

---

# EEG Thinking1 Datasets: Think-Count-Recall (TCR) and Read-Write-Type (RWT)

---

**Xiaodong Qu, Peiyan Liu**  
Department of Computer Science  
Brandeis University  
Waltham, MA 02453  
xiqu,peiyanliu@brandeis.edu

## Abstract

1 EEG-based Brain-Computer Interfaces (BCI) have been widely used in clinical  
2 and non-clinical research. In this paper, we present a framework to collect a  
3 large amount of EEG data with easy-to-use experiment setup, using non-invasive,  
4 wireless, and affordable hardware. Interpretable feedback generated by benchmark  
5 machine learning algorithms have been provided to the researchers and end-users.  
6 Two existing datasets are used as case studies for the framework: Read-Write-Type  
7 (RWT) and Think-Count-Recall (TCR). The goal is to inspire new machine learning  
8 approaches for decoding behavior from large-scale EEG data. The framework  
9 of experimental design, data collection, data analysis, feedback generation, and  
10 community building could pave the way towards a future when everyone can easily  
11 use BCI systems every day, similar to smartphones nowadays.

## 12 1 Introduction

13 Neural interfaces are becoming of increasing interest to industry and having large available datasets  
14 could be useful for students and researchers to tease out signals from noisy data. Brain Computer  
15 Interfaces (BCI) have been widely used for both clinical and non-clinical applications (Lotte et al.  
16 [2018a], Craik et al. [2019]), such as diagnosis of abnormal states, evaluating the effect of the  
17 treatments, helping patients with motor disabilities to move a mouse or to control a motorized  
18 wheelchair, mental workload, seizure detection, motor imagery tasks (Devlaminck et al. [2010]), BCI  
19 based games (Coyle et al. [2013]) and passive BCI. Previous research has reviewed existing datasets  
20 in the BCI field, such as Schalk et al. [2004], Lotte et al. [2007], Zhang et al. [2020], Roy et al. [2019],  
21 Miller [2019], Kaya et al. [2018], most of the datasets mentioned are collected in research labs or  
22 clinical settings with expensive medical equipment and time-consuming setup procedure, under the  
23 supervision of clinical professionals. The data collection framework we proposed allows non-expert  
24 participants to run the experiment by themselves at home, whenever they have a small amount of time,  
25 such as twenty minutes. The visual feedback generated by benchmark machine learning algorithms  
26 could help them to perform better in the future sessions.

27 Considering classic datasets in other domains, such as ImageNet for image classification, or MNIST  
28 for handwritten digit recognition, more data can be generated directly from the non-expert end users,  
29 and more general patterns could be recognized based on such large scale data. With the motivation  
30 to gather EEG data with a cheaper, easier and faster approach, we designed a pilot study towards  
31 building a large-scale EEG data set, for multi-class classification of user-centered tasks, generated by  
32 non-expert end-users. Results of classification with the proposed new data and machine models show  
33 a reasonable accuracy (70% to the random 20%), indicating the potential of this framework.

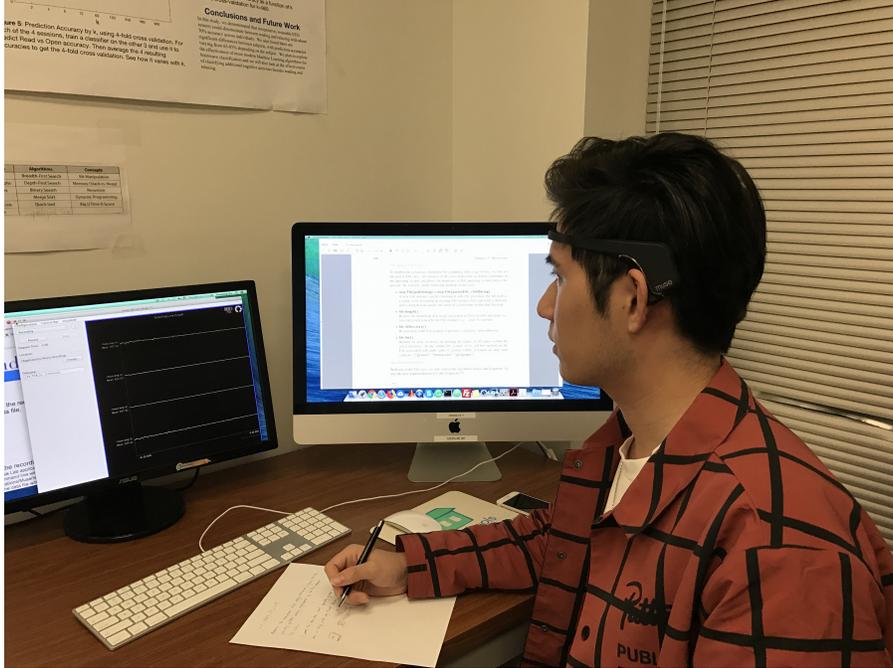


Figure 1: A researcher demonstrating the task "Write", wearing EEG headset.

34 This is an ongoing project and this 'Thinking1' repository currently has four datasets: Read-Write-  
 35 Type (RWT, Qu et al. [2020b]), Think-Count-Recall (TCR, Qu et al. [2020a]), Python-Math (Qu et al.  
 36 [2018b]), and GRE-Relax (Qu et al. [2018a]). In this paper, we use the two recent experiments, RWT  
 37 and TCR, as examples to explain the approach. Details about the data collection, including how the  
 38 subjects were recruited under IRB requirements, how long each session was, what kind of visualized  
 39 feedback is provided to the subjects, how many EEG sessions were recorded, data cleaning, feature  
 40 extraction, and research from benchmark algorithms, are in our previous papers, we also attached an  
 41 updated version in the appendix section of this paper to allow readers to replicate these experiments.

## 42 1.1 Read-Write-Type (RWT)

43 Previous studies (Bird et al. [2018], Qu et al. [2018a]) demonstrated that EEG signals could success-  
 44 fully distinguish several kinds of cognitive tasks. Such as programming in Python vs. solving Math  
 45 problems; solving Math problems (GRE) vs. solving Reading problems (GRE). These experiments  
 46 focused on distinguishing different cognitive tasks, but not on whether different communication  
 47 modes may also have a distinguishable impact on EEG patterns. The experiment RWT (Qu et al.  
 48 [2020b]) in this data set was designed to test the hypothesis of whether AI based EEG markers  
 49 could distinguish both between two modes of communication: typing vs. writing, and between three  
 50 cognitive states: reading vs. copying vs. answering. The five tasks are described in Figure 2.

## 51 1.2 Think-Count-Recall (TCR)

52 Other studies (Lotte et al. [2018a], Lotte [2015], Bird et al. [2018], Qu et al. [2018a], Craik et al.  
 53 [2019]) demonstrated that EEG classification was successfully used to distinguish multiple cognitive  
 54 tasks. In the TCR experiment (Qu et al. [2020a]), we designed these five user-centered tasks as shown  
 55 in Figure 3, abbreviated them as Think (T), Count (C), Recall (R), Breathe (B) and Draw (D). The  
 56 task selection is motivated by human memory experiments such as Kahana et al. [2018].

## 57 2 Methods

58 Such datasets are suitable for machine learning due to its high dimensional and noisy nature, similar  
 59 to image recognition problems. There is great potential to provide higher accuracy and more

60 interpretable feedback to both researchers and end-users. For example, in each data point of 1/10  
61 second, the raw EEG data is a 4 x 5 matrix, which represents four electrodes and five frequency  
62 bands. Such twenty-dimensional data performs well enough (compare to 64 or 128 electrodes medical  
63 devices) when applied to mainstream EEG-related machine learning or deep learning algorithms.

64 Each session of these experiments are reproducible with twenty minutes of effort for non-experienced  
65 end-users. These human-in-the-loop experimental designs motivated by (LaRocco et al. [2020], Lotte  
66 et al. [2018b]), have several advantages. First, the tasks are selected more from the end-users, less  
67 from the researchers, similar to the smartphone usage situation now. Secondly, the role of the EEG  
68 coach can make the end-user experience much better. Last but not the least, easy-to-understand user  
69 feedback could be helpful for the end-user to reduce the noise and focus more on the designed tasks.  
70 More details in the previous papers and the appendix section of this paper.

## 71 2.1 Experimental Design

72 The experimental design is easy to adapt, and the three hundred dollars or less wireless hardware, as  
73 mentioned in Ienca et al. [2018], makes it affordable to a broader audience. For example, our research  
74 lab has expanded the experiments from just targeting less than twenty students, to a community of  
75 more than one hundred students, each of them starts with little or no computer science or neuroscience  
76 background, and usually, after at most two to three twenty-minute sessions, they can learn to how to  
77 control the noise level, and achieve the desired experimental goals with high accuracy.

78 The sensor hardware research and development have grown rapidly recently (Kübler et al. [2014],  
79 Tabar and Halici [2016] ), so does the trend of making it more affordable to the non-expert users.  
80 After comparing several options, such as devices mentioned in Ienca et al. [2018], we chose the Muse  
81 Headset for our experiments, with an affordable price of less than three hundred for each wireless  
82 headset. For the design of the tasks, previous research has shown deep learning works well in emotion  
83 recognition, motor imagery, mental workload, and seizure detection areas (Craik et al. [2019]), we  
84 tried learning, motor-imagery tasks, sleep, and entertainment tasks. In this study, we focused on the  
85 learning related tasks college students perform often in their daily lives.

## 86 2.2 Data collection

87 Data was collected in non-clinical settings, partly in the reserved classrooms or conference rooms in  
88 the universities, partly at the participants' home. The size of the data usually is 15 to 20 subjects,  
89 five to six sessions for each subject, each sessions varies from five minutes to twenty minutes. For  
90 example, the TCR (16 subjects) and RWT (14 subjects) experiment each includes six sessions, each  
91 session is five minutes long. Comparing with existing experiments on cognitive tasks mentioned  
92 in Craik et al. [2019], Gabard-Durnam et al. [2018], Roy et al. [2019], Pernet et al. [2019], our  
93 experimental design and data collection is easier, cheaper and faster. With twenty-minute training,  
94 most participants can generate hours of EEG recording data at home with interpretable feedback.

95 The non-invasive, wireless EEG headset usually needs a training session to reduce the noise level.  
96 The role of EEG coach was created to smooth the learning curve for first time end-users. The  
97 end-users and EEG coaches are fairly compensated under the IRB requirement. More details such as

**Read (R)** Subjects were asked to read a PDF file displayed on the monitor silently, the PDF file is a  
computer science textbook on Data Structures (Sierra and Bates [2003]).

**Write Copy (WC)** Subjects wrote on a blank white paper with a pen, copying the text from the  
same textbook PDF file display on the monitor. As shown in Figure 1.

**Write Answer (WA)** Subjects wrote an essay using a pen on a blank paper, answering the question:  
'Why did you choose your major?'

**Type Copy (TC)** Subjects copied text from the same textbook PDF file, into a text entry box on the  
screen, by typing on a keyboard.

**Type Answer (TA)** Subjects typed their answers to the question 'What is your academic plan for  
this semester?' into a text entry box on the screen.

Figure 2: Tasks in experiment Read-Write-Type (RWT).

- Think (T)** Subjects were asked to think of several (six, seven, eight) random objects, these objects need to be independent of each other. For example, (Sun, Fish, Flower, Table, Student, Car), is a valid set, but (computer, keyboard, monitor, speaker, phone, TV) is not a valid set.
- Count (C)** Subjects counted numbers aloud, from 200 towards 0, each time subtracting by 7, e.g. 200, 193, 186, 179, with eyes open, eyes and jaws movement minimized.
- Recall (R)** Subject recalled the objects they had typed in the Think (T) task, in the correct order, if possible, and entered them in a similar text entry box with a keyboard.
- Breathe (B)** Subjects were instructed to breathe deeply with their eyes open. They were asked NOT to think about any other tasks in this experiment, or anything else except their breath.
- Draw (D)** Subjects were asked to draw the objects they thought about in the earlier task Think (T), with a pen, on a blank A4 paper. The objects text they just entered in T was displayed on the monitor, so they did not need to recall, just focus on drawing.

Figure 3: Tasks in experiment Think-Count-Recall (TCR).

98 IRB approval and instructions given to the participants are included in the appendix section of this  
 99 paper. Each headset was connected to a mainstream personal computer through Bluetooth. We use  
 100 the software package that comes with the EEG headset (Muse-io and MuseLab) to record the raw  
 101 EEG data to the computers. Then the data was processed and Analyze using machine learning and  
 102 deep learning algorithms. The visualized feedback is provided to the end-users, EEG coaches, and  
 103 researchers to improve the next round of data collection.

104 Before the experiment, the EEG coach helps the end-users to understand the IRB requirements and  
 105 make sure they sign the informed consent forms, then explain in detail to the end-users what they are  
 106 expected to see and to do during each step. During each session of the experiments, the EEG coach  
 107 leads the end-users to the experiment website to fill out the pre-experiment survey, then helps the  
 108 end-users to connect the EEG headsets and conduct a test recording for one minute before the official  
 109 EEG recording starts, A time-boxed online survey style guide was then used to give the end-user  
 110 step-by-step prompt during the experiment, the EEG coach is there for any possible questions. After  
 111 the experiment, The EEG coach makes sure the end-user fill out the post-experiment survey and  
 112 help them better understand the visual feedback. Also, the EEG coach keeps track of the notes for  
 113 the entire process and communicates with the researchers regularly to deal with pop-up issues and  
 114 maintain a frequent-asked-question (FAQ) list.

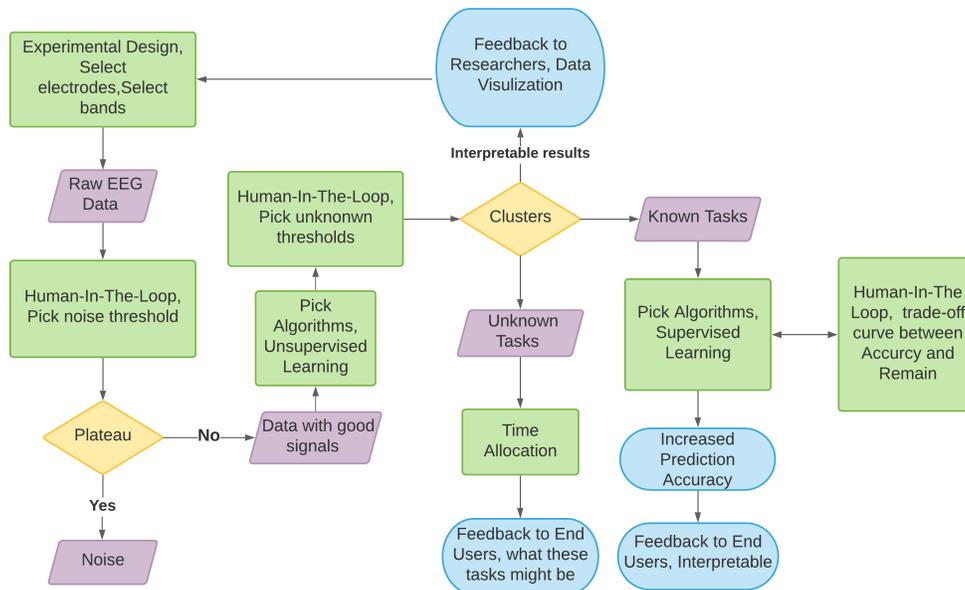


Figure 4: Data analysis framework.

115 **2.3 Data analysis framework**

116 New ways to analyze EEG data have been developed recently, such as Chevalier et al. [2020], Roc  
 117 et al. [2020], Sabbagh et al. [2019], Tu et al. [2019]. In our experiments, as Figure 4 shows, the raw  
 118 EEG data were first visualized to allow researchers to define the threshold to remove the noise, here  
 119 we used the plateau threshold to determine whether a certain time window of signals is considered  
 120 noise. Then we first used unsupervised learning algorithms, such as K-means, to cluster the data  
 121 points, then we visualized the clustering result and made it interpretable to the researchers and  
 122 end-user. For the designed tasks, we then used supervised learning, such as Random Forest or Long  
 123 Short-Term Memory (LSTM) to predict the tasks. For the unknown tasks, we put a marker on the  
 124 visual feedback to the end-user to ask what may happen during that time period. More details about  
 125 these steps are in the appendix section of this paper.

126 By Human-In-The-Loop (HITL) it means the researcher, EEG coaches, and end-users and making  
 127 data-driven decisions based on visualized feedback. The researchers bring together the existing  
 128 knowledge about this experiment and help the EEG coaches and end-users to perform better in  
 129 the next session. The EEG coaches guided the end-users to master how to use EEG hardware and  
 130 software necessary to perform the designed tasks. The end-users learn from the researchers and  
 131 the EEG coaches to use the EEG-based on BCI in their daily life as the experiments designed, and  
 132 provide valuable feedback to the researchers to develop new and improved experiments.

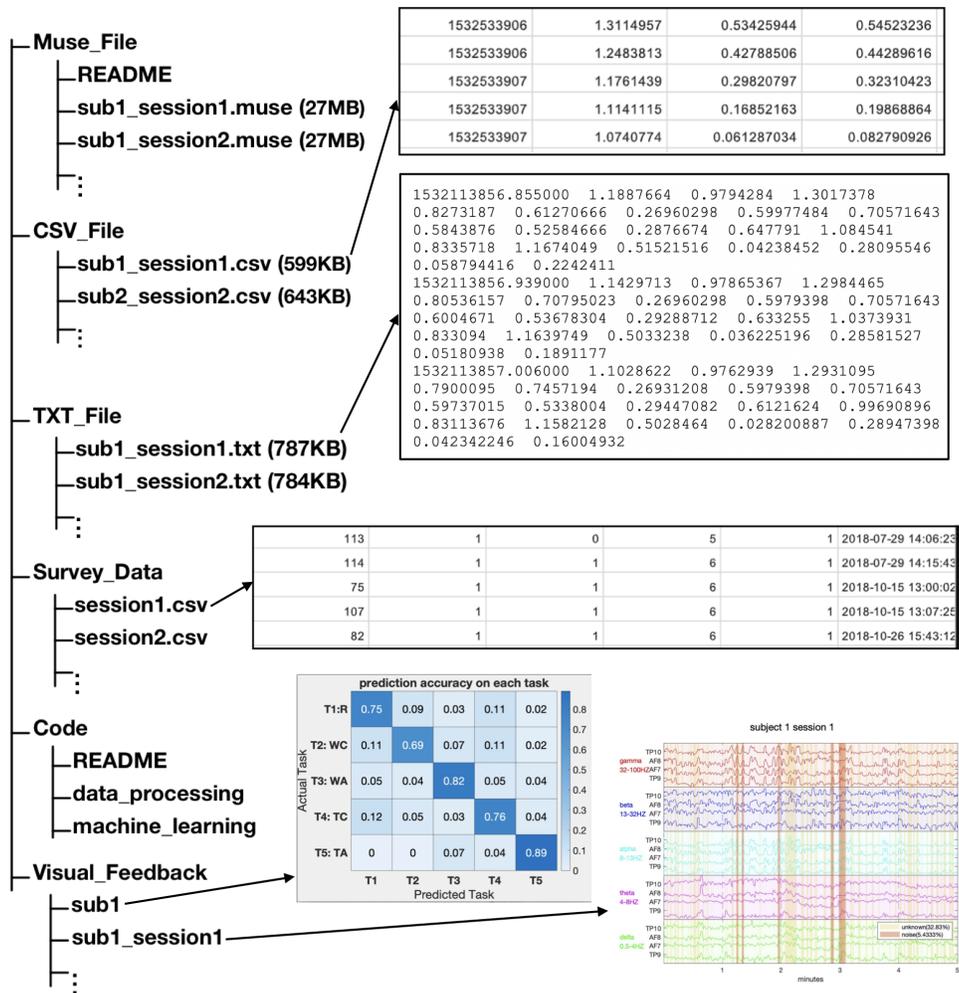


Figure 5: File structure of Our dataset

## 133 **2.4 Data format**

134 Pernet et al. [2019], and Nichols et al. [2017] have presented several recommended practices about  
135 EEG data formats and sharing. Our data set, as Figure 5 shows, consists of the original MUSE files,  
136 and CSV files, TXT files after pre-processing. Also, the metadata collected through Qualtrics online  
137 survey system has been included, as well as the code has been implemented for this dataset. For  
138 example, for the Read-Write-Type (RWT) experiments, for each subject, each five-minute session,  
139 there is a MUSE data file size of around 27M, and after pre-processing, the MUSE file is converted  
140 to a CSV or TXT file for further analysis, with a much smaller size of about 700K. Then there are  
141 folders of suvery metadata and related code.

## 142 **2.5 Machine Learning applications**

143 In this paper, we introduce a machine learning benchmark for predicting the task humans are engaged  
144 in from the EEG. We presented what machine learning and deep learning algorithms have been  
145 applied to these datasets, and suggest several recommended practices for these datasets.

146 For the pre-processing part, data visualization is helpful for noise detection. For the multi-class  
147 classification, ensemble methods, such as random forest, and Recurrent Neural Network (RNN), such  
148 as Long Short-Term Memory (LSTM) consistently outperformed other classifiers (Qu et al. [2020b]),  
149 we suggest using them as benchmark algorithms. Building on top of that, we proposed our algorithm,  
150 Time-Continuity-Voting (TCV, Qu et al. [2020a]), which achieved the highest prediction accuracy for  
151 these datasets. More details are in the appendix section of this paper.

## 152 **2.6 Community Forum and further support**

153 We established an online forum for the community who works on these datasets, including researchers,  
154 EEG coaches, end-users, and clinical professionals. Due to our IRB requirements, this forum is  
155 invitation-only at this time. Through our BCI forum, we connected to three computer science labs,  
156 two neuroscience labs, two clinical research labs, and two hospitals during the last three years, as  
157 well as get more than a hundred undergraduate students involved as experiment participants, eight of  
158 them later became EEG coaches.

159 We held discussions on how to improve EEG experimental design and dataset development. Further  
160 support on how to explore the potential of such an EEG-based BCI system is encouraged based on  
161 community members' availability. Also, we are presenting these research papers and this forum to  
162 more college students in the computer science and Neuroscience courses we lectured each semester.

163 Participating in the existing BCI community and bridging our own small EEG-based BCI community  
164 to a broad network is also an important direction.

## 165 **2.7 Availability and Ethical considerations**

166 To make sure these datasets would be used ethically and responsibly, we adapted several recommended  
167 practices of sharing BCI data, such as Gabard-Durnam et al. [2018], Pernet et al. [2019]. According  
168 to our IRB requirement, these data sets are available upon written request, we review the request to  
169 make sure it is coming from a reputable research institution and the requester is willing to sign a  
170 Non-Disclosure Agreement. Previous studies have reviewed freely available EEG datasets, such as  
171 Zhang et al. [2020], Roy et al. [2019], Miller [2019], Kaya et al. [2018], Craik et al. [2019], we are  
172 amending our IRB to find acceptable ways of data anonymization to share it more freely.

## 173 **3 Results**

174 We develop feedback for different user roles. For example, the figures that compare different end-  
175 users or different machine learning algorithms are more for the researchers, optional for the end-users.  
176 Here are some sample feedback figures we provide to our researchers, EEG coaches and end-users.

177 **3.1 For Researchers**

178 For cross-subject comparison, as Figure 6 shows, although there are individual differences, the task  
 179 prediction accuracy is reasonably high. Together with Figure 7, (both figures X-axis is subject id  
 180 ordered by prediction accuracy), we observed the noise and unknown tasks vary across different

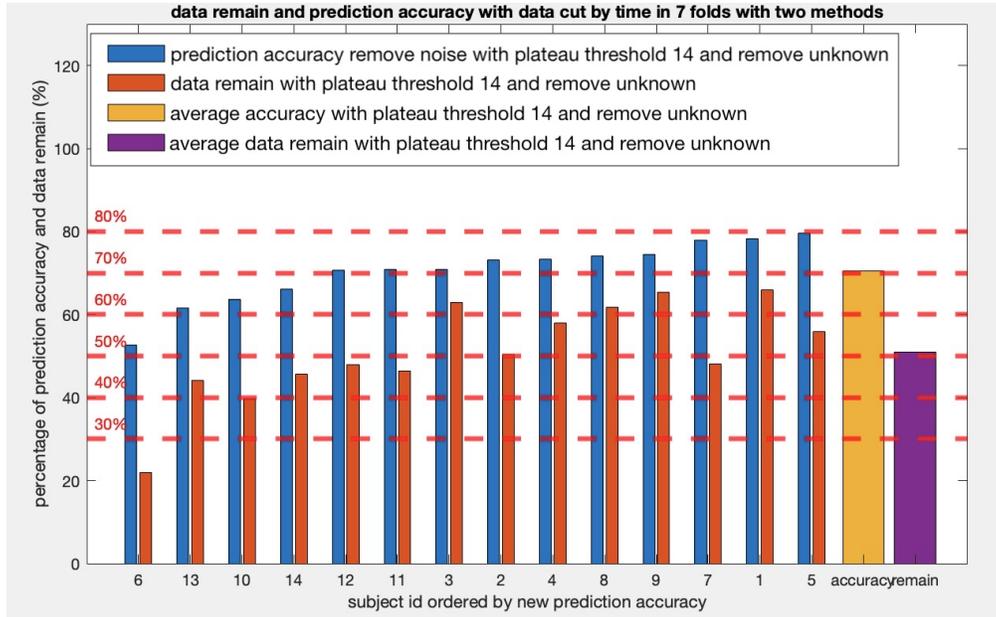


Figure 6: Experiment RWT: task prediction accuracy and data remain.

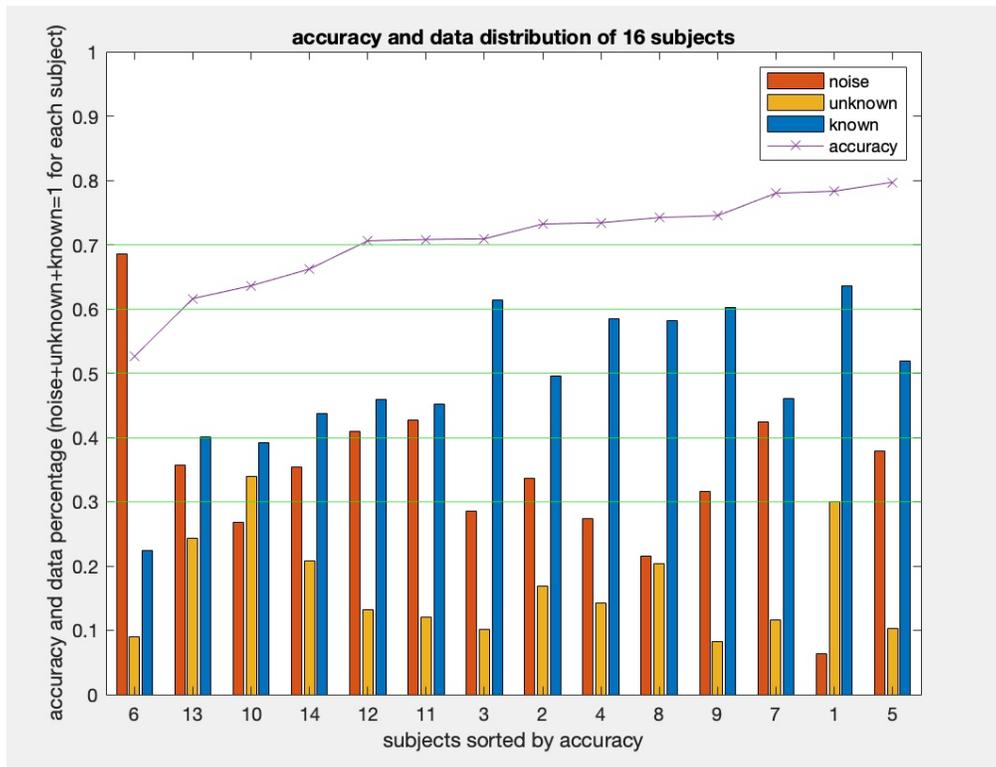


Figure 7: Experiment RWT: noise, unknown, and known tasks percentage.

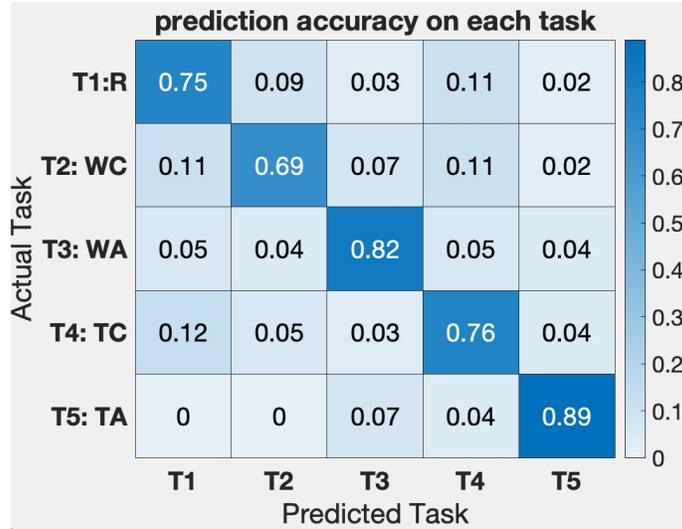


Figure 8: Diagonal Accuracy

181 subjects. Thus end-user training is necessary for better controlling the noise and unknown tasks. The  
 182 role of EEG coach is created for this purpose.

### 183 3.2 For End-Users

184 Figure 8 shows for subject one in experiment RWT, how the accuracy of each task is predicted over all  
 185 six sessions. This feedback may guide the further task selections. Each individual has a unique task  
 186 set that is easy to be recognized with this EEG-based BCI experimental design and data collection  
 187 framework. Thus it has the potential to be used as personal EEG fingerprint. Figure 9 shows the noise

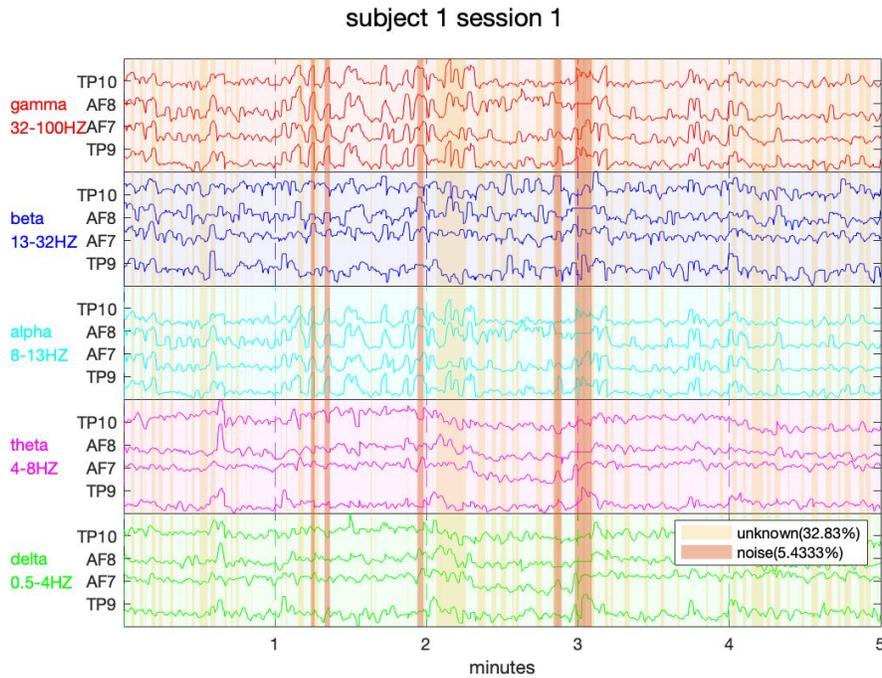


Figure 9: Noise, unknown, and known tasks, experiment RWT: subject one, session one.

188 and unknown task locations in each session, such feedback is helpful for the end-user to reflect what  
189 happened around a certain time spot.

## 190 **4 Discussion**

191 The main goal for this paper is to provide a framework of experimental design and data collection to  
192 gather EEG data through cheap means and non-expert participants. Intepretable feedback generate by  
193 benchmark machine learning algorithms can speed up this process. Comparing with the traditional  
194 data collection methods, as mentioned in Craik et al. [2019], Gabard-Durnam et al. [2018], Roy  
195 et al. [2019], Pernet et al. [2019], our approach is faster and cheaper to gather more EEG data with  
196 non-expert participants. Our efforts are made toward the future of everyone can use EEG-based BCI  
197 in their daily life, similar to the current everyday usage of smartphones. Although the limitation  
198 of sensory accuracy will remain for a while, the research related to the non-invasive BCI shows a  
199 growing potential to reach out to non-expert end-users.

200 The datasets we present are an early exploration of how to map the healthy subjects' daily activities to  
201 their personal EEG signal patterns. Based on the currently available sensory hardware, tasks without  
202 too much moving or talking could be a good start. A unique role of EEG coach could be helpful in  
203 such experiments to encourage more end-users to get involved in such experiments. The short-term  
204 goal of these datasets is to inspire new machine learning approaches for decoding behavior from  
205 EEG.

### 206 **4.1 Future Work**

207 Neural interfaces are becoming of increasing interest to industry and having large available datasets  
208 could be useful for students and researchers to tease out signals from noisy data. As non-invasive  
209 neural recordings become ubiquitous, there is a greater need for such algorithms and datasets. We  
210 will continue to focus on developing a framework to make it easier for non-expert end-users to use  
211 EEG-based BCI. Our short-term exploration includes developing more specific role sets for the BCI  
212 research and development framework, with the emphasis on the role of EEG coach and an online  
213 EEG experience community. The impact of continuous feedback to end-users is also a topic we  
214 are working on. Also, the idea of step out of the lab to home, starting with encouraging end-users  
215 to record EEG during tasks of her/his choice as many times as possible at home, is an interesting  
216 direction we are heading to.

### 217 **4.2 Broader Impact**

218 This approach could contribute to the building of a large-scale EEG dataset using low-cost tools and  
219 simple experimental settings at home. Our framework could be illuminated to a broader audience  
220 of other time-serious human sensory data collection. After all, brain signals are just one type of  
221 sensory health signals, the development of wearable devices are expanding rapidly to provide more  
222 perspective about human health information from both real-time monitoring and afterward data  
223 analysis.

224 Together with other human sensory data, EEG-based BCI has the potential to significantly change  
225 the ways of human interaction with the rest of the world, including both other individuals, and all  
226 the technology devices we developed. The human brain is a type of high-speed neural network, and  
227 the current AI-enhanced internet is also a high-speed network, how to connect the two high-speed  
228 networks could be an interesting long-term research direction. Our pilot study of quickly gathering  
229 large-scale EEG data could be a baby step moving towards this direction.

## 230 **5 Conclusion**

231 In this paper, we present a framework to gather large-scale EEG data through cheap means and  
232 non-expert participants, including experimental design, data collection, data analysis, and community  
233 building approaches. Two existing datasets are used as case studies for the framework: Think-Count-  
234 Recall (TCR) and Read-Write-Type (RWT). This could be a building block towards the future of  
235 everyone using non-invasive, wireless, and affordable BCI systems every day, similar to current  
236 smartphone usage for the general non-expert population.

237 **References**

- 238 J. J. Bird, L. J. Manso, E. P. Ribeiro, A. Ekart, and D. R. Faria. A study on mental state classification  
239 using eeg-based brain-machine interface. In *2018 International Conference on Intelligent Systems*  
240 *(IS)*, pages 795–800. IEEE, 2018.
- 241 J.-A. Chevalier, A. Gramfort, J. Salmon, and B. Thirion. Statistical control for spatio-temporal  
242 meg/eeg source imaging with desparsified multi-task lasso. *arXiv preprint arXiv:2009.14310*,  
243 2020.
- 244 D. Coyle, J. Principe, F. Lotte, and A. Nijholt. Guest editorial: Brain/neuronal-computer game  
245 interfaces and interaction. *IEEE Transactions on Computational Intelligence and AI in games*, 5  
246 (2):77–81, 2013.
- 247 A. Craik, Y. He, and J. L. Contreras-Vidal. Deep learning for electroencephalogram (eeg) classification  
248 tasks: a review. *Journal of neural engineering*, 16(3):031001, 2019.
- 249 D. Devlaminck, W. Waegeman, B. Bauwens, B. Wyns, P. Santens, and G. Otte. From circular  
250 ordinal regression to multilabel classification. In *Proceedings of the 2010 Workshop on Preference*  
251 *Learning (European Conference on Machine Learning, ECML)*, page 15, 2010.
- 252 L. J. Gabard-Durnam, A. S. Mendez Leal, C. L. Wilkinson, and A. R. Levin. The harvard automated  
253 processing pipeline for electroencephalography (happe): standardized processing software for  
254 developmental and high-artifact data. *Frontiers in neuroscience*, 12:97, 2018.
- 255 M. Ienca, P. Haselager, and E. J. Emanuel. Brain leaks and consumer neurotechnology. *Nature*  
256 *biotechnology*, 36(9):805–810, 2018.
- 257 M. J. Kahana, E. V. Aggarwal, and T. D. Phan. The variability puzzle in human memory. *Journal of*  
258 *Experimental Psychology: Learning, Memory, and Cognition*, 44(12):1857, 2018.
- 259 M. Kaya, M. K. Binli, E. Ozbay, H. Yanar, and Y. Mishchenko. A large electroencephalographic  
260 motor imagery dataset for electroencephalographic brain computer interfaces. *Scientific data*, 5(1):  
261 1–16, 2018.
- 262 A. Kübler, E. M. Holz, A. Riccio, C. Zickler, T. Kaufmann, S. C. Kleih, P. Staiger-Sälzer, L. Desideri,  
263 E.-J. Hoogerwerf, and D. Mattia. The user-centered design as novel perspective for evaluating the  
264 usability of bci-controlled applications. *PLoS One*, 9(12):e112392, 2014.
- 265 J. LaRocco, M. D. Le, and D.-G. Paeng. A systemic review of available low-cost eeg headsets used  
266 for drowsiness detection. *Frontiers in neuroinformatics*, 14, 2020.
- 267 F. Lotte. Signal processing approaches to minimize or suppress calibration time in oscillatory  
268 activity-based brain–computer interfaces. *Proceedings of the IEEE*, 103(6):871–890, 2015.
- 269 F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi. A review of classification algorithms  
270 for EEG-based brain–computer interfaces. *Journal of neural engineering*, 4(2):R1, 2007.
- 271 F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger. A  
272 review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update.  
273 *Journal of neural engineering*, 15(3):031005, 2018a.
- 274 F. Lotte, C. Jeunet, J. Mladenović, B. N’Kaoua, and L. Pillette. A bci challenge for the signal  
275 processing community: considering the user in the loop, 2018b.
- 276 K. J. Miller. A library of human electrocorticographic data and analyses. *Nature human behaviour*, 3  
277 (11):1225–1235, 2019.
- 278 T. E. Nichols, S. Das, S. B. Eickhoff, A. C. Evans, T. Glatard, M. Hanke, N. Kriegeskorte, M. P.  
279 Milham, R. A. Poldrack, J.-B. Poline, et al. Best practices in data analysis and sharing in  
280 neuroimaging using mri. *Nature neuroscience*, 20(3):299–303, 2017.
- 281 C. R. Pernet, S. Appelhoff, K. J. Gorgolewski, G. Flandin, C. Phillips, A. Delorme, and R. Oostenveld.  
282 Eeg-bids, an extension to the brain imaging data structure for electroencephalography. *Scientific*  
283 *data*, 6(1):1–5, 2019.

- 284 X. Qu, M. Hall, Y. Sun, R. Sekuler, and T. J. Hickey. A personalized reading coach using wearable  
285 EEG sensors-a pilot study of brainwave learning analytics. In *CSEDU (2)*, pages 501–507, 2018a.
- 286 X. Qu, Y. Sun, R. Sekuler, and T. Hickey. EEG markers of stem learning. In *2018 IEEE Frontiers in*  
287 *Education Conference (FIE)*, pages 1–9. IEEE, 2018b.
- 288 X. Qu, P. Liu, Z. Li, and T. Hickey. Multi-class time continuity voting for eeg classification. In  
289 *International Conference on Brain Function Assessment in Learning*, pages 24–33. Springer,  
290 2020a.
- 291 X. Qu, Q. Mei, P. Liu, and T. Hickey. Using eeg to distinguish between writing and typing for the  
292 same cognitive task. In *International Conference on Brain Function Assessment in Learning*, pages  
293 66–74. Springer, 2020b.
- 294 A. Roc, L. Pillette, J. Mladenovic, C. Benaroch, B. N’Kaoua, C. Jeunet, and F. Lotte. A review  
295 of user training methods in brain computer interfaces based on mental tasks. *Journal of Neural*  
296 *Engineering*, 2020.
- 297 Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert. Deep learning-based  
298 electroencephalography analysis: a systematic review. *Journal of neural engineering*, 16(5):  
299 051001, 2019.
- 300 D. Sabbagh, P. Ablin, G. Varoquaux, A. Gramfort, and D. A. Engemann. Manifold-regression to  
301 predict from meg/eeg brain signals without source modeling. *arXiv preprint arXiv:1906.02687*,  
302 2019.
- 303 G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw. Bci2000: a general-  
304 purpose brain-computer interface (bci) system. *IEEE Transactions on biomedical engineering*, 51  
305 (6):1034–1043, 2004.
- 306 K. Sierra and B. Bates. *Head first java*. " O’Reilly Media, Inc.", 2003.
- 307 Y. R. Tabar and U. Halici. A novel deep learning approach for classification of eeg motor imagery  
308 signals. *Journal of neural engineering*, 14(1):016003, 2016.
- 309 T. Tu, J. Paisley, S. Haufe, and P. Sajda. A state-space model for inferring effective connectivity of  
310 latent neural dynamics from simultaneous eeg/fmri. *Advances in Neural Information Processing*  
311 *Systems*, 32:4662–4671, 2019.
- 312 X. Zhang, L. Yao, X. Wang, J. J. Monaghan, D. Mcalpine, and Y. Zhang. A survey on deep  
313 learning-based non-invasive brain signals: recent advances and new frontiers. *Journal of Neural*  
314 *Engineering*, 2020.