uCLIMB Summary Report 2017-2019

Updated 8/15/19 by MJM

**Table of Contents**

1. Subject counts by training cohort
   1. uCLIMB\_Short
   2. uCLIMB\_Python
   3. Both trainings
2. Measures collected by training cohort
   1. uCLIMB\_Short
   2. uCLIMB\_Python
3. Data locations and directory structure explanation by cohort
   1. Data structure overview
4. EEG Analyses
   1. Cleaned EEG data
   2. Directory structure
      1. Data directory structure
      2. Analysis directory structure
   3. Cleaning and analysis procedures
5. fMRI Analyses
   1. Analyses ran to-date
   2. Directory structure: imaging data
   3. Directory structure: behavioral data
6. sEBR Analyses
   1. Directory structure & analyses ran to-date
7. ***Subject Counts by Training Cohort***
   1. *Subject Count – uCLIMB\_Short*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Total Subjects\* | Drop-outs | Usable eyes-closed EEG | Usable eyes-open EEG | Usable learning EEG | Usable structural  MRI | Usable resting state MRI | Usable CatLoc MRI |
| 54 | 7 | 42 | 47 | 45 | 46 | 44 | 44 |

\*Drop-outs are included in total count

* 1. *Subject Count – uCLIMB\_Python*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Total Subjects\* | Drop-outs | Usable eyes-closed EEG | Usable eyes-open EEG | Usable learning EEG | Usable structural  MRI | Usable resting state MRI | Usable CatLoc MRI |
| 42 | 5 | 36 | 34 | 34 | 39 | 37 | 36 |

\*Drop-outs are included in total count

* 1. *Subject Count – Both Trainings*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Total Subjects\* | Drop-outs | Usable eyes-closed EEG | Usable eyes-open EEG | Usable learning EEG | Usable structural  MRI | Usable resting state MRI | Usable CatLoc MRI |
| 20 | 0 | 17 | 16 | 15 | 19 | 17 | 17 |

\*Drop-outs are included in total count

\*Complete list of usable subjects by cohort, entitled *uCLIMB\_Usable\_Subject\_List.xlsx*, can be found here [G:\My Drive\CCDL Shared\Expt\UCLIMB](../UCLIMB)

1. ***Measures Collected by Training Cohort***
   1. *uCLIMB Short*

* Pre-Test
  + MLAT
* Nelson Denny Comprehension
* Nelson Denny Vocabulary
* Nelson Denny Reading Rate
* Ravens Advanced Progressive Matrices (shortened)
* Reading Span (shortened, from different paper than Op and Sym Span)
* Symmetry Span (shortened)
* Operation Span (shortened)
* Simon (25%/75%)
* 3Back
* PSS
* Attentional Blink
* LLAMA (only 32 subjects completed)
* Handedness Inventory
* ALEAP-Q
* EEG
  + Eyes-closed resting state (5 minutes)
  + Eyes-open resting state (7 minutes for most subjects a few are only 5 minutes from the beginning of the cohort)
  + Lang 2 learning data (time variable)
    - Lang 1 (from lesson 1) data also collected, by Lang2 is a better estimate of learning
  + sEBR (7 minutes for most subjects a few are only 5 minutes from the beginning of the cohort, middle 5 minutes analyzed)
* fMRI
  + MPRAGE Structural (5 minutes)
  + Resting State (7 minutes)
  + CatLoc (2 Blocks, 8 minutes each)
* OLCTS Training
  + Eight 30 minute sessions
  + Quiz scores and speaking accuracy recorded
* OLCTS Outcome Measures
  + L2 learning rate
  + French Post-Test
    - F2E 🡪 Free response French to English translation questions; 28 questions
    - E2F 🡪 Multiple choice English to French translation questions; 28 questions
    - GR 🡪 Multiple choice grammar questions; 14 questions
    - SC 🡪 Multiple choice sociocultural questions; 14 questions
    - \*Test questions span lessons 1-10 in the OLCTS software
    - \*\* Because this was a set-level test (everyone took the same test regardless of how many lessons they completed), we computed several kinds of accuracy: “Total” refers to the total number items a subject got correct, “Old” refers to the number of questions that a subject got correct that they had been exposed to (i.e. the question was derived from a lesson that the subject had completed during the training), and “New” refers to the number of questions that a subject got correct that they had not previously been exposed to (i.e. the question was derived from a lesson that the subject had not completed during training)
    - \*\*\*The “Vocab” measure averages scores for the F2E and E2F portions of the test.
    - \*\*\*\*SC questions did not correlate or predict much of anything so these are usually left out of analyses.
  1. *uCLIMB Python*
* Pre-Test
  + MLAT
  + Nelson Denny Vocabulary
  + Ravens Advanced Progressive Matrices (shortened)
  + Reading Span (shortened, from different paper than Op and Sym Span)
  + Symmetry Span (shortened)
  + Operation Span (shortened)
  + Simon (25%/75%)
  + 3Back
  + PSS
  + Attentional Blink
  + LLAMA (only 12 subjects completed, 8 additional only LLAMA F)
  + Numeracy (only 33 completed)
  + Handedness Inventory
  + ALEAP-Q
* EEG
  + Eyes-closed resting state (5 minutes)
  + Eyes-open resting state (7 minutes for most subjects a few are only 5 minutes from the beginning of the cohort)
  + Python learning data (time variable)
  + sEBR (7 minutes for most subjects a few are only 5 minutes from the beginning of the cohort, middle 5 minutes analyzed)
* fMRI
  + MPRAGE Structural (5 minutes)
  + Resting State (7 minutes)
  + CatLoc (2 Blocks, 8 minutes each)
* Python Training
  + Ten 45-minute sessions
  + Screen recordings collected (used to verify accurate help use reporting)
  + Help resources: hints, community forums, and solution button uses
  + Quiz scores recorded
* Python Outcome Measures
  + Python learning rate
  + Rock, Paper, Scissors coding project (30 minutes allotted)
  + Multiple choice python test (30 minutes allotted)
    - “Semantic” questions; 25 questions
    - “Syntax” questions; 25 questions
* STAG (14 subjects participated from uCLIMB\_Python cohort)
  + Artificial Grammar Task
  + Working Memory Reinforcement Learning Task
  + Weather Prediction Task

1. ***Data Locations and Directory Structure***
   1. *Overview of uCLIMB Directory Structure*

* UCLIMB Directory: All uCLIMB related analysis files are stored in the following directory: [G:\My Drive\CCDL Shared\Expt\UCLIMB](file:///G:\My%20Drive\CCDL%20Shared\Expt\UCLIMB)
  + Any files that contain information about more than one training cohort are located in the general UCLIMB directory hyperlinked above (this includes the usable subject list, summary sheets for subjects completing both Python and Language Training, and summary sheets for all subjects
* There are four sub-directories listed within the UCLIMB folder:
  + uCLIMB\_Long(OLCTS)
  + uCLIMB\_Short(OLCTS)
  + uCLIMB\_Python
  + uCLIMB Team Repositories
* Raw data are stored in the following directories:
  + [G:\My Drive\CCDL Data Backup\uCLIMB\_Python](../../../CCDL%20Data%20Backup/uCLIMB_Python) Python subjects (283\*\* & 292\*\*)
  + [G:\My Drive\CCDL Data Backup\uCLIMB\_Short](../../../CCDL%20Data%20Backup/uCLIMB_Short) OLCTS subjects (282\*\*)
  + [G:\My Drive\CCDL Data Backup\uCLIMB\_Winter19](../../../CCDL%20Data%20Backup/uCLIMB_Winter19) OLCTS returning subjects (293\*\*)
  + There should also be copies of all raw data files in the following locations:
    - [G:\My Drive\CCDL Shared\Expt\UCLIMB\uCLIMB\_Python\Tasks](uCLIMB_Python/Tasks)
    - [G:\My Drive\CCDL Shared\Expt\UCLIMB\uCLIMB\_Short(OLCTS)\Tasks](uCLIMB_Short(OLCTS)/Tasks)

1. ***EEG Analyses*** 
   1. *Cleaned EEG Data*

\*\*\*The most recent data summary including the cleaned EEG output (individual channel power, network power and coherence, and alpha laterality for each EEG recording type) is labeled [Cohort]\_Summary\_AllEEG\_[date].xlsx and can be found in the ‘Analyses’ directory of each cohort or in the general ‘UCLIMB’ directory for a summary that includes both cohorts.

* 1. *Overview of EEG Directories*

*4.1.1. Cleaned Data Summary*

* EEG data is stored in the following directory for uCLIMB\_Short participants: [G:\My Drive\CCDL Shared\Expt\UCLIMB\UCLIMB\_Short(OLCTS)\EEG\Data](UCLIMB_Short(OLCTS)/EEG) and in the following directory for uCLIMB\_Python participants: [G:\My Drive\CCDL Shared\Expt\UCLIMB\UCLIMB\_Python](UCLIMB_Python/EEG)
  + Data is then further separated by recording type (Eyes-Closed, Eyes-Open, Learning Data), stored in the Analyzed directory, then separated by subject, and then separated by Spectrum and Coherence.

*4.1.2. Analysis Directories*

* The main EEG Analysis directory contains scripts and analysis templates and can be found here for uCLIMB\_Short participants: [G:\My Drive\CCDL Shared\Expt\UCLIMB\UCLIMB\_Short(OLCTS)\EEG\Analysis](UCLIMB_Short(OLCTS)/EEG/Analysis) and here for uCLIMB\_Python participants: [G:\My Drive\CCDL Shared\Expt\UCLIMB\UCLIMB\_Python\EEG\Analysis](UCLIMB_Python/EEG/Analysis)
  + Analyses are then further separated into subdirectories by recording type (i.e. analyses pertaining to eyes-closed data).
  + The Visual Inspection subdirectory is organized by data type, and contains converted .edf files which can be visualized using edf browser (see cleaning protocol for more details).
  1. *Data Cleaning Procedure*
* Copies of all original cleaning and analysis scripts and instructions on how to use them are located in the following directory: [G:\My Drive\CCDL Shared\Expt\UCLIMB\EEG Cleaning & Analysis Scripts](EEG%20Cleaning%20&%20Analysis%20Scripts)
* 1) All EEG data were first ran through Andrea’s R Script (eeg.analysis.3.1.3.R.txt).
  + [G:\My Drive\CCDL Shared\Expt\UCLIMB\EEG Cleaning & Analysis Scripts\Cleaning Scripts\Step1\_AndyCleaningScript](EEG%20Cleaning%20&%20Analysis%20Scripts/Cleaning%20Scripts/Step1_AndyCleaningScript)
* 2) All raw EEG files must be converted to .edf files and visualized using EDF Browser to identify bad channels.
  + [G:\My Drive\CCDL Shared\Expt\UCLIMB\EEG Cleaning & Analysis Scripts\Cleaning Scripts\Step2\_IdentifyBadChannels](EEG%20Cleaning%20&%20Analysis%20Scripts/Cleaning%20Scripts/Step2_IdentifyBadChannels)
* 3) Concatenate the output from Andrea’s script to enter into the cleaning template
  + [G:\My Drive\CCDL Shared\Expt\UCLIMB\EEG Cleaning & Analysis Scripts\Cleaning Scripts\Step3\_ConcantenateSummary](EEG%20Cleaning%20&%20Analysis%20Scripts/Cleaning%20Scripts/Step3_ConcantenateSummary)
* 4) The concatenated summary output from the script was entered into the EEG Cleaning Template, and bad channels are removed based on the previously completed visual inspection in EDF Browser (Step 2).
  + [G:\My Drive\CCDL Shared\Expt\UCLIMB\EEG Cleaning & Analysis Scripts\Cleaning Scripts\Step4\_CleaningTemplate](EEG%20Cleaning%20&%20Analysis%20Scripts/Cleaning%20Scripts/Step4_CleaningTemplate)
* Networks were sometimes calculated using EEG\_Network\_Syntax script in SPSS. These scripts can be found here [G:\My Drive\CCDL Shared\Expt\UCLIMB\EEG Cleaning & Analysis Scripts\Network Scripts](EEG%20Cleaning%20&%20Analysis%20Scripts/Network%20Scripts)
* Individualized bands based on alpha peak frequency can also be calculated using the formula here [G:\My Drive\CCDL Shared\Expt\UCLIMB\EEG Cleaning & Analysis Scripts\IAF Scripts](EEG%20Cleaning%20&%20Analysis%20Scripts/IAF%20Scripts)

1. ***fMRI Analyses***
   1. *Analyses ran to date*

* CatLoc
  + Preprocessing & artifact correction
  + First level – individual subjects
  + Second level – group data
  + Correlations – GLM individual differences analyses
  + ROI analyses – extract betas
  + Timecourses – extract percent signal change in ROIs
* Resting State
  + Preprocessing & artifact correction
  + Seed analyses (Margarita)
  + Other??
* Structural VBM
  1. *Directory structure - data on Google Drive*
* The ‘uCLIMB fMRI Analysis’ folder contains two subdirectories: ‘Analysis Instructions’ contains instructions on how to run analyses in SPM using the batch scripts, and ‘Analysis Outputs’ contains analysis outputs separated by analysis type. The directory can be found here: [G:\My Drive\CCDL Shared\Expt\UCLIMB\uCLIMB fMRI Analyses](uCLIMB%20fMRI%20Analyses)
* CatLoc behavioral data is stored in the following directory for OLCTS subjects [G:\My Drive\CCDL Shared\Expt\UCLIMB\uCLIMB\_Short(OLCTS)\fMRI\Experiments\CategoryLocation](uCLIMB_Short(OLCTS)/fMRI/Experiments/CategoryLocation) and the following directory for Python subjects [G:\My Drive\CCDL Shared\Expt\UCLIMB\uCLIMB\_Python\fMRI\Experiments\Category Location](uCLIMB_Python/fMRI/Experiments/Category%20Location)
  1. *Directory structure – data on Ceres*
* All of the data for both the OLCTS and Python cohorts is stored in the following directory on the server: /projects/uCLIMB\_Short
  + Raw scan data is stored in the ‘scans’ directory
* All CatLoc data is stored in the following directory /projects/uCLIMB\_Short/CatLoc
* All analyses in CatLoc have been run using the contrasts\_MM.txt file (this file has several additional contrasts included that are not in the original contrasts.txt file)
* Because there were some systematic differences between OLCTS\_Long subjects ran at DISC and OLCTS\_Short/Python subjects ran at BMIC. I used the BMIC group for my analyses, the BMIC\_Subjects file should detail which subjects were used for these analyses.
* To determine which subjects were usable check the uCLIMB\_Usable\_Subject\_List\_[date].xlsx found here [G:\My Drive\CCDL Shared\Expt\UCLIMB](../UCLIMB). Reasons subjects were excluded from analyses included motion artifacts, scanner artifacts, and failing to meet CatLoc behavioral criteria.
  + If using the folder view on Ceres you can see that some individual subject folders are red (these subjects have bad scan data) and some are yellow (these subjects have good scan data but did not meet our behavioral criteria for CatLoc).

1. ***sEBR Analyses***
   1. *Directory Structure & Analyses Ran To-Date*

* All sEBR data is stored in the following directory: [G:\My Drive\CCDL Shared\Expt\UCLIMB\uCLIMB sEBR](uCLIMB%20sEBR), there are also copies of this data split by cohort in the individual cohort directories.
  + Raw sEBR is stored in either the ‘SavedData\_5minute’ of the ‘SavedData\_7minute’ directory. Subjects from early in the OLCTS cohort only have a 5minute sEBR collection, we decided to increase this to 7minutes partway through the OLCTS cohort, so all of the Python subjects should have 7minute recordings.
  + There are two versions of the blink detect script ‘blinkDetect\_v4.4.ipynb’ for the 5 minute data files and ‘blinkDetect7mins\_v4.4.ipynjb’ for the 7 minute data files (the 7 minute script analyses only the first 5 minutes of data)
    - These scripts run through Jupiter Notebook and output plots with blink counts to the ‘tsPlots’ directory. See the ‘Read Me’ file for more detailed instructions about using the blink detect script.
* Unfortunately, the script does not perfectly detect blinks so some manual recounting may need to be done. We were never able to come to a consensus of how to do the manual recounting so I only included sEBR counts in the data summaries for subjects whose counts were properly counted by the blink detection algorithm.
  + A list of subjects that needed to be manually recounted (and initial attempts at recounting) can be found here. https://docs.google.com/spreadsheets/d/1Nz1NJf57QVOy2uOdDwket6Dju3-f4An0RBBFef\_Jhdo/edit?usp=sharing