

RESEARCH ARTICLE

Is Neural Activity Detected by ERP-Based Brain-Computer Interfaces Task Specific?

Markus A. Wenzel^{1*}, Inês Almeida², Benjamin Blankertz¹

1 Neurotechnology Group, Technische Universität Berlin, Berlin, Germany, **2** Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal

* markus.wenzel@tu-berlin.de



OPEN ACCESS

Citation: Wenzel MA, Almeida I, Blankertz B (2016) Is Neural Activity Detected by ERP-Based Brain-Computer Interfaces Task Specific? PLoS ONE 11 (10): e0165556. doi:10.1371/journal.pone.0165556

Editor: Daniele Marinazzo, Ghent University, BELGIUM

Received: March 29, 2016

Accepted: October 13, 2016

Published: October 28, 2016

Copyright: © 2016 Wenzel et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: EEG and behavioral data are available from the repository of the Technische Universität Berlin (<http://dx.doi.org/10.14279/depositonce-5523>).

Funding: The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013; <http://cordis.europa.eu/fp7/>) under grant agreement n° 611570. The work of BB was additionally funded by the Bundesministerium für Bildung und Forschung (<https://www.bmbf.de/>) under contract 01GQ0850. The funders had no role in study design, data

Abstract

Objective

Brain-computer interfaces (BCIs) that are based on event-related potentials (ERPs) can estimate to which stimulus a user pays particular attention. In typical BCIs, the user silently counts the selected stimulus (which is repeatedly presented among other stimuli) in order to focus the attention. The stimulus of interest is then inferred from the electroencephalogram (EEG). Detecting attention allocation implicitly could be also beneficial for human-computer interaction (HCI), because it would allow software to adapt to the user's interest. However, a counting task would be inappropriate for the envisaged implicit application in HCI. Therefore, the question was addressed if the detectable neural activity is specific for silent counting, or if it can be evoked also by other tasks that direct the attention to certain stimuli.

Approach

Thirteen people performed a silent counting, an arithmetic and a memory task. The tasks required the subjects to pay particular attention to target stimuli of a random color. The stimulus presentation was the same in all three tasks, which allowed a direct comparison of the experimental conditions.

Results

Classifiers that were trained to detect the targets in one task, according to patterns present in the EEG signal, could detect targets in all other tasks (irrespective of some task-related differences in the EEG).

Significance

The neural activity detected by the classifiers is not strictly task specific but can be generalized over tasks and is presumably a result of the attention allocation or of the augmented workload. The results may hold promise for the transfer of classification algorithms from BCI research to implicit relevance detection in HCI.

collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Introduction

If a person pays special attention to a stimulus, a particular neural response is evoked that can be detected as event-related potential (ERP) in the electroencephalogram (EEG). This phenomenon is used in brain-computer interfacing (BCI) in order to establish a communication and control channel, which is purely based on neural activity and does not involve any muscle movements. Typical ERP-based BCI systems repeatedly present different stimuli one by one to a person who selects one stimulus of interest (*target*) and silently counts its appearance and ignores other stimuli (*distractors*) [1–11]. Counting helps to direct the attention to one of several stimulus types. Targets evoke a detectable neural response in comparison to distractors (an augmented late positive going centroparietal EEG component referred to as P300) [12–14].

Recently, it was suggested that BCI technology could be transferred to relevance detection in human-computer interaction (HCI), because EEG combined with an eye tracker can be used to predict which of several items displayed at the same time on the screen are task-relevant for the user [15–20]. The application of BCI technology to HCI is presumably most useful and convenient for the user if the information (e.g. about the relevance of words or pictograms) can be inferred implicitly from the neural activity when the user pursues different activities. Accordingly, the question was addressed if the detectable, target-related neural activity is specific for the silent counting task, or if it is also present in other tasks that direct the attention to target stimuli. If the activity is not task specific, it is presumably a result of the attention allocation or of the augmented workload. Generalizability of the neural patterns would be promising for the envisaged application case, where the user performs different tasks while focusing on, and expending more mental effort on relevant screen content. While silent counting is legitimate to enhance performance in most BCI applications, relevance detection is not feasible if silent counting is essential to elicit a neural response that can be detected in single (or few) trials of EEG.

Materials and Methods

Experimental Design

Thirteen people performed a silent counting, an arithmetic and a memory task. The tasks required the subject to pay particular attention to target stimuli of a color that was randomly changed after each task repetition. The stimulus presentation was the same in all three tasks, which allowed a direct comparison of the experimental conditions. Squares in the colors magenta, yellow, red, blue and green flashed one by one for 500 ms each, interleaved by 500 ms blank screen, in a five-times-five grid in pseudo-random order and arrangement (cf. Fig 1). The probability of the appearance of each color was the same, such that the ratio of the random target color to the other colors was approximately one to four, resulting in eight to thirteen targets among the 47 to 50 colored squares in total per stimulus sequence.

Condition C constitutes the original version where stimuli of the target color had to be counted while stimuli of other colors—the distractors—could be ignored. In the arithmetic task of condition A, ‘ten’ had to be added for targets and ‘one’ for the more frequent distractors. In condition M, the position of the targets on the screen had to be memorized. The target color magenta appeared twice among four distractors in the short exemplary stimulus sequence given in Fig 1. The correct result would be $1 + 1 = 2$ for condition C, $1 + 1 + 10 + 1 + 1 + 10 = 24$ for condition A and ‘row 1, column 3’ and ‘row 3, column 5’ for condition M.

The task and the random target color were introduced before each stimulus sequence. After the presentation of the stimulus sequence, the result had to be entered with keyboard (numbers) or mouse (coordinates) and, finally, the correct answer was shown on the screen. The three tasks took turns and were repeated twenty times each (cf. Fig 1).

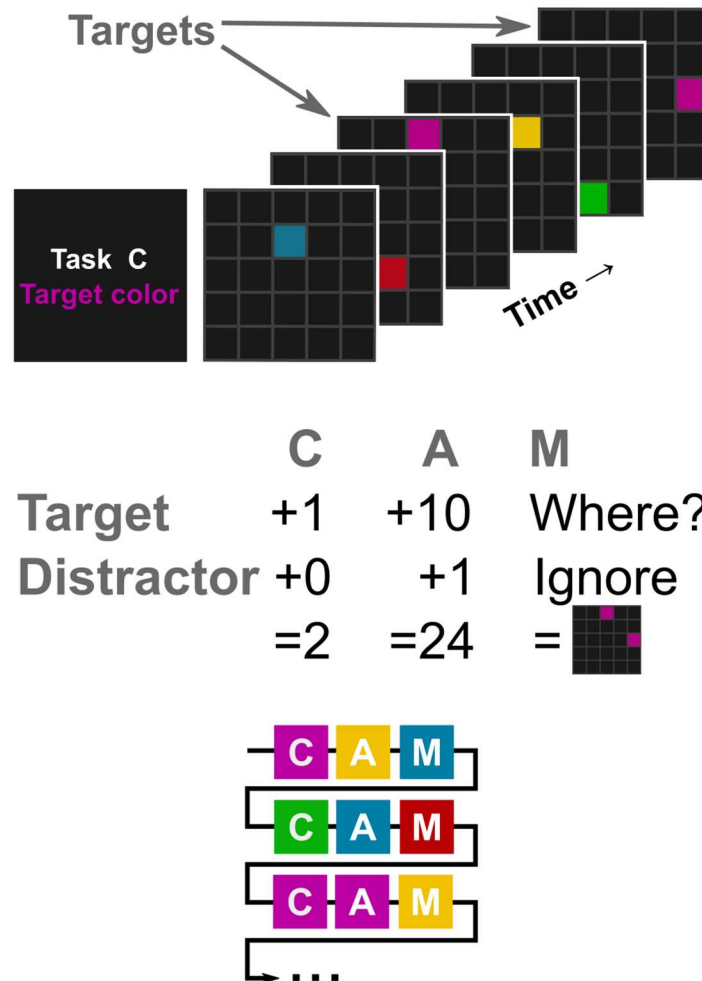


Fig 1. Short exemplary stimulus sequence (top), experimental tasks C, A and M (center) and sequence of the tasks and random target colors (bottom). The participants looked at a random stimulus sequence, where 47 to 50 squares of five colors flashed (with equal probabilities) in a grid for 500 ms each, interleaved by 500 ms blank screen. Before each stimulus sequence, the task and a random target color were assigned. The respective target color required a particular mental operation, depending on the task. Every participant performed task C (counting targets), A (arithmetic for targets and distractors), and M (memorizing target positions) twenty times each. The result had to be entered after the stimulus sequence.

doi:10.1371/journal.pone.0165556.g001

Stimuli of the target color did not stand out systematically, e.g., with respect to salience or frequency. Targets distinguished themselves only due to the preceding definition as target for the present task repetition, because each color appeared with equal probability and the target color was frequently changed.

Experimental Setup

Participants sat at a viewing distance of approximately eighty centimeters from the screen (refresh rate 60 Hz, resolution 1920 x 1200 pixels, size 52 cm x 32.5 cm, visual angle 33° in horizontal and 22° in vertical direction) and had access to a keyboard and a mouse. EEG signals were recorded with 64 active EEG electrodes arranged according to the international 10–20 system (*ActiCap*, *BrainAmp*, *BrainVision Recorder*, BrainProducts, Munich, Germany; sampling frequency of 1000 Hz). The ground electrode was placed on the forehead, the reference electrode on the left mastoid,

one of the regular EEG electrodes on the right mastoid for later re-referencing to the linked-mastoids and another electrode below the left eye for electrooculography (EOG). Electrode impedance was set at values of 5 k Ω or less, which was possible in more than 95% of the cases. If an optimal impedance between an electrode and the scalp could not be achieved despite considerable effort, this non-optimal impedance was accepted and the experiment was started. Maximum impedance at start time was 7 k Ω at the ground electrode, 9 k Ω at the reference electrode and 26 k Ω at a scalp electrode. Stimuli were presented with in-house software written in *Processing* (version 2.2.1, <https://processing.org>) controlled by *Matlab* (MathWorks, Natick, USA).

Data Acquisition

Five female and eight male subjects with normal or corrected to normal vision, no report of eye or neurological diseases and ages ranging from 18 to 65 years (mean of 31.2 years) participated in the study. The tasks were introduced and trained at the beginning of the experiment of two hours. The participants gave their informed written consent to take part in the experiment. The study was approved by the ethics committee of the Department of Psychology and Ergonomics of the Technische Universität Berlin (reference BL_02_20140520).

The EEG data were re-referenced to the linked mastoids and band-pass filtered between 0.5 Hz and 40 Hz with an infinite impulse response forward-backward filter. The continuous multi-channel data were segmented in one second long epochs aligned to the flashing of targets and distractors, starting at 100 ms before the respective stimulus onset. Baseline correction was applied using the data within the 100 ms long interval before stimulus onset.

The participants repeated each task twenty times and viewed 47 to 50 stimuli per task repetition. The first eight markers per repetition that indicated the stimulus onset had to be discarded due to a jitter, i.e. an imprecision, in the stimulus presentation. As result, there were 165 ± 5 target and 648 ± 13 distractor epochs (mean \pm std) available per participant and experimental condition.

Data Analysis

Single-Trial Classification. The question was addressed if the neural response to target stimuli is specific to the silent counting task or if it can be also evoked by other tasks. The problem was approached by asking the subjects to perform three tasks that required to pay attention to certain stimuli. The stimuli were classified either as targets or distractors based on the immediate neural response to them. The classifiers were trained with data recorded when the subject performed one of the three tasks and tested on separate data acquired when a different task was requested. Classifiers trained in one experimental condition should be able to detect targets in different experimental conditions if the target-related neural activity is not task specific. Training and testing was performed on all possible pair-wise combinations of the three conditions. As additional reference level, every condition was inspected separately and served both for training and testing. In this case, the classification performance was assessed by splitting the data in test and training sets in a ten-fold cross-validation [21].

Spatio-temporal features for the classifications were extracted from each EEG epoch within the interval from 100 ms to 800 ms. The EEG signal was downsampled to 20 Hz in order to improve classification performance via a reduction of the dimensionality of the features [22]. A 930 dimensional feature vector was obtained for each EEG epoch by concatenating the EEG potentials measured at all 62 channels and 15 time points within the 700 ms long epoch. Classifications were performed with regularized linear discriminant analysis where the shrinkage parameter was determined analytically [23–25]. Performance was assessed with the area under the curve (AUC) of the receiver operating characteristic [26].

Single-trial classifications were performed with all samples including trials potentially corrupted by artifacts. Accepting this challenge is expected to be useful for online operation in prospective applications. Moreover, the employed multivariate methods are able to project out artifacts of various kinds.

The previously introduced classifications were conducted separately for each participant (*within*-participants). Besides, an *across*-participants classification scheme was employed in order to investigate if a transfer of the predictor is possible between subjects, which would allow to skip a time-consuming individual calibration session (cf. section 'Discussion/Single-Trial Classification'). For this purpose, classifiers were trained on the data of all participants but one and tested on the data of the respective withheld participant. The procedure was iterated such that the data of every participant were tested. Again, all combinations of training and testing condition were assessed. Moreover, the effect of the number of training subjects on the classification performance was determined. The data of one to twelve subjects were used to train a classifier (to discriminate between targets and distractors) that was tested on the data of each withheld participant. In this analysis, all experimental conditions were merged for the sake of conciseness and in view of the envisaged application case where the users are expected to perform various tasks. The training subjects were drawn at random if there existed several possibilities.

Spatio-Temporal Dynamics. Additionally, the spatio-temporal dynamics of the neural responses to the flashing of target and distractor stimuli were inspected. While the main hypothesis under investigation was tested with the classification approach detailed above, this inspection allows a better understanding of the underlying reasons for success or failure of the classifications. The measured potentials were averaged over the single EEG epochs of all participants, separately for each experimental condition, class (targets/distractors), channel and time point.

The difference between the two classes was assessed by computing the correlation between the potentials of the single EEG epochs and the class label, 1 for targets and 0 for distractors, separately for each channel and time point. The yielded correlation coefficients were squared while retaining the original sign (signed r^2 values). Again, averages across participants were calculated. The coefficients were Fisher z -transformed before averaging to make them approximately Gaussian distributed, which was reversed after averaging to bring them back to the original unit [27]. A significance threshold was not employed in order to keep the full spatio-temporal pattern including potentially subtle differences that might be exploited by the multivariate classifier, which was introduced above.

In order to ensure a clean and undisturbed visualization of the neural responses, artifact epochs had been rejected beforehand based on a maximum-minimum criterion of $100 \mu\text{V}$ for the EEG channels and of $200 \mu\text{V}$ for the EOG channel, within the post-stimulus interval. Around 133 ± 30 target (mean \pm std) and 489 ± 150 distractor epochs remained per participant and experimental condition.

Behavioral Performance. It was checked that every participant complied with the instructions and performed the tasks. For this purpose, the numbers entered and the positions clicked at were compared with the correct numbers and positions and it was statistically assessed whether the results were more accurate than it can be expected if the participants answered randomly. The distances between the correct and the entered numbers were calculated in the conditions C and A. It was assessed with Mann-Whitney U tests if the resulting distances were significantly smaller than random distances, which had been generated by shuffling the relations between correct and entered numbers a thousand times. In the condition M, the accuracies of selecting the correct target positions were computed. Mann-Whitney U tests checked if these accuracies were significantly greater than random accuracies, which had been determined by moving the targets to random positions a thousand times.

Analysis and visualization of the EEG and behavioral data were performed with Python (version 3.5.2, <http://www.python.org>), the MNE-Python software, pandas, scikit-learn and seaborn [28–33].

Results

Single-Trial Classification

Fig 2 displays the results of the within-participant classifications of target versus distractor EEG epochs. The classification performance was assessed with the AUC. This metric represents both the sensitivity and the specificity of the classifier and is insensitive to class imbalances [26]. An AUC of 0.5 constitutes the chance level of the classification. For all combinations of training and testing conditions and for every participant, the AUC was consistently better than it can be expected from random guessing. Wilcoxon signed-rank tests showed that the results were on the population level significantly above an AUC of 0.5 ($p \leq 0.05$, Bonferroni corrected for the nine combinations of training and testing conditions).

The cross-validation results (values on the diagonal of the matrix in Fig 2) might not be directly compared with the results obtained by training on one condition and testing on a different condition (on the off-diagonal of the matrix in Fig 2).

Fig 3 displays the results of the classifications across-participants. Classification performance was on the population level significantly better than chance in all cases but one (C → M, Wilcoxon statistic as above).

Data of more participants used for the classifier training resulted in a better performance when transferred to a different participant (cf. Fig 4; the three conditions were merged for this

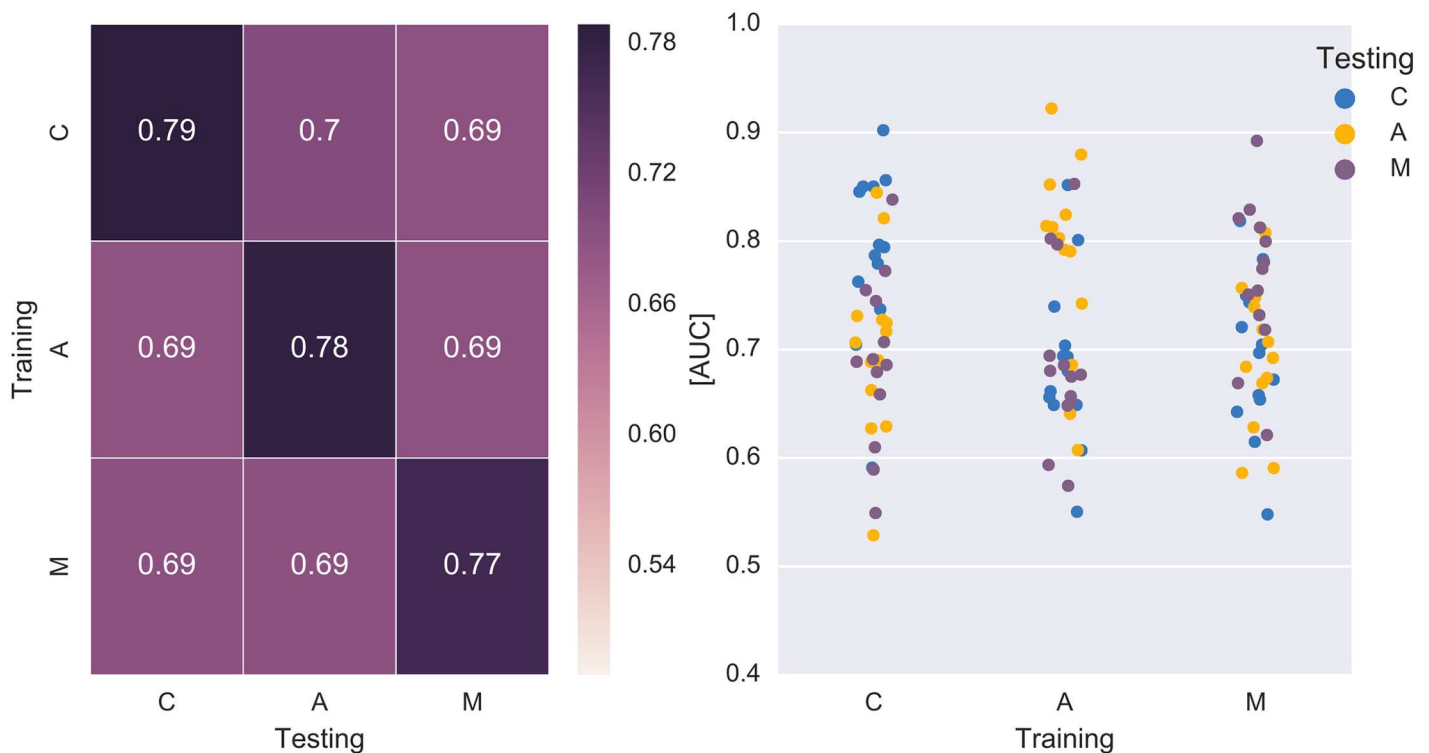


Fig 2. Average (left) and single participant (right) results of the classifications *within*-participants for all combinations of training and testing condition, measured as area under the curve of the receiver operating characteristic. All results were on the population level significantly better than random guessing ($p \leq 0.05$).

doi:10.1371/journal.pone.0165556.g002

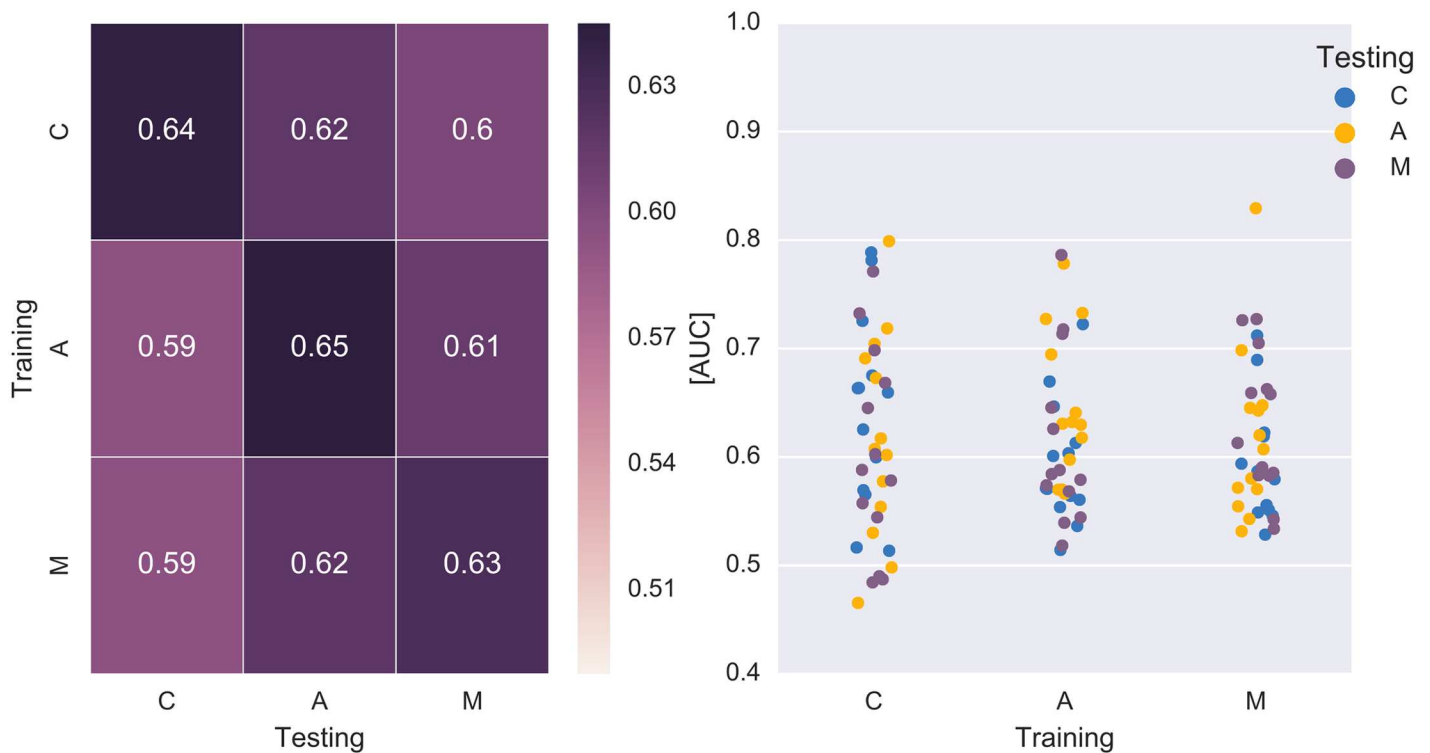


Fig 3. Average (left) and single participant (right) results of the classifications across-participants. The results were on the population level significantly better than chance, except in one case (C → M).

doi:10.1371/journal.pone.0165556.g003

analysis as motivated in section ‘Data Analysis/Single-Trial Classification’). The number of training subjects was significantly correlated with the AUC (correlation coefficients were calculated for every subject, average: $\rho = 0.50$, t-test: $p \leq 0.05$). Nevertheless, a ceiling effect can be observed for $n \geq 6$.

Spatio-Temporal Dynamics

The spatio-temporal dynamics of the neural responses to the flashing of target and distractor stimuli are visualized in Figs 5, 6 and 7. Averages across participants are displayed separately for the conditions C, A and M. Fig 5 shows the time course of the EEG potential measured at frontal, central and parietal positions along the midline of the head. Fig 6 depicts the time courses at all electrodes in color code, separately for targets (top) and distractors (center). The lower row shows the difference between the two classes. Fig 7 presents the data as scalp topographies.

Behavioral Performance

Every participant entered numbers and clicked at positions that were significantly more accurate as it can be expected by chance ($p \leq 0.05$).

Discussion

Single-Trial Classification

EEG epochs that were either aligned to targets or to distractors could be discriminated significantly better than it can be expected by chance (AUC of 0.5) for all combinations of training and testing conditions (within-participant classifications; cf. Fig 2). Discrimination based on

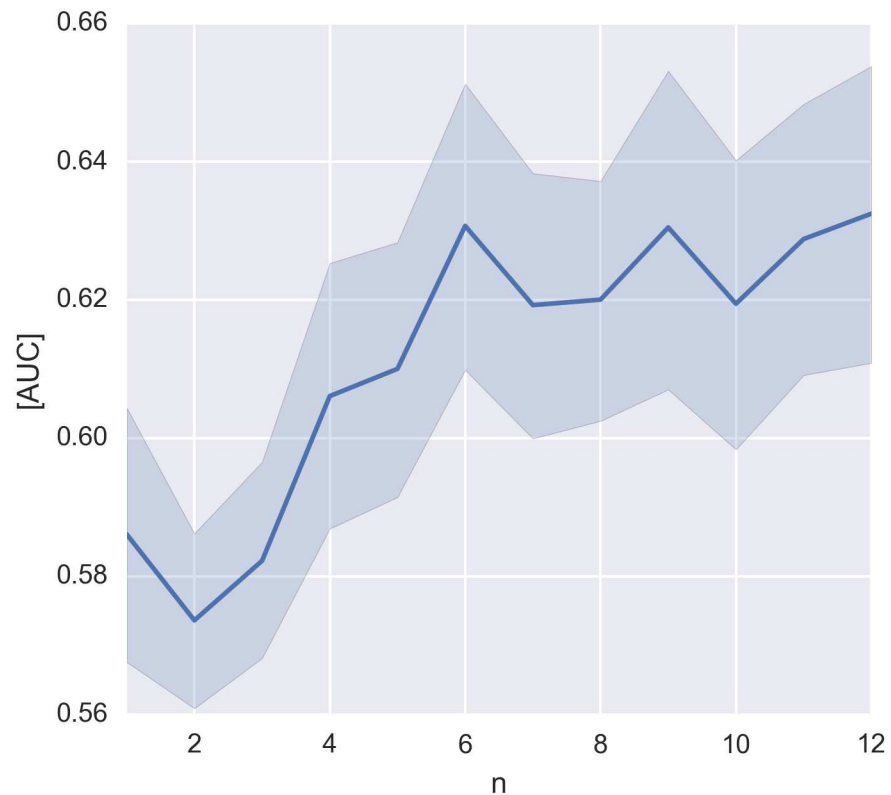


Fig 4. Performance (AUC) of the classification across-participants depending on the number (n) of participants used to train the classifier. The three experimental conditions were merged for this analysis (cf. section 'Data Analysis/Single-Trial Classification'). Bootstrapping, a resampling method, was used to estimate the 68% confidence intervals (equivalent to ± 1 standard deviation in the Gaussian case) of the mean across participants [34].

doi:10.1371/journal.pone.0165556.g004

EEG data was not only possible in the classic counting variation (C) but also when both targets and distractors required arithmetic (A) or when the positions of the targets had to be memorized (M).

Each classifier could predict targets in every experimental condition and not only in the condition where it had been trained. This successful transfer suggests that a substantial part of the neural activity evoked by targets is neither specific to the silent counting, nor to the arithmetic, nor to the memory task. Both the target recognition itself, as a result of the attention allocation, and the augmented cognitive effort are equally plausible causes for the findings, because targets required a more demanding task than distractors (at least in the conditions C and M where distractors could be simply ignored).

Tailoring the classifier to each individual person, as it is typically done in BCI experiments, would be a hindering factor for the application in HCI. A time-consuming calibration session constitutes a hurdle for the users to adopt EEG-based technology for the every-day interaction with a computer. Interestingly, however, it was possible to skip the individual classifier training and predict the task-relevant stimuli with a classifier that was trained on the data of other participants (across-participants classifications; cf. Fig 3) even if the performance was significantly inferior ($p \leq 0.05$, Wilcoxon signed-rank test) to the classification within-subjects (cf. Fig 2). Acquiring data of more participants improved the predictive performance until a ceiling level was reached for $n \geq 6$. (cf. Fig 4). Transfer learning methods could further improve the transferability between subjects [35–40].

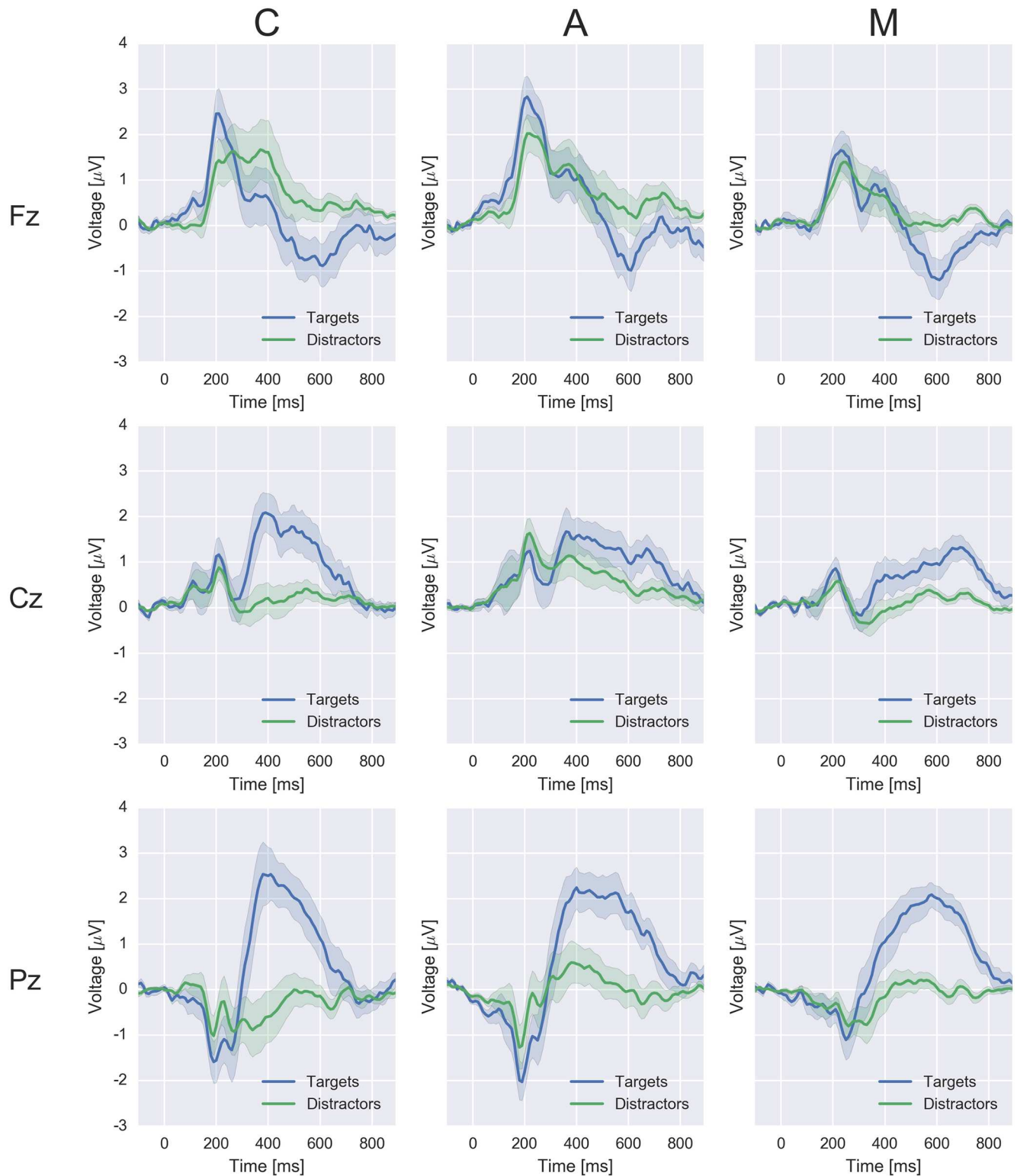


Fig 5. Time courses of the EEG responses to targets and distractors at the midline electrodes Fz, Cz and Pz in the experimental conditions C, A and M (averages over all epochs of all subjects). The respective stimulus-onset is situated at $t=0$ ms. The 68% confidence intervals were calculated with bootstrapping.

doi:10.1371/journal.pone.0165556.g005

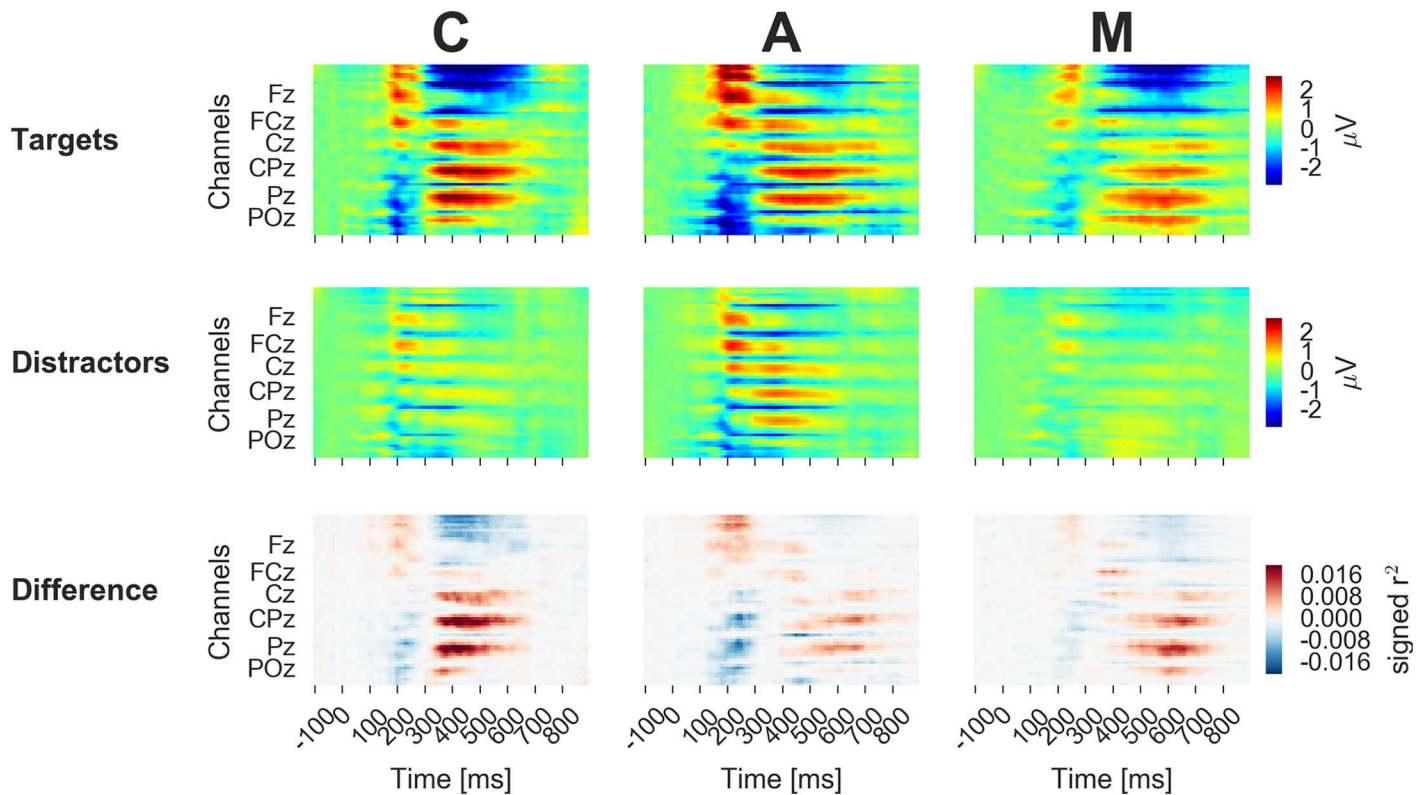


Fig 6. The time courses of the EEG responses to targets and distractors and of the corresponding difference (top, center, bottom) are displayed for every channel in color code.

doi:10.1371/journal.pone.0165556.g006

Spatio-Temporal Dynamics

The patterns in the neural data that allow differentiating between targets and distractors were inspected in order to uncover the reason for the successful classifications. Targets evoked, in all experimental conditions, an augmented late positive component in comparison to distractors in particular at the midline centroparietal and parietal electrodes (cf. Figs 5, 6 and 7), which is typical for the P300 wave [13].

Some differences between the conditions can be noted (cf. Figs 5, 6 and 7): condition C, the classic variation with silent counting, featured a comparably large difference between the potentials evoked by targets and distractors. Condition A shows a comparably large late positive deflection for distractors. In this condition, all stimuli including the distractors required arithmetic and, thus, a certain amount of attention and neural processing. Finally, the discriminative neural activity lasted longer in A and M than in C (cf. Fig 7). Presumably, the memory encoding was more variable in time in these two conditions.

Behavioral Performance

The behavioral results show that all participants complied with the instructions and performed the tasks.

Limitations

Single stimuli popped up in succession in this experiment. However, it can be expected that several words or pictograms are shown in parallel in a more realistic setting. The combination

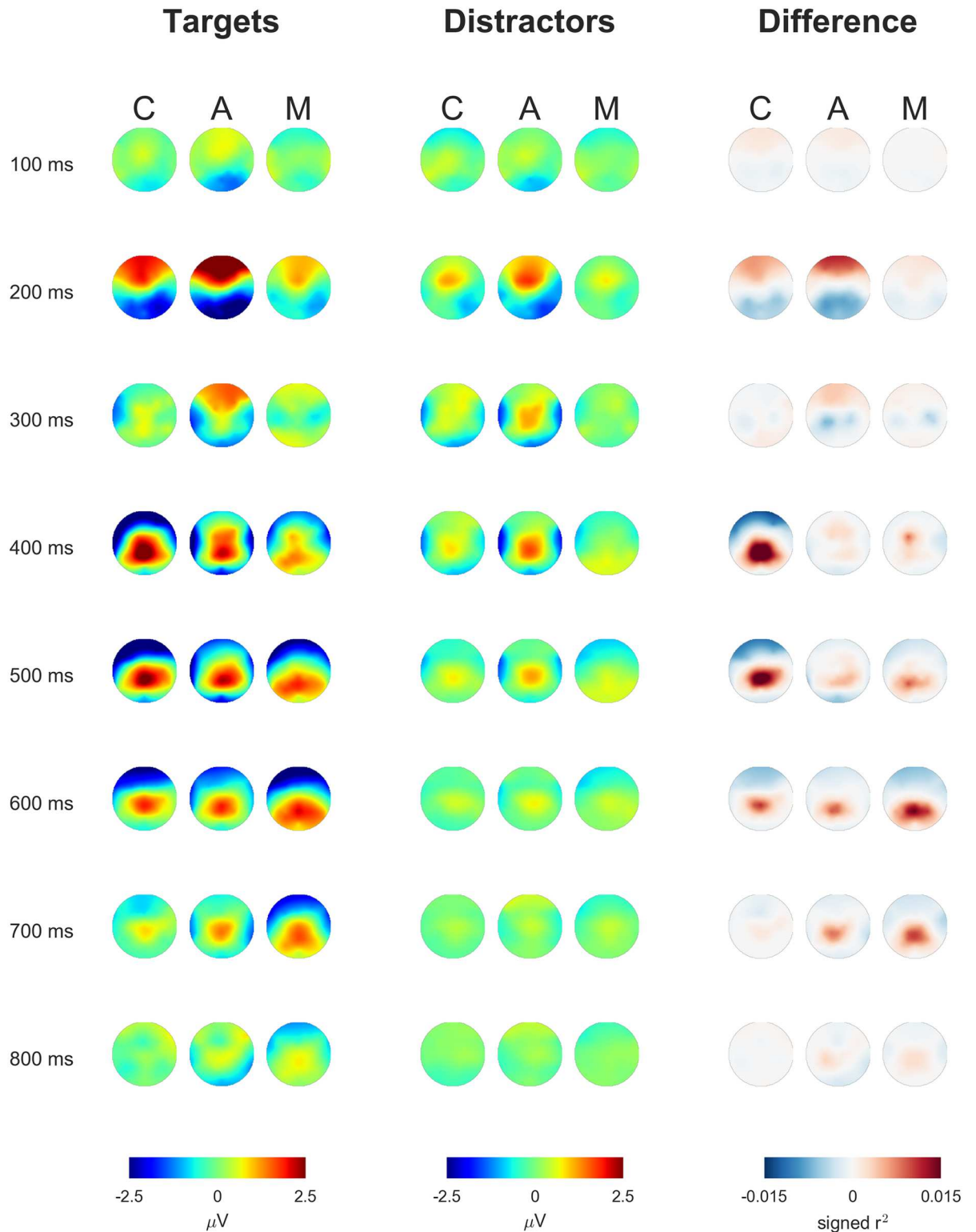


Fig 7. EEG responses to targets and distractors and the corresponding difference (left, center, right) are depicted as scalp topographies (head from above with the nose on top, average values over 50 ms long intervals around 100 ms, 200 ms, . . . , 800 ms post-stimulus-onset).

doi:10.1371/journal.pone.0165556.g007

of EEG with an eye tracker would allow to relate the neural activity to each pictogram or word [15–20]. Eye movements towards the items could be used as time points of reference for the EEG segmentation in epochs, instead of the onset of stimuli popping up on the screen. With this approach, a relevance score could be assigned to every item displayed.

We showed that the detectable neural activity evoked by targets is not specific to any of the three well-defined tasks employed in the experiment. However, it still has to be shown that relevance information can be collected implicitly in the background during the ‘natural’ interaction with a computer in the absence of precisely defined tasks.

Moreover, the stimuli used here were squares and differed only in their color. The decision if a stimulus was a target was simple and could be performed immediately. In contrast, various pictograms and words can be presented on the screen in a realistic scenario. The decision if a pictogram or a word is of interest can need sometimes less and sometimes more time. Accordingly, a variable latency of the neural response can be expected, which makes relevance estimation based on neurophysiological data more difficult [18, 19].

All stimuli were similar with respect to their salience in this experiment. Yet, in a more realistic scenario, particularly salient but not necessarily relevant stimuli could elicit a passive P300, which would result in false positive estimates (even though the passive P300 is evoked rather by auditory stimuli than by visual stimuli) [41–43].

Possible Application in the Future

Decoding the cognitive state of computer users from neuro-, peripheral-physiological or behavioral measurements (such as EEG, electrodermal activity, facial electromyography, eye movement patterns, or pupil size) became more and more of interest recently [17–19, 44–55]. The resulting information, e.g. about the user’s attention allocation and interest, is implicitly contained in the sensor data, can be recorded in the background without any effort on the part of the user, and could augment ‘traditional’ input devices such as mouse and keyboard. Until recently, the physiological measurement devices were bulky and expensive and the set-up was inconvenient and time consuming. Yet, the situation is improving at present due to technological innovations such as miniaturized, gel-free, and in-ear as well as around-ear electrodes [56–61] and a considerable drop in the price of eye trackers [62]. Further impulses can be expected from large tech companies that launch wearable physiological sensors as parts of their products, like heart-rate sensors in smart watches, or that are working on miniaturized glucose sensors in contact lenses.

Conclusion

Based on EEG data, screen content could be classified as task relevant or irrelevant, even when different mental operations were performed than during classifier training. The results suggest that the neural activity detected by the classifiers is not strictly task specific, and that presumably attention allocation or cognitive effort can be inferred from the EEG data, at least under the controlled conditions of this experimental study. This outcome may hold promise for a future technology transfer from brain-computer interfacing, where the users typically count the stimuli of interest, to relevance detection in human-computer interaction, where the users do not limit themselves to pursuing a specific activity.

Author Contributions

Conceptualization: MW BB.

Data curation: MW.

Formal analysis: MW BB.
Funding acquisition: BB.
Investigation: MW IA.
Methodology: MW BB.
Project administration: BB.
Resources: BB.
Software: MW IA BB.
Supervision: BB.
Validation: MW.
Visualization: MW.
Writing – original draft: MW.
Writing – review & editing: BB IA.

References

1. Farwell LA, Donchin E. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*. 1988; 70(6):510–523. doi: [10.1016/0013-4694\(88\)90149-6](https://doi.org/10.1016/0013-4694(88)90149-6) PMID: [2461285](https://pubmed.ncbi.nlm.nih.gov/2461285/)
2. Sellers EW, Donchin E. A P300-based brain–computer interface: initial tests by ALS patients. *Clinical Neurophysiology*. 2006; 117(3):538–548. doi: [10.1016/j.clinph.2005.06.027](https://doi.org/10.1016/j.clinph.2005.06.027) PMID: [16461003](https://pubmed.ncbi.nlm.nih.gov/16461003/)
3. Kanoh S, Miyamoto Ki, Yoshinobu T. A brain-computer interface (BCI) system based on auditory stream segregation. In: 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE; 2008. p. 642–645.
4. Guger C, Daban S, Sellers E, Holzner C, Krausz G, Carabalona R, et al. How many people are able to control a P300-based brain–computer interface (BCI)? *Neuroscience Letters*. 2009; 462(1):94–98. doi: [10.1016/j.neulet.2009.06.045](https://doi.org/10.1016/j.neulet.2009.06.045) PMID: [19545601](https://pubmed.ncbi.nlm.nih.gov/19545601/)
5. Brouwer AM, Van Erp JB. A tactile P300 brain-computer interface. *Frontiers in Neuroscience*. 2010; 4:19. doi: [10.3389/fnins.2010.00019](https://doi.org/10.3389/fnins.2010.00019) PMID: [20582261](https://pubmed.ncbi.nlm.nih.gov/20582261/)
6. Treder MS, Schmidt NM, Blankertz B. Gaze-independent brain–computer interfaces based on covert attention and feature attention. *Journal of Neural Engineering*. 2011; 8(6):066003. doi: [10.1088/1741-2560/8/6/066003](https://doi.org/10.1088/1741-2560/8/6/066003) PMID: [21975312](https://pubmed.ncbi.nlm.nih.gov/21975312/)
7. Schreuder M, Rost T, Tangermann M. Listen, you are writing! Speeding up online spelling with a dynamic auditory BCI. *Frontiers in Neuroscience*. 2011; 5:112. doi: [10.3389/fnins.2011.00112](https://doi.org/10.3389/fnins.2011.00112) PMID: [22016719](https://pubmed.ncbi.nlm.nih.gov/22016719/)
8. Liu Y, Zhou Z, Hu D. Gaze independent brain–computer speller with covert visual search tasks. *Clinical Neurophysiology*. 2011; 122(6):1127–1136. <http://dx.doi.org/10.1016/j.clinph.2010.10.049>. PMID: [21163695](https://pubmed.ncbi.nlm.nih.gov/21163695/)
9. Manyakov NV, Chumerin N, Combaz A, Van Hulle MM. Comparison of classification methods for P300 brain-computer interface on disabled subjects. *Computational Intelligence and Neuroscience*. 2011; 2011:2. doi: [10.1155/2011/519868](https://doi.org/10.1155/2011/519868) PMID: [21941530](https://pubmed.ncbi.nlm.nih.gov/21941530/)
10. Acqualagna L, Blankertz B. Gaze-independent BCI-spelling using rapid serial visual presentation (RSVP). *Clinical Neurophysiology*. 2013; 124(5):901–908. doi: [10.1016/j.clinph.2012.12.050](https://doi.org/10.1016/j.clinph.2012.12.050) PMID: [23466266](https://pubmed.ncbi.nlm.nih.gov/23466266/)
11. An X, Höhne J, Ming D, Blankertz B. Exploring combinations of auditory and visual stimuli for gaze-independent brain-computer interfaces. *PLOS ONE*. 2014; 9(10):e111070. doi: [10.1371/journal.pone.0111070](https://doi.org/10.1371/journal.pone.0111070) PMID: [25350547](https://pubmed.ncbi.nlm.nih.gov/25350547/)
12. Sutton S, Braren M, Zubin J, John ER. Evoked-potential correlates of stimulus uncertainty. *Science (New York, NY)*. 1965; 150(3700):1187–1188. doi: [10.1126/science.150.3700.1187](https://doi.org/10.1126/science.150.3700.1187)

13. Picton TW. The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology: Official Publication of the American Electroencephalographic Society*. 1992; 9(4):456–479. doi: [10.1097/00004691-199210000-00002](https://doi.org/10.1097/00004691-199210000-00002) PMID: [1464675](https://pubmed.ncbi.nlm.nih.gov/1464675/)
14. Polich J. Updating P300: An Integrative Theory of P3a and P3b. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*. 2007; 118(10):2128–2148. doi: [10.1016/j.clinph.2007.04.019](https://doi.org/10.1016/j.clinph.2007.04.019) PMID: [17573239](https://pubmed.ncbi.nlm.nih.gov/17573239/)
15. Brouwer AM, Reuderink B, Vincent J, van Gerven MAJ, van Erp JBF. Distinguishing between target and nontarget fixations in a visual search task using fixation-related potentials. *Journal of Vision*. 2013; 13(3):17. doi: [10.1167/13.3.17](https://doi.org/10.1167/13.3.17) PMID: [23863335](https://pubmed.ncbi.nlm.nih.gov/23863335/)
16. Kaunitz LN, Kamienkowski JE, Varatharajah A, Sigman M, Quiroga RQ, Ison MJ. Looking for a face in the crowd: Fixation-related potentials in an eye-movement visual search task. *NeuroImage*. 2014; 89:297–305. doi: [10.1016/j.neuroimage.2013.12.006](https://doi.org/10.1016/j.neuroimage.2013.12.006) PMID: [24342226](https://pubmed.ncbi.nlm.nih.gov/24342226/)
17. Kauppi JP, Kandemir M, Saarinen VM, Hirvenkari L, Parkkonen L, Klami A, et al. Towards brain-activity-controlled information retrieval: Decoding image relevance from MEG signals. *NeuroImage*. 2015; 112:288–298. doi: [10.1016/j.neuroimage.2014.12.079](https://doi.org/10.1016/j.neuroimage.2014.12.079) PMID: [25595505](https://pubmed.ncbi.nlm.nih.gov/25595505/)
18. Wenzel MA, Golenia JE, Blankertz B. Classification of eye fixation related potentials for variable stimulus saliency. *Frontiers in Neuroprosthetics*. 2016; 10(23). doi: [10.3389/fnins.2016.00023](https://doi.org/10.3389/fnins.2016.00023) PMID: [26912993](https://pubmed.ncbi.nlm.nih.gov/26912993/)
19. Uščumlić M, Blankertz B. Active visual search in non-stationary scenes: coping with temporal variability and uncertainty. *Journal of Neural Engineering*. 2016; 13(1):016015. doi: [10.1088/1741-2560/13/1/016015](https://doi.org/10.1088/1741-2560/13/1/016015) PMID: [26726921](https://pubmed.ncbi.nlm.nih.gov/26726921/)
20. Finke A, Essig K, Marchioro G, Ritter H. Toward FRP-Based Brain-Machine Interfaces—Single-Trial Classification of Fixation-Related Potentials. *PLOS ONE*. 2016; 11(1):e0146848. doi: [10.1371/journal.pone.0146848](https://doi.org/10.1371/journal.pone.0146848) PMID: [26812487](https://pubmed.ncbi.nlm.nih.gov/26812487/)
21. Lemm S, Blankertz B, Dickhaus T, Müller KR. Introduction to machine learning for brain imaging. *NeuroImage*. 2011; 56(2):387–399. doi: [10.1016/j.neuroimage.2010.11.004](https://doi.org/10.1016/j.neuroimage.2010.11.004) PMID: [21172442](https://pubmed.ncbi.nlm.nih.gov/21172442/)
22. Blankertz B, Lemm S, Treder M, Haufe S, Müller KR. Single-trial analysis and classification of ERP components—A tutorial. *NeuroImage*. 2011; 56(2):814–825. doi: [10.1016/j.neuroimage.2010.06.048](https://doi.org/10.1016/j.neuroimage.2010.06.048) PMID: [20600976](https://pubmed.ncbi.nlm.nih.gov/20600976/)
23. Friedman JH. Regularized Discriminant Analysis. *Journal of the American Statistical Association*. 1989; 84(405):165. doi: [10.1080/01621459.1989.10478752](https://doi.org/10.1080/01621459.1989.10478752)
24. Ledoit O, Wolf M. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*. 2004; 88(2):365–411. doi: [10.1016/S0047-259X\(03\)00096-4](https://doi.org/10.1016/S0047-259X(03)00096-4)
25. Schäfer J, Strimmer K. A Shrinkage Approach to Large-Scale Covariance Matrix Estimation and Implications for Functional Genomics. *Statistical Applications in Genetics and Molecular Biology*. 2005; 4(1). doi: [10.2202/1544-6115.1175](https://doi.org/10.2202/1544-6115.1175) PMID: [16646851](https://pubmed.ncbi.nlm.nih.gov/16646851/)
26. Fawcett T. An Introduction to ROC Analysis. *Pattern Recognition Letters*. 2006; 27(8):861–874. doi: [10.1016/j.patrec.2005.10.010](https://doi.org/10.1016/j.patrec.2005.10.010)
27. Silver NC, Dunlap WP. Averaging correlation coefficients: should Fisher’s z transformation be used? *Journal of Applied Psychology*. 1987; 72(1):146. doi: [10.1037/0021-9010.72.1.146](https://doi.org/10.1037/0021-9010.72.1.146)
28. Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, et al. MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*. 2013; 7(267). doi: [10.3389/fnins.2013.00267](https://doi.org/10.3389/fnins.2013.00267) PMID: [24431986](https://pubmed.ncbi.nlm.nih.gov/24431986/)
29. Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, et al. MNE software for processing MEG and EEG data. *NeuroImage*. 2014; 86:446–460. <http://dx.doi.org/10.1016/j.neuroimage.2013.10.027>. PMID: [24161808](https://pubmed.ncbi.nlm.nih.gov/24161808/)
30. McKinney W. *pandas: a Foundational Python Library for Data Analysis and Statistics*; 2011.
31. McKinney W. *Data Structures for Statistical Computing in Python*. In: van der Walt S, Millman J, editors. *Proceedings of the 9th Python in Science Conference*; 2010. p. 51–56.
32. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. 2011; 12:2825–2830.
33. Waskom M, Botvinnik O, Hobson P, Warmenhoven J, Cole JB, Halchenko Y, et al. *seaborn: v0.6.0* (June 2015); 2015. Available from: <http://dx.doi.org/10.5281/zenodo.19108>.
34. Efron B. *Bootstrap Methods: Another Look at the Jackknife*. vol. 7. The Institute of Mathematical Statistics; 1979. Available from: <http://dx.doi.org/10.1214/aos/1176344552>.
35. Lu S, Guan C, Zhang H. Unsupervised Brain Computer Interface Based on Intersubject Information and Online Adaptation. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*. 2009; 17(2):135–145. doi: [10.1109/TNSRE.2009.2015197](https://doi.org/10.1109/TNSRE.2009.2015197) PMID: [19228561](https://pubmed.ncbi.nlm.nih.gov/19228561/)

36. Fazli S, Popescu F, Danóczy M, Blankertz B, Müller KR, Grozea C. Subject-independent mental state classification in single trials. *Neural Networks*. 2009; 22(9):1305–1312. <http://dx.doi.org/10.1016/j.neunet.2009.06.003>. PMID: 19560898
37. Fazli S, Danóczy M, Schelldorfer J, Müller KR. ℓ_1 -penalized linear mixed-effects models for high dimensional data with application to BCI. *NeuroImage*. 2011; 56(4):2100–2108. <http://dx.doi.org/10.1016/j.neuroimage.2011.03.061>. PMID: 21463695
38. Kindermans PJ, Tangermann M, Müller KR, Schrauwen B. Integrating dynamic stopping, transfer learning and language models in an adaptive zero-training ERP speller. *Journal of Neural Engineering*. 2014; 11(3):035005. doi: [10.1088/1741-2560/11/3/035005](https://doi.org/10.1088/1741-2560/11/3/035005) PMID: 24834896
39. Jayaram V, Alamgir M, Altun Y, Schölkopf B, Grosse-Wentrup M. Transfer learning in brain-computer interfaces. *IEEE Computational Intelligence Magazine*. 2016; 11(1):20–31. doi: [10.1109/MCI.2015.2501545](https://doi.org/10.1109/MCI.2015.2501545)
40. Koyamada S, Shikauchi Y, Nakae K, Koyama M, Ishii S. Deep learning of fMRI big data: a novel approach to subject-transfer decoding. *arXiv preprint arXiv:150200093*. 2015.
41. Squires NK, Squires KC, Hillyard SA. Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*. 1975; 38(4):387–401. doi: [10.1016/0013-4694\(75\)90263-1](https://doi.org/10.1016/0013-4694(75)90263-1) PMID: 46819
42. Bennington JY, Polich J. Comparison of P300 from passive and active tasks for auditory and visual stimuli. *International Journal of Psychophysiology*. 1999; 34(2):171–177. [http://dx.doi.org/10.1016/S0167-8760\(99\)00070-7](http://dx.doi.org/10.1016/S0167-8760(99)00070-7). PMID: 10576401
43. Jeon YW, Polich J. P3a from a passive visual stimulus task. *Clinical Neurophysiology*. 2001; 112(12):2202–2208. [http://dx.doi.org/10.1016/S1388-2457\(01\)00663-0](http://dx.doi.org/10.1016/S1388-2457(01)00663-0). PMID: 11738190
44. Zander TO, Kothe C. Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general. *Journal of Neural Engineering*. 2011; 8(2):025005. doi: [10.1088/1741-2560/8/2/025005](https://doi.org/10.1088/1741-2560/8/2/025005) PMID: 21436512
45. Eugster MJA, Ruotsalo T, Spapé MM, Kosunen I, Barral O, Ravaja N, et al. Predicting Term-relevance from Brain Signals. In: *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*. SIGIR'14. New York, NY, USA: ACM; 2014. p. 425–434. Available from: <http://doi.acm.org/10.1145/2600428.2609594>.
46. Wenzel M, Moreira C, Lungu IA, Bogojeski M, Blankertz B. Neural Responses to Abstract and Linguistic Stimuli with Variable Recognition Latency. In: Blankertz B, Jacucci G, Gamberini L, Spagnolli A, Freeman J, editors. *Symbiotic Interaction*. vol. 9359 of *Lecture Notes in Computer Science*. Springer International Publishing; 2015. p. 172–178. Available from: http://dx.doi.org/10.1007/978-3-319-24917-9_19.
47. Golenia JE, Wenzel M, Blankertz B. Live Demonstrator of EEG and Eye-Tracking Input for Disambiguation of Image Search Results. In: Blankertz B, Jacucci G, Gamberini L, Spagnolli A, Freeman J, editors. *Symbiotic Interaction*. vol. 9359 of *Lecture Notes in Computer Science*. Springer International Publishing; 2015. p. 81–86. Available from: http://dx.doi.org/10.1007/978-3-319-24917-9_8.
48. Oliveira FTP, Aula A, Russell DM. Discriminating the Relevance of Web Search Results with Measures of Pupil Size. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI'09. New York, NY, USA: ACM; 2009. p. 2209–2212.
49. Hardoon DR, Pasupa K. Image Ranking with Implicit Feedback from Eye Movements. In: *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*. ETRA'10. New York, NY, USA: ACM; 2010. p. 291–298. Available from: <http://doi.acm.org/10.1145/1743666.1743734>.
50. Haji Mirza SNH, Proulx M, Izquierdo E. Gaze Movement Inference for User Adapted Image Annotation and Retrieval. In: *Proceedings of the 2011 ACM Workshop on Social and Behavioural Networked Media Access*. SBNMA'11. New York, NY, USA: ACM; 2011. p. 27–32.
51. Cole MJ, Gwizdka J, Belkin NJ. Physiological Data as Metadata. In: *SIGIR 2011 Workshop on Enriching Information Retrieval (ENIR 2011)*, Beijing, China; 2011.
52. Cole MJ, Gwizdka J, Liu C, Bierig R, Belkin NJ, Zhang X. Task and user effects on reading patterns in information search. *Interacting with Computers*. 2011; 23(4):346–362. doi: [10.1016/j.intcom.2011.04.007](https://doi.org/10.1016/j.intcom.2011.04.007)
53. Gwizdka J, Cole MJ. Inferring cognitive states from multimodal measures in information science. In: *ICMI 2011 Workshop on Inferring Cognitive and Emotional States from Multimodal Measures (ICMI'2011 MMCogEmS)(Alicante:)*; 2011.
54. Hajimirza SN, Proulx MJ, Izquierdo E. Reading Users' Minds From Their Eyes: A Method for Implicit Image Annotation. *IEEE Transactions on Multimedia*. 2012; 14(3):805–815. doi: [10.1109/TMM.2012.2186792](https://doi.org/10.1109/TMM.2012.2186792)

55. Barral O, Eugster MJA, Ruotsalo T, Spapé MM, Kosunen I, Ravaja N, et al. Exploring Peripheral Physiology As a Predictor of Perceived Relevance in Information Retrieval. In: Proceedings of the 20th International Conference on Intelligent User Interfaces. IUI'15. New York, NY, USA: ACM; 2015. p. 389–399. Available from: <http://doi.acm.org/10.1145/2678025.2701389>.
56. Nikulin VV, Kegeles J, Curio G. Miniaturized electroencephalographic scalp electrode for optimal wearing comfort. *Clinical Neurophysiology*. 2010; 121(7):1007–1014. doi: [10.1016/j.clinph.2010.02.008](https://doi.org/10.1016/j.clinph.2010.02.008) PMID: [20227914](https://pubmed.ncbi.nlm.nih.gov/20227914/)
57. Grozea C, Voinescu CD, Fazli S. Bristle-sensors—Low-cost Flexible Passive Dry EEG Electrodes for Neurofeedback and BCI Applications. *Journal of Neural Engineering*. 2011; 8:025008. doi: [10.1088/1741-2560/8/2/025008](https://doi.org/10.1088/1741-2560/8/2/025008) PMID: [21436526](https://pubmed.ncbi.nlm.nih.gov/21436526/)
58. Zander TO, Lehne M, Ihme K, Jatzev S, Correia J, Kothe C, et al. A dry EEG-system for scientific research and brain–computer interfaces. *Frontiers in Neuroscience*. 2011; 5(53):1–10. doi: [10.3389/fnins.2011.00053](https://doi.org/10.3389/fnins.2011.00053) PMID: [21647345](https://pubmed.ncbi.nlm.nih.gov/21647345/)
59. Guger C, Krausz G, Allison BZ, Edlinger G. Comparison of dry and gel based electrodes for P300 brain-computer interfaces. *Frontiers in Neuroscience*. 2012; 6(60):1–7. doi: [10.3389/fnins.2012.00060](https://doi.org/10.3389/fnins.2012.00060) PMID: [22586362](https://pubmed.ncbi.nlm.nih.gov/22586362/)
60. Looney D, Kidmose P, Mandic DP. Ear-EEG: user-centered and wearable BCI. In: *Brain-Computer Interface Research*. Springer; 2014. p. 41–50. doi: [10.1007/978-3-642-54707-2_5](https://doi.org/10.1007/978-3-642-54707-2_5)
61. Debener S, Emkes R, De Vos M, Bleichner M. Unobtrusive ambulatory EEG using a smartphone and flexible printed electrodes around the ear. *Scientific reports*. 2015; 5. doi: [10.1038/srep16743](https://doi.org/10.1038/srep16743) PMID: [26572314](https://pubmed.ncbi.nlm.nih.gov/26572314/)
62. Dalmaijer E. Is the low-cost EyeTribe eye tracker any good for research? *PeerJ PrePrints*; 2014. e585v1.