A preliminary qEEG study on Augmented Reality Guidance of Manual Tasks

1th Gianluca Rho
Dip. di ingegneria dell'informazione
University of Pisa
Pisa, Italy
gianluca.rho@phd.unipi.it

2th Alejandro Luis Callara
Research Center "E. Piaggio"
University of Pisa
Pisa, Italy
alejandro.callara@ing.unipi.it

3nd Sara Condino
Dip. di ingegneria dell'informazione,
EndoCAS Center, University of Pisa
Pisa, Italy
sara.condino@endocas.unipi.it

4rd Shadi Ghiasi
Research Center "E. Piaggio"
University of Pisa
Pisa, Italy
shadi.ghiasi@centropiaggio.unipi.it

5st Mimma Nardelli
Research Center "E. Piaggio"
University of Pisa
Pisa, Italy
m.nardelli@ing.unipi.it

6th Marina Carbone
EndoCAS Center
University of Pisa
Pisa, Italy
vincenzo.ferrari@unipi.it

7th Vincenzo Ferrari
Dip. di ingegneria dell'informazione,
EndoCAS Center, University of Pisa
Pisa, Italy
vincenzo.ferrari@unipi.it

8th Alberto Greco
Dip. di ingegneria dell'informazione
University of Pisa
Pisa, Italy
orcid.org/0000-0002-4822-5562

9th Enzo Pasquale Scilingo
Dip. di ingegneria dell'informazione
University of Pisa
Pisa, Italy
enzo.scilingo@unipi.it

Abstract—Augmented reality head mounted display devices (HMDs) provide user's natural view of the real world with enhanced experience through optical superposition of virtual data. Manual-task guidance applications exploiting these systems are particularly suited in computer-aided-surgery. However, the typical working distance of commercial devices is higher than user reachable space, limiting the purpose of manual-task guidance. Specifically, known issues such as the "vergence-accomodation-conflict" and the "focus-rivalry" may lead to visual fatigue and mental workload worsening task performance. Here, we exploit EEG recordings during a "connecting-the-dots" task performed with and without AR to evaluate the mental workload associated with AR-related visual fatigue. First, we quantify the reduction of users' performance based on starting and end points gap errors. Then, we investigate the effects on AR usage on cortical activity through the analysis of EEG power and Frontal Alpha Asymmetry (FAA) index. Although preliminary, our results suggest that mental workload associated with AR usage may derive from enhanced difficulty associated with the task. Furthermore, a shift in FAA from controlateral to ipsilateral regions seems to confirm this hypothesis.

Index Terms—Augmented-Reality (AR), quantitative-EEG, visual fatigue, mental workload, power analysis, frontal alpha asymmetry

I. INTRODUCTION

In the recent years, augmented reality (AR) head-mounted display devices (HMDs) have gained popularity over a large number of fields, ranging from consumer-level products for gaming applications to BCI systems and computer-aided surgery [1], [2], [3], [4]. One great advantage of these systems is to augment operator's natural view of the real world, allowing for manual tasks guidance and real-time estimation of hand location during task execution. In this context, several limitations may arise, ranging from technological (i.e. device obtrusiveness, low luminance, small field-of-view) to perceptual (i.e. the conflict between 2D virtual content on the surface of projection and 3D real world). Perceptual conflict is certainly one of the most critical factors of HMDs, leading to two common phenomena: the vergence-accomodation conflict (VAC), that arises when maintaining a single binocular vision [5], and the focus rivalry (FR), that occurs when focusing real and virtual content at the same time [6].

VAC and FR effects are particularly evident when the real content is located in the user's peripersonal space (i.e. the space containing reachable objects), limiting their use in precision manual tasks such as AR guided-surgery [1]. Specifically, VAC and FR give rise to visual fatigue [7], [8], which turns out in an increased visual discomfort and mental workload on user performance during the execution of AR guided tasks. In particular, user performance was observed to be better during naked eye tests with respect to AR-guided tasks [9].

Visual fatigue is usually assessed by exploiting questionnaires gathering users' evaluations before and after the experiment [10], [11]. Nevertheless, questionnaires are naturally biased by inter-subject variability in the answers, due to the subject's mood at the time of the experiment as well as to their past experience or to their level of understanding of the questions [12]. On the other hand, alternative approaches may exploit electrophysiological features for assessing the visual fatigue, such as those provided by quantitative-electroencephalography (qEEG) [13], [14]. Previous studies involving power spectrum indexes [12] as well as Event-Related Potentials (ERPs) latency and amplitude successfully identified visual fatigue and discomfort in VR environments during passive tasks involving visual stimuli [15]. A step further may be that of evaluating the mental workload associated with visual fatigue during manual task execution with respect to simple passive visual tasks.

Therefore, in this preliminary study we estimate brain activity correlates of mental workload during a manual task with and without AR. Power spectrum and Frontal Alpha Asymmetry (FAA) [16] were analyzed in order to highlight statistically significant differences in brain activity related to the AR-environment. In particular, we estimate these two electrophysiological measures to discriminate among the same manual task executed with or without the Microsoft HoloLens device. Results are integrated with performance error outcomes of the manual task in both conditions.

II. MATERIAL AND METHODS

A. Subjects

Twelve healthy volunteers (age 25 ± 2.85, 5 males, all right-handed) underwent an augmented reality guided "connecting-the-dots" precision task using Microsoft HoloLens. Subjects were selected based on their visual acuity, assessed through the Digital Acuity LogMAR Charts from Chart2020. Specifically, only those subjects with normal or corrected-to-normal visual acuity participated in the study. All subjects gave their informed consent to take part to the study, self-reporting no history of clinical cardiovascular and/or mental diseases.

B. Experimental protocol

During the experiment, each subject sat in front of a desk, with their chin placed over a "chin-rest" in order to fix the target/user distance at $0.5m$ (Fig. 1). The protocol consisted of two modalities:

- AR guided (AR);
- Naked eye (NK).

and for each modality, the "connecting-the-dots" task was performed 3 times. Specifically, the experiment consisted of 120s of rest followed by 3 repetitions of each task, with an inter-task interval of 30s as illustrated in Fig. 2. Finally, modality was randomized across subjects.

For each task, subjects connected a sequence of 15 numbered dots with a straight line (Fig. 3). Dots were randomly displayed on an A4 paper put on a vertical physical support. During the AR modality, numbered dots were displayed in front of a blank paper: this forces the subject's eyes to focus both on the virtual (dots) and the real objects (paper, pen). For the NK modality, HoloLens were removed and the numbered dots were printed on the paper.

At the end of the experiment subjects filled in the Likert questionnaire on how was the experience with AR during the experiment (e.g., level of discomfort, sharpness of the dots, level of fatigue).

C. Augmented Reality device

The selected HMD is the 1st Generation Microsoft HoloLens, an Optical See-Through HMD (OST-HMD) with self-contained computing power, based on an undisclosed Intel 32-bit processor, with a custom-built Microsoft Holographic Processing Unit (HPU 1.0) which supports Universal Windows Platform (UWP) apps. HoloLens features 2 GB of RAM and 64 GB of flash memory, and network connectivity via Wi-Fi 802.11ac and Bluetooth 4.1 LE wireless technology. The sensory system of the device includes: one depth camera, four grayscale tracking cameras, and one world-facing photo/video camera (2 MP), one ambient light sensor, one inertial measurement unit (IMU) to track head movements, and four microphones.

D. Manual task performance analysis

Each "connecting-the-dots" sequence was analyzed to measure subject's performance in connecting dots. Given that eye-to-display calibration errors may lead to a distortion of the perceived virtual content, and therefore they can play a major role in the misperception of line lengths. The performance was evaluated in term of gaps (G_{ij}) between the end (End_i) and the starting points ($Start_j$) of each pair (i,j) of consecutive lines. This measure indeed cannot be related to calibration inaccuracies. The line endpoints were automatically detected with the Harris Corner Detector, as described in [9], processing the image drawn by the subject with MATLAB
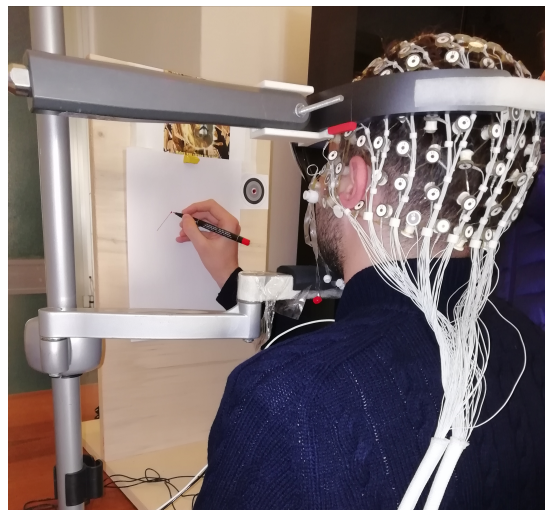

Fig. 1. Experimental setup. The subjects sat in front of a desk and placed their chin over a "chin-rest" in order to fix the target/user distance at $0.5m$.

Version R2017b. For each trial the following parameters were calculated: maximum gap (GMAX), mean gap (GMEAN), and total gap (GTOT), i.e. the sum of all the gaps (G_{ij}) measured in the trial.

E. EEG data acquisition and preprocessing

EEG signals were recorded using 128-channel Geodesic EEG System 300 from Electrical Geodesics, Inc. (EGI). Channels were referenced to Cz and acquired with a sample frequency of 500Hz.

EEG signals were preprocessed with EEGLAB [17]. First, the data was band-pass filtered in the range [1-40]Hz. Then, 50Hz line noise and the 100Hz and 150Hz harmonics were removed by applying three notch filters with a stop-band of $4Hz$. Bad channels were estimated by evaluating the correlation coefficient between channels. Specifically, we removed those channels whose correlation with their neighbours was less than the 80% as they are likely to be associated with artifacts [18]. Then, EEG data were visually inspected and bad data periods (e.g. those affected by movement artifacts, EMG artifacts, and others non-stereotyped phenomena) were removed by hand, together with bad channels not identified in the previous step. The deleted channels were recovered by interpolation prior to referencing them to signal average. Finally, EEG signals were decomposed into sets of maximally independent components by Independent Component Analysis (ICA) [19]. The obtained components corresponded to statistically independent time-courses associated with static ICA maps. Components are representative of both brain activity and different types of artifacts. Here, we exploited ICA decomposition for removing ICs representing eye blinks artifacts, saccadic movements and ECG artifacts.

F. EEG power spectral density estimation

For each electrode, the power spectral density of preprocessed EEG signals was estimated using the Welch method. Moving windows were $5s$-long and their overlap was the 75%. Electrodes were further grouped into eight regions of interest according to their location over the scalp (i.e. Frontal Left (FL), Frontal Right (FR), Temporal Left (TL), Temporal Right (TR), Parietal Left (PL), Parietal Right(PR), Occipital Left (OL) and Occipital Right (OR)). For each ROI and for each modality, the average power was estimated. Average power estimates were then analyzed for 3 frequency bands of interest: $\theta[4-8]Hz$, $\alpha[8-13]Hz$ and $\beta[13-30]Hz$. Furthermore, we evaluated left and right hemispheres' power for each modality.
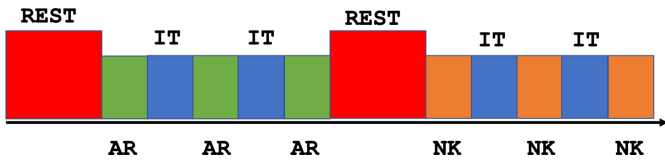

Fig. 2. Experimental timeline. Each modality consisted of $120sec$ of rest, followed by 3 repetitions of the task. The inter-task interval was of $30sec$.

Then, we evaluated the Frontal Alpha Asymmetry (FAA) for each modality, as this can be related to approach-withdrawal behaviour related to the use of AR. FAA was evaluated as the difference between the log mean power of the frontal right and left ROIs (1):

$$FAA = 10\,Log_{10}(P_{\text{FR}}) - 10\,Log_{10}(P_{\text{FL}}) \qquad (1)$$

G. EEG statistical analysis

Statistically significant differences in between AR and NK modalities were assessed with a non-parametric Wilcoxon sign-rank test. This procedure was applied to performance outcomes and power estimates. For the latter, differences were evaluated within the θ, α and β bands and for each one of the 8 ROIs on the scalp. Multiple comparison was controlled with False-Discovery-Rate (FDR) Benjamini-Yekutieli correction for multiple testing under dependency [20]. We further estimated differences in power, for each frequency band, between the left and right hemisphere by grouping ROIs based on their position on the scalp (i.e. FL, TL, PL and OL vs. FR, TR, PR and OR) with a Friedman's test with replicates. Finally, FAA significant differences were evaluated by comparing their medians with the Wilcoxon sign-rank test.

III. RESULTS

A. Performance Results

All the subjects did not experience any perceivable jitter/drift of the virtual content, and successfully completed the six tasks. Mean and standard deviation values of GMAX, GMEAN and GTOT are reported in Table I, showing that, on average, subjects performed better during the Naked-eye sessions. The Wilcoxon signed-rank test showed significant differences in subject performances depending on the test modalities ($p < 0.05$).

B. EEG power analysis

The Wilcoxon signed-rank test did not show any significant change in power between AR and NK modalities for none of the analyzed ROIs and for none of the frequency bands considered. On the other hand, considering the power in
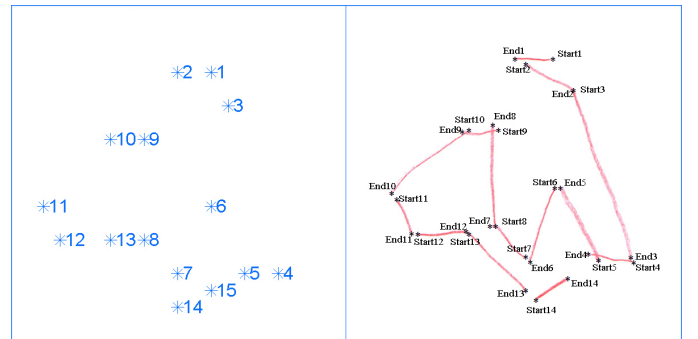

Fig. 3. Example of the task performed. On the left, the virtual content visualized by the subject; on the right, lines drawn by the subject processed to evaluate the performance: the endpoints ($Start_i$, End_i) of each line are represented with black stars.

the left and right hemispheres from grouped ROIs analysis, significant differences were observed between modalities ($p < 0.05$, $FDR - corrected$). Specifically, we found higher alpha power during AR with respect to NK in both hemispheres (Fig.4).

Finally, significant differences were observed in the FAA index between AR and NK modalities. In particular, while limited lateralization occurred during AR, unbalanced power was observed during the NK condition. Specifically, the power in FL was higher with respect to the power in the FR (Fig.5).
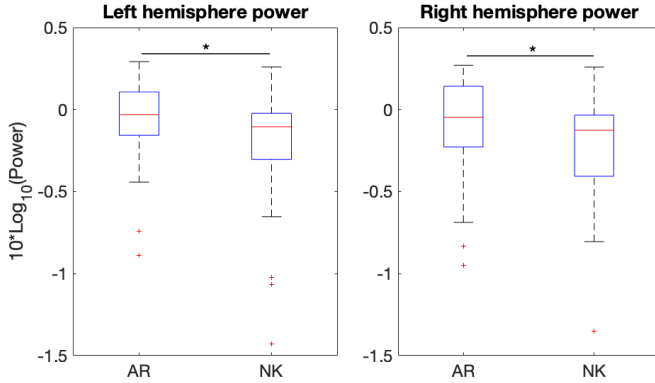

Fig. 4. Box-plot of α-power differences between AR and NK modality. Left: comparison between the distribution of the α-power in the left hemisphere between AR and NK conditions. Right: comparison of the α-power in the right hemisphere between AR and NK conditions. The median value of α-power is greater in the AR modality than in NK in both hemispheres. Outliers are marked with a '+' sign (* = $p < 0.001$).

IV. DISCUSSION AND CONCLUSION

In this preliminary study, we investigated the brain electrical activity provided by EEG measurements during an AR-guided manual task and evaluated relevant differences with respect to the same task executed without AR. Specifically, we aimed at highlighting potential differences in EEG features associated with mental workload deriving from visual fatigue.

Twelve healthy volunteers used the Microsoft HoloLens device at a lower distance with respect to the minimum focal distance, enhancing VAC and FR effects. Our results are in line with previous studies that quantified the worse performance of manual tasks using AR [9]. This may derive from an increased mental workload associated with AR use. Here, we showed that such workload can be quantified through the use of electrophysiological measures.

We observed that EEG power does not change between AR and NK modalities for eight brain ROIs and for the three

TABLE I
PERFORMANCE EVALUATION RESULT.

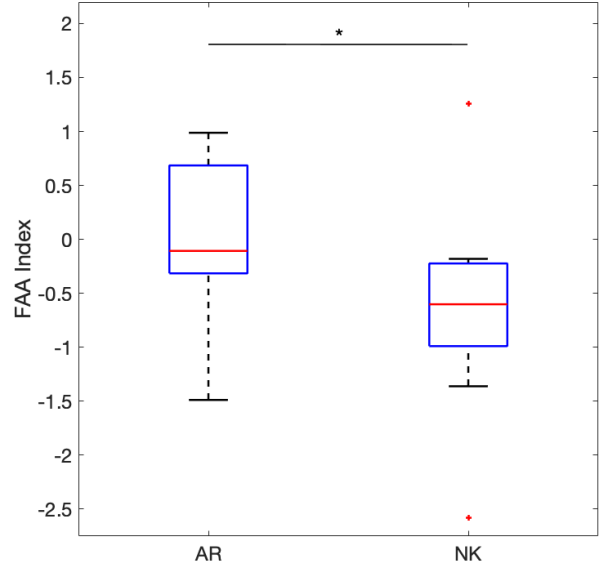| 2*Feature	Naked-eye		AR		2*p-values
	Mean	STD	Mean	STD	
GMAX [mm]	0.2	0.1	0.3	0.1	0.013
GMEAN [mm]	0.1	0.5	0.1	0.1	0.011
GTOT [mm]	0.8	0.5	1.8	0.8	0.004

Fig. 5. Distributions of the FAA index computed in the AR condition (left) and in the NK condition (right). The power in the NK condition is shifted toward the left hemisphere, while in the AR condition the power is almost symmetrical. Outliers are marked with a '+' sign (* = $p < 0.05$).

frequency bands considered. Nevertheless, we observed statistically significant differences between modalities in the α-power when the left and right hemispheres were considered as a group. In this view, we can hypothesize that local differences may not be as high as global ones. Indeed, previous studies involving tasks of increasing difficulty showed that alpha power increases as the difficulty of the task does, and that such increase is independent of hemisphere considered [21]. Although in such work the experiments were not specifically manual tasks, the impact of task-difficulty may reflect in a higher mental workload. This seems to be confirmed also by the performance analysis, since all subjects committed higher errors during the AR-guided task with respect to NK condition.

Significant differences in FAA were observed when performing the task in NK condition with respect to using AR. This index is known to characterize an approach-withdrawal behaviour [22]. Thus, it seems reasonable to assume that such a behaviour may distinguish the two modalities. Specifically, left-lateralization was observed to promote approach whereas increased right activity was associated with withdrawal behaviour [23]. In this view, our results suggest that during NK an approach-like behaviour is present. On the other hand, no evident lateralization in frontal EEG power was observed for the AR. However, another consideration must be done. Indeed, increased neural ipsilateral recruitment has been observed as tasks become more complex [24], [25]. In this view, considering the left-lateralization observed during NK condition, and that all subjects that participated in the study were right-handed, we can suppose that the shift in FAA

towards the right hemisphere may be due to an increase in task complexity, that could also reflect a higher mental workflow.

Although preliminary, our results suggest that mental workload associated with visual fatigue during AR-guided manual task can be measured with EEG derived electrophysiological features. It is our opinion that more outstanding results will be possible in the future by enlarging the number of subjects participating to the study as well as including other relevant physiological signals, as for instance those related to autonomic nervous function. Indeed, a not-negligible limitation is related to the limited number of subjects involved in the study. Another fundamental distinction may consider user's familiarity with AR systems that can potentially bias results, as experienced users may reasonably feel less discomfort in AR-environments. Nevertheless, this study offers interesting perspectives for further characterizing the neural correlates associated with AR environments.

REFERENCES

[1] W. Birkfellner, K. Huber, F. Watzinger, M. Figl, F. Wanschitz, R. Hanel, D. Rafolt, R. Ewers, and H. Bergmann, "Development of the Varioscope AR. A see-through HMD for computer-aided surgery," in *Proceedings IEEE and ACM International Symposium on Augmented Reality (ISAR 2000)*, pp. 54–59, Oct. 2000. ISSN: null.

[2] A. Bornik, R. Beichel, B. Reitinger, G. Gotschuli, E. Sorantin, F. W. Leberl, and M. Sonka, "Computer-aided liver surgery planning: an augmented reality approach," in *Medical Imaging 2003: Visualization, Image-Guided Procedures, and Display*, vol. 5029, pp. 395–406, International Society for Optics and Photonics, May 2003.

[3] I. Hupont, J. Gracia, L. Sanagustín, and M. A. Gracia, "How do new visual immersive systems influence gaming QoE? A use case of serious gaming with Oculus Rift," in *2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX)*, pp. 1–6, May 2015. ISSN: null.

[4] M. Wang, R. Li, R. Zhang, G. Li, and D. Zhang, "A Wearable SSVEP-Based BCI System for Quadcopter Control Using Head-Mounted Device," *IEEE Access*, vol. 6, pp. 26789–26798, 2018.

[5] G. Kramida, "Resolving the Vergence-Accommodation Conflict in Head-Mounted Displays," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, pp. 1912–1931, July 2016.

[6] K. Oshima, K. R. Moser, D. C. Rompapas, J. E. Swan, S. Ikeda, G. Yamamoto, T. Taketomi, C. Sandor, and H. Kato, "SharpView: Improved clarity of defocussed content on optical see-through head-mounted displays," in *2016 IEEE Virtual Reality (VR)*, pp. 253–254, Mar. 2016. ISSN: 2375-5334.

[7] E. Peli, "The visual effects of head-mounted display (HMD) are not distinguishable from those of desk-top computer display," *Vision Research*, 1998.

[8] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence-accommodation conflicts hinder visual performance and cause visual fatigue," *Journal of Vision*, vol. 8, pp. 33.1–30, Mar. 2008.

[9] S. Condino, M. Carbone, R. Piazza, M. Ferrari, and V. Ferrari, "Perceptual Limits of Optical See-Through Visors for Augmented Reality Guidance of Manual Tasks," *IEEE transactions on bio-medical engineering*, vol. 67, no. 2, pp. 411–419, 2020.

[10] J. Berry, D. Budgen, and N. Holliman, "Evaluating Subjective Impressions of Quality Controlled 3D Films on Large and Small Screens," *Journal of Display Technology*, vol. 11, pp. 927–938, Nov. 2015.

[11] T. Shibata, J. Kim, D. M. Hoffman, and M. S. Banks, "The zone of comfort: Predicting visual discomfort with stereo displays," *Journal of Vision*, vol. 11, p. 11, July 2011.

[12] B. Zou, Y. Liu, M. Guo, and Y. Wang, "EEG-Based Assessment of Stereoscopic 3D Visual Fatigue Caused by Vergence-Accommodation Conflict," *Journal of Display Technology*, vol. 11, pp. 1076–1083, Dec. 2015.

[13] C. Mai, M. Hassib, and R. Königbauer, "Estimating Visual Discomfort in Head-Mounted Displays Using Electroencephalography," in *Human-Computer Interaction – INTERACT 2017* (R. Bernhaupt, G. Dalvi, A. Joshi, D. K. Balkrishan, J. O'Neill, and M. Winckler, eds.), Lecture Notes in Computer Science, (Cham), pp. 243–252, Springer International Publishing, 2017.

[14] H.-C. O. Li, J. Seo, K. Kham, and S. Lee, "Measurement of 3D Visual Fatigue Using Event-Related Potential (ERP): 3D Oddball Paradigm," in *2008 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video*, pp. 213–216, May 2008. ISSN: 2161-203X.

[15] J. Frey, L. Pommereau, F. Lotte, and M. Hachet, "Assessing the Zone of Comfort in Stereoscopic Displays using EEG," *Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems - CHI EA '14*, pp. 2041–2046, 2014. arXiv: 1404.6222.

[16] E. E. Smith, S. J. Reznik, J. L. Stewart, and J. J. B. Allen, "Assessing and conceptualizing frontal EEG asymmetry: An updated primer on recording, processing, analyzing, and interpreting frontal alpha asymmetry," *International Journal of Psychophysiology*, vol. 111, pp. 98–114, Jan. 2017.

[17] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, pp. 9–21, Mar. 2004.

[18] T. R. Mullen, C. A. E. Kothe, Y. M. Chi, A. Ojeda, T. Kerth, S. Makeig, T.-P. Jung, and G. Cauwenberghs, "Real-Time Neuroimaging and Cognitive Monitoring Using Wearable Dry EEG," *IEEE transactions on bio-medical engineering*, vol. 62, pp. 2553–2567, Nov. 2015.

[19] J. A. Palmer, K. Kreutz-Delgado, and S. Makeig, "AMICA : An Adaptive Mixture of Independent Component Analyzers with Shared Components," 2011.

[20] Y. Benjamini and D. Yekutieli, "The control of the false discovery rate in multiple testing under dependency," *The Annals of Statistics*, vol. 29, pp. 1165–1188, Aug. 2001.

[21] D. Galin, J. Johnstone, and J. Herron, "Effects of task difficulty on EEG measures of cerebral engagement," *Neuropsychologia*, vol. 16, pp. 461–472, Jan. 1978.

[22] R. J. Davidson, P. Ekman, C. D. Saron, J. A. Senulis, and W. V. Friesen, "Approach-withdrawal and cerebral asymmetry: Emotional expression and brain physiology: I.," *Journal of Personality and Social Psychology*, vol. 58, no. 2, pp. 330–341, 1990.

[23] B. B. Briesemeister, S. Tamm, A. Heine, and A. M. Jacobs, "Approach the Good, Withdraw from the Bad—A Review on Frontal Alpha Asymmetry Measures in Applied Psychological Research," *Psychology*, vol. 04, p. 261, Mar. 2013.

[24] T. Verstynen, J. Diedrichsen, N. Albert, P. Aparicio, and R. B. Ivry, "Ipsilateral motor cortex activity during unimanual hand movements relates to task complexity," *Journal of Neurophysiology*, vol. 93, pp. 1209–1222, Mar. 2005.

[25] A. Van Impe, J. P. Coxon, D. J. Goble, N. Wenderoth, and S. P. Swinnen, "Ipsilateral coordination at preferred rate: Effects of age, body side and task complexity," *NeuroImage*, vol. 47, pp. 1854–1862, Oct. 2009.