UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

fMRI protocol optimization for simultaneously studying small subcortical and cortical areas at 7 T

Miletić, S.; Bazin, P.-L.; Weiskopf, N.; van der Zwaag, W.; Forstmann, B.U.; Trampel, R.

DOI
10.1016/j.neuroimage.2020.116992

Publication date
2020

Document Version
Final published version

Published in
NeuroImage

License
CC BY

Link to publication

Citation for published version (APA):
Miletić, S., Bazin, P.-L., Weiskopf, N., van der Zwaag, W., Forstmann, B. U., & Trampel, R. (2020). fMRI protocol optimization for simultaneously studying small subcortical and cortical areas at 7 T. *NeuroImage*, *219*, Article 116992. https://doi.org/10.1016/j.neuroimage.2020.116992

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

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:09 Nov 2026

NeuroImage 219 (2020) 116992

Contents lists available at ScienceDirect

NeuroImage

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

fMRI protocol optimization for simultaneously studying small subcortical and cortical areas at 7 T

Steven Miletić [a,*], Pierre-Louis Bazin [a,b], Nikolaus Weiskopf [b,c], Wietske van der Zwaag [d], Birte U. Forstmann [a], Robert Trampel [b]

[a] University of Amsterdam, Department of Psychology, Nieuwe Achtergracht 129B, Amsterdam, the Netherlands
[b] Max Planck Institute for Human Cognitive and Brain Sciences, Department of Neurophysics, Stephanstraße 1A, Leipzig, Germany
[c] Felix Bloch Institute for Solid State Physics, Faculty of Physics and Earth Sciences, Leipzig University, Linnéstraße 5, Leipzig, Germany
[d] Spinoza Center for Neuroimaging, Meibergdreef 75, Amsterdam, the Netherlands

ARTICLE INFO

Keywords:
Multi echo EPI
Single echo EPI
fMRI protocol optimization
Subcortex
Cortex
Stop-signal response task

ABSTRACT

Most fundamental cognitive processes rely on brain networks that include both cortical and subcortical structures. Studying such networks using functional magnetic resonance imaging (fMRI) requires a data acquisition protocol that provides blood-oxygenation-level dependent (BOLD) sensitivity across the entire brain. However, when using standard single echo, echo planar imaging protocols, researchers face a tradeoff between BOLD-sensitivity in cortex and in subcortical areas. Multi echo protocols avoid this tradeoff and can be used to optimize BOLD-sensitivity across the entire brain, at the cost of an increased repetition time. Here, we empirically compare the BOLD-sensitivity of a single echo protocol to a multi echo protocol. Both protocols were designed to meet the specific requirements for studying small, iron rich subcortical structures (including a relatively high spatial resolution and short echo times), while retaining coverage and BOLD-sensitivity in cortical areas. The results indicate that both sequences lead to similar BOLD-sensitivity across the brain at 7 T.

1. Introduction

Most fundamental cognitive processes rely on brain networks that include both cortical and subcortical structures (Forstmann et al., 2017). A prominent example is the ability to inhibit an already initiated motor action, which is impaired in a range of disorders including attention deficit/hyperactivity disorder (Lijffijt et al., 2005), obsessive-compulsive disorder (Chamberlain et al., 2006), and Parkinson's disease (Gauggel et al., 2004). The brain network underlying this ability is thought to include the inferior frontal gyrus (IFG), pre-supplementary motor area (preSMA), primary motor cortex (M1), as well as basal ganglia structures including the striatum (STR), subthalamic nucleus (STN) and globus pallidus externa (GPe) and interna (GPi) (Aron et al., 2014; Aron and Poldrack, 2006; De Hollander et al., 2017; Li et al., 2008; Ray et al., 2012; Sebastian et al., 2018; Wiecki and Frank, 2012). For a complete understanding of those brain networks, it is vital to study all components of the network simultaneously.

When studying such networks using functional magnetic resonance imaging (fMRI), cortical and subcortical structures put different and sometimes conflicting requirements on imaging protocols, causing researchers to face a tradeoff in BOLD-sensitivity between cortex and subcortex. An important factor underlying these conflicting requirements is the neurochemical composition of brain structures: Subcortical structures such as the STN, GPe, and GPi contain high concentrations of iron (Aquino et al., 2009; De Hollander et al., 2014; Deistung et al., 2013; Keuken et al., 2017, 2014), which leads to substantially shorter T_2^* relaxation times compared to cortical areas (Peters et al., 2007). The interregional variability in T_2^* values is relevant because for optimal BOLD-sensitivity, the echo time (TE) of a single echo echo-planar imaging (EPI) protocol should be equal to the T_2^* of the tissue of interest (Posse et al., 1999), which is impossible to achieve across the entire brain, even when varying the echo time of the EPI acquisition between slices (Stöcker et al., 2006). As a consequence, using protocols with TEs optimized for cortex can lead to limited or even no BOLD-sensitivity in iron rich subcortical structures (e.g., De Hollander et al., 2017). Vice versa, theory predicts that protocols using a short TE have suboptimal BOLD-sensitivity in cortical areas.

The tradeoff between BOLD-sensitivity in cortex and subcortex can be avoided by acquiring images at multiple echo times after each excitation

* Corresponding author. Nieuwe Achtergracht 129B, 1001NK Amsterdam, the Netherlands.
E-mail address: s.miletic@uva.nl (S. Miletić).

https://doi.org/10.1016/j.neuroimage.2020.116992
Received 10 December 2019; Received in revised form 14 May 2020; Accepted 20 May 2020
Available online 29 May 2020
1053-8119/© 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

pulse using a single-shot multi echo EPI protocol (Gowland and Bowtell, 2007; Poser et al., 2006; Posse et al., 1999; Speck and Hennig, 1998; Weiskopf et al., 2005). For subsequent statistical analyses, the multiple echoes can then be recombined into a single timeseries. Depending on the T_2^* of the tissue and the echo times of the protocol, the recombination of multiple echoes can lead to increased contrast-to-noise across the entire brain (Gowland and Bowtell, 2007; Kettinger et al., 2016; Poser et al., 2006; Poser and Norris, 2009; Posse et al., 1999).

Previous studies empirically tested the BOLD-sensitivity of multi echo protocols compared to a more common single echo counterpart (e.g., Kirilina et al., 2016; Poser and Norris, 2009; Puckett et al., 2018). For example, one study (Kirilina et al., 2016) compared four imaging protocols at 3 T, including two multi echo sequences. The main conclusion was that protocol choice had limited influence on the final results in random effects group analyses, except in brain areas where the contrast-to-noise was very limited for some protocols (e.g., orbitofrontal cortex due to dropout). In such areas, the final results benefitted from multi echo acquisitions. Since contrast-to-noise is relatively limited in small, iron rich subcortical nuclei, we hypothesized that such nuclei may also particularly benefit from multi echo acquisitions. This hypothesis was corroborated by a second recent protocol comparison study (Puckett et al., 2018), which concluded at 7 T that a multi echo protocol provided increased BOLD-sensitivity in subcortical areas compared to a single echo protocol optimized for cortex.

While these results are promising, the voxel sizes of protocols used in earlier studies were large relative to the volume of many subcortical areas (e.g., Keuken et al., 2014; Zwirner et al., 2017). As a result, the number of voxels per such structure is limited, and voxels likely contain tissue from adjacent structures (De Hollander et al., 2015; Mulder et al., 2019). In voxel-wise analyses using the general linear model (GLM), this issue is exacerbated by the spatial smoothing that is typically imposed in fMRI preprocessing (De Hollander et al., 2015; Geissler et al., 2005; Stelzer et al., 2014). As a consequence, voxels contain signals that originate from a mixture of sources, which limits the conclusions that can be drawn about the function of small nuclei as independent nodes in a network.

Here, we provide a new comparison between a multi echo protocol and a single echo protocol at 7 T. The comparison differs from earlier experiments in two ways. First, we use a high spatial resolution in both protocols (1.6 mm isotropic voxel size). Although ultra-high field MRI can be used for functional data acquisition with even sub-millimeter spatial resolution (e.g., Heidemann et al., 2012), especially multi echo protocols require short echo train lengths which limits further increasing spatial resolution. Acquiring multiple echoes at a spatial resolution of 1.6 mm requires more parallel acceleration than would be required for a single echo protocol with the same TR, and we aim to test whether the multi echo protocol still provides advantages under such amounts of acceleration. Second, we explicitly choose to optimize both protocols for studying the subcortex, rather than cortex, by adapting the echo time(s) accordingly. We compare protocols in their ability to detect BOLD-responses in small subcortical nuclei, as well as larger cortical areas. Data from both protocols were analyzed using signal-to-noise and contrast-to-noise analyses, signal deconvolution, and general linear models, as detailed below.

2. Method

2.1. Participants

18 participants (9 female, age 23.7y [SD 3.2y], all right handed) were recruited from the subject pool of the Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig. The study was approved by the local ethical committee. All participants gave written informed consent. All participants had normal or corrected-to-normal vision, and no history of neurological, major medical, or psychiatric disorders.

2.2. MR protocols

Each participant was scanned twice in counterbalanced sessions on a 7 T MRI system (Siemens Healthineers, Germany) using a 32-channel phased array headcoil (Nova Medical Inc, USA). Each session consisted of the acquisition of three functional runs, a B0 field map, and an anatomical image from an MP2RAGE sequence (Marques et al., 2010). The sessions differed only in the protocol that was used to acquire the functional data. In one session, a single echo EPI protocol optimized for subcortex was used (TR = 3000 ms, TE = 14 ms, 1.6 mm isotropic, GRAPPA 3 in phase encoding (PE) direction, excitation flip angle 70°, receiver bandwidth 1436 Hz/Pixel, partial Fourier 6/8, FOV 192 mm × 192 mm x 112 mm, imaging matrix 120 × 120, 70 slices). The timing of a single readout gradient waveform was as follows: duration = 0.80 ms (= echo spacing), duration of ADC = 0.70 ms (= 1/bandwidth), duration of flat top = 0.56 ms (Gradient amplitude during flat top = 22.4 mT/m), ramp time = 0.12 ms (slew rate = 189 T/m/s), ramp time while ADC was switched on = 0.07 ms. This resulted in a ramp-sampling percentage of 19%. The protocol was based on earlier work where it demonstrated good BOLD-sensitivity in basal ganglia nuclei (De Hollander et al., 2017; Mestres-Missé et al., 2017), but with a slightly increased voxel size to obtain near whole-brain coverage (with the exception of the tips of the temporal lobes and cerebellum). In the other session, a multi echo EPI sequence developed by the CMRR (https://www.cmrr.umn.edu/multiband/) was used, in which we minimized echo times (TR = 3000 ms, $TE_{1\text{-}3}$ = 9.66, 24.87, 40.08 ms, 1.6 mm isotropic, GRAPPA 4 in PE direction, multi-band factor 2 with CAIPI shift of FOV/3, excitation flip angle 65°, receiver bandwidth 1984 Hz/Pixel, partial Fourier 6/8, FOV 192 mm × 192 mm x 112 mm, imaging matrix = 120 × 120, 70 slices). The timing of a single readout gradient waveform was as follows: duration = 0.63 ms (= echo spacing), duration of ADC = 0.50 ms (= 1/bandwidth), duration of flat top = 0.27 ms (Gradient amplitude during flat top = 34.6 mT/m), ramp time = 0.18 ms (slew rate = 189 T/m/s), ramp time while ADC was switched on = 0.12 ms. This resulted in a ramp-sampling percentage of 48%. Increasing the receiver bandwidth and the GRAPPA factor for the multi echo protocol was necessary to achieve the same spatial resolution, coverage and TR as for the single echo sequence, as well as to achieve the fast first echo time of 9.66 ms.

The B0 field map was acquired with the same FOV as the functional protocols (TR = 1500 ms, TE 1 = 6.00 ms, TE 2 = 7.02 ms, voxel size 3.2 mm isotropic). The MP2RAGE (Marques et al., 2010) was acquired with the following parameters: voxel size = 0.7 mm isotropic, FOV = 224 mm × 224 mm x 168 mm, TR = 5000 ms, TE = 2.45 ms, TI 1 = 900 ms, TI 2 = 2750 ms, flip angle 1 = 5°, flip angle 2 = 3°.

2.3. Theoretical contrast-to-noise (CNR) estimations

Based on both protocols' readout bandwidths, in-plane parallel imaging factors, and *g*-factor estimations (see Appendix C), we can approximate the expected between-protocol ratio of signal-to-noise ratio (SNR) and BOLD CNR. Since $SNR \propto \frac{\sqrt{nLines}}{g\sqrt{bandwidth}}$, where *nLines* is the number of acquired lines in k-space, the ratio between SNRs $\frac{SNR_{multi}}{SNR_{single}} \propto \frac{\sqrt{bandwidth_{single}} * g_{single} * \sqrt{nLines_{multi}}}{\sqrt{bandwidth_{multi}} * g_{multi} * \sqrt{nLines_{single}}}$. Assuming *g*-factors of 1.15 and 1.5 for the single echo and multi echo protocol, respectively (corresponding to the *g*-factor increase in the STN area, see Appendix C), this leads to an estimated $\frac{SNR_{multi}}{SNR_{single}}$ of approximately 0.57 - indicating that the SNR of the multi echo protocol is estimated to be approximately 57% of the SNR of the single echo protocol in the STN. Assuming *g*-factors of approximately 1.1 and 1.3, (for the single echo and multi echo protocol, respectively) for cortical regions, $\frac{SNR_{multi}}{SNR_{single}} = 0.62$.

Assuming mono-exponential decay of transverse magnetization, the CNR per unit change in T_2^* for the single echo protocol, is given by (e.g., Poser et al., 2006; Posse et al., 1999):

$$CNR(TE) = \frac{S_0}{\sigma_0} TE exp\left(\frac{-TE}{T_2^*}\right) \quad (1)$$

where S_0 is the signal at TE = 0 ms. The expected CNR of the multi echo protocol depends on the method for combining the echoes (Gowland and Bowtell, 2007; Kettinger et al., 2016; Poser et al., 2006; Poser and Norris, 2009; Posse et al., 1999; Weiskopf et al., 2005). When using the optimal combination (OC) method (Posse et al., 1999), the weighting factor w of echo n is determined by the echo time TE and the estimated T_2^* of the voxel, according to:

$$w_n = \frac{TE_n exp\left(\frac{-TE_n}{T_2^*}\right)}{\sum_n^N TE_n exp\left(\frac{-TE_n}{T_2^*}\right)} \quad (2)$$

Assuming that noise is uncorrelated across echoes, and TE-independent, the corresponding CNR is then given by Poser et al. (2006):

$$CNR_{oc} = \frac{\sum_{n=1}^{N} w_n * TE_n * S_n / \sigma}{\sqrt{\sum_{n=1}^{N} w_n^2}} \quad (3)$$

When calculating the expected CNR of both protocols for an iron-rich area with a T_2^* value of 15 ms, and assuming that the SNR of the multi echo protocol is 0.57 times the SNR of the single echo protocol, we find that $\frac{CNR_{oc}}{CNR_{single}} = 0.78$. For cortical areas with T_2^* values of 33 ms (as is typical for cortical grey matter at 7 T, Van der Zwaag et al., 2016), and assuming that the SNR of the multi echo protocol is 0.62 times the SNR of the single echo protocol, $\frac{CNR_{oc}}{CNR_{single}} = 1.19$. Thus, these calculations suggest that the single echo protocol should have higher CNR in iron-rich subcortical areas, whereas the multi echo protocol should have higher CNR in cortical areas.

Several cautionary notes apply to these calculations, which could cause discrepancies between the theoretical estimations above and empirically observed data. Firstly, the calculations depend on a range of assumptions. The CNR estimations assume mono-exponential decay of transverse magnetization, and in the case of the multi echo protocol, uncorrelated noise across echoes and TE-independence of noise. Furthermore, it does not take into account susceptibility-gradient echo time shifts (see Volz et al., 2019, for an overview of these effects), which may further impact CNR.

Secondly, the calculations ignore the potential additional effects of multiband acceleration on SNR for the multi echo protocol (more on this in the Discussion), and the effects of ramp sampling. The latter may cause two effects which are not necessary or practically impossible, respectively, to take into account in these calculations. For one, a certain part of the outer *k*-space is acquired with a varying effective bandwidth due to the sampling with varying readout gradient amplitudes modifying the effective SNR and noise correlations in this area of k-space. However, since only the outer k-space is affected, which typically contributes little to the SNR compared to the k-space center, this would only have a small effect on our SNR and CNR metrics. Furthermore, ramp sampling can cause increased Nyquist (N/2) ghosting, since readout polarity dependent gradient imperfections are pronounced at points of dynamic gradient changes (e.g., at the transition between the ramp and flat top of the readout gradient). The ghosting can significantly degrade image quality. However, the dependence of the ghosting artifact on the extent of ramp sampling is non-trivial and cannot be predicted from the ramp sampling percentage based on established models, but depends on the details of the EPI implementation and scanner hardware. Therefore, we performed a visual inspection of the EPI time series in order to avoid severe Nyquist ghosting in order to minimize the second effect of ramp sampling.

Thirdly, the estimation cannot take parameters into account which depend on the specific subject. Certain scanning techniques, however, for instance parallel imaging or ramp sampling, may especially suffer from subject motion and/or breathing. Therefore, the multi echo protocol may depend more on subject behavior than the single-echo protocol. This results in a small effect for highly compliant volunteers but could be a significant cause for image degradation when scanning MR-naive subjects or patient groups.

2.4. Behavioral paradigm

The task used was an auditory stop-signal response task (Logan et al., 1984; Verbruggen et al., 2019). This task was chosen because it is known to elicit BOLD-responses both in cortical areas and in small, subcortical nuclei such as the subthalamic nucleus (e.g., Aron and Poldrack, 2006; De Hollander et al., 2017). In this task, the participant is presented with an arrow pointing left or right, and is instructed to respond as fast as possible without making a mistake, by pressing a button with the corresponding hand. On a subset of the trials (25%), an auditory tone is presented shortly after the arrow. On these trials, the participant is instructed to withhold responding. The time between the presentation of the arrow and the auditory stop signal (the stop-signal delay; SSD) is adjusted to the participant's performance on the task using a staircase procedure to obtain a successful inhibition rate of 50% (Verbruggen et al., 2019). Response times and response directions are recorded on each trial.

Trials are categorized as *go trials* (no stop signal presented), *failed stop* (a stop signal was presented, but the participant was not able to inhibit their response), or *successful stop* (a stop signal was presented, and the participant successfully inhibited their response). Due to the simplicity of the go trials, incorrect go responses (e.g., a left button press in response to an arrow pointing rightwards) are infrequently observed and are typically excluded from statistical analyses. In fMRI research using the stop-signal response task (e.g., Aron and Poldrack, 2006; De Hollander et al., 2017; Li et al., 2008), three contrasts are of main interest to study which brain areas are involved in stopping an initiated response: *failed stop* > *go*, *successful stop* > *go*, and *failed stop* > *successful stop*.

2.5. Procedure and exclusions

Before each session, the participant received instructions and practiced the task for approximately 10 min. Each session started with a short localizer scan, followed by two functional runs, a field map, the third functional run, and the anatomical scan. Each functional run took 17:09 min (343 vol). Each trial took 9 s (3 vol), and consisted of a fixation cross (with a jittered duration of 750 ms, 1500 ms, 2250 ms, or 3000 ms to decorrelate trial onsets), the stimulus (1 s), followed by a blank screen for the remainder of the trial duration. The total duration of a full session was approximately 1 h.

After the second run in the multi echo session, one participant reported not being able to hear the auditory stop signal. This participant was excluded from all further analyses. Further, one other participant did not perform the task significantly above chance level (50% correct answers in *go* trials) in the third run during both sessions due to fatigue. These runs were excluded from analyses. For one other participant, the third run was aborted after 12 min due to technical issues during the multi echo session. Of this participant, the data from the third runs of both sessions were excluded from analyses.

For five participants, there was no time left at the end of one of the sessions to finish the MP2RAGE scan. It was ensured that an MP2RAGE scan was always acquired in at least one of the two sessions. Finally, for one participant, the flip angle in the multi echo sequence had to be lowered to 60° to remain below the specific absorption rate (SAR) limit.

2.6. MRI data preprocessing

All functional data were preprocessed using *fMRIprep* (version 1.2.6; Esteban et al., 2019). The pipeline included slice timing correction using `3dTshift` from AFNI 20160207 (Cox and Hyde, 1997), head motion

correction using `mcflirt` (FSL 5.0.9, Jenkinson et al., 2002), susceptibility distortion correction using a custom workflow, realignment using `bbregister` (FreeSurfer; Greve and Fischl, 2009), and registration to the ICBM 152 Nonlinear Asymmetrical template version 2009c (henceforth "MNI2009c", Fonov et al., 2009) using `antsRegistration` (ANTs 2.2.0, Avants et al., 2008). Full pipeline details are included in Appendix A.

Prior to running *fMRIprep*, background noise in the MP2RAGE T1-weighted images was reduced using the robust MP2RAGE combination method (O'Brien et al., 2014). This was required to ensure that the skull-stripping routine included in *fMRIprep* can process the data. For the multi echo data, *fMRIprep* was modified to process each echo separately. For the second and third echoes, however, head motion correction and realignment were based on the parameters obtained from processing the first echo. This was done because the first echo has highest signal intensity and should lead to best estimation of head motion correction and realignment parameters (Speck and Hennig, 2001).

After running *fMRIprep*, all data were temporally filtered using a highpass filter (cut-off 1/128 Hz) to remove slow drifts. The multi echo data were then recombined into a single timeseries using a weighted sum, using the optimal combination (OC) method (Posse et al., 1999). In this method, the weighting factor w of echo n is determined by the echo time TE and the estimated T_2^* of the voxel, according to Equation (1). Voxelwise T_2^*-values were estimated by fitting a mono-exponential decay curve through the mean signal (across time) of the three echoes. Initial explorations showed that the results and conclusions would be highly similar using the parallel acquired inhomogeneity desensitized (PAID; Poser et al., 2006) method of combination (see also Kettinger et al., 2016).

2.7. Behavioral analyses

Trials in which the participants responded faster than 150 ms or slower than the stimulus duration of 1 s were excluded from analyses (0.20% of all trials). Then, for both sessions, we calculated the median RT on *go* trials and on *failed stop* trials, response accuracy, mean stop-signal delay (SSD), and proportion of successful stops. Further, for each participant, we estimated the amount of time it took to stop an initiated response: The stop-signal response time (SSRT). SSRTs were estimated using the mean method, which assumes that *go* and *stop* processes race independently from each other to a common threshold in a so-called horse race (Aron and Poldrack, 2006; Logan et al., 1984). Under this assumption, the SSRT can be calculated by finding the percentile of go RTs that corresponds to the proportion of successful stops (approximately 50%), and subtracting the mean SSD from this value.

2.8. Regions of interest (ROIs) analyses

Alongside the whole-brain analyses reported below, we are interested in several ROIs of the response inhibition network. We specifically study IFG, M1, preSMA, STN, GPe, and GPi. For all ROIs except preSMA, masks from probabilistic atlases were used. The masks of STN, GPe, GPi, and STR were obtained from the ATAG atlas (Young Adults; Keuken et al., 2014). The mask of primary motor cortex was obtained from the Brainnetome atlas ("upper limb area", Fan et al., 2016). The mask of IFG was extracted from the Harvard-Oxford cortical atlas (Desikan et al., 2006), identified as the union of the pars opercularis and pars triangularis. Finally, a discrete preSMA mask was based on coordinates provided by Johansen-Berg et al. (2004).

All masks were registered to MNI2009c space (and down sampled to 1.6 mm isotropic) using tools embedded in *nighres* (Huntenburg et al., 2018). For brevity, ROI analyses reported below are based on the right hemisphere only, as it has been suggested that the stopping network is right lateralized (Aron and Poldrack, 2006). Explorations of the left hemisphere ROIs suggest, however, similar patterns of BOLD-activity as their contralateral counterparts (see also Fig. 5 below).

2.9. Temporal signal to noise ratios (tSNR) and contrast to noise ratios (CNR)

For both protocols, as well as for the individual echoes of the multi echo protocol, we calculated the tSNR as the voxelwise ratio of the mean and the standard deviation of the signal across the timeseries. tSNR values for each ROI are a weighted mean of each voxels' tSNR, with the weighting factor equal to the voxels' probability of belonging to the ROI. For each subject and session, tSNR-values were first calculated per run, after which the tSNR-values were averaged across runs. When using the standard deviation of the residuals of the GLM analyses reported below as a noise measure (e.g., Weiskopf et al., 2005), the results are qualitatively similar and lead to the same conclusions.

Secondly, we estimated the expected relative BOLD contrast to noise ratio (CNR), where BOLD contrast is defined as the amount of signal change per unit change in T_2^* due to fluctuations in tissue oxygenation (Poser et al., 2006). For a single echo sequence, the CNR per unit change in T_2^* as a function of TE is given by Equation (1). Assuming mono-exponential decay, the measured signal decays with:

$$S(TE) = S_0 exp\left(\frac{-TE}{T_2^*}\right) \tag{4}$$

From (1) and (4), and under the assumption that noise is independent from TE, it follows that CNR can be rewritten as $CNR = \frac{S^*TE}{\sigma}$, which implies that CNR can be calculated as the voxelwise product of the tSNR and TE of the sequence (Poser et al., 2006). We note that we did not adjust for susceptibility-gradient related echo time shifts (Volz et al., 2019), which may further impact the BOLD sensitivity.

For a combined multi echo image, the CNR depends on the echo combination method (Gowland and Bowtell, 2007; Kettinger et al., 2016; Poser et al., 2006; Poser and Norris, 2009; Posse et al., 1999; Weiskopf et al., 2005). When using a weighted sum, the contrast can be calculated by taking the weighted sum of the contrast in all echoes (estimated in the same way as the contrast in a single echo image). The noise term can be estimated by taking the standard deviation of time series of the combined signal (Poser et al., 2006).

Note that these analyses depend on a range of simplifying assumptions (e.g., signal decay is mono-exponential, no susceptibility-gradient related echo time shifts, noise is Gaussian), which likely do not all hold. As a consequence, the BOLD-sensitivity estimates from the GLM analyses reported below can deviate from these analyses.

2.10. Signal deconvolution

To estimate the shape and size of the signal responses to each trial type in the experimental paradigm, we performed signal deconvolution analyses (Gitelman et al., 2003; Glover, 1999; Poldrack et al., 2011). The signals used were weighted mean timeseries within ROIs (without spatially smoothing the data), where voxel contributions to timeseries were weighted by their probability of belonging to the region according to the probabilistic masks. ROI time series were rescaled to percentage signal change by dividing each timepoint's signal value by the mean signal value (as baseline), multiplying by 100, and subtracting 100. Data from the three runs (per session per participant) were concatenated. Identical analyses were performed for the single echo data and the optimally combined multi echo data, as well as each of the individual echoes in the multi echo protocol.

For deconvolution, Fourier basis sets were used (Bullmore et al., 1996; Gitelman et al., 2003). Fourier basis sets provide a smoothness constraint, which decreases variance in the parameter estimates compared to finite impulse response (FIR) basis sets, at the cost of minimal parameter estimation bias (Liu et al., 2017; Poldrack et al., 2011). Per trial type, five basis sets were used (an intercept, two sines, and two

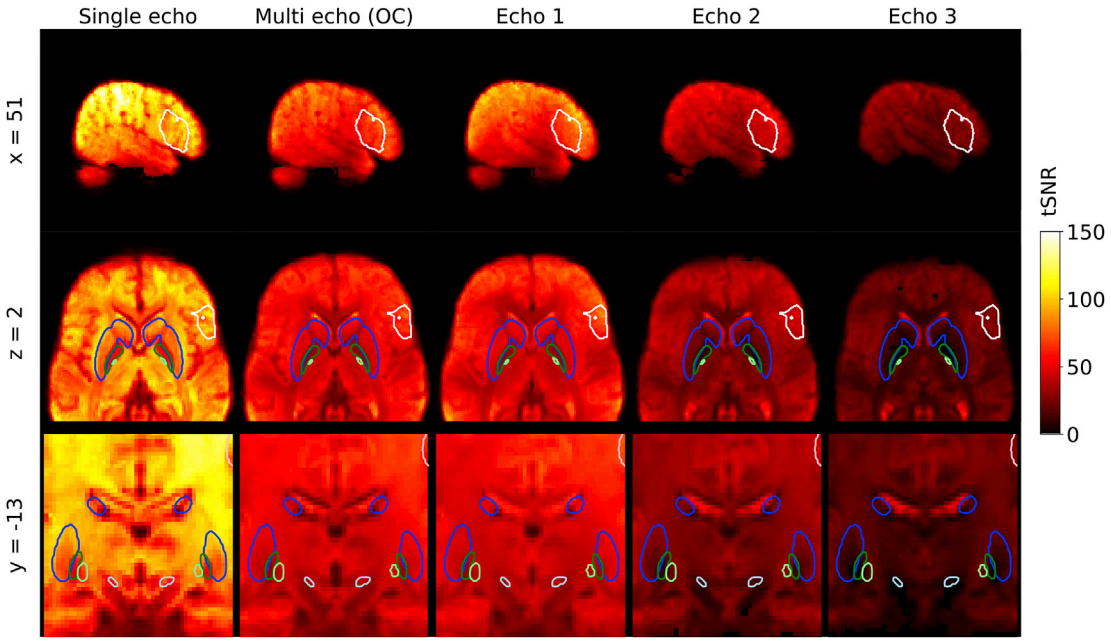

Fig. 1. Whole-brain tSNR values. Contours indicate regions of interest, thresholded at a 30% probability level (rIFG in white, striatum in blue, GPe and GPi in dark and light green, respectively, and STN in light blue). Slice locations are in MNI2009c space (1 mm).

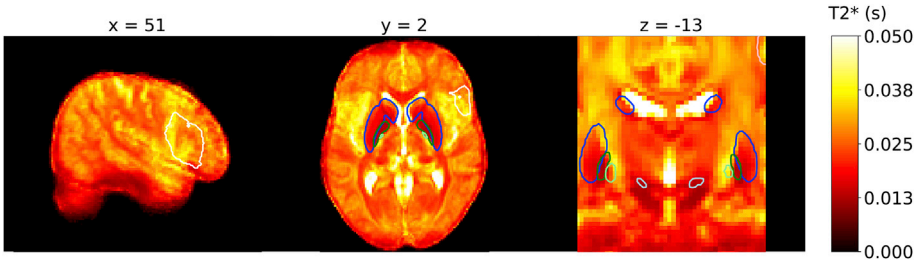

Figure 2. T_2^* values (in seconds) across the brain, obtained by fitting a mono-exponential decay curve through the mean signal (across volumes per run) per echo, and averaged across participants. Contours indicate regions of interest, thresholded at a 30% probability level (rIFG in white, striatum in blue, GPe and GPi in dark and light green, respectively, and STN in light blue). Slice locations are in MNI2009c space.

cosines), totaling 16 regressors (three trial types plus an overall intercept) in the design matrix. The model was fit to the data using ordinary least squares. We also fit a second deconvolution model with seven additional motion regressors (rotation and translation in three dimensions, and framewise displacement), but as the results were highly similar, we report only the results without these additional motion regressors. Deconvolution analyses were performed using *nideconv* (https://github.com/VU-Cog-Sci/nideconv).

We visualize signal time courses (in seconds after trial onset) expressed as percentage signal change. Due to the TE-dependence of percentage signal change (Kundu et al., 2012; Menon et al., 1993), it is expected that percentage signal change is higher in the optimally combined multi echo data (which has, on average, higher TEs) than in the single echo data. However, since the amount of noise in the data is also expected to increase with TE, the percentage signal change itself is not a pure estimate of statistical detection power. To quantify the noise component, we calculated σ^2 (the sum of squared errors divided by the degrees of freedom) of the models. Formal statistical comparisons between the two protocols are performed using GLMs, which we turn to in the next section.

2.11. General linear models (GLMs)

To statistically assess differences between protocols, we first fit a set of GLMs using the canonical double-gamma hemodynamic response function (HRF) and temporal derivatives (Friston et al., 1998, 1995; 1994; Glover, 1999). In the design matrix, the three trial types of the experimental paradigm were included as regressors of interest. Further, six motion parameters (translation and rotation in three dimensions) obtained from head motion correction and the framewise displacement were included in the design matrix as confounding variables. Including the intercept, this leads to a total number of 13 regressors. The main contrasts of interest were *failed stop* > *go*, *successful stop* > *go*, and *failed stop* > *successful stop*.

GLM analyses were done both whole-brain and ROI-based (using the weighted mean of the signal timeseries within each ROI). Whole-brain GLMs were performed using FSL FEAT (Jenkinson et al., 2012; Woolrich et al., 2001) using the FILM method to account for autocorrelated residuals. Prior to fitting the GLMs (for the whole-brain analyses only) the data were spatially smoothing using a Gaussian smoothing kernel (FWHM = 5 mm). This spatial smoothing decreases the effective resolution considerably, but it allows for comparison of the overall whole-brain GLM with previous studies (notably, De Hollander et al., 2017). Fixed-effects GLMs were used to combine data from different runs per subject and session. FLAME1 (Woolrich et al., 2004) was subsequently used to analyze the group-level results per session separately. Statistical parametric maps (SPMs) were subsequently created to visualize the whole-brain results of the group-level models. All *p*-values were

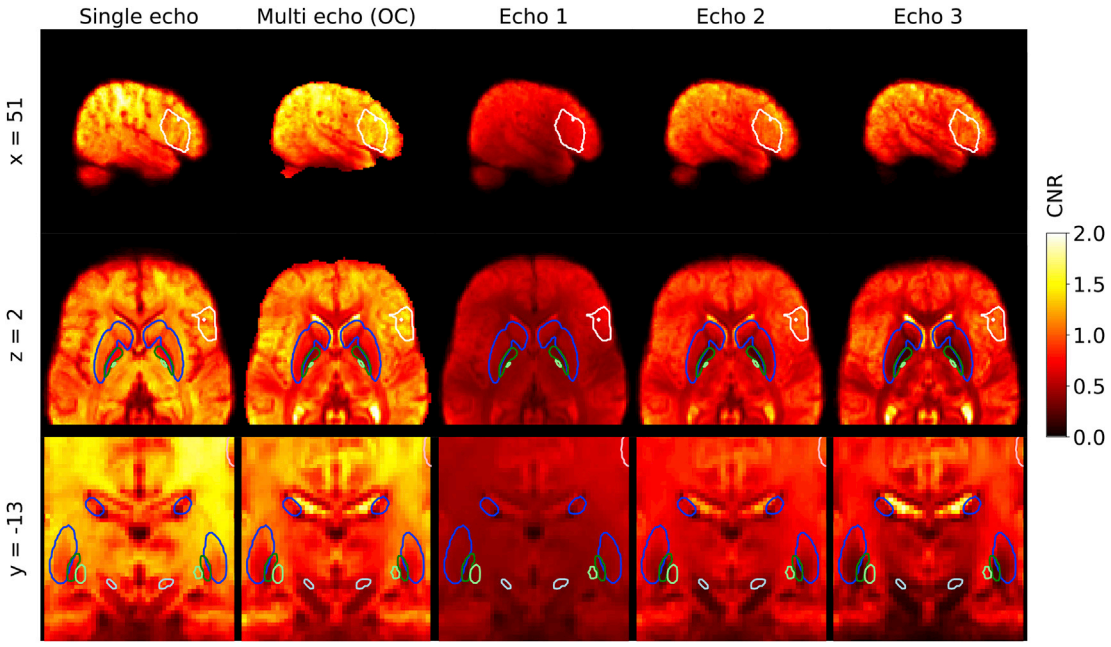

Fig. 3. Whole-brain CNR values. Contours indicate regions of interest, thresholded at a 30% probability level (rIFG in white, striatum in blue, GPe and GPi in dark and light green, respectively, and STN in light blue). Slice locations are in MNI2009c space.

corrected for the false discovery rate (FDR; Yekutieli and Benjamini, 1999), and we used a critical value of $q < 0.05$ to determine significance.

ROI-based GLMs were done by first concatenating all runs within a session per participant (after rescaling each run's timeseries to percent signal change). Unlike the whole-brain GLMs, these data were not spatially smoothed to ensure that signal within each ROI is not contaminated with signal from outside the ROI. Then, we fit the GLM using an *AR(1)*-model as implemented in Python package *statsmodels* (Seabold and Perktold, 2010) to account for temporal autocorrelation in the timeseries. The whitened residuals were inspected to ensure there were no between-protocol differences in the remaining autocorrelation, which could bias *t*-values. GLMs were fit to data from both sessions separately.

Statistical analyses comparing the two protocols were based on the analyses presented in Puckett et al. (2018). From the fitted GLMs, we extracted *t*-statistics per subject and session (from the fixed effects analysis) as a measure of BOLD-sensitivity for each trial type (cf. Puckett et al., 2018). To statistically compare protocols, we used a linear mixed effects model (Gelman and Hill, 2007) with these *t*-values dependent variables, and ROI, trial type, and protocol as independent variables including all possible interaction terms. ROI was included as an independent variable to be able to study whether the protocols differ in their BOLD-sensitivity across ROIs. Trial type was included in as an independent variable because it is unlikely that all ROIs show the same BOLD-response in all trial types. Alternatively, it would have been possible to analyze only one of the three trial types, but this would require an arbitrary choice as to which trial type is analyzed. A second linear mixed effects model with the same independent variables was fit using the *t*-values of the contrasts between trial types as the dependent variable, to study whether contrasts (which have smaller effect sizes than events against baseline) can be detected more easily with either protocol.

Random intercepts were included for subjects. To ensure larger *t*-values always indicate a larger effect, the sign of *t*-values of terms for which the group effect was negative was flipped. Homogeneity of the residuals as well as the normality of residuals were assessed. Follow-up paired *t*-tests (one per combination of ROI, trial type and data type) were used to provide detail on the possible interaction terms. The reported *q*-values are *p*-values adjusted for the false discovery rate (critical value $q < 0.05$).

In frequentist tests, the lack of a significant finding cannot be interpreted as evidence against the presence an effect (e.g., Wetzels and Wagenmakers, 2012). To allow for quantifying evidence against effects, we repeated the statistical analyses using Bayesian linear models (Liang et al., 2008; Rouder and Morey, 2012) as implemented in the R package BayesFactor (version 0.9.12–4.2; Morey et al., 2018).

Exploratory analyses were included to assess how BOLD-sensitivity changes with TE in the various ROIs. To do so, we first fit, for every participant and ROI, a linear regression model predicting the *t*-value using the echo time as predictor. The slope of the resulting model, $\Delta t/TE$, indicates the amount of change in *t*-values per millisecond increase in TE. Values of $\Delta t/TE$ were calculated for each trial type (vs baseline), and then averaged across trial types per participant and ROI. Then, for every participant, we used the T_2^* per ROI (obtained by fitting monoexponential decay functions through ROI timeseries per echo) to predict $\Delta t/TE$. A one-sample *t*-test was used to test whether the slope of this model is different from 0 on a group level.

3. Results

3.1. Behavioral analyses

Descriptive statistics summarizing the behavioral data are presented in Table 1. In both sessions, the data are typical for the stop-signal response task (e.g., De Hollander et al., 2017; White et al., 2014). The percentages of successful inhibition are close to 50%, indicating that the staircasing procedure successfully adjusted SSDs to participants' SSRTs. Between sessions, no significant differences were observed in any of the dependent variables after multiple comparison correction (using the FDR with critical value $q < 0.05$). Further, median *go* RTs were not correlated with SSRTs, in line with the assumption of independence between the go and stop processes in the horse-race model (Logan et al., 1984).

3.2. tSNR and CNR

As a first assessment of functional data quality, we analyzed the temporal signal-to-noise ratios. Fig. 1 illustrates the whole-brain tSNR for each of the protocols. tSNR values within each ROI are also summarized in Table 2. The single echo tSNRs are in line with earlier findings using a

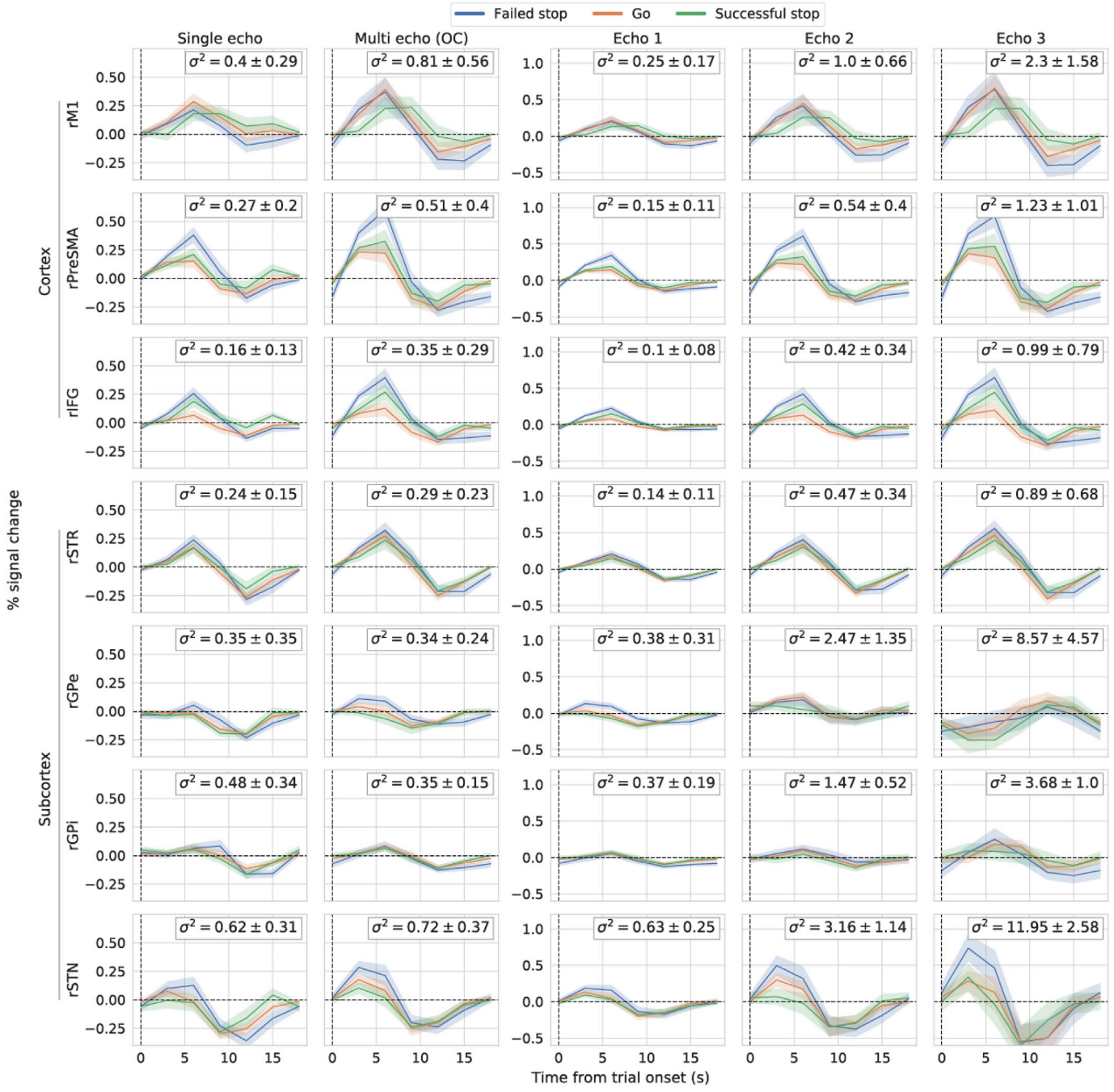

Fig. 4. Deconvolved hemodynamic responses to the three trial types of interest, expressed as percentage signal change. Shaded areas indicate a 67% confidence interval, obtained by bootstrapping with 1000 bootstrap replications. σ^2 values are mean across participants (±SD).

similar protocol (De Hollander et al., 2017). Across the all the ROIs, the tSNR of the single echo protocol is higher than the tSNR of the optimally combined multi echo protocol (smallest difference observed in rIFG, $t(16) = 11.08$, $p < 0.001$). In the multi echo protocol, the decrease in tSNR across the three echoes deviated only marginally from the decrease in tSNR that is expected based on T_2^* decay (results not shown).

Contrast-to-noise depends on both the tSNR and the ratio between the TE and the T_2^* of the tissue (Eq. (1)). The short TE in the single echo sequence may have resulted in a high tSNR, but the TE is much shorter than the T_2^* in many brain areas (Fig. 2 provides a T_2^* map to illustrate the between-region variability in T_2^*). For such brain areas, the single echo protocol samples from the T_2^*-decay curve too early, which can lead to decreased BOLD contrast-to-noise. The BOLD CNR of the optimally combined multi echo data is expected to profit from the additional samples from the T_2^*-decay curve.

To provide a comparison between the protocols that considers both tSNR and TEs, we estimated the theoretically expected relative CNR-values. Fig. 3 shows the voxel-wise CNR-values across the entire brain. Table 3 summarizes these CNR-values for the ROIs. The results are shown in Table 3 and Fig. 3. The expected CNR in rPreSMA and rIFG is higher in the optimally combined multi echo data ($t(16) = 6.54$, $p < 0.001$ and $t(16) = 4.76$, $p < 0.001$, respectively), but the single echo data has a higher expected CNR in subcortical areas (smallest difference for rSTR, $t(16) = 8.15$, $p < 0.001$).

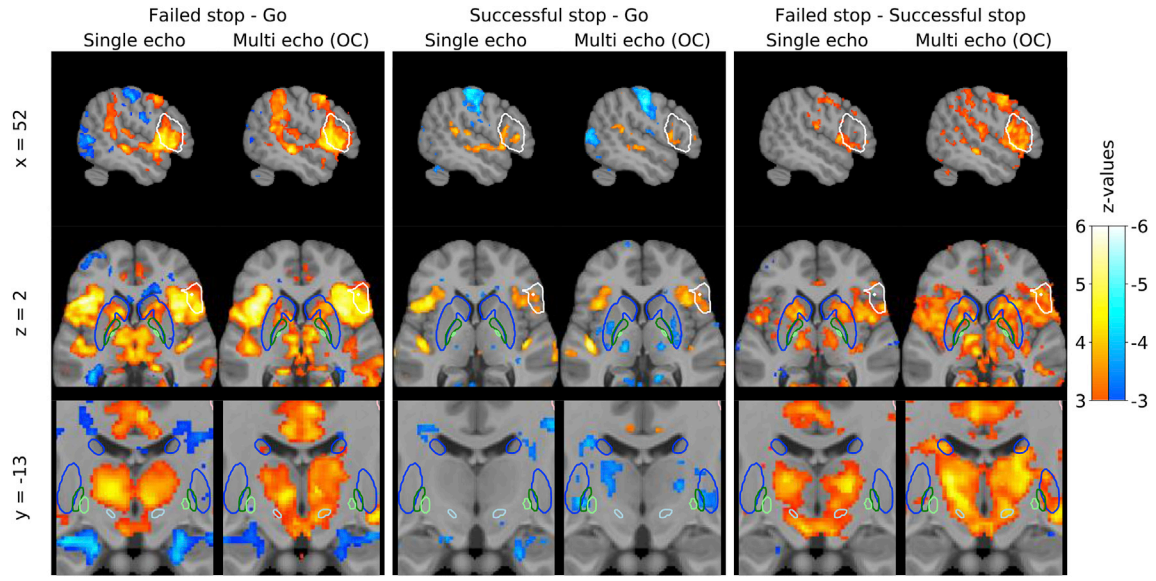

Fig. 5. Group-level SPMs of the three main contrasts of interest for both the single echo and optimally combined multi echo data. Colors indicate *z*-values. Thresholds were set using the FDR method ($q < 0.05$). Since this method determines the critical *z*-value based on the distribution of observed *z*-values, the critical values obtained for each contrast differs between protocols. To prevent within-contrast varying thresholds between protocols, the highest of the two critical *z*-values (per contrast) was chosen as the threshold. Contours indicate regions of interest, thresholded at a 30% probability level (rIFG in white, striatum in blue, GPe and GPi in dark and light green, respectively, and STN in light blue). Slice locations are in MNI2009c space (1 mm).

Table 1
Descriptives of the dependent variables of interest for the behavioral analyses.

	Median go RT (SD), ms	Median failed stop RT (SD), ms	% errors (SD)	SSRT (SD), ms	SSD (SD), ms	% successful inhibition (SD)
Single echo	438.82 (63.54)	424.70 (56.64)	0.64 (0.68)	233.50 (83.54)	206.79 (94.49)	50.95 (7.07)
Multi echo	463.42 (78.49)	446.30 (65.32)	0.46 (0.63)	229.56 (91.54)	232.30 (108.48)	48.81 (7.08)

3.3. Signal deconvolution

Deconvolution analyses were performed to estimate the signal responses to the three trial types. Deconvolved responses are shown in Fig. 4. In every ROI, the deconvolved responses, while not identical, look qualitatively similar between protocols. With the exception of GPe and GPi, the amount of signal increases until approximately 7–10 s after stimulus onset, followed by a post-stimulus undershoot until approximately 15 s after stimulus onset. The deconvolved signal of the individual echoes show clear TE-dependence (again with the exception of GPe and GPi) with higher percentages signal change observed at later TEs. The TE-dependence of the observed responses indicates that the signal changes are due to changes in T_2^*, and thus likely BOLD-responses (Kundu et al., 2012; Menon et al., 1993). Since there appears to be limited TE-dependency in GPe and GPi, the data observed (at least in the later echoes) are likely largely noise and not BOLD-signal.

Table 3
Mean (SD) CNR across participants, specified per ROI and per protocol.

	Cortex			Subcortex			
Data/ROI	rM1	rPreSMA	rIFG	rSTR	rGPe	rGPi	rSTN
Single echo	1.30 (0.22)	1.22 (0.19)	1.16 (0.22)	1.01 (0.15)	0.78 (0.13)	0.82 (0.13)	0.84 (0.12)
Multi echo (OC)	1.34 (0.18)	1.40 (0.18)	1.28 (0.20)	0.81 (0.09)	0.46 (0.05)	0.51 (0.07)	0.53 (0.05)
Echo 1	0.69 (0.07)	0.65 (0.06)	0.64 (0.07)	0.43 (0.04)	0.31 (0.04)	0.31 (0.04)	0.33 (0.04)
Echo 2	1.01 (0.11)	1.05 (0.11)	0.98 (0.14)	0.57 (0.07)	0.27 (0.04)	0.31 (0.05)	0.36 (0.04)
Echo 3	1.00 (0.11)	1.11 (0.14)	0.97 (0.15)	0.49 (0.07)	0.22 (0.02)	0.26 (0.03)	0.25 (0.03)

Interestingly, the data suggest a difference in the post-stimulus undershoot amplitude between the cortical and subcortical responses. Whereas the amplitude of the undershoot in cortical areas is approximately 1/5th of the peak amplitude (in line with the canonical HRF; Glover, 1999), the undershoot amplitude in subcortical areas (rSTN, rSTR) is roughly equal to the peak. This pattern is present in the single echo data, in the optimally combined multi echo data, and in each of the individual echoes where it scales with TE. Further, we also observed it in the left hemisphere ROIs (results not shown), and it appears present in earlier findings as well (De Hollander et al., 2017). We briefly return to this observation in the Discussion.

A second notable aspect of the responses is that there appears to be a

Table 2
Mean (SD) tSNR-values across participants, specified per ROI and per protocol.

	Cortex			Subcortex			
Data/ROI	rM1	rPreSMA	rIFG	rSTR	rGPe	rGPi	rSTN
Single echo	91.06 (16.25)	86.74 (13.48)	81.63 (15.52)	72.13 (10.86)	55.38 (9.76)	58.55 (9.02)	59.87 (8.29)
Multi echo (OC)	60.49 (8.63)	60.43 (7.51)	58.21 (8.96)	43.52 (4.15)	31.54 (3.66)	32.54 (4.16)	34.56 (3.43)
Echo 1	72.66 (8.09)	67.08 (6.58)	65.87 (7.60)	44.82 (3.88)	32.03 (3.82)	32.37 (4.31)	34.55 (4.07)
Echo 2	40.61 (4.29)	42.18 (4.43)	39.56 (5.47)	23.11 (2.72)	10.89 (1.55)	12.46 (1.89)	14.28 (1.59)
Echo 3	24.95 (2.86)	27.60 (3.58)	24.08 (3.79)	12.28 (1.81)	5.55 (0.60)	6.47 (0.80)	6.16 (0.73)

delay in the hemodynamic response in rM1 during Successful Stop trials (onset roughly 3 s later compared to the other two trial types). Since the delay appears present in rM1 only, but in both protocols, we assume it is both task and region-specific. Furthermore, the delay also appears present in the deconvolved hemodynamic response of the same contrast observed by Aron and Poldrack (2006, their Fig. 2).

Due to the (on average) higher TEs of the multi echo sequence, the optimally combined multi echo BOLD-responses reach a higher peak-level percent signal change than the responses in the single echo data. However, this mean increase in signal is accompanied by an increase in model misfit, as quantified by σ^2. As a consequence, the higher percentage signal change does not necessarily translate into an increase in detection power. In the next section, we formally compare protocols.

3.4. GLMs: Whole-brain results

We first fit GLMs using the canonical HRF to each voxel. Group-level whole-brain statistical parametric maps (SPMs) of the contrasts of interest (*failed stop – successful stop*, *successful stop – go,* and *failed stop – go*) are shown in Fig. 5.

Both protocols lead to highly similar SPMs. Clusters of differential BOLD-responses between *failed stop – successful stop* are present in IFG, insula, preSMA, dorsal anterior cingulate cortex (dACC), and parietal cortex. The single echo data further appears to show a negative cluster around M1, which is not visible in the combined multi echo data. Subcortically, a differential BOLD-response is present in thalamus, extending to substantia nigra (SN). The subcortical BOLD-responses are widespread and not confined to individual subcortical nuclei, as has been observed before (De Hollander et al., 2017).

The *successful stop – go* contrast shows relatively little differential BOLD-responses. There is a small positive cluster in IFG and insula, which appears slightly stronger in the single echo data than in the combined multi echo data. Both data sets show a negative cluster around M1 for this contrast. Finally, the pattern of clusters in the *failed stop – successful stop* cluster is similar to (although somewhat weaker than) the *failed stop – go* contrast, with clusters of activation present in IFG, preSMA, ACC, thalamus, STN, and SN. The clusters in this contrast appear more widespread and stronger in the combined multi echo data than in the single echo data. Note, however, that the *failed stop – successful stop* contrast is based on least trials due to the design of the stop-signal response paradigm (approximately 12.5% of all trials are failed stops, and another 12.5% of all trials are successful stops).

3.5. GLMs: ROI analyses

To statistically quantify the differences between the protocols, we first fit a set of GLMs using the canonical HRF to the timeseries within each ROI. Per session, for each participant and ROI, we calculated the *t*-values of each trial type regressor against baseline (Fig. 6, diagonal). Using a linear mixed effects model, we subsequently tested whether ROI, protocol, and trial type (and interaction effects) had an effect on the *t*-value. The results are shown in Table 4 (upper half). Reported *q*-values are *p*-values adjusted for the false discovery rate (Yekutieli and Benjamini, 1999), although the exact same conclusions would be reached without correcting for multiple comparisons. Most importantly, there was no evidence for a main effect of protocol on *t*-values, in line with the marginal difference between mean *t*-values for both protocols (95% confidence intervals: single echo [2.48, 5.20], combined multi echo [2.47, 5.19]). There was also no evidence for an interaction effect of protocol with other terms in the model.

Furthermore, follow-up paired *t*-tests per trial type and ROI showed no significant differences between protocols after multiple comparison corrections. Only without multiple comparisons corrections (and assuming an alpha-level of 0.05), one would have concluded that there was an increase in *t*-values in the multi echo protocol compared to the single echo protocol in rPreSMA for *failed stop* ($t(16) = -2.242$, $p = 0.04$) and *successful stop* ($t(16) = -2.50$, $p = 0.024$) trials. Vice versa, without multiple comparison correction, one would have concluded there was a decrease in *t*-values in the multi echo protocol in rGPe for *go* ($t(16) = 3.184$, $p = 0.006$) and *successful stop* trials ($t(16) = 2.493$, $p = 0.024$).

To quantify the strength of the evidence against an effect of protocol on first-level *t*-values, we computed Bayes factors by comparing linear

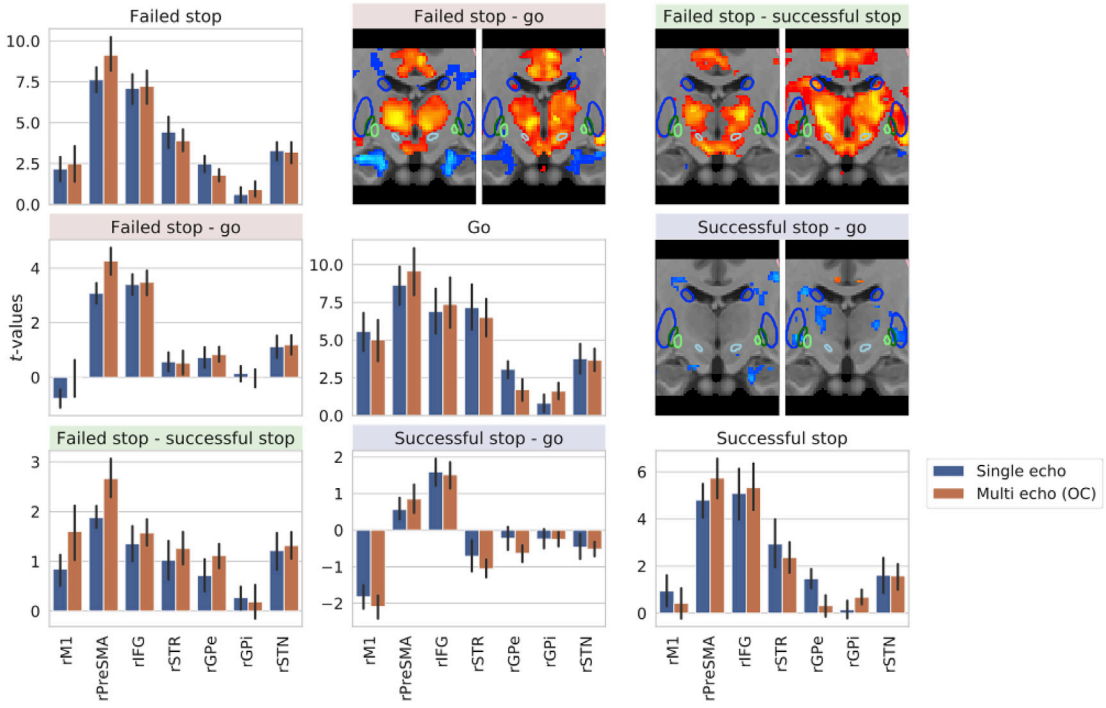

Fig. 6. Comparison of *t*-values between protocols. Bar graphs illustrate *t*-values per ROI for the main effects (diagonal) and contrasts (lower left). Error bars indicate 67% confidence intervals, obtained by bootstrapping with 1000 bootstrap replications. SPMs (y = −13) are included from Fig. 5 for comparison.

Table 4
Frequentist *F*-tests on the effect of terms in the linear mixed effects models, using Satterthwaite's method (Satterthwaite, 1941) to approximate the denominator degrees of freedom. Both models included random intercepts for subjects. *q*-values are *p*-values adjusted for the false discovery rate, and are considered significant if $q < 0.05$. The same conclusions would be reached without any multiple comparison correction.

Dependent variable	Term	Frequentist test
t-value of trial type	Protocol	$F(1, 656) = 0.0001$, $q = 0.99$
	Trial type	$F(2, 656) = 56.47$, $q < 0.001$
	ROI	$F(6, 656) = 80.23$, $q < 0.001$
	Protocol x Trial type	$F(2, 656) = 0.19$, $q = 0.99$
	Protocol x ROI	$F(6, 656) = 1.73$, $q = 0.20$
	ROI x Trial type	$F(12, 656) = 3.13$, $q < 0.001$
	ROI x Trial type x Protocol	$F(12, 656) = 0.09$, $q = 0.99$
t-value of contrasts	Protocol	$F(1, 656) = 3.16$, $q = 0.13$
	Contrast	$F(2, 656) = 7.92$, $q < 0.001$
	ROI	$F(6, 656) = 26.25$, $q < 0.001$
	Protocol x Contrast	$F(2, 656) = 0.53$, $q = 0.68$
	Protocol x ROI	$F(6, 656) = 0.85$, $q = 0.68$
	ROI x Contrast	$F(12, 656) = 9.99$, $q < 0.001$
	ROI x Contrast x Protocol	$F(12, 656) = 0.50$, $q = 0.92$

models (with subjects as random effects). The results are presented in Table 5 (upper half). The model with the overall highest Bayes factor (against a model including the intercept only) included main effects of trial type and ROI, and an interaction between trial type and ROI. Model comparison showed that there was substantial to very strong evidence against an additional effect of protocol and/or interactions between protocol and ROI on *t*-values.

Finally, we repeated both the frequentist as the Bayesian analyses using the three contrasts of interest as independent variables, and the corresponding *t*-values as dependent variable (Fig. 6, lower left triangle). The main results are the same as in the previous analyses. The frequentist tests (Table 4, lower half) showed no evidence for an effect of protocol (or interactions between other terms) on these *t*-values (95% confidence intervals: single echo [0.82, 1.34], combined multi echo [1.02, 1.54]). Similarly, the Bayesian model with the highest overall Bayes Factor included a main effect of ROI, contrast, and an interaction between ROI and contrast. Bayesian model comparison indicated that there was anecdotal to decisive evidence against an additional effect of protocol and/or interactions between protocol and ROI. In summary, the differences between the two protocols were marginal.

Using the individual echoes of the multi echo protocol (Fig. 7, left panel), we tested the effect of echo time on BOLD-sensitivity for each ROI. To do so, for each subject and ROI, we first used a linear model to estimate $\Delta t/TE$; the change in *t*-values per millisecond increase in TE. Then, for each subject, we used a second linear model to predict $\Delta t/\ TE$ based on T_2^* value of the ROI. Note that the true effect of echo time on *t*-values is likely not linear; however, with three echoes, a linear model is the best approximation we can obtain with the current data.

The results are shown in Fig. 7 (right panel), and indicate an overall positive relation (one-sample *t*-test of the slope parameter: $t(16) = 4.369$, $p < 0.001$). For regions with low T_2^* (e.g., subcortical areas), BOLD-sensitivity decreased with TE, whereas for areas with higher T_2^*, BOLD-sensitivity increased (or remained approximately the same) with higher TE. Note, however, that visual inspection of Fig. 7 suggests that this linear effect is mainly driven by ROIs with low T_2^* values. This seems supported by follow-up *t*-tests per ROI, which show (after correction for the FDR) significantly negative $\Delta t/TE$ for rGPe ($t(16) = -3.11$, $q = 0.024$), rGPi ($t(16) = -3.47$, $q = 0.022$), and rSTN ($t(16) = -2.72$, $q = 0.04$), but no significant changes in *t*-values across TEs for rSTR, rM1, and rPreSMA. Only for rIFG, there was a significantly positive $\Delta t/\ TE$ ($t(16) = 2.55$, $q = 0.037$). In summary, the current data show similar *t*-values across TEs for most cortical (high T_2^*) areas, which is in line with what has been observed before (Gorno-Tempini et al., 2002; Hyde et al., 2001; Weiskopf et al., 2005). This may explain why the single echo sequence still provides good BOLD-sensitivity in cortical areas despite the low TE.

Table 5
Results of the Bayesian linear mixed effects models. All models included random intercepts for subjects. H_0 was the overall winning model (against a model including only an intercept). Bayes Factors (BF_{01}) quantify the relative likelihood of the data under two models; e.g., a $BF_{01} = 10$ entails that the data is 10 times more likely under the null hypothesis H_0 than under the alternative hypothesis H_1. Verbal qualifications of the strength evidence derived from the Bayes Factors ("anecdotal" to "decisive" evidence) are based on taken from Wetzels and Wagenmakers (2012), who adjusted Jeffreys' (1961) original formulation.

Dependent variable	Model	Bayesian test
t-value of trial type	H_0: $t \sim$ ROI + Trial type + ROI x Trial type	
	H_1: $t \sim$ ROI + Trial type + ROI x Trial type + Protocol	$BF_{01} = 11.78 \pm 1.52\%$ 'strong evidence' for H_0
	H_1: $t \sim$ ROI + Trial type + ROI x Trial type + Protocol x ROI	$BF_{01} = 7.98 \pm 0.71\%$, 'substantial evidence' for H_0
	H_1: $t \sim$ ROI + Trial type + ROI x Trial type + Protocol + Protocol x ROI	$BF_{01} = 98.67 \pm 0.88\%$, 'very strong evidence' for H_0
t-value of contrasts	H_0: $t \sim$ ROI + Contrast + ROI x Contrast	
	H_1: $t \sim$ ROI + Contrast + ROI x Contrast + Protocol	$BF_{01} = 2.50 \pm 0.58\%$, 'anecdotal evidence' for H_0
	H_1: $t \sim$ ROI + Contrast + ROI x Contrast + Protocol x ROI	$BF_{01} = 62.43 \pm 0.83\%$, 'very strong evidence' for H_0
	H_1: $t \sim$ ROI + Contrast + ROI x Contrast + Protocol + Protocol x ROI	$BF_{01} = 155.72 \pm 0.88\%$, 'decisive evidence' for H_0

4. Discussion

When studying both cortical and subcortical structures using conventional single echo 2D EPI fMRI protocols, researchers are faced with a tradeoff between BOLD-sensitivity in the cortex and in the subcortex. Here, we studied whether a multi echo EPI protocol can be used to avoid this tradeoff and obtain optimal BOLD-sensitivity across the brain. We directly compared a multi echo protocol with a single echo protocol, both optimized for subcortex by using a relatively high spatial resolution combined with short echo times. Our results showed only marginal differences between protocols, which both led to similar conclusions about BOLD-activity in the stop-signal task. The relatively small difference between protocols appears to be caused by two factors.

Firstly, the advantage of combining multiple volumes in the multi echo sequence is counteracted by the relatively low tSNR per echo. Contrary to earlier empirical studies that employed different experimental designs (Poser and Norris, 2009; Posse et al., 1999; Puckett et al., 2018), this resulted in a CNR that was similar to (in cortical areas) or even lower than (in subcortical areas) the CNR of the single echo sequence. An important difference with previous experiments is the amount of acceleration used. Earlier studies typically kept the amount of acceleration identical between sequences and minimized TR, whereas we kept the TRs identical and minimized the amount of acceleration. Acquiring multiple echoes after one excitation pulse with an isotropic spatial resolution of 1.6 mm required the otherwise disadvantageous usages of highly accelerated imaging and an increase in the receiver bandwidth. An increased acceleration factor not only decreases the number of acquired k-space lines but also increases the so-called g-factor penalty (Larkman and Nunes, 2007) especially at the center of the brain. Increasing the GRAPPA acceleration from iPAT = 3 to 4 while using our 32-channel phased array headcoil led to a g-factor increase of approximately 30% at the position of the STN (see Appendix C). This is in line with similar technical setups using GRAPPA (Salomon et al., 2014) or SENSE (Sengupta et al., 2016) for unfolding the aliased data (although differences in slice location, phase encoding direction, and iPAT factor hinder an exact comparison). As a result, both the reduced amount of

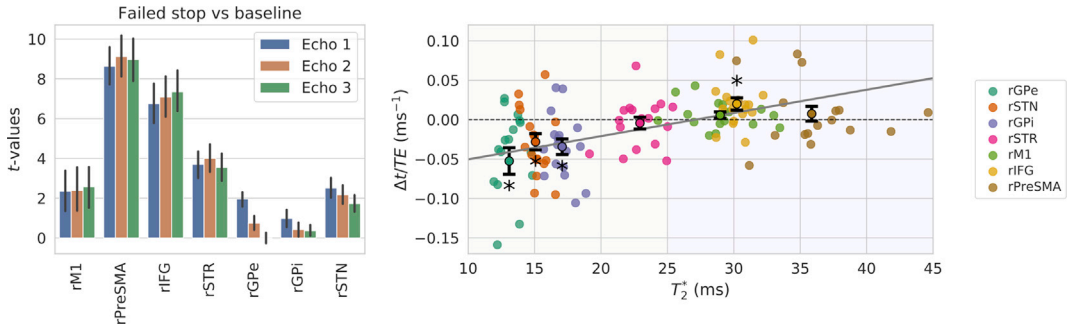

Fig. 7. Left panel: First-level ROI-wise GLM *t*-values for each echo separately, for the *failed stop vs baseline* contrast. The other contrasts show similar patterns (results not shown). Right panel: The effect of T_2^* on the change in first-level *t*-values per millisecond increase in TE. Negative values on the y-axis indicate that *t*-values decreased with TE, positive values indicate increases in *t*-values with TE. Colors indicate ROIs, background colors roughly separate subcortex and cortex. The group-

acquired k-space and the increased g-factor penalty lead to reduced SNR in highly accelerated acquisitions (see equation 8 in Larkman and Nunes, 2007) and finally to reduced *t*-values in the activation maps obtained with the multi echo approach. The multiband factor 2 in the multi echo protocol likely also had a disadvantageous effect on the SNR, although previous studies at different field strengths showed that the effect of low multiband factors is relatively small, especially when CAIPI shifts are used (Demetriou et al., 2018; Preibisch et al., 2015; Setsompop et al., 2012; Todd et al., 2017, 2016; Uğ;urbil et al., 2013; Xu et al., 2013), compared to the effect of increasing iPAT from 3 to 4. Also in combination with iPAT 2, the effect of low multiband factors of 2–3 has been shown to be relatively limited (Demetriou et al., 2018; Setsompop et al., 2012), although more impact on g-factors is expected in combination with iPAT factors of 3–4. In addition to in-plane and slice acceleration, the increased receiver bandwidth in the multi echo protocol (compared to the single echo protocol) resulted in an increased amount of ramp sampling. This effect is discussed in section 2.3. The complex interactions between acceleration and physiology leads to highly non-uniform effects on BOLD-sensitivity across the brain, often with a large decline in BOLD-sensitivity in subcortical areas with higher acceleration (Todd et al., 2017).

Secondly, the CNR of the single echo protocol in cortical areas was higher than what may have been expected based on its short TE. Per-echo analyses of the multi echo data also showed similar first-level *t*-values across TEs for most cortical regions, in line with previous reports (Gorno-Tempini et al., 2002; Hyde et al., 2001; Weiskopf et al., 2005). In particular, it has been suggested that physiological noise underlies the relatively small effect of echo time on BOLD-sensitivity (Van De Moortele et al., 2008).

We specifically compared the BOLD-sensitivity of two protocols with high spatial resolution and whole-brain coverage. Our task-based analyses indicate that, under these requirements, the protocols provided similar BOLD-sensitivity. However, the single echo protocol used in our study could be optimized further and has the potential for additional gains in statistical power. With the amount of acceleration used in the present study, for example, the temporal resolution can already be increased, and gains in statistical power may potentially be obtained by the use of multiband acceleration (Larkman et al., 2001; Moeller et al., 2010). In contrast, although the contrast-to-noise of the multi echo protocol clearly benefitted from the combination of multiple echoes, the individual echoes themselves had relatively low tSNR. Further increasing the amount of acceleration by increasing the multiband factor to 3 (or higher), while retaining the GRAPPA factor of 4 (which could not be decreased without substantially increasing the echo times), would likely worsen the conditioning of the parallel imaging solution and increase temporal instabilities and therefore decrease the tSNR. This would thus be disadvantageous for the contrast-to-noise, although gains in statistical power might be obtained from the increased total amount of volumes.

One may have specific reasons to prefer multi echo acquisition. One example is the reduced susceptibility-related signal loss compared to standard single echo 2D EPI sequences (Kirilina et al., 2016). Another reason may be the promise of denoising techniques. Recent papers (Kundu et al., 2017, 2012) suggested that denoising methods based on multiple echoes can significantly benefit statistical analyses by filtering out components in signal that are not TE-dependent and thus unlikely to be caused by fluctuations in tissue T_2^*. Unfortunately, the application of the current publicly available code for this technique on our data led to no detected BOLD-signals. We speculate that this may be caused by the combination of high spatial resolution and low temporal resolution (1.6 mm and 3 s vs. 2.5 mm and 1.8 s in Kundu et al., 2017), which leads to lower tSNR per voxel, as well as fewer volumes (data points) to assess TE-dependence. Thus, we think that the inability to detect BOLD-responses is likely due to the peculiar type of data we collected, more than a shortcoming of the denoising approach, and hope that the techniques are further developed to handle this type of data.

Deconvolved cortical and subcortical responses showed an interesting difference in both experiments: The undershoots of subcortical BOLD-responses appear large relative to the undershoots of the cortical BOLD-responses. This was observed in the data of both sequences and in each echo separately. It is also in line with the deconvolved BOLD-responses observed in De Hollander et al. (2017, their Fig. 6). Substantial interindividual differences (Aguirre et al., 1998; D'Esposito et al., 1999; Handwerker et al., 2004) and interregional differences (Boillat and Van der Zwaag, 2019; Handwerker et al., 2004; Miezin et al., 2000) in the HRF have been reported before. It is important to investigate in future research whether the large undershoots are a consequence of the specific neural response evoked by the task used in the current study, or perhaps a signature of a different hemodynamic response than often assumed, for example caused by differences in vasculature. Such studies would require specifically designed experiments, preferably using multiple behavioral paradigms. If a more substantial investigation of subcortical hemodynamics would confirm these observations, the canonical HRF may be adjusted for these areas. Since mismatches between the assumed hemodynamic response function in a GLM and the actual response inflates the noise term of the model, this may improve statistical power. However, because the larger undershoots were present in both protocols of the present study, it is unlikely that they biased the protocol comparison in the present study.

The GLM results with respect to the stop-signal task replicate earlier results using the same paradigm (Aron and Poldrack, 2006; De Hollander et al., 2017; Li et al., 2008). For example, we replicated the typical finding that rIFG (as well as insula) are more active during failed stops compared to go trials and successful stops (Aron et al., 2014). However, we observed an increase in STN activity in failed stop trials compared to successful stop and go trials. This is in line with some (De Hollander et al., 2017; Li et al., 2008), but not all (Aron and Poldrack, 2006) earlier

studies, and it is an open question whether this observation can be reconciled with current theories of the STN's functioning in the stop-signal task (Aron and Poldrack, 2006).

A potential limitation of the current study is that we did not correct for potential cardiac and respiratory artefacts. High-frequency noise such as cardiac and respiratory artefacts are prominent in inferior parts of the brain (e.g., Barry et al., 2013; Todd et al., 2017). However, these effects would only bias the obtained *t*-values if the cardiac and respiratory responses are spuriously correlated with the task design. Furthermore, in an earlier study with a very similar single echo protocol and the same behavioral paradigm, it was shown that removing respiratory and cardiac effects from the signal using RETROICOR (Glover et al., 2000) had only marginal effects on tSNR and task-based *t*-statistics (De Hollander et al., 2017). This is in line with an earlier report showing that only a small part of the signal variance in midbrain regions could be explained by cardiac and respiratory effects (Barry et al., 2013).

In conclusion, we observed at best marginal differences between the BOLD-sensitivity of a protocol optimized for subcortex and a multi echo protocol that both retain whole-brain coverage and have a high spatial resolution. A single echo protocol attuned to subcortical areas (with a reduced TE and high spatial resolution) retained adequate sensitivity in the cortex. This single echo protocol can be optimized further (e.g., increased temporal resolution, with potential benefits from multi-band acceleration), and the resulting data are better suited for established procedures of data processing and analyses. Therefore, when studying small, subcortical nuclei in parallel with cortical areas with current typical 7 T hardware, we recommend to use an optimized single echo sequence unless one has additional reasons (e.g., reduced susceptibility-related signal loss or when one plans to use denoising techniques that rely on multiple contrasts) to collect multiple echoes.

Code & data availability

All analysis code can be found at https://github.com/IMCN-UvA/fMRI_protocol_comparison_se_me. All data will be uploaded to https://osf.io/6qwjz/ upon publication.

Declarations of competing interest

The Max Planck Institute for Human Cognitive and Brain Sciences has an institutional research agreement with Siemens Healthcare. NW was a speaker at an event organized by Siemens Healthcare and was reimbursed for the travel expenses.

CRediT authorship contribution statement

Steven Miletić: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Pierre-Louis Bazin:** Methodology, Investigation, Software, Supervision, Writing - review & editing. **Nikolaus Weiskopf:** Resources, Supervision, Project administration, Funding acquisition, Writing - review & editing. **Wietske van der Zwaag:** Resources, Investigation, Writing - review & editing. **Birte U. Forstmann:** Conceptualization, Methodology, Resources, Supervision, Project administration, Funding acquisition, Writing - review & editing. **Robert Trampel:** Conceptualization, Methodology, Investigation, Resources, Software, Project administration, Writing - review & editing, Supervision.

Acknowledgements

We would like to thank Anke Kummer, Mandy Jochemko, Simone Wipper, and Domenica Wilfling for their help collecting the data. We would like to thank Felix Breuer and Renzo Huber for insightful discussions. This work was supported by an NWO-Vici grant (BUF). NW received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement n° 616905; from the European Union's Horizon 2020 Research and Innovation Programme under the grant agreement No 681094; from the BMBF (01EW1711A & B) in the framework of ERA-NET NEURON.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2020.116992.

References

Aguirre, G.K., Zarahn, E., D'Esposito, M., 1998. The variability of human, BOLD hemodynamic responses. Neuroimage 8, 360–369. https://doi.org/10.1006/nimg.1998.0369.

Aquino, D., Bizzi, A., Grisoli, M., Garavaglia, B., Bruzzone, M.G., Nardocci, N., Savoiardo, M., Chiapparini, L., 2009. Age-related iron deposition in the basal ganglia: quantitative analysis in healthy subjects. Radiology 252, 165–172.

Aron, A.R., Poldrack, R.A., 2006. Cortical and subcortical contributions to stop signal response inhibition: role of the subthalamic nucleus. J. Neurosci. 26, 2424–2433. https://doi.org/10.1523/JNEUROSCI.4682-05.2006.

Aron, A.R., Robbins, T.W., Poldrack, R.A., 2014. Inhibition and the right inferior frontal cortex: one decade on. Trends Cognit. Sci. 18, 177–185. https://doi.org/10.1016/j.tics.2013.12.003.

Avants, B., Epstein, C., Grossman, M., Gee, J., 2008. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med. Image Anal. 12, 26–41. https://doi.org/10.1016/j.media.2007.06.004.

Barry, R.L., Coaster, M., Rogers, B.P., Newton, A.T., Moore, J., Anderson, A.W., Zald, D.H., Gore, J.C., 2013. On the origins of signal variance in FMRI of the human midbrain at high field. PloS One 8, 1–14. https://doi.org/10.1371/journal.pone.0062708.

Boillat, Y., Van der Zwaag, W., 2019. Whole brain measurements of the positive BOLD response variability during a finger tapping task at 7 T show regional differences in its profiles. Magn. Reson. Med. 81, 2720–2727. https://doi.org/10.1002/mrm.27566.

Bullmore, E., Brammer, M., Williams, S.C.R., Rabe-Hesketh, S., Janot, N., David, A., Mellers, J., Howard, R., Sham, P., 1996. Statistical methods of estimation and inference for functional MR image analysis. Magn. Reson. Med. 35, 261–277. https://doi.org/10.1002/mrm.1910350219.

Chamberlain, S.R., Fineberg, N.A., Blackwell, A.D., Robbins, T.W., Sahakian, B.J., 2006. Motor inhibition and cognitive flexibility in obsessive-compulsive disorder and trichotillomania. Am. J. Psychiatr. 163, 1282–1284. https://doi.org/10.1176/appi.ajp.163.7.1282.

Cox, R.W., Hyde, J.S., 1997. Software tools for analysis and visualization of fMRI data. NMR Biomed. 10, 171–178. https://doi.org/10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L.

De Hollander, G., Keuken, M.C., Bazin, P.L., Weiss, M., Neumann, J., Reimann, K., Wähnert, M., Turner, R., Forstmann, B.U., Schäfer, A., 2014. A gradual increase of iron toward the medial-inferior tip of the subthalamic nucleus. Hum. Brain Mapp. 35, 4440–4449. https://doi.org/10.1002/hbm.22485.

De Hollander, G., Keuken, M.C., Forstmann, B.U., 2015. The subcortical cocktail problem; Mixed signals from the subthalamic nucleus and substantia nigra. PloS One 10. https://doi.org/10.1371/journal.pone.0120572.

De Hollander, G., Keuken, M.C., van der Zwaag, W., Forstmann, B.U., Trampel, R., 2017. Comparing functional MRI protocols for small, iron-rich basal ganglia nuclei such as the subthalamic nucleus at 7 T and 3 T. Hum. Brain Mapp. 38, 3226–3248. https://doi.org/10.1002/hbm.23586.

Deistung, A., Schäfer, A., Schweser, F., Biedermann, U., Turner, R., Reichenbach, J.R., 2013. Toward in vivo histology: a comparison of quantitative susceptibility mapping (QSM) with magnitude-, phase-, and R2* -imaging at ultra-high magnetic field strength. Neuroimage 65, 299–314. https://doi.org/10.1016/j.neuroimage.2012.09.055.

Demetriou, L., Kowalczyk, O.S., Tyson, G., Bello, T., Newbould, R.D., Wall, M.B., 2018. A comprehensive evaluation of increasing temporal resolution with multiband-accelerated protocols and effects on statistical outcome measures in fMRI. Neuroimage 176, 404–416. https://doi.org/10.1016/j.neuroimage.2018.05.011.

Desikan, R.S., Se, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31, 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021.

D'Esposito, M., Zarahn, E., Aguirre, G.K., Rypma, B., 1999. The effect of normal aging on the coupling of neural activity to the bold hemodynamic response. Neuroimage 10, 6–14. https://doi.org/10.1006/nimg.1999.0444.

Esteban, O., Markiewicz, C.J., Blair, R.W., Moodie, C.A., Isik, A.I., Erramuzpe, A., Kent, J.D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S.S., Wright, J., Durnez, J., Poldrack, R.A., Gorgolewski, K.J., 2019. fMRIPrep: a robust preprocessing pipeline for functional MRI. Nat. Methods 16, 111–116. https://doi.org/10.1038/s41592-018-0235-4.

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A.R., Fox, P.T., Eickhoff, S.B., Yu, C., Jiang, T., 2016. The human brainnetome atlas: a new brain atlas based on connectional architecture. Cereb. Cortex 3508–3526. https://doi.org/10.1093/cercor/bhw157.

Fonov, V., Evans, A., McKinstry, R., Almli, C., Collins, D., 2009. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. Neuroimage 47, S102. https://doi.org/10.1016/S1053-8119(09)70884-5.

Forstmann, B.U., De Hollander, G., Van Maanen, L., Alkemade, A., Keuken, M.C., 2017. Towards a mechanistic understanding of the human subcortex. Nat. Rev. Neurosci. 18, 57–65.

Friston, K.J., Jezzard, P., Turner, R., 1994. Analysis of functional MRI time-series. Hum. Brain Mapp. 1, 153–171. https://doi.org/10.1002/hbm.460010207.

Friston, K.J., Holmes, a.P., Worsley, K.J., Poline, J.-P., Frith, C.D., Frackowiak, R.S.J., 1995. Statistical parametric maps in functional imaging: a general linear approach. Hum. Brain Mapp. 2, 189–210. https://doi.org/10.1002/hbm.460020402.

Friston, K.J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M.D., Turner, R., 1998. Event-related fMRI: characterizing differential responses. Neuroimage 7, 30–40. https://doi.org/10.1006/nimg.1997.0306.

Gauggel, S., Rieger, M., Feghoff, T.-A., 2004. Inhibition of ongoing responses in patients with Parkinson's disease. J. Neurol. Neurosurg. Psychiatry 75, 539–544. https://doi.org/10.1136/jnnp.2003.016469.

Geissler, A., Lanzenberger, R., Barth, M., Tahamtan, A.R., Milakara, D., Gartus, A., Beisteiner, R., 2005. Influence of fMRI smoothing procedures on replicability of fine scale motor localization. Neuroimage 24, 323–331. https://doi.org/10.1016/j.neuroimage.2004.08.042.

Gelman, A., Hill, J., 2007. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, Cambridge.

Gitelman, D.R., Penny, W.D., Ashburner, J., Friston, K.J., 2003. Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. Neuroimage 19, 200–207. https://doi.org/10.1016/S1053-8119(03)00058-2.

Glover, G.H., 1999. Deconvolution of impulse response in event-related BOLD fMRI. Neuroimage 9, 416–429.

Glover, G.H., Li, T., Ress, D., 2000. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. Magn. Reson. Med. 44, 162–167. https://doi.org/10.1002/1522-2594(200007)44:1<162::AID-MRM23>3.0.CO;2-E.

Gorno-Tempini, M.L., Hutton, C., Josephs, O., Deichmann, R., Price, C., Turner, R., 2002. Echo time dependence of BOLD contrast and susceptibility artifacts. Neuroimage 15, 136–142. https://doi.org/10.1006/nimg.2001.0967.

Gowland, P.A., Bowtell, R., 2007. Theoretical optimization of multi-echo fMRI data acquisition. Phys. Med. Biol. 52, 1801–1813. https://doi.org/10.1088/0031-9155/52/7/003.

Greve, D.N., Fischl, B., 2009. Accurate and robust brain image alignment using boundary-based registration. Neuroimage 48, 63–72. https://doi.org/10.1016/j.neuroimage.2009.06.060.

Handwerker, D.A., Ollinger, J.M., D'Esposito, M., 2004. Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. Neuroimage 21, 1639–1651. https://doi.org/10.1016/j.neuroimage.2003.11.029.

Heidemann, R.M., Ivanov, D., Trampel, R., Fasano, F., Meyer, H., Pfeuffer, J., Turner, R., 2012. Isotropic submillimeter fMRI in the human brain at 7 T: Combining reduced field-of-view imaging and partially parallel acquisitions. Magn. Reson. Med. 68, 1506–1516. https://doi.org/10.1002/mrm.24156.

Huntenburg, J.M., Steele, C.J., Bazin, P.-L., 2018. Nighres: processing tools for high-resolution neuroimaging. GigaScience 7, 1–9. https://doi.org/10.1093/gigascience/giy082.

Hyde, J.S., Biswal, B.B., Jesmanowicz, A., 2001. High-resolution fMRI using multislice partial k-space GR-EPI with cubic voxels. Magn. Reson. Med. 46, 114–125. https://doi.org/10.1002/mrm.1166.

Jeffreys, H., 1961. The Theory of Probability, third ed. Oxford University Press, Oxford.

Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17, 825–841. https://doi.org/10.1006/nimg.2002.1132.

Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M., 2012. FSL. Neuroimage 62, 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015.

Johansen-Berg, H., Behrens, T.E.J., Robson, M.D., Drobnjak, I., Rushworth, M.F.S., Brady, J.M., Smith, S.M., Higham, D.J., Matthews, P.M., 2004. Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. Proc. Natl. Acad. Sci. U.S.A. 101, 13335–13340. https://doi.org/10.1073/pnas.0403743101.

Kettinger, Á., Hill, C., Vidnyánszky, Z., Windischberger, C., 2016. Investigating the group-level impact of advanced dual-echo fMRI combinations. Front. Neurosci. 10, 1–14. https://doi.org/10.3389/fnins.2016.00571.

Keuken, M.C., Bazin, P.L., Crown, L., Hootsmans, J., Laufer, A., Müller-Axt, C., Sier, R., van der Putten, E.J., Schäfer, A., Turner, R., Forstmann, B.U., 2014. Quantifying inter-individual anatomical variability in the subcortex using 7T structural MRI. Neuroimage 94, 40–46. https://doi.org/10.1016/j.neuroimage.2014.03.032.

Keuken, M.C., Bazin, P.L., Backhouse, K., Beekhuizen, S., Himmer, L., Kandola, A., Lafeber, J.J., Prochazkova, L., Trutti, A., Schäfer, A., Turner, R., Forstmann, B.U., 2017. Effects of aging on T1 , T2* , and QSM MRI values in the subcortex. Brain Struct. Funct. 222, 2487–2505. https://doi.org/10.1007/s00429-016-1352-4.

Kirilina, E., Lutti, A., Poser, B.A., Blankenburg, F., Weiskopf, N., 2016. The quest for the best: the impact of different EPI sequences on the sensitivity of random effect fMRI group analyses. Neuroimage 126, 49–59. https://doi.org/10.1016/j.neuroimage.2015.10.071.

Kundu, P., Inati, S.J., Evans, J.W., Luh, W.M., Bandettini, P.A., 2012. Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI. Neuroimage 60, 1759–1770. https://doi.org/10.1016/j.neuroimage.2011.12.028.

Kundu, P., Voon, V., Balchandani, P., Lombardo, M.V., Poser, B.A., Bandettini, P.A., 2017. Multi-echo fMRI: a review of applications in fMRI denoising and analysis of BOLD signals. Neuroimage 154, 59–80. https://doi.org/10.1016/j.neuroimage.2017.03.033.

Larkman, D.J., Nunes, R.G., 2007. Parallel magnetic resonance imaging. Phys. Med. Biol. 52, R15–R55. https://doi.org/10.1088/0031-9155/52/7/R01.

Larkman, D.J., Hajnal, J.V., Herlihy, A.H., Coutts, G.A., Young, I.R., Ehnholm, G., 2001. Use of multicoil arrays for separation of signal from multiple slices simultaneously excited. J. Magn. Reson. Imag. 13, 313–317. https://doi.org/10.1002/1522-2586(200102)13:2<313::AID-JMRI1045>3.3.CO;2-N.

Li, C.R., Yan, P., Sinha, R., Lee, T.W., 2008. Subcortical processes of motor response inhibition during a stop signal task. Neuroimage 41, 1352–1363. https://doi.org/10.1016/j.neuroimage.2008.04.023.

Liang, F., Paulo, R., Molina, G., Clyde, M.A., Berger, J.O., 2008. Mixtures of g priors for Bayesian variable selection. J. Am. Stat. Assoc. 103, 410–423. https://doi.org/10.1198/016214507000001337.

Lijffijt, M., Kenemans, J.L., Verbaten, M.N., Van Engeland, H., 2005. A meta-analytic review of stopping performance in attention-deficit/hyperactivity disorder: deficient inhibitory motor control? J. Abnorm. Psychol. 114, 216–222. https://doi.org/10.1037/0021-843X.114.2.216.

Liu, J., Duffy, B.A., Bernal-Casas, D., Fang, Z., Lee, J.H., 2017. Comparison of fMRI analysis methods for heterogeneous BOLD responses in block design studies. Neuroimage 147, 390–408. https://doi.org/10.1016/j.neuroimage.2016.12.045.

Logan, G.D., Cowan, W.B., Davis, K.A., 1984. On the ability to inhibit simple and choice reaction time responses: a model and a method. J. Exp. Psychol. Hum. Percept. Perform. 10, 276–291. https://doi.org/10.1080/02699930600625081.

Marques, J.P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P.F., Gruetter, R., 2010. MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. Neuroimage 49, 1271–1281. https://doi.org/10.1016/j.neuroimage.2009.10.002.

Menon, R.S., Ogawa, S., Tank, D.W., Uğurbil, K., 1993. 4 Tesla gradient recalled echo characteristics of photic stimulation-induced signal changes in the human primary visual cortex. Magn. Reson. Med. 30, 380–386.

Mestres-Missé, A., Trampel, R., Turner, R., Kotz, S.A., 2017. Uncertainty and expectancy deviations require cortico-subcortical cooperation. Neuroimage 144, 23–34. https://doi.org/10.1016/j.neuroimage.2016.05.069.

Miezin, F.M., Maccotta, L., Ollinger, J.M., Petersen, S.E., Buckner, R.L., 2000. Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. Neuroimage 11, 735–759. https://doi.org/10.1006/nimg.2000.0568.

Moeller, S., Yacoub, E., Olman, C.A., Auerbach, E., Strupp, J., Harel, N., Uğurbil, K., 2010. Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain FMRI. Magn. Reson. Med. 63, 1144–1153. https://doi.org/10.1002/mrm.22361.

Morey, R.D., Rouder, J.N., Jamil, T., Urbanek, S., Forner, K., Ly, A., 2018. BayesFactor.

Mulder, M.J., Keuken, M.C., Bazin, P., Alkemade, A., Forstmann, B.U., 2019. Size and shape matter: the impact of voxel geometry on the identification of small nuclei. PloS One 14, e0215382. https://doi.org/10.1371/journal.pone.0215382.

O'Brien, K.R., Kober, T., Hagmann, P., Maeder, P., Marques, J., Lazeyras, F., Krueger, G., Roche, A., 2014. Robust T1-weighted structural brain imaging and morphometry at 7T using MP2RAGE. PloS One 9. https://doi.org/10.1371/journal.pone.0099676.

Peters, A.M., Brookes, M.J., Hoogenraad, F.G., Gowland, P.A., Francis, S.T., Morris, P.G., Bowtell, R., 2007. T2* measurements in human brain at 1.5, 3 and 7 T. Magn. Reson. Imaging 25, 748–753. https://doi.org/10.1016/j.mri.2007.02.014.

Poldrack, R.A., Mumford, J.A., Nichols, T.E., 2011. Statistical modeling: single subject analysis. In: Handbook of Functional MRI. Cambridge University Press, Cambridge, pp. 70–99.

Poser, B.A., Norris, D.G., 2009. Investigating the benefits of multi-echo EPI for fMRI at 7 T. Neuroimage 45, 1162–1172. https://doi.org/10.1016/j.neuroimage.2009.01.007.

Poser, B.A., Versluis, M.J., Hoogduin, J.M., Norris, D.G., 2006. BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: parallel-acquired inhomogeneity-desensitized fMRI. Magn. Reson. Med. 55, 1227–1235. https://doi.org/10.1002/mrm.20900.

Posse, S., Wiese, S., Gembris, D., Mathiak, K., Kessler, C., Grosse-Ruyken, M.L., Elghahwagi, B., Richards, T., Dager, S.R., Kiselev, V.G., 1999. Enhancement of BOLD-contrast sensitivity by single-shot multi-echo functional MR imaging. Magn. Reson. Med. 42, 87–97.

Preibisch, C., Castrillón, G.J.G., Bührer, M., Riedl, V., 2015. Evaluation of multiband EPI acquisitions for resting state fMRI. PloS One 10, 1–14. https://doi.org/10.1371/journal.pone.0136961.

Puckett, A.M., Bollmann, S., Poser, B.A., Palmer, J., Barth, M., Cunnington, R., 2018. 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 172, 886–895. https://doi.org/10.1016/j.neuroimage.2017.12.005.

Ray, N.J., Brittain, J.S., Holland, P., Joundi, R.A., Stein, J.F., Aziz, T.Z., Jenkinson, N., 2012. The role of the subthalamic nucleus in response inhibition: evidence from local field potential recordings in the human subthalamic nucleus. Neuroimage 60, 271–278. https://doi.org/10.1016/j.neuroimage.2011.12.035.

Rouder, J.N., Morey, R.D., 2012. Default Bayes factors for model selection in regression. Multivariate Behav. Res. 47, 877–903. https://doi.org/10.1080/00273171.2012.734737.

Salomon, R., Darulova, J., Narsude, M., Van Der Zwaag, W., 2014. Comparison of an 8-hannel and a 32-channel coil for high-resolution fMRI at 7 T. Brain Topogr. 27, 209–212. https://doi.org/10.1007/s10548-013-0298-6.

Satterthwaite, F.E., 1941. Synthesis of variance. Psychometrika 6, 309–316. https://doi.org/10.1007/BF02288586.

Seabold, S., Perktold, J., 2010. Statsmodels: econometric and statistical modeling with Python. In: 9th Python Sci. Conf., pp. 57–61.

Sebastian, A., Forstmann, B.U., Matzke, D., 2018. Towards a model-based cognitive neuroscience of stopping – a neuroimaging perspective. Neurosci. Biobehav. Rev. 90, 130–136. https://doi.org/10.1016/j.neubiorev.2018.04.011.

Sengupta, S., Roebroeck, A., Kemper, V.G., Poser, B.A., Zimmermann, J., Goebel, R., Adriany, G., 2016. A specialized multi-transmit head coil for high resolution fMRI of the human visual cortex at 7T. PloS One 11. https://doi.org/10.1371/journal.pone.0165418.

Setsompop, K., Gagoski, B.A., Polimeni, J.R., Witzel, T., Wedeen, V.J., Wald, L.L., 2012. Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. Magn. Reson. Med. 67, 1210–1224. https://doi.org/10.1002/mrm.23097.

Speck, O., Hennig, J., 1998. Functional imaging by I0 and T2* parameter mapping using multi-image EPI. Magn. Reson. Med. 40, 243–248.

Speck, O., Hennig, J., 2001. Motion correction of parametric fMRI data from multi-slice single-shot multi-echo acquisitions. Magn. Reson. Med. 46, 1023–1027. https://doi.org/10.1002/mrm.1291.

Stelzer, J., Lohmann, G., Mueller, K., Buschmann, T., Turner, R., 2014. Deficient approaches to human neuroimaging. Front. Hum. Neurosci. 8, 1–16. https://doi.org/10.3389/fnhum.2014.00462.

Stöcker, T., Kellermann, T., Schneider, F., Habel, U., Amunts, K., Pieperhoff, P., Zilles, K., Shah, N.J., 2006. Dependence of amygdala activation on echo time: results from olfactory fMRI experiments. Neuroimage 30, 151–159. https://doi.org/10.1016/j.neuroimage.2005.09.050.

Todd, N., Moeller, S., Auerbach, E.J., Yacoub, E., Flandin, G., Weiskopf, N., 2016. Evaluation of 2D multiband EPI imaging for high-resolution, whole-brain, task-based fMRI studies at 3T: sensitivity and slice leakage artifacts. Neuroimage 124, 32–42. https://doi.org/10.1016/j.neuroimage.2015.08.056.

Todd, N., Josephs, O., Zeidman, P., Flandin, G., Moeller, S., Weiskopf, N., 2017. Functional sensitivity of 2D simultaneous multi-slice echo-planar imaging: effects of acceleration on g-factor and physiological noise. Front. Neurosci. 11, 1–14. https://doi.org/10.3389/fnins.2017.00158.

Uğurbil, K., Xu, J., Auerbach, E.J., Moeller, S., Vu, A.T., Duarte-Carvajalino, J.M., Lenglet, C., Wu, X., Schmitter, S., Van de Moortele, P.F., Strupp, J., Sapiro, G., De Martino, F., Wang, D., Harel, N., Garwood, M., Chen, L., Feinberg, D.A., Smith, S.M., Miller, K.L., Sotiropoulos, S.N., Jbabdi, S., Andersson, J.L.R., Behrens, T.E.J., Glasser, M.F., Van Essen, D.C., Yacoub, E., 2013. Pushing spatial and temporal resolution for functional and diffusion MRI in the Human Connectome Project. Neuroimage 80, 80–104. https://doi.org/10.1016/j.neuroimage.2013.05.012.

Van De Moortele, P.-F., Ugurbil, K., Lehericy, S., 2008. Is T2* always the optimum echo time in BOLD FMRI? Challenging a common concept with a new contrast to noise ratio BOLD model. Proc. Int. Soc. Magn. Reson. Med. 6, 70800.

Van der Zwaag, W., Schäfer, A., Marques, J.P., Turner, R., Trampel, R., 2016. Recent applications of UHF-MRI in the study of human brain function and structure: a review. NMR Biomed. 29 (9), 1274–1288. https://doi.org/10.1002/nbm.3275.

Verbruggen, F., Aron, A.R., Band, G.P., Beste, C., Bissett, P.G., Brockett, A.T., Brown, J.W., Chamberlain, S.R., Chambers, C.D., Colonius, H., Colzato, L.S., Corneil, B.D., Coxon, J.P., Dupuis, A., Eagle, D.M., Garavan, H., Greenhouse, I., Heathcote, A., Huster, R.J., Jahfari, S., Kenemans, J.L., Leunissen, I., Logan, G.D., Matzke, D., Morein-Zamir, S., Murthy, A., Li, C.-S.R., Paré, M., Poldrack, R.A., Ridderinkhof, K.R., Robbins, T.W., Roesch, M., Rubia, K., Schachar, R.J., Schall, J.D., Stock, A.-K., Swann, N.C., Thakkar, K.N., van der Molen, M.W., Vermeylen, L., Vink, M., Wessel, J.R., Whelan, R., Zandbelt, B.B., Boehler, C.N., 2019. A consensus guide to capturing the ability to inhibit actions and impulsive behaviors in the stop-signal task. Elife 8, 1–25. https://doi.org/10.7554/eLife.46323.

Volz, S., Callaghan, M.F., Josephs, O., Weiskopf, N., 2019. Maximising BOLD sensitivity through automated EPI protocol optimisation. Neuroimage 189, 159–170. https://doi.org/10.1016/j.neuroimage.2018.12.052.

Weiskopf, N., Klose, U., Birbaumer, N., Mathiak, K., 2005. Single-shot compensation of image distortions and BOLD contrast optimization using multi-echo EPI for real-time fMRI. Neuroimage 24, 1068–1079. https://doi.org/10.1016/j.neuroimage.2004.10.012.

Wetzels, R., Wagenmakers, E.J., 2012. A default Bayesian hypothesis test for correlations and partial correlations. Psychon. Bull. Rev. 19, 1057–1064. https://doi.org/10.3758/s13423-012-0295-x.

White, C.N., Congdon, E., Mumford, J.A., Karlsgodt, K.H., Sabb, F.W., Freimer, N.B., London, E.D., Cannon, T.D., Bilder, R.M., Poldrack, R.A., 2014. Decomposing decision components in the stop-signal task: a model-based approach to individual differences in inhibitory control. J. Cognit. Neurosci. 26, 1601–1614. https://doi.org/10.1162/jocn_a_00567.

Wiecki, T.V., Frank, M.J., 2012. A Computational Model of Inhibitory Control in Frontal Cortex and Basal Ganglia 1–66.

Woolrich, M.W., Ripley, B.D., Brady, M., Smith, S.M., 2001. Temporal autocorrelation in univariate linear modeling of FMRI data. Neuroimage 14, 1370–1386. https://doi.org/10.1006/nimg.2001.0931.

Woolrich, M.W., Behrens, T.E.J., Beckmann, C.F., Jenkinson, M., Smith, S.M., 2004. Multilevel linear modelling for FMRI group analysis using Bayesian inference. Neuroimage 21, 1732–1747. https://doi.org/10.1016/j.neuroimage.2003.12.023.

Xu, J., Moeller, S., Auerbach, E.J., Strupp, J., Smith, S.M., Feinberg, D.A., Yacoub, E., Uğurbil, K., 2013. Evaluation of slice accelerations using multiband echo planar imaging at 3T. Neuroimage 83, 991–1001. https://doi.org/10.1016/j.neuroimage.2013.07.055.

Yekutieli, D., Benjamini, Y., 1999. Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics. J. Stat. Plann. Inference 82, 171–196. https://doi.org/10.1016/s0378-3758(99)00041-5.

Zwirner, J., Möbius, D., Bechmann, I., Arendt, T., Hoffmann, K.T., Jäger, C., Lobsien, D., Möbius, R., Planitzer, U., Winkler, D., Morawski, M., Hammer, N., 2017. Subthalamic nucleus volumes are highly consistent but decrease age-dependently—a combined magnetic resonance imaging and stereology approach in humans. Hum. Brain Mapp. 38, 909–922. https://doi.org/10.1002/hbm.23427.