Lindenberg, A., Lundervold, A., 2009. Evaluation of automated brain MR image segmentation and volumetry methods. Hum. Brain Mapp. 30, 1310–1327.
Kleim, J.A., Cooper, N.R., VandenBerg, P.M., 2002. Exercise induces angiogenesis but does not alter movement representations within rat motor cortex. Brain Res. 934, 1–6.
Kleim, J.A., Markham, J.A., Vij, K., Freese, J.L., Ballard, D.H., Greenough, W.T., 2007. Motor learning induces astrocytic hypertrophy in the cerebellar cortex. Behav. Brain Res. 178, 244–249.
Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B., Chiang, M.C., Christensen, G.E., Collins, D.L., Gee, J., Hellier, P., 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. NeuroImage 46, 786–802.
Knott, G.W., Holtmaat, A., Wilbrecht, L., Welker, E., Svoboda, K., 2006. Spine growth precedes synapse formation in the adult neocortex *in vivo*. Nat. Neurosci. 9, 1117–1124.
Kolster, H., Peeters, R., Orban, G.A., 2010. The retinotopic organization of the human middle temporal area MT/V5 and its cortical neighbors. J. Neurosci. 30, 9801–9820.
Kriegeskorte, N., Simmons, W., Bellgowan, P., Baker, C., 2009. Circular analysis in systems neuroscience: the dangers of double dipping. Nat. Neurosci. 12, 535–540.
Kriegeskorte, N., Lindquist, M.A., Nichols, T.E., Poldrack, R.A., Vul, E., 2010. Everything you never wanted to know about circular analysis, but were afraid to ask. J. Cereb. Blood Flow Metab. 30, 1551–1557.
Kwok, V., Niu, Z., Kay, P., Zhou, K., Mo, L., Jin, Z., So, K.F., Tan, L.H., 2011. Learning new color names produces rapid increase in gray matter in the intact adult human cortex. Proc. Natl. Acad. Sci. U. S. A. 108, 6686–6688.
Landi, S.M., Baguear, F., Della-Maggiore, V., 2011. One week of motor adaptation induces structural changes in primary motor cortex that predict long-term memory one year later. J. Neurosci. 31, 11808–11813.
Lazar, N.A., Luna, B., Sweeney, J.A., Eddy, W.F., 2002. Combining brains: a survey of methods for statistical pooling of information. NeuroImage 16, 538–550.
Lerch, J.P., Evans, A.C., 2005. Cortical thickness analysis examined through power analysis and a population simulation. NeuroImage 24, 163–173.
Lerch, J.P., Yiu, A.P., Martinez-Canabal, A., Pekar, T., Bohbot, V.D., Frankland, P.W., Henkelman, R.M., Josselyn, S.A., Sled, J.G., 2011. Maze training in mice induces MRI detectable brain shape changes specific to the type of learning. NeuroImage 54, 2086–2095.
Lövdén, M., Bodammer, N.C., Kühn, S., Kaufmann, J., Schütze, H., Tempelmann, C., Heinze, H.J., Düzel, E., Schmiedek, F., 2010. Experience-dependent plasticity of white-matter microstructure extends into old age. Neuropsychologia 48, 3878–3883.
Maguire, E., Woollett, K., Spiers, H., 2006. London taxi drivers and bus drivers: a structural MRI and neuropsychological analysis. Hippocampus 16, 1091–1101.
Mangin, J.F., Poupon, C., Clark, C., Le Bihan, D., Bloch, I., 2002. Distortion correction and robust tensor estimation for MR diffusion imaging. Med. Image Anal. 6, 191–198.
Marenco, S., Rawlings, R., Rohde, G.K., Barnett, A.S., Honea, R.A., Pierpaoli, C., Weinberger, D.R., 2006. Regional distribution of measurement error in diffusion tensor imaging. Psychiatry Res. Neuroimaging 147, 69–78.
Markham, J.A., Herting, M.M., Luszpak, A.E., Juraska, J.M., Greenough, W.T., 2009. Myelination of the corpus callosum in male and female rats following complex environment housing during adulthood. Brain Res. 1288, 9–17.
May, A., 2011. Experience-dependent structural plasticity in the adult human brain. Trends Cogn. Sci. 15, 475–482.
May, A., Gaser, C., 2006. Magnetic resonance-based morphometry: a window into structural plasticity of the brain. Curr. Opin. Neurol. 19, 407–411.
May, A., Gaser, C., 2012. Response to Thomas and Baker: the structural adaptation of the brain to training. Trends Cogn. Sci. 16, 97–98.
McBride, T.J., Rodriguez-Contreras, A., Trinh, A., Bailey, R., Debello, W.M., 2008. Learning drives differential clustering of axodendritic contacts in the barn owl auditory system. J. Neurosci. 28, 6960–6973.
Mizrahi, A., Katz, L.C., 2003. Dendritic stability in the adult olfactory bulb. Nat. Neurosci. 6, 1201–1207.
Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. 15, 1–25.
Nieuwenhuis, S., Forstmann, B.U., Wagenmakers, E.J., 2011. Erroneous analyses of interactions in neuroscience: a problem of significance. Nat. Neurosci. 14, 1105–1107.
Pereira, A.C., Huddleston, D.E., Brickman, A.M., Sosunov, A.A., Hen, R., McKhann, G.M., Sloan, R., Gage, F.H., Brown, T.R., Small, S.A., 2007. An *in vivo* correlate of exercise-induced neurogenesis in the adult dentate gyrus. Proc. Natl. Acad. Sci. U. S. A. 104, 5638–5643.
Pierpaoli, C., 2011. Artifacts in Diffusion MRI. Diffusion MRI Oxford Publications, pp. 303–318.
Pierpaoli, C., Basser, P.J., 1996. Toward a quantitative assessment of diffusion anisotropy. Magn. Reson. Med. 36, 893–906.
Pierpaoli, C., Marenco, S., Rohde, G.K., Jones, D.K., Barnett, A., 2003. Analyzing the contribution of cardiac pulsation to the variability of quantities derived from the diffusion tensor. Proc. Int. Soc. Magn. Reson. Med. 11, 70.
Poldrack, R., 2000. Imaging brain plasticity: conceptual and methodological issues — a theoretical review. NeuroImage 12, 1–13.
Poldrack, R.A., Mumford, J.A., 2009. Independence in ROI analysis: where is the voodoo? Soc. Cogn. Affect. Neurosci. 4, 208–213.
Rakic, P., 2002. Neurogenesis in adult primate neocortex: an evaluation of the evidence. Nat. Rev. Neurosci. 3, 65–71.
Ridgway, G.R., Henley, S., Rohrer, J.D., Scahill, R.I., Warren, J.D., Fox, N.C., 2008. Ten simple rules for reporting voxel-based morphometry studies. NeuroImage 40, 1429–1435.
Saad, Z., Reynolds, R., Argall, B., Japee, S., Cox, R.W., 2004. SUMA: an interface for surface-based intra- and inter-subject analysis with AFNI. Proc. IEEE Int. Symp. Biomed. Imaging 1510–1513.
Salimi-Khorshidi, G., Smith, S.M., Keltner, J.R., Wager, T.D., Nichols, T.E., 2009. Meta-analysis of neuroimaging data: a comparison of image-based and coordinate-based pooling of studies. NeuroImage 45, 810–823.
Salmond, C., Ashburner, J., Vargha-Khadem, F., Connelly, A., Gadian, D., Friston, K., 2002. Distributional assumptions in voxel-based morphometry. NeuroImage 17, 1027–1030.
Saxe, R., Brett, M., Kanwisher, N., 2006. Divide and conquer: a defense of functional localizers. NeuroImage 30, 1088–1096.
Schmidt-Wilcke, T., Rosengarth, K., Luerding, R., Bogdahn, U., Greenlee, M., 2010. Distinct patterns of functional and structural neuroplasticity associated with learning Morse code. NeuroImage 51, 1234–1241.
Scholz, J., Klein, M., Behrens, T., Johansen-Berg, H., 2009. Training induces changes in white-matter architecture. Nat. Neurosci. 12, 1370–1371.
Silver, M., Montana, G., Nichols, T.E., 2011. False positives in neuroimaging genetics using voxel-based morphometry data. NeuroImage 54, 992–1000.
Smith, S.M., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P., Federico, A., De Stefano, N., 2002. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. NeuroImage 17, 479–489.
Stettler, D.D., Yamahachi, H., Li, W., Denk, W., Gilbert, C.D., 2006. Axons and synaptic boutons are highly dynamic in adult visual cortex. Neuron 49, 877–887.
Takeuchi, H., Sekiguchi, A., Taki, Y., Yokoyama, S., Yomogida, Y., Komuro, N., Yamanouchi, T., Suzuki, S., Kawashima, R., 2010. Training of working memory impacts structural connectivity. J. Neurosci. 30, 3297–3303.

Tang, Y., Lu, Q., Geng, X., Stein, E., Yang, Y., Posner, M., 2010. Short-term meditation induces white matter changes in the anterior cingulate. Proc. Natl. Acad. Sci. U. S. A. 107, 15649–15652.
Taubert, M., Draganski, B., Anwander, A., Muller, K., Horstmann, A., Villringer, A., Ragert, P., 2010. Dynamic properties of human brain structure: learning-related changes in cortical areas and associated fiber connections. J. Neurosci. 30, 11670–11677.
Thomas, C., Baker, C.I., 2012. Remodeling human cortex through training: comment on May. Trends Cogn. Sci. 16, 96–97.
Thomas, A., Marrett, S., Saad, Z., Ruff, D., Martin, A., Bandettini, P., 2009. Functional but not structural changes associated with learning: an exploration of longitudinal voxel-based morphometry (VBM). NeuroImage 48, 117–125.
Tofts, P., 2003. In: Tofts, Paul (Ed.), Quantitative MRI of the Brain: Measuring Changes Caused by Disease. John Wiley & Sons Ltd.
Trachtenberg, J., Chen, B., Knott, G., Feng, G., Sanes, J., Welker, E., Svoboda, K., 2002. Long-term *in vivo* imaging of experience-dependent synaptic plasticity in adult cortex. Nature 420, 788–794.
Van Praag, H., 2008. Neurogenesis and exercise: past and future directions. Neuromolecular Med. 10, 128–140.
Van Praag, H., Kempermann, G., Gage, F.H., 1999. Running increases cell proliferation and neurogenesis in the adult mouse dentate gyrus. Nat. Neurosci. 2, 266–270.
Van Praag, H., Shubert, T., Zhao, C., Gage, F.H., 2005. Exercise enhances learning and hippocampal neurogenesis in aged mice. J. Neurosci. 25, 8680–8685.
Vul, E., Harris, C., Winkielman, P., Pashler, H., 2009. Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. Perspect. Psychol. Sci. 4, 274.
Wager, T.D., Lindquist, M., Kaplan, L., 2007. Meta-analysis of functional neuroimaging data: current and future directions. Soc. Cogn. Affect. Neurosci. 2, 150.
Walker, L., Chang, L.C., Koay, C.G., Sharma, N., Cohen, L., Verma, R., Pierpaoli, C., 2011. Effects of physiological noise in population analysis of diffusion tensor MRI data. NeuroImage 54, 1168–1177.
Woollett, K., Maguire, E.A., 2011. Acquiring "the knowledge" of London's layout drives structural brain changes. Curr. Biol. 21, 2109–2114.
Worsley, K.J., Evans, A., Marrett, S., Neelin, P., 1992. A three-dimensional statistical analysis for CBF activation studies in human brain. J. Cereb. Blood Flow Metab. 12, 900–918.
Xu, L., Johnson, T.D., Nichols, T.E., Nee, D.E., 2009a. Modeling inter-subject variability in fMRI activation location: a Bayesian hierarchical spatial model. Biometrics 65, 1041–1051.
Xu, T., Yu, X., Perlik, A.J., Tobin, W.F., Zweig, J.A., Tennant, K., Jones, T., Zuo, Y., 2009b. Rapid formation and selective stabilization of synapses for enduring motor memories. Nature 462, 915–919.
Yamahachi, H., Marik, S.A., McManus, J.N.J., Denk, W., Gilbert, C.D., 2009. Rapid axonal sprouting and pruning accompany functional reorganization in primary visual cortex. Neuron 64, 719–729.