**Reliability, Effect size, And Data quality In EEG (READIE) toolbox**

Wenyi Xu

READIE Toolbox User Manual

Updated July 30th 2024

1. **Introduction**

The READIE Toolbox is a user-friendly computational solution designed to enhance the reporting and analysis of EEG data quality metrics. Developed with a user-friendly interface, it has been integrated into HAPPE EEG preprocessing pipelines (Gabard-Durnam et al., 2018; Monachino et al., 2022) facilitating the automated estimation of reliability, effect size, and Standard Measurement Error (SME) (Luck et al., 2021)for EEG datasets. This innovative tool simplifies the computational process and aids in interpreting data quality metrics, such as effect size, reliability, and SME, providing insights into the optimal number of trials for detecting significant effects and establishing reliable measures. By offering overall and bootstrapped reliability, effect size across increasing numbers of trials, and SME estimates at the participant level, we aim to expand the current set of toolboxes for EEG data quality metric analysis, providing a toolbox that is intuitive and user-friendly for researchers.

A diagram of a software

Description automatically generated

Figure 1: READIE Overall Flowchart

**Download the Toolbox**

The READIE Toolbox can be downloaded here: https://github.com/Bead-Lab/The-READIE-Toolbox-Reliability- Effect-size-And-Data-quality-In-EEG. The toolbox has been tested with MATLAB R2022 on macOS Ventura Version 13.4.1(c).

For READIE to run properly, **Version MATLAB R2022** should be used, and MATLAB add-ons, including **Parallel Computing Toolbox, Statistics and Machine Leaning Toolbox, and Signal Processing Toolbox** must be downloaded ahead before running READIE.

The updated version of HAPPE with data quality metrics calculation capabilities can be found here: https://github.com/PINE-Lab/HAPPE.

1. **Analyzing Data**

**2.1 Preparing Data**

The READIE Toolbox works with single-subject, single-trial ERP mean amplitude measurements. It can easily read in the HAPPE + ER output (Gabard-Durnam et al., 2018; Monachino et al., 2022) or any output format similar to HAPPE's: data folder with single-subject, trial-level information.

**For datasets that’s been processed by the generateERP script**: the choices for generateERP script should be as follows so READIE can read them:

Trial type: Individual

Choose your export format: Multiple .csv files

Subjects: Rows as trials, columns as values split by subject

**For datasets that has been preprocessed by other pipeline**: READIE can take any output format similar to HAPPE's: data folder with single-subject, trial-level information.

For demonstration purposes, we used publicly available Visual-Evoked Potential (VEP) Event-Related Files from the General Anesthesia and Brain Activity (GABA) Study dataset (Monachino et al., 2021) that can be found here: <https://zenodo.org/records/5172962>. The data has been preprocessed using HAPPE + ER software, and the processed data can be found within the datafile: ERP\_calculateVals\_data folder.

Within this folder, each CSV file contains the trial-level information for each participant (Figure 1).

A screenshot of a computer

Description automatically generated

Figure 2: Each csv file represents one participant.

\*Note: For datasets with more than one condition, the condition information must be consistently present in the filename with observable pattern, e.g., if we have two conditions, correct and incorrect, the filenames should look something like this:

1212\_generatedERPvals\_correct.csv

1213\_generatedERPvals\_correct.csv

1212\_generatedERPvals\_incorrect.csv

1213\_generatedERPvals\_incorrect.csv

Within each participant's individual CSV files, each line should correspond to one trial, and there can be multiple measurements per line, e.g. mean amplitude for windows 75 – 130 and 100 – 230 (Figure 2).

A table of numbers and symbols

Description automatically generatedFigure 2: Participant Trial Level Example

**2.2 Start the Toolbox**

To start the READIE Toolbox, open the downloaded MATLAB file master.m.

A screenshot of a computer program

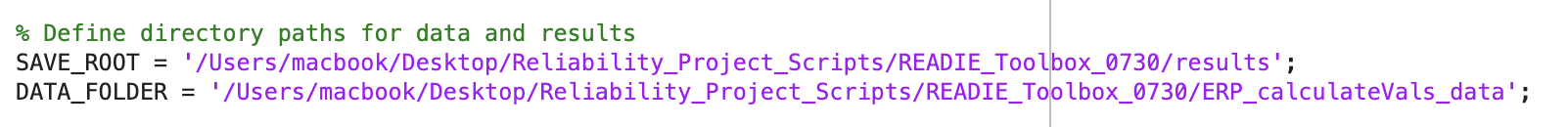
Description automatically generated

Figure 3: The READIE Toolbox Setup Script

In the setup script, you will be inputting the following information:

**Folder and file setup:**

SAVE\_ROOT: the location where you want to save the files.

DATA\_FOLDER: the location where your participant trial-level data is stored.

FILENAME\_DIVIDER: any text before the divider will be treated as the subject name. For example:

% Extract participant name from file names using the specified divider

% e.g., for "2\_191\_49685484\_3\_20220901\_110819\_generatedERPvals\_27-02-2024.csv",

% participant name is "2\_191\_49685484\_3\_20220901\_110819"

FILENAME\_DIVIDER = "\_generatedERPvals";

IGNORE\_CONTAINS: Put the name of the files that you want to exclude from analysis. Don’t include the .csv extension.

**Conditions in file:**

CONDITIONS: List ALL conditions of your dataset.

If you only have one condition in your dataset, DON'T put anything into it.

A close-up of a computer screen

Description automatically generated

**Column(s) of interest:**

VALUE\_COLUMNS: Specify columns that you are interested in for data analysis.A close-up of a window

Description automatically generated

**Parameters for bootstrapping:**

Set the number of iterations you want for data analysis.

NUM\_ITERATIONS = 1000; A close-up of a white background

Description automatically generated

**Parameters for reliability estimates:**

N\_FROM = 10; \* Start of bootstrap trials.

N\_TO = 100; \* End of bootstrap trials.

N\_BY = 10; \* Increments of bootstrapping.

In this example, reliability and effect size will be calculated at intervals of 10, 20, 30, ... up to 100 trials.

A close-up of a white background

Description automatically generated

**Other Parameters:**

Other parameters, including color for graph plotting and the number of CPUs, have also been set in the script (no need to modify).

**A screenshot of a computer code

Description automatically generated**

After updating all the information, click **'Run'** in MATLAB.

In the command window, you will see the progress of reliability, effect size and SME computation. A screenshot of a computer program

Description automatically generated

**2.3 Review Results**

Go to your designated folder to view results based on your selection. A complete set of results of the sample dataset includes:

Each column of interest is the folder name: **A screenshot of a phone

Description automatically generated**

Figure 6: Sample Output Folder Layout

Each folder contains:

Effect size (overall and trial level).

Summary of effect size (overall and trial level).

Reliability (overall and trial level).

Summary of reliability (overall and trial level).

Graphs of trial-level reliability and effect size.

SME per participant.

**A screenshot of a computer

Description automatically generated**

Figure 7: Sample Output

**Citations**

Gabard-Durnam, L. J., Mendez Leal, A. S., Wilkinson, C. L., & Levin, A. R. (2018). The Harvard Automated Processing Pipeline for Electroencephalography (HAPPE): Standardized processing software for developmental and high-artifact data. *Frontiers in Neuroscience*, *12*, 97.

Luck, S. J., Stewart, A. X., Simmons, A. M., & Rhemtulla, M. (2021). Standardized measurement error: A universal metric of data quality for averaged event‐related potentials. *Psychophysiology*, *58*(6), e13793. https://doi.org/10.1111/psyp.13793

Monachino, A. D., Lopez, K. L., Pierce, L. J., & Gabard-Durnam, L. J. (2022). The HAPPE plus Event-Related (HAPPE+ER) software: A standardized preprocessing pipeline for event-related potential analyses. *Developmental Cognitive Neuroscience*, *57*, 101140. https://doi.org/10.1016/j.dcn.2022.101140

Monachino, A. D., Lopez, K. L., Underwood, E., Tao, A., Nelson, C., Berde, C., Cornelissen, L., Hensch, T., & Gabard-Durnam, L. (2021). *Visual-Evoked Potential (VEP) Event-Related Files from the General Anesthesia and Brain Activity (GABA) Study and Infant Sibling Project (ISP)* (HAPPE 2.0) [dataset]. Zenodo. https://doi.org/10.5281/zenodo.5931539