Published in final edited form as:

Neuroimage. 2018 July 15; 175: 100–110. doi:10.1016/j.neuroimage.2018.03.071.

Statistical power comparisons at 3T and 7T with a GO/NOGO task

Salvatore Torrisi¹, Gang Chen², Daniel Glen², Peter A. Bandettini³, Chris I. Baker⁴, Richard Reynolds², Jeffrey Yen-Ting Liu¹, Joseph Leshin¹, Nicholas Balderston¹, Christian Grillon¹, and Monique Ernst¹

¹Section on Neurobiology of Fear and Anxiety, NIMH, Bethesda MD

²Scientific and Statistical Computing Core, NIMH, Bethesda MD

³Section on Functional Imaging Methods, NIMH, Bethesda MD

⁴Section on Learning and Plasticity, NIMH, Bethesda MD

Abstract

The field of cognitive neuroscience is weighing evidence about whether to move from standard field strength to ultra-high field (UHF). The present study contributes to the evidence by comparing a cognitive neuroscience paradigm at 3 Tesla (3T) and 7 Tesla (7T). The goal was to test and demonstrate the practical effects of field strength on a standard GO/NOGO task using accessible preprocessing and analysis tools. Two independent matched healthy samples (N=31 each) were analyzed at 3T and 7T. Results show gains at 7T in statistical strength, the detection of smaller effects and group-level power. With an increased availability of UHF scanners, these gains may be exploited by cognitive neuroscientists and other neuroimaging researchers to develop more efficient or comprehensive experimental designs and, given the same sample size, achieve greater statistical power at 7T.

Keywords

7 Tesla; UHF; cognitive neuroscience; group studies

Introduction

Beyond benefits to clinical research and practice (Beisteiner et al., 2011; van der Kolk et al., 2013), functional MRI at ultra-high field (UHF, 7 Tesla) has much to offer cognitive neuroscience. As the number of human UHF scanners steadily increases across the world, more research centers and laboratories face the decision of whether to transition from standard 3 Tesla field strength. Such decisions will naturally be based on the benefits UHF offers to particular scientific questions and on whether those benefits outweigh

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

disadvantages. Although a body of important research comparing field strengths has been performed by experts in physics and engineering (see below), additional empirical evidence using neurocognitive tasks may also aid the cognitive neuroscientist's decision to invest in UHF.

Statistical power is critical in cognitive neuroscience research. It is the probability of rejecting a null hypothesis, given a true alternative, and motivates each experimental step, from choices about experimental design, preprocessing, analysis and the number of subjects to acquire for group studies. Based on physics principles alone, a higher field strength should provide greater fMRI signal and thereby higher statistical power. However, neuroimaging researchers, including neurosurgeons and psychiatric neuroimagers, are mindful of potential complications and may worry about undesirable trade-offs, worsening of common limitations, emergence of new issues, or increases in various kinds of noise. In addition, solutions to remediate these factors may not be widely known. Fortunately, UHF, as well as the computational infrastructure needed to handle its data, have advanced spectacularly in the last few years. While issues still exist, the technology has progressed to the point where UHF fMRI offers definite advantages to answering most, if not all cognitive neuroscientific questions. The purpose of the present work is to directly compare statistical power between 3T and 7T with a well-known task, using a standard analysis and sample size.

As stated, the present work is not without precedent. Researchers have already compared 3T and UHF (usually 7T) fMRI. For example, one benefit of 7T is that it allows the acquisition of higher resolution data and, therefore, less partial voluming of tissues of interest (Bandettini et al., 2012; Newton et al., 2012). However, smaller voxels also have intrinsically less image signal to noise ratio (SNR) (Edelstein et al., 1986), which perhaps makes the benefits of 7T fMRI initially not obvious. In actuality, for typical fMRI studies, the image SNR is secondary to temporal SNR (TSNR), the signal to noise of the time series (Murphy et al., 2007). At high image SNR, TSNR tends to be the determining parameter of sensitivity. TSNR values are influenced by different kinds of physiological and non-physiological noise, which in turn are affected by field strength (Krüger et al., 2001). Therefore, much research has focused on teasing apart these factors. Studies have investigated the relation of spatial resolution with TSNR (Triantafyllou et al., 2005; 2006; van der Zwaag et al., 2009), and with contrast to noise ratio (CNR) in specific regions of interest (Beisteiner et al., 2011; Sladky et al., 2013). Recent work has additionally outlined the effects on TSNR by technologies other than field strength, such as accelerated imaging and head coil design (Triantafyllou et al., 2011; Wald and Polimeni, 2017). Some of these studies are theoretically and empirically grounded but use unconventional, custom, or outdated acquisitions and analyses. Therefore, for the cognitive neuroscientist, it may also prove illustrative to test the effects of field strength on statistical power with state-of-the-art data and conventional analyses.

Specifically, this study investigates group-level statistical power differences in BOLD activation between 3T and 7T in two conditions of a classic response inhibition GO/NOGO paradigm. The GO condition probes the motor network during button pressing, and the NOGO condition probes response inhibition following the buildup of prepotent motor

response. Both the GO and NOGO conditions may therefore serve as general examples of sensorimotor and cognitive processes, respectively. Because “cognitive neuroscience” is a broad and multifarious term, the present work focuses on activation analyses using the general linear model (GLM), arguably the most common and robust approach to functional brain mapping (Poline and Brett, 2012).

Hypotheses were twofold: First, at group-level, both the effect estimates and their *t*-statistic values in known GO and NOGO regions would be larger at 7 Tesla than at 3 Tesla. This was expected to be observable in within- and between-field whole-brain statistical maps. Secondly, it was hypothesized that, using unbiased ROIs, power analyses between field strengths would result in a median fewer subjects at 7T needed to produce comparable effects at 3T.

Methods

Methods were shared between the 7T and 3T acquisitions and analyses, yet there were some necessary differences. The similarities and differences are described separately below.

Consistent Methods Between Field Strengths

Participants (inclusion/exclusion criteria)—Right-handed, healthy adult volunteers were recruited from a mixed urban and suburban population through Internet listservs, flyers and print advertisements. Exclusion criteria included the following: (a) current or past Axis I psychiatric disorder as assessed by SCID-I/NP (First et al., 2007), (b) first-degree relative with a known psychotic disorder, (c) brain abnormality on MRI as assessed by a radiologist, (d) positive toxicology screen, (e) MRI contraindication, or (f) excessive head motion during the functional scans. Head motion exclusion criteria were identical between studies: subjects with more than 10% of their TRs censored were excluded, and censoring on a data point was defined as a between-TR movement of >0.3mm Euclidean norm distance. Of the subjects that were included in the present work, there was no statistically significant difference between the average fraction of TRs censored in both runs at 3T vs 7T (averages(SD)% of 3T/7T = 3.1(2.4)/2.1(2.4); $p=0.10$) or the absolute number of censored TRs (averages(SD) of 3T/7T = 24.4(19)/20.0(21); $p=0.38$). Written, informed consent was approved by the National Institute of Mental Health Combined Neuroscience Institutional Review Board and obtained from all subjects. The number of subjects was $n = 31$ for both analyses.

Task design and behavior—The event-related task employed was the Sustained Attention to Response Task (SART) (Robertson et al., 1997), which is essentially a GO/NOGO paradigm (91% GO trials with the “=” symbol indicating button push and 9% NOGO trials with the “o” symbol indicating no push). The high proportion of GO vs. NOGO trials builds a prepotent response that is challenging to overcome. The same number of trials for each condition were in both analyses: 574 GOs and 54 NOGOs.

Of note, this data was originally collected for different analyses from those reported here, e.g. (Torrìsi et al., 2016). A threat of unpredictable electric shock procedure was also “overlaid” on the SART, based on a translational anxiety induction paradigm (Schmitz and Grillon, 2012). Overall, the task consisted of six, 2-minute periods of threat of shock vs.

safety from shock (condition indicated by colored border around stimuli). Randomly-timed shocks were 1 per run (2 runs total). Because only main effects of GO and NOGO trials are reported in the present manuscript, collapsing across threat and safe conditions, the threat component of the experiment will not be discussed and main effects will simply be called “effects”.

Behavioral performance was also quantified. First, NOGO and GO accuracies were averaged across threat/safe conditions for main effects. GO reaction times were also averaged across threat/safe conditions. These NOGO and GO behavioral effects were then separately compared between field strength cohorts with two-tailed, two-sample *t*-tests.

Data Acquisition—Both whole brain, gradient EPI functional acquisitions were collected with Siemens scanners, 32-channel head coils, a flip angle of 70° and TR of 2 seconds. Data for both datasets were acquired ascending interleaved, without multiband/SMS, and with partial Fourier turned off (Supplemental Table 1). Subjects were given earplugs, a pillow beneath their knees for comfort, and foam padding was used to mitigate head movement.

Preprocessing and Analyses—Preprocessing and analyses across datasets were performed within AFNI (Cox, 1996) using *afni_proc.py* (see supplemental material for script). Although the 3T data were previously published, those data were reanalyzed for the present manuscript to match the normalization and smoothing of the 7T study. The spatial normalization and smoothing preprocessing steps are the only changes to the previously reported methods (Torrìsi et al., 2016). Nonlinear normalization used *3dQwarp* (Cox and Glen, 2013) and was performed to conform to the ICBM 2009 template in MNI space (Fonov et al., 2011). Because smoothing can alter the effect sizes and statistics of group studies (Mikl et al., 2008), AFNI’s iterative tool *3dBlurToFWHM* was applied to equate the smoothness across the datasets at a FWHM of 6mm.

For both 3T and 7T subject-level analyses using *3dDeconvolve*, shocks were modeled as effects of no interest along with each subject’s 6 head motion parameters. Data were additionally scaled by the voxel-wise mean during preprocessing so effect estimates would proximately represent percent signal change (and an effect of interest). Two-way ANOVA was performed at the group level using *3dANOVA3* for both field strengths separately, with GO and NOGO as two within-subject factors. Two types of statistical analyses were performed. The first, Analysis 1, focused on main effects of GO trials on the motor network (correct GO trials regardless of condition), while the second analysis, Analysis 2, focused on the main effects of NOGO trials (correct NOGO trials regardless of condition). When shown in figures as whole-brain results, statistical thresholding was matched between field strengths within the same group mask (cluster defining threshold $p=0.0001$) and corrected at the cluster level for whole brain $p<0.05$ FWE using the AFNI team’s recommended mixed model autocorrelation function (Cox et al., 2017a).

A direct, voxel-level comparison between field strengths was also performed. Each subject’s GO and NOGO main effects statistical maps were resampled to an intermediate spatial resolution between the acquired resolutions: $(1.9+3)/2 = 2.5\text{mm}$ isotropic. This enabled direct, 2-sample, 2-tailed *t*-tests for each condition. Correction for multiple comparisons for

this test was performed using non-parametric permutation testing using the *-Clustsim* option in *3dtest++* (Cox et al., 2017b). Resulting maps were set at a cluster defining threshold of $p=0.001$, $k=32$ for a strict, whole-brain $p<0.05$.

It was desirable to obtain a range of power estimates and not to limit each analysis to a single region of interest (ROI) (Mumford, 2012). The averaged (mean) effect values were extracted from a number of 8-mm radius spheres formed around coordinates reported in two published neuroimaging meta-analyses. For Analysis 1, the effect of GO trials, 10 ROIs were first chosen that had been reported as evidencing "...the main effects of all finger-tapping task variations..." (Witt et al., 2008). This meta-analysis included 38 studies. The 10 regions consisted of: bilateral sensorimotor cortices, supplementary motor area (SMA), left ventral premotor cortex, bilateral inferior parietal cortices, bilateral basal ganglia, and bilateral anterior cerebellum. For Analysis 2, the effect of NOGO trials, 16 ROIs were chosen directly from Table 2 of reported NOGO peaks (Swick et al., 2011). This meta-analysis included 48 studies reporting activation in NOGO response inhibition tasks. Because this table also reported cluster sub-peaks, the first coordinates for every cluster were chosen as centers for spherical ROIs. These 16 regions consisted of: right insula, right inferior parietal lobule, right medial frontal gyrus, left claustrum, left inferior parietal lobule, left middle frontal gyrus, right precuneus, right superior temporal gyrus, left supramarginal gyrus, right inferior occipital gyrus, left fusiform gyrus, left inferior frontal gyrus, right middle frontal gyrus, right caudate, right cingulate gyrus and left inferior parietal lobule. Because both meta-analyses reported peaks in Talairach space, prior to ROI creation the *tal2icbm_spm* transform was used to convert coordinates to MNI space (Lancaster et al., 2007). Locations of all ROIs are shown in Supplemental Figure 1.

Notably, our task did not activate some of these ROIs, which were removed from the analyses. This was determined by overlaying the spheres on liberally-thresholded and binarized activation maps. Thresholding was whole-brain corrected at a cluster defining threshold of $p<0.01$. ROIs showing no activation overlap in any voxels in an ROI for *both* 3T and 7T voxel-level group maps were considered outliers (Supplemental Figure 2). This resulted in the removal of three ROIs for the GO analyses (R basal ganglia, R inferior parietal and R sensorimotor), leaving 7, and the removal of one ROI for the NOGO analyses (R middle frontal gyrus), leaving 15.

Statistics between field strengths for each ROI were performed on averaged effects at the individual subject level, with group *t*-tests performed on those values. Bonferroni corrections for multiple ROI comparisons were determined at $p=0.05/7=0.007$ for Analysis 1 and $p=0.05/15=0.0033$ for Analysis 2.

An exploratory comparison was also performed to test whether small effects were more visible at 7T. A given region that activates robustly and with large effect is usually one identified in a quantitative meta-analysis, such as the ones used presently for ROI definitions. We asked if any of the NOGO clusters identified in the direct 7T>3T voxel-level comparison were not represented by the NOGO ROIs, indicating that they may typically be of small effect and not statistically detectable at normal field strength.

Finally, group-level power analyses between field strengths for both analyses and for each ROI were calculated. T-statistics were calculated from averaged effects value in a given ROI (as described above) and then used to calculate a Cohen's d using the formula $t\text{-stat}/\sqrt{n}$, where n = sample size of 31. This Cohen's d was then used with the $pwr()$ function in R, version 3.3.3 (R Core Team, 2016), to perform the power analyses (Cohen, 1988) with projected group sample sizes prescribed for each ROI as outcome measures. These were determined at: 80% power, $p < 0.001$, and as one-sample, two-sided t -tests. Results were separately plotted as slopegraphs for each ROI (Tufte, 1983) as well as median sample sizes for the GO and NOGO analyses, including interquartile ranges (IQR). The IQR function in R calculates IQR as $\text{quantile}(x, 75\%) - \text{quantile}(x, 25\%)$, based on (Hyndman and Fan, 1996). Medians are reported rather than means because they are less sensitive to outliers, which were expected given the approach taken to derive ROIs from coordinate-based meta-analyses.

Differing Methods Between Field Strengths

There were some differences across field strength acquisitions and analyses. Some were minor, others necessary and still others, such as higher spatial resolution at 7T, were chosen to reflect the current and perhaps inevitable trend of UHF researchers acquiring at increased resolution. None of these differences are expected to substantially affect the main findings in a way that would not affect most cognitive neuroscientists moving to UHF.

Participants—3T and 7T results are derived from two different cohorts of 31 subjects (62 total individuals), who were matched on age, sex and head motion. For the 3 Tesla study the average(SD) age was 28.0(6) years with 17 male and 14 female participants. For the 7 Tesla study the average age was 27.8(6); (17 male, 14 female).

Task design—The inter-stimulus interval (ISI) between 500ms GO and NOGO events ranged from a randomly-chosen 1500 to 2500 ms for the 3T study and from 1500 to 3000 ms for the 7T study. Because the same number of trials were retained between each field strengths, this minor difference in ISIs increased the total scan length at 7T. At 3T each run was 13min 12s (26 min 24s total) and at 7T each run was 14 min 20 sec, (28 min 40s total). Notably, this difference in ISI could have an effect on resolving the hemodynamic response function, however this would affect the precision and not the magnitude of parameter estimates. Because precision was not brought to the group-level analyses, this difference should have negligible effect on statistical comparisons.

Data acquisition—For the 3T study a standard functional resolution of 3 mm, isotropic voxels with a TE=30ms was acquired in a Siemens Magnetom Skyra. A T1-weighted 1 mm isotropic MPRAGE was also collected. For the 7T study a functional resolution of $1.78 \times 1.78 \times 1.98$ mm voxels was acquired with a TE=27ms in a Siemens Magnetom, and during analyses was resampled to 1.75mm isotropic. Parallel imaging (in-plane acceleration) was $2\times$ at 3T and $3\times$ at 7T. Further acquisition specifications can be found in (Torrise et al., 2016) and in Supplemental Table 1. Prior to acquisition of the 7T data, a 3rd-order shimming procedure was implemented to correct for inhomogeneities. A T1-weighted 0.7 mm

isotropic MPRAGE was collected for 7T subjects. Reverse phase encoded EPI (30 seconds) was also acquired at 7T for geometric distortion correction (Holland et al., 2010).

Finally, for the 7T subjects only, respiratory and cardiac data were collected and sampled at 500hz with a BioPac MP150 system (www.biopac.com). Because this data were not collected in the 3T study, physiological data were not used to remove physiological noise from the imaging data. Nevertheless, the 7T recordings were used for a manipulation check between conditions. Paired *t*-tests were used to test whether threat had an appreciable effect on respiration or cardiac cycles, which could potentially affect physiological noise and thus differentially affect the 7T data. Custom scripts were written to pull relevant measures from AFNI's *PeakFinder.py*.

Results

Behavior and physiology

Cohorts (3T vs. 7T) did not differ on NOGO accuracy (Mean(SD)% of 3T NOGOs = 71(20)%, 7T NOGOs = 76(15)%; $p=0.24$) but differed on GO accuracy: (3T GOs = 99(3)%, 7T GOs=95(7)%; $p=0.01$). Groups did not differ on GO reaction times for the threat ($p=0.62$) or safe conditions ($p=0.48$). For the 7T data, respiration ($t(27)=1.46$, $p=0.16$) and cardiac cycles ($t(27)=0.2$, $p=0.85$) did not change significantly between conditions.

Neuroimaging: percent signal change and *t*-statistics

Analysis 1 probed the effects of correct GO trials. Figure 1A shows the whole brain results at 3T and Figure 1B shows identically-thresholded whole brain results at 7T. Averaged effect estimates within the 7 GO ROIs, scaled to represent percent signal change (Figure 2A), and the group-level *t*-statistics of these effects (Figure 2B), are generally higher at 7T.

Analysis 2 probed the effects of correct NOGO trials. Figures 3A and 3B display identically-thresholded whole brain results for 3T and 7T, respectively. At 7T, averaged effect estimates within the 15 NOGO ROIs, scaled to represent percent signal change (Figure 4A), and the group-level *t*-statistics of these effects (Figure 4B), are also generally higher than at 3T. Supplemental figures 3A and 3B additionally show the maximum *t*-statistics in each ROI for both analyses 1 and 2.

Table 1 reports the effects sizes (percent signal change) and Cohen's *d* (Chen et al., 2017a) as well as the *t*-tests and statistical significance for all ROIs between field strengths. Many of the averaged voxel effects are significantly stronger at 7T, surviving Bonferroni correction for multiple tests.

The direct comparison of field strengths following intermediate resampling shows significant task-relevant regions despite strict, non-parametric permutation testing. Figure 5 (upper two panels) shows the 7T>3T contrast for GO trials, where the primary motor and ipsilateral cerebellar regions survived correction. Figure 5 (lower six panels) shows the 7T>3T contrast for NOGO trials and surviving regions such as the medial frontal gyrus, left occipital gyrus, left mid insula, bilateral anterior insula, right dlPFC and inferior parietal

lobules. No regions were significantly more activated at 3T. Table 2 reports peak voxel information for all significant clusters.

The exploratory analysis looking for evidence that 7T more robustly activates regions with generally smaller effect sizes is shown as Supplemental Figure 4, for the NOGO contrast. With the possible exception of regions ‘c’ and ‘e’, most clusters that crossed strict, permutation control, representing stronger activations at 7T than 3T, are not ones canonically activated by NOGO tasks. This is indicated by their non-overlap and spatial distance from the meta-analysis-derived ROIs used.

Neuroimaging: power analyses

Individual power analyses were performed for each GO or NOGO ROI. For both conditions, the median of both power analyses prescribed fewer subjects at 7T for the same level of statistical power as 3T (Figure 6). Additionally, Analysis 2 shows generally less variance of prescribed subjects. Figures 6A and 6C display how less subjects were prescribed at 7T for most regions, where one can follow the change for each ROI. Figures 6B and 6D separately plot the median and interquartile range for the individual ROI power analyses. For Analysis 1 (GO trials) at 3T, a median (IQR) of 19.6 (41.5) subjects were prescribed following the power analysis of the 7 ROIs that survived outlier exclusion. At 7T, 15.7 (43.8) subjects were prescribed following the power analysis with those 7 ROIs. For Analysis 2 (NOGO trials) at 3T, 38.5 (136.6) subjects were prescribed following the power analysis of the 15 ROIs that survived outlier exclusion. At 7T, 25.3 (22.2) subjects were prescribed following the power analysis of those 15 ROIs. We note that a direct statistical comparison between subject N prescriptions is not appropriate, as degrees of freedom are directly dependent on the number of ROIs used for each condition type. The ROI-based group power comparisons described here, reported as median and IQR values, should therefore be considered qualitative.

Discussion and Conclusion

Advances in hardware and software now make UHF sufficiently user-friendly to enable clinical or cognitive neuroscientists to easily adapt standard neuroimaging analyses, such as those performed here. This is the first group-level fMRI study reporting on the statistical power of 3T and 7T scanners using a conventional cognitive paradigm, analysis, and sample size. Field strength effects on the two conditions of a GO/NOGO paradigm were examined using unbiased ROIs. Results demonstrate that at 7T, effect sizes and statistics were significantly higher and generally fewer subjects were prescribed at the group level to match the same power at 3T. The following discussion explores the implications of these main findings.

Results suggest that by moving to UHF without altering experimental design and analysis, one can obtain higher fMRI effect sizes and levels of significance. Furthermore, it appears that smaller effects can also be detected at 7T. Higher significance levels overall, for a fixed false positive rate level, would benefit the cognitive neuroscience field, which has replicability issues partly driven by low statistical power (Poldrack et al., 2017). UHF is

generally not yet a part of the replicability conversation, however, and hopefully the present work will contribute to that discussion.

There is an obvious trade-off in functional imaging between signal strength partially enhanced by spatial blurring, and anatomical specificity (Mikl et al., 2008). Higher field strength will generally deliver better localization, particularly together with nonlinear warping of subjects to a common, well-defined template. Depending on the application, using UHF to achieve higher signal and corresponding statistics, but at the expense of anatomical specificity, may not be recommended. Future 7T cognitive neuroscience studies will need to strike a balance between desired anatomical precision and results that are correctable. It is likely that spatial smoothing will play a part in this balance, with less smoothing than currently applied becoming the norm. To some extent, this may require a cultural shift in functional neuroimaging with greater emphasis placed on precise anatomical reporting (Aguirre, 2014; Devlin and Poldrack, 2007; Turner and Geyer, 2014), and concomitant changes to conceptualizing the boundaries of functional localizations. As such, it is possible that future UHF cognitive neuroscience research will avoid general descriptors like “superior frontal gyrus” or “inferior parietal lobule” and instead consistently report cortical results at, for example, the level of a particular bank of a sulcus (Petrides, 2011).

Naturally, what UHF will offer to functional anatomy is still under active investigation, but most signs point to more precise anatomical investigations and reporting. For example, at the cortical level, work is ongoing to resolve the cortex with fMRI at the level of columns (Yacoub et al., 2008) and layers (Huber et al., 2017; 2016) which could have profound effects on the field. Nonetheless, specific functional boundaries of cortical subregions (e.g. digits of the hand) can have substantial intersubject variability (Kolasinski et al., 2016). The situation is likely to be more complex for areas of the association cortex (Laumann et al., 2015). Work at UHF has therefore also focused on subcortical structures (De Martino et al., 2013; Torrìsi et al., 2017) which may possess less intersubject variability. Accordingly, 7T benefits could be even greater for subcortical nuclei than for the cortex, at least with averaged group maps like those used here.

Finally, we performed power analyses across sets of regions for Analyses 1 and 2 of both the 3T and 7T data. These resulted in the 7T analyses prescribing a lower median number of subjects than at 3T and with generally less variance of the NOGO estimates across regions. This lower variance implies greater reliability when running power analyses at 7T. However, readers are urged not to make too much of the absolute numbers of subjects reported and shown in figure 6; instead, it is the relative median difference across ROIs, and the difference in variance, that are considered the more important findings. These findings could serve as a general guide for those considering how their own paradigms and procedures may be impacted by UHF.

To some extent, suggesting that future UHF cognitive neuroscientists may collect less data runs counter to recent admissions that research studies are often underpowered (Button et al., 2013) and the related push towards big data (Biswal et al., 2010). Aside from an investigator needing a large enough cohort to sufficiently sample their population of interest, very large sample sizes will still be necessary, such as with neuroimaging genetics (de

Zubicaray et al., 2008), when performing brain-behavior correlations (Yarkoni, 2009) or with dimensional analyses across subclinical and clinical constructs (Elton et al., 2016). Therefore, the results presented here are not a panacea for all future UHF imaging.

Crucially, greater signal and statistical power also allows for more flexibility in experimental designs. While directly moving an experimental paradigm from 3T to UHF to acquire fewer subjects may be desirable, one can also “trade” this power for a similar N but with other elements altered. For example, another benefit of UHF may be faster throughput for testing and piloting novel experimental paradigms. UHF could also enable investigators to make existing paradigms shorter, allowing for more different paradigms to be performed in a given session, or, alternatively, shorten the conditions within a paradigm to allow for a fewer number of trials and greater number of conditions, contrasts, or parametric levels.

Limitations

This study has some potential limitations. Although our contrasts concern the effects of GO and NOGO, the paradigm employed was not a pure GO/NOGO. The threat of shock component of the task may have therefore ‘colored’ the results in an unpredictable, or difficult to interpret, manner. However, our group previously observed that the threat of shock condition largely enhances the activation of regions that are already active during safety, and does not necessarily introduce newly active regions (Torrise et al., 2016). Furthermore, this effect would be the same between field strengths. It also appears, based on data at 7T, that respiratory and cardiac measures do not significantly change between threat and safe conditions, implying that field-dependent differences in noise would not be exacerbated by the threat conditions. Additionally, the ROIs used for the effects extractions and power analyses were based on clusters derived from previous research involving only finger tapping and NOGO trials, without a threat component. Therefore, even if additional brain regions were recruited during threat of shock they would likely not be captured among the ROIs used here.

Another limitation of the present study, alluded to above, is that the issue of physiological noise was not directly examined. Physiological noise increases with field strength (Hutton et al., 2011; Krüger and Glover, 2001) and, at our nominal functional resolution, physiological noise, relative to thermal noise, may affect the data (Triantafyllou et al., 2006; 2005). For example, due to the *t*-statistic equation and its conversion to a Cohen’s *d*, noise did influence the effect size estimates for the group power analyses. Future field strength comparisons should investigate the effects of noise on group power estimates. Our approach to probe power in meta-analysis-derived ROIs was not conducive to also investigating the contributions of physiological noise, because of the non-uniformity of noise across the imaging space and the inevitable inclusion of variable quantities of white matter in the spherical ROIs. Therefore, a study that probes group-level power and field strength effects on physiological noise effects may also benefit from power analyses at the voxel level rather than with ROIs (Durnez et al., 2016; Joyce and Hayasaka, 2012).

It is also a limitation that the cohorts were different between field strengths. Although both *n*=31 cohorts were also matched on age, sex, head motion, smoothing and most task performance, results could partly be due to differences in the subjects scanned. The present

results, however, are in line with other studies that tested identical cohorts at 3T and 7T (Beisteiner et al., 2011), suggesting that the main findings would remain if we collected the same subjects at both field strengths. Furthermore, habituation and order effects are avoided by collecting different cohorts. Nevertheless, having different individuals in each group remains a limitation because some analyses that could further gauge the consistency or agreement between field strengths could not be performed, such as an intraclass correlation (ICC) analysis (Chen et al., 2017b).

Finally, and concerning anatomical issues outlined above, we were unable to demonstrate potential benefits of 7T to explore small subcortical nuclei, such as the subthalamic nuclei involved in the motor inhibition component of NOGO trials (Aron et al., 2016). Indeed, our particular data were suboptimal to detect BOLD changes in these nuclei because of their high iron concentration (Puckett et al., 2017).

Conclusion

The present fMRI study examined a standard GO/NOGO paradigm at 3T and 7T field strengths. The goal was to test and demonstrate the effects of field strength on a cognitive neuroscience paradigm using accessible preprocessing and analysis tools. Results show greater statistical strength and greater power at 7T. Assuming an increase in availability of UHF scanners in the coming years, this power may be used in UHF cognitive neuroscience studies to present statistically reliable work using fewer participants or to leverage this power toward different and more efficient or comprehensive experimental designs. We hope that the present work will temper skepticism and uncertainties among some investigators about UHF, and we look forward to a future where UHF will complement the work done with commonly available technologies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Additional thanks to Paul Taylor, David Pagliaccio, Joshua Gowin and Wei Guo. This work utilized the computational resources of the NIH HPC Biowulf cluster (<http://hpc.nih.gov>). This work was supported by the Intramural Research Program of the National Institutes of Mental Health, project number ZIAMH002798 (clinical protocol 02-M-0321, NCT00047853) to CG, project number ZICMH002888 to GC, DG and RR, project number ZIAMH 00290 to CB and an NIH IRP sponsor to PB. The authors report no competing interests.

References

- Aguirre GK. Functional Neuroimaging: Technical, Logical, and Social Perspectives. Hastings Center Report. 2014; 44:S8–S18. DOI: 10.1002/hast.294
- Aron AR, Herz DM, Brown P, Forstmann BU, Zaghoul K. Frontosubthalamic Circuits for Control of Action and Cognition. Journal of Neuroscience. 2016; 36:11489–11495. DOI: 10.1523/JNEUROSCI.2348-16.2016 [PubMed: 27911752]
- Bandettini PA, Bowtell R, Jezzard P, Turner R. Ultrahigh field systems and applications at 7 T and beyond: Progress, pitfalls, and potential. Magn Reson Med. 2012; 67:317–321. DOI: 10.1002/mrm.23151 [PubMed: 22083719]

- Beisteiner R, Robinson S, Wurnig M, Hilbert M, Merksa K, Rath J, Höllinger I, Klinger N, Marosi C, Trattnig S, Geißler A. Clinical fMRI: Evidence for a 7T benefit over 3T. 2011; 57:1015–1021. DOI: 10.1016/j.neuroimage.2011.05.010
- Biswal BB, Mennes M, Zuo XN, Gohel S, Kelly C, Smith SM, Beckmann CF, Adelstein JS, Buckner RL, Colcombe S, Dogonowski AM, Ernst M, Fair D, Hampson M, Hoptman MJ, Hyde JS, Kiviniemi VJ, Kotter R, Li SJ, Lin CP, Lowe MJ, Mackay C, Madden DJ, Madsen KH, Margulies DS, Mayberg HS, McMahon K, Monk CS, Mostofsky SH, Nagel BJ, Pekar JJ, Peltier SJ, Petersen SE, Riedl V, Rombouts SARB, Rypma B, Schlaggar BL, Schmidt S, Seidler RD, Siegle GJ, Sorg C, Teng GJ, Veijola J, Villringer A, Walter M, Wang L, Weng XC, Whitfield-Gabrieli S, Williamson P, Windischberger C, Zang YF, Zhang HY, Castellanos FX, Milham MP. Toward discovery science of human brain function. *Proc Natl Acad Sci USA*. 2010; 107:4734–4739. DOI: 10.1073/pnas.0911855107 [PubMed: 20176931]
- Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, Munafò MR. Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci*. 2013; 14:365–376. DOI: 10.1038/nrn3475 [PubMed: 23571845]
- Chen G, Taylor PA, Cox RW. Is the statistic value all we should care about in neuroimaging? 2017a; 147:952–959. DOI: 10.1016/j.neuroimage.2016.09.066
- Chen G, Taylor PA, Haller SP, Kircanski K, Stoddard J, Pine DS, Leibenluft E, Brotman MA, Cox RW. Intraclass correlation: Improved modeling approaches and applications for neuroimaging. *Hum Brain Mapp*. 2017b; 39:1187–1206. DOI: 10.1002/hbm.23909 [PubMed: 29218829]
- Cohen, J. Statistical power analysis for the behavioral sciences. Lawrence Earlbaum Associates; Hillsdale, NJ: 1988.
- Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res*. 1996; 29:162–173. [PubMed: 8812068]
- Cox RW, Chen G, Glen DR, Reynolds RC, Taylor PA. fMRI clustering and false-positive rates. *Proc Natl Acad Sci USA*. 2017a; 201614961–2. doi: 10.1073/pnas.1614961114
- Cox RW, Chen G, Glen DR, Reynolds RC, Taylor PA. FMRI Clustering in AFNI: False-Positive Rates Redux. *Brain Connectivity*. 2017b; 7:152–171. DOI: 10.1089/brain.2016.0475 [PubMed: 28398812]
- Cox, RW., Glen, DR. Nonlinear Warping in AFNI. Poster presented at the 19th Annual Meeting of the Organization for Human Brain Mapping; Seattle, WA, USA. 2013.
- De Martino F, Moerel M, van de Moortele PF, Ugurbil K, Goebel R, Yacoub E, Formisano E. Spatial organization of frequency preference and selectivity in the human inferior colliculus. *Nat Comms*. 2013; 4:1386.doi: 10.1038/ncomms2379
- de Zubizaray GI, Chiang MC, McMahon KL, Shattuck DW, Toga AW, Martin NG, Wright MJ, Thompson PM. Meeting the Challenges of Neuroimaging Genetics. *Brain Imaging and Behavior*. 2008; 2:258–263. DOI: 10.1007/s11682-008-9029-0 [PubMed: 20016769]
- Devlin J, Poldrack RA. In praise of tedious anatomy. 2007; 37:1033–1041. DOI: 10.1016/j.neuroimage.2006.09.055
- Durnez J, Degryse J, Moerkerke B, Seurinck R, Sochat V, Poldrack R, Nichols TE. Power and sample size calculations for fMRI studies based on the prevalence of active peaks. *bioRxiv*. 2016; :1–24. DOI: 10.1101/049429
- Edelstein WA, Glover GH, Hardy CJ, Redington RW. The intrinsic signal-to-noise ratio in NMR imaging. *Magnetic Resonance Medicine*. 1986; 3:604–618. DOI: 10.1002/mrm.1910030413
- Elton A, Di Martino A, Hazlett HC, Gao W. Neural Connectivity Evidence for a Categorical-Dimensional Hybrid Model of Autism Spectrum Disorder. *BPS*. 2016; 80:120–128. DOI: 10.1016/j.biopsycho.2015.10.020
- First MB, Spitzer RL, Gibbon M, Williams J. Structured Clinical Interview for DSM-IV-TR Axis I Disorders – Non-Patient Edition (SCID-I/NP, 1/2007 Revision). 2007
- Fonov V, Evans AC, Botteron K, Almli CR, McKinstry RC, Collins DL, Group TBDC. Unbiased average age-appropriate atlases for pediatric studies. 2011; 54:313–327. DOI: 10.1016/j.neuroimage.2010.07.033

- Holland D, Kuperman JM, Dale AM. Efficient correction of inhomogeneous static magnetic field-induced distortion in Echo Planar Imaging. 2010; 50:175–183. DOI: 10.1016/j.neuroimage.2009.11.044
- Huber L, Handwerker DA, Jangraw DC, Chen G, Hall A, Stüber C, Gonzalez-Castillo J, Ivanov D, Marrett S, Guidi M, Goense J, Poser BA, Bandettini PA. High-Resolution CBV-fMRI Allows Mapping of Laminar Activity and Connectivity of Cortical Input and Output in Human M1. *Neuron*. 2017; 96:1253–1263.e7. DOI: 10.1016/j.neuron.2017.11.005 [PubMed: 29224727]
- Huber L, Ivanov D, Handwerker DA, Marrett S, Guidi M, Uludağ K, Bandettini PA, Poser BA. Techniques for blood volume fMRI with VASO: From low-resolution mapping towards sub-millimeter layer-dependent applications. *NeuroImage*. 2016; 133:1–13. doi: 10.1016/j.neuroimage.2016.11.039
- Hutton C, Josephs O, Stadler J, Featherstone E, Reid A, Speck O, Bernarding J, Weiskopf N. The impact of physiological noise correction on fMRI at 7T. 2011; 57:101–112. DOI: 10.1016/j.neuroimage.2011.04.018
- Hyndman RJ, Fan Y. Sample Quantiles in Statistical Packages. *The American Statistician*. 1996; 50:361–365. DOI: 10.2307/2684934
- Joyce KE, Hayasaka S. Development of PowerMap: a Software Package for Statistical Power Calculation in Neuroimaging Studies. *Neuroinformatics*. 2012; 10:351–365. DOI: 10.1007/s12021-012-9152-3 [PubMed: 22644868]
- Kolasinski J, Makin TR, Jbabdi S, Clare S, Stagg CJ, Johansen-Berg H. Investigating the Stability of Fine-Grain Digit Somatotopy in Individual Human Participants. *Journal of Neuroscience*. 2016; 36:1113–1127. DOI: 10.1523/JNEUROSCI.1742-15.2016 [PubMed: 26818501]
- Krüger G, Glover GH. Physiological noise in oxygenation-sensitive magnetic resonance imaging. *Magnetic Resonance Medicine*. 2001; 46:631–637. DOI: 10.1002/mrm.1240
- Krüger G, Kastrup A, Glover GH. Neuroimaging at 1.5 T and 3.0 T: Comparison of oxygenation-sensitive magnetic resonance imaging. *Magnetic Resonance Medicine*. 2001; 45:595–604. DOI: 10.1002/mrm.1081
- Lancaster JL, Tordesillas-Gutiérrez D, Martinez M, Salinas F, Evans A, Zilles K, Mazziotta JC, Fox PT. Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Hum Brain Mapp*. 2007; 28:1194–1205. DOI: 10.1002/hbm.20345 [PubMed: 17266101]
- Laumann TO, Gordon EM, Adeyemo B, Snyder AZ, Joo SJ, Chen MY, Gilmore AW, McDermott KB, Nelson SM, Dosenbach NUF, Schlaggar BL, Mumford JA, Poldrack RA, Petersen SE. Functional System and Areal Organization of a Highly Sampled Individual Human Brain. *Neuron*. 2015; 87:657–670. DOI: 10.1016/j.neuron.2015.06.037 [PubMed: 26212711]
- Mikl M, Marek R, Hlušík P, Pavlicová M, Drastich A, Chlebus P, Brázdil M, Krupa P. Effects of spatial smoothing on fMRI group inferences. *Magnetic Resonance Imaging*. 2008; 26:490–503. DOI: 10.1016/j.mri.2007.08.006 [PubMed: 18060720]
- Mumford JA. A power calculation guide for fMRI studies. *Social Cognitive and Affective Neuroscience*. 2012; 7:738–742. DOI: 10.1093/scan/nss059 [PubMed: 22641837]
- Murphy K, Bodurka J, Bandettini PA. How long to scan? The relationship between fMRI temporal signal to noise ratio and necessary scan duration. *NeuroImage*. 2007; 34:565–574. DOI: 10.1016/j.neuroimage.2006.09.032 [PubMed: 17126038]
- Newton AT, Rogers BP, Gore JC, Morgan VL. Improving measurement of functional connectivity through decreasing partial volume effects at 7T. 2012; 59:2511–2517. DOI: 10.1016/j.neuroimage.2011.08.096
- Petrides, M. The human cerebral cortex: an MRI atlas of the sulci and gyri in MNI stereotaxic space. 1. Academic Press; 2011.
- Poldrack RA, Baker CI, Durnez J, Gorgolewski KJ, Matthews PM, Munafò MR, Nichols TE, Poline JBP, Vul E, Yarkoni T. Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nat Rev Neurosci*. 2017; 18:115–126. DOI: 10.1038/nrn.2016.167 [PubMed: 28053326]
- Poline J-BP, Brett M. The general linear model and fMRI: Does love last forever? 2012; 62:871–880. DOI: 10.1016/j.neuroimage.2012.01.133
- Puckett AM, Bollmann S, Poser BA, Palmer J, Barth M, Cunningham R. Using multi-echo simultaneous multi-slice (SMS) EPI to improve functional MRI of the subcortical nuclei of the

- basal ganglia at ultra-high field (7T). *NeuroImage*. 2017; :1–10. DOI: 10.1016/j.neuroimage.2017.12.005
- R Core Team. R: A language and environment for statistical computing. 2016.
- Schmitz A, Grillon C. Assessing fear and anxiety in humans using the threat of predictable and unpredictable aversive events (the NPU-threat test). *Nature Protocols*. 2012; 7:527–532. DOI: 10.1038/nprot.2012.001 [PubMed: 22362158]
- Sladky R, Baldinger P, Kranz GS, Tröstl J, Höflich A, Lanzenberger R, Moser E, Windischberger C. High-resolution functional MRI of the human amygdala at 7T. *European Journal of Radiology*. 2013; 82:728–733. DOI: 10.1016/j.ejrad.2011.09.025 [PubMed: 22138120]
- Swick D, Ashley V, Turken U. Are the neural correlates of stopping and not going identical? Quantitative meta-analysis of two response inhibition tasks. 2011; 56:1655–1665. DOI: 10.1016/j.neuroimage.2011.02.070
- Torrìsi SJ, Nord CL, Balderston N, Roiser JP, Grillon C, Ernst M. Resting state connectivity of the human habenula at ultra-high field. 2017; 147:872–879. DOI: 10.1016/j.neuroimage.2016.10.034
- Torrìsi SJ, Robinson OJ, O’Connell K, Davis A, Balderston N, Ernst M, Grillon C. The neural basis of improved cognitive performance by threat of shock. *Social Cognitive and Affective Neuroscience*. 2016; :nsw088–10. DOI: 10.1093/scan/nsw088
- Triantafyllou C, Hoge RD, Krueger G, Wiggins CJ, Potthast A, Wiggins GC, Wald LL. Comparison of physiological noise at 1.5 T, 3 T and 7 T and optimization of fMRI acquisition parameters. 2005; 26:243–250. DOI: 10.1016/j.neuroimage.2005.01.007
- Triantafyllou C, Hoge RD, Wald LL. Effect of spatial smoothing on physiological noise in high-resolution fMRI. *NeuroImage*. 2006; 32:551–557. DOI: 10.1016/j.neuroimage.2006.04.182 [PubMed: 16815038]
- Triantafyllou C, Polimeni JR, Wald LL. Physiological noise and signal-to-noise ratio in fMRI with multi-channel array coils. *NeuroImage*. 2011; 55:597–606. DOI: 10.1016/j.neuroimage.2010.11.084 [PubMed: 21167946]
- Tufte, ER. *The Visual Display of Quantitative Information*. 1. 1983.
- Turner R, Geyer S. Comparing Like with Like: The Power of Knowing Where You Are. *Brain Connectivity*. 2014; 4:547–557. DOI: 10.1089/brain.2014.0261 [PubMed: 24999746]
- van der Kolk AG, Hendrikse J, Zwanenburg JJM, Visser F, Luijten PR. Clinical applications of 7T MRI in the brain. *European Journal of Radiology*. 2013; 82:708–718. DOI: 10.1016/j.ejrad.2011.07.007 [PubMed: 21937178]
- van der Zwaag W, Francis S, Head K, Peters A, Gowland P, Morris P, Bowtell R. fMRI at 1.5, 3 and 7 T: Characterising BOLD signal changes. 2009; 47:1425–1434. DOI: 10.1016/j.neuroimage.2009.05.015
- Wald LL, Polimeni JR. Impacting the effect of fMRI noise through hardware and acquisition choices – Implications for controlling false positive rates. *NeuroImage*. 2017; :1–8. DOI: 10.1016/j.neuroimage.2016.12.057
- Witt ST, Laird AR, Meyerand ME. Functional neuroimaging correlates of finger-tapping task variations: an ALE meta-analysis. 2008; 42:343–356. DOI: 10.1016/j.neuroimage.2008.04.025
- Yacoub E, Harel N, Ugurbil K. High-field fMRI unveils orientation columns in humans. *Proceedings of the National Academy of Sciences*. 2008; 105:10607–10612. DOI: 10.1073/pnas.0804110105
- Yarkoni T. Big Correlations in Little Studies: Inflated fMRI Correlations Reflect Low Statistical Power-Commentary on Vul et al. (2009). *Perspect Psychol Sci*. 2009; 4:294–298. DOI: 10.1111/j.1745-6924.2009.01127.x [PubMed: 26158966]

Figure 1.

A: The t-statistic values of GO trials at 3 Tesla. B: The effects of GO trials at 7 Tesla. Both sets of results are whole-brain corrected at cluster $p < 0.05$. Cluster defining threshold is matched between results ($p < 0.0001$); however, different spatial resolutions required different cluster sizes for correction: $k=5$ voxels for 3T and $k=25$ for 7T (whole-brain corrected at $p < 0.05$). In this figure and Figures 3 and 5, statistical results are overlaid on axial slices of subjects' averaged T1s after nonlinear normalization. Results here are shown within a gray matter mask that matches the normalization template.

Figure 2.

Analysis 1, GO effects (right-handed index finger pressing). A: Effect differences in peak percent signal change for 7 regions of interest between 3 and 7 Tesla. B: Peak (or minimum) T-statistics within ROIs. Sphere locations chosen from coordinates reported in (Witt et al., 2008). L=left, R=right, ant=anterior, inf=inferior, suppl= supplementary. Error bars = SEM.

Figure 3.

A: The t-statistic values of NOGO trials at 3 Tesla. B: T-statistic values of NOGO trials at 7 Tesla. Cluster defining thresholds are matched with the GO trial analysis ($p < 0.0001$), however, different resolutions required cluster sizes of $k=5$ for 3T and $k=25$ for 7T.

Figure 4.

Analysis 2, NOGO effects. A: Averaged percent signal change differences (scaled regression betas) for 15 regions of interest between 3 and 7 Tesla. B: Averaged T-statistics within ROIs. Sphere locations chosen from coordinates reported in (Swick et al., 2011). L=left, R=right, inf=inferior, mid=middle, supram=supramarginal, par=parietal, temp=temporal, occip=occipital, med=medial, g=gyrus. Error bars=SEM.

Figure 5.

Direct comparison between 7T and 3T datasets, resampled to an intermediate 2.5mm isotropic resolution. Results are corrected with a non-parametric permutation test, CDT: $p < 0.001$, $k=32$, whole brain 5% false positive rate. What is shown is a subset of all significant clusters (see Table 2 for all).

Figure 6.

Power analyses for individual ROIs estimated for each field strength. A: Slopegraph visualization of 7 ROIs derived from Witt, et al representing GO effects, shown as group sample size prescriptions for each ROI. B: The same ROIs plotted as points with median (IQR). C: Slopegraph visualization of 15 ROIs derived from Swick, et al representing NOGO effects, shown as group sample size prescriptions for each ROI. D: The same ROIs plotted as points with median (IQR). Y axes = log base 2. **Colors in A and C match Supplemental Figure 1.** In B and D closed circles are horizontally jittered for clarity.

Table 1

Comparisons between field strengths for all ROIs; GO and NOGO analyses. Averaged effects are tested between field strengths. Effect values correspond to those graphed in figures 2A and 4A. DoF = degrees of freedom. Greyed p-values indicate Bonferroni uncorrected.

ROI source (trial type)	ROI (as labeled by source)	3T mean $\pm$ SD of average effects (% sig change)	7T mean $\pm$ SD of average effects (% sig change)	3T Cohen's d	7T Cohen's d	7T vs 3T average effects: T statistics	7T vs 3T average effects: DoF	7T vs 3T average effects: p values
Witt et al. (GO)	left anterior cerebellum	0.029 $\pm$ 0.07	0.083 $\pm$ 0.14	0.40	0.55	1.90	43.9	0.063563
	right anterior cerebellum	0.116 $\pm$ 0.1	0.265 $\pm$ 0.14	1.11	2.38	4.60	54.4	0.000026
	left basal ganglia	0.033 $\pm$ 0.05	0.036 $\pm$ 0.09	0.66	0.56	0.16	46.1	0.873512
	left inferior parietal	0.077 $\pm$ 0.07	0.191 $\pm$ 0.14	1.09	1.41	3.97	43.3	0.000270
	left ventral premotor	0.037 $\pm$ 0.07	0.07 $\pm$ 0.11	0.50	0.61	1.32	51.4	0.193927
	left sensorimotor	0.17 $\pm$ 0.14	0.327 $\pm$ 0.17	1.20	1.98	3.89	57.6	0.000261
	supplementary motor	0.111 $\pm$ 0.09	0.237 $\pm$ 0.22	1.22	1.28	2.95	40.0	0.005308
Swick et al. (NOGO)	left fusiform gyrus	0.096 $\pm$ 0.12	0.225 $\pm$ 0.13	0.79	1.67	3.96	59.3	0.000205
	right insula	0.114 $\pm$ 0.11	0.214 $\pm$ 0.24	1.01	0.90	2.12	42.7	0.039899
	right inferior occipital gyrus	0.034 $\pm$ 0.15	0.176 $\pm$ 0.15	0.23	1.15	3.70	59.9	0.000465
	left claustrum	0.06 $\pm$ 0.07	0.14 $\pm$ 0.14	0.87	0.95	2.74	42.5	0.008889
	right superior temporal gyrus	0.059 $\pm$ 0.14	0.165 $\pm$ 0.24	0.28	0.31	2.11	49.6	0.039856
	right caudate body	0.064 $\pm$ 0.09	0.12 $\pm$ 0.18	0.40	0.69	1.51	43.3	0.139394
	left inferior frontal gyrus	0.009 $\pm$ 0.1	0.034 $\pm$ 0.2	0.72	0.65	0.62	43.8	0.541197
	right cingulate gyrus	0.065 $\pm$ 0.12	0.206 $\pm$ 0.26	0.09	0.17	2.71	41.3	0.009827
	left middle frontal gyrus	0.093 $\pm$ 0.1	0.245 $\pm$ 0.21	0.55	0.77	3.57	42.8	0.000899
	left supramarginal gyrus	0.093 $\pm$ 0.12	0.256 $\pm$ 0.24	0.92	1.14	3.32	45.2	0.001777
	right precuneus	0.019 $\pm$ 0.12	0.148 $\pm$ 0.21	0.74	1.06	2.92	49.5	0.005259
	right inferior parietal lobule	0.072 $\pm$ 0.1	0.237 $\pm$ 0.25	0.15	0.71	3.33	39.8	0.001891
	left inferior parietal lobule	0.085 $\pm$ 0.12	0.246 $\pm$ 0.2	0.68	0.93	3.83	48.5	0.000370
	right medial frontal gyrus	0.137 $\pm$ 0.17	0.325 $\pm$ 0.26	0.72	1.22	3.30	51.3	0.001750
	left inferior parietal lobule	0.098 $\pm$ 0.19	0.227 $\pm$ 0.26	0.80	1.22	2.25	54.9	0.028748

Table 2

Regions and peak coordinates from direct 7T>3T comparison, corresponding to Figure 5. Results whole brain corrected by non-parametric permutation. BA = (approximate) Brodmann's area. k = number of (2.5mm) voxels in a given cluster. Coordinates in MNI space.

Contrast, trial	region label	k	X	Y	Z	Z-stat
7T>3T, GOs						
	L primary motor cortex (BA 4)	171	-38.5	-22.8	58.5	6.11
	R cerebellum (lobule VI)	114	21.5	-60.2	-26.5	4.99
	L middle occipital gyrus (BA 18)	64	-31	-95.2	1	4.75
	R angular gyrus white matter	40	41.5	-57.8	31	5.79
7T>3T, NOGOs						
	L SMA (BA 6)	351	-6	-2.8	53.5	6.01
	L middle occipital gyrus (BA 18)	99	-33.5	-92.8	-1.5	5.21
	L middle insula (BA 13)	87	-38.5	7.2	-4	4.83
	R middle frontal gyrus (BA 9)	80	34	37.2	33.5	5.11
	L supramarginal gyrus (BA 40)	78	-53.5	-42.8	36	4.63
	L primary motor cortex (BA 4)	68	-38.5	-20.2	58.5	4.93
	R anterior insula (BA 13)	65	36.5	19.8	6	4.65
	R supramarginal gyrus (BA 40)	65	66.5	-45.2	26	4.73
	R middle occipital gyrus (BA 18)	43	34	-92.8	-1.5	4.51
	L superior middle insula (BA 13)	42	-38.5	-2.8	13.5	4.54
	R middle frontal gyrus (BA 6)	37	44	-2.8	51	4.51
	R anterior cingulate cortex (BA 32)	33	9	27.2	26	4.37