WU-Minn HCP MEG Initial Data Release: Reference Manual

4 March 2014

Table of Contents

Introduction and overview

This document provides information and guidance on how to use magnetoencephalography (MEG) data released by the WU-Minn HCP consortium in the MEG Initial Data Release (MEG1, March 2014). Given the richness of the HCP datasets and their utility for a wide range of research purposes, it is important that potential users understand what data are in the current release, how the datasets are organized, how they can be accessed and what's in store for future releases.

The HCP aims to study 400 subjects per year, with a target of 1,200 subjects studied before the grant ends in 2015. Participants from families with twins and non-twin siblings are being scanned on the same equipment using the same protocol for every subject. A subset of 50 same-sex monozygotic twin pairs (100 participants) are targeted to be additionally studied using both resting-state and task MEG. These 100 subjects also participate in the rest of the HCP study design including structural and functional magnetic resonance image studies (MRI, fMRI), diffusion tensor imaging (dMRI) and behavioral and genetic testing (see Van Essen *et al.*, 2013). Significant effort has been made to ensure that MEG and fMRI are performed on the same subjects using the same tasks performed, as closely as possible, with the same timing (see Larson-Prior *et al.*, 2013). Together with behavioral and genetic information on these subjects, these data will facilitate the study of individual variability across a large population.

This is the first HCP release of MEG data, including resting-state and task-based modalities. All future HCP releases will include MEG data.

The MEG Initial (MEG1) release includes data from 14 MEG-scanned participants, which together comprise 10 complete datasets of resting state (rMEG) and 10 of each of the three task protocols (tMEG: motor, working memory, and language processing). To provide cross-modal information, the MEG task paradigms are a subset of the 7 HCP tfMRI paradigms collected on all subjects (Barch *et al.*, 2013). The MEG paradigms are matched in temporal structure to fMRI tasks where plausible to maximize the scientific value of the multi-modal design of the HCP (Van Essen *et al.*, 2012) (see tMEG scripts and protocol details below).

Concurrent with the MEG1 data release, we are also releasing our first version of the HCP MEG processing pipelines. The pipelines use freely available software from the *FieldTrip* analysis toolbox and custom scripts that are collectively referred to as "megconnectome". Both megconnectome software and the FieldTrip toolbox are available for download on the HCP website. See MEG Channel-level processing pipelines below for descriptions of analysis pipeline scripts included in megconnectome v1.0 release.

MRI and behavioral/individual difference data have also been collected on the MEG1 release participants and is either currently released (3 subjects) or will be released (11 subjects) in the

upcoming 500 Subjects Release scheduled for late spring 2014. On the MEG Download page described in How to download MEG1 data below, links to the MRI and behavioral data is provided for the 3 subjects for whom this data is currently available. Detailed descriptions of the HCP MRI modalities, behavioral/individual difference measures, and instructions on exploring and downloading these types of data can be found in the HCP Q3 Reference Manual.

Since all HCP MEG data are collected from monozygotic twin pair subjects, it will be important for many analyses to know the family status of the subjects under study. Because of this study design and the limited amount of MEG data collected to date, *several of the subjects included in the MEG1 release are monozygotic (MZ) twin pairs*. Thus, for all tasks and resting state data involving multiple subjects, family structure may confound independent analysis.

To protect the privacy of HCP subjects, information on family structure is restricted. The Open Access MEG1 data do not distinguish related subjects from those that are unrelated. Users who would like to incorporate family structure in their analysis may apply for restricted data access (see below for more information).

If you are interested in accessing the HCP MEG1 data, please continue reading the following important information.

What's in the HCP MEG Initial data release? The MEG1 release includes high quality MEG scans from 14 healthy adults (all members of monozygotic twin pairs) collected at rest (rMEG) and during 3 tasks (tMEG) that measure sensory, motor, and cognitive task performance. As mentioned above, to achieve 10 full datasets each of rMEG and each tMEG task, data from more than 10 subjects are included. This includes 3 subjects with MEG data, plus released data for the rest of the HCP protocol (T1w and T2w MRI, rfMRI, tfMRI, dMRI, and behavioral measures). MRI and behavioral measures for the remaining 11 subjects will be released as part of the upcoming HCP 500 Subjects Release scheduled for late spring 2014.

The release includes:

- raw, unprocessed MEG data in 4D Neuroimaging format

- co-registration information (in the form of transformation matrices) that allow coordinate transformations between individual subject MEG coordinate systems and the MNI coordinate system

- volume conduction model of the head (in MATLAB format)

- regular 3-D source models (in MATLAB format)

- E-Prime log files (tab-delimited and as Excel spreadsheets)

- lists of bad channels and bad segments

- sets of independent components (each comprising the time course and the sensor map) together with the related classification output

- cleaned channel-level processed data aligned to stimulus and response onsets (for the task data) or segmented in short pieces (2 sec epochs for the resting state data)

- averaged event-related fields and time-frequency estimates of power (for the task data)

- averaged power spectra (for the resting state data)

Also available are the scripts HCP used to process the MEG1 release data, collectively called "megconnectome software version 1.0". See http://humanconnectome.org/documentation/MEG1/meg-pipeline.html for more information.

Levels of data processing. The MEG1 data includes the outputs of a number of data processing pipelines. The utility pipelines include quality control and artifact identification to remove "bad" channels, segments, and physiological artifacts, and an anatomy/source reconstruction pipeline to coregister the MEG to structural MRI (sMRI) and dMRI. The resting-state and task preprocessing pipelines result in sets of components obtained from ICA together with the related classification into artifacts and non-artifacts, cleaned and segmented data with channel-level Event-Related fields (ERFs), Time-Frequency responses (TFRs) and power-spectra.

The methods used for this processing, (see MEG Channel-level processing pipelines below), were implemented by the HCP consortium through an intensive three-year refinement and optimization process. However, refinement of processing pipelines is ongoing, therefore:

Remember to be sure to keep track of version numbers when analyzing HCP data. Version numbers of the relevant HCP pipelines used to process the data in each HCP dataset are available in the release notes directory packaged in each dataset archive. For example, the archive 177746_MEG_Motort_preproc.zip contains the file release-notes/MEG.txt:

177746_MEG_Motort_preproc.zip

Tue Feb 25 07:11:49 CST 2014

The following data processing pipelines were executed with megconnectome v1.0
-datacheck
-anatomy
-baddata
-icaclass
-icaclass_qc
-rmegpreproc
-powavg
-tmegpreproc

```
                    -tfavg
                    -eravg
```

How are HCP MEG1 data accessed? The released MEG1 data are available for download from the [ConnectomeDB](#) splash page (requires registration, login and signing of the HCP [Open Access Data Use Terms](#)). The total amount of data for the MEG1 release is about 280 gigabytes (about 28 GB per subject) and is available in packages of each dataset type per individual subject or as a single multidataset archive per individual. See [MEG Initial Release Data Sizes](#) below.

Users interested in accessing the HCP MRI data (~7 terabytes total data available) can do so by downloading packages for groups of subjects from the ConnectomeDB or via "[Connectome in a Box](#)" (see [HCP Q3 Reference Manual](#) pp.16-23.) Beginning with the upcoming HCP 500 Subjects Release (scheduled for late spring 2014), MEG1 data will be included in Connectome in a Box distributions.

How are MEG1 data analyzed and visualized?

MEG data from the HCP can be analyzed and visualized using existing MEG analysis platforms such as *FieldTrip, MNE, SPM* and *Brainstorm*.

Given the extensive MRI-based data also available for all HCP subjects, a wide range of multimodal analyses involving combined MEG and MRI will be of interest to explore. Connectome Workbench (http://www.humanconnectome.org/connectome/connectome-workbench.html) currently has many features for multimodal visualization of the MRI-based modalities available in HCP data, including functional and structural connectivity as well as task-fMRI. During the coming year, we plan to incorporate additional features that bring extensively analyzed MEG data into the Workbench multimodal visualization framework.

What's in store for future releases? Future releases will not only add to the total number of subjects studied, but will include the more extensively analyzed datasets. We plan to produce source reconstructions for the MEG data registered to the individual's cortical sheet that may be overlaid with data from MRI in a future release; at the time of our first release the registration is more coarse and therefore direct links are not recommended across modalities.

The next multimodality HCP data release is scheduled for late spring 2014 and will include data from 500 subjects scanned in the first 1.5 years of the HCP production phase. This change from the quarterly release cycle (that was followed up to the Q3 release) is enabling the generation and release of more extensively processed data, processing pipelines, and new data mining features on all released subjects, in addition to more datasets on newly scanned subjects.

How were the HCP datasets acquired and processed? Besides the documentation provided below (see Summary of MEG imaging protocols, Data in the Initial MEG release, and MEG Channel-level processing pipelines) there is a published description of various aspects of HCP MEG data collection and processing (Larson-Prior *et al.*, 2013). Descriptions of other HCP methods for behavioral/individual difference tests and MRI modalities are available in the HCP Q3 Reference Manual (and its appendices) and in seven papers recently published in NeuroImage (Van Essen *et al.*, 2013, Ugurbil *et al.*, 2013, Glasser *et al.,* 2013, Smith *et al.*, 2013, Barch *et al.*, 2013, Sotiropoulos *et al.*, 2013, Marcus *et al.*, 2013. See http://www.humanconnectome.org/documentation/citations.html for updated lists of publications.

Can I adapt HCP MEG protocols to my own scanner and research projects? Appendix I: MEG Session Protocol Details provides advice and suggestions on this issue, which depends on the particular scanner system you are using as well as your research objectives.

Open access and restricted data - IMPORTANT!

The HCP provides imaging, behavioral, and demographic data from a large population of healthy adults. This poses special challenges for protecting the privacy of participants, especially because it is a family study including twins and their siblings. Unless these data are properly managed, there is a risk that some participants might be recognizable to family members and others (restricted data). In addition, some of the data elements (sensitive data) collected might harm or embarrass participants if they were to be inadvertently identified.

To protect the privacy of our participants, the HCP has implemented a two-tiered plan for data sharing, with different provisions for handling Open Access data and Restricted Data.

Open access data (all imaging data and most of the behavioral data) are available to those who register and agree to the Open Access Data Use Terms. This includes agreement to comply with institutional rules and regulations (see below), **which may require you to get your IRB's approval to use the data.**

Restricted data elements that could be potentially used to identify subjects include family structure (twin or non-twin status and number of siblings); birth order; age by year; handedness; ethnicity and race; body height, weight, and BMI; and a number of other categories (see HCP Q3 Reference Manual). Restricted data also includes sensitive data (also called tier 2 restricted data) that could harm or embarrass participants if they were to be inadvertently identified, such as psychiatric information from the SSAGA telephone diagnostic interview, drug test results, and physiological data. We have built the distinction between the restricted and sensitive classes of restricted data elements into the user interface of the ConnectomeDB dashboard used to explore and filter on the behavioral data (see HCP Q3 Reference Manual).

Each qualified investigator wanting to use Restricted Data must agree to the [Restricted Data Use Terms](#). These terms explain how Restricted Data may and may not be used and shared, and they reiterate the need for compliance with institutional requirements. **They include major limitations on how Restricted Data can be incorporated into publications and public presentations**. Click [here](#) for additional information about required investigator qualifications and exemplar use case scenarios.

You must comply with your institutional rules and regulations regarding research on human subjects. Different institutions operate under different national and state laws. Before you use any HCP data (either *Open Access or Restricted Data)* for research, you should ask your institution's Institutional Review Board (IRB), Ethics Committee, or office that oversees research on human subjects to determine whether you need institutional approval in order to use the data. On request, HCP can provide a certificate stating that you have accepted the Data Use Terms (Open Access or Restricted Access, as appropriate).

What about technical support, bug reports, and feature requests?

We anticipate a wide range of questions, suggestions, and discussion points as HCP data and software become freely available to the community. Users are strongly encouraged to join the HCP Users mailing list (hcp-users@humanconnectome.org) by signing up at http://www.humanconnectome.org/contact/ or by checking the appropriate box when registering to download HCP data.

Contributions to the hcp-users mailing list will be monitored and responded to by investigators and staff in the WU-Minn HCP consortium. Often this will entail prompt responses to answer questions or suggest solutions to technical problems. As with mailing lists for other brain mapping platforms (e.g., FieldTrip, MNE, SPM), investigators outside the HCP consortium are encouraged to respond as well. Bug reports and feature requests will be entered into the issue tracking system used by HCP software developers by trained HCP staff.

If you are not currently a member of the hcp-users mailing list, you can submit a feature request, question, or suggestion directly at http://humanconnectome.org/contact/feature-request.php.

HCP Data Release Updates Wiki

As mentioned above, a [wiki page](#) has been created to document Known Bugs/Missing files and Planned fixes in the HCP data released to date. This wiki will be updated frequently as we discover new bugs and files missing in the datasets. We recommend checking the wiki page before you do any major analyses with HCP data so that you are aware of any issues with files you may want to use. On the wiki, we also describe our plans to fix bugs/redistribute missing files and provide links to patches the HCP releases periodically to replace the files affected.

MEG Initial Release Data Sizes

The HCP MEG data are notably high in temporal resolution. Coupled with the long scan duration for each modality, the resultant data files are very large. This "big data" is generally a good thing, enabling researchers to conduct types of data analyses that have never before been possible. However, the sheer size of the data generated by the HCP protocol requires some special considerations:

Prepare for long download times. The HCP uses an optimized download client (Aspera) to make downloading as fast as possible, even when multiple users are downloading at the same time. However, depending on the amount of data you are downloading and your internet bandwidth, you may experience lengthy download times.

Housing the data will require a large amount of disk space. Downloading data on even a few subjects (total data/subject: ~25 GB) could fill up your hard drive quickly! Consider just how much data you need to "try out" the data, or do the analysis you want to do, and download just what you need.

MEG Initial Release Data Sizes (per Subject)		
Dataset	**Format**	**.zip File Size**
Anatomy Models	Preprocessed	4 MB
Resting State	Unprocessed	2 GB
	Preprocessed	2 GB
Task-- Working memory	Unprocessed	3 GB
	Preprocessed	5 GB
Task-- Language Processing (Story-Math)	Unprocessed	2 GB
	Preprocessed	5 GB
Task-- Motor	Unprocessed	3 GB
	Preprocessed	6 GB
Noise	Unprocessed	800 MB
Total (per Subject)	Unprocessed	~11 GB
	Preprocessed	~18 GB
	Both	~29 GB
Total (All available MEG1 datasets)	Unprocessed	~110 GB
	Preprocessed	~180 GB
	Both	~290 GB

Consider organizing investigators at your institution to obtain a single copy of the HCP data for local distribution. In many cases, several groups at a research institution will want access to HCP data. We encourage users to organize efforts to obtain a local archive of the data via a single download that can be copied and used by many within your institution. With a modest amount of planning, this faster distribution may save valuable research time.

Important cautionary note about distributing HCP data: If you distribute HCP data in any way, <u>even within a single laboratory or institution</u>, remember that **all data users**, even those who access HCP data through a local archive and not through ConnectomeDB, must [register](#) and sign the [Open Access Data Use Terms](#) before HCP data are used for research purposes. This step is required to be in compliance with regulations on human subjects research (other approval procedures may be required by your institution – see [above](#)). Those who distribute HCP data or use HCP data without signing the Data Use Terms are at risk for violation of their institution's human research policies and possible legal repercussions.

How to access HCP MEG data

Go to the HCP data home page:
http://humanconnectome.org/data/

Here, shown at right, you can get an overview of the MEG Initial Data Release and other current data releases and find links to resources for using HCP data.

In the panel at the right, find the link to ConnectomeDB, the HCP database platform where you will be able to explore and download MEG1 data.

Click on the ConnectomeDB icon. This takes you to the portal (bottom right) where you need to create a ConnectomeDB account.
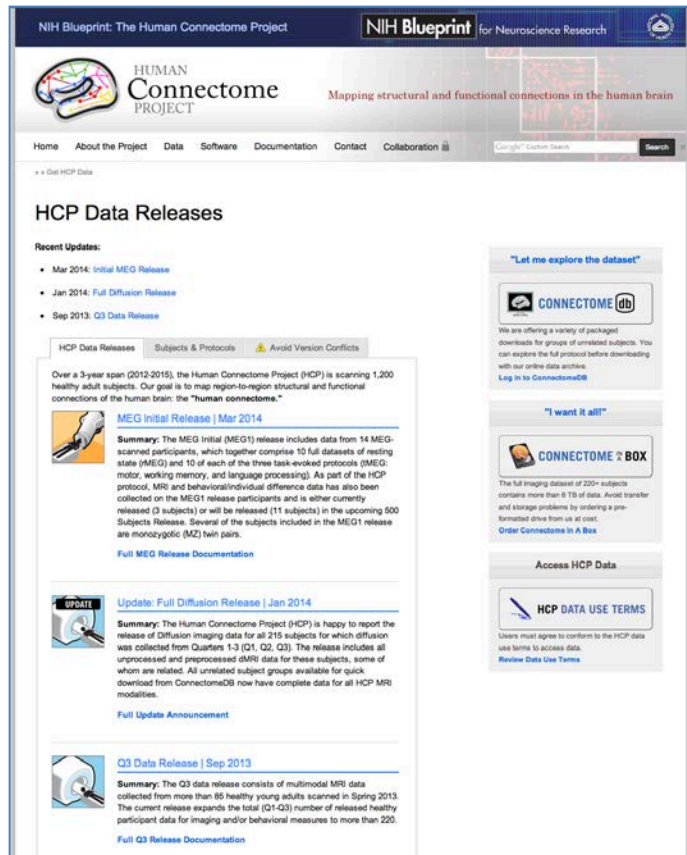
After you have registered, you will be sent a confirmation email. Click on the link in the email to complete your registration.

Go back to the ConnectomeDB portal at https://db.humanconnectome.org and log in.
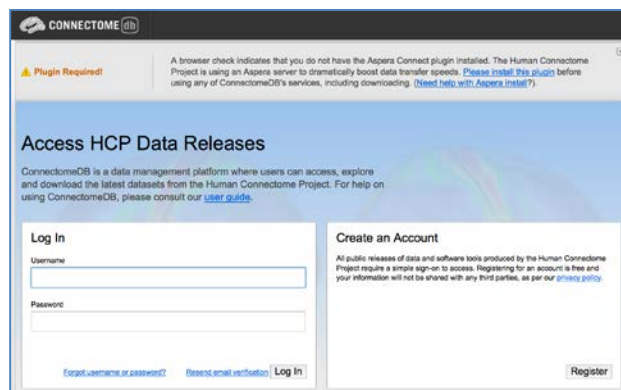
When you log in to ConnectomeDB for the first time, you must agree to the HCP Open Access Data Use Terms (required for access to data).

Once these steps are complete, you will launch the "HCP Data Releases" page

(a.k.a. ConnectomeDB splash page, shown on next page) where you can begin to explore and download HCP data (see How to download MEG1 data below).

Note: Your login to ConnectomeDB times out every 30 minutes. Click the "renew" link to reset the time. (Auto-logout resets ConnectomeDB to the HCP Data Release page).

How to download MEG1 data

Due to the large file sizes of the data, (see HCP Data Sizes above), the HCP uses an Aspera server to dramatically boost data transfer speeds. Therefore, downloading data from the HCP requires you to have the Aspera plugin installed on your browser.
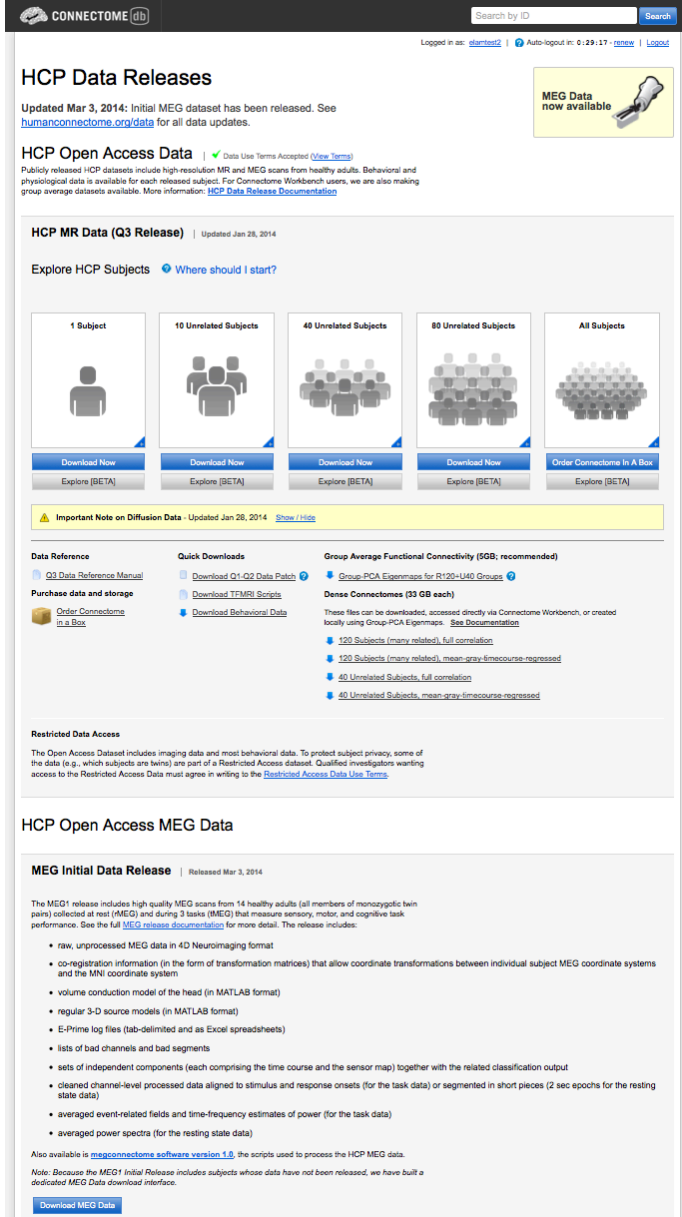
If you have not installed the plugin yet, a warning message will appear at the top of the browser window reminding you that you must Install the Aspera plugin before using any of ConnectomeDB's services, including downloading (see the second panel in How to access HCP data). Note: Installing the Aspera plugin to your browser requires you to restart your browser, therefore, we recommend installing the plugin *before* you explore HCP data in ConnectomeDB.

An Aspera install help guide is available on the HCP website. Note: One common problem with installing Aspera on the Mac platform for non-US users is that you must add US English to the Languages list (using the checkbox) to your System Preferences>Personal>Language & Text settings.

Upon login to ConnectomeDB, you are brought to the HCP Data Releases splash page (top right). About halfway down this page is information and a link to the download page for the MEG Initial Release.

In order to facilitate access, we have prepackaged MEG1 data into packages by data type and by individual subject for download from ConnectomeDB.

If you click the "Download MEG Data" link, you will launch the Download Initial MEG Data Release page designed for

downloading the MEG data as zipped packages (below). Here you can choose exactly which subjects and data packages you want to queue for download.

Beside the Subject ID for 3 of the listed subjects is an "other data" link. These links take you to the subject specific page in ConnectomeDB for the behavioral/individual difference and MRI data collected on that subject. More information about the options for downloading HCP MRI and behavioral data can be found in the HCP Q3 Reference Manual.
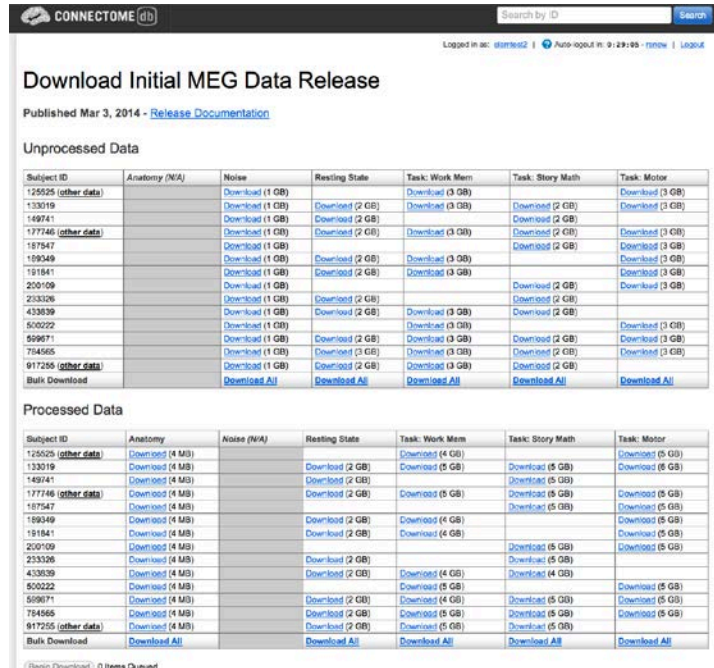
First time Aspera Download Setup

If this is your first time downloading data from ConnectomeDB, **before you start selecting packages to download**, we highly recommend setting up the preferences in your Aspera Connect plugin to download to the appropriate place on your local or network hard drive.

Note: the default Aspera Connect download location is your Desktop, which is likely not where you want your HCP data to go!
(Unfortunately HCP has no control over the Aspera defaults).

To set your Aspera Connect preferences:

- Search your computer for Aspera Connect using the Spotlight or Search functions in your operating system. Click on "Aspera Connect" to open the application.

- Open the Aspera Connect: Transfers window:

  o On Mac, in the Menu bar click on Window > Transfers, or use the shortcut **^T** (control + T)

  o On PC, in the Notification area at the very bottom right of your Desktop, click the ▲ button, then click the Aspera Connect icon:

o On Linux, extract the .gzip Aspera Connect plugin file you downloaded. Run the extracted script in a terminal window to launch Aspera Connect. The Aspera Connect icon should appear in your program notification area (often at the top or bottom right of your desktop), click on this icon and choose Transfers in the dropdown.

- Click the gear ⚙ icon at the bottom left of the Aspera Connect — Transfers window to launch the Aspera preferences window.

- Click on the "Transfers" tab to setup where in your hard drive or network you would like your HCP data to download.
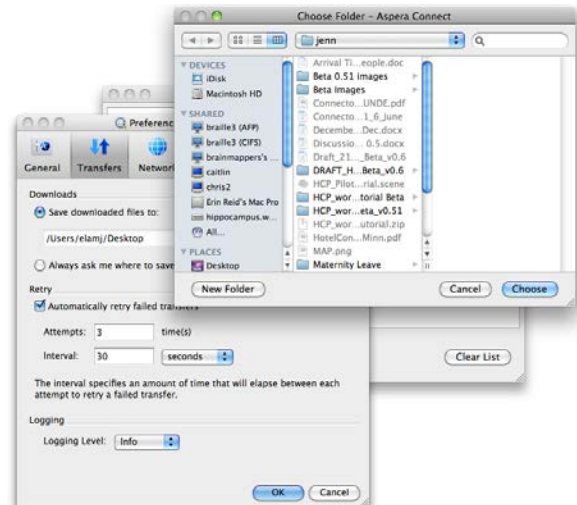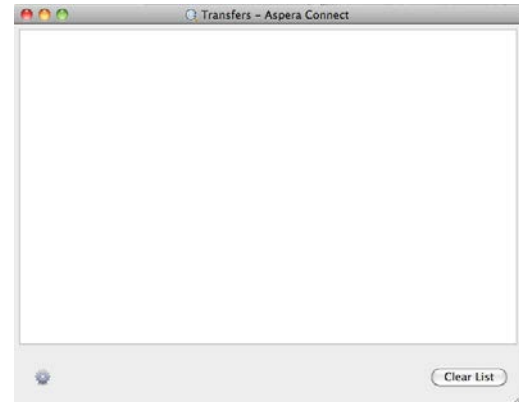
Downloading MEG1 data packages

Now that you have Aspera setup to download to the correct location, go back to the "Download MEG Packages" page in your browser.

Here, you can choose to download unprocessed and processed packages separately for each subject and for each type of data: Anatomy, Noise, Resting State, Working Memory, Language (Story-Math), and Motor tasks. Using the "Download All" links at the bottom of the columns for each data type, you also may download data of that type for all subjects with that type of data available.

Click the Begin Download button to launch the download. A dialog box will appear asking you to confirm the connection between your computer and the HCP's Aspera server. Click Allow.

If you selected more than one file for download, this dialog box may pop up and ask you to confirm the connection for each file you selected.

The download(s) should start automatically in the Aspera Connect Transfers window.

Note: If you have not setup a new download directory in Aspera Connect, as described [above](#), by default the data will start downloading to your Desktop. If this not where you want the data to go, we recommend stopping the download by clicking the "X" button to the right of the download progress bar. Click the gear icon at the bottom left of the Aspera Connect: Transfers window to launch the Aspera preferences window (see above). Once you set the preferred download location, you will need to relaunch your download in ConnectomeDB, using your browser's back button to return to the "Download Initial MEG Data Release" page.

If the download does not start automatically, use the link on the Download Package Launcher page in your browser to restart.

Once downloading has started, the green bar in the Aspera Connect: Transfers window will show your progress (per subject file, not your overall progress).

Obviously, the more subjects and packages you selected for download, the longer your download will take (remember, this is big data!).
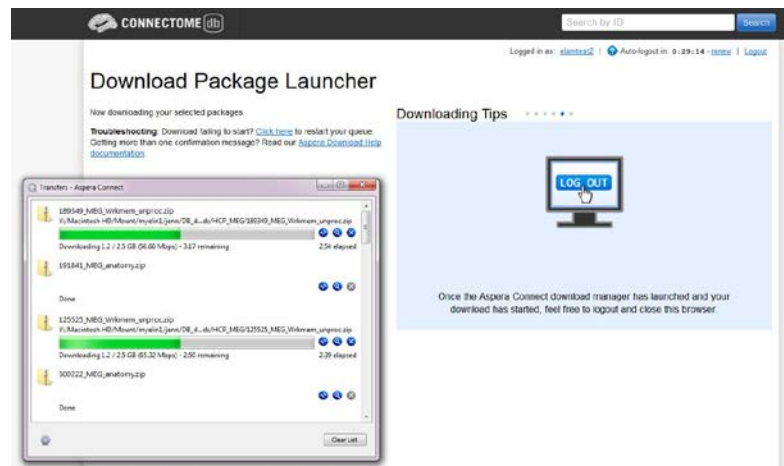
Handling downloaded HCP MEG data

Now that you have downloaded the data you want, here's some tips for handling the data:

- Each download package consists of a .zip archive, and an md5 checksum. After you download the data you want, you can use the md5 file to verify the integrity of your downloaded file.

- To unzip the downloaded .zip archives, for Windows users we recommend using a utility such as 7-zip (available free) so that the directory structure of the unzipped files remains intact. The built-in unzip capabilities of Linux and MacOSX can be used, as they do not affect the directory structure of the unzipped files.

- Downloaded MEG1 data unzip to a set directory structure that is detailed below in Directory structure for unprocessed MEG data, Directory structure for anatomical models for MEG source estimation, Directory structure for Channel-level processed MEG data, and in Appendix III: File Names and Directory Structure for Unprocessed, Anatomical Model, and Channel-Level Processed MEG Data.

- If you have downloaded a large amount of HCP data, we recommend sharing the data locally at your research institution. This will save download and organization time for your colleagues and relieve some of the burden on the ConnectomeDB download bandwidth available to other HCP users. Again, if you distribute HCP data in any way, **all users** (those who access HCP data) must register and sign the Open Access Data Use Terms in order to comply with regulations on human subjects research (see above).

- We are compiling a set of best practices for long-term data storage, which can be seen in our online documentation.

MEG hardware specifications and sensor locations

Scanner hardware and parameters

All HCP subjects being scanned by MEG are scanned on a whole head MAGNES 3600 (4D Neuroimaging, San Diego, CA) system housed in a magnetically shielded room, located at the Saint Louis University (SLU) medical campus.

The MEG system includes 248 magnetometer channels together with 23 reference channels (18 magnetometer and 5 first-order gradiometer channels). The root mean squared (RMS) noise of the magnetometers is ~5 fT/sqrt (Hz) on average in the white noise range (above 2 Hz). Data was recorded at 2034.5101 Hz sampling rate. The electrooculography (EOG, 2 channels), electrocardiography (ECG, 1 channel), and electromyography (EMG, 4 channels) recordings are synchronized with the MEG. All electrode impedances are maintained below 10 kOhms. Other parameter settings that are used for all HCP MEG recordings are: Bandwidth of 400Hz, High pass filter: DC, Acquisition mode: continuous, Data encoding: Delta.

A Polhemus FASTRAK-III is used for spatial digitization of three anatomical landmarks (nasion and L/R periauricular points), five localizer coils, and subjects' head shape (about 2400 points).

A Sun Blade running 64 bit Solaris 8 Unix and communicating with a dedicated UltraAX-12 system is used for data collection and sensor configuration. Raw data are stored in the native 4D file format. Following data acquisition, the data are uploaded via https to the Washington University in St. Louis (WashU) internal HCP database for quality control and processing steps.

Sensor locations

HCP MEG participants are positioned in the MEG scanner supine, with the crown of the head touching but not pressing into the back of the MEG dewar. This position sets the eyebrows inside the dewar, leaving the rest of the face free of the scanner, as in the picture below.

To co-register MEG data to the MRI structural scans, a subject's reference system and locator coils are needed. The reference system is defined through 3 points, the nasion (gray dot in diagram below) and the peri-auricular points. These are placed where the ear folds to the scalp (red oval in photo below), and they are touched with a stylus perpendicularly to the participant's scalp a few times, to record their position on the head.
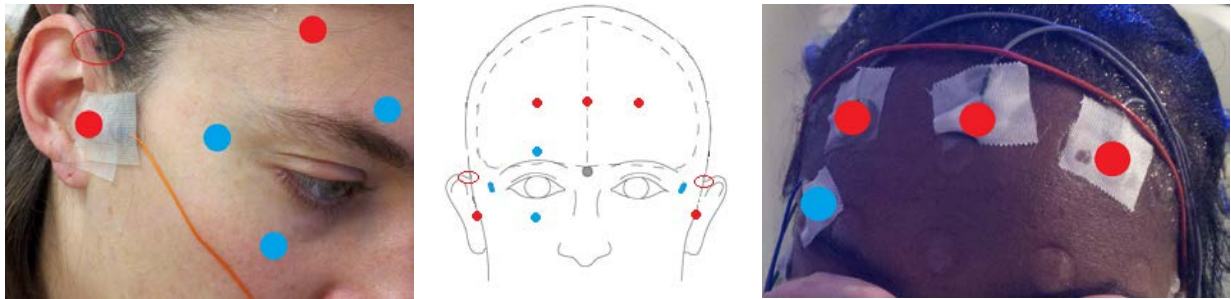
Five MEG position-sensor coils are affixed to the head between 1 and 5 cm from the hairline: at each tragus, midline above nasion, and above each eyebrow (as in red dots in photos).

Coil 1 – in front of participant's LEFT tragus (not shown)

Coil 2 – in front of participant's RIGHT tragus (red dot in photo)

Coil 3 – center of forehead near hairline (red dot in center of rightmost photo)

Coil 4 – left side of forehead near hairline in line with Coil 3 (as in rightmost photo)

Coil 5 – right side of forehead near hairline in line with Coil 3 (as in rightmost photo)

To reject cardiac and ocular activity from MEG data (see below for artifact removal pipelines), vertical eye leads (blue dots in photo) are applied above and below participant's right eye, centered on the pupil (channels 3 & 4). Horizontal eye leads (blue dot) are applied about 1cm lateral to the participant's outer canthus of each eye (right as in photo, left not shown, channels 5 & 6). ECG (electrocardiography) sensors are applied to the bilateral collarbone, or to lower chest if the collarbone signal is poor.

A ground electrode (GND channel) and a reference electrode (REF1 channel) are applied to the participant's collarbone.

EMG (electromyography) electrode stickers are applied to the participant's skin as in the photos below. The participant is asked to extend the toes and the electrode is placed by palpation over the active muscle.

Sensors are placed on the lateral superior surface on the extensor digitorum brevis muscle (orange ovals) and near the medial malleolus of each foot.
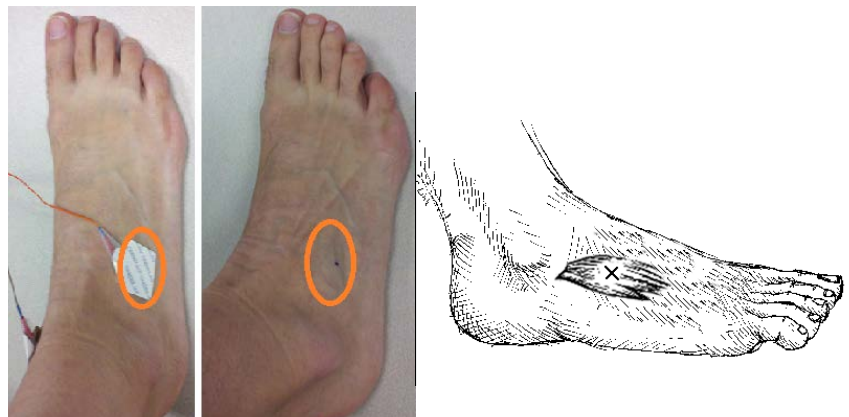
Sensors are placed on the first dorsal interosseus muscle between thumb and index finger, and the styloid process of the ulna at the wrist of each hand.

Stimulus presentation information

Stimuli for task MEG are generated using E-Prime 2.0 Pro on a HP personal computer. The stimulus computer presents visual stimulus through an LCD projector (ImagePro 8935, DUKANE) onto a mirror ~3 feet above the participant at 1024x768 resolution for viewing.

Participant button press responses are recorded via fiber optic as the Response channel in the MEG data. Motor-task motion responses, EOG and ECG are recorded via electrodes as described above.

Audio stimuli are output to an Audiosource amplifier that drives two 300-ohm Etymotic Research (ER-30) transducers. Sound is transferred to the subject through flexible hollow tubes (~7 m), with a measured auditory delay of 20 ms. The delay is compensated in software by offsetting the trigger from the auditory presentation, providing a coincident trigger and stimulus.

For visual stimuli, the TTL triggers and another digital trigger signal from a photodiode installed on the right-top corner of the screen are combined and recorded as the Trigger channel in the MEG data.

All TTL triggers (from the parallel port) of the stimulus computer are recorded as the Trigger channel in the MEG data.

Summary of MEG acquisition protocols

MEG acquisitions were collected for each subject in one 3 hour recording session at Saint Louis University. Our acquisition protocol is included in the standard operating procedures (SOP) Appendix II. Subjects lie supine with body supports to encourage staying still, and are positioned with the top of their head touching but not pressing against the dewar, and the back of their head cushioned by one towel and a thin layer of bubble wrap for comfort.

The MEG recordings for a subject are split over several runs or scans, e.g. three runs for the resting state data and 2 runs for the motor task. At the start and end of each run the head position is recorded. There is no continuous recording of the head position.

The following provides basic parameters for the main scan types in the session. A more complete set of scanning parameters can be found in the protocol exports from the scanner, available in Appendix I.

Empty room scans

Empty room measurements (5 min) are acquired daily and used to monitor the MEG system for hardware malfunctions or excessive environmental noise. These data are compared to a set of reference scans collected to estimate baseline noise levels and corrective action is taken appropriately. These data are not contained in the MEG1 release.

Participant noise scans

Participants are scanned for magnetic artifacts (1 min) and are degaussed and rescanned, if necessary. If multiple participant noise scans were made for a subject, only the final scan is uploaded and shared.

Resting-state MEG (rMEG)

rMEG data were acquired in three runs of approximately 6 minutes each. Subjects are instructed to remain still during scanning, delays between subject instructions and scan start may result in scan duration exceeding six minutes. During the scan the subject is supine, padded for comfort, instructed to relax with eyes open and maintain fixation on a projected red crosshair on a dark background (presented in a darkened room). ECG and EOG electrodes are attached to the subject for cardiac and oculomotor monitoring and off line artifact rejection. Resting-state scans always follow the empty room noise scan and the participant noise scan, and precede the task scans.

Condition	Runs	Run Duration (min:sec)
REST (Resting-state)	3	~6:00

Task MEG (tMEG)

Following completion of rMEG, subjects were asked to complete tasks that were designed to activate a variety of cortical and subcortical networks. The following table provides a listing of the tMEG scans collected.

tMEG data were acquired with the same parameters as rMEG, except for the run duration information listed below. Duration of each scan is somewhat flexible per run to ensure complete capture of the E-Prime scripts for each of the three tasks.

Task	Runs	Blocks per run	Run Duration (min:sec)
Working Memory	2	16	10:00
Story vs. Math (Language Processing)	2	7 story, 15 math	7:00
Motor	2	32	14:00

An E-Prime <Subject_ID>_MEG_<Task_run>.tab (tab-delimited text file) and Excel spreadsheet provide output task parameters that can be used in various tMEG analysis packages; see Task MEG files and protocol details.

Full scanning protocols (PDF)

Complete scanning protocols for each MEG acquisition can be found in Appendix I: HCP MEG Session Protocol Details.

Standard Operating Procedures (SOPs)

A major effort has been made to establish consistent procedures for all aspects of data acquisition and data processing. These are described in a set of Standard Operating Procedures (SOPs) that are included in Q3 Reference Manual Appendix 4 and in Appendix II: HCP MEG Protocol SOPs. These SOPs provide a useful reference for investigators wanting to know more about exactly what was done. Any outdated versions of SOPs are retained and are available on request.

Standard schedule for subject visits

Most subjects complete the full HCP protocol during a two-day visit. If a subject is one of the 100 MEG protocol subjects, the full protocol is typically completed in a three-day visit. Typical 2-day and 3-day schedules are shown below. In the subset of participants who receive MEG sessions, a MEG day will be scheduled before HCP MRI Day 1 because the superconducting sensors in the MEG scanner can be affected by normally imperceptible residual magnetization from the MR scanner's static field. In many cases, the participant's schedule necessitates condensing the MEG, behavioral, and MRI sessions into 2 long acquisition days, but MEG acquisitions always precede MRI sessions.

Breaks are included in the schedule to provide the researchers the opportunity to adjust experimental equipment, and the participant a chance to stretch and relax their eyes, as well as relieve any stiffness or pressure from lying so completely still in the scanner

MEG Subjects WU-Minn HCP Schedule

MEG Day

Time	Activity
7:30	Set-up and MEG scanner QC
8:30	Subject consent at WU and travel to SLU
	SLU Subject consent and Task practice
9:30	Empty Room and Participant Noise Scan
	Head Surface Digitization
10:30	rMEG scans (25 min)
	tMEG WM scans (25 min)
11:30	tMEG Story/Math scans (20 min)
	Break and application of EMG electrodes
12:30	tMEG Motor scans (30 min)
	Satisfaction survey
1:30	Return to WU

Typical 2-day WU-Minn HCP Subject Schedule

Time	Day 1	Day 2
7:30	Set-up and scanner QC	Set-up and scanner QC
8:30	Consent and Intake tests*	MR check and Diffusion session
	Mock scanner practice	
9:30	MR check and Structural session	NIH Toolbox Behavioral Tests
10:30		
	Non-Toolbox Behavioral Tests	
11:30		Lunch break
	Lunch break	Task practice and MR check
12:30	Task practice and MR check	R-fMRI session #2 (30 min)
	R-fMRI session #1 (30 min)	T-fMRI session #2 (30 min)
1:30	T-fMRI session #1 (30 min)	Drug/alcohol tests, tobacco/alcohol 7d retrospective, satisfaction survey
2:30	Recognition task (ex-scanner)	

*Drug/alcohol tests, surveys: mini mental state, sleep quality, menstrual info, HbA1c, TSH, hematocrit, blood for genotyping.

Feb. 25, 2013

Data in the Initial MEG release

Data in the HCP Initial MEG release is comprised of 10 individual subject scans each of resting-state MEG (rMEG), and each of 3 task MEG (tMEG) paradigms (working memory, language processing [story-math], and motor) in unprocessed and channel-level processed formats. Task data also include minimally processed E-Prime output information for users. Task order was not counterbalanced, all were run in the same order: working memory first, then language processing [story-math], then motor.

The 10 scans of each type were collected on a total of 14 subjects, in quarters 3, 4, and 5 (Q3-Q5) of 2013-2014. Because we recruit monozygotic twin pairs, some of these subjects are related twins. Anatomical models for MEG source estimation and scans of baseline empty room noise and baseline participant noise are available included in the release for all 14 subjects. The modalities of unprocessed and processed data available for each subject included in the release are indicated in the table below:

Subject ID	Noise	Anatomy	Resting State	Working Memory	Story-Math	Motor
125525*	✓	✓	—	✓	—	✓
133019	✓	✓	✓	✓	✓	✓
149741	✓	✓	✓	—	✓	—
177746*	✓	✓	✓	✓	✓	✓
187547	✓	✓	—	—	✓	✓
189349	✓	✓	✓	✓	—	✓
191841	✓	✓	✓	✓	—	✓
200109	✓	✓	—	—	✓	✓
233326	✓	✓	✓	—	✓	—
433839	✓	✓	✓	✓	✓	✓
500222	✓	✓	—	✓	—	✓
599671	✓	✓	✓	✓	✓	✓
784565	✓	✓	✓	✓	✓	✓
917255*	✓	✓	✓	✓	✓	—

*In addition to MEG data, these subjects currently have behavioral and MR imaging data available (links are provided on the MEG Download page) in ConnectomeDB. See the HCP Q3 Reference Manual for details on downloading these data.

The sections on Directory Structure for the Unprocessed, Anatomical Model, and Channel-level Processed MEG data provide an overview of the content of the release. A detailed file-by-file listing of the content of the release is provided in the Appendix III: File Names and Directory Structure for Unprocessed, Anatomical Models, and Channel-Level Processed MEG Data.

Standard MEG session structure

Each subject was scanned in one 3 hour MEG session, with the following shorthand labels:

MEG Day: MEG Session

- Empty room scan
- Participant Noise Scan
- Head surface digitization
- Three 6-min rMEG scans
- Six tMEG scans (2 scans x 3 tasks: Working Memory, Language Processing [Story-Math], Motor)

The tMEG was collected as six scans, two scans per task type, to acquire enough repetitions to create averages without causing participant fatigue.

In some cases, scans were aborted and re-acquired within the same session. This was done if the subject did not comply with task instructions or if technical problems prevented a full set of acquisitions. In these cases, the restart was noted on the acquisition data sheet. Only full length scans are included in the released data.

Physiological Data

We acquire electrooculography (EOG) to measure eye movement, electrocardiography (ECG), to measure electrical activity of the heart, and electromyography (EMG) to measure electrical activity produced by skeletal muscles movement signals associated with each scan.

Data Conversion: 4D import to FieldTrip

The raw data are stored on disk in a binary file named 'c,rfDC', which consists of the actual data (16-bits per sample), followed by a 'footer' that contains the metadata needed to interpret the signals. In addition, each data file is accompanied by a file named 'config', which contains additional metadata needed to interpret the signals. The raw binaries can be read with the custom software that comes with the 4D-system, or with FieldTrip. The latter approach requires a MATLAB license. FieldTrip contains the low-level code that is needed to interpret the metadata. More information about using FieldTrip for importing the data into MATLAB can be found on the FieldTrip website http://fieldtriptoolbox.org.

Quality control and artifact identification

Quality control and artifact identification steps are included in the pre-processing of raw MEG data using the "datacheck", "baddata", and "icaclass" pipelines. These utility pipelines are described in more detail in the MEG Channel-level processing pipelines section below.

All scans are uploaded from the scanner at SLU and permanently stored in an internal database (IntraDB). Scans that are considered usable for preprocessing are transferred to the public-facing ConnectomeDB based on information encoded in a set of scan-specific fields in IntraDB. This process also entails a reorganization of data from session/modality-specific directories to a composite subject-specific directory in ConnectomeDB, with subdirectories appropriate for the scan modality. Files of a given modality are grouped with ancillary files from the same session as needed in order to facilitate preprocessing using standardized scripts and pipelines. Many file names are modified and standardized as part of this transfer and reorganization process.

Directory structure for unprocessed MEG data

Unprocessed datasets downloaded from ConnectomeDB are organized into subject-specific and scan-specific archives. A full list of file names in each directory and subdirectory is provided in Appendix III: File Names and Directory Structure for Unprocessed, Anatomical Models, and Channel-Level Processed MEG Data, Section A.

Unprocessed data for exemplar subject 012345 unpacks to the following directory structure:

```
012345/unprocessed/MEG/
        1-Rnoise/4D
        2-Pnoise/4D
        3-Restin/4D
        4-Restin/4D
        5-Wrkmem/4D
        6-Wrkmem/4D
        6-Wrkmem/EPRIME
        7-Wrkmem/4D
        7-Wrkmem/EPRIME
        8-StoryM/4D
        8-StoryM/EPRIME
        9-StoryM/4D
        9-StoryM/EPRIME
        10-Motort/4D
        10-Motort/EPRIME
        11-Motort/4D
        11-Motort/EPRIME
```

The MEG/ subdirectory signifies that these data were acquired in the MEG lab at Saint Louis University (SLU). Since all subjects will also be scanned at 3T at Wash U, the 3T data (separately available on ConnectomeDB and in Connectome in a Box) will unpack to a 3T/ subdirectory. For the subjects that are later scanned at 7T (200 of the 1200), the 7T data will unpack to a 7T/ subdirectory.

Within each subdirectory, files are named as indicated by the exemplars below for several scan types.

The MEG data was acquired in 11 scans. Scan 1 and 2 are empty room (R) and participant (P) noise measurements, scan 3-5 are resting state measurements, scans 6-11 are task measurements. The 4D subdirectory signifies that the data originates from the 4D Magnes MEG scanner. The EPRIME subdirectory contains the log files from the stimulus presentation software (only for the task measurements).

For the noise scans, the 2 scans are downloaded into separate subdirectories (1 and 2), because the initial preprocessing is carried out separately for each scan. For example, for the first noise scan:

```
1-Rnoise/4D
        config
        c,rfDC
```

The config file contains additional header information in a mixture of textual and binary representation and the c,rfDC file contains the raw data in binary format (see [Data Conversion: 4D import to FieldTrip](#) above for more information). Note that the two noise scans (1 and 2) do not have head localization data.

The three (6 min) rMEG scans for each subject are downloaded into separate subdirectories (3-5), because the initial preprocessing is carried out separately for each scan. For example, for the first rMEG scan:

```
3-Restin/4D
        config
        c,rfDC
        e,rfhp1.0Hz,COH
        e,rfhp1.0Hz,COH1
        hs_file
```

As above, the config file contains additional header information and the c,rfDC file contains the raw data. The e,rfhp1.0Hz,COH file contains the head localization data at the start of the scan, the e,rfhp1.0Hz,COH1 file contains the head localization data at the end of the scan, and the hs_file contains a 3D-point cloud describing the participant's head surface, registered to the MEG coordinate system.

For tMEG scans, the 6 scans (3 tasks x 2 runs each in subdirectories 6-11) are downloaded into separate subdirectories, because the initial preprocessing is carried out separately for each scan. EPRIME log files for each scan are available in ASCII tab-delimited format (*.tab) and in Microsoft Excel (*.xlsx) format. For example, for the first working memory task scan:

```
6-Wrkmem/4D
        config
        c,rfDC
        e,rfhp1.0Hz,COH
        e,rfhp1.0Hz,COH1
        hs_file

6-Wrkmem/EPRIME
        012345_MEG_Wrkmem_run1.xlsx
        012345_MEG_Wrkmem_run1.tab
```

MEG Anatomy processing pipeline

The MEG anatomy pipeline generates data, constructed from anatomical information, that are needed for MEG/EEG source reconstruction:

1. coregistration information that defines transformations between the different relevant coordinate systems: voxel-based, MEG-sensor based coordinate system, and an ACPC-based RAS coordinate system, where the origin is defined in the anterior commissure, and the X, Y and Z axes run in right, anterior and posterior direction, respectively.
2. generation of a single shell volume conduction model.
3. generation of a set of source models (cortical sheet and 3-D grids) defined on a regular 3D grid in normalized space.

The single shell volume conduction model is defined in the MEG-system based head co-ordinate system and was obtained by segmentation of the anatomical MRI, using SPM8.

The source models are defined on a regular 3D grid in normalized MNI-space. For each participant the anatomical image was non-linearly volumetrically warped to a template image (using SPM8), and the inverse of this warp was applied to the positions on the grid. As a consequence, different source reconstructions that are defined on this grid (e.g. from different sessions or subjects) can be directly averaged, because each position in the grid represents the same location in normalized space.

How is the MEG coregistered to the anatomical MRI and volume conduction model?

Coregistration is performed using an unmasked 1mm resolution T1-weighted anatomical image (which, due to anonymity requirements, is not part of the release). The geometric information that is needed for source reconstruction (i.e. MEG sensor positions and information derived from the anatomical MRI) is expressed in the head coordinate system, which is defined as being relative to the participant's head and based on a set of landmarks on the head, and is thus subject-specific. This coordinate system is expressed in meters, with the principal (X, Y, Z) axes going through external anatomical landmarks on the head (i.e. fiducials). The details are:

- the origin is exactly between the left and right pre-auricular points (LPA and RPA)
- the X-axis goes towards the nasion (NAS)
- the Y-axis goes approximately towards LPA, orthogonal to X and in the plane spanned by the fiducials
- the Z-axis goes approximately towards the vertex, orthogonal to X and Y

The magnetometer locations in the raw and processed channel level data are represented in this coordinate system. The "anatomy" pipeline performs the coregistration of the MEG with the

MRI data and results in in a volume conduction model and various source models (cortical sheet and 3-D grids), which are all expressed in this head coordinate system. The "anatomy" pipeline also produces the 4x4 homogenous transformation matrices to convert between individual head coordinates, MNI coordinates and voxel coordinates. These transformation matrices are represented in the 012345_MEG_anatomy_transform.txt files.

What is the interpretation of the different homogeneous transformation matrices?

The transformation matrices represented in the 012345_MEG_anatomy_transform.txt files represent the transformation between different coordinate systems. Each matrix is named [from]2[to] where it denotes the transformation from coordinate system [from] into coordinate system [to]. Since the release does not contain the original anatomical images that were used to create the transformation matrices, all of the matrices that contain 'vox' are of limited use to external users. The transformation matrices that are of relevance are the bti2spm and spm2bti, which convert between the MEG-system based coordinate system (bti) and a coordinate system based on the ACPC-based RAS (spm).

How are the MEG source models defined?

The 3-dimensional grids that can be used for (beamformer) source reconstruction are based on 3-dimensional grids with a regular spacing of the dipoles, defined in normalized MNI space. The spacing between the dipoles (in normalized space) is represented in the filename (i.e. SUBJECTID_MEG_anatomy_sourcemodel_3d4mm means that the dipoles are 4mm apart). For each participant, the template grid has been warped to individual head space in order to express the dipoles' coordinates in the head coordinate system that is used for source reconstruction (see above). This warp is a non-linear one, and is based on the inverse of the nonlinear volumetric spatial normalization from individual head space to normalized MNI space. Consequently, the dipoles' coordinates are not placed on a regular 3-dimensional grid in individual head space. Yet, the source models allow for easy averaging across subjects, without the need for an intermediate interpolation step, because individual grid points coincide in normalized space.

Directory structure for anatomical models for MEG source estimation

Anatomical model datasets downloaded from ConnectomeDB are organized into subject-specific archives. When unpacked, they are moved into the <subject_id>/MEG/anatomy subdirectory. A full list of file names in each directory and subdirectory is provided in Appendix III: File Names and Directory Structure for Unprocessed, Anatomical Models, and Channel-Level Processed MEG Data, Section B.

Anatomical model data for exemplar subject 012345 unpacks to the following directory structure:

```
012345/MEG/anatomy
        012345_MEG_anatomy_fiducials.txt
        012345_MEG_anatomy_landmarks.txt
        012345_MEG_anatomy_transform.txt
        012345_MEG_anatomy_headmodel.mat
        012345_MEG_anatomy_sourcemodel_3d4mm.mat
        012345_MEG_anatomy_sourcemodel_3d6mm.mat
        012345_MEG_anatomy_sourcemodel_3d8mm.mat

        provenance/
                012345_MEG_anatomy_fiducials.txt.xml
                012345_MEG_anatomy_landmarks.txt.xml
                012345_MEG_anatomy_transform.txt.xml
                012345_MEG_anatomy_headmodel.mat.xml
                012345_MEG_anatomy_sourcemodel_3d4mm.mat.xml
                012345_MEG_anatomy_sourcemodel_3d6mm.mat.xml
                012345_MEG_anatomy_sourcemodel_3d8mm.mat.xml

        figures/
                012345_MEG_anatomy_headmodel.png
                012345_MEG_anatomy_sourcemodel_3d4mm.png
                012345_MEG_anatomy_sourcemodel_3d6mm.png
                012345_MEG_anatomy_sourcemodel_3d8mm.png

        figures/provenance/
                012345_MEG_anatomy_headmodel.png.xml
                012345_MEG_anatomy_sourcemodel_3d4mm.png.xml
                012345_MEG_anatomy_sourcemodel_3d6mm.png.xml
                012345_MEG_anatomy_sourcemodel_3d8mm.png.xml
```

The anatomical coregistration information is contained in the *fiducials.txt, *landmarks.txt, and *transform.txt files, *headmodel.mat contains the volume conduction model (headmodel), and the source models at different resolutions are contained in the *.sourcemodel_3d*mm files.

Provenance information (in Extensible Markup Language, i.e. *.xml), quality control figures (in Portable Network Graphics format, i.e. *.png) and provenance information for the figures. are contained in the provenance/, figures/, and figures/provenance/ subdirectories.

Overview of MEG Channel-level processing pipelines
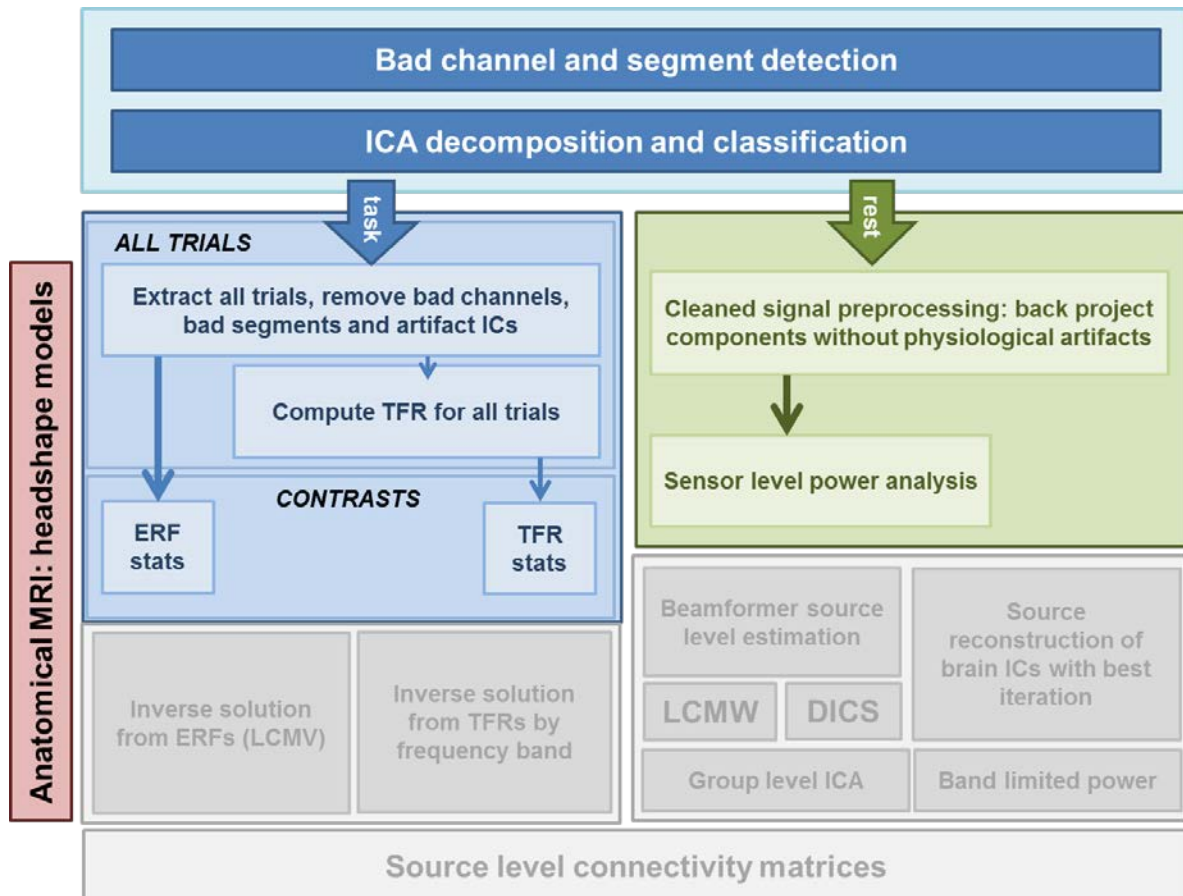
Overview of HCP MEG Processing

HCP MEG datasets were processed by a set of custom pipelines aimed at providing high quality data. The figure above shows the overall workflow for MEG data processing in the HCP. Artifacts, bad channels, and bad segments are identified and removed from rMEG and tMEG data which are processed by an ICA-based pipeline to identify artifacts (environmental and subject's artifacts) and non-artifacts components of the overall signal. The rMEG pipelines available in this release produce: i) a set of non-artifact ICs comprising time-courses and sensor maps; ii) cleaned data (raw minus artifacts) at the channel level in 2 sec chunks; iii) channel-level power spectrum densities. The tMEG pipelines available in this release provide: i) clean data for the Task experimental paradigms, split in various groups of trials according to different conditions of interest, ii) Event Related Field averages at the sensor level for a number of conditions and for comparisons between pairs of conditions for each Task paradigm, and iii) Time-Frequency Representation averages at the sensor level for a number of conditions and for comparisons between pairs of conditions.

A summary of the major processing steps for each pipeline is provided below. The following section, Directory structure for Channel-level processed MEG data summarizes the directory structure created by the processing pipelines. Appendix 3C contains a complete list of all resulting processed file names and subdirectories.

The pipelines use freely available software from the *FieldTrip* analysis toolbox and custom scripts developed by the HCP that are collectively referred to as "megconnectome". Both megconnectome software and the FieldTrip toolbox are available for download on the HCP website. The megconnectome software consists of analysis scripts, analysis functions and template data, e.g. the channel layout and a canonical MRI. Each analysis script represents one of the steps of the analysis pipeline as outlined below and in Larson-Prior *et al.*, 2013.

Improvements in MEG processing pipelines that occur over time may warrant retrospective application of updated pipelines to previous data releases. The HCP will document any such changes in processing pipelines between versions and indicate which versions are used in the release notes provided with each dataset.

MEG Data Processing Output File Formats

The MEG data processing pipelines output files in .txt, .mat, and .xml file formats. Below are details on each of these formats.

*.txt Files

The files ending with the *.txt extension are ASCII text files with 8-bit Unicode Transformation Format (UTF-8) content. The content of the files is structured to be self-explanatory. An example is given below:

```
 badsegment.zscore = [
419334 426841
427909 428408
];
badsegment.ica = [
1 113472
417760 427193
617686 620126
702465 704442
743761 745389
928217 979475
];
badsegment.manual = [
1 113472
```

928217 979475
];

Although the files have the extension *.txt, their content can be directly interpreted by MATLAB as a valid script. Since the file names start with numbers and have special characters in them, such as the '-' (minus) sign, you cannot simply rename the extension from *.txt to *.m, as MATLAB script file names cannot start with a number and cannot have a '-' in their name. The megconnectome/analysis_functions/hcp_read_ascii.m function can be used to read the content into MATLAB workspace, just like the standard MATLAB "load" command would be used for *.mat files.

*.mat Files

The files ending with the *.mat extension are binary MATLAB files. The file format of these files is specified at http://www.mathworks.com/help/pdf_doc/matlab/matfile_format.pdf. Over the years, MATLAB has introduced a number of file formats. The following overview is derived from the MATLAB documentation:

- v4: Can only contain two-dimensional double, character, and sparse arrays.
- v6: Supports all version 4 features plus N-dimensional arrays, cell and structure arrays, and variable names greater than 19 characters.
- v7: Supports all version 6 features plus data compression and Unicode character encoding.
- v7.3: Supports all version 7.0 features plus support for data items greater than or equal to 2GB on 64-bit systems

From mat-file format version 7.3 onward, the files are based on the Hierarchical Data Format (HDF) standard. Although this makes them in principle more accessible in other platforms, in practice the v7 files can more easily be accessed, as it is a simpler file format for which the specification is available. Furthermore, the v7 format is in general slightly smaller than the v7.3 format, hence we have chosen to store data in the v7 file format.

In MATLAB you can read these files using the "load" command.

In Python you can read these files using

>>> import scipy.io
>>> data = scipy.io.loadmat('file.mat')

*.xml Files

The files ending with the *.xml extension contain provenance information in Extensible Markup Language (XML) format. These files always accompany another file, e.g. the file example.mat is

accompanied by the provenance file example.mat.xml. The content of the XML file is encoded with 8-bit Unicode Transformation Format (UTF-8) and can be read in any text editor. The format should be self-explanatory, an example is given below:

```
<?xml version="1.0" encoding="utf-8"?>
<megconnectome xmlns:xsi="http://www.w3.org/2001XMLSchema-instance"
xsi:noNamespaceSchemaLocation="megconnectome.xsd">
  <version>
    <matlab>
      <Name>MATLAB</Name>
      <Version>8.0</Version>
      <Release>(R2012b)</Release>
      <Date>20-Jul-2012</Date>
    </matlab>
    <megconnectome>
      <Name>megconnectome</Name>
      <Version>0.7</Version>
      <Release>www.humanconnectome.org</Release>
      <Date>11-Dec-2013</Date>
    </megconnectome>
    <fieldtrip>
      <Name>FieldTrip</Name>
      <Version>r9020</Version>
      <Release>fieldtriptoolbox.org</Release>
      <Date>11-Dec-2013</Date>
    </fieldtrip>
  </version>
  <compiled>true</compiled>
  <username>schoffelen</username>
  <hostname>node165</hostname>
  <architecture>glnxa64</architecture>
  <buildtimestamp>11-Dec-2013 17:50:06</buildtimestamp>
  <pwd>/home/schoffelen</pwd>
  <matlabstack>
In hcp_write_provenance at 37
In hcp_write_matlab at 62
In megconnectome at 103
</matlabstack>
  <filename>233326_MEG_anatomy_sourcemodel3d6mm.mat</filename>
  <md5sum>31c9746a16fa7ac53dfd285e9bc3112c</md5sum>
</megconnectome>
```

MEG Preprocessing Pipelines

A set of (utility) preprocessing pipelines are run on all MEG data prior to further resting-state or task-specific analysis. Overall, these pipelines perform various sanity checks and quality assurance tasks.

"datacheck" pipeline

In the "datacheck" pipeline, the header and events of each MEG dataset are checked for basic features (such as sampling rate, number of channels and whether EOG and ECG are present). The dataset check pipeline results in a text file and figures that provide qualitative information on MEG data quality (see exemplar files in **Directory structure for Channel-level Processing Pipelines-** Datacheck). It does not perform a detailed analysis of artifacts, bad channels or data segments. Measures of data quality are examined (output in six figures per run):

1. SQUID jumps: (values ≤~1e12), to identify bad sensors or acute noise increases during a run.

2. Trigger / Response events: (values 0-400), to check for appropriate numbers and values of trigger events and response events for a given run type.

3. Power Line Noise: (values <1.5e-27 T^2/Hz), to check for sensors with unusually great power, (e.g. A246 is usually noisy indicating a bad sensor).

4. MEG Powerspectra – to confirm three features: rough linearity in the range of sampling frequency, 1/f coarse overall pattern, and expected 60Hz noise. Deviant channels (unusual spikes or unusual power) are noted.

5. MEGREF Powerspectra - to confirm rough linearity in the range of sampling frequency and 1/f coarse overall pattern. Deviant channels (unusual spikes or unusual power) are noted.

6. Neighbor Correlation – to check for smooth transitions between neighbors. Correlations are expected to be greater than 0.6 (except for empty room).

"baddata" pipeline

MEG processing requires that artifact signals are removed from data. The artifacts displaying excessive signal amplitude are generated by non-working channels and bursts of noise and are referred to as "bad channels" and "bad segments", respectively. The purpose of this pipeline is to detect these bad channels and bad segments.

The pipeline input consists of the raw MEG data and a list of manually defined bad channels and bad segments. Different strategies for detection of bad channels and segments are applied,

and their outputs can be combined or used separately for further processing of the data. At the end of the computation, output QC figures are produced. This QC step is mainly used to evaluate the quality of the scan, in terms of number of bad channels and number and duration of bad segments. Additionally, QC figures allow one to identify whether further non-working channels or bad segments not detected by the pipeline are still depleting the data, so that the pipeline should be run again. In this case, if any channel or segment is manually selected, the baddata pipeline is run again using these selected channels/segments as input.

Epochs at the beginning and at the end of each resting state fixation session are automatically selected as bad, since in these epochs the subject's noise (eye blinks and saccades) is considerably larger than during the central part of the run (see below).

For task runs, data between the first and last triggers limit the portion of data to be preprocessed by the baddata pipeline. The output of the baddata pipeline consists of two ASCII files, called subject_session_baddata_badchannels.txt and subject_session_baddata_badsegments.txt (see below for samples).

- **What does it mean if a channel is marked as "bad"?**

    A channel is marked as "bad" when the signal recorded by that channel is abnormal and cannot be used for further processing. The presence of these channels would have a negative impact on both sensor level analysis and source reconstruction. There might be a set of reasons invalidating the "bad" channel output: i) the SQUID is not working; ii) the SQUID is extremely noisy; iii) the channel read out electronics is not working; iv) the channel read out electronics is not locked (that is, the flux locked loop scheme is temporarily not working and should be reset) during that recording and the output is meaningless.

- **Why do subsequent scans of the same subject have different bad channels?**

    If the SQUID is damaged or noisy, or if the readout electronics is not working, the related channel will always be classified as "bad". Conversely, if the channel loop was not temporarily locked, the channel may not be flagged as "bad" in the subsequent recording. Indeed, the channel array is reset before every run, and this operation might recover the "former bad" channel. This implies that the set of "bad" channels will include both permanently and temporary non-working channels, and the temporary ones might be different for each run.

- **What does it mean if a segment is marked as "bad"?**

    During the recordings, short time epochs of high variance might occur. During these epochs, sources lying outside the brain (noise sources) generate a magnetic field which is several times larger than the brain magnetic field. Those fields might be generated by the subject him/herself, as for example eye squints, or might be related to external

sources, such as car traffic inducing mechanical vibrations of the shielded room or power supply bursts.

- ***What is the consequence that in some Resting-state sessions the data acquisition was started prior to starting the Resting-state E-Prime script?***

    During the Resting-state scans of some of the subjects, the data acquisition was started prior to starting the E-Prime script. As a consequence the initial segment of MEG data may be noisy, which could reflect the fact that the participant was not yet in the resting-state. In these instances it is advisable to discard for analysis any data that has been specified manually as bad segment in the accompanying ASCII files (as in e.g. 012345_MEG_3-Restin_baddata_manual_badsegments.txt). The manual bad segments have been specified in order to realize a 300-second long scan, where any data in excess of this has manually been specified to be a bad segment. 2/3 of this data in excess of 300 was specified at the beginning of the scan, the remaining 1/3 at the end.

The pipeline can be represented by the following logical steps (see figure below):

1. Bad channels are selected through the estimation of the correlation between each channel and its neighbors. This approach is based on the assumption that the signal of nearby sensors is very similar due to the spread in the propagation of the brain magnetic field (Winter *et al.*, 2007). A bad channel is then identified by searching for outliers in the neighbor correlation distribution by means of a statistical threshold (3σ below the mean neighbor correlation from all the channels). Once these channels are identified they are removed from subsequent analysis by this pipeline.

2. Each MEG channel is scanned for the identification of noisy segments based on the Z-score of each data point relative to the statistical characteristics of the entire data time-series of a channel. First the Z-score of each data point is computed and the data points that have a Z-score higher than a threshold (20) is flagged as bad. Then each channel is scanned for clipping
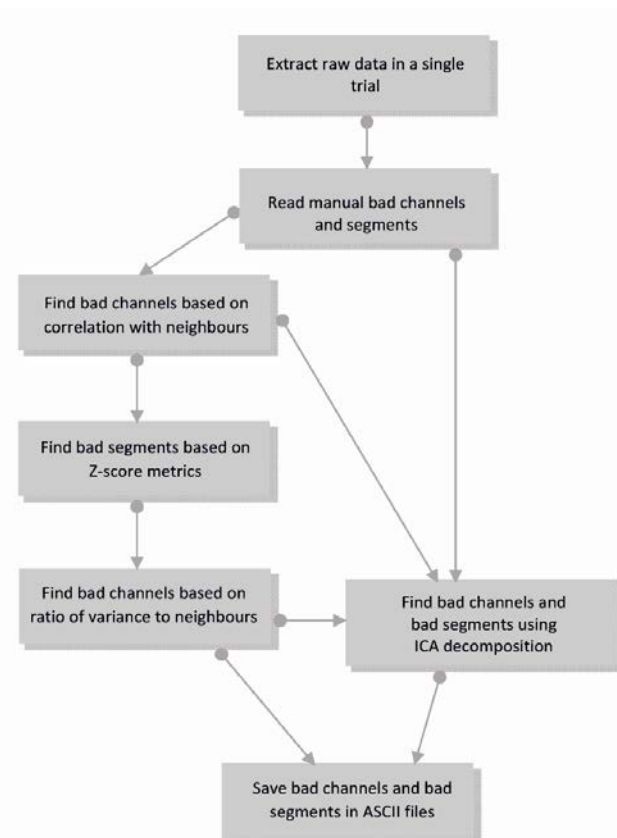
Steps of the baddata pipeline.

segments, during which the signal is flat. This is performed using FieldTrip's built in function ft_artifact_clip.m. Finally each channel is scanned for segments with muscle related artifacts. This is performed using FieldTrip's built in function ft_artifact_muscle.m. Once these segments are identified they are removed from data used in step 3 and are stored in the output.

3. Bad channel identification is performed by calculating the ratio between the standard deviations of each channel and its neighbors. This approach is based on the assumption that the signal in nearby sensors should have a comparable variance (Winter *et al.*, 2007). A bad channel is then identified if its standard deviation is more the 50% higher than that of its neighbors. Once these channels are identified they are removed from subsequent analysis by this pipeline.

4. An iterative approach based on the Independent Component Analysis (ICA) is used to decompose the sensor signals into independent components (ICs) in order to find components associated with non-working channels or epochs affected by artifacts such as large spikes and bumps. This approach uses the raw data and the manual list of bad segments (obtained in step 2) and the list of bad channels (obtained in steps 1 and 3) as input.

First, one run of ICA is applied to the data and bad channels are selected based on the sensor map topographies associated with the ICs produced by the approach. Indeed, sensor maps associated with non-working channels show a very large value on the problematic channel and values close to zero for the other channels. Automatic selection is then achieved by a statistical threshold (10σ) applied on the distribution of the sensor map values.

Bad segment selection is performed on each IC in two main steps. First, the IC time course is segmented in epochs with 1 s duration and the variance of the time course is estimated for each epoch. Epochs with high variance are selected through a statistical threshold (12σ). Second, to identify the bad segment with a finer time resolution, the power time course of the IC is estimated in a window of 10 seconds around the epochs with high variance. A statistical threshold is applied to the distribution of the power time course. This ICA approach is iterative, as it is run every time a bad channel or a bad segment is identified, with upgraded input, until no further bad channels and bad segments are found. The outputs of the baddata pipeline can be merged together (as it is done in the task analysis and channel level rest analysis) or can be used separately (as it is done in the ICA based analysis of resting state data, using only the bad segments produced by the ICA based approach).

- ***How are bad channels represented?***

  Bad channels are detected for every individual scan and represented in an ASCII file with extension *.txt. The content of the text file can directly be parsed by MATLAB or can be imported using the megconnectome hcp_read_ascii function. An example is given here:

badchannel.neigh_corr = {

'A2', 'A90'

};

badchannel.neigh_stdratio = {

'A195', 'A20', 'A246', 'A41', 'A73'

};

badchannel.all = {

'A195', 'A2', 'A20', 'A246', 'A41', 'A73', 'A90'

};

- ***How are bad segments represented?***

    Bad segments are detected for every individual scan and represented in an ASCII file with extension *.txt. The content of the text file can directly be parsed by MATLAB or can be imported using the megconnectome hcp_read_ascii function. Segments are indicated by the begin and end sample in the raw data file, which starts with sample 1. An example is given here:

```
badsegment.zscore = [
291544 297604
589700 591042
591748 595895
];

badsegment.ica = [
1 89094
384791 395319
398478 405186
561693 570009
573970 576941
699447 743995
];


badsegment.manual = [
1 89094
699447 743995
```

```
    ];

    badsegment.all = [
    1 89094
    291544 297604
    384791 395319
    398478 405186
    561693 570009
    573970 576941
    589700 591042
    591748 595895
    699447 743995


    ];
    .
```

The directory structure including the results from the baddata pipeline is reported in in Appendix III: File Names and Directory Structure for Unprocessed, Anatomical Models, and Channel-Level Processed MEG Data, Section C.

"icaclass" pipeline

This pipeline performs decomposition of MEG raw data into brain and non-brain components. This is achieved by means of an Independent Component Analysis followed by an automatic classification of the obtained Independent Components (ICs).

Differently from the baddata pipeline, which is not classifying ICs and is only detecting bad channels and bad segments, here the IC decomposition is repeated several times starting from different initial conditions, and the decomposition with the lowest artifact residual and the largest number of brain components is retained as the best decomposition. This best decomposition is selected through an automatic classification of the ICs, which is based on the evaluation of 6 parameters quantifying temporal and spectral properties of the IC signals (3 parameters) and contribution of the eye or heart magnetic signals (3 parameters). The IC time courses and the related sensor maps related to the best decomposition are stored as the output. The pipeline produces also a summary txt file and output figures useful for the quality check step. Each figure shows summary plots of a single IC as well as the values of the classification parameters and the corresponding flag (brain or artifact).

The pipeline can be represented by the following logical steps (see figure below):

1. Data are filtered through a band-pass (1.3-150 Hz) and a notch (59-61/119-121 Kz) Butterworth filters. Bad channels and bad segments are removed from data.

2. ECG and EOG data are preprocessed with the same procedures as in 1.

3. Twenty iterations ICA decomposition (FastICA with deflation approach) of the preprocessed
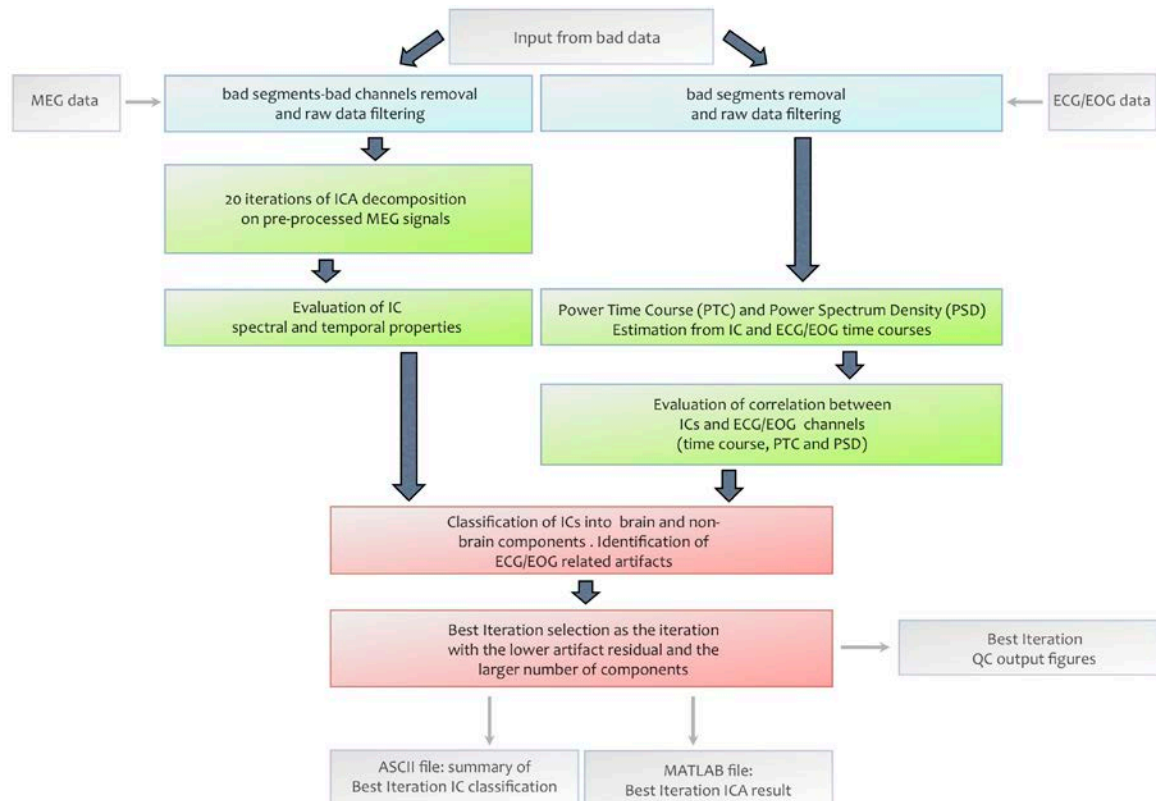
Steps of the icaclass pipeline.

data are run from different initializations.

4. Power Spectrum Density (PSD) and Power Time Course (PTC) are estimated from the IC time courses and the EOG/ECG time courses. Three correlation coefficients between: i) the IC and the reference time courses, ii) the IC and the reference PTC, and iii) the IC and the reference PSD are estimated.

5. Additional 3 parameters are estimated from the kurtosis of the IC time courses, the 1/f trend and the flatness in the PSD of each IC. These parameters allow classification of environmental and instrumental artifacts.

6. For each iteration of the ICA decomposition, classification into brain and non-brain components is carried out as follows: an IC is classified as a brain component is identified if none of the 6 parameters exceed its corresponding threshold. Moreover, if at least one of the 3

parameters based on the correlation with the EOG/ECG channels exceeds its threshold, the IC is flagged as ECG/EOG related artifact. The ECG/EOG components will be removed from data in the task pipelines. In the resting state pipeline only the brain components are used for further analysis.

7. The best iteration is selected as the one with the largest value of the product between the number of brain ICs and a number quantifying the quality of artifact rejection, defined as 1-artifact residual. The artifact residual is estimated as the maximum correlation between the time courses and PTC of all the IC and the EOG/ECG channels.

8. Results are saved as MATLAB and ASCII files. A summary on the icaclass output, including the number of total components, which ones are brain IC and which are not, and which ones are flagged as ECG/EOG artifacts, is in the file subject_session_icaclass.txt. A sample of the summary included in this file is given below:

vs.total_ic_number = 39;
vs.brain_ic_vs_number = 37;
vs.brain_ic_vs = [
 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 32 33 34 35 36 37 38 39
];
vs.ecg_eog_ic = [
 1 31
];

The directory structure including the results from the icaclass pipeline is reported in in Appendix III: File Names and Directory Structure for Unprocessed, Anatomical Models, and Channel-Level Processed MEG Data, Section C.

"icaclass_qc" pipeline

The purpose of this pipeline is to verify the automatic classification of independent components identified in the icaclass pipeline (above) by visual inspection. In the cases in which a component is misclassified, this step allows one to manually modify its classification. Misclassification can occur for a number of reasons. As an example, reference electrical channels might not be correctly placed on the subject or might not work. EOG/ECG reference channels are specifically used to identify artifacts that will be discarded in further analysis. If the ECG/EOG signals are inappropriate, the ICs capturing the heart and eye-related muscular activity may be classified like brain activity or as non-brain (but not ECG/EOG) components.

Another reason for misclassification might be that the EOG channel is capturing brain activity due to positioning (one of the electrodes is on the forehead) and thus some brain components are flagged as non-brain due to correlation with the EOG channel.

Finally, automatic classification of ICs is based on a set of thresholds, which were estimated from a ROC analysis on more than three hundred components classified by visual inspection. These components were obtained from Phase 1 subjects. Nevertheless, for statistical reasons a component can be misclassified because it slightly exceeds one of the thresholds. As an example, this might occur when the magnetic noise due to car traffic is much larger than in the ROC set, thus the 1/f threshold would not be adequate in this case.

If a component is modified, according to the new classification, the label "Brain Corrected", "Artifact corrected" or "Physiological Artifact Corrected" is shown in the output figure. In the file subject_session_icaclass_vs.txt, the field vs.flag = 1. The indexes of the corrected ICs are displayed in the fields: vs.good, vs.bad, and vs.physio for re-classified brain, non-brain and ECG/EOG components, respectively. Result files and figures include _vs in their name (see in Appendix III: File Names and Directory Structure for Unprocessed, Anatomical Models, and Channel-Level Processed MEG Data, Section C). If the automatic classification is able to properly classify all the components, then the above fields are empty.

rMEG Channel-level Analysis Pipelines

The rMEG Channel-level analysis pipelines 1) use the outputs of the "baddata" and "icaclass" pipelines (above) to clean and segment the rMEG data, and 2) generate the averaged resting state power spectrum for each sensor (channel).
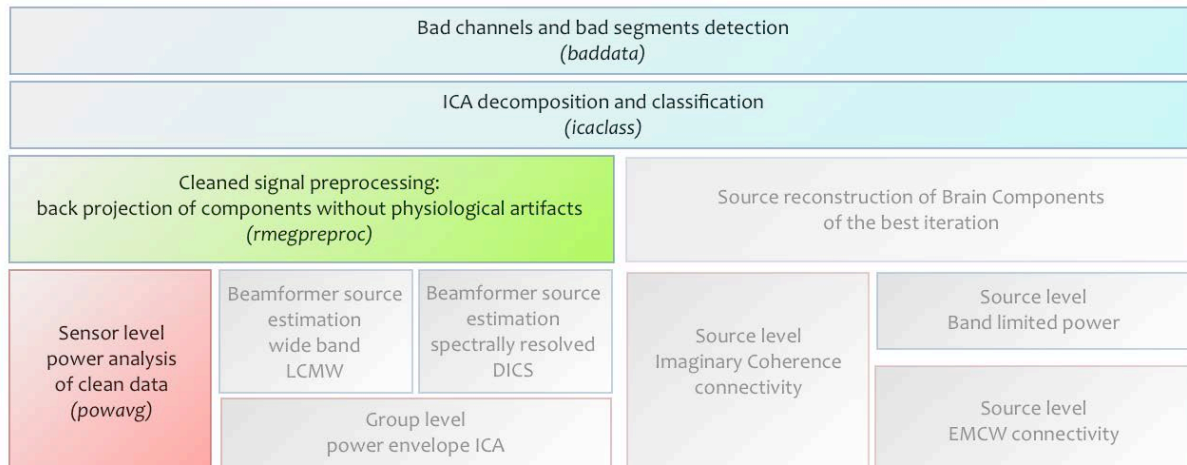
Steps of the rMEG processing pipelines.

"rmegpreproc" pipeline

The resting state preprocessing pipeline results in a cleaned representation of the sensor-level MEG data. The raw data are segmented into 2-second pieces. Bad channels and segments are removed. ICA components that have been classified as artifacts are projected out of the data. The cleaned data are represented as FieldTrip structure in a MATLAB file.

"powavg" pipeline

This pipeline results in the averaged power spectrum of the resting state MEG data at the sensor level. The continuous data are segmented, Hanning-tapered, Fourier-transformed and the estimated power is averaged over all segments. The power spectrum is represented as FieldTrip structure in a MATLAB file.

What are the filter settings used in the analysis of the resting-state MEG?

In the resting state pipeline, data filtering is performed by means of a Butterworh filter both in band-pass and band-stop case. Raw data are first band-pass filtered by means of a two-step procedure performing consequently high-pass and low-pass filtering (cutoff frequency 1.3 Hz and 150 Hz). Notch filters are then applied to the band-pass filtered data in order to remove the power line noise (cutoff frequencies 59-61 Hz and 119-121 Hz).

tMEG scripts and protocol details

Task-Evoked Brain Activity

Our primary goals in including task MEG (tMEG) in the HCP are to: 1) help identify well-characterized task-related brain areas or "nodes" that can guide, validate and help interpret the results of the connectivity analyses that will be conducted on resting state functional MRI (rfMRI), resting state MEG (rMEG) and diffusion data; 2) to allow a comparison of network connectivity in a task context to connectivity results generated using rMEG and rfMRI; and 3) to relate signatures of activation magnitude or location in key network nodes to individual differences in performance, psychometric measures, or other phenotypic traits. To accomplish these goals, for MEG we use task paradigms testing three domains that are a subset of those used in the HCP for tfMRI: Working Memory, Language processing (Story-Math), and Motor. Significant effort has been made to ensure that MEG and fMRI are performed on the same subjects using the same tasks performed, as closely as possible, with the same timing (see Larson-Prior *et al.*, 2013). Together with behavioral and genetic information on these subjects, these data will facilitate the study of individual variability across a large population.

These tasks are described in more detail below and in the cited publications. Stimuli are projected onto a mirror ~1 meter above the participant for viewing within the magnetically shielded room (MSR) in which all the MEG recordings are acquired.

tMEG scripts and data files

Script files are run in E-Prime 2.0 Pro to present task MEG stimuli and collect behavioral responses in the scanner. The E-Prime scripts can be made available upon request via email to Tracy at tracyn@npg.wustl.edu for researchers wishing to run HCP tasks in their own research. Please put "HCP-MEG EPRIME" in the subject line of your email. Tab-delimited text versions of E-Prime data files (in *.tab and *.xlsx formats) will be included in the released data. A brief description of the key variables in those files can be found in Appendix IV: Task MEG E-Prime Key Variables. The original *.edat2 files cannot be shared as they may contain identifying information.

MEG vs. MRI Naming Conventions

Despite the fact that the MEG and fMRI HCP data are collected using the same task paradigms (with slightly different protocol details) and the same eyes-open fixation resting state, differences in the scripting used in the processing pipelines between the MRI data and MEG data have necessitated slightly different naming conventions for each modality. These are shown in the table below.

Paradigm	MEG abbreviation	fMRI abbreviation
Resting-state	Restin	REST1 or REST2
Working Memory	Wrkmem	WM
Language Processing	StoryM	LANGUAGE
Motor	Motort	MOTOR

Details of the tMEG Working Memory task

Working memory is assessed using an N-back task in which participants are asked to monitor sequentially presented pictures. Participants are presented with blocks of trials that consisted of pictures of tools or faces. Within each run, the 2 different stimulus types are presented in separate blocks. Also, within each run, ½ of the blocks use a 2-back working memory task and ½ use a 0-back working memory task (as a working memory comparison). Participants are instructed to press a button for every picture. If the currently presented picture matches the cued picture (0-Back) or the same picture that was presented two pictures before (2-Back), subjects press one button with their right index finger. For non-matching pictures, participants press a second button with their right middle finger. Two runs are performed, 16 blocks each, with a bright fixation "rest" on dark background for 15 seconds between blocks.

Stimulus Overview

In the working memory task, participants are instructed to retain images in their visual working memory and compare them with subsequently presented images. There are 2 memory load conditions.

The first is a match-to-sample condition (termed in the following text as 0-Back) during which a cue "Target" image is presented at the beginning of a block and which the subject has been instructed to memorize. Then a sequence of 10 images is presented. Each of these

A. Illustration of the block sequence in the first run of the Working Memory Task. B. Example of the sequence of stimuli in a 2-Back block.

images is displayed for 2000 ms. At the end of this interval, the subject must respond on a button pad with the index or middle finger of the right hand if this image matched the Target or not, respectively. The response must take place within 500 ms, during which a fixation cross is presented on the center of the screen.

The second condition is a 2-Back condition. During such a block, the subject is presented with a sequence of 10 images and must respond if each image is identical to the one 2 positions earlier or not (figure, right). At the beginning of the block there is a cue screen informing the subject that the upcoming stimuli are part of the 2-Back protocol. The timing of the cue screen, the presentation of the 10 stimulus images and of the response interval are identical to that of the 0-Back condition.

There are 2 different categories of images used in this experiment: images of faces and tools. Each block contains images from a single category. Some of the images in the non-matched trials have been characterized as "Lure". These images have been selected so that they have common features with the target image, but are still different. These trials as flagged as "Lure".

In between the 0-Back and 2-Back blocks are interleaved Fixation Blocks, during which the subject has been instructed to remain fixated on a fixation cross at the center of the screen. The duration of each of these fixation blocks is 15 seconds.

Block/Trial Overview

This experiment is performed in 2 runs with a small break between them. The block/trial breakdown is identical in both runs although the presented stimuli are variable.

The following information describes the Block/Trial breakdown per run:

Total Number of Blocks = 24;
Number of 0-Back Blocks= 8;
Number of 2-Back Blocks = 8;
Number of Faces Blocks = 8;
Number of Tools Blocks = 8;
Number of Fixation Blocks= 8;

In the following per run trial breakdown the cuing stimuli are not included as trials:

Total number of image trials = 160;
Number of 0-Back trials= 80;
Number of 2-Back trials= 80;
Number of Faces trials= 80;
Number of Tools trials= 80;

Trigger Overview

The signal on the trigger channel consists of 2 superimposed trigger sequences. One from the Stimulus PC running the E-prime protocol and one from a photodiode placed on the stimulus presentation screen. The trigger channel for one Working Memory task run is shown, top right.

The photodiode is activated on the onset of each of the 10 stimulus images within each block. It is deactivacted when each of these stimulus images is removed from the screen. The trigger value for on is 255 and the trigger value for off is 0. The photodiode trigger sequence extracted from the trigger channel of one Working Memory task run is shown, middle right.

The Stim-PC triggers are more detailed. As in the initial trigger channel, these triggers are superimposed on the photodiode triggers, in the following description it is assumed that the photodiode triggers have been subtracted from the trigger channel so that only the triggers from the E-Prime stimulation protocol remain. Such an E-Prime trigger sequence extracted from the trigger channel is shown, bottom right.

Original Trigger channel sequence from one run of Working Memory Task. E-prime and Photodiode triggers are superimposed.
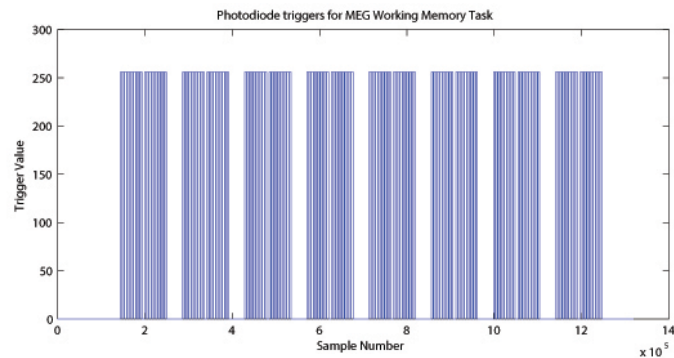
Photodiode trigger sequence extracted from the Trigger channel of one Working Memory Task run.

E-Prime trigger sequence extracted from the Trigger channel of one Working Memory Task run.

Because of the high number of combinations of Memory Load, Image category, and Match category, the trigger values by E-Prime follow the following algorithmic rules:

Memory Type Trigger Base = { 4   for   0-Back

                      68   for   2-Back }
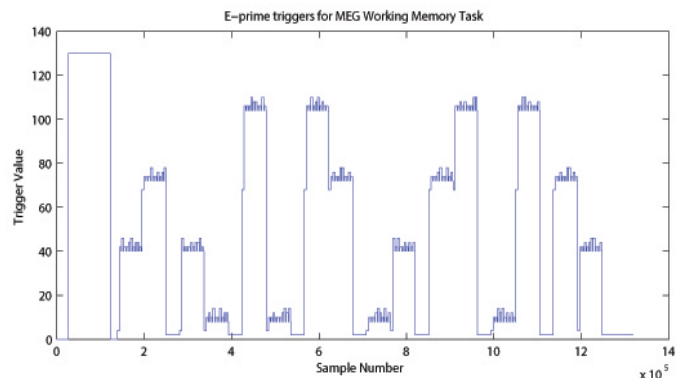
Image Category Trigger Pad=[{ 4   for Faces

36 for Tools }

Match Type Trigger Pad= { 2   for Non-target

4   for Lure

6   for Target   }

The trigger encoding is based on the following protocol:

Trigger at Onset of Cue Screen = (Memory Type Trigger Base)

Trigger at Onset of Stimulus Image =

(Memory Type Trigger Base)+(Image Category Trigger Pad)+(Match Type Trigger Pad)

Trigger at Offset of Stimulus Image =

(Memory Type Trigger Base)+(Image Category Trigger Pad)

The trigger during the fixation blocks is fixed to: Fixation Trigger = 2;

For descriptions of variables (column headers) to sync the tab-delimited E-Prime output for each run see Appendix IV: Task MEG E-Prime Key Variables.

Details of the tMEG Language Processing (Story-Math) task

This task was developed by Binder and colleagues (Binder *et al.* 2011). The task consists of two runs that each interleave 7 blocks of a story task and 15 blocks of a math task. The lengths of the blocks vary (average of approximately 30 seconds), but the task was designed so that the math task blocks match the length of the story task blocks, with some additional math trials at the end of the task to complete the ~7 minute run as needed.

The story blocks present participants with brief auditory stories (5-9 sentences) adapted from Aesop's fables, followed by a 2-alternative forced-choice question that asks participants about the topic of the story. The example provided in the original Binder paper (p. 1466) is "*For example, after a story about an eagle that saves a man who had done him a favor, participants were asked, "Was that about revenge or reciprocity?"* The math task also presents trials aurally and requires subjects to complete addition and subtraction problems. The trials present subjects with a series of arithmetic operations (e.g., "fourteen plus twelve"), followed by "equals" and then two choices (e.g., "twenty-nine or twenty-six"). Participants push a button to select either the first or the second answer. The math task is adaptive to try to maintain a similar level of difficulty across participants.

Stimulus Overview

In the Story-Math task participants listened either to an auditory sequence of sentences comprising Story narratives (Aesop fables) or matched-duration simple arithmetic Math problems (sequential addition/subtraction) (Binder *et al.*, 2011). Each of the stimulus sequences were followed by a 2-alternative forced choice question. In the Story case, the distinction of 2 alternative answers is qualitative (e.g. Was this story about loyalty or betrayal?) In the Math case, the distinction is quantitative (e.g. What is the correct answer: 3 or 7?).

Participants responded with the index and middle finger of the right hand on a button pad for the first or the second alternative, respectively. The 2 alternatives are presented in sequence and NOT at the same time (e.g. screen 1: What is the correct answer?, screen 2: 3, screen 3: or, screen4: 7). Each Story lasted about 30 seconds. Each Individual Math problem is much shorter, so multiple Math problems are put in sequence in order to form a block of matched duration of about 30 seconds. The number of operations within each Math problem was variable and was dynamically adjusted between blocks in order to adapt to the mathematical fluency of the subject. Typically, the number of operations varied between 3 and 5.

Block/Trial Overview

This experiment is performed in 2 runs with a small break between them. Run 1 consisted of 4 Story blocks and 4 Math blocks. Run 2 consisted of 5 Story blocks and 5 Math blocks.

Trigger channel sequence from one entire run of Story Math Task.

Trigger Overview

The trigger channel contains the trigger sequence from the E-Prime experimental protocol which runs on the stimulus presentation PC. The trigger channel for one Story/Math task run is shown, top right.

An example trigger sequence during a Story block is shown, bottom right. Before the beginning of a Story block, there is a first header trigger block. In this header trigger block, trigger values go from 0 to 128. Then there

Trigger channel sequence during a Story block.

is a following trigger going from 128 to a higher value and then returns back to 128. This excess value over 128, is the ID number of the current Story, multiplied by 2.

Then another trigger pulse follows within the header block. The excess value of this second trigger pulse over 128 represents the level of the current story multiplied by 2. The 'story level' is a quantitative representation of 4 different levels of difficulty for the participant to distinguish between the two different options presented at the end of each Story. Therefore, each story has 4 different pairs of options, one for each level of difficulty. For example, these 4 different levels can be:

"This story was about:"

Level 1: a bull or a mule?

Level 2: squeezing or kicking?

Level 3: aging or power?

Level 4: inexperience or deliberation?

At the end of the header trigger block the trigger returns to 0.

When the actual stimuli in the Story Block are about to start, and for its entire duration, the trigger baseline changes from 0 to 32, upon which all other triggers are superimposed.

Trigger value	Meaning	Duration
96	Start of story block	transient
32	Story block offset	for the duration of the block
24 + offset	Start of narrative	transient
26 + offset	Start of single story	transient
16 + offset	End of narrative	transient
14 + offset	Start of question ("That story was about ?"	transient
12 + offset	Start of correct option (e.g. "Loyalty")	transient
10 + offset	Start of incorrect options (e.g. "Betrayal")	transient
8 + offset	Start of "Or" between the 2 above options	transient
2 + offset	Start of response period	for the duration of the response period

An example trigger sequence during a Math block is shown below. Before the beginning of a Math problem block, there is a first header trigger block. In this header trigger block, trigger values go from 0 to 128. Then there is a following trigger going from 128 to a higher value and then returns back to 128. This excess value over 128, is the Difficulty Level of the current Math problem multiplied by 2.

Then there are 2 more trigger pulses within the header block superimposed on the base value of 128. The values of these 2 trigger pulses encode the Number of each math problem. The numerical amount over 128 of the value of the first trigger pulse, represents the value of the 4 Most Significant Bits of the Math problem number divided by 8. The numerical amount over 128 of the value of the second trigger pulse, represents the value of the 4 Least Significant Bits of the Math problem number multiplied by 2.
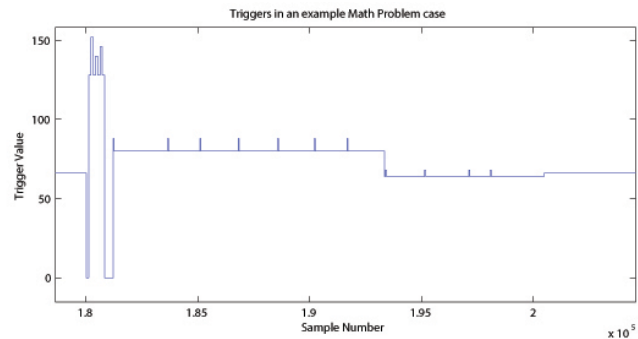
Trigger channel sequence during a Math problem block.

For example, if these 2 trigger pulses take values 142 and 156, then after the header block base value (128) is subtracted these values become 14 and 28, respectively. By multiplying the first by 8 and dividing the second by 2, these values become 112 and 14, respectively. The Math problem number is then derived by summing these numbers, so in this example the Math ID number is 112+14 =126.

The same Math problem number can exist for different difficulty levels. So the unique identification of a used problem requires both the difficulty level number and the Math problem number. So if in the example above the difficulty level was 10, then in order to uniquely identify this math problem one would need to know the pair of values [10, 126]. This is also reflected in the naming of the wav files used for the presentation of a specific math problem. So for the above example the wav file used by the E-Prime protocol would be called:

math-level10-126-Q.wav

At the end of the header trigger block the trigger returns to 0.

When the actual Math Problem Block starts, and for its entire duration, the trigger baseline changes from 0 to 64, upon which all other triggers are superimposed.

Trigger value	Meaning	Duration
112	Start of math block	transient
64	Math block offset	for the duration of the block
16 + offset	Math block instruction offset	for the duration of the instruction
8 + instruction offset	Presentation of number or arithmetic operand	transient
0 + offset	Math block question period offset in case the **first** option provided is the	for the duration of the question

Trigger value	Meaning	Duration
	correct answer	
4 + offset	Math block question period offset in case the **second** option provided is the correct answer	for the duration of the question
4 + question period offset	Presentation of question elements	transient
2 + offset	Start of response period	for the duration of the response period

The first part of the Math problem block is the narration of the arithmetic sequence. An example of an instruction would be: "3 + 2 + 6 + 4", where the onset of the presentation of each number and operand is marked by a transient trigger on top of the math block instruction offset. An example of the question would be: "equals 15 or 17", where the onset of the presentation of each of these elements is marked by a transient trigger on top of the question period offset.

The transient triggers last about 20 msec. At the end of the response period, the Math Block ends and the trigger returns to 0.

For descriptions of variables (column headers) to sync the tab-delimited E-Prime output for each run, see Appendix IV: Task MEG E-Prime Key Variables.

References for Language Processing Task: Reliable across subjects and robust activation (Binder *et al.* 2011).

Details of the tMEG Motor task

Sensory-motor processing is assessed using a task in which participants are presented with visual cues instructing the movement of either the right hand, left hand, right foot, or left foot. Movements are paced with a visual cue, which is presented in a blocked design. This task was adapted from the one developed by Buckner and colleagues (Buckner *et al.* 2011; Yeo *et al.* 2011). Participants are presented with visual cues that ask them to either tap their left or right index and thumb fingers or squeeze their left or right toes. Each block of a movement type lasts 12 seconds (10 movements), and is preceded by a 3 second cue. In each of the two runs, there are 32 blocks, with 16 of hand
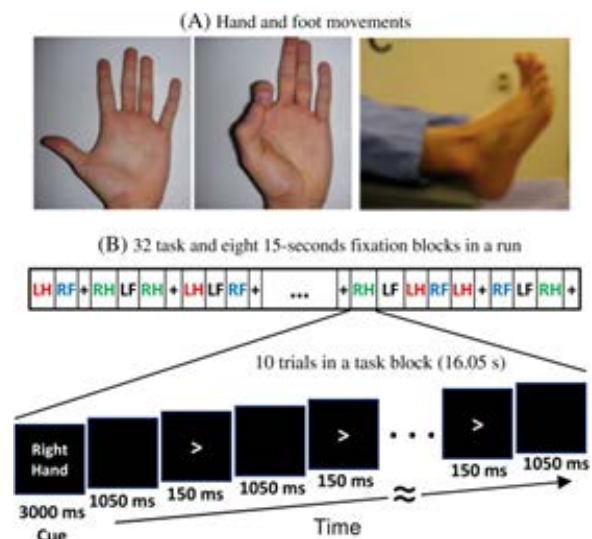
A. Hand and Foot movements during the Motor Task. B. Example sequence of stimuli in a block of Right Hand motor movements.

movements (8 right and 8 left), and 16 of foot movements (8 right and 8 left). In addition, there are nine 15-second fixation blocks per run. EMG signals were used for onset of event for hand and foot movement. EMG electrode stickers are applied as shown in <u>MEG hardware specifications and sensor locations</u> to the skin to the lateral superior surface of the foot on the extensor digitorum brevis muscle and near the medial malleolus, also the first dorsal interosseus muscle between thumb and index finger, and the styloid process of the ulna at the wrist.

Stimulus Overview

In the Motor task, participants executed a simple hand or foot movement. The limb and the side were instructed by a visual cue, and the timing of each movement was controlled by a pacing arrow presented on the center of the screen (A, right). The paradigm included movement and rest blocks (B, right).

Block/Trial Overview

Each block started with an instruction screen, indicating the side (left, right) and the limb (hand, foot) to be used by the subject in the current block. Then, 10 pacing stimuli were presented in sequence, each one instructing the participant to make a brisk movement. The pacing stimulus consisted of a small arrow in the center of the screen pointing to the side of the limb movement (left or right, above). The interval between consecutive stimuli was fixed to 1200 msec. The arrow stayed on the screen for 150 msec and for the remaining 1050 msec the screen was black.

In addition to the blocks of limb movements there were 10 interleaved resting blocks, each one of 15 sec duration. During these blocks the screen remained black. The last block was always a resting block after the last limb movement block.

The experiment was performed in 2 runs with a small break between them. The block/trial breakdown was identical in both runs. Each of the runs consisted of 42 blocks. 10 of these blocks were resting blocks, and there were 8 blocks of movement per motor effector. This yielded in total 80 movements per motor effector.

In addition to the recorded MEG channels, EMG activity was recorded from each limb. Also ECG and EOG electrodes were used to record heart- and eye movement-related electrophysiological activity.

Trigger Overview

The signal on the trigger channel consists of 2 superimposed trigger sequences. One from the Stimulus PC running the E-Prime protocol and one from a photodiode placed on the stimulus presentation screen. The trigger channel for one Motor task run is shown, top right.

The photodiode was activated whenever a cueing stimulus or pacing arrow was presented on the display. It was deactivated when the display was black. The trigger value for on is 255 and the trigger value for off is 0. The photodiode trigger sequence extracted from the trigger channel of one Motor task run is shown, middle right.

The E-Prime triggers contain the information about the experimental sequence. These triggers are superimposed on the photodiode triggers, in the following description it is assumed that the photodiode triggers have been subtracted from the trigger channel so that only the triggers from the E-Prime stimulation protocol remain. Such an E-Prime trigger sequence, extracted from the trigger channel is shown, bottom right.

Original Trigger channel sequence from one run of Motor Task. E-Prime and Photodiode triggers are superimposed.
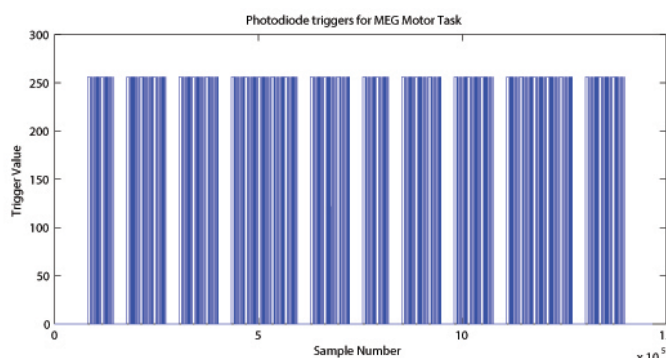
Photodiode trigger sequence extracted from the Trigger channel of one Motor Task run.

E-Prime trigger sequence extracted from the Trigger channel of one Motor Task run.

For descriptions of variables (column headers) to sync the tab-delimited E-Prime output for each run see Appendix IV: Task MEG E-Prime Key Variables. This task contains the following events, each of which is computed against the fixation baseline.

References for Motor Task: Localizer (Morioka *et al.* 1995; Bizzi *et al.* 2008; Buckner *et al.* 2011; Yeo *et al.* 2011).
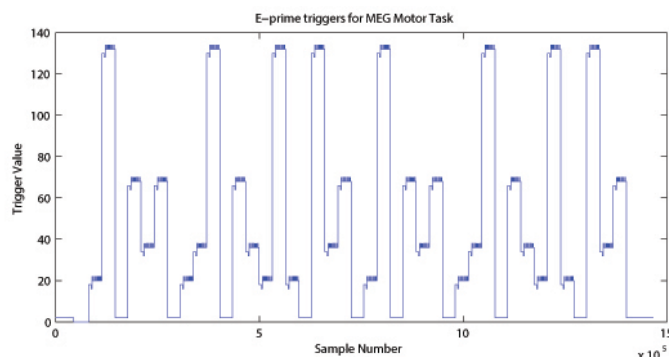
Trigger value	Meaning	Duration
2	resting block	entire duration of epoch
16	epoch offset left hand	entire duration of epoch
32	epoch offset left foot	entire duration of epoch
64	epoch offset right hand	entire duration of epoch
128	epoch offset right foot	entire duration of epoch
2 + epoch offset	cue stimulus onset	transient
4 + epoch offset	movement epoch offset	entire duration of movement epoch
2 + movement epoch offset	pacing arrow onset	transient

Conditions (Blocked)	Trigger codes
Left finger blocks	22
Right finger blocks	70
Left toe blocks	38
Right toe blocks	134

tMEG Channel-level Analysis Pipelines

The tMEG Channel-level analysis pipelines tmegpreproc, eravg, and tfavg perform the following analysis steps:

1) divide the data from each tMEG experiment into groups of trials that represent various events of interest

2) use the outputs of the previously run preprocessing pipeline steps to clean the tMEG data (now grouped as events of interest)

3) compute the average (across trials/runs) event related fields and calculate averaged condition contrasts, at the sensor (channel) level

4) compute the average Time-Frequency response at the sensor level for conditions and contrasts of conditions in the Task-MEG scans

How are different conditions and contrasts between conditions of Task scans analyzed?

Contrasts are used in the tMEG pipelines so that all different conditions and comparisons of conditions at different stages of analysis may be processed with the same pipelines.

A contrast is a MATLAB structure representing the settings for performing a specific type for a specific condition or between 2 different conditions of a single experiment.

The settings for the analysis of a single condition or comparison between 2 conditions can vary in terms of the time periods, the baseline or comparison operation and other settings. So in order to add the flexibility for varying such parameters for a single condition or for the comparison of different conditions, more than one contrast are used for each single condition or for the comparison of different conditions.

The table below defines and describes the fields of the Contrast structure for the eravg and tfavg pipelines.

Field Name	Print Mnemonic prefix	Description	Example	Possible Values	Default value
.pipeline		name of pipeline where contrast is to be run	'eravg'	'eravg','tfavg'	[ ]
.lockmode	LM-	describes data group from which trials for contrast have been selected	{'TIM'}	different datagroups within each task, currently: For Motort: 'TEMG','TFLA'. For Workmem: 'TIM', 'TRESP'. For StoryM: 'TEV','TRESP','BSENT', 'BU'	{ [ ] }
.mnemtrl	(The trial set mnemonics are appended after the lockmode mnemonic)	a mnemonic describing the specific sets of trials (conditions) in the contrast	{'0B' , '2B'} (i.e. 0-Back, 2-Back)	trial set mnemonics for each of the different conditions used in the contrasts A table presenting all the currently considered contrasts is presented further below in this page. Check there for all the available contrasts.	{ [ ] }
.operation	OP-	describes how metrics from the 2 different conditions will be compared – only meaningful where 2 conditions are compared within a contrast	'diff','rel', or 'relch'	'diff' (Difference),'rel' (relative:ratio),'relch' (relative change:ratio-1)	[ ]
.timeperiods	Not part of naming	describes time points or periods for which analysis will be performed If a one column vector (one column, many rows), analysis will be performed for each of these time points. If two columns exist, this is interpreted as time windows (each row a time window).	[-0.5 0 0 0.5 0.5 1]	no specific constraint This is only constrained by the data time information	{ [ ] }
.timedef	TD-	used when time periods are defined in windows (see above) It defines how the individual time points within each window	'concat'	'concat' (compute metric by all time points in windows concatanated for all trials) 'avg' (take the avg of all times within the window before computing the metric)	[ ]

Field Name	Print Mnemonic prefix	Description	Example	Possible Values	Default value
		will be processed during the analysis. For example, when computing the coherence between brain areas in time window [0 0.5], coherence can computed by estimating on spectral value for the entire window for each trial. Or the spectral coefficients can be computed for all time points within the window and they can all be used in the computation of coherence.		'sing' ('Pick a single value at the center of the window for each trial to perform analysis')	
.baseline	Not part of naming	contains the time window to be used as baseline in the contrast -- mostly used in single condition contrasts	{ [0 0.5] }	no specific constraint This is only constrained by the data time information	{ [ ] }
.baselinetype	BT-	defines by what operation the baseline will be applied	'diff'	'diff' (Difference) 'rel' (relative:ratio) 'relch' (relative change:ratio-1)	[ ]
.selection	Not part of naming	contains indices of the trials for the specific contrast conditions from the pool of all trials for the specific data group	{'[1:20]' ,... [21:40]'}	trial indices for each contrast are selected by the contrast function of each experiment	{ [ ] }
.description	Not part of naming	contains a very short description of the contrast in a form that can be used for printing in a report -- automatically produced by the contrast function of each experiment	'memberTypes : Left_Hand\n'		[ ]

Field Name	Print Mnemonic prefix	Description	Example	Possible Values	Default value
.mnemprint		Print Mnemonic constructed from all the above fields and used for the naming of the produced flies by the pipeline	'tfavg_[LM-TFLA-LH]_[BT-diff]'		[ ]

All the available conditions and contrasts between conditions for each Task paradigm are extracted using a dedicated MATLAB function named as:

contrast_$SCANMNEMONIC.m

where $SCANMNEMONIC can be "Wrkmem", 'Motort' or 'StoryM'.

These functions extract the list of all contrasts for all pipelines.

"tmegpreproc" pipeline

The hcp_tmegpreproc.m pipeline script is the first level of processing of the task specific analysis. There are 3 different Task paradigms in the HCP MEG protocol, very different from each other in content and structure. In each of these task experiments there are different events whose perception by and effect on the operation of the human brain are subjects of research.

For example in the motor task, one type of event is the onset of the flashing symbol instructing the subject to move his/her fingers or toes. The processing and perception of this visual instruction is a fast process well time locked to the onset of the visual stimulus. However, the onset of the movement is not as well time locked to the corresponding instruction. Most of this time jittering is introduced by anticipation of the instruction timing, by the relatively short interstimulus interval (ISI) translated into the motor system not being entirely settled before the next instruction and by the brain networks that make sense of a perceived instruction and translate it into a motor command. A more appropriate event for studying the onset of the movement is the onset of the EMG signal from the activated muscle. The above example simply demonstrates that within the same dataset one can cut trials relative to different events in order to study different aspects of brain's operation. The same holds true also for the working memory and story/math (language processing) task.

In order to accommodate this diversity in the analysis of task MEG data, the first step is to divide the data from each experiment into groups of trials according to the different events of interest. Each of these groups is then cleaned from artifacts according to the results from the preceding analysis of hcp_baddata.m and hcp_icaclass.m. Each group of trials is then saved with a characteristic mnemonic and used by all subsequent analysis.

The following table outlines the existing division into groups of trials locked on different events in each experiment.

Experiment	Data group mnemonic	Defining event	Trial duration	What can be studied
Motor Task	TFLA	onset of flashing cross that instructs subject to perform movement by hand or foot	-1.2 to 1.2 sec relative to onset of flashing cross	• brain networks of processing of the presented instruction • activation flow from visual to motor areas • brain networks of anticipation of the flashing cross onset
Motor Task	TEMG	onset of the EMG signal from hand or foot recorded muscles	-1.2 to 1.2 sec relative to onset of EMG signal	• brain networks employed during movement execution • functional mechanism and hierarchy between central and peripheral nervous system
Working Memory	TIM	onset of an image that the subject has to match or not with the target image	-1.5 to 2.5 sec relative to the onset of the image	• networks employed during image processing (faces vs tools) • networks of memory recall for different memory loads
Working Memory	TRESP	onset of button press by the subject	-1.5 to 1.5 sec relative to button press	• mechanism of memory recall and decision making just prior to the response mechanism of memory storage after response when the memory load is dynamic (2-back) as compared to static • motor network
Story/Math	TEV	onset of any task event during the question and option period in stories and math problems	-1.5 to 1.5 secs relative to story math event	• networks employed during math calculations • compare networks during first parts of a calculation and later parts with heavier load • networks during the onset of a story sentence relative to that of a math problem sentence • networks during the presentation of the correct answer as compared to the wrong one • activated areas when a number is presented as compared to when an operand is presented

Experiment	Data group mnemonic	Defining event	Trial duration	What can be studied
Story/Math	TRESP	onset of button press by the subject	-1.5 to 1.5 sec relative to button press	• motor network when a finger has already been picked for response • compare immediate responses with later responses which can mean uncertainty in the answer
Story/Math	BSENT	trials contain entire sentences For stories, this is a sentence during narration without the option sentence at the end of the story. For math, this is the sentence of the math problem excluding the option sentence at the end.	1 sec before onset of sentence and ends 1 sec after its offset	• compare how brain networks change at the later parts of the story as compared to earlier parts of a story • compare how brain networks change at the later math problems in a block as compared to earlier math problems in a block
Story/Math	BUN	trials containing entire Blocks of stimulus Units As stimulus unit is defined an entire story or an entire math problem including the option part.	1 sec before onset of unit and ends 1 sec after its offset	• compare the overall networks deployed during the processing of stories vs the processing of math problems The logic behind selecting entire units as trials comes from the fact that the total duration of stories is balanced with the total duration of math problems. However the number of math problems is significantly higher than the number of stories and the inverse holds for their duration. So actual comparison between stories and math can be performed at the block level and not at event level.

The definition of the parameters for all the groups of a specific scan are described in a configuration file named after the convention:

alltrialdefparameters_$SCANMNEMONIC.m, where $SCANMNEMONIC can be, for example, 'Wrkmem','StoryM' or 'Motort'. These files are located in the megconnectome/trial_functions/ folder.

These parameters are fed through ft_definetrial.m into the corresponding trial definition functions located in megconnectome/trial_functions/ and named in a corresponding manner as:

trialfun_$SCANMNEMONIC.m

This function returns the starting and ending samples as well as the offset of each trial in a data group for the specific scan.

This information is then used to extract the group trials data by ft_preprocessing.m

The trial definition function also extracts additional information about each trial, which is stored in the data structure in the trialinfo field.

All the above data groups whose mnemonic starts with T* have fixed trial length, which is relatively short. For this reason the trials that coincide even partially with bad data segments are completely removed. The data groups whose mnemonic starts with B* have variable trial length and the represent longer blocks of data. For these data sets the number of trials is small. They represent entire blocks of stimuli, i.e. an entire Story or an entire Math problem. As the number of trials is small, trials that coincide with bad data segments are not entirely removed. In order to preserve the trial definition structure inside the data, these segments are replaced by NaNs (Not a Number).

The tmegpreproc pipeline extracts and saves the trial definition and additional information for each group.

The data for each group based on the above information are extracted and cleaned for bad segments, bad channels and eye and heart related Independent Components.

The data are resampled to one fourth of the original sampling frequency in order to reduce the size of the dataset. The new sampling frequency is 508.63 Hz.

The clean data for each data group are saved in a separate file to be used by the later steps of the pipeline.

The output file $EXPERIMENTID_$SCANID_tmegpreproc_$GROUPMNEMONIC.mat contains the clean data split in trials for the specific data group in FieldTrip format.

The output file $EXPERIMENTID_$SCANID_tmegpreproc_trialinfo.mat contains a structure 'trlInfo' which contains the trial definition and additional information for each of the data groups for a specific scan. This trial information describes the clean set of trials within each data group.

The ASCII output file $EXPERIMENTID_$SCANID_tmegpreproc_rawtrialinfoQC.txt contains in MATLAB readable format general trial information about the raw data file. For example, it contains the number of trials in each of the main conditions, the number of trials in which there was a response, etc. For the EMG defined trials in the motor task, it contains how many trials were identified by the EMG. This information can be used as Quality Control in order to identify that the triggers have been read correctly and that the correct number of trials has been identified.

"eravg" pipeline

The purpose of this pipeline is the computation of the average of the event related fields and for calculation of condition differences (contrasts) at the sensor (channel) level.

This function inputs the data created by the hcp_tmegpreproc pipeline. This data contains the entire set of trials for a given data group, i.e. for Working Memory the data group TIM contains all the trials locked on the onset of a stimulus image. Within this data structure there is also information, in the form of a matrix, about the sequence and type of each trial.

The eravg pipeline uses the contrast function for the given experiment, e.g., Working Memory (see above, ***How are different conditions and contrasts between conditions of Task scans analyzed?***), to select the trials corresponding to the conditions that comprise a contrast. The contrast functions returns a default set of pre-defined experimental comparisons and information with respect to the pipeline(s) in which it is to be computed. The same comparison between conditions (e.g. 2-Back vs 0-Back) can be made with different specifications (e.g. stimulus-versus response-locked) at various levels of analysis (e.g. channel or source). For the above-mentioned example, the average of the event related fields at the sensor level can be computed for every original data sample in the trial, however for time frequency analysis a lower time resolution suffices, e.g. 25 ms steps. Additionally, the frequencies for TFR are defined, e.g. 1 to 100 Hz in steps of 1 Hz. For connectivity analysis, this high time and frequency resolution is again more resolution than is necessary for analysis, therefore the analysis is performed in a few time windows and in a few frequency bands. These are examples of the wide range of algorithmic specifications for the same comparison between conditions at the different analysis levels. To accommodate for this in the data analysis pipelines, a single contrast defines not only the selection of trials, but also the settings for the desired analyses in each of the pipelines.

The 'eravg' pipeline starts by loading the entire list of contrasts for a given experiment. Subsequently it selects the ones for the 'eravg' pipeline and feeds them into the core analysis function, named 'hcp_eravg_contrasts.m'.

The first task that this core function performs is to rearrange the list of contrasts so that those from the same data group are combined. This helps to avoid loading the same data files more than once, as this requires significant time. As the contrasts are clustered according to the data group, the data containing all the trials for one data group are only loaded once.

Once the data for all trials is loaded, the trials corresponding to the specific conditions of a contrast are selected and the event related average is computed and saved to disk.

At this time, the only operation implemented to compare different conditions in a contrast is to compute the difference in the mean. Furthermore, a contrast to the baseline is implemented for single conditions and for two-condition-contrasts, in which case the baseline correction is

applied prior to the computation of the average of each individual condition. The only type of baseline operation implemented at this time is subtraction, i.e. absolute difference.

Prior to baseline correction, the ERF is band pass filtered between 0.5 and 60 Hz.

The 'eravg' pipeline computes the event related average for both the axial magnetic field and for the magnitude of the estimated planar gradient. The planar gradient magnitude provides a more interpretable localization of the underlying sources, on the sensor level.

Additionally the 'eravg' pipeline has been designed to combine both scans from a given paradigm, e.g. "10-Motort" and "11-Motort", computing the average over all selected trials in both scans. The results are saved using a SCANID that omits the individual run number, e.g. "Motort". If one of the scans is not available, the pipeline will operate on a single scan.

For the eravg pipelines, the following conditions and contrasts between conditions are computed:

Motor Task: The following contrasts are computed for both **data groups TFLA** (Trials locked on the onset of the Flashing Cross that instructs the subject to perform a movement) and **TEMG** (Trials locked on the onset of the EMG signal).

Cond1	Cond2	Cond 1 Description	Cond 2 Description	Details	Naming of Output Files
LH		Left Hand		-1.2 to 1.2 sec with 508.63 Hz resolution baseline Period= [-0.3 0]; baseline Type = 'diff'	$EXPID_Motort_eravg_[LM-TFLA-LH]_[BT-diff]_[MODE-mag] (magnetic field) $EXPID_Motort_eravg_[LM-TFLA-LH]_[BT-diff]_[MODE-planar](planar gradient) (for data group TEMG , LM-TFLA in the naming is replaced by LM-TEMG)
RH		Right Hand		same as above	$EXPID_Motort_eravg_[LM-TFLA-RH]_[BT-diff]_[MODE-mag] (magnetic field) $EXPID_Motort_eravg_[LM-TFLA-RH]_[BT-diff]_[MODE-planar](planar gradient) (for data group TEMG , LM-TFLA in the naming is replaced by LM-TEMG)
LF		Left Foot		same as above	$EXPID_Motort_eravg_[LM-TFLA-LF]_[BT-diff]_[MODE-mag] (magnetic field) $EXPID_Motort_eravg_[LM-TFLA-LF]_[BT-diff]_[MODE-planar](planar gradient) (for data group TEMG , LM-TFLA in the naming is replaced by LM-TEMG)

Cond1	Cond2	Cond 1 Description	Cond 2 Description	Details	Naming of Output Files
RF		Right Foot		same as above	$EXPID_Motort_eravg_[LM-TFLA-RF]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_Motort_eravg_[LM-TFLA-RF]_[BT-diff]_[MODE-planar](planar gradient)
					(for data group TEMG , LM-TFLA in the naming is replaced by LM-TEMG)

The same contrasts are performed for the data group TEMG. For this group, LM-TFLA is replaced by LM-TEMG in the naming of the resultant files.

Working Memory: data group **TIM** (Trials locked on the onset of the image)

Cond1	Cond2	Cond 1 Description	Cond 2 Description	Details	Naming of Output Files
0B		0-Back Memory Trials		[-1.5 to 2.5] sec with 508.63 Hz resolution baseline Period=[-0.5 0]; baseline Type = 'diff'	$EXPID_Wrkmem_eravg_[LM-TIM-0B]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_Wrkmem_eravg_[LM-TIM-0B]_[BT-diff]_[MODE-planar](planar gradient)
2B		2-Back Memory Trials		same as above	$EXPID_Wrkmem_eravg_[LM-TIM-2B]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_Wrkmem_eravg_[LM-TIM-2B]_[BT-diff]_[MODE-planar] (planar gradient)
face		face image trials		same as above	$EXPID_Wrkmem_eravg_[LM-TIM-face]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_Wrkmem_eravg_[LM-TIM-face]_[BT-diff]_[MODE-planar] (planar gradient)
tool		tool image trials		same as above	$EXPID_Wrkmem_eravg_[LM-TIM-tool]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_Wrkmem_eravg_[LM-TIM-tool]_[BT-diff]_[MODE-planar] (planar gradient)
0B	**2B**	0-Back Memory Trials	2-Back Memory Trials	[-1.5 to 2.5] sec with 508.63 Hz resolution baseline Period=[-0.5 0]; baseline Type = 'diff' operation:'diff'	$EXPID_Wrkmem_eravg_[LM-TIM-0B-versus-2B]_[OP-diff]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_Wrkmem_eravg_[LM-TIM-0B-versus-2B]_[OP-diff]_[BT-diff]_[MODE-mag] (planar gradient)

Cond1	Cond2	Cond 1 Description	Cond 2 Description	Details	Naming of Output Files
face	**tool**	face image trials	tool image trials	same as above	$EXPID_Wrkmem_eravg_[LM-TIM-face-versus-tool]_[OP-diff]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_Wrkmem_eravg_[LM-TIM-face-versus-tool]_[OP-diff]_[BT-diff]_[MODE-planar] (planar gradient)

Working Memory: data group **TRESP** (Trials locked on the onset of the button press by the subject). Similar to the TIM data group above, with the only difference in the baseline interval used. In each of the eravg contrasts of the TRESP data group the baseline is set to interval [-0.75 to -0.25] seconds. For this group in the naming of the result files the part LM-TIM is replaced by LM-TRESP.

Story/Math: data group **TEV** (Onset of any task event during the question and option period in stories and math problems).

Cond1	Cond2	Cond 1 Description	Cond 2 Description	Details	Naming of Output Files
mathnumque		Onset of numbers in math questions		[-1.5 to 4] sec with 508.63 Hz resolution baseline Period=[-0.5 0]; baseline Type = 'diff'	$EXPID_StoryM_eravg_[LM-TEV-mathnumque]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_StoryM_eravg_[LM-TEV-mathnumque]_[BT-diff]_[MODE-planar] (planar gradient)
storsentnon		Onset of sentence in stories (not option part)		same as above	$EXPID_StoryM_eravg_[LM-TEV-storsentnon]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_StoryM_eravg_[LM-TEV-storsentnon]_[BT-diff]_[MODE-planar] (planar gradient)
mathsentnon		Onset of sentence in math question (not option part)		same as above	$EXPID_MEG_StoryM_eravg_[LM-TEV-mathsentnon]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_MEG_StoryM_eravg_[LM-TEV-mathsentnon]_[BT-diff]_[MODE-planar] (planar gradient)

Cond1	Cond2	Cond 1 Description	Cond 2 Description	Details	Naming of Output Files
mathnumopt		Onset of numbers in math options		same as above	$EXPID_StoryM_eravg_[LM-TEV-mathnumopt]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_StoryM_eravg_[LM-TEV-mathnumopt]_[BT-diff]_[MODE-planar] (planar gradient)
mathoper		Onset of operand (+ -) in math questions		same as above	$EXPID_StoryM_eravg_[LM-TEV-mathoper]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_StoryM_eravg_[LM-TEV-mathoper]_[BT-diff]_[MODE-planar] (planar gradient)
storsentnon	**mathsentnon**	Onset of sentence in stories (not option part)	Onset of sentence in math question (not option part)	[-1.5 to 4] sec with 508.63 Hz resolution baseline Period= [-0.5 0]; baseline Type = 'diff' operation :'diff'	$EXPID_StoryM_eravg_[LM-TEV-storsentnon-versus-mathsentnon]_[OP-diff]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_StoryM_eravg_[LM-TEV-storsentnon-versus-mathsentnon]_[OP-diff]_[BT-diff]_[MODE-planar] (planar gradient)
mathnumquelate	**mathnumqueearly**	Onset of the latter half of numbers in math questions	Onset of the earlier half of numbers in math questions	same as above	$EXPID_StoryM_eravg_[LM-TEV-mathnumquelate-versus-mathnumqueearly]_[OP-diff]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_StoryM_eravg_[LM-TEV-mathnumquelate-versus-mathnumqueearly]_[OP-diff]_[BT-diff]_[MODE-planar] (planar gradient)
mathnumoptcor	**mathnumoptwro**	Onset of correct option numbers in math options	Onset of wrong option numbers in math options	same as above	$EXPID_StoryM_eravg_[LM-TEV-mathnumoptcor-versus-mathnumoptwro]_[OP-diff]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_StoryM_eravg_[LM-TEV-mathnumoptcor-versus-mathnumoptwro]_[OP-diff]_[BT-diff]_[MODE-planar] (planar

Cond1	Cond2	Cond 1 Description	Cond 2 Description	Details	Naming of Output Files
					gradient)
storoptcor	**storoptwro**	Onset of correct option in story options	Onset of wrong option in story options	same as above	$EXPID_StoryM_eravg_[LM-TEV-storoptcor-versus-storoptwro]_[OP-diff]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_StoryM_eravg_[LM-TEV-storoptcor-versus-storoptwro]_[OP-diff]_[BT-diff]_[MODE-planar] (planar gradient)
mathnumque	**mathoper**	Onset of numbers in math questions	Onset of operand (+ -) in math questions	same as above	$EXPID_StoryM_eravg_[LM-TEV-mathnumque-versus-mathoper]_[OP-diff]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_StoryM_eravg_[LM-TEV-mathnumque-versus-mathoper]_[OP-diff]_[BT-diff]_[MODE-planar] (planar gradient)

Story/Math: data group **TRESP** (Onset of Button press in response).

Cond1	Cond2	Cond 1 Description	Cond 2 Description	Details	Naming of Output Files
all		all response trials		[-1.5 to 1.5] sec with 508.63 Hz resolution baseline Period= [-0.75 -0.25]; baseline Type = 'diff'	$EXPID_StoryM_eravg_[LM-TRESP-all]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_StoryM_eravg_[LM-TRESP-all]_[BT-diff]_[MODE-planar] (planar gradient)

"tfavg" pipeline

The purpose of this pipeline is the computation of the average Time-Frequency response at sensor level for a set of conditions and contrasts of conditions in the Task-MEG scans.

This function inputs the data created by the hcp_tmegpreproc pipeline. These data contain the entire set of trials for a given data group, i.e. for Working Memory the data group "TIM" contains all the trials locked on the onset of a stimulus image. Within this data structure there is also information, in the form of a matrix, about the sequence and type of each trial.

The tfavg uses the contrast function for the given experiment, e.g., Working Memory, to select the trials corresponding to the conditions that comprise a contrast. The contrast functions returns a default set of pre-defined experimental comparisons and information with respect to the pipeline(s) in which it is to be computed. The same comparison between conditions (e.g. 2-Back vs 0-Back) can be made with different specifications (e.g. stimulus versus response locked) at various levels of analysis (e.g. channel or source). For the above mentioned example, the average of the Time-Frequency response at the sensor level can be computed for every original data sample in the trial, however a lower time resolution suffices, e.g. 25 ms steps. Additionally the frequencies for TFR are defined in the range 1 to 100 Hz in steps of 1 Hz.

The 'tfavg' pipeline starts by loading the entire list of contrasts for a given experiment. Subsequently it selects the ones for the 'tfavg' pipeline and feeds them into the core analysis function, named 'hcp_tfavg_contrasts.m'.

 The first task that this core function performs is to rearrange the list of contrasts so that those from the same data group are clustered so that the data containing all the trials for one data group is only loaded once. This helps to avoid loading the same data files multiple times, as this requires significant time. When the data for all trials are loaded, the trials corresponding to the specific conditions of a contrast are selected and the event related average is computed and stored to disk.

At this time there are 2 operations implemented for comparison between different conditions: difference and ratio. Furthermore, a contrast to the baseline is implemented for single conditions and for two-condition contrasts, in which case the baseline correction is applied prior to the computation of the average of each individual condition. The only type of baseline operation implemented at this time is ratio.

The 'tfavg' pipeline computes the average Time-Frequency for both the axial magnetic field and for the magnitude of the estimated planar gradient. The planar gradient magnitude provides on the sensor level a more interpretable localization of the underlying sources.

Additionally the 'tfavg' pipeline has been designed to combine both scans from a given paradigm, e.g. "10-Motort" and "11-Motort", computing the average over all selected trials in both scans. The results are saved using a SCANID that omits the individual run number, e.g. "Motort". In case one of the scans is not available, the pipeline will operate on a single scan.

Time-frequency analysis is performed using the following settings:

TFR analysis settings		
Frequencies	f	1 to 100 Hz in steps of 1 Hz
Time Step	tstep	0.025 sec (time resolution)
Time Window	5./f	For each different frequency a different time window length is used for the computation of TFR at a time point. The length of the time window is set to 4 periods of the given frequency. For low frequencies this time window length

					becomes long so a threshold has been set to 500 msec as the longest window.
TFR method	'mtmconvol'				multitaper time-frequency transformation based on multiplication in the frequency domain
TFR taper type	'hanning'				Hanning window taper

For the tfavg pipelines the following conditions and contrasts between conditions are computed:

Motor Task: The following contrasts are computed for both data groups **TFLA** (Trials locked on the onset of the Flashing Cross that instructs the subject to perform movement) and **TEMG** (Trials locked on the onset of the EMG signal).

Cond1	Cond2	Cond1 Description	Cond2 Description	Details	Naming of Output Files
LH		Left Hand		Time points =[-1.2:0.025:1.2] sec; 100 frequencies=[1:1:100]; baseline Period=[-0.3 0]; baseline type:'diff'	$EXPID_Motort_tfavg_[LM-TLFA-LH]_[BT-diff]_[MODE-mag] (magnetic field) $EXPID_Motort_tfavg_[LM-TLFA-LH]_[BT-diff]_[MODE-planar] (planar gradient)
RH		Right Hand		same as above	$EXPID_Motort_tfavg_[LM-TLFA-RH]_[BT-diff]_[MODE-mag] (magnetic field) $EXPID_Motort_tfavg_[LM-TLFA-RH]_[BT-diff]_[MODE-planar] (planar gradient)
LF		Left Foot		same as above	$EXPID_Motort_tfavg_[LM-TLFA-LF]_[BT-diff]_[MODE-mag] (magnetic field) $EXPID_Motort_tfavg_[LM-TLFA-LF]_[BT-diff]_[MODE-planar] (planar gradient)
RF		Right Foot		same as above	$EXPID_Motort_tfavg_[LM-TLFA-RF]_[BT-diff]_[MODE-mag] (magnetic field) $EXPID_Motort_tfavg_[LM-TLFA-RF]_[BT-diff]_[MODE-planar] (planar gradient)
LH		Left Hand		Time points = [-1.2:0.025:1.2] sec; 100 frequencies=[1:1:100]; Connectivity Metric: emgcoh (Coherence with EMG signal)	$EXPID_Motort_tfavg_[LM-TLFA-LH]_[CM-emgcoh]_[MODE-mag] (magnetic field) $EXPID_Motort_tfavg_[LM-TLFA-LH]_[CM-emgcoh]_[MODE-planar]

					(planar gradient)
RH	Right Hand	same as above			$EXPID_Motort_tfavg_[LM-TLFA-RH]_[CM-emgcoh]_[MODE-mag] (magnetic field)
					$EXPID_Motort_tfavg_[LM-TLFA-RH]_[CM-emgcoh]_[MODE-planar] (planar gradient)
LF	Left Foot	same as above			$EXPID_Motort_tfavg_[LM-TLFA-LF]_[CM-emgcoh]_[MODE-mag] (magnetic field)
					$EXPID_Motort_tfavg_[LM-TLFALF]_[CM-emgcoh]_[MODE-planar] (planar gradient)
RF	Right Foot	same as above			$EXPID_Motort_tfavg_[LM-TLFA-RF]_[CM-emgcoh]_[MODE-mag] (magnetic field)
					$EXPID_Motort_tfavg_[LM-TLFA-RF]_[CM-emgcoh]_[MODE-planar] (planar gradient)

The same contrasts are performed for the data group **TEMG**. For this group in the naming of the result files the part LM-TFLA is replaced by LM-TEMG.

Working Memory: data group: **TIM** (Trials locked on the onset of the image)

Cond1	Cond2	Cond1 Description	Cond1 Description	Details	Naming of Output Files
0B		0-Back Memory Trials		100 frequencies=[1:1:100]; Time points = [-1:0.025:2] sec baseline Period=[-0.5 0]; baseline type:'diff'	$EXPID_Wrkmem_tfavg_[LM-TIM-0B]_[MODE-mag] (magnetic field)
					$EXPID_Wrkmem_tfavg_[LM-TIM-0B]_[MODE-planar] (planar gradient)
2B		2-Back Memory Trials		same as above	$EXPID_Wrkmem_tfavg_[LM-TIM-2B]_[MODE-mag] (magnetic field)
					$EXPID_Wrkmem_tfavg_[LM-TIM-2B]_[MODE-planar] (planar gradient)
face		face image trials		same as above	$EXPID_Wrkmem_tfavg_[LM-TIM-face]_[MODE-mag] (magnetic field)
					$EXPID_Wrkmem_tfavg_[LM-TIM-face]_[MODE-planar] (planar gradient)

Cond1	Cond2	Cond1 Description	Cond1 Description	Details	Naming of Output Files
tool		tool image trials		same as above	$EXPID_Wrkmem_tfavg_[LM-TIM-tool]_[MODE-mag] (magnetic field)
					$EXPID_Wrkmem_tfavg_[LM-TIM-tool]_[MODE-planar] (planar gradient)
0B	**2B**	0-Back Memory Trials	2-Back Memory Trials	100 frequencies=[1:1:100]; Time points = [-1:0.025:2] sec operation:'diff'	$EXPID_Wrkmem_tfavg_[LM-TIM-0B-versus-2B]_[OP-diff]_[MODE-mag] (magnetic field)
					$EXPID_Wrkmem_tfavg_[LM-TIM-0B-versus-2B]_[OP-diff]_[MODE-planar] (planar gradient)
face	**tool**	face image trials	tool image trials	same as above	$EXPID_Wrkmem_tfavg_[LM-TIM-face-versus-tool]_[OP-diff]_[MODE-mag] (magnetic field)
					$EXPID_Wrkmem_tfavg_[LM-TIM-face-versus-tool]_[OP-diff]_[MODE-planar] (planar gradient)

Working Memory: data group **TRESP** (Trials locked on the onset of the button press by the subject)

Similar to the TIM data group above, with only difference in the time interval used. In each of the tfavg contrasts of the TRESP data group there are 121 timepoints in the interval [-1.25 to 1.75] seconds with time step of 25 mseconds. For this group in the naming of the result files the part LM-TIM is replaced by LM-TRESP.

Story/Math: data group: **TEV** (Onset of any task event during the question and option period in stories and math problems.)

Cond1	Cond2	Cond 1 Description	Cond 2 Description	Details	Naming of Output Files
mathnumque		Onset of numbers in math questions		100 frequencies= [1:1:100]; Time points = [-1:0.025:3.5] sec baseline Period=[-0.5 0]; baseline Type = 'diff'	$EXPID_StoryM_eravg_[LM-TEV-mathnumque]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_StoryM_eravg_[LM-TEV-mathnumque]_[BT-diff]_[MODE-planar] (planar gradient)

Cond1	Cond2	Cond 1 Description	Cond 2 Description	Details	Naming of Output Files
storsentnon		Onset of sentence in stories (not option part)		same as above	$EXPID_StoryM_tfavg_[LM-TEV-storsentnon]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_StoryM_tfavg_[LM-TEV-storsentnon]_[BT-diff]_[MODE-planar] (planar gradient)
mathsentnon		Onset of sentence in math question (not option part)		same as above	$EXPID_MEG_StoryM_tfavg_[LM-TEV-mathsentnon]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_MEG_StoryM_tfavg_[LM-TEV-mathsentnon]_[BT-diff]_[MODE-planar] (planar gradient)
mathnumopt		Onset of numbers in math options		same as above	$EXPID_StoryM_tfavg_[LM-TEV-mathnumopt]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_StoryM_tfavg_[LM-TEV-mathnumopt]_[BT-diff]_[MODE-planar] (planar gradient)
mathoper		Onset of operand (+ -) in math questions		same as above	$EXPID_StoryM_tfavg_[LM-TEV-mathoper]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_StoryM_tfavg_[LM-TEV-mathoper]_[BT-diff]_[MODE-planar] (planar gradient)
storsentnon	**mathsentnon**	Onset of sentence in stories (not option part)	Onset of sentence in math question (not option part)	100 frequencies= [1:1:100]; Time points = [-1:0.025:3.5] sec operation :'diff'	$EXPID_StoryM_tfavg_[LM-TEV-storsentnon-versus-mathsentnon]_[OP-diff]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_StoryM_tfavg_[LM-TEV-storsentnon-versus-mathsentnon]_[OP-diff]_[BT-diff]_[MODE-

Cond1	Cond2	Cond 1 Description	Cond 2 Description	Details	Naming of Output Files
					planar] (planar gradient)
mathnumquelate	**mathnumqueearly**	Onset of the latter half of numbers in math questions	Onset of the earlier half of numbers in math questions	same as above	$EXPID_StoryM_tfavg_[LM-TEV-mathnumquelate-versus-mathnumqueearly]_[OP-diff]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_StoryM_tfavg_[LM-TEV-mathnumquelate-versus-mathnumqueearly]_[OP-diff]_[BT-diff]_[MODE-planar] (planar gradient)
mathnumoptcor	**mathnumoptwro**	Onset of correct option numbers in math options	Onset of wrong option numbers in math options	same as above	$EXPID_StoryM_tfavg_[LM-TEV-mathnumoptcor-versus-mathnumoptwro]_[OP-diff]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_StoryM_tfavg_[LM-TEV-mathnumoptcor-versus-mathnumoptwro]_[OP-diff]_[BT-diff]_[MODE-planar] (planar gradient)
storoptcor	**storoptwro**	Onset of correct option in story options	Onset of wrong option in story options	same as above	$EXPID_StoryM_tfavg_[LM-TEV-storoptcor-versus-storoptwro]_[OP-diff]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_StoryM_tfavg_[LM-TEV-storoptcor-versus-storoptwro]_[OP-diff]_[BT-diff]_[MODE-planar] (planar gradient)
mathnumque	**mathoper**	Onset of numbers in math questions	Onset of operand (+ -) in math questions	same as above	$EXPID_StoryM_tfavg_[LM-TEV-mathnumque-versus-mathoper]_[OP-diff]_[BT-diff]_[MODE-mag] (magnetic field)
					$EXPID_StoryM_tfavg_[LM-TEV-mathnumque-versus-mathoper]_[OP-diff]_[BT-diff]_[MODE-planar] (planar gradient)

Story/Math: data group TRESP (Onset of Button press in response.)

Cond1	Cond2	Cond 1 Description	Cond 2 Description	Details	Naming of Output Files
all		all response trials		100 frequencies=[1:1:100]; Time points = [-1.25:0.025:1.25] sec baseline Period= [-0.75 -0.25]; baseline Type = 'diff'	$EXPID_StoryM_tfavg_[LM-TRESP-all]_[BT-diff]_[MODE-mag] (magnetic field) $EXPID_StoryM_tfavg_[LM-TRESP-all]_[BT-diff]_[MODE-planar] (planar gradient)

How was the E-Prime trigger sequence reproduced for the scans in which no parallel port triggers were recorded?

In the Working Memory and Motor task scans, the signal in the TRIGGER channel consists of 2 superimposed bit patterns. One from the stimulus presentation computer that is running the E-Prime protocol, the other is from a photodiode placed in a corner on the stimulus presentation screen.

The E-Prime triggers encode all the events that are of interest within a scan. The photodiode triggers contain a subset of these events, corresponding mainly to the onset of images on the screen. For the Working Memory scans, the photodiode is on when a stimulus image is presented and is off otherwise. For the Motor Task scans, the photodiode is on when the Cue Screen (e.g. Left Hand) or the pacing flashing cross screen for each movement are presented, and the photodiode is off otherwise.

The photodiode triggers provide more accurate timing than the E-Prime triggers, because they measure the instance of the actual stimulus onset on the screen and do not represent the time at which the display command was sent by the software. However, the value of the photodiode trigger can take only two values: 255 for ON and 0 for OFF. In contrast, the E-Prime triggers values are diverse and encode multiple aspects of the experimental sequence, such as stimulus content and task demands.

As a consequence of the two superimposed triggers, the event code sequence is extracted based on the E-Prime triggers, but its timing is extracted from the corresponding photodiode triggers.

Due to hardware problems, for some subjects during the Working Memory and Motor tasks, only the photodiode triggers were recorded in the TRIGGER channel in the raw MEG data file. This happened for the following subjects:

158035
189349
191841

250427
352132
352738
500222

The E-Prime trigger information is also present in the E-Prime log files that are created during stimulus presentation. These files describe the sequence of events with time stamps referenced to the beginning of the stimulation script run. These time stamps are represented with 1 Khz resolution. Based on these timestamps the parallel port trigger sequence can be reproduced.

As the start of the E-Prime stimulation script does not correspond precisely to the start of the MEG acquisition, the log file timing needs to be aligned to the MEG data file TRIGGER channel by means of the photodiode triggers.

First the sequence of events from the E-Prime log file was extracted. The trigger events for which the corresponding photodiode events exist were identified. Then the E-Prime timestamps and the photodiode trigger timestamps for these events were referenced to the very first of these events, so that the first event time stamp was 0 for both. Subsequently the slope between the E-Prime and 4D timestamps for these events was computed. Based on this slope the E-Prime event sequence timing was transformed to the 4D sampling resolution of 2KHz. Then based on their common events the e-prime and photodiode sequences were aligned in time and an accurate mapping from E-Prime timestamps to 4D timestamps was derived.

The reconstructed E-Prime trigger sequence (expressed in MEG timestamps) was superimposed on the existing photodiode trigger sequence and then written in the TRIGGER channel of a new version of the raw data file. This version is the one available for download.

Directory structure for Channel-level processed MEG data

All channel-level processed MEG data for each subject should unpack to the MEG/ directory under the <SubjectID> directory in a subdirectory for each type of experiment. A list of file names in each directory and subdirectory is provided in Appendix III: File Names and Directory Structure for Unprocessed, Anatomical Models, and Channel-Level Processed MEG Data, Section C.

Under each of the experimental conditions, the directory structure represents the analysis pipelines that have been executed on the data.

For the resting state dataset, the sequence of pipelines consists of datacheck->baddata->icaclass->rmegpreproc->powavg.

For the three task datasets, the sequence of pipelines consists of datacheck->baddata->icaclass->tmegpreproc->eravg&tfavg.

For the empty-room and subject noise datasets, the only applicable pipeline is datacheck. The noise datacheck pipeline results do not comprise a separate package but are included in the packages for the unprocessed empty room data.

Channel-level processed MEG data for exemplar subject 012345 unpacks to the following directory structure:

```
012345/MEG
        Rnoise
        Pnoise
        Restin
        Wrkmem
        StoryM
        Motort
```

Under each of the experimental conditions, the directory structure represents the analysis pipelines that have been executed on the data.

Channel-level processed MEG data for exemplar subject 012345 unpacks to the directory structure that is listed below for each of the pipelines. Each of the *.txt, *.mat and *.png data files that are listed below is accompanied with a similarly named *.xml file in the provenance directory, which details the version of the software used to produce the results. These XML files are not listed below, but are present in the release packages.

Datacheck

The results of the Datacheck pipeline for exemplar subject 012345 unpack to the following directory structure:

MEG/Rnoise/datacheck/
        012345_MEG_1-Rnoise_datacheck_info.txt

        figures/
           012345_MEG_1-Rnoise_datacheck_jumps.png
           012345_MEG_1-Rnoise_datacheck_MEG_lowfreq_power.png
           012345_MEG_1-Rnoise_datacheck_MEG_powerline_noise.png
           012345_MEG_1-Rnoise_datacheck_MEG_powspctrm.png
           012345_MEG_1-Rnoise_datacheck_MEGREF_powspctrm.png
           012345_MEG_1-Rnoise_datacheck_neighb_correlation.png
           012345_MEG_1-Rnoise_datacheck_triggers.png

There are similar results for the other scans, each with the corresponding scan type and number in the directory and in the file names:

MEG/Pnoise/datacheck/
MEG/Restin/datacheck/
MEG/Wrkmem/datacheck/
MEG/StoryM/datacheck/
MEG/Motort/datacheck/

Baddata

The results of Baddata pipeline for exemplar subject 012345 unpack to the following directory structure:

MEG/Restin/baddata/
        012345_MEG_3-Restin_baddata_badchannels.txt
        012345_MEG_3-Restin_baddata_badsegments.txt
        012345_MEG_3-Restin_baddata_manual_badchannels.txt
        012345_MEG_3-Restin_baddata_manual_badsegments.txt
        012345_MEG_4-Restin_baddata_badchannels.txt
        etc.

        figures/
           012345_MEG_3-Restin_baddata_badchan_cor_scatter.png
           012345_MEG_3-Restin_baddata_badchan_cor_topo3D.png
           012345_MEG_3-Restin_baddata_badchan_cor_topo.png
           012345_MEG_3-Restin_baddata_badchan_std_scatter.png
           012345_MEG_3-Restin_baddata_badchan_std_topo.png
           012345_MEG_3-Restin_baddata_icaqc_badchannel_A88.png
           012345_MEG_3-Restin_baddata_icaqc_badchannel_A246.png

```
                 etc.

          012345_MEG_3-Restin_baddata_icaqc_badsegment_1.png
          012345_MEG_3-Restin_baddata_icaqc_badsegment_2.png
          012345_MEG_3-Restin_baddata_icaqc_badsegment_3.png
                 etc.

          012345_MEG_3-Restin_baddata_icaqc_results_1.png
          012345_MEG_3-Restin_baddata_icaqc_results_2.png
          012345_MEG_3-Restin_baddata_icaqc_results_3.png
                 etc.
```

There are similar results for the other scans, each with the corresponding scan type and number in the directory and in the file names:

```
MEG/Wrkmem/baddata/
MEG/StoryM/baddata/
MEG/Motort/baddata/
```

Icaclass and Icaclass_qc

The results of the Icaclass and Icaclass_qc pipelines for exemplar subject 012345 unpack to the following directory structure:

```
MEG/Restin/icaclass/
          012345_MEG_3-Restin_icaclass.mat
          012345_MEG_3-Restin_icaclass.txt
          012345_MEG_3-Restin_icaclass_vs.txt
          012345_MEG_3-Restin_icaclass.mat
          012345_MEG_3-Restin_icaclass.txt
          012345_MEG_3-Restin_icaclass_vs.txt
          012345_MEG_4-Restin_icaclass.mat
          etc.

      figures/
          012345_MEG_3-Restin_icaclass_refch.png
          012345_MEG_3-Restin_icaclass_1.png
          012345_MEG_3-Restin_icaclass_2.png
          012345_MEG_3-Restin_icaclass_3.png
              etc.

          012345_MEG_3-Restin_icaclass_vs_1.png
          012345_MEG_3-Restin_icaclass_vs_2.png
          012345_MEG_3-Restin_icaclass_vs_3.png
              etc.
```

There are similar results for the other scans, each with the corresponding scan type and number in the directory and in the file names:

MEG/Wrkmem/icaclass/
MEG/StoryM/icaclass/
MEG/Motort/icaclass/

Rmegpreproc

The results of the Rmegpreproc pipeline for exemplar subject 012345 unpack to the following directory structure:

MEG/Restin/rmegpreproc/
      012345_MEG_3-Restin_rmegpreproc.mat
      012345_MEG_4-Restin_rmegpreproc.mat
      012345_MEG_5-Restin_rmegpreproc.mat

Powavg

The results of the Powavg pipeline for exemplar subject 012345 unpack to the following directory structure:

MEG/Restin/powavg/
      012345_MEG_3-Restin_powavg.mat
      012345_MEG_4-Restin_powavg.mat
      012345_MEG_5-Restin_powavg.mat

      figures/
         012345_MEG_3-Restin_powavg_multiplot.png
         012345_MEG_3-Restin_powavg_singleplot.png
         012345_MEG_4-Restin_powavg_singleplot.png
         012345_MEG_5-Restin_powavg_multiplot.png
         012345_MEG_5-Restin_powavg_singleplot.png

Eravg

The results of the Eravg pipeline for exemplar subject 012345 unpack to the following directory structure for the Working Memory task:

MEG/Wrkmem/eravg/
      012345_MEG_Wrkmem_eravg_[LM-TIM-0B]_[BT-diff]_[MODE-mag].mat
      012345_MEG_Wrkmem_eravg_[LM-TIM-0B]_[BT-diff]_[MODE-planar].mat
      012345_MEG_Wrkmem_eravg_[LM-TIM-0B-versus-2B]_[OP-diff]_[BT-diff]_[MODE-mag].mat
      012345_MEG_Wrkmem_eravg_[LM-TIM-0B-versus-2B]_[OP-diff]_[BT-diff]_[MODE-planar].mat
      012345_MEG_Wrkmem_eravg_[LM-TIM-2B]_[BT-diff]_[MODE-mag].mat
      012345_MEG_Wrkmem_eravg_[LM-TIM-2B]_[BT-diff]_[MODE-planar].mat
      012345_MEG_Wrkmem_eravg_[LM-TIM-face]_[BT-diff]_[MODE-mag].mat
      012345_MEG_Wrkmem_eravg_[LM-TIM-face]_[BT-diff]_[MODE-planar].mat
      012345_MEG_Wrkmem_eravg_[LM-TIM-face-versus-tool]_[OP-diff]_[BT-diff]_[MODE-mag].mat
      012345_MEG_Wrkmem_eravg_[LM-TIM-face-versus-tool]_[OP-diff]_[BT-diff]_[MODE-planar].mat
      012345_MEG_Wrkmem_eravg_[LM-TIM-tool]_[BT-diff]_[MODE-mag].mat

012345_MEG_Wrkmem_eravg_[LM-TIM-tool]_[BT-diff]_[MODE-planar].mat
012345_MEG_Wrkmem_eravg_[LM-TRESP-0B]_[BT-diff]_[MODE-mag].mat
012345_MEG_Wrkmem_eravg_[LM-TRESP-0B]_[BT-diff]_[MODE-planar].mat
012345_MEG_Wrkmem_eravg_[LM-TRESP-0B-versus-2B]_[OP-diff]_[BT-diff]_[MODE-mag].mat
012345_MEG_Wrkmem_eravg_[LM-TRESP-0B-versus-2B]_[OP-diff]_[BT-diff]_[MODE-planar].mat
012345_MEG_Wrkmem_eravg_[LM-TRESP-2B]_[BT-diff]_[MODE-mag].mat
012345_MEG_Wrkmem_eravg_[LM-TRESP-2B]_[BT-diff]_[MODE-planar].mat
012345_MEG_Wrkmem_eravg_[LM-TRESP-face]_[BT-diff]_[MODE-mag].mat
012345_MEG_Wrkmem_eravg_[LM-TRESP-face]_[BT-diff]_[MODE-planar].mat
012345_MEG_Wrkmem_eravg_[LM-TRESP-face-versus-tool]_[OP-diff]_[BT-diff]_[MODE-mag].mat
012345_MEG_Wrkmem_eravg_[LM-TRESP-face-versus-tool]_[OP-diff]_[BT-diff]_[MODE-planar].mat
012345_MEG_Wrkmem_eravg_[LM-TRESP-tool]_[BT-diff]_[MODE-mag].mat
012345_MEG_Wrkmem_eravg_[LM-TRESP-tool]_[BT-diff]_[MODE-planar].mat

figures/
    012345_MEG_Wrkmem_eravg_[LM-TIM-0B]_[BT-diff]_[MODE-mag]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TIM-0B]_[BT-diff]_[MODE-planar]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TIM-0B-versus-2B]_[OP-diff]_[BT-diff]_[MODE-mag]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TIM-0B-versus-2B]_[OP-diff]_[BT-diff]_[MODE-planar]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TIM-2B]_[BT-diff]_[MODE-mag]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TIM-2B]_[BT-diff]_[MODE-planar]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TIM-face]_[BT-diff]_[MODE-mag]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TIM-face]_[BT-diff]_[MODE-planar]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TIM-face-versus-tool]_[OP-diff]_[BT-diff]_[MODE-mag]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TIM-face-versus-tool]_[OP-diff]_[BT-
    012345_MEG_Wrkmem_eravg_[LM-TIM-tool]_[BT-diff]_[MODE-mag]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TIM-tool]_[BT-diff]_[MODE-planar]_plot.pn
    012345_MEG_Wrkmem_eravg_[LM-TRESP-0B]_[BT-diff]_[MODE-mag]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TRESP-0B]_[BT-diff]_[MODE-planar]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TRESP-0B-versus-2B]_[OP-diff]_[BT-diff]_[MODE-mag]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TRESP-0B-versus-2B]_[OP-diff]_[BT-diff]_[MODE-planar]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TRESP-2B]_[BT-diff]_[MODE-mag]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TRESP-2B]_[BT-diff]_[MODE-planar]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TRESP-face]_[BT-diff]_[MODE-mag]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TRESP-face]_[BT-diff]_[MODE-planar]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TRESP-face-versus-tool]_[OP-diff]_[BT-diff]_[MODE-
        mag]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TRESP-face-versus-tool]_[OP-diff]_[BT-diff]_[MODE-
        planar]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TRESP-tool]_[BT-diff]_[MODE-mag]_plot.png
    012345_MEG_Wrkmem_eravg_[LM-TRESP-tool]_[BT-diff]_[MODE-planar]_plot.png

The results of the Eravg pipeline (with the appropriate event-related condition and contrast detail differences) for the Language Processing (Story-Math) and Motor tasks similarly unpack to the following directories:

MEG/StoryM/eravg/
        figures/

MEG/Motort/eravg/

figures/

Tfavg

The results of the Tfavg pipeline for exemplar subject 012345 unpack to the following directory structure:

MEG/Wrkmem/tfavg/
    012345_MEG_Wrkmem_tfavg_[LM-TIM-0B]_[MODE-planar].mat
    012345_MEG_Wrkmem_tfavg_[LM-TIM-2B]_[MODE-planar].mat
    012345_MEG_Wrkmem_tfavg_[LM-TRESP-0B]_[MODE-planar].mat
    012345_MEG_Wrkmem_tfavg_[LM-TRESP-2B]_[MODE-planar].mat
    012345_MEG_Wrkmem_tfavg_[LM-TIM-0B]_[MODE-mag].mat
    012345_MEG_Wrkmem_tfavg_[LM-TIM-0B-versus-2B]_[OP-diff]_[MODE-mag].mat
    012345_MEG_Wrkmem_tfavg_[LM-TIM-0B-versus-2B]_[OP-diff]_[MODE-planar].mat
    012345_MEG_Wrkmem_tfavg_[LM-TIM-2B]_[MODE-mag].mat
    012345_MEG_Wrkmem_tfavg_[LM-TIM-face]_[MODE-mag].mat
    012345_MEG_Wrkmem_tfavg_[LM-TIM-face]_[MODE-planar].mat
    012345_MEG_Wrkmem_tfavg_[LM-TIM-face-versus-tool]_[OP-diff]_[MODE-mag].mat
    012345_MEG_Wrkmem_tfavg_[LM-TIM-face-versus-tool]_[OP-diff]_[MODE-planar].mat
    012345_MEG_Wrkmem_tfavg_[LM-TIM-tool]_[MODE-mag].mat
    012345_MEG_Wrkmem_tfavg_[LM-TIM-tool]_[MODE-planar].mat
    012345_MEG_Wrkmem_tfavg_[LM-TRESP-0B]_[MODE-mag].mat
    012345_MEG_Wrkmem_tfavg_[LM-TRESP-0B-versus-2B]_[OP-diff]_[MODE-mag].mat
    012345_MEG_Wrkmem_tfavg_[LM-TRESP-0B-versus-2B]_[OP-diff]_[MODE-planar].mat
    012345_MEG_Wrkmem_tfavg_[LM-TRESP-2B]_[MODE-mag].mat
    012345_MEG_Wrkmem_tfavg_[LM-TRESP-face]_[MODE-mag].mat
    012345_MEG_Wrkmem_tfavg_[LM-TRESP-face]_[MODE-planar].mat
    012345_MEG_Wrkmem_tfavg_[LM-TRESP-face-versus-tool]_[OP-diff]_[MODE-mag].mat
    012345_MEG_Wrkmem_tfavg_[LM-TRESP-face-versus-tool]_[OP-diff]_[MODE-planar].mat
    012345_MEG_Wrkmem_tfavg_[LM-TRESP-tool]_[MODE-mag].mat
    012345_MEG_Wrkmem_tfavg_[LM-TRESP-tool]_[MODE-planar].mat

    figures/
        012345_MEG_Wrkmem_tfavg_[LM-TIM-0B]_[MODE-planar]_plot.png
        012345_MEG_Wrkmem_tfavg_[LM-TIM-2B]_[MODE-planar]_plot.png
        012345_MEG_Wrkmem_tfavg_[LM-TRESP-0B]_[MODE-planar]_plot.png
        012345_MEG_Wrkmem_tfavg_[LM-TRESP-2B]_[MODE-planar]_plot.png
        012345_MEG_Wrkmem_tfavg_[LM-TIM-0B]_[MODE-mag]_plot.png
        012345_MEG_Wrkmem_tfavg_[LM-TIM-0B-versus-2B]_[OP-diff]_[MODE-mag]_plot.png
        012345_MEG_Wrkmem_tfavg_[LM-TIM-0B-versus-2B]_[OP-diff]_[MODE-planar]_plot.png
        012345_MEG_Wrkmem_tfavg_[LM-TIM-2B]_[MODE-mag]_plot.png
        012345_MEG_Wrkmem_tfavg_[LM-TIM-face]_[MODE-mag]_plot.png
        012345_MEG_Wrkmem_tfavg_[LM-TIM-face]_[MODE-planar]_plot.png
        012345_MEG_Wrkmem_tfavg_[LM-TIM-face-versus-tool]_[OP-diff]_[MODE-mag]_plot.png
        012345_MEG_Wrkmem_tfavg_[LM-TIM-face-versus-tool]_[OP-diff]_[MODE-planar]_plot.png
        012345_MEG_Wrkmem_tfavg_[LM-TIM-tool]_[MODE-mag]_plot.png
        012345_MEG_Wrkmem_tfavg_[LM-TIM-tool]_[MODE-planar]_plot.png
        012345_MEG_Wrkmem_tfavg_[LM-TRESP-0B]_[MODE-mag]_plot.png
        012345_MEG_Wrkmem_tfavg_[LM-TRESP-0B-versus-2B]_[OP-diff]_[MODE-mag]_plot.png

012345_MEG_Wrkmem_tfavg_[LM-TRESP-0B-versus-2B]_[OP-diff]_[MODE-planar]_plot.png
012345_MEG_Wrkmem_tfavg_[LM-TRESP-2B]_[MODE-mag]_plot.png
012345_MEG_Wrkmem_tfavg_[LM-TRESP-face]_[MODE-mag]_plot.png
012345_MEG_Wrkmem_tfavg_[LM-TRESP-face]_[MODE-planar]_plot.png
012345_MEG_Wrkmem_tfavg_[LM-TRESP-face-versus-tool]_[OP-diff]_[MODE-mag]_plot.png
012345_MEG_Wrkmem_tfavg_[LM-TRESP-face-versus-tool]_[OP-diff]_[MODE-planar]_plot.png
012345_MEG_Wrkmem_tfavg_[LM-TRESP-tool]_[MODE-mag]_plot.png
012345_MEG_Wrkmem_tfavg_[LM-TRESP-tool]_[MODE-planar]_plot.png

The results of the Tfavg pipeline (with the appropriate time-frequency response conditions and contrast detail differences) for the Language Processing (Story-Math) and Motor tasks similarly unpack to the following directories:

MEG/StoryM/tfavg/
      figures/

MEG/Motort/tfavg/
      figures/

References

Achenbach TM (2009). <u>The Achenbach System of Empirically Based Assessement (ASEBA): Development, Findings, Theory, and Applications</u>. Burlington, VT: University of Vermont Research Center for Children, Youth and Families.

Andersson JLR, Skare S, Ashburner J (2003). How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage* 20(2): 870-888.

Arditi A (2005). Improving the design of the letter contrast sensitivity test. *Invest Ophthalmol Vis Sci* 46(6): 2225-2229.

Barch DM, Burgess GC, Harms MP, Petersen SE, Schlaggar BL, Corbetta M, Glasser MF, Curtiss S, Dixit S, Feldt C, Nolan D, Bryant E, Hartley T, Footer O, Bjork JM, Poldrack R, Smith SM, Johansen-Berg H, Snyder AZ, Van Essen DC (2013). [Function in the human connectome: Task-fMRI and individual differences in behavior](). *NeuroImage* 80: 169-189.

Beck RW, Moke PS, Turpin AH, Ferris FL, SanGiovanni JP, Johnson CA, Birch EE, Chandler DL, Cox TA, Blair RC, Kraker RT (2003). A computerized method of visual acuity testing: adaptation of the early treatment of diabetic retinopathy study testing protocol. *Am J Ophthalmol* 135: 194-205.

Beckmann CF, Smith SM (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Trans Med Imaging.* 23(2): 137-152.

Bilker WB, Hansen JA, Brensinger CM, Richard J, Gur RE, Gur RC (2012). Development of abbreviated nine-item forms of the Raven's standard progressive matrices test. *Assessment* 19(3): 354-369.

Binder JR, Gross WL, Allendorfer JB, Bonilha L, Chapin J, Edwards JC, Grabowski TJ, Langfitt JT, Loring DW, Lowe MJ, Koenig K, Morgan PS, Ojemann JG, Rorden C, Szaflarski JP, Tivarus ME, Weaver KE (2011). Mapping anterior temporal lobe language areas with fMRI: a multicenter normative study. *NeuroImage* 54(2): 1465-1475.

Bizzi A, Blasi V, Falini A, Ferroli P, Cadioli M, Danesi U, Aquino D, Marras C, Caldiroli D, Broggi G (2008). Presurgical functional MR imaging of language and motor functions: validation with intraoperative electrocortical mapping. *Radiology* 248(2): 579-589.

Bucholz KK, Cadoret R, Cloninger CR, Dinwiddie SH, Hesselbrock VM, Nurnberger Jr. JI, Reich T, Schmidt I, Schuckit MA (1994). A new, semi-structured psychiatric interview for use in

genetic linkage studies: a report on the reliability of the SSAGA. *J Stud Alcohol* 55: 149-158.

Buckner RL, Krienen FM, Castellanos A, Diaz JC, Yeo BT (2011). The organization of the human cerebellum estimated by intrinsic functional connectivity. *J Neurophysiol* 106(5): 2322-2345.

Buysse DJ, Reynolds III CF, Monk TH, Berman SR, Kupfer DJ (1989). The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res* 28: 193-213.

Caceres A, Hall DL, Zelaya FO, Williams SC, Mehta MA (2009). Measuring fMRI reliability with the intra-class correlation coefficient. *Neuroimage* 45(3): 758-768.

Castelli F, Frith C, Happé F, Frith U (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain* 125(Pt 8): 1839-1849.

Castelli F, Happé F, Frith U, Frith C (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage* 12(3): 314-325.

Christoff K, Prabhakaran V, Dorfman J, Zhao Z, Kroger JK, Holyoak KJ, Gabrieli JD (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *NeuroImage* 14(5): 1136-1149.

Cole BL (2007). Assessment of inherited colour vision defects in clinical practice. *Clin Exp Optom* 90(3): 157-175.

Conway AR, Kane MJ, Bunting MF, Hambrick DZ, Wilhelm O, Engle RW (2005). Working memory span tasks: A methodological review and user's guide. *Psychon Bull Rev* 12(5): 769-786.

Crum RM, Anthony JC, Bassett SS, Folstein MF (1993). Population-based norms for the Mini-Mental State Examination by age and educational level. *JAMA* 269: 2386-2391.

Dale AM, Fischl B, Sereno MI (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage* 9(2): 179-194.

Delgado MR, Nystrom LE, Fissell C, Noll DC, Fiez JA (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *J Neurophysiol* 84: 3072-3077.

Dougherty BE, Flom RE, Bullimore MA (2005). An evaluation of the Mars Letter Contrast Sensitivity Test. *Optom Vis Sci* 82(11): 970-975.

Downing PE, Jiang Y, Shuman M, Kanwisher N (2001). A cortical area selective for visual processing of the human body. *Science* 293(5539): 2470-2473.

Drobyshevsky A, Baumann SB, Schneider W (2006). A rapid fMRI task battery for mapping of visual, motor, cognitive, and emotional function. *NeuroImage* 31(2): 732-744.

Duncan J (2003). Intelligence tests predict brain response to demanding task events. *Nat Neurosci* 6(3): 207-208.

Duncan J (2005). Frontal lobe function and general intelligence: why it matters. *Cortex* 41(2): 215-217.

Duncan J, Seitz RJ, Kolodny J, Bor D, Herzog H, Ahmed A, Newell FN, Emslie H (2000). A neural basis for general intelligence. *Science* 289(5478): 457-460.

Estle SJ, Green L, Myerson J, Holt DD (2006). Differential effects of amount on temporal and probability discounting of gains and losses. *Mem Cognit* 34(4): 914-928.

Folstein MF, Folstein SE, McHugh PR (1975). Mini-mental state: a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12,189-198.

Forbes EE, Hariri AR, Martin SL, Silk JS, Moyles DL, Fisher PM, Brown SM, Ryan ND, Birmaher B, Axelson DA, Dahl RE (2009). Altered striatal activation predicting real-world positive affect in adolescent major depressive disorder. *Am J Psychiatry* 166(1): 64-73.

Fox CJ, Iaria G, Barton JJ (2009). Defining the face processing network: optimization of the functional localizer in fMRI. *Hum Brain Mapp* 30(5): 1637-1651.

Glasser MF, Van Essen DC (2011). Mapping human cortical areas in vivo based on myelin content as revealed by T1-and T2-weighted MRI. *J Neurosci* 31:11597-11616.

Glasser MF, Sotiropoulos SN, Wilson JA, Coalson T, Fischl B, Andersson JL, Xu J, Jbabdi S, Webster M, Polimeni JR, Van Essen DC, Jenkinson M (2013). The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* 80:105-124.

Goldberg LR (1993). The structure of phenotypic personality traits. *Am Psychol* 48: 26-34.

Gray JR, Chabris CF, Braver TS (2003). Neural mechanisms of general fluid intelligence. *Nat Neurosci* 6(3): 316-322.

Gray JR, Burgess GC, Schaefer A, Yarkoni T, Larsen RJ, Braver TS (2005). Affective personality differences in neural processing efficiency confirmed using fMRI. *Cogn Affect Behav Neurosci* 5(2): 182-190.

Green L, Myerson J, Shah AK, Estle SJ, Holt DD (2007). Do adjusting-amount and adjusting-delay procedures produce equivalent estimates of subjective value in pigeons? *J Exp Anal Behav* 87(3): 337-347.

Griffanti L, Salimi-Khorshidi G, Beckmann CF, Bijsterbisch J, Auerbach EJ, Douaud G, Zsoldos E, Ebmeier KP, Filippini N, Mackay C, Moeller S, Xu J, Yacoub E, Baselli G, Ugurbil K, Miller KL, Smith SM (2013). Automated artefact removal and accelerated fMRI acquisition for improved Resting State Network imaging (in preparation).

Gur RC, Ragland JD, Moberg PJ, Turner TH, Bilker WB, Kohler C, Siegel SJ, Gur RE (2001). Computerized neurocognitive scanning: I. Methodology and validation in healthy people. *Neuropsychopharmacology* 25(5): 766-776.

Gur RC, Ragland JD, Moberg PJ, Bilker WB, Kohler C, Siegel SJ, Gur RE (2001). Computerized neurocognitive scanning: II. The profile of schizophrenia. *Neuropsychopharmacology* 25(5): 777-788.

Gur RC, Richard J, Hughett P, Calkins ME, Macy L, Bilker WB, Brensinger C, Gur RE (2010). A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: standardization and initial construct validation. *J Neurosci Methods* 187(2): 254-262.

Hariri AR, Brown SM, Williamson DE, Flory JD, de Wit H, Manuck SB (2006). Preference for immediate over delayed rewards is associated with magnitude of ventral striatal activity. *J Neurosci* 26: 13213-13217.

Haymes SA, Roberts KF, Cruess AF, Nicolela MT, LeBlanc RP, Ramsey MS, Chauhan BC, Artes PH (2006). The letter contrast sensitivity test: clinical evaluation of a new design. *Invest Ophthalmol Vis Sci* 47(6): 2739-2745.

Heatherton TF, Kozlowski LT, Frecker RC, Fagerstrom KO (1991). The Fagerstrom Test for Nicotine Dependence: a revision of the Fagerstrom Tolerance Questionnaire. *Br J Addict* 86: 1119-1127.

Heine SJ, Buchtel EE (2009). Personality: the universal and the culturally specific. *Annu Rev Psychol* 60: 369-394.

Hesselbrock M, Easton C, Bucholz KK, Schuckit M, Hesselbrock V (1999). A validity study of the SSAGA— a comparison with the SCAN. *Addiction* 94: 1361-1370.

Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM (2012). FSL. *NeuroImage* 62: 782-90.

Jovicich J, Czanner S, Greve D, Haley E, Kouwe A, Gollub R, Kennedy D, Schmitt F, Brown G, MacFall J, Fischl B, Dale A (2006). Reliability in Multi-Site Structural MRI Studies: Effects of Gradient Non-linearity Correction on Phantom and Human Data. *NeuroImage* 30(2): 436-43.

Kozlowski LT, Porter CQ, Orleans CT, Pope MA, Heatherton T (1994). Predicting smoking cessation with self-reported measures of nicotine dependence: FTQ, FTND, and HSI. *Drug Alcohol Depend* 34: 211-216.

Kung CC, Peissig JJ, Tarr MJ (2007). Is region-of-interest overlap comparison a reliable measure of category specificity? *J Cogn Neurosci* 19(12): 2019-2034.

Larson-Prior LJ, Oostenveld R, Della Penna S Michalareas G, Prior F, Babajani-Feremi, Schoffelen J-M, Marzetti L, de Pasquale F, Di Pompeo F, Stout J, Moolrich M, Luo Q, Bucholz R, Fries P, Pizzella V, Romani GL, Corbetta M, Snyder AZ (2013). Adding dynamics to the Human Connectome Project with MEG. *NeuroImage* 80: 190-201.

Manuck SB, Brown SM, Forbes EE, Hariri AR (2007). Temporal stability of individual differences in amygdala reactivity. *Am J Psychiatry* 164(10): 1613-1614.

Mantini D, Della Penna S, Marzetti L, de Pasquale F, Pizzella V, Corbetta M, Romani GL (2011). A signal-processing pipeline for magnetoencephalography resting-state networks. Brain Cogn 1: 49–59.

Marcus DS, Harms MP, Snyder AZ, Jenkinson M, Wilson JA, Glasser MF, Barch DM, Archie KA, Burgess GC, Ramaratnam M, Hodge M, Horton W, Herrick R, Olsen T, McKay M, House M, Hileman M, Reid E, Harwell J, Coalson T, Schindler J, Elam JS, Curtiss SW, Van Essen DC (2013). Human Connectome Project informatics: Quality control, database services, and data visualization. *NeuroImage* 80: 144-168.

May JC, Delgado MR, Dahl RE, Stenger VA, Ryan ND, Fiez JA, Carter C (2004). Event-related functional magnetic resonance imaging of reward-related brain circuitry in children and adolescents. *Biol psychiatry* 55(4): 359-366.

McCrae RR, Costa PT (2004). A contemplated revision of the NEO Five Factor Inventory. *Pers Individ Diff* 36(3): 587-596.

McCrae RR, Costa PT (2008). The five factor theory of personality. *Handbook of Personality: Theory and Research*. O. P. John, R. W. Robins and L. A. Pervin. New York, Guilford: 159-181.

Milchenko M, Marcus D (2013). Obscuring surface anatomy in volumetric imaging data. *Neuroinformatics* 11: 65-75.

Moke PS, Turpin AH, Beck RW, Holmes JM, Repka MX, Birch EE, Hertle RW, Kraker RT, Miller JM, Johnson CA (2001). Computerized method of visual acuity testing: adaptation of the amblyopia treatment study visual acuity testing protocol. *Am J Ophthalmol* 132: 903-909.

Morioka T, Yamamoto T, Mizushima A, Tombimatsu S, Shigeto H, Hasuo K, Nishio S, Fujii K, Fukui M (1995). Comparison of magnetoencephalography, functional MRI, and motor evoked potentials in the localization of the sensory-motor cortex. *Neurol Res* 17(5): 361-367.

Myerson J, Green L, Warusawitharana M (2001). Area under the curve as a measure of discounting. *J Exp Anal Behav* 76(2): 235-243.

Oldfield RC (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1): 97-113.

Oostenveld R, Fries P, Maris E, and Schoffelen J-M (2011). [FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data](). *Computational Intelligence and Neuroscience*, vol. 2011, Article ID 156869, 9 pages, 2011. doi:10.1155/2011/156869

Peelen MV, Downing PE (2005). Within-subject reproducibility of category-specific visual activation with functional MRI. *Hum Brain Mapp* 25(4): 402-408.

Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, Petersen SE (2011). Functional network organization of the human brain. *Neuron* 72(4): 665-78.

Prabhakaran V, Smith JA, Desmond JE, Glover GH, Gabrieli JD (1997). Neural substrates of fluid reasoning: an fMRI study of neocortical activation during performance of the Raven's Progressive Matrices Test. *Cogn Psychol* 33(1): 43-63.

Rilling JK, Glasser MF, Jbabdi S, Andersson J, Preuss TM (2012). Continuity, divergence, and the evolution of brain language pathways. *Front Evol Neurosci* 3: 11.

Salimi-Khorshidi G, Douaud G, Beckmann CF, Bijsterbisch J, Duff E, Glasser MF, Griffanti L, Smith SM (2013). Automatic Denoising of Functional MRI Data: Combining Independent Component Analysis and Hierarchical Fusion of Classifiers (in preparation).

Satterthwaite TD, Elliott MA, Gerraty RT, Ruparel K, Loughead J, Calkins ME, Eickhoff SB, Hakonarson H, Gur RC, Gur RE, Wolf DH (2013). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *Neuroimage*, 64(0): 240-256.

Smith R, Keramatian K, Christoff K (2007). Localizing the rostrolateral prefrontal cortex at the individual level. *Neuroimage* 36(4): 1387-1396.

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23(S1): 208-19.

Smith SM, Andersson J, Auerbach EJ, Beckmann CF, Bijsterbosch J, Douaud G, Duff E, Feinberg DA, Griffanti L, Harms MP, Kelly M, Laumann T, Miller KL, Moeller S, Petersen SE, Power J, Salimi-Khorshidi G, Snyder AZ, Vu A, Woolrich MW, Xu J, Yacoub E, Ugurbil K, Van Essen DC, Glasser MF (2013). Resting-state fMRI in the Human Connectome Project. *NeuroImage* 80: 144-168.

Sotiropoulos SN, Moeller S, Jbabdi S, Xu J, Andersson JL, Auerbach EJ, Yacoub E, Feinberg D, Setsompop K, Wald LL, Behrens TEJ, Ugurbil K, Lenglet C (2013). Effects of Image reconstruction on fibre orientation mapping from multichannel diffusion MRI: Reducing the noise floor using SENSE. *Mag Res Med* EPub ahead of print 7 Feb 2013.

Sotiropoulos SN, Jbabdi S, Xu J, Andersson JL, Moeller S, Auerbach EJ, Glasser MF, Hernandez M, Sapiro G, Jenkinson M, Feinberg D, Yacoub E, Lenglet C, Van Essen DC, Ugurbil K, Behrens TEJ (2013). Advances in diffusion MRI acquisition and processing in the Human Connectome Project. *NeuroImage* 80: 125-143.

Taylor JC, Wiggett AJ, Downing PE (2007). Functional MRI analysis of body and body part representations in the extrastriate and fusiform body areas. *J Neurophysiol* 98(3): 1626-1633.

Thayaparan K, Crossland MD, Rubin GS (2007). Clinical assessment of two new contrast sensitivity charts. *Br J Ophthalmol* 91(6): 749-752.

Tricomi EM, Delgado MR, Fiez JA (2004). Modulation of caudate activity by action contingency. *Neuron* 41(2): 281-292.

Ugurbil K, Xu J, Auerbach EJ, Moeller S, Vu A, Duarte-Carvajalino JM, Lenglet C, WU X, Schmitter S, Van de Moortele PF, Strupp J, Sapiro G, De Martino F, Wang D, Harel N, Garwood M, Chen L, Feinberg DA, Smith SM, Miller KL, Sotiropoulos SN, Jbabdi S, Andersson JL, Behrens TEJ, Glasser MF, Van Essen DC, Yacoub E (2013). Pushing spatial and temporal resolution for functional and diffusion MRI in the Human Connectome Project. *NeuroImage* 80: 80-104.

Van Essen DC, Glasser MF, Dierker DL, Harwell J, Coalson T (2012). Parcellations and hemispheric asymmetries of human cerebral cortex analyzed on surface-based atlases. *Cereb Cortex* 22(10): 2241-62.

Van Essen DC, Smith SM, Barch DM, Behrens TEJ, Yacoub E, Ugurbil K (2013). The WU-Minn Human Connectome Project: An Overview. *NeuroImage* 80: 62-79.

Wheatley T, Milleville SC, Martin A (2007). Understanding animate agents: distinct roles for the social network and mirror system. *Psychol Sci* 18(6): 469-474.

White SJ, Coniston D, Rogers R, Frith U (2011). Developing the Frith-Happe animations: a quick and objective test of Theory of Mind for adults with autism. *Autism Res* 4(2): 149-154.

Winter WR, Nunez PL, Ding J, Srinivasan R (2007). Comparison of the effect of volume conduction on EEG coherence with the effect of field spread on MEG coherence. *Stat Med* 26: 3946–3957.

Wendelken C, Nakhabenko D, Donohue SE, Carter CS, Bunge SA (2008). Brain is to thought as stomach is to ??: investigating the role of rostrolateral prefrontal cortex in relational reasoning. *J Cogn Neurosci* 20(4): 682-693.

Woolrich MW, Jbabdi S, Patenaude B, Chappell M, Makni S, Behrens T, Beckmann C, Jenkinson M, Smith SM (2009). Bayesian analysis of neuroimaging data in FSL. *NeuroImage* 45(1 Suppl): S173-S186.

Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, Fischl B, Liu H, Buckner RL (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* 106(3): 1125-1165.

Zaitsev M, Dold, Sakas G, Hennig J, Speck O (2006). Magnetic resonance imaging of freely moving objects: prospective real-time motion correction using an external optical motion tracking system. *NeuroImage* 31:1038–1050.

** Further details on the HCP data acquisition and analysis methods as of March 2013 are available in eight manuscripts published in a special Connectome issue of *NeuroImage* (Larson-Prior *et al.*, 2013, Van Essen *et al.*, 2013, Ugurbil *et al.*, 2013, Glasser *et al.* 2013, Smith *et al.*, 2013, Barch *et al.*, 2013, Sotiropoulos *et al.*, 2013, and Marcus *et al.*, 2013).

Appendices

Appendix 1. HCP MEG Session Protocol Details

Download Appendix 1 here:
http://humanconnectome.org/documentation/MEG1/HCP_MEG1_Release_Appendix_I.pdf

Appendix 2. HCP MEG Protocol Standard Operating Procedures (SOPs)

Download Appendix 2 here:
http://humanconnectome.org/documentation/MEG1/HCP_MEG1_Release_Appendix_II.pdf

Appendix 3. File Names and Directory Structure for Unprocessed, Anatomical Model, and Channel-Level Processed MEG Data

Download Appendix 3 here:
http://humanconnectome.org/documentation/MEG1/HCP_MEG1_Release_Appendix_III.pdf

Appendix 4. Task MEG E-Prime Key Variables

Download Appendix 4 here:
http://humanconnectome.org/documentation/MEG1/HCP_MEG1_Release_Appendix_IV.pdf

Appendix 5. tMEG: Extracting subsets of trials from the cleaned data produced by tmegpreproc pipeline

Download Appendix 5 here:
http://humanconnectome.org/documentation/MEG1/HCP_MEG1_Release_Appendix_V.pdf