Courtois NeuroMod

Release 0.0

Mar 19, 2020

Contents:

> **Warning:** This documentation corresponds to the cneuromod-2020-alpha, unstable release, with access limited to the NeuroMod team. The first public and stable data release is scheduled for June 2020. Thank you for your patience!

The Courtois project on Neural Modelling (Courtois NeuroMod) aims at training artificial neural networks using extensive experimental data on individual human brain activity and behaviour. Six subjects (three women, three men) are getting scanned weekly for five years, for a total of 500 hours of functional data per subject, including functional localizers (vision, language, memory, emotion), movies and video game play. Functional neuroimaging data are collected with both functional magnetic resonance imaging, magnetoencephalography and a variety of sensors (including electrodermal activity and occulometry).

Courtois NeuroMod data are freely shared with the scientific community to advance research at the interface of neuroscience and artificial intelligence.

Datasets

1.1 hcptrt

This `cneuromod` dataset is called HCP test-retest (`hcptrt`), because participants repeated 15 times the functional localizers developed by the Human Connectome Project, for a total of approximately 10 hours of functional data per subject. The protocol consisted of seven tasks, described below (text adapted from the HCP protocol). Before each task, participants were given detailed instructions and examples, as well as a practice run. A session was typically composed either of two repetitions of the HCP localizers, or one resting-state run and one HCP localizer. The eprime scripts for preparation and presentation of the stimuli can be found in the HCP database. Stimuli and e-prime scripst were provided by the Human Connectome Project, U-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research, and by the McDonnell Center for Systems Neuroscience at Washington University. Note that in the `cneuromod` DataLad, functional runs are named `func_sub-<participant>_ses-<sess>_task-<task>_run-<run>`, where the `<participant>` tag includes `sub-01`, `sub-02`, `sub-03` and `sub-05`. For each functional run, a companion file `_events.tsv` contains the timing and type of events presented to the subject. Session tags `<sess>` are `001`, `002` etc, and the number and composition of sessions vary from subject to subject. The `<task>` tags are `restingstate`, `gambling`, `motor`, `social`, `wm`, `emotion`, `language` and `relational`. Tasks that were repeated twice have separate `<run>` tags (`01`, `02`).

1.1.1 Gambling

`gambling` duration: approximately 4 minutes. Participants were asked to guess whether a hidden number (represented by a "?" during 1500ms) was above or below 5 (Delgado et al. 2000). They indicated their choice using a button press, and were then shown the actual number. If they guessed correctly they were told they won money (+$1.00, `win` trial), if they guessed incorrectly they were told they lost money (-$0.50, `loss` trial), and if the number was exactly 5 they were told that they neither won or lost money ($0, `neutral` trial). Note that no money was actually given to the participants and, as such, this task may not be an accurate reproduction of the HCP protocol. The conditions were presented in blocks of 8 trials of type `reward` (6 `win` trials pseudo randomly interleaved with either 1 `neutral` and 1 `loss` trial, 2 `neutral` trials, or 2 `loss` trials) or of type `punishment` (6 `loss` trials pseudo-randomly

interleaved with either 1 `neutral` and 1 `win` trial, 2 `neutral` trials, or 2 `win` trials). There were four blocks per run (2 `reward` and 2 `punishment`), and two runs in total.

1.1.2 Motor

`motor` duration: approximately 5 minutes. This task was adapted from (Buckner et al. 2011; Yeo et al. 2011). Participants were presented a visual cue, and were asked to either tap their left or right fingers (event types `left_hand` and `right_hand`, resp.), squeeze their left or right toes (event types `left_foot` and `right_foot`, resp.), or move their tongue to map motor area (event type `tongue`). Each movement lasted 12 seconds, and in total there were 13 blocks, with 2 of `tongue` movements, 4 of hand movements (2 `right_hand` and 2 `left_hand`), and 4 of foot movements (2 `right_foot` and 2 `left_foot`), and three 15 second fixation blocks where participants were instructed not to move anything. There were two runs in total, and 13 blocks per run.

1.1.3 Language processing

`language` duration: approximately 5 minutes. Participants were presented with two types of events. During `story` events, participants listened to an auditory story (5-9 sentences, about 20 seconds), followed by a two-alternative forced-choice question. During `math` events, they listened to a math problem (addition and subtraction only, varies in length), and were instructed to push a button to select the first or the second answer as being correct. The task was adaptive so that for every correct answer the level of difficulty increased. The math task was designed this way to maintain the same level of difficulty between participants. There were 2 runs, each with 4 `story` and 4 `math` blocks, interleaved.

1.1.4 Social cognition

`social` duration: approximately 5 minutes. Participants were presented with short video clips (20 seconds) of objects (squares, circles, triangles) that either interacted in some way (event type `mental`), or moved randomly on the screen (event type `random`) (Castelli et al. 2000; Wheatley et al. 2007). Following each clip, participants were asked to judge whether the objects had a "Mental interaction" (an interaction that appeared as if the shapes were taking into account each other's feelings and thoughts), whether the were "Not Sure", or if there was "No interaction". Button presses were used to record their responses. In each of the two runs, participants viewed 5 `mental` videos and 5 `random` videos, and had 5 fixation blocks of 15 seconds each.

1.1.5 Relational processing

`relational` duration: approximately 5 minutes. Participants were shown 6 different shapes filled with 1 of 6 different textures (Smith et al. 2007). There were two conditions: relations processing (event type `relational`), and control matching condition (event type `control`). In the `relational` events, 2 pairs of objects were presented on the screen, with one pair at the top of the screen, and the other pair at the bottom. Participants were instructed to decide what dimension differed in the top pair (shape or texture), and then decide if the bottom pair differed, or not, on the same dimension (i.e. if the top pair differed in shape, did the bottom pair also differ in shape). Their answers were recorded by one of two button presses: "a" differ on same dimension; "b" don't differ on same dimension. In the `control` events, participants were shown two objects at the top of the screen, and one object at the bottom of the screen, with a word in the middle of the screen (either "shape" or "texture").They were told to decide whether the bottom object matched either of the top two objects on that dimension (i.e., if the word is "shape", did the bottom object have the same shape as either of the top two objects). Participants responded "yes" or "no" using the button box. For the `relational` condition, the stimuli were presented for 3500 ms, with a 500 ms ITI, and there were four trials per block. In the `control`condition, stimuli were presented for 2800 ms, with a 400 ms ITI, and there were 5 trials per block. In total there were two runs, each with three `relational` blocks, three `control` blocks and three 16-second fixation blocks.

1.1.6 Emotion processing

`emotion` duration: approximately 4 minutes. Participants were shown triads of faces (event type `face`) or shapes (event type `shape`), and were asked to decide which of the shapes at the bottom of the screen matches the target face/ shape at the top of the screen (adapted from Smith et al. 2007). Faces had either an angry or fearful expression. Faces, and shapes were presented in three blocks of 6 trials (3 `face` and 3 `shape`), with each trial lasting 2 seconds, followed by a 1 second inter-stimulus interval. Each block was preceded by a 3000 ms task cue ("shape" or "face"), so that each block was 21 seconds long, including the cue. In total there were two runs, three `face` blocks and three `shape` blocks, with 8 seconds of fixation at the end of each run.

1.1.7 Working memory

`wm` duration: approximately 5 minutes. There were two subtasks: a category specific representation, and a working memory task. Participants were presented with blocks of either places, tools, faces, and body parts. Within each run, all 4 types of stimuli were presented in block, with each block being labelled as a 2-back task (participants needed to indicate if they saw the same image two images back), or a version of a 0-back task (participants were shown a target at the start of the trial and they needed to indicate if the image that they were seeing matched the target). There were thus 8 different event types `<stim>_<back>`, where `<stim>` was one of `place`, `tools`, `face` or `body`, and `<back>` was one of `0back` or `2back`. Each image was presented for 2 seconds, followed by a 500 ms ITI. Stimuli were presented for 2 seconds, followed by a 500 ms inter-task interval. Each of the 2 runs included 8 event types with 10 trials per type, as well as 4 fixations blocks (15 secs).

1.1.8 Resting state

`restingstate` duration: 15 minutes. In every other session, one resting-state fMRI run was acquired, giving 5 runs per participant. Participants were asked to have their eye open, be looking at fixation cross in the middle of the screen and be instructed to not fall asleep. A total of five resting-state fMRI runs were acquired per subject.

1.2 movie10

This dataset includes about 10 hours of functional data for all 6 participants. The python & psychopy scripts for preparation and presentation of the clips can be found in `src/tasks/video.py` of the following github repository. Session tags `<sess>` were `vid001`, `vid002` etc, and the number and composition of sessions varied from subject to subject. The `<task>` tags used in DataLad corresponded to each movie (`bournesupremacy`, `wolfofwallstreet`, `life`, `hiddenfigures`). Each movie was cut into roughly ten minutes segments (tags `<seg>` in `01`, `02`, etc) presented in a separate run. Exact cutting points were manually selected to not interrupt the narrative flow. Fade out to a black screen was added at the end of each clip, and with a few seconds overlap between the end of a clip and the beginning of the next clip. The movie segments can be found under `movie10/stimuli/<movie>/<movie>_seg<seg>.mkv`, and the functional runs are named `func_sub-<participant>_ses-<sess>_task-<movie>_run-<seg>`, where the `<participant>` tag ranges from `sub-01` to `sub-06`. A companion file `_events.tsv` contains the timing and type of conditions presented to the subject.

The participants watched the following movies (cogatlas):

- `<task>` name `bournesupremacy`: The Bourne supremacy. Duration ~100 minutes.

- `<task>` name `wolfofwallstreet`: The wolf of wall street. Duration ~170 minutes.

- `<task>` name `hiddenfigures`: Hidden figures. Duration ~120 minutes. This movie was presented twice, for a total duration of ~240 minutes.

- `<task>` name `life`: Life Disc one of four: "Challenges of life, reptiles and amphibian mammals". DVD set was narrated by David Attenborough Duration, and lasted ~50 minutes. This movie was presented twice, for a total duration of ~100 minutes.

Access

2.1 Ethics

The Courtois NeuroMod project has been approved by the institutional research ethics board of the CIUSSS du Centre-Sud-de-l'île-de-Montréal. The CIUSSS is a large governmental health organization, and the core NeuroMod team is based at the Research Centre of the Montreal Geriatric Institute (CRIUGM), which is a part of the CIUSSS, and affiliated with the University of Montreal. The ethics documentation may be useful for research teams to receive ethics approval for secondary analysis by their local institutions, if required.

- approval_irb_ciusss_french.pdf: letter of approval from the institutional review board for the Courtois Neuro-Mod project (in French).

- courtois_neuromod_project_description.pdf: scientific overview of the Courtois NeuroMod project (in English).

- consent_form_english.pdf: the informed consent form signed by participants (English version).

- consent_form_french.pdf: the informed consent form signed by participants (French version).

2.2 Downloading the dataset

All data are made available as a DataLad collection (login required) in the UNF git. DataLad is a tool for versioning a large data structure in a git repository. The dataset can be explored without downloading the data, and it is easy to only download the subset of the data you need for your project. See the DataLad handbook for further information.

We recommend creating an SSH key (if not already present) on the machine on which the dataset will be installed using `ssh-keygen` and following the instructions. Then you can go to the SSH key settings of your account an add the SSH key, by pasting you public keys contained in the file `~/.ssh/id_rsa.pub`.

To obtain the data, you need to install a recent version of the DataLad software, available for Linux, OSX and Windows. Note that you need to have valid login credentials to access the NeuroMod git as well as the NeuroMod Amazon S3 fileserver. Once you have obtained these credentials, you can proceed as follows in a terminal:

```
# Install recursively the dataset and subdataset of the current project.
# If using ssh git clone as follow, you can set your public SSH key in the present⌴
→git to ease future updates.
datalad install -r git@git.unf-montreal.ca:cneuromod/cneuromod.git
# If errors show up relative to .heudiconv subdataset/submodule, this is OK, they are⌴
→not published (will be cleaned up in the future).
cd cneuromod

# You will most likely want to checkout a stable release tag for your analysis.
# You can do this while creating a new branch with the name you like.s
# In this branch you can run you analysis and commit your work.
# For instance (but this tag doesn't exist yet):
git checkout -b <myanalysisbranch> cneuromod-2020

# We now set as environment variable the credentials to the file server.
# The s3 access_key and secret_key will be provided upon request by the data manager.
# This needs to be set in your `bash` everytime you want to download data.
export AWS_ACCESS_KEY_ID=<s3_access_key>  AWS_SECRET_ACCESS_KEY=<s3_secret_key>

# Now you can get data using:
datalad get -r <any/file/in/the/dataset.example>

# You can also run reproductible analysis using
datalad run -i <the files needed for my analysis> ../path/to/my/analysis.script
# which will download the required files and store the command launched and the⌴
→results in the datalad dataset.
```

2.3 Updates

The dataset will be updated with new releases so you might want to get these changes (unless you are running analyses, or trying to reproduce results). The master branch will evolve with the project, and can be unstable or messy. Thus, we recommend using specific release tags. There is one stable release per year, e.g. `cneuromod-2020`, which is preceded by alpha (e.g. `cneuromod-2020-alpha`), beta (e.g. `cneuromod-2020-beta`) and release candidate (e.g. `cneuromod-2020-rc`). To update your dataset to the latest version, use:

```
# update the dataset recursively
datalad update -r --merge
```

Once your local dataset clone is updated, you might need to pull new data, as some files could have been replace.

CHAPTER 3

MRI

3.1 Image acquisition

3.1.1 Scanner

Magnetic resonance imaging (MRI) for the Courtois neuromod project is being acquired at the functional neuroimaging unit (UNF), located at the "Centre de Recherche de l'Institut Universitaire de Gériatrie de Montréal" (CRIUGM) and affiliated with University of Montreal as well as the CIUSSS du Centre-Sud-de-l'île-de-Montréal. The scanner is a Siemens Prisma Fit, equipped with a 2-channel transmit body coil and a 64-channel receive head/neck coil. Most imaging in the Courtois Neuromod project are composed solely of functional MRI runs. Periodically, an entire session is dedicated to anatomical scans.

3.1.2 Personalized head cases

In order to minimize movement, each participant wears a custom-designed, personalized headcase during scanning, built by Caseforge. The headcases are milled based on a head scan of each participant generated using a handheld 3D scanner, and the shape of the MRI coil. Caseforge mills the personalized headcases in polystyrene foam blocks.

3.2 BIDS formating

All functional and anatomical data has been formated in BIDS, for more information visit the Brain Imaging Data Structure documentation site.

3.2.1 Functional sequences

The parameters of the functional MRI sequence relevant for data analysis can be found in the NeuroMod DataLad. The functional acquisition parameters are all identical to the one used in the `hcptrt` dataset. The Siemens exam card can be found here, and is briefly recapitulated below. Functional MRI data was acquired using an accelerated simultaneous multi-slice, gradient echo-planar imaging sequence (Xu et al., 2013) developed at the Center for Magnetic Resonance

Research (CMRR) University of Minnesota, as part of the Human Connectome Project (Glasser et al., 2016). The sequence is available on the Siemens PRISMA scanner at UNF through a concept to production (C2P) agreement, and was used with the following parameters: slice acceleration factor = 4, TR = 1.49 s, TE = 37 ms, flip angle = 52 degrees, voxel size = 2 mm x 2 mm x 2 mm, 60 slices, acquisition matrix 96x96. In each session, a short acquisition (3 volumes) with reversed phase encoding direction was run to allow retrospective correction of B0 field inhomogeneity-induced distortion.

3.2.2 Brain anatomical sequences

The parameters of the brain anatomical MRI sequences relevant for data analysis can be found in the NeuroMod DataLad. The acquisition parameters are identical for all anatomical sessions. The Siemens pdf exam card of the anatomical sessions can be found here, and is briefly recapitulated below. A standard (brain) anatomical session started with a 21 s localizer scan, and then included the following sequences:

- T1-weighted MPRAGE 3D sagittal sequence (duration 6:38 min, TR = 2.4 s, TE = 2.2 ms, voxel size = 0.8 mm isotropic, R=2 acceleration)

- T2-weighted FSE (SPACE) 3D sagittal sequence (duration 5:57 min, TR = 3.2 s, TE = 563 ms, voxel size = 0.8 mm isotropic, R=2 acceleration)

- Diffusion-weighted 2D axial sequence (duration 4:04 min, TR = 2.3 s, TE = 82 ms, 57 slices, voxel size = 2 mm isotropic, phase-encoding P-A, SMS=3 through-plane acceleration, b-max = 3000 s/mm2). The same sequence was run with phase-encoding A-P to correct for susceptibility distortions.

- gradient-echo magnetization-transfer 3D sequence (duration 3:34 min, 28 = ms, TE = 3.3 ms, flip angle = 6 deg, voxel size = 1.5 mm isotropic, R=2 in-plane GRAPPA, MT pulse Gaussian shape centered at 1.2 kHz offset).

- gradient-echo proton density 3D sequence (same parameters as above, without the MT pulse).

- gradient-echo T1-weighted 3D sequence (same parameters as above, except: TR = 18 ms, flip angle = 20 deg).

- B1+ field map (duration 0:21 min, voxel size = 6 mm isotropic)

- MP2RAGE 3D sequence (duration 7:26 min, TR = 4 s, TE = 1.51 ms, TI1 = 700 ms, TI2 = 1500 ms, voxel size = 1.2 mm isotropic, R=2 acceleration)

- Susceptibility-weighted 3D sequence (duration 4:54 min, TR = 27 ms, TE = 20 ms)

3.2.3 Spinal cord anatomical sequences

The parameters of the spinal cord anatomical MRI sequences relevant for data analysis can be found in the BIDS dataset, and included metadata. The acquisition parameters are identical for all anatomical sessions, and follow a community spinal cord standard imaging protocol. The Siemens pdf exam card of the anatomical sessions can be found here, and is briefly recapitulated below. A standard (spinal cord) anatomical session starts with a 21 s localizer scan, and then includes the following sequences:

- T1-weighted 3D sagittal sequence (duration 4:44 min, TR = 2 s, TE = 3.72 ms, FA = 9 deg, voxel size = 1.0 mm isotropic, R=2 acceleration)

- T2-weighted 3D sagittal sequence (duration 4:02 min, TR = 1.5 s, TE = 120 ms, FA = 120 deg, voxel size = 0.8 mm isotropic, R=3 acceleration)

- Diffusion-weighted 2D axial sequence (cardiac-gated with pulseOx, approximate duration 3 min, TR = 620 ms, TE = 60 ms, voxel size = 0.9 x 0.9 x 0.5 mm, phase-encoding A-P, b-max = 800 s/mm2)

- Gradient-echo magnetization-transfer 3D axial sequence (duration 2:12 min, TR = 35 ms, TE = 3.13 ms, FA = 9 deg, voxel size = 0.9 x 0.9 x 0.5 mm, R=2 acceleration, with MT Gaussian pulse)

- Gradient-echo proton-density weighted 3D axial sequence (same parameters as above, without the MT pulse).

- Gradient-echo T1-weighted 3D axial sequence (same parameters as above, except: TR = 15 ms, flip angle = 15 deg).

- gradient-echo ME (duration 4:45 min, TR = 600 ms, effective TE = 14 ms (this is a multi-echo sequence), FA = 30 deg, voxel size = 0.9 x 0.9 x 0.5 mm, R=2 acceleration)

3.3 Stimuli

3.3.1 Visual presentation

All visual stimuli were projected onto a blank screen located in the MRI room, through a waveguide.

3.3.2 Auditory system

For functional sessions, participant wore MRI compatible S15 Sensimetric headphone inserts, proving high-quality acoustic stimulation and substantial attenuation of background noise. On the computer used for stimuli presentation, a custom impulse response of the headphones is applied with an online finite impulse response filter using the LADSPA DSP to all the presented stimuli.This impulse response was provided by the manufacturer. Sounds was amplified using an AudioSource AMP100V amplifier, situated in the control room. Participants also wear earmuffs, adapted from commercially available model. The commercial earmuffs have been cut and glued back together in order to be thin enough to be used with the 64-channels MRI coil.

3.3.3 Stimuli presentation

For the HCP-trt dataset, Eprime scripts provided by the Human Connectome project were adapted for our presentation system, and run using Eprime 2.0. For all other tasks, a custom overlay on top of the psychopy library was used to present the different tasks and synchronize task with the scanner trigger pulses. This software also allowed to trigger the start of the eyetracking system, and onset the stimuli presentation. Trigger pulses were also recorded in the AcqKnowledge software. All task stimuli scripts are available through github.

3.4 Physiological measures

3.4.1 Biopac

During all sequences, electrophysiological signals were recorded using a Biopac M160 MRI compatible systems and amplifiers. Measurements were acquired at 1000 Hz. Recodings were synchronized to the MRI scans via trigger pulses. All measurements were recorded and monitored using Biopac's AcqKnowledge sofware.

3.4.2 Plethysmograph

Participant cardiac pulse was measured using an MRI compatible plethysmograph. A Biopac TSD200-MRI photo-plethysmogram transducer was placed on the foot or toe of the participants to obtain beat-by-beat estimates of heart rate.

3.4.3 Skin conductance

Skin conductance, was measured using two electrodes, one applied to the sole of the foot and the other to the ankle, to record the participant electrodermal response.

3.4.4 Electrocardiogram

An electrocardiogram (ECG) was used to measure the electrical activity generated by the heart. The ECG was recorded using three MRI compatible electrodes that were placed adjacent to one aother, on the lower left rib cage, just under the heart.

3.4.5 Respiration

Participant's respiration was measured using a custom MRI compatible respiration belt. The respiration system consisted of: a pressure cuff taken from a blood pressure monitor (PhysioLogic blood), a pressure sensor (MPXV5004GC7U, NXP USA Inc), and flexible tubing. The cuff was attached to the participant upper abdomen using Velcro strap, and then connected to the pressure sensor, located outside the scanner room, using tubing passed through a waveguide. The pressure signal was recorded using an analog input on the Biopac system, and monitored using AcqKnowledge software.

Derivatives

4.1 fMRIPrep

4.1.1 Overview

The functional data was preprocessed using the fMRIprep pipeline version: 1.5.0. FmriPrep is an fMRI data preprocessing pipeline that requires minimal user input, while providing error and output reporting. It performs basic processing steps (coregistration, normalization, unwarping, noise component extraction, segmentation, skullstripping etc.) and provides outputs that can be easily submitted to a variety of group level analyses, including task-based or resting-state fMRI, graph theory measures, surface or volume-based statistics, etc. The fmriprep pipeline uses a combination of tools from well-known software packages, including FSL, ANTs, FreeSurfer and AFNI. For additional information regarding fMRIPrep installion, worflow and outputs, please visit the documentation page. Note that the `slicetiming` and `recon-all` options were disabled (i.e. fMRIprep was invoked with the flags `--fs-no-reconall --ignore slicetiming`).

4.1.2 Outputs

The outputs of fMRIprep can be found under the folder of each dataset (e.g. `movie10`) `derivatives/fmriprep` in the Courtois NeuroMod datalad. The description of participant, session, task and event tags can be found in the Datasets section. Each participant folder (`sub-*`) contains:

- `anat` folder with T1 preprocessed and segmented in native and MNI space, registration parameters

- `ses-*/func` containing for each fMRI run of that session file prefixed with:

    - `_boldref.nii.gz` : a BOLD single volume reference.

    - `_*-brain_mask.nii.gz` : the brain mask in fMRI space.

    - `_*-preproc_bold.nii.gz` : the preprocessed BOLD timeseries.

    - `_*-confounds_regressors.tsv` : a tabular tsv file, containing a large set of confounds to use in analysis steps (eg. GLM). Note that regressors are likely correlated, thus it is recommended to use a subset of these regressors. Also note that preprocessed time series have not been corrected for any confounds,

but simply realigned in space, and it is therefore critical to regress some of the available confounds prior to analysis. For python users, we recommend using this tool fmriprep_confound_loader to load confounds from the fMRIprep outputs, using with the `minimal` strategy. In particular, as the NeuroMod data consistently exhibits low levels of motion, we recommend against removing time points with excessive motion (aka scrubbing).

4.1.3 Pipeline description

The following boilerplate text was automatically generated by fMRIPrep with the express intention that users should copy and paste this text into their manuscripts *unchanged*. It is released under the CC0 license. All references in the text link to a `.bib` file with detailed reference list, ready to be incorporated in a `LaTeX` document.

Results included in this manuscript come from preprocessing performed using *fMRIPrep* 1.5.0 (fmriprep1; fmriprep2; RRID:SCR_016216), which is based on *Nipype* 1.2.2 (nipype1; nipype2; RRID:SCR_002502).

Anatomical data preprocessing

The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with `N4BiasFieldCorrection` [n4], distributed with ANTs 2.2.0 [ants, RRID:SCR_004757], and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a *Nipype* implementation of the `antsBrainExtraction.sh` workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using `fast` [FSL 5.0.9, RRID:SCR_002823, fsl_fast]. Volume-based spatial normalization to one standard space (MNI152NLin2009cAsym) was performed through nonlinear registration with `antsRegistration` (ANTs 2.2.0), using brain-extracted versions of both T1w reference and the T1w template. The following template was selected for spatial normalization: *ICBM 152 Nonlinear Asymmetrical template version 2009c* [mni152nlin2009casym, RRID:SCR_008796; Template-Flow ID: MNI152NLin2009cAsym].

Functional data preprocessing

For each of the BOLD runs found per subject (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. A deformation field to correct for susceptibility distortions was estimated based on two echo-planar imaging (EPI) references with opposing phase-encoding directions, using `3dQwarp` afni (AFNI 20160207). Based on the estimated susceptibility distortion, an unwarped BOLD reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using `flirt` [FSL 5.0.9, flirt] with the boundary-based registration [bbr] cost-function. Co-registration was configured with nine degrees of freedom to account for distortions remaining in the BOLD reference. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using `mcflirt` [FSL 5.0.9, mcflirt]. The BOLD time-series (including slice-timing correction when applied) were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. These resampled BOLD time-series will be referred to as *preprocessed BOLD in original space*, or just *preprocessed BOLD*. The BOLD time-series were resampled into standard space, generating a *preprocessed BOLD run in ['MNI152NLin2009cAsym'] space*. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Several confounding time-series were calculated based on the *preprocessed BOLD*: framewise displacement (FD), DVARS and three region-wise global signals. FD and DVARS are calculated for each functional run, both using their implementations in *Nipype* [following the definitions by power_fd_dvars]. The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction [*CompCor*, compcor]. Principal components are estimated after high-pass filtering the *preprocessed BOLD* time-series (using a discrete cosine filter with 128s cut-off) for the two *CompCor* variants:

temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 5% variable voxels within a mask covering the subcortical regions. This subcortical mask is obtained by heavily eroding the brain mask, which ensures it does not include cortical GM regions. For aCompCor, components are calculated within the intersection of the aforementioned mask and the union of CSF and WM masks calculated in T1w space, after their projection to the native space of each functional run (using the inverse BOLD-to-T1w transformation). Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the *k* components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each [confounds_satterthwaite_2013]. Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardised DVARS were annotated as motion outliers. All resamplings can be performed with *a single interpolation step* by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using `antsApplyTransforms` (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels [lanczos]. Non-gridded (surface) resamplings were performed using `mri_vol2surf` (FreeSurfer).

Many internal operations of *fMRIPrep* use *Nilearn* 0.5.2 [nilearn, RRID:SCR_001362], mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in *fMRIPrep*'s documentation.

4.1.4 Missing outputs

The following outputs of the fMRIprep pipeline are missing because of issues with the computing server at the time of the `cneuromod-2020-alpha` release. This will be fixed for the `cneuromod-2020-beta` release, scheduled in April 2020.

Movie-10

```
sub-05/ses-vid010/func/sub-05_ses-vid010_task-hiddenfigures_run-11_bold.nii.gz
sub-06/ses-vid005/func/sub-06_ses-vid005_task-bournesupremacy_run-02_bold.nii.gz
sub-06/ses-vid005/func/sub-06_ses-vid005_task-life_run-04_bold.nii.gz
sub-06/ses-vid006/func/sub-06_ses-vid006_task-bournesupremacy_run-06_bold.nii.gz
sub-06/ses-vid009/func/sub-06_ses-vid009_task-wolfofwallstreet_run-02_bold.nii.gz
```

HCP-trt

```
sub-01/ses-002/func/sub-01_ses-002_task-emotion_run-02_bold.nii.gz
sub-01/ses-003/func/sub-01_ses-003_task-emotion_run-01_bold.nii.gz
sub-01/ses-003/func/sub-01_ses-003_task-language_run-01_bold.nii.gz
sub-01/ses-003/func/sub-01_ses-003_task-social_run-01_bold.nii.gz
sub-01/ses-004/func/sub-01_ses-004_task-gambling_run-01_bold.nii.gz
sub-01/ses-004/func/sub-01_ses-004_task-motor_run-01_bold.nii.gz
sub-01/ses-004/func/sub-01_ses-004_task-restingstate_run-01_bold.nii.gz
sub-01/ses-007/func/sub-01_ses-007_task-emotion_run-01_bold.nii.gz
sub-01/ses-007/func/sub-01_ses-007_task-relational_run-01_bold.nii.gz
sub-01/ses-007/func/sub-01_ses-007_task-social_run-01_bold.nii.gz
sub-01/ses-009/func/sub-01_ses-009_task-language_run-01_bold.nii.gz
sub-01/ses-009/func/sub-01_ses-009_task-relational_run-01_bold.nii.gz
sub-01/ses-010/func/sub-01_ses-010_task-gambling_run-02_bold.nii.gz
sub-01/ses-010/func/sub-01_ses-010_task-motor_run-01_bold.nii.gz
```

```
sub-01/ses-010/func/sub-01_ses-010_task-motor_run-02_bold.nii.gz
sub-02/ses-001/func/sub-02_ses-001_task-motor_run-01_bold.nii.gz
sub-02/ses-001/func/sub-02_ses-001_task-relational_run-01_bold.nii.gz
sub-02/ses-003/func/sub-02_ses-003_task-motor_run-01_bold.nii.gz
sub-02/ses-003/func/sub-02_ses-003_task-social_run-01_bold.nii.gz
sub-02/ses-004/func/sub-02_ses-004_task-gambling_run-01_bold.nii.gz
sub-02/ses-004/func/sub-02_ses-004_task-gambling_run-02_bold.nii.gz
sub-02/ses-004/func/sub-02_ses-004_task-language_run-01_bold.nii.gz
sub-02/ses-004/func/sub-02_ses-004_task-language_run-02_bold.nii.gz
sub-02/ses-004/func/sub-02_ses-004_task-motor_run-01_bold.nii.gz
sub-02/ses-004/func/sub-02_ses-004_task-relational_run-02_bold.nii.gz
sub-02/ses-004/func/sub-02_ses-004_task-social_run-01_bold.nii.gz
sub-02/ses-004/func/sub-02_ses-004_task-wm_run-02_bold.nii.gz
sub-02/ses-005/func/sub-02_ses-005_task-emotion_run-01_bold.nii.gz
sub-02/ses-005/func/sub-02_ses-005_task-restingstate_run-01_bold.nii.gz
sub-02/ses-006/func/sub-02_ses-006_task-gambling_run-01_bold.nii.gz
sub-02/ses-006/func/sub-02_ses-006_task-language_run-01_bold.nii.gz
sub-02/ses-006/func/sub-02_ses-006_task-relational_run-01_bold.nii.gz
sub-02/ses-006/func/sub-02_ses-006_task-relational_run-02_bold.nii.gz
sub-02/ses-006/func/sub-02_ses-006_task-social_run-01_bold.nii.gz
sub-02/ses-006/func/sub-02_ses-006_task-wm_run-01_bold.nii.gz
sub-02/ses-006/func/sub-02_ses-006_task-wm_run-02_bold.nii.gz
sub-02/ses-007/func/sub-02_ses-007_task-emotion_run-01_bold.nii.gz
sub-02/ses-007/func/sub-02_ses-007_task-language_run-01_bold.nii.gz
sub-02/ses-008/func/sub-02_ses-008_task-gambling_run-02_bold.nii.gz
sub-02/ses-008/func/sub-02_ses-008_task-language_run-02_bold.nii.gz
sub-02/ses-008/func/sub-02_ses-008_task-motor_run-01_bold.nii.gz
sub-02/ses-008/func/sub-02_ses-008_task-relational_run-01_bold.nii.gz
sub-02/ses-008/func/sub-02_ses-008_task-wm_run-02_bold.nii.gz
sub-02/ses-009/func/sub-02_ses-009_task-emotion_run-01_bold.nii.gz
sub-02/ses-009/func/sub-02_ses-009_task-motor_run-01_bold.nii.gz
sub-02/ses-009/func/sub-02_ses-009_task-relational_run-01_bold.nii.gz
sub-02/ses-010/func/sub-02_ses-010_task-gambling_run-01_bold.nii.gz
sub-02/ses-010/func/sub-02_ses-010_task-gambling_run-02_bold.nii.gz
sub-02/ses-010/func/sub-02_ses-010_task-motor_run-01_bold.nii.gz
sub-02/ses-010/func/sub-02_ses-010_task-social_run-01_bold.nii.gz
sub-02/ses-010/func/sub-02_ses-010_task-social_run-02_bold.nii.gz
sub-02/ses-010/func/sub-02_ses-010_task-wm_run-02_bold.nii.gz
sub-03/ses-002/func/sub-03_ses-002_task-emotion_run-01_bold.nii.gz
sub-03/ses-002/func/sub-03_ses-002_task-motor_run-01_bold.nii.gz
sub-03/ses-002/func/sub-03_ses-002_task-relational_run-01_bold.nii.gz
sub-03/ses-003/func/sub-03_ses-003_task-emotion_run-02_bold.nii.gz
sub-03/ses-003/func/sub-03_ses-003_task-gambling_run-01_bold.nii.gz
sub-03/ses-003/func/sub-03_ses-003_task-gambling_run-02_bold.nii.gz
sub-03/ses-003/func/sub-03_ses-003_task-language_run-02_bold.nii.gz
sub-03/ses-003/func/sub-03_ses-003_task-motor_run-01_bold.nii.gz
sub-03/ses-003/func/sub-03_ses-003_task-motor_run-02_bold.nii.gz
sub-03/ses-003/func/sub-03_ses-003_task-relational_run-02_bold.nii.gz
sub-03/ses-004/func/sub-03_ses-004_task-emotion_run-01_bold.nii.gz
sub-03/ses-004/func/sub-03_ses-004_task-relational_run-01_bold.nii.gz
sub-03/ses-004/func/sub-03_ses-004_task-restingstate_run-01_bold.nii.gz
sub-03/ses-004/func/sub-03_ses-004_task-wm_run-01_bold.nii.gz
sub-03/ses-005/func/sub-03_ses-005_task-language_run-01_bold.nii.gz
sub-03/ses-005/func/sub-03_ses-005_task-relational_run-01_bold.nii.gz
sub-03/ses-005/func/sub-03_ses-005_task-restingstate_run-01_bold.nii.gz
sub-03/ses-006/func/sub-03_ses-006_task-gambling_run-02_bold.nii.gz
sub-03/ses-006/func/sub-03_ses-006_task-motor_run-01_bold.nii.gz
```

```
sub-03/ses-006/func/sub-03_ses-006_task-motor_run-02_bold.nii.gz
sub-03/ses-006/func/sub-03_ses-006_task-social_run-01_bold.nii.gz
sub-03/ses-007/func/sub-03_ses-007_task-emotion_run-01_bold.nii.gz
sub-03/ses-007/func/sub-03_ses-007_task-gambling_run-01_bold.nii.gz
sub-03/ses-007/func/sub-03_ses-007_task-language_run-01_bold.nii.gz
sub-03/ses-007/func/sub-03_ses-007_task-motor_run-01_bold.nii.gz
sub-03/ses-007/func/sub-03_ses-007_task-social_run-01_bold.nii.gz
sub-03/ses-008/func/sub-03_ses-008_task-emotion_run-01_bold.nii.gz
sub-03/ses-008/func/sub-03_ses-008_task-emotion_run-02_bold.nii.gz
sub-03/ses-008/func/sub-03_ses-008_task-gambling_run-02_bold.nii.gz
sub-03/ses-008/func/sub-03_ses-008_task-language_run-02_bold.nii.gz
sub-03/ses-008/func/sub-03_ses-008_task-motor_run-01_bold.nii.gz
sub-03/ses-008/func/sub-03_ses-008_task-motor_run-02_bold.nii.gz
sub-03/ses-008/func/sub-03_ses-008_task-relational_run-01_bold.nii.gz
sub-03/ses-009/func/sub-03_ses-009_task-emotion_run-02_bold.nii.gz
sub-03/ses-009/func/sub-03_ses-009_task-motor_run-01_bold.nii.gz
sub-03/ses-009/func/sub-03_ses-009_task-relational_run-01_bold.nii.gz
sub-03/ses-009/func/sub-03_ses-009_task-relational_run-02_bold.nii.gz
sub-03/ses-009/func/sub-03_ses-009_task-social_run-02_bold.nii.gz
sub-03/ses-009/func/sub-03_ses-009_task-wm_run-01_bold.nii.gz
sub-05/ses-001/func/sub-05_ses-001_task-gambling_run-01_bold.nii.gz
sub-05/ses-001/func/sub-05_ses-001_task-relational_run-02_bold.nii.gz
sub-05/ses-001/func/sub-05_ses-001_task-social_run-02_bold.nii.gz
sub-05/ses-001/func/sub-05_ses-001_task-wm_run-01_bold.nii.gz
sub-05/ses-001/func/sub-05_ses-001_task-wm_run-02_bold.nii.gz
sub-05/ses-003/func/sub-05_ses-003_task-language_run-01_bold.nii.gz
sub-05/ses-003/func/sub-05_ses-003_task-social_run-01_bold.nii.gz
sub-05/ses-003/func/sub-05_ses-003_task-wm_run-01_bold.nii.gz
sub-05/ses-004/func/sub-05_ses-004_task-gambling_run-01_bold.nii.gz
sub-05/ses-004/func/sub-05_ses-004_task-gambling_run-02_bold.nii.gz
sub-05/ses-004/func/sub-05_ses-004_task-language_run-01_bold.nii.gz
sub-05/ses-004/func/sub-05_ses-004_task-language_run-02_bold.nii.gz
sub-05/ses-004/func/sub-05_ses-004_task-motor_run-01_bold.nii.gz
sub-05/ses-004/func/sub-05_ses-004_task-social_run-01_bold.nii.gz
sub-05/ses-004/func/sub-05_ses-004_task-social_run-02_bold.nii.gz
sub-05/ses-005/func/sub-05_ses-005_task-gambling_run-01_bold.nii.gz
sub-05/ses-005/func/sub-05_ses-005_task-motor_run-01_bold.nii.gz
sub-05/ses-005/func/sub-05_ses-005_task-relational_run-01_bold.nii.gz
sub-05/ses-005/func/sub-05_ses-005_task-social_run-01_bold.nii.gz
sub-05/ses-006/func/sub-05_ses-006_task-gambling_run-01_bold.nii.gz
sub-05/ses-006/func/sub-05_ses-006_task-restingstate_run-01_bold.nii.gz
sub-05/ses-008/func/sub-05_ses-008_task-gambling_run-02_bold.nii.gz
sub-05/ses-008/func/sub-05_ses-008_task-language_run-02_bold.nii.gz
sub-05/ses-008/func/sub-05_ses-008_task-motor_run-02_bold.nii.gz
sub-05/ses-008/func/sub-05_ses-008_task-relational_run-01_bold.nii.gz
sub-05/ses-008/func/sub-05_ses-008_task-social_run-02_bold.nii.gz
sub-05/ses-008/func/sub-05_ses-008_task-wm_run-02_bold.nii.gz
sub-05/ses-009/func/sub-05_ses-009_task-motor_run-01_bold.nii.gz
sub-05/ses-009/func/sub-05_ses-009_task-social_run-01_bold.nii.gz
sub-05/ses-009/func/sub-05_ses-009_task-wm_run-01_bold.nii.gz
sub-05/ses-010/func/sub-05_ses-010_task-emotion_run-01_bold.nii.gz
sub-05/ses-010/func/sub-05_ses-010_task-gambling_run-01_bold.nii.gz
sub-05/ses-010/func/sub-05_ses-010_task-language_run-01_bold.nii.gz
sub-05/ses-010/func/sub-05_ses-010_task-relational_run-02_bold.nii.gz
```

Authors

5.1 Overview

The Courtois NeuroMod project originated from the laboratory for brain simulation and exploration (SIMEXP), and collaborators are located at the Centre de Recherche de l'Institut de Gériatrie de Montréal (CRIUGM), CIUSSS du Centre-Sud-de-l'île-de-Montréal, as well as the Psychology Departement of University of Montreal (UdeM). The team has grown to include individuals from various institutions, and in particular the Computer Science and Operational Research (DIRO) Department at UdeM and the MILA.

5.2 Funding

The Courtois NeuroMod project was made possible by a 6.3M CAD (2018-23, PI Bellec) donation from the Courtois foundation. These funds are administered by the Fondation Institut Gériatrie Montréal (FIGM), part of CIUSSS du Centre-Sud-de-l'île-de-Montréal, as well as University of Montreal. Courtois NeuroMod also includes support for two separate consortia, called CIMAQ (early identification of Alzheimer's disease), and PRISME (looking for brain correlates of the evolution of symptoms in individuals with psychosis) based at the Institut Universitaire en Santé Mentale de Montréal (IUSMM).

5.3 Team

5.3.1 Core

- Pierre Bellec, scientific director (CRIUGM, Psychology, UdeM).

- Julie A Boyle, project manager (CRIUGM).

- Basile Pinsard, data manager (CRIUGM).

5.3.2 Modelling

- Guillaume Lajoie, Principal Investigator (Mathematics, UdeM & MILA)
- François Paugam, PhD student (CRIUGM & DIRO, UdeM).
- Pravish Sainath, Master's student (DIRO, UdeM & MILA)
- Yu Zhang, Post-Doctoral Fellow (CRIUGM, IVADO & Psychology, UdeM)
- Amal Boukhdir, PhD student (CRIUGM & DIRO, UdeM)

5.3.3 Vision

- Sana Ahmadi, Student (CRIUGM & Concordia).
- Norman Kong, Bachelor Student (McGill).

5.3.4 Memory

- Sylvie Belleville, Principal Investigator (CRIUGM & Psychology, UdeM).
- Samie-Jade Allard, Bachelor Student (Psychology, UdeM).
- François Nadeau, Bachelor Student (Psychology, UdeM).

5.3.5 Emotions

- Pierre Rainville, Principal Investigator (CRIUGM & Stomatology, UdeM).
- François Lespinasse, Master's Student (Psychology, UdeM).

5.3.6 Language

- Simona Brambatti, Principal Investigator (CRIUGM & Psychology, UdeM).
- Jonathan Armoza, Research Associate (CRIUGM & NYU).
- James Martin Floreani, Summer Intern 2019 (École Polytechnique, France).

5.3.7 Audition

- Adrian Fuente, Collaborator (CRIUGM & Audiology, UdeM).
- Nicolas Farrugia, Collaborator (IMT Atlantique, France).
- Maëlle Freteault, PhD student (IMT Atlantique, France & Psychology, UdeM).

5.3.8 Video games

- Paul-Henri Mignot, Research Associate 2018-19 (CRIUGM & IMT Atlantique, France).
- Maximilien Le Clei, Master's Student (CRIUGM & DIRO, UdeM & MILA).

5.3.9 MRI

- Julien Cohen-Adad, Principal Investigator (CRIUGM & Polytechnique Montréal).
- Eva Alonso, Post-Doctoral Fellow (Polytechnique Montréal).

5.3.10 MEG

- Karim Jerbi, Principal Investigator (CRIUGM & Psychology, UdeM).
- Yann Harel, PhD Student (Psychology, UdeM).

5.3.11 Computing

- Tristan Glatard, Collaborator (Computer Science and Software Engineering, Concordia University).