WU-Minn HCP Quarter 1 (Q1) Data Release: Reference Manual

5 March 2013

Table of Contents

Introduction and overview

This document provides information and guidance on how to use the Quarter 1 (Q1) data released by the WU-Minn HCP consortium in March 2013. 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 quarterly releases.

Should I be interested in the HCP Q1 data release? Over the next several years, the HCP will make available an enormous amount of data to enable exploration of human brain circuits, their relationship to individual behavior, and their heritability and genetic underpinnings. In the long term, the HCP data will thus be of interest to investigators in many fields, including neuroimaging, neuroanatomy, psychology, and network modeling - to name just a few.

The HCP aims to study 400 subjects per year, with a target of 1,200 subjects studied before the grant ends in 2015. Subjects from families with twins and non-twin siblings are being scanned on the same scanner using the same protocol for every subject. Data will be released quarterly, starting with the current Q1 release.

The Q1 data release is limited in terms of the number of subjects extensively scanned (68 vs the eventual target of 1,200 subjects), the amount of highly processed data that is available, and the data mining capabilities currently implemented. Consequently, we anticipate that the Q1 data release will be of particular interest to investigators who are prepared to apply their own neuroimaging analysis tools to this 'starter kit' of imaging data – especially by using the 'minimally preprocessed' datasets described below. Investigators interested in extensively processed data can access a 20-subject group-average dataset that includes multiple modalities.

If you are interested in accessing the Q1 data release, please continue reading the following important information.

What's in the HCP Q1 data release? The Q1 data include high-resolution MR scans from 68 healthy adults and four imaging modalities: structural images (T1w and T2w), resting-state fMRI (rfMRI), task-fMRI (tfMRI), and high angular resolution diffusion imaging (dMRI). Some of the behavioral data acquired from each subject is also available. Since the subjects include twins and their non-twin siblings, it will be important for many analyses to know the family status of the subjects under study. **Family status and other potentially sensitive information are part of the Restricted Access data that is available only to qualified investigators (see below)**. **To protect the privacy of our HCP subjects, anyone wanting to use the Restricted Access data MUST understand and agree to comply with the important constraints that are imposed.** To enable investigators to analyze data without being concerned about family structure issues, we have prepackaged groups of 5 and of 20 unrelated subjects for download (see How to download HCP data).

Multiple levels of data processing. The Q1 imaging data includes "**unprocessed**" data (in NIFTI format) plus "**minimally preprocessed**" datasets. We encourage investigators to use the preprocessed datasets, in which spatial distortions have been minimized and data have been aligned across modalities and across subjects. The methods used for this preprocessing, (see Preprocessing Pipelines section below, were implemented by the HCP consortium through an intensive two-year refinement and optimization process. If your specific research interests make it important to use the unprocessed data, it is critical to be aware of certain characteristics that require special preprocessing (see Gradient Nonlinearities in Unprocessed Datasets).

The preprocessed structural datasets include cortical surfaces, myelin maps, and other data that are suitable for a variety of morphometric analyses. In addition, the HCP has analyzed resting-state fMRI and task-fMRI data on the group of 20 unrelated Q1 subjects. The Q1 data release includes demonstration datasets based on group-average functional connectivity and task-fMRI. The extensively processed data for the individual subjects used in this analysis are not being released at this time because key methods are still being evaluated and refined (e.g., temporal preprocessing, denoising, and correcting for residual motion artifacts).

How do I access HCP data? Many HCP imaging files are large, and they must be maintained in a well-organized directory structure in order to prevent serious confusion or corruption when analyzing the data. For the Q1 release, the total amount of imaging data for all modalities is about 2 terabytes (about 30 GB per subject).

HCP provides three ways to access part or all of these data:

(i) ***"Connectome in a Box"***. Investigators wanting access to data from all Q1 subjects and most or all imaging modalities are encouraged to use the Connectome in a Box option. A hard drive containing all of the available imaging data will be shipped within two weeks of receiving payment for the hardware and delivery costs (see *Connectome in a Box* below). Institutions or imaging centers having many investigators who want access to the data should consider this option to facilitate sharing of HCP open access data.

(ii) ***Download Packages from ConnectomeDB***. The HCP provides an easy way to select standard packages based on modality of interest and a standard set of subjects. This includes selection of datasets by modality for one subject, five unrelated subjects, and 20 unrelated subjects. Downloading is mediated by software that enables high-speed file transfer. Download times will depend on bandwidth and overall user demand, and there is a 500 GB limit on archive size that can be requested (see How to download HCP data).

(iii) The **ConnectomeDB** database user interface allows filtering of subjects based on a range of selection criteria (see How to access HCP data). This will be of limited utility for the Q1 release, but will become much more important as the number of subjects increases and the capabilities of the data mining interface expand in future quarterly releases.

How can I analyze and visualize HCP data? Most of the preprocessed HCP data files are in standard formats supported by major neuroimaging platforms, including NIFTI (Neuroimaging Informatics Technology Initiative, see http://nifti.nimh.nih.gov) format for standard 3D or 4D volume data and GIFTI (Geometry Informatics Technology Initiative) format for surface data. Importantly, the preprocessing pipelines for fMRI data (resting-state and task-fMRI) also include files that are in a 'grayordinate' representation that combines cortical surface vertices and subcortical gray matter voxels into a single file using the recently introduced CIFTI data format (see Glasser *et al*., submitted). These grayordinate CIFTI files offer major advantages in terms of efficiency in handling very large and complex datasets.

The Connectome Workbench platform can be used to visualize the outputs of HCP preprocessing and advanced processing pipelines, including CIFTI, NIFTI and GIFTI files. Connectome Workbench is freely available (beta version 0.8 at http://humanconnectome.org/connectome/get-connectome-workbench.html) and includes a tutorial and associated dataset. Connectome Workbench includes many command-line options (wb_command) that support a variety of analysis steps, including processing of CIFTI files.

What's in store for future quarterly releases? Future quarterly releases will not only add to the total number of subjects studied, but will include the more extensively analyzed datasets and will have progressively enhanced data mining and visualization capabilities.

A subset of 100 HCP subjects will have MEG/EEG scans acquired at St. Louis University, commencing in the spring of 2013. Another subset of 200 HCP subjects will be scanned using a 7T scanner at the University of Minnesota, commencing in the fall of 2013. Data from these additional modalities will be made available in future quarterly releases.

How were the HCP datasets acquired and processed? Besides the documentation provided below, there are published descriptions of various aspects of the HCP (Van Essen *et al*., 2012; Marcus *et al*., 2011). Detailed descriptions of many HCP methods will be available in eight papers to appear in a special issue of NeuroImage in 2013. See http://www.humanconnectome.org/documentation/citations.html for updated lists of publications.

What if I want to adapt HCP protocols to my own scanner and research projects? Appendix I: Protocol Guidance and HCP Session Protocols 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 access databases - IMPORTANT!

The HCP is designed to provide investigators with a rich dataset allowing comparison of different imaging modalities and the opportunity to study the heritability of brain connectivity and behavior. However, the study design also poses some special challenges for the protection of subject privacy. Since this is a family study in which family structures can sometimes be complex, and since a wide variety of other types of data have also been collected, there is a risk that subjects could be recognizable to family members and others unless these data are properly managed. In addition, the dataset includes information that could harm or embarrass subjects if they were to be inadvertently identified.

To protect our subjects' privacy, HCP investigators have developed a two-tiered plan for sharing data, with different provisions for the Open Access data releases and the Restricted Access data releases to qualified investigators.

Open Access Data, available to those who register and sign the Open Access Data Use Terms, include imaging data, sex, subject age in 5-year ranges, and most behavioral data. In the Q1 data release, Restricted Access Data comprise family structure (twin or non-twin status and subject number of siblings); birth order; exact age; color vision test results and visual correction measures; handedness; ethnicity and race; body height, weight, and BMI; all data from the SSAGA telephone diagnostic interview; drug test results; HbA1c and TSH results; Information on endocrine disorders and age of onset; and any psychiatric and neurologic illnesses of subjects' parents.

Each investigator who wants to be able to view and use Restricted Access Data must apply for and be granted access. Importantly, as part of the application process each must also agree to an additional, Restricted Access Data Use Agreement that explains how Restricted Access Data may and may not be used. This includes significant limitations on how Restricted Access data can be presented in publications. More information about required investigator qualifications, use of Restricted Access Data, and the application process can be found on the HCP site.

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., FSL, FreeSurfer), 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. Comments submitted this way will be posted to the hcp-users mailing list so that all can benefit from the information provided.

Gradient nonlinearities in unprocessed datasets

If you intend to download the unprocessed NIFTI data, read on! Unusually large gradient nonlinearities are present in the raw images for ALL modalities, and it is important to correct for the spatial distortions they cause.

All HCP imaging data for this Q1 data release were acquired on a Siemens Skyra 3T scanner with a customized SC72 gradient insert that greatly improves the quality of diffusion imaging scans (the 'Connectome Skyra'). Higher performing gradients require compromises in bore diameter and gradient nonlinearities. Further, in custom-fitting the higher performing gradient set into a standard clinical system, technical limitations prevent centering of the subjects' heads in the bore isocenter. Consequently, the gradient nonlinearities associated with all Connectome Skyra scans exceed those of a conventional clinical 3T scanner. In the HCP processed datasets for all scan modalities (structural, fMRI, and dMRI), these distortions have been corrected for by spatially warping the images using gradient field information specific to the Connectome Skyra. The gradient unwarping code is available at https://github.com/ksubramz/gradunwarp/blob/master/Readme.md (Jovicich *et al.*, 2006). The gradient field nonlinearity coefficients for the Connectome Skyra are considered by Siemens to be proprietary information. To request access to these coefficients, please contact your Siemens collaboration manager or email Dingxin Wang at dingxin.wang@siemens.com.

For diffusion MRI, the gradient nonlinearities also cause voxel-by-voxel changes in the strength and orientation of the diffusion encoding gradients. Consequently, the effective b-values and b-vectors in all the primary data that you can download exhibit small variations from voxel to voxel. When analyzing the primary (unprocessed) datasets or the minimally preprocessed dMRI data you may use the code provided in Appendix 2: Matlab code for voxel-wise correction of dMRI gradients in your fitting routine in order to account for gradient spatial nonlinearities during fiber orientation estimation.

Relation to Initial Open Access Data Release

An Initial Open Access Data Release based on 12 subjects was released in October 2012 (version 1). This was updated to version 2 in December 2012 to include diffusion MRI data and to incorporate improvements in preprocessing of fMRI volumes (see http://www.mail-archive.com/hcp-users@humanconnectome.org/msg00009.html).

The 12 subjects from the Initial Open Access Data Release are all included as a subset of the current HCP Q1 Data Release. The unprocessed datasets are unchanged, and the preprocessed data are unchanged relative to version 2 (December 2012), except that preprocessing of the dMRI scans used an improved method of compensating for eddy currents and spatial distortions.

For all HCP data releases, be sure to keep track of version numbers when analyzing HCP data. Improvements in processing pipelines that occur over time may warrant retrospective application of updated pipelines to previous data releases. The HCP will document any such changes associated with modified processing pipelines between versions.

HCP Data Sizes

The MRI images collected for each HCP subject are notably high in spatial resolution and (for fMRI) in temporal resolution as well. Coupled with the long scan duration for each modality, the resultant image 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 imaging data generated by the HCP protocol (total Q1 release data ~2TB) requires some special considerations:

If you choose to obtain your data via download, prepare for long download times. The HCP will use 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. Even data on a few subjects (total data/subject: ~10 GB unprocessed, ~16 GB preprocessed) 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. If you want all the data, consider ordering “Connectome in a Box” (see below).

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 joint ordering of Connectome in a Box (recommended) or a single download that can be copied and used by many within your institution. With a modest

HCP Data Sizes (per Subject)

Session	Format	.zip File Size
Structural	Unprocessed	70.99 MB
	Preprocessed	1.19 GB
Resting State fMRI (each of 2 sessions)	Unprocessed	2 GB
	Preprocessed	3.24 GB
Task fMRI (avg per Task)	Unprocessed	490 MB
	Preprocessed	771 MB
(all 7 Tasks)	Unprocessed	3.43 GB
	Preprocessed	5.4 GB
Diffusion	Unprocessed	2.18 GB
	Preprocessed	2.81 GB
Group-Average on Unrelated 20	Additionally Processed	296 MB
Total (per Subject)	Unprocessed	9.81 GB
	Preprocessed	15.77 GB
	Both	25.58 GB
Total (5 Subjects)	Unprocessed	62.16 GB
	Preprocessed	78.83 GB
	Both	141 GB
Total (20 Subjects)	Unprocessed	247.34 GB
	Preprocessed	315.05 GB
	Both	562.39 GB
Total (68 Subjects)	Unprocessed	815.4 GB
	Preprocessed	1.058 TB
	Both	1.873 TB

amount of planning, groups that chose to do this will enjoy faster distribution and save valuable research time.

With these realities in mind, the HCP is distributing image data in two ways:

1. Via download through ConnectomeDB. HCP has made convenient data packages in 1 subject, 5 subject, or 20 unrelated subject groups by modality type to allow the user to "try out" the data without incurring a large download or data storage burden. The subjects in these set groups are unrelated to facilitate its use to investigators who want to analyze data without being concerned about family structure issues. The ConnectomeDB interface also allows the user to create their own subject groups of interest and download only the modalities necessary for analysis of that group.

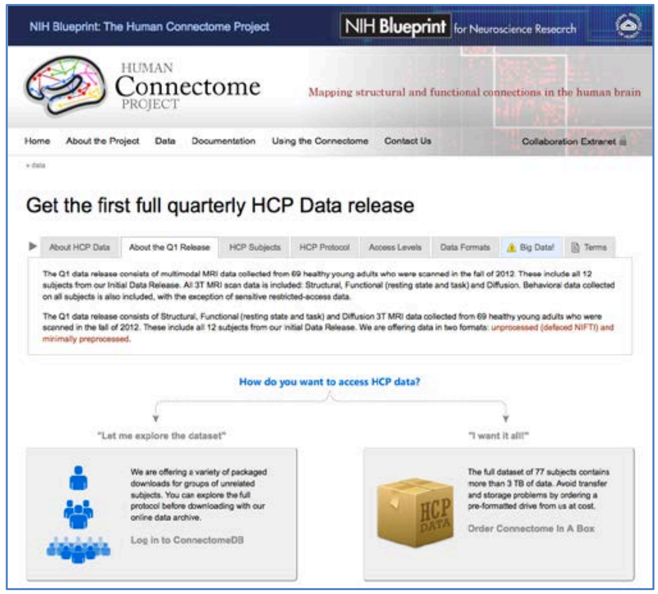
2. Q1 "Connectome in a Box". A complete Q1 Data Release preloaded onto a formatted 4TB hard drive, that can be ordered and shipped to you for a minimal cost (~$200, the cost of the hard drive + shipping).

In addition to the image data, the HCP is releasing additionally processed group-average resting-state fMRI and task-fMRI data acquired from 20 unrelated subjects. This data is available both through download (296 MB) and as part of Q1 Connectome in a Box. Links to two ~30 GB demonstration datasets based on extensively processed group-average functional connectivity are available for visualization through Connectome Workbench beta v0.8. An associated Connectome Workbench Tutorial for the Q1 Data Release is packaged with the Group Average data available for download on ConnectomeDB.

As a reminder, the extensively processed data for the individual subjects used in this analysis are not being released for download at this time because key methods are still being evaluated and refined (e.g., temporal preprocessing, denoising, and correcting for residual motion artifacts).

How to access HCP data

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



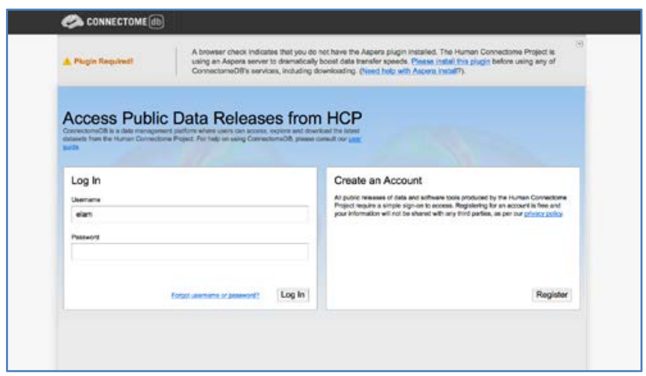
Here, you can get an overview of the Q1 data release, read the [Open Access Data Use Terms](), explore links to details of the HCP protocols, and access more useful resources for using HCP data.

Just below the top panel, you can decide if you would like to:

- Explore and/or download data in our database platform, ConnectomeDB

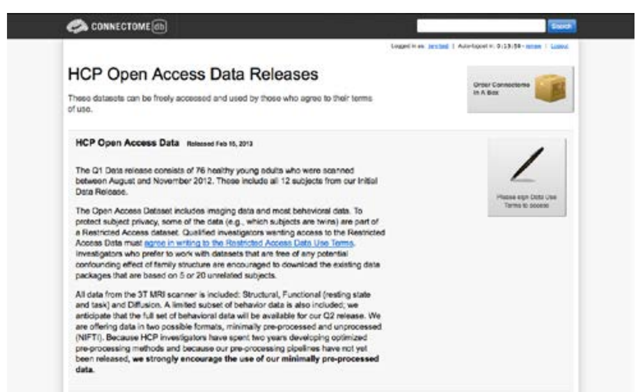
  or

- Order all the Q1 data via the "Connectome in a Box" option.



Clicking on either option takes you to a dialog to register for a ConnectomeDB account.

After you have created an account and are logged in, you must agree to the HCP Open Access Data Use Terms (required for access).



Once these steps are complete, you will launch the HCP Open Access Data Releases page where you can begin to explore and download HCP data by selecting options for groups of subjects from the Q1 release (see [How to download HCP data]() below).

Previously released HCP data is also available in the 'More Datasets' section.

Note: Your login to ConnectomeDB times out every 15 minutes. Click the "renew" link to reset the time. (Auto-logout resets ConnectomeDB to the HCP Open Access Data Release page).

How to download HCP data

In order to facilitate access to the very large data files generated on every subject (see HCP Data Sizes above), we have prepackaged data into convenient group data packages. Datasets can be downloaded from ConnectomeDB in set packages or users may order all the Q1 data to be sent to them on a hard drive (Connectome in a Box, see below).

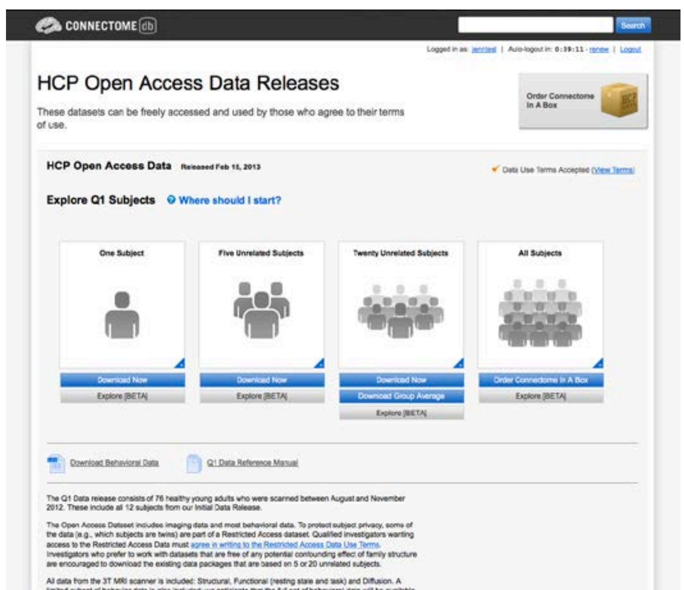
Due to the large file sizes of the data, the HCP uses an Aspera server to dramatically boost data transfer speeds. Therefore, downloading data from the HCP requires you to have the Aspera plugin installed on your browser.

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 on page 12). 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.

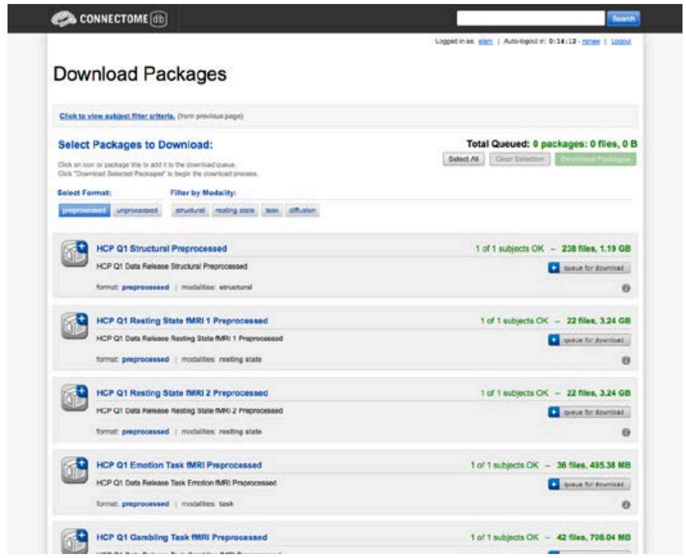
Upon login to ConnectomeDB, you are brought to the HCP Open Access Data Releases landing page, where you can:

- Download image data for groups of 1, 5, or 20 unrelated subjects
  - Hover over the illustrations of the subjects to get a short description to help you decide which group is best for your purposes
- Explore the data for these groups in the ConnectomeDB user interface (currently in Beta testing, not all planned functionality is yet implemented)
- Download additionally processed group-average data for the group of 20 unrelated subjects



- Download behavioral data on all subjects
- Order all Q1 image data via Connectome in a Box
- Load a custom group of subjects that you created in a previous ConnectomeDB session
- Browse and download other HCP data that has been released previously

If you choose one of the “Download Now” options for 1, 5, or 20 subjects, you will launch the Download Packages page.



First time Aspera Download Setup

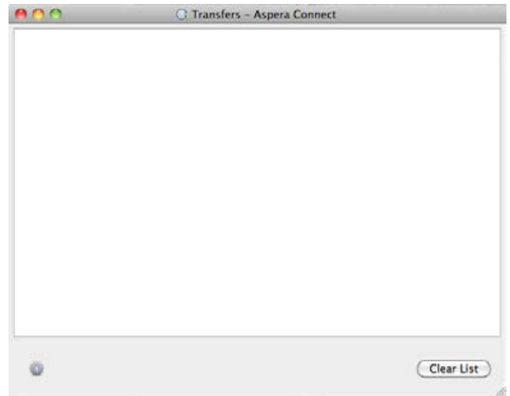
If this is your first time downloading data from ConnectomeDB, **before you start selecting packages to download**, we highly recommend setting up the preferences in your Aspera Connect plugin to download to the appropriate place on your local or network hard drive.

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:

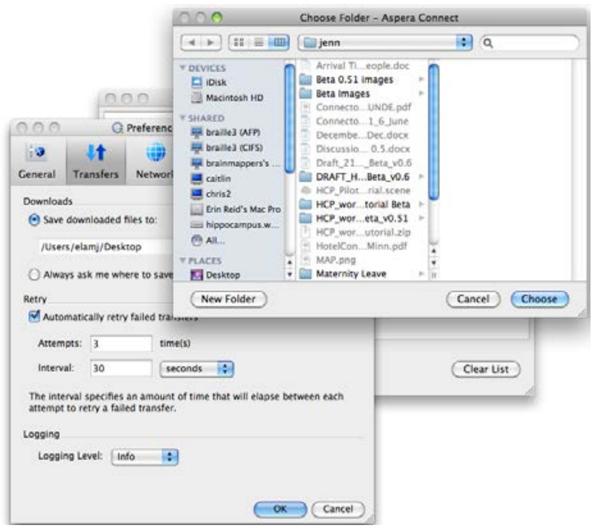
- Search your computer for Aspera Connect using the Spotlight or Search functions in your operating system. Click on “Aspera Connect” to open the application.
- Open the Aspera Connect: Transfers window:
  - On Mac, in the Menu bar click on Window > Transfers, or use the shortcut **^T** (control + T)
  - On PC, in the Notification area at the very bottom right of your Desktop, click the ▲ button, then click the Aspera Connect icon
  - 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.
- Now, each time you download HCP data it will go to the directory you selected (unless you change the preference again).



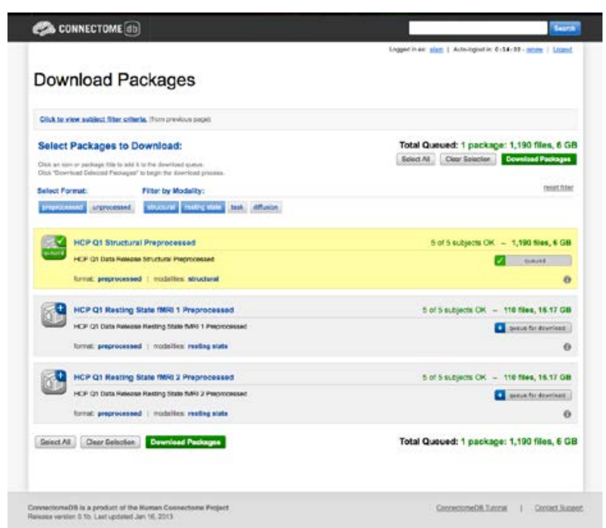
Downloading HCP data packages

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



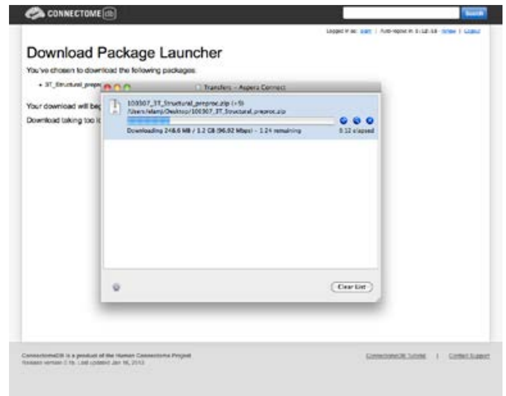
Here, you can filter the image data available for download by format and modality by clicking the "Select Format" and "Filter by modality" buttons. To make selections for download, click the "add to queue" icons to the left of each dataset.

The total size of the data you have queued for download is at the upper and bottom right. Use this calculation to check against your available hard drive space to be sure you have enough space available for the files you intend to download.



When you are done making your selections, click the Download Packages button at the upper right. A popup will remind you again what packages you selected and the total size of the files. Clicking "Download Now" will automatically open Aspera Connect and launch the download.

Note: If you have not setup 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 "Downloading Packages" 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 blue 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!).

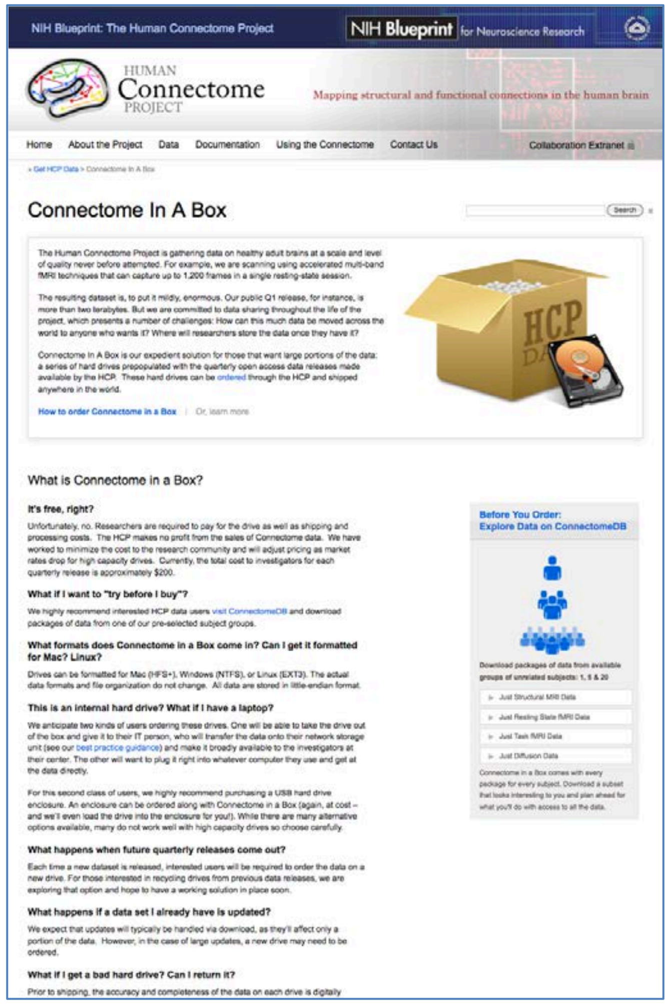
This is why we are offering the "Connectome in a Box" option for those who want data on many subjects (see below).

Handling downloaded HCP 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.
- Once the archives are downloaded and unzipped, many of the component files are gzipped (.gz files). You will need to unzip all.gz files in order to use them. For this, you need an application that is compatible with gzip. For Windows users, we again recommend 7-zip. Linux has support for gzip built in, and Mac users can use the Mac Gzip utility.
- Downloaded HCP data unzips to a set directory structure that is detailed below in Directory structure for unprocessed data, Directory structure for processed data, and in Appendix 3: File Names and Directory Structure for Unprocessed and Preprocessed HCP Q1 Data. The directory structure for downloaded HCP data is identical to what is provided in HCP_Q1 Connectome in a Box.
- 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.
- We are compiling a set of best practices for long-term data storage, which can be seen in our online documentation.

How to order Q1 Connectome in a Box



An attractive option for getting Q1 HCP data is “Connectome in a Box”. This option allows users to order a hard drive containing the complete Q1 imaging data from the HCP (HCP_Q1 data) “at cost” (approximately $200/Quarter of data, including shipping costs) and have it shipped to their address.

HCP_Q1 Connectome in a Box can be securely ordered from the HCP website.

Connectome in a Box contains both unprocessed and preprocessed NIFTI image data formats. The data is loaded on a SATA hard drive formatted to your choice of operating system (MacOSX, Linux, or Windows).

Once the user (or group of users) receive the drive, it can be added into a network storage unit and made available to all on the user’s local network. Alternatively, the drive can be simply plugged right into a computer and used as you would an external hard drive. For this option, we recommend purchasing a hard drive enclosure. We have an enclosure available (at cost) that you can purchase when you order Connectome in a Box, although you may wish to buy your own elsewhere.

Each time a new dataset is released, users may wish to order another drive with that quarter’s HCP data loaded onto it. We are currently exploring the option to recycle previous Connectome in a Box drives for subsequent releases (more on that when Q2 is released).

As HCP data is updated, depending on the amount of data and related logistics, we may recommend updating the HCP_Q1 Connectome in a Box via download from ConnectomeDB.

MR scanner and other hardware

Scanner hardware. All data in this Q1 release were acquired on a Siemens Skyra 3T scanner housed at Washington University in St. Louis. The scanner has a customized SC72 gradient insert and a customized body transmitter coil with 56 cm bore size (*diffusion*: Gmax = 100 mT/m, max slew rate = 91 mT/m/ms; *readout/imaging*: Gmax = 42 mT/m, max slew rate = 200 mT/m/ms). The HCP Skyra has the standard set of Siemen's shim coils (up to 2nd order) and the HCP is using Siemen's standard 32 channel head coil.

Visual projection and E-Prime computer. Visual stimuli were presented and participant responses were collected using a Dell Optiplex 790 computer, running an Intel Core i3-2100 with 8GB of RAM and 64-bit Windows 7 Enterprise SP1. The E-Prime version was E-Prime 2.0 Professional Production Release (2.0.10.242). Visual stimuli were projected with a NEC V260X projector onto a lucite screen at 1024x768 resolution, and viewed by the participant using a mirror mounted on the top of the head coil. Participant responses were registered on a customized fiber-optic button box.

Summary of imaging protocols

Structural, fMRI, and dMRI acquisitions were collected over 4 total imaging sessions, each approximately 1 hour in duration. Resting-state and task-fMRI data was collected in two sessions. Each session consisted of two resting-state acquisitions of approximately 15 minutes each, followed by task-fMRI acquisitions of varying durations (see below).

Vitamin E capsule on right side. A capsule of vitamin E was taped to the subject's right temple in every scan session, to enable definitive determination of the right side in the image data.

The following provides basic parameters for the main scan types in each session, and pertinent details about each session. A more complete set of imaging parameters can be found in the protocol exports from the scanner, available in Appendix 1. FOV positioning in all runs was handled in an automated manner using Siemens AutoAlign feature.

Structural session

Type	Series Description	Description	TR (ms)	TE (ms)	TI (ms)	Flip Angle	FOV (mm)	Voxel Size	BW (Hz/Px)	iPAT	Acquisition Time (min:sec)
T1w	T1w_MPR1	3D MPRAGE	2400	2.14	1000	8 deg	224x224	0.7 mm isotropic	210	2	7:40
T2w	T2w_SPC1	3D T2-SPACE	3200	565		variable	224x224	0.7 mm isotropic	744	2	8:24

Resting-state fMRI (rfMRI)

rfMRI data were acquired in four runs of approximately 15 minutes each, two runs in one session and two in another session, with eyes open with relaxed fixation on a projected bright cross-hair on a dark background (and presented in a darkened room). Within each session, oblique axial acquisitions alternated between phase encoding in a right-to-left (RL) direction in one run and phase encoding in a left-to-right (LR) direction in the other run.

Resting state images were collected with the following parameters:

Parameter	Value
Sequence	Gradient-echo EPI
TR	720 ms
TE	33.1 ms
flip angle	52 deg
FOV	208x180 mm (RO x PE)

Parameter	Value
Matrix	104x90 (RO x PE)
Slice thickness	2.0 mm; 72 slices; 2.0 mm isotropic voxels
Multiband factor	8
Echo spacing	0.58 ms
BW	2290 Hz/Px

Condition	Runs	Frames per run	Run Duration (min:sec)
REST (Resting-state)	4	1200	14:33

Task-evoked fMRI (tfMRI)

Following completion of rfMRI in each of the two fMRI scanning sessions, 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 fMRI scans collected. For each scan type, one run was acquired with right-to-left phase encoding, and a second run with left-to-right phase encoding (in-plane FOV [field of view] rotation obtained by inverting both the RO (readout) and PE [phase encoding] gradient polarity).

T-fMRI data were acquired with the same EPI pulse sequence parameters as R-fMRI, except for the run duration information listed below. There are seven tasks (14 tfMRI runs) totaling one hour of total tfMRI scan time, with 3 tasks collected in one session and the remaining 4 tasks collected in another session.

Task	Runs	Frames per run	Run Duration (min:sec)
Working Memory	2	405	5:01
Gambling	2	253	3:12
Motor	2	284	3:34
Language	2	316	3:57
Social Cognition	2	274	3:27
Relational Processing	2	232	2:56
Emotion Processing	2	176	2:16

Scans necessary for field mapping correction were also acquired in each of the fMRI scan sessions.

EV (Explanatory Variable) files included in the dataset provide a set of conditions (and their associated timing) that can be used in the analysis of each task. An EPRIME_TAB.txt (tab-delimited text spreadsheet) describes task parameters that can be used in various tfMRI analysis packages; see Task-fMRI files and protocol details.

Diffusion imaging (dMRI)

Parameter	Value
Sequence	Spin-echo EPI
TR	5520 ms
TE	89.5 ms
flip angle	78 deg
refocusing flip angle	160 deg
FOV	210x180 (RO x PE)
matrix	168x144 (RO x PE)
slice thickness	1.25 mm, 111 slices, 1.25 mm isotropic voxels
Multiband factor	3
Echo spacing	0.78 ms
BW	1488 Hz/Px
Phase partial Fourier	6/8
b-values	1000, 2000, and 3000 s/mm^2

Other details: Diffusion gradients are monopolar. Oblique axial acquisitions alternate between right-to-left and left-to-right phase encoding directions in consecutive runs. Image reconstruction uses SENSE1 multi-channel.

A full dMRI session includes 6 runs (each approximately 9 minutes and 50 seconds), representing 3 different gradient tables, with each table acquired once with right-to-left and left-to-right phase encoding polarities, respectively. Each gradient table includes approximately 90 diffusion weighting directions plus 6 b=0 acquisitions interspersed throughout each run. Diffusion weighting consisted of 3 shells of b=1000, 2000, and 3000 s/mm^2 interspersed with an approximately equal number of acquisitions on each shell within each run. The diffusion directions were obtained using a toolbox available from INRIA that returns uniformly distributed directions in multiple q-space shells. The directions are optimized so that every subset of the first M directions is also isotropic. References and the INRIA toolbox can be found at: http://www-sop.inria.fr/members/Emmanuel.Caruyer/q-space-sampling.php.

Full scanning protocols (PDF)

Complete scanning protocols for each imaging modality can be found in Appendix 1.

Summary of behavioral measures

HCP Subjects undergo many behavioral tests that are part of the NIH Toolbox battery and several Non-Toolbox behavioral measures. The Q1 data release includes only Non-NIH Toolbox behavioral measures.

Non-Toolbox measures

The Non-Toolbox behavioral measures included in this release are:

- Visual processing – Farnsworth color vision test*****, Mars contrast sensitivity test
- Personality - Costa and McRae Neuroticism/Extroversion/Openness Five Factor Inventory
- Self-regulation (delay discounting)
- Sustained attention
- Verbal episodic memory.
- Emotion processing
- Spatial processing
- Fluid intelligence
- Self-reported function

Detailed descriptions of these measures can be found in Details of Behavioral Measures.

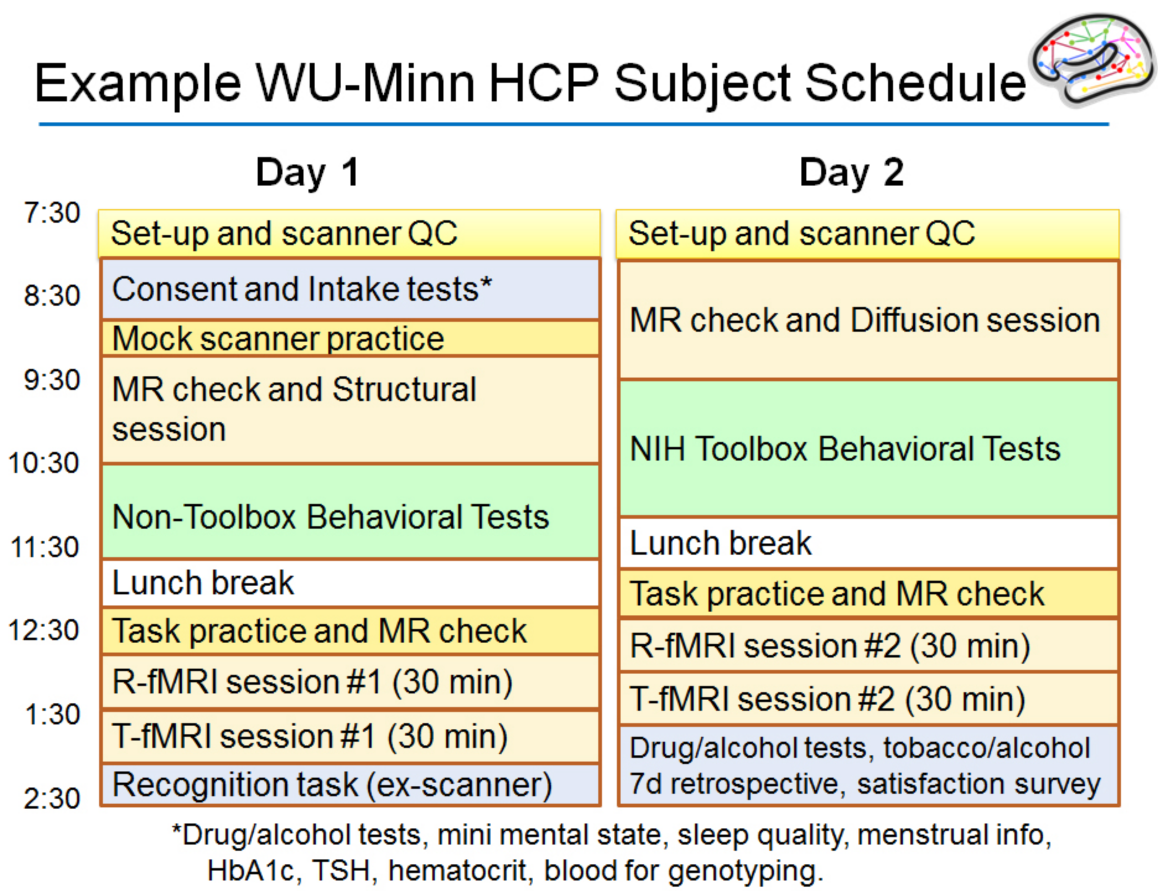
*****Released as part of the Q1 Restricted Access Data.

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 Appendix 4. 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 two-day schedule for subject visits

Most subjects complete the full HCP protocol during a two-day visit. A typical schedule is shown below. Scan sessions are always done in the order shown below, unless problems with a scan necessitate a rescan in an extra imaging session.



Data in this release

76 healthy adult subjects in the age range 22 – 35 participated in the first quarter of data collection. These include 68 subjects with data from all or nearly all modalities, 3 with structural MRI scans and behavioral tests, and 5 with behavioral tests only. (For several Q1 subjects, dMRI scans are not in the current release but may be included in a future release.) Subjects include both left- and right-handed individuals and both men and women.

Standard session structure

Each subject was scanned in four regular sessions (~4 hours total), with the following shorthand labels:

Day 1:

- Session 1: Structural scans (two T1w and two T2w scans)
- Session 2: Two 15-min R-fMRI scans (RL and LR phase encoding) and three T-fMRI tasks (one RL and one LR scan for each task)

Day 2:

- Session 3: dMRI scan
- Session 4: Two 15-min R-fMRI scans (RL and LR phase encoding) and the four remaining T-fMRI tasks (one RL and one LR scan for each task)

Note: Counterbalancing the ordering of the different phase-encoding acquisitions for the resting-state fMRI scans (RL followed by LR in the first fMRI session; LR followed by RL in the second session) was adopted on October 1, 2012. Prior to that, both sets of rfMRI scans were acquired using the RL followed by LR order. The temporal order of actual scan acquisitions for each subject is reflected in the scan number in ConnectomeDB (and also as an explicit ScanOrder variable).

In some cases, scans were acquired in an extra session (coded session 5 or higher as appropriate). This was done if the data quality was inadequate in the regular scan or if technical problems prevented a full set of acquisitions in the regular session.

Image Reconstruction and DICOM to NIFTI conversion

Image reconstruction of multiband images was implemented in ICE (the image reconstruction environment of the Siemens scanners) using a customized set of image reconstruction algorithms (see Ugurbil *et al.* 2013). While a fMRI or dMRI scan is ongoing, data are streamed to an offline ("remote") reconstruction computer for unaliasing of the multiband data, which streams the processed data back to the scanner for final processing and viewing. In some cases of technical problems, reconstruction was done at a later time using a "retro" reconstruction.

The reconstructed DICOM files are uploaded to the IntraDB database. Image defacing is carried out (Milchenko and Marcus, 2013) on all DICOM files containing potentially recognizable facial features. This includes structural (T1w, T2w), Bias receive and Bias transmit scans.

Conversion of DICOM files to NIFTI format was carried out using the dcm2nii utility. This utility is a component of the MRIcron suite of tools developed by Chris Rorden: (http://www.nitrc.org/frs/?group_id=152).

Initial QC and File Transfer to ConnectomeDB

Structural scans. An experienced rater has evaluated each structural scan and designated it as excellent, good, fair, poor, or unusable based on criteria related to tissue contrast, blurriness, or banding artifacts. To be included in the data release and in the structural pipeline processing, at least one T1w and one T2w rated as good or excellent must have been acquired in the same session (without the subject having exited the scanner). If these criteria are not met using the scans from the initial structural session, an extra scan session is acquired during the subject's visit or (if necessary) during a subsequent visit.

Usability of other scans. Scans from other modalities (fMRI and Diffusion) are included in the data release only if they meet specified criteria that make them useable in conjunction with other scans of the same modality acquired from the subject.

All scans are 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-specific directories (reflecting the fact that data are acquired in multiple scan sessions) 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 (e.g., bias fields and field maps) 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, again to facilitate preprocessing.

A subject-specific spreadsheet is included with the structural download packages, containing information that may be useful when analyzing the data. Appendix 5: Data Acquisition Information for an Exemplar Subject shows a representative example.

Directory structure for unprocessed data

Unprocessed datasets downloaded from ConnectomeDB are organized into subject-specific and modality-specific archives. When unpacked, they are moved into a <subject_id>/unprocessed/3T/ subdirectory (by creating it if not present, or placing into existing directories if present). A full list of file names in each directory and subdirectory is provided in Appendix 3: File Names and Directory Structure for Unprocessed and Preprocessed HCP Q1 Data, Section A.

Unprocessed data for exemplar subject 100307 unpacks to the following directory structure:

```
100307/unprocessed/3T/
        100307_3T.csv
        Diffusion
        rfMRI_REST1_LR
        rfMRI_REST1_RL
        rfMRI_REST2_LR
        rfMRI_REST2_RL
        T1w_MPR1
        T2w_SPC1
        tfMRI_EMOTION_LR
        tfMRI_EMOTION_RL
        tfMRI_GAMBLING_LR
        tfMRI_GAMBLING_RL
        tfMRI_LANGUAGE_LR
        tfMRI_LANGUAGE_RL
        tfMRI_MOTOR_LR
        tfMRI_MOTOR_RL
        tfMRI_RELATIONAL_LR
        tfMRI_RELATIONAL_RL
        tfMRI_SOCIAL_LR
        tfMRI_SOCIAL_RL
        tfMRI_WM_LR
        tfMRI_WM_RL
```

The 3T/ subdirectory signifies that these data were acquired on the 3T Connectome Skyra at Wash U. For the subjects that are later scanned at 7T (200 of the 1200), the 7T data will unpack to a 7T/ subdirectory. The .csv file contains the subject-specific spreadsheet mentioned in the preceding section, which provides useful metadata about the scans acquired for this particular subject.

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

```
T1w_MPR1/
        100307_3T_AFI.nii.gz
        100307_3T_BIAS_32CH.nii.gz
        100307_3T_BIAS_BC.nii.gz
        100307_3T_FieldMap_Magnitude.nii.gz
        100307_3T_FieldMap_Phase.nii.gz
        100307_3T_T1w_MPR1.nii.gz
```

T2w_SPC1/
- 100307_3T_AFI.nii.gz
- 100307_3T_BIAS_32CH.nii.gz
- 100307_3T_BIAS_BC.nii.gz
- 100307_3T_FieldMap_Magnitude.nii.gz
- 100307_3T_FieldMap_Phase.nii.gz
- 100307_3T_T2w_SPC1.nii.gz

The FieldMap, BIAS, and AFI scans are ancillary files copied into multiple structural subdirectories, in order to facilitate using standardized scripts and pipelines.

If multiple structural scans for a given subject passed the QC criteria and are included in ConnectomeDB, they will be in separate subdirectories T1w_MPR2/ and/or T2w_SPC2, with uniquely named structural files (e.g. 100307_3T_T1w_MPR2.nii.gz and/or 100307_3T_T2w_SPC2.nii.gz). Each subdirectory will include copies of relevant ancillary files needed for preprocessing.

The four 15-min rfMRI scans for each subject are downloaded into separate subdirectories, because the initial preprocessing is carried out separately for each scan.

rfMRI_REST1_LR
- 100307_3T_BIAS_32CH.nii.gz
- 100307_3T_BIAS_BC.nii.gz
- 100307_3T_SpinEchoFieldMap_LR.nii.gz
- 100307_3T_SpinEchoFieldMap_RL.nii.gz
- 100307_3T_rfMRI_REST1_LR.nii.gz
- 100307_3T_rfMRI_REST1_LR_SBRef.nii.gz

rfMRI_REST1_RL
- 100307_3T_BIAS_32CH.nii.gz
- 100307_3T_BIAS_BC.nii.gz
- 100307_3T_SpinEchoFieldMap_LR.nii.gz
- 100307_3T_SpinEchoFieldMap_RL.nii.gz
- 100307_3T_rfMRI_REST1_RL.nii.gz
- 100307_3T_rfMRI_REST1_RL_SBRef.nii.gz

Here, the ancillary files are SpinEchoFieldMap and BIAS scans acquired at the beginning of the rfMRI session and copied into both the rfMRI_REST1_LR/ and rfMRI_REST1_RL/ directories to simplfy preprocessing scripts.

For tfMRI scans, the 14 scans (7 tasks x two phase encoding directions) are downloaded into separate subdirectories, because the initial preprocessing is carried out separately for each scan. For example, the working memory (WM) LR phase encoding imaging data:

tfMRI_WM_LR
- 100307_3T_BIAS_32CH.nii.gz
- 100307_3T_BIAS_BC.nii.gz
- 100307_3T_SpinEchoFieldMap_LR.nii.gz
- 100307_3T_SpinEchoFieldMap_RL.nii.gz
- 100307_3T_tfMRI_WM_LR.nii.gz
- 100307_3T_tfMRI_WM_LR_SBRef.nii.gz

and the relevant E-Prime data are in appropriate subdirectories:

tfMRI_WM_LR/LINKED_DATA/EPRIME
 100307_3T_REC_run2_TAB.txt
 100307_3T_WM_run2_TAB.txt

tfMRI_WM_LR/LINKED_DATA/EPRIME/EVs
 0bk_body.txt
 0bk_cor.txt
 0bk_err.txt
 0bk_faces.txt
 0bk_nlr.txt
 etc.…

Note that inclusion of an ancillary file in any given directory does not necessarily mean that it is used as part of the current HCP preprocessing pipelines. For example, the BIAS and AFI scans are not used currently (for any modality) for the preprocessed data of the Q1 data release.

The 6 dMRI scans (3 b-values x two phase encoding directions) are in a single Diffusion/ subdirectory:

Diffusion/
 100307_3T_BIAS_32CH.nii.gz
 100307_3T_BIAS_BC.nii.gz
 100307_3T_DWI_dir95_LR_SBRef.nii.gz
 100307_3T_DWI_dir95_LR.bval
 100307_3T_DWI_dir95_LR.bvec
 100307_3T_DWI_dir95_LR.nii.gz
 100307_3T_DWI_dir95_RL_SBRef.nii.gz
 100307_3T_DWI_dir95_RL.bval
 100307_3T_DWI_dir95_RL.bvec
 100307_3T_DWI_dir95_RL.nii.gz
 100307_3T_DWI_dir96_LR_SBRef.nii.gz
 100307_3T_DWI_dir96_LR.bval
 100307_3T_DWI_dir96_LR.bvec
 100307_3T_DWI_dir96_LR.nii.gz
 100307_3T_DWI_dir96_RL_SBRef.nii.gz
 100307_3T_DWI_dir96_RL.bval
 100307_3T_DWI_dir96_RL.bvec
 100307_3T_DWI_dir96_RL.nii.gz
 100307_3T_DWI_dir97_LR_SBRef.nii.gz
 100307_3T_DWI_dir97_LR.bval
 100307_3T_DWI_dir97_LR.bvec
 100307_3T_DWI_dir97_LR.nii.gz
 100307_3T_DWI_dir97_RL_SBRef.nii.gz
 100307_3T_DWI_dir97_RL.bval
 100307_3T_DWI_dir97_RL.bvec
 100307_3T_DWI_dir97_RL.nii.gz

Preprocessing pipelines

Structural MR and fMRI datasets were processed by a set of 6 pipelines aimed at providing high quality volume and surface data. These pipelines use freely available software from the FSL (Jenkinson M *et al*., 2012; http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/) and FreeSurfer (Dale AM *et al*., 1999; http://surfer.nmr.mgh.harvard.edu/) and Connectome Workbench image analysis suites, and will be discussed in detail in a forthcoming special issue of *NeuroImage* (Glasser *et al*., submitted).

Each dataset downloaded via ConnectomeDB (see How to access HCP data) includes Release Notes that state the version number(s) of the relevant HCP pipelines used to process the data. For example, the archive 100307_3T_tfMRI_EMOTION_preproc.zip contains release-notes-tfMRI_EMOTION_preprocess.txt:

100307_3T_tfMRI_EMOTION_preproc.zip
    Structural Pipeline v1.0
    fMRI Pipeline v1.0
    Execution 1
    Fri Feb 23 08:21:00 CDT 2013

We anticipate making additional refinements to the HCP minimal preprocessing pipelines. When that occurs, all datasets, starting with the Q1 data, will be reprocessed using the revised pipelines and will include appropriately updated release notes.

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

A) Structural image processing pipeline

1. Gradient distortion correction.
2. Coregistration and averaging of T1w and T2w runs.
3. ACPC (i.e. 6 dof) registration for T1w and T2w (FLIRT plus a customized script). The acpc registered distortion corrected T1w image is in the “native” volume space.
4. Initial brain extraction for T1w and T2w (FNIRT-atlas-based-brain-mask).
5. Field map distortion correction and registration of T2w to T1w using a customized FLIRT BBR algorithm.
6. Bias field correction using sqrt(T1w X T2w) (Rilling *et al*., 2011).

7. Atlas registration (FLIRT linear + FNIRT nonlinear to MNI152).

B) FreeSurfer pipeline

1. Downsampling of 0.7mm T1w to 1mm using splines.
2. Initial FreeSurfer steps (autorecon1 except -skullstrip).
3. Initialize FreeSurfer with skull registration using PreFreeSurfer brain mask for a more robust registration.
4. FreeSurfer skullstripping.
5. Further early FreeSurfer processing steps (-autorecon2 -nosmooth2 -noinflate2).
6. Adjustment of white matter surface placement using hires T1w (full 0.7mm data) with a customized mri_normalize, mris_make_surfaces, and fine tuning of T2w to T1w registration using bbregister.
7. Middle FreeSurfer steps (-smooth2 -inflate2 -sphere -surfreg -jacobian_white -avgcurv –cortparc).
8. Creation of pial surface using hires T1w (full 0.7mm data) and adjustment of pial surface placement using hires T2w (full 0.7mm data) with a customized mris_make_surfaces to remove vessels and dura. Then grey matter intensity normalization of T1w, regeneration of pial surface and adjustment with T2w.
9. Final FreeSurfer steps (-surfvolume -parcstats -cortparc2 -parcstats2 -cortribbon -segstats -aparc2aseg -wmparc -balabels -label-exvivo-ec).

C) Post FreeSurfer processing pipeline

1. Creation of Caret5 and Connectome Workbench datafiles and spec files, creation of FreeSurfer segmentation brain mask, surface registration and downsampling, volume transformation of surfaces nonlinearly into MNI space.
2. Creation of FreeSurfer ribbon files with full 0.7mm resolution data (masks of grey matter and white matter).
3. Myelin mapping and combination of all transforms for one step resampling of T1w and T2w images from original to MNI space (Glasser and Van Essen, 2011).

D) Volume generic fMRI processing pipeline (Volume-based analysis starts at the end of this pipeline)

1. Gradient distortion correction.

2. FLIRT-based motion correction using the "SBRef" volume as the target (Smith SM *et al*., submitted).
3. TOPUP-based field map preprocessing using Spin echo field map (for each day of each BOLD run)
4. Distortion correction and EPI to T1w registration of the "SBRef" volume using a customized FLIRT BBR algorithm for distortion correction and bbregister to fine tune the registration. A Jacobian file representing the intensity modulations of the distortion correction is stored but not applied to the data.
5. One step spline resampling from the original EPI frames to atlas space including all transforms (motion, EPI distortion, EPI to T1w from FLIRT BBR, fine tuning of EPI to T1w with bbregister, nonlinear T1w to MNI).
6. Intensity normalization to mean of 10000 (like in FEAT) and bias field removal. Brain mask based on FreeSurfer segmentation.

E) Surface generic fMRI processing pipeline (Surface-based analysis starts at the end of this pipeline)

1. Cortical ribbon-based volume to surface mapping (from MNI space 2mm volume to MNI space native mesh surface).
    1. Voxels between white and pial surface are included in mapping. Voxels that are only partially between white and pial surface are weighted according to their partial volume that is between the two surfaces.
    2. Voxels with a temporal coefficient of variance greater than 0.5 standard deviations of their local neighborhood (sigma=5mm) are excluded from volume to surface mapping. The practical effect of this is to remove any small vessels and brain rim voxels remaining in the ribbon.
    3. Transformation of timeseries from native mesh to fs_LR registered 32k mesh (2mm average vertex spacing) in a single step.
2. Surface-based smoothing (2mm FWHM).
3. Subcortical parcel-constrained smoothing and atlas resampling (2mm FWHM).
4. Creation of dense timeseries (on a standard set of brainordinates).

F) Diffusion processing pipeline

1. Basic preprocessing: Intensity normalization across runs, preparation for later modules.
2. 'TOPUP' algorithm for EPI distortion correction.

3. ‘EDDY’ algorithm for eddy current and motion correction.
4. Gradient nonlinearity correction, calculation of gradient bvalue/bvector deviation.
5. Registration of mean b0 to native volume T1w with FLIRT BBR+bbregister, and generation of diffusion to MNI transform.

Note: The next release will put diffusion data into a 1.25mm space rigidly aligned with the structural data.

Directory structure for processed data

Preprocessing generates thousands of files, many of which are of little or no use. Therefore, HCP releases a subset of files that are likely to be of general use to investigators. A list of file names in each directory and subdirectory is provided in Appendix 3B.

Structural data

- T1w/ contains T1w and T2w volume data
- T1w/Native/ contains FreeSurfer surfaces in their native mesh and original dimensions after rigid-body rotation to AC-PC alignment.
- MNINonLinear/ contains cortical surfaces and other data volumetrically registered to MNI152 space (using nonlinear FNIRT) followed by surface registration to Conte69 '164k_fs_LR' mesh (Van Essen *et al*. 2011) (via FreeSurfer fsaverage as an intermediate).
- MNINonLinear/Native/ replicates some of the files in T1w/Native/ but contains additional files used during surface-based registration.
- MNINonLinear/xfms/ contains files encoding the transformation between acpc and MNINonLinear volumetric space
- MNINonLinear/fsaverage_LR32k contains files spatially downsampled to a 32k mesh (average vertex spacing of ~2 mm), which is useful for analyses of rfMRI and dMRI connectivity data.

fMRI data

- MNINonLinear/Results/ contains volumetric and CIFTI grayordinates data for rfMRI scans (15 min each) and motion parameters in four subdirectories,
    - rfMRI_REST1_RL
    - rfMRI_REST1_LR
    - rfMRI_REST2_RL
    - rfMRI_REST2_LR

  plus volumetric and CIFTI grayordinates data for 7 pairs of T-fMRI scans (each task run once with right-to-left and once with left-to-right phase encoding.

    - tfMRI_EMOTION_RL
    - tfMRI_EMOTION_LR
    - tfMRI_GAMBLING_RL
    - tfMRI_GAMBLING_LR

- o tfMRI_LANGUAGE_RL
- o tfMRI_LANGUAGE_LR
- o tfMRI_MOTOR_RL
- o tfMRI_MOTOR_LR
- o tfMRI_RELATIONAL_RL
- o tfMRI_RELATIONAL_LR
- o tfMRI_SOCIAL_RL
- o tfMRI_SOCIAL_LR
- o tfMRI_WM_RL
- o tfMRI_WM_LR

Each of the directories above will contain an .fsf file and a "EVs" directory containing explanatory variables. These files can be used to run lower-level analyses in FSL 5. The directories above will also contain a TAB.txt file containing the timing of events, so that researchers might create their own explanatory variables of interest.

The MNINonLinear/Results/ directory will contain seven other directories, one for each task:

- o tfMRI_EMOTION
- o tfMRI_GAMBLING
- o tfMRI_LANGUAGE
- o tfMRI_MOTOR
- o tfMRI_RELATIONAL
- o tfMRI_SOCIAL
- o tfMRI_WM

These directories contain an .fsf file that can be used to run a higher-level analysis across the two runs of each task. They also contain a script (prepare_level2_feat_analysis.sh) that must be run first, to set up appropriate registration matrices in the lower-level outputs.

- **Motion parameters.** Estimates of motion parameters are saved into two different files: Movement_Regressors.txt and Movement_Regressors_dt.txt. The first file (Movement_Regressors.txt) contains 12 variables. The first six variables are the motion parameters estimates from a rigid-body transformation to the SBRef image acquired at the start of each fMRI scan.

  - o trans_x (mm)
  - o trans_y (mm)
  - o trans_z (mm)
  - o rot_x (deg)
  - o rot_y (deg)
  - o rot_z (deg)

The second six variables are temporal derivatives of those motion parameters

- trans_dx
- trans_dy
- trans_dz
- rot_dx
- rot_dy
- rot_dz

The second file (Movement_Regressors_dt.txt) contains 12 variables derived by removing the mean and linear trend from each variable in Movement_Regressors.txt

Diffusion Data

- Diffusion data includes diffusion weighting, direction, time series, brain mask, and gradient nonlinearity data, with the following files and directory structure:
  - Diffusion/data/bvals
    (contains the diffusion weighting (b-value) for each volume)
  - Diffusion/data/bvecs
    (contains the diffusion direction (b-vector) for each volume)
  - Diffusion/data/data.nii.gz
    (preprocessed diffusion time series file)
  - Diffusion/data/nodif_brain_mask.nii.gz
    (brain mask in diffusion space)
  - Diffusion/data/grad_dev.nii.gz
    (contains the effects of gradient nonlinearities on the bvals and bvecs for each voxel)
- Transforms for diffusion analysis are contained in the following files:
  - T1w/xfms/diff2str.mat
    (diffusion space to native [structural] volume space 6 degrees of freedom [dof] transform)
  - T1w/xfms/str2diff.mat
    (native [structural] volume space to diffusion space 6 dof transform)
  - MNINonLinear/xfms/diff2standard.nii.gz
    (diffusion space to MNI volume space nonlinear transform)
  - MNINonLinear/xfms/standard2diff.nii.gz
    (MNI volume space to diffusion space nonlinear transform)

Additionally processed Group-average Data

A number of processing pipelines are currently being implemented and refined by the HCP that make use of the minimally processed data. The HCP Q1 Data Release includes one freely available dataset based on group-average data acquired from 20 unrelated subjects (HCP_Q1_GroupAvgUnrelated20.zip, available by clicking "Download Group Average" under the Unrelated 20 section of the HCP Open Access Data Releases page). It includes structural scans (T1w, T2w), group-average cortical surfaces and myelin maps, and group-average Task-fMRI results for 7 tasks and many task contrasts.

The dataset is customized for viewing in Connectome Workbench (http://www.humanconnectome.org/connectome/get-connectome-workbench.html), but most of the files are in standard NIFTI and GIFTI format and can also be viewed using other brain-mapping software platforms. Some files, including two of the Task-fMRI files, are in a relatively new CIFTI file format (http://www.nitrc.org/projects/cifti/) that can incorporate surface vertices (both hemispheres) plus subcortical gray-matter voxels into a single file format.

Task-fMRI group analysis. Full details are described in Barch *et al*., (2013, submitted). Each of the seven tasks was processed using initial volume-based smoothing and the other using 'grayordinates-constrained' smoothing that respects the topology of the cortical sheet and the boundaries of subcortical anatomical parcels.

For the volume-based analysis, spatial smoothing was applied using an unconstrained 3D Gaussian kernel of FWHM=4mm. Activity estimates were computed for the preprocessed functional time series from each run using a general linear model (GLM) implemented in FSL's FILM (Woolrich *et al.*, 2001). Predictors were convolved with a double gamma "canonical" hemodynamic response function (Glover, 1999) to generate the main model regressors. To compensate for slice-timing differences and variability in the HRF delay across regions, temporal derivative terms derived from each predictor were added to each GLM and were treated as confounds of no interest. Subsequently, both the 4D time series and the GLM design were temporally filtered with a Gaussian-weighted linear highpass filter with a (soft) cutoff of 200 s. Finally, the time series was prewhitened within FILM to correct for autocorrelations in the fMRI data.

The grayordinates-based analysis began with outputs of the HCP "fMRISurface" pipeline (see above) in which the data from the cortical gray matter ribbon were projected onto the surface and then onto registered surface meshes with a standard number of vertices. Subcortical data were also projected to a set of subcortical grey matter parcel voxels, and when combined with the surface data formed the standard grayordinates space (Glasser *et al*. 2013, submitted). The grayordinates-based run-level analysis was carried out identically to the volume-based analysis described above aside from spatial smoothing steps. Smoothing of the left and right hemisphere time series and autocorrelation estimates (from FILM) were done on the surface using a geodesic Gaussian algorithm. Subcortical gray matter time series were smoothed within defined gray matter parcels. Smoothing of the grayordinate data was done in two steps, but for a total of 4mm FWHM. Surface-based autocorrelation estimate smoothing was incorporated into FSL's FILM at a sigma of 5mm.

The GLM Model Design for each task is described in Barch *et al.* (2013). Fixed-effects analyses were conducted using FSL's FEAT to estimate the average effects across runs within-subjects. Mixed-effects analyses treating subjects as random effects were conducted using FSL's FLAME to estimate the average effects of interest for the group. Volume-based group-level analyses were carried out using voxelwise comparisons in MNI space. The grayordinates-based participant-level and group-level analyses were done identically to the volume-based analysis except that cross-run and cross-subject statistical comparisons occurred in the better-aligned standard grayordinates space. As in the individual analysis, NIFTI-1 matrices were processed separately for left and right surface and subcortical volume data, and surface outputs were converted to GIFTI at the conclusion of analysis. Subject-level and group-level z-statistic maps were combined from left and right hemisphere cortical and subcortical gray matter into the recently introduced CIFTI data format (http://www.nitrc.org/projects/cifti/; see (Glasser *et al.* 2013, submitted) for visualization using the Connectome Workbench platform (see Marcus *et al.* 2013, submitted).

Group-average functional connectivity. An analysis of resting-state fMRI data from the same HCP Q1 unrelated 20 subjects has yielded two versions of a group-average functional connectivity matrix ("dense" functional connectome). These dense connectomes are not part of the downloadable Q1 dataset, because the methods are still under refinement and have not been finalized. Nonetheless, the data are accessible by interactive remote access to the ConnectomeDB database (see "Connectome Workbench Tutorial for the Q1 Data Release"; http://www.humanconnectome.org/documentation/tutorials/)

Each of the four resting state runs was cleaned (denoised) using 24-parameter motion regression followed by ICA+FIX denoising (Smith *et al.* 2013, submitted). The resulting cleaned runs were combined across the 20 unrelated subjects using a novel iterative group PCA procedure to reduce the dimensionality of the data from 96,000 timepoints to 4,500 PCA components. This set of PCA components was further reduced to 250 independent components, which were then reconstructed into the original PCA space to remove residual Gaussian noise. One version of the dense connectome uses the full correlation matrix. For the other, the mean gray signal was removed prior to grayordinatewise cross-correlation before generating the dense connectome.

File Names and Directory Structure for Additionally Processed Data

The additionally processed dataset (HCP_Q1_GroupAvgUnrelated20.zip) unpacks to HCP_Q1_GroupAvgUnrelated20/. It includes four sets of files:

- Group-average surface and volume data generated from the T1w and T2w images:
  - HCP_PhaseII_Q1_Unrelated20.L.curvature.32k_fs_LR.shape.gii
  - HCP_PhaseII_Q1_Unrelated20.L.inflated.32k_fs_LR.surf.gii
  - HCP_PhaseII_Q1_Unrelated20.L.midthickness.32k_fs_LR.surf.gii
  - HCP_PhaseII_Q1_Unrelated20.L.pial.32k_fs_LR.surf.gii
  - HCP_PhaseII_Q1_Unrelated20.L.sphere.32k_fs_LR.surf.gii
  - HCP_PhaseII_Q1_Unrelated20.L.sulc.32k_fs_LR.shape.gii
  - HCP_PhaseII_Q1_Unrelated20.L.thickness.32k_fs_LR.shape.gii
  - HCP_PhaseII_Q1_Unrelated20.L.very_inflated.32k_fs_LR.surf.gii
  - HCP_PhaseII_Q1_Unrelated20.L.white.32k_fs_LR.surf.gii
  - HCP_PhaseII_Q1_Unrelated20.MyelinMap_BC.32k_fs_LR.dtseries.nii
  - HCP_PhaseII_Q1_Unrelated20.R.curvature.32k_fs_LR.shape.gii
  - HCP_PhaseII_Q1_Unrelated20.R.inflated.32k_fs_LR.surf.gii
  - HCP_PhaseII_Q1_Unrelated20.R.midthickness.32k_fs_LR.surf.gii
  - HCP_PhaseII_Q1_Unrelated20.R.pial.32k_fs_LR.surf.gii
  - HCP_PhaseII_Q1_Unrelated20.R.sphere.32k_fs_LR.surf.gii
  - HCP_PhaseII_Q1_Unrelated20.R.sulc.32k_fs_LR.shape.gii
  - HCP_PhaseII_Q1_Unrelated20.R.thickness.32k_fs_LR.shape.gii
  - HCP_PhaseII_Q1_Unrelated20.R.very_inflated.32k_fs_LR.surf.gii
  - HCP_PhaseII_Q1_Unrelated20.R.white.32k_fs_LR.surf.gii
  - HCP_PhaseII_Q1_Unrelated20.SmoothedMyelinMap_BC.32k_fs_LR.dscalar.nii
  - HCP_PhaseII_Q1_Unrelated20_AverageT1w_restore.nii.gz
  - HCP_PhaseII_Q1_Unrelated20_AverageT2w_restore.nii.gz

- Group-average Task-fMRI files (CIFTI and NIFTI format):
  - **TaskFMRI_HCP_Q1_GroupAvg_unrelated20.dscalar.nii** is a CIFTI file in which spatial smoothing (4mm FWHM) was constrained to 'grayordinates' (cortical surface vertices and subcortical voxels). By avoiding spatial blurring across anatomical compartments (gray matter vs white matter and CSF) and across sulcal banks, this provides the highest spatial fidelity attainable using these data.
  - **TaskFMRI_HCP_Q1_GroupAvg_unrelated20_vol-smooth.dscalar.nii** is also in CIFTI format, but spatial smoothing (4mm FWHM) was done in standard

volume space that includes blurring across sulcal banks and between gray-matter and non-gray-matter voxels.

- **TaskFMRI_HCP_Q1_GroupAvg_unrelated20_vol-smooth.nii.gz** is the same as the preceding file but was converted from CIFTI to NIFTI format to enable viewing with platforms that are not yet CIFTI-compliant

- Reference files generated from other studies:
  - parcellations_VGD11b.L.32k_fs_LR.border
  - parcellations_VGD11b.L.32k_fs_LR.label.gii
  - parcellations_VGD11b.R.32k_fs_LR.border
  - parcellations_VGD11b.R.32k_fs_LR.label.gii
  - Parcels_GV11.L.32k_fs_LR.foci
  - Parcels_GV11.R.32k_fs_LR.foci
  - RSN-networks.L.32k_fs_LR.label.gii
  - RSN-networks.R.32k_fs_LR.label.gii

The parcellations_VGD11b.*.32k_fs_LR.* files are derived from a composite cortical parcellation containing 52 distinct areas accurately mapped to the fs_LR atlas surface and based on architectonic or retinotopic fMRI maps (Van Essen *et al.* 2011). Abbreviations associated with labels of cortical areas and border classes (FRB08, OFP03, etc.) refer to the publication that defined a particular cortical area in the composite map (cf. Table 3 of Van Essen *et al*. 2011).

The Parcels_GV11.*.32k_fs_LR.foci files are for display of foci located at the center of mass of the probabilistic cytoarchitectonic areas (originally defined by Amunts and Zilles) used by Glasser and Van Essen 2011 for comparison to areas defined by myelin mapping.

More information on the RSN-networks label files will be available in the next HCP data release.

- Miscellaneous files:
  - HCP_Q1_Unrelated20_DATA.32k_fs_LR.wb.spec
  - HCP_Q1_Unrelated20_demo.scene
  - ReleaseNotes_HCP_Q1_GroupAvgUnrelated20.txt
  - AllTaskContrasts.txt

Task-fMRI files and protocol details

Task-Evoked Functional Brain Activity

We assessed seven major domains that we think sample the diversity of neural systems that will be of interest to a wide range of individuals in the field, including: 1) visual, motion, somatosensory, and motor systems; 2) category specific representations; 3) working memory/cognitive control systems; 4) language processing (semantic and phonological processing); 5) social cognition (Theory of Mind); 6) relational processing; and 7) emotion processing. These tasks are described in more detail below. Stimuli were projected onto a computer screen behind the subject's head within the imaging chamber. The screen was viewed by a mirror positioned approximately 8 cm above the subject's face.

tfMRI scripts and data files

Script files in E-Prime are used to present stimuli and collect behavioral responses in the scanner. If you would like to run HCP tasks in your own research project, these files can be obtained by contacting Greg Burgess via email at burgessg@pcg.wustl.edu, please put "HCP E-Prime scripts" in the subject line of your message.

Tab-delimited versions of the E-Prime data files (TAB.txt) are included in this release. TAB.txt files are named according to the task condition that they describe and are contained within the directories for each of the two runs within each task (each phase encoding direction). A brief description of the key variables in those files can be found in Appendix 6: Task fMRI E-Prime Key Variables. The original edat files will not be available, because they may contain identifying information.

In addition to the TAB.txt files defined above, this release contains EV .txt files derived from those TAB.txt files. EV files are explanatory variables (predictors) in FSL format (3-columns: onset, duration, and amplitude). There is a separate EV directory for each of the two runs within each task. Examples of the EV files for each task are detailed below.

The release also includes .fsf files for each task. The fsf file is the setup or configuration file for running GLM-based fMRI analyses in 'FEAT' (FMRIB's Expert Analysis Tool: http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FEAT). The Lev1 fsf files contain setup information necessary to run GLM analyses on the timeseries data for an individual scan run. Lev2 fsf files contain setup information to run GLM analyses combining multiple scan runs for an individual participant. Lev3 fsf files (not included in the release) can be created to setup GLM analyses across multiple participants.

Here are examples of the EV files for each task and phase encoding direction in the appropriate tfMRI_[task_phaseencodingdirection]/LINKED_DATA/EPRIME directory (e.g. tfMRI_WM_LR/LINKED_DATA/EPRIME):

Working Memory

EVs/0bk_body.txt	BLOCKED	onset of 0Back body block condition
EVs/0bk_faces.txt	BLOCKED	onset of 0Back faces block condition
EVs/0bk_places.txt	BLOCKED	onset of 0Back places block condition
EVs/0bk_tools.txt	BLOCKED	onset of 0Back tools block condition
EVs/2bk_body.txt	BLOCKED	onset of 2Back body block condition
EVs/2bk_faces.txt	BLOCKED	onset of 2Back faces block condition
EVs/2bk_places.txt	BLOCKED	onset of 2Back places block condition
EVs/2bk_tools.txt	BLOCKED	onset of 2Back tools block condition
EVs /0bk_cor.txt	EVENT	onset of correct trials in 0Back blocks
EVs /0bk_err.txt	EVENT	onset of error trials in 0Back blocks
EVs /0bk_nlr.txt	EVENT	onset of trials in 0Back blocks with no response
EVs /2bk_cor.txt	EVENT	onset of correct trials in 2Back blocks
EVs /2bk_err.txt	EVENT	onset of error trials in 2Back blocks
EVs /2bk_nlr.txt	EVENT	onset of trials in 0Back blocks with no response
EVs /all_bk_cor.txt	EVENT	onset of correct trials in 0- and 2Back blocks
EVs/all_bk_err.txt	EVENT	onset of error trials in both 0- and 2Back blocks

Gambling

EVs/win.txt	BLOCKED	Onset of mostly reward blocks
EVs /loss.txt	BLOCKED	Onset of mostly loss blocks
EVs /win_event.txt	EVENT	Onset of reward trials
EVs /loss_event.txt	EVENT	Onset of loss trials
EVs /neutral.txt	EVENT	Onset of neutral trials

Motor

EVs /cue.txt	BLOCKED	Onset of task cues
EVs /lf.txt	BLOCKED	Onset of left foot blocks
EVs /rf.txt	BLOCKED	Onset of right foot blocks
EVs /lh.txt	BLOCKED	Onset of left hand blocks
EVs /rh.txt	BLOCKED	Onset of right hand blocks
EVs /t.txt	BLOCKED	Onset of tongue blocks

Language

EVs /story.txt	BLOCKED	Onset of story blocks
EVs /math.txt	BLOCKED	Onset of math blocks

Social Cognition

EVs /mental.txt	BLOCKED	Onset of mental interaction blocks
EVs /rnd.txt	BLOCKED	Onset of random interaction blocks
EVs /mental_resp.txt	EVENT	Onset of trials rated as mental interaction
EVs /other_resp.txt	EVENT	Onset of trials not rated as mental interaction

Relational Processing

EVs /relation.txt	BLOCKED	Onset of relational blocks
EVs /match.txt	BLOCKED	Onset of match blocks

Emotion Processing

EVs /fear.txt	BLOCKED	Onset of emotional face blocks
EVs /neut.txt	BLOCKED	Onset of shape blocks

Details of tfMRI tasks

Working Memory

The category specific representation task and the working memory task are combined into a single task paradigm. Participants were presented with blocks of trials that consisted of pictures of places, tools, faces and body parts (non-mutilated parts of bodies with no “nudity”). Within each run, the 4 different stimulus types were 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). A 2.5 second cue indicates the task type (and target for 0-back) at the start of the block. Each of the two runs contains 8 task blocks (10 trials of 2.5 seconds each, for 25 seconds) and 4 fixation blocks (15 seconds). On each trial, the stimulus is presented for 2 seconds, followed by a 500 ms inter-task interval (ITI).

Conditions (Blocked)	
0-back faces	2-back faces
0-back places	2-back places

0-back tools	2-back tools
0-back body parts	2-back body parts

Conditions (Event-Related)	
0-back correct trials	2-back correct trials
0-back error trials	2-back error trials
0-back no response trials	2-back no response trials

Additional Contrasts. These event types can be combined to create two categories of contrasts.

Working Memory Contrasts
0-back contrast (activity combined across conditions 1-4)
2-back contrast (activity combined across conditions 5-8)
2-back versus 0-back contrast (2-back contrast minus 0-back contrast)

Category Contrasts	
Faces contrast	(0-back faces plus 2-back faces)
Places contrast	(0-back places plus 2-back places)
Tools contrast	(0-back tools plus 2-back tools)
Body contrast	(0-back body plus 2-back body)

Potential Additional Event Related Contrasts: Researchers can also use the TAB.txt E-Prime data files to generate the following potential event-related contrasts:

1. Targets
    a. For 2-back tasks, targets are 2-back repeats
    b. For 0-back tasks, targets match the cue stimulus
2. Non-targets
    a. For 2-back tasks, non-targets are novel items
    b. For 0-back tasks, non-targets do not match the cue stimulus
3. Lures
    a. For 2-back tasks, lures are 1-back or 3-back repeats
    b. For 0-back tasks, lures are repeated stimuli that do not match the cue stimulus

Gambling

This task was adapted from the one developed by Delgado and Fiez (Delgado *et al*. 2000). Participants play a card guessing game where they are asked to guess the number on a mystery card (represented by a “?”) in order to win or lose money. Participants are told that potential card numbers range from 1-9 and to indicate if they think the mystery card number is more or less than 5 by pressing one of two buttons on the response box. Feedback is the number on the card (generated by the program as a function of whether the trial was a reward, loss or neutral trial) and either: 1) a green up arrow with “$1” for reward trials, 2) a red down arrow next to -$0.50 for loss trials; or 3) the number 5 and a gray double headed arrow for neutral trials. The “?” is presented for up to 1500 ms (if the participant responds before 1500 ms, a fixation cross is displayed for the remaining time), following by feedback for 1000 ms. There is a 1000 ms ITI with a “+” presented on the screen. The task is presented in blocks of 8 trials that are either mostly reward (6 reward trials pseudo randomly interleaved with either 1 neutral and 1 loss trial, 2 neutral trials, or 2 loss trials) or mostly loss (6 loss trials pseudo-randomly interleaved with either 1 neutral and 1 reward trial, 2 neutral trials, or 2 reward trials). In each of the two runs, there are 2 mostly reward and 2 mostly loss blocks, interleaved with 4 fixation blocks (15 seconds each).

Conditions (Blocked)
Mostly reward blocks
Mostly loss blocks

Conditions (Event-Related)
Reward trials
Loss trials
Neutral trials

References for Gambling Task: Reliable across subjects and robust activation in fMRI (Delgado *et al*. 2000; May *et al*. 2004; Tricomi *et al*. 2004; Forbes *et al*. 2009)

Motor

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 fingers, or squeeze their left or right toes, or move their tongue to map motor areas. Each block of a movement type lasted 12 seconds (10 movements), and is preceded by a 3 second cue. In each of the two runs, there are 13 blocks, with 2 of tongue movements, 4 of hand movements (2 right and 2 left), and 4 of foot movements (2 right and 2 left). In addition, there are 3 15-second fixation blocks per run. This task contains the following events, each of which is computed against the fixation baseline.

Conditions (Blocked)
Left finger blocks

Right finger blocks
Left toe blocks
Right toe blocks
Tongue movement

References for Motor Task: Localizer (Morioka *et al*. 1995; Bizzi *et al*. 2008; Buckner *et al*. 2011; Yeo *et al*. 2011).

Language Processing

This task was developed by Binder and colleagues (Binder *et al.* 2011) and uses the E-prime scripts provided by these investigators. The task consists of two runs that each interleave 4 blocks of a story task and 4 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 3.8 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 auditorially 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. For more details on the task, please see (Binder *et al.* 2011).

Conditions (Blocked)
Story
Math

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

Social Cognition (Theory of Mind)

Participants were presented with short video clips (20 seconds) of objects (squares, circles, triangles) that either interacted in some way, or moved randomly on the screen. These videos were developed by either Castelli and colleagues (Castelli *et al.* 2000) or Martin and colleagues (Wheatley *et al.* 2007). After each video clip, participants judge whether the objects had a mental interaction (an interaction that appears as if the shapes are taking into account each other's feelings and thoughts), Not Sure, or No interaction (i.e., there is no obvious interaction between the shapes and the movement appears random). Each of the two task runs has 5 video blocks (2 Mental and 3 Random in one run, 3 Mental and 2 Random in the other run) and 5 fixation blocks (15 seconds each).

Conditions (Blocked)
Random interaction
Mental interaction

References for the Social Cognition Task: Reliable across subjects and robust activation (Castelli *et al*. 2000; Castelli *et al.* 2002; Wheatley *et al.* 2007; White *et al.* 2011).

Relational Processing

This task was adapted from the one developed by Christoff and colleagues (Smith *et al.* 2007). The stimuli are 6 different shapes filled with 1 of 6 different textures. In the relational processing condition, participants are presented with 2 pairs of objects, with one pair at the top of the screen and the other pair at the bottom of the screen. They are told that they should first decide what dimension differs across the top pair of objects (differed in shape or differed in texture) and then they should decide whether the bottom pair of objects also differ along that same dimension (e.g., if the top pair differs in shape, does the bottom pair also differ in shape). In the control matching condition, participants are shown two objects at the top of the screen and one object at the bottom of the screen, and a word in the middle of the screen (either "shape" or "texture"). They are told to decide whether the bottom object matches either of the top two objects on that dimension (e.g., if the word is "shape", is the bottom object the same shape as either of the top two objects. For both conditions, the subject responds yes or no using one button or another. For the relational condition, the stimuli are presented for 3500 ms, with a 500 ms ITI, and there are four trials per block. In the matching condition, stimuli are presented for 2800 ms, with a 400 ms ITI, and there are 5 trials per block. Each type of block (relational or matching) lasts a total of 18 seconds. In each of the two runs of this task, there are 3 relational blocks, 3 matching blocks and 3 16-second fixation blocks.

Conditions (Blocked)
Relational processing
Matching

References for the Relational Processing Task: Localizer (Smith *et al.* 2007).

Emotion Processing

This task was adapted from the one developed by Hariri and colleagues (Smith *et al.* 2007). Participants are presented with blocks of trials that either ask them to decide which of two faces presented on the bottom of the screen match the face at the top of the screen, or which of two shapes presented at the bottom of the screen match the shape at the top of the screen. The faces have either an angry or fearful expression. Trials are presented in blocks of 6 trials of the same task (face or shape), with the stimulus presented for 2000 ms and a 1000 ms ITI. Each block is preceded by a 3000 ms task cue ("shape" or "face"), so that each block is 21 seconds including the cue. Each of the two runs includes 3 face blocks and 3 shape blocks, with 8 seconds of fixation at the end of each run.

Conditions (Blocked)
Face
Shape

Note: A bug was written into the E-prime script for the EMOTION task, such that the task stopped short of the last three trials of the last task block in each run. This bug was not discovered until data had been collected on several participants. Consequently, the BOLD images and E-Prime data for the EMOTION task are shorter than our original design described above.

References for the Emotion Processing Task: Localizer (Hariri *et al.* 2002); Moderate reliability across time (Manuck *et al*. 2007).

Details of behavioral measures

NIH Toolbox behavioral measures

We are collecting measures developed for the NIH Toolbox to assess several domains. NIH Toolbox behavioral measure data will be described in detail and included in future releases. NIH Toolbox measures we are collecting as part of the HCP include:

Domain	Subdomain (Measure Name)
Cognition	Episodic Memory (Picture Sequence Memory)
	Executive Function/Cognitive Flexibility (Dimensional Change Card Sort)
	Executive Function/Inhibition (Flanker Task)
	Language/Vocabulary Comprehension (Picture Vocabulary)
	Processing Speed (Pattern Completion Processing Speed)
	Working Memory (List Sorting)
	Language/Reading Decoding (Oral Reading Recognition)
Emotion*	Negative Affect (Sadness, Fear, Anger)
	Psychological Well-being (Positive Affect, Life Satisfaction, Meaning and Purpose)
	Social Relationships (Social Support, Companionship, Social Distress, Positive Social Development)
	Stress and Self Efficacy (Perceived Stress, Self-Efficacy)
Motor	Dexterity (9-hole Pegboard)
	Endurance (2 minute walk test)
	Locomotion (4-meter walk test)
	Strength (Grip Strength Dynamometry)
Sensory	Audition (Words in Noise)
	Olfaction (Odor Identification Test)
	Taste (Taste Intensity Test)
	Pain (Pain Intensity and Interference Surveys)

***** All emotion measures and the pain measures are self-report.
From Barch *et al.* 2013 (submitted), Table 2.

Details on these NIH Toolbox behavioral measures can be found at: http://www.nihtoolbox.org. NIH Toolbox data are not ready for inclusion in the Q1 release, but should be available for all subjects by the time of the Q2 release (May 2013).

Non-NIH Toolbox behavioral measures

We collect additional measures to assess several domains not covered by the NIH Toolbox.

Additional Behavioral and Individual Difference Measures Included in the HCP

Domain	Subdomain (Measure Name)
Visual Processing	Visual Acuity (Electronic Visual Acuity System)
	Color Vision (Farnsworth Test)
	Contrast Sensitivity (Mars Contrast Sensitivity)
Personality	Five Factor Model (NEO-FFI)
Cognition	Self-regulation/Impulsivity (Delay Discounting)
	Sustained Attention (Short Penn Continuous Performance Test)
	Verbal Episodic Memory (Penn Word Memory Test)
	Spatial Orientation (Variable Short Penn Line Orientation Test)
	Fluid Intelligence (Penn Progressive Matrices)
Emotion	Emotion Recognition (Penn Emotion Recognition Test)
Psychiatric, Substance Abuse, and Life Function	Life Function (Achenbach Adult Self-Report)
	Psychiatric Clinical Symptoms (Semi-Structured Assessment for the Genetics of Alcoholism)
	Nicotine Dependence (Fagerstrom Test for Nicotine Dependence)
	Current Substance Use (breathalyzer, Urine Drug Screen, Self-Report)
Physical Function	Hematocrit Levels
	Menstrual Cycle and Hormonal Status
	Thyroid Function (Thyroid Stimulating Hormone Levels)
	Glucose Function (Hemoglobin A1c)
Other	Cognitive Status (Mini Mental Status Exam)
	Sleep (Pittsburgh Sleep Questionnaire)

From Barch *et al.* 2013 (submitted), Table 3.

Visual Processing

The NIH Toolbox does not measure color vision or contrast sensitivity. Thus, we are assessing color vision using the Farnsworth Test, a valid and reliable measure that provides more quantitative information than the commonly used Ishihara Test (Cole *et al*. 2007). In this task, participants order 15 colored blobs as a function of what they think are the closest matching colors. Based on the results, participants are classified as having Normal color vision, Protan (reduced sensitivity to red light), Deutan (reduced sensitivity to green light) or Tritian (reduced sensitivity to blue light) color vision problems. We are assessing contrast sensitivity using the

Mars Contrast Sensitivity Test (Arditi *et al*. 2005), a brief, valid and reliable measure that improves upon the traditional Pelli-Robson measure (Dougherty *et al.* 2005; Haymes *et al.* 2006; Thayaparan *et al.* 2007).

Personality

There is consensus that a five factor model captures the major facets of human personality across cultures (Heine and Buchtel 2009): a) neuroticism; b) extroversion/introversion; c) agreeableness; d) openness; and e) conscientiousness (Goldberg 1993; McCrae and Costa 2008). We are administering the 60 item version of the Costa and McRae Neuroticism/Extroversion/Openness Five Factor Inventory (NEO-FFI), which has shown excellent reliability and validity (McCrae and Costa 2004). This measure was available as part of the Penn Computerized Cognitive Battery (Gur *et al*. 2001a; Gur *et al.* 2010).

Self-Regulation (Delay Discounting)

Delay discounting describes the undervaluing of rewards that are delayed in time. It is illustrated by the fact that humans (and other animals) will often choose a smaller immediate reward over an objectively larger, but delayed reward. We use a version of the discounting task that identifies ‘indifference points’ at which a person is equally likely to choose a smaller reward (e.g., $100) sooner versus a larger reward later (e.g., $200 in 3 years). Based on the work of Green and Myerson (Estle *et al*. 2006; Green *et al*. 2007), we use an adjusting-amount approach, in which delays are fixed and reward amounts are adjusted on a trial-by-trial basis based on participants’ choices, to rapidly hone in on indifference points. This approach has been repeatedly validated to provide reliable estimates of delay discounting (Estle *et al.* 2006). As a summary measure, we use an area-under-the-curve discounting measure (AUC) that provides a valid and reliable index of how steeply an individual discounts delayed rewards (Myerson *et al.* 2001). See below for exact details on the parameters of the Delay Discounting Task.

Sustained Attention

We measure continuous sustained attention using the Short Penn Continuous Performance Test (CPT, Number/Letter Version) (Gur *et al.* 2001; Gur *et al.* 2001; Gur *et al.* 2010). Participants see vertical and horizontal red lines flash on the computer screen. In one block, they must press the spacebar when the lines form a number and in the other block they press the spacebar when the lines form a letter. The lines are displayed for 300 ms followed by a 700 ms ITI. Each block contains 90 stimuli and lasts for 1.5 minutes.

Verbal Episodic Memory

The NIH Toolbox contains a measure of non-verbal episodic memory. Thus, we are assessing verbal episodic memory using Form A of the Penn Word Memory Test (Gur *et al.* 2001a; Gur *et al.* 2010). Participants are shown 20 words and asked to remember them for a subsequent memory test. They are then shown 40 words (the 20 previously presented words and 20 new words matched on memory related characteristics). They decide whether they have seen the word previously by choosing among “definitely yes,” “probably yes,” “probably no,” and “definitely no.”

Spatial Orientation

The NIH Toolbox does not contain any measures of visual-spatial processing. Thus, we are measuring spatial orientation processing using the Variable Short Penn Line Orientation Test (Gur *et al.* 2001a; Gur *et al.* 2010). Participants are shown two lines with different orientations. They have to rotate one of the lines (a moveable blue one) so that is parallel to the other line (a fixed red line). The rotation of the blue line is accomplished by clicking buttons on the keyboard that rotate the lines either clockwise or counterclockwise. Across trials, the lines vary in their relative location on the screen, though the distance between the centers of the two lines is always the same. The length of the red line is always the same, but the length of the blue line can be either short or long. There are a total of 24 trials.

Emotion Processing

The NIH Toolbox contains only self-report measures of emotional function. Thus, in order to obtain a behavioral measure of emotion processing, we are using the Penn Emotion Recognition Test (Gur *et al.* 2001a; Gur *et al.* 2010). Participants are presented with 40 faces, one at a time. They are asked to choose what emotion the face is showing from five choices: Happy, Sad, Angry, Scared and No Feeling. Half of the faces are males and half are females. There are 8 faces each that have a happy, sad, angry, scared or no feeling expression.

Fluid Intelligence

Although the Toolbox contains measures of crystallized IQ (e.g., vocabulary acquisition), an aspect of IQ strongly influenced by educational opportunities, and measures of executive function (which are both theoretically and empirically related to fluid intelligence), it does not contain a specific measure of fluid intelligence. This construct is strongly linked to specific functional outcomes and to variations in neuronal structure and function in humans (Duncan *et al.* 2000; Duncan 2003; Duncan 2005). The most commonly used measure of fluid intelligence is Raven’s Progressive Matrices (Prabhakaran *et al.* 1997; Christoff *et al.* 2001; Gray *et al.* 2003; Conway *et al.* 2005; Gray *et al.* 2005; Wendelken *et al.* 2008). We use Form A of an abbreviated version of the Raven's developed by Gur and colleagues (Bilker *et al.* 2012). Participants are presented with patterns made up of 2x2, 3x3 or 1x5 arrangements of squares, with one of the squares missing. The participant must pick one of five response choices that best fits the missing square on the pattern. The task has 24 items and 3 bonus items, arranged in order of increasing difficulty. However, the task discontinues if the participant makes 5 incorrect responses in a row.

Self-Reported Function

The NIH toolbox contains self-report measures of a number of important domains of experience, including positive and negative affect, stress, anxiety, depression and social support. To obtain additional self-report information on an even broader variety of domains, we also administer the Achenbach Adult Self-Report (ASR) for Ages 18-59 (Achenbach 2009). Specifically, we administer the 123 items from Section VIII. These can be used to generate the ASR Syndrome Scales and the ASR DSM-Oriented Scales.

Detailed Description of Delay Discounting Task

In this task, participants are presented with two choices on each trial – a smaller amount "today" or a larger amount at a later point in time. Participants make choices at each of 6 delays (1 month, 6 months, 1 year, 3 years, 5 years and 10 years) and for two delayed amounts ($200 and $40,000). For each combination of delay and amount of delayed reward (e.g., $200 in 1 month or $40,000 in 6 months), participants make 5 choices, and the value that would have been used for the immediate amount in a 6th choice is taken as the indifference point for that condition. The participants make all five choices for a particular combination of delay and amount before moving on to the next combination of delay and amount. The order is as follows:

Delayed amount of $200 dollars

- Today versus 6 months
- Today versus 3 years
- Today versus 1 month
- Today versus 5 years
- Today versus 10 years
- Today versus 1 year

Delayed amount of $40,000 dollars

- Today versus 6 months
- Today versus 3 years
- Today versus 1 month
- Today versus 5 years

- Today versus 10 years
- Today versus 1 year

The first choice at each delay is between the delayed amount ($200 or $40,000) and an immediate amount equal to ½ the delayed amount (e.g., $100 today or $200 in 1 month, $20,000 today or $40,000 in one month). The size of the adjustment after the first choice is always ½ the amount of the immediate value on the first choice (e.g., a change of $50 if the first immediate amount is $100). If the subject chooses the immediate amount, then the immediate amount is reduced on the next choice (e.g., $50 today versus $200 in 1 month). If the subject chooses the delayed amount, then the immediate amount is increased (e.g., $150 today versus $200 in 1 month). The amount of change on each subsequent choice is ½ the amount of the prior change (e.g., $25 on the 3rd trial), regardless of whether the subject chooses the immediate or the delayed amount. This procedure rapidly hones in on the amount of immediate gain that is close to the subjective value of the delayed gain.

This design means that for all the choices with $200 dollars as the delayed amount, the first choice will always be between $100 today, and $200 in the specified time period. The second choice will always increment or decrement the immediate value by $50. The third choice will always increment or decrement the immediate value by $25. The fourth choice will always increment or decrement the immediate value by $12.50. The fifth choice will always increment or decrement the immediate value by $6.25. The “sixth” choice value, which is never presented to the subject, but is entered in the database, is always an increment or decrement of $3.125 from the immediate value on the 5th choice. Similarly, for all the choices with $40,000 dollars as the delayed amount, the first choice will always be between $20,000 today, and $40,000 in XX time period. The second choice will always increment or decrement the immediate value by $10,000. The third choice will always increment or decrement the immediate value by $5,000. The fourth choice will always increment or decrement the immediate value by $2,500. The fifth choice will always increment or decrement the immediate value by $1,250. The “sixth” choice value, which is never presented to the subject, but is entered in the database, will always be an increment or decrement of $625 from the immediate value on the 5th choice.

Thus, for the $200 amount, we will have 6 Subject Values:

- $SV_{1mo.2}$
- $SV_{6mo.2}$
- $SV_{1yr.2}$
- $SV_{3yr.2}$
- $SV_{5yr.2}$
- $SV_{10yr.2}$

Thus, for the $40,000 amount, we will have 6 Subject Values:

- $SV_{1mo.40}$
- $SV_{6mo.40}$
- $SV_{1yr.40}$
- $SV_{3yr.40}$
- $SV_{5yr.40}$
- $SV_{10yr.40}$

We compute an Area under the curve measure for each of the two amounts as described below.

Area under the Curve for $200 = $((1+SV_{1mo.2})/(120*200)) + ((SV_{1mo.2}+SV_{6mo.2})/(48*200)) + ((SV_{6mo.2}+SV_{1yr.2})/(40*200)) + ((SV_{1yr.2}+SV_{3yr.2})/(10*200)) + ((SV_{3yr.2}+SV_{5yr.2})/(10*200)) + ((SV_{5yr.2}+SV_{10yr.2})/(4*200))$

Area under the Curve for $40,000 = $((1+SV_{1mo.4})/(120*40,000)) + ((SV_{1mo.4}+SV_{6mo.4})/(48*40,000)) + ((SV_{6mo.4}+SV_{1yr.4})/(40*40,000)) + ((SV_{1yr.4}+SV_{3yr.4})/(10*40,000)) + ((SV_{3yr.4}+SV_{5yr.4})/(10*40,000)) + ((SV_{5yr.4}+SV_{10yr.4})/(4*40,000))$

References

Achenbach TM (2009). The Achenbach System of Empirically Based Assessement (ASEBA): Development, Findings, Theory, and Applications. 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.

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.

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

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: a journal of neurology* 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.

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

Delgado MR, Nystrom LE, Fissell C, Noll DC, Fiez JA (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of Neurophysiology* 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. *Am J Ophthalmol* 130(5): 687

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.

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. *The American journal of 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. *The Journal of Neuroscience* 31:11597-11616.

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

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.

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

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.

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.

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.

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

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.

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.

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

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. *Biological psychiatry* 55(4): 359-366.

McCrae RR, Costa PT (2004). A contemplated revision of the NEO Five Factor Inventory. *Personality and Individual Differences* 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.

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. *Neurological research* 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.

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

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.

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.

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.

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.

Wheatley T, Milleville SC, Martin A (2007). Understanding animate agents: distinct roles for the social network and mirror system. *Psychological science* 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.

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:S173-86.

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.

** Further details on the data acquisition and analysis methods are available in five manuscripts submitted in February 2013, to a special issue of *NeuroImage* (Barch *et al.* 2013; Glasser *et al.* 2013, Smith *et al.* 2013, Ugurbil *et al.* 2013, Marcus *et al.* 2013).

Appendices

Appendix 1. HCP scan protocols

Download Appendix 1 here: http://humanconnectome.org/documentation/data-release/Q1_Release_Appendix_I.pdf

Appendix 2. Matlab code for voxel-wise correction of dMRI gradients

Download Appendix 2 here: http://humanconnectome.org/documentation/data-release/Q1_Release_Appendix_II.pdf

Appendix 3. File Names and Directory Structure for Unprocessed and Preprocessed HCP Q1 Data

Download Appendix 3 here: http://humanconnectome.org/documentation/data-release/Q1_Release_Appendix_III.pdf

Appendix 4. HCP Protocol Standard Operating Procedures (SOPs)

Download Appendix 4 here: http://humanconnectome.org/documentation/data-release/Q1_Release_Appendix_IV.pdf

Appendix 5: Data Acquisition Information for an Exemplar Subject

Download Appendix 5 here: http://humanconnectome.org/documentation/data-release/Q1_Release_Appendix_V.pdf

Appendix 6. Task fMRI E-Prime Key Variables

Download Appendix 6 here: http://humanconnectome.org/documentation/data-release/Q1_Release_Appendix_VI.pdf