



Medizinische Fakultät · Geissweg 5 · 72076 Tübingen ·

Confidential Ombuds-Matter
Tübingen, 22. November 2018

xxxxx

- **Examination of the allegation of scientific misconduct with regard to the authors of an original work published in 2017 in the journal PlosBiol.**

-

-

- With Email dated 16 April 2018 to Prof. xxx (confidant of Med. Faculty or Clinical Research) and myself as the MFT's confidant for laboratory research, the suspicion of scientific misconduct in the context of an original paper published in 2017 in the journal PLoSBiol 15(1) e1002593 entitled Brain computer interface-based communication in the completely locked-in state by the authors Chaudhary U, Xia B., Silvoni, S., Cohen, L. G. and Birbaumer N. (see Appendix A and B).

- Previously, the whistleblower had sent an identical e-mail on 9 April 2018 to the confidant of the Mathematical and Natural Science Faculty of the University of Tübingen, Prof. Böhme. Prof. Böhme saw the substance of the suspicion as given and referred the whistleblower to the confidantes of the Medical Faculty Tübingen, because the suspicion concerned members of the MFT (see also Appendix A).

- The corresponding authors of the incriminated PLoSBiol publication are together

- Ujwal Chaudhary (first author) and Niels Birbaumer (last author and head of the working group). Both authors are members of the medical faculty. At the time of publication

- . .

- The authors were listed below the following affiliations:

- - Ujwal Chaudhary: Institute of Medical Psychology and Behavioral Neuroscience, University of Tübingen

- - Bin Xia: Shanghai Maritime University, Shanghai, China

- - Stefano Silvoni: Dept. of Cognitive and Clinical Neuroscience, Central Institute of Mental Health, Mannheim

- - Leonardo G. Cohen: Human Cortical Physiology and Stroke Neurorehabilitation Section, National Institute of Neurological Disorders and Stroke, National Institute of Health, Bethesda, Maryland, USA-

- Niels Birbaumer: Institute of Medical Psychology and Behavioral Neuroscience, University of Tübingen

In various personal as well as telephone conversations with the ombudspersons as well as electronically transmitted documents the whistleblower has given detailed reasons for his accusation of scientific misconduct.

Due to different scheduling availability of the ombudspersons and the whistleblower a first confidential personal meeting could only be held on 17 May 2018 and a second one on 7 June 2018. In the aftermath of these talks the whistleblower also transmitted various documents, which give insight, not only into the facts described by him, but in particular also into the discussions that were held between the whistleblower, Prof. Birbaumer and Dr. Chaudhary as well as other persons familiar with the situation regarding the different view of the 2017 PLoSBiol publication.

In the meantime (on 26.4.2018) xxx, German Research Foundation, Scientific Integrity Department, informed ombudsperson xxx by telephone that the DFG had also been informed by the whistleblower about the accusation of scientific misconduct in the context of the PLoSBiol publication of the Birbaumer working group. xxx explained that the DFG was not prepared to set up a DFG Committee but would first await the outcome of the local commissions. In this context it should be mentioned that the study that is presented in the incriminated PLoSBiol publication was financed with a material grant of the DFG to Prof. Birbaumer (DFG-Az: Bi.195). In addition, the publication also mentions other funding initiatives: Foundation Volkswagen, Federal Ministry of Education and Research-BMBF (grant number 136W0053), Baden-Württemberg-Stiftung, Eva und Horst Köhler-Stiftung, National Natural Science Foundation of China, and the European Union (application no. 686764).

- **Background to the results and outcomes of the PLoSBiol publication:**

Prof. Birbaumer's working group has been working for years on the development of a brain-computer interface to enable communication with patients in the Complete Locked-in State (CLIS). The Locked-In-Syndrome (LIS) is based on a degenerative disease of the brain, especially of the pons. The pons is a part of the brain that belongs together with the cerebellum to the mesencephalon (hindbrain). The selective and complete damage of the motor nerve tracts of this area of the brain with preservation of the remaining tissue causes the unusual appearance of the LIS. Accordingly, this leads to paralysis of the four extremities and the horizontal motor function of the eye, while the vertical view, controlled by the rostral mesencephalon, is still functioning. The function of the caudal cranial nerves (swallowing, speaking, mimic functions) also fails. These degeneration, associated with the loss of innervation of the musculoskeletal system, leads to a gradual loss of patient mobility. After a certain degree of severity, patients can only be kept alive through artificial feeding and ventilation. In LIS patients, the only way to communicate with the outside world is through movements of the eyes, whose muscle groups can still be controlled up to the transition to the Complete Locked In state (CLIS). The Complete-Locked-in-State is when, through further brain stem lesions above the pons additional parts of the midbrain are affected and thus the complete loss of the eyelid and vertical eye movements ensues. CLIS patients are thus deprived of any possibility of movement and communication while fully conscious.

Against this background the working group of Prof. Birbaumer is working on the development of a brain computer interface (BCI), which uses different patterns of the brain activity and translates them into a voice command. Accordingly, with the help of a BCI, the brain activities of the CLIS patients, which stand for imagined "yes" or "no" answers could be derived, and thus communication with the patients

could be enabled on the basis of questions that can be answered by "Yes" or "No". The prerequisite for this, however, is first and foremost the clean recording and analysis of brain activity in the CLIS patients on questions that can only be answered with "Yes" or "No" and patients can only think "Yes" or "No". Changes in the frontocentral oxygenation of the brain, which can be measured with the method of online near infrared spectroscopy (fNIRS) serve as measurement parameters. The determination the fNIRS pattern (fNIRS classification), which uniquely stands for "Yes" or "No" answers must be done with the patients in training sessions with a correspondingly high number of "Yes" and "No" training questions. Subsequently, the corresponding brain signals must be analysed to determine if statistically significant differences between "yes" and "no" responses can be detected in the data.

After the development of clearly differentiable "Yes" or "No" fNIRS patterns, the patients are then presented with open questions, which can neither be answered with "Yes" or "No" by the patients. Based on the fNIRS-Classification the training data of the training sessions can then be used to create a model by means of machine learning, which is applied to a separate test data set to determine a classification accuracy that enables the accuracy with which the yes/no answers can be distinguished. If the accuracy of the answers is significantly higher than the guessing probability, it can be concluded that the BCI technology works and that CLIS patients are thus able to communicate reliably with their environment.

The incriminated PLoS Biol publication of the Birbaumer working group describes studies in 4 CLIS patients suffering from advanced amyotrophic lateral sclerosis (ALS). Three of these patients underwent 46 and one patient 20 questioning sessions. The results of this publication show that the online fNIRS classifications of personal questions with Yes/No answers and .of the open questions using SVM (linear support vector machine) led to a correctness probability of over 70%, i.e. the corresponding questions were answered with an accuracy significantly above the guessing probability. According to the authors, the results presented are thus the first step towards eliminating the CLIS stage, at least in ALS patients.

Scientific background and allegations of the whistleblower

• Scientific background

The whistleblower is a proven expert in the research area of Brain-Computer Interfaces (BCI) and has been active at an institute that has contributed significantly to the development of the Brain-Computer Interface. The whistleblower was not and is not a member of the working group of Prof. Birbaumer, but has had a scientific cooperation on various aspects of BCI technology with the working group of Prof. Birbaumer for 5 years. These cooperations resulted in a total of 9 Pubmed-listed publications between 2014 and 2018. In two of these publications on the topics Comparing metrics to evaluate performance of regression methods for decoding of neural signals (2015) and Decoding of motor intentions from epidural E oG recordings in severely paralyzed chron7c stroke patients (2014), the author was also the first author (see also Annex 1: PubMed publication listing). Against this scientific background, the author is very familiar with the topic, study design, methodology, data collection and data analysis described in the publication with the inclusion of the study. Due to the existing cooperation with the working group of Prof. Birbaumer, the author also has insight into the methodology used and the data sets collected on which the results written in the PlosBiol publication are based.

It is not correct that the whistleblower is “a proven expert in the research area of Brain-Computer Interfaces (BCI)”. The whistleblower never participated in the collection of data from completely locked in patients nor did he ever analyze data collected with functional near infrared spectroscopy. The whistleblower only analyzed data collected and preprocessed by experts in the respective fields in cooperation with them. He is not at all knowledgeable in the method of data collection and data analysis of the study he criticized.

- Allegations of the whistleblower:

The accusations of the whistleblower that address scientific misconduct in the context of the PlosBiol publication

1. purposeful and selective data elimination with regard to statistical evaluation of the offline analysis.

None of the valid feedback and training data were eliminated. In this paper only those sessions were not included in the model building, which yielded less than 65% accuracy – otherwise a BCI would not be feasible. In these severely disabled patients, moreover, the model has to be built anew every day, since there is high variability in the brain activation data. Building a model over several days is not feasible and against the state of the art. In addition, test sessions that are automatically recorded by the equipment and sessions with technical errors, that could not be evaluated, were excluded. Finally, all feedback sessions were included in the data analysis, none of these sessions was eliminated.

2. the deliberate representation of online data that does not correspond to the data actually collected.

We reported all data that were relevant for the results as requested by PLoS Biology but not all data the whistle blower had access to. The whistle blower received sessions before and after the data inclusion deadline because these patients are continuously assessed. In addition, the whistle blower had access to data that were eliminated due to technical failure, which was visible in the data due to wrong triggers. All these data were given to him in the course of his request to cooperate in determining if there had been any problems in data collection and analysis. It is possible that the whistle blower may have included eliminated data in his data analysis although he could have seen that there are trigger problems. This may have led to insignificant results in addition to the fact that the whistle blower wrongly averaged the near infrared spectroscopy data (see below).

3. The dissemination and publication of the results of the PlosBiol application by the working group leader in the form of media reports despite the information provided to him by the whistleblower that the study was incorrect and that the published data could not be reproduced on the basis of the accessible original data.

The paper was published January 31, 2017 but the whistle blower only sent a report with allegations of wrong data analysis on October 9, 2017. Therefore this is a misrepresentation of the facts.

- Justification of the accusations by the whistleblower:

For the ombudspersons (who, however, are not specialists in the subject in question in view of the background of their field of expertise), the whistleblower gives well founded reasons related to the statistical methods required in the field of BCI and the subject of BCI:

➤ Errors were made in the statistical analyses. The statistical method used shows highly significant results even with completely random data. When using valid statistical methods, there is no significant difference in brain activity in yes/no responses.

It should be noted that neither the verdict of the university nor that of the DFG found incorrect data analysis procedures in the analysis performed by Chaudhary et al. 2017 and 2019. A lack of replicability occurs often in science and would be of some concern had the whistleblower used the exact same method as the authors of the PLoS Biology 2017 paper and the exact same data base. This was not the case. Therefore obtaining different results with different methods should be a matter of scientific discourse not of scientific misconduct. In addition, there is evidence in the 12-page report of the whistle blower that he did not preprocess the functional near infrared spectroscopy data in the proper manner and that he did not eliminate trials that could not be used for model building in this patient population. We commented on this in great detail in our response to the DFG.

➤ In the required offline analysis, training and validation data sets were most likely not properly separated. A total of 273 different model parameters tested. This led to 273 different results, from

these 273 results however only the best result corresponding to the hypothesis was reported in the PLoSBiol publication.

We have stated that we separated the training and the validation data sets and this was accepted by both the University and the DFG commissions. We never obtained 273 results and we never selected the best result. These allegations could not be substantiated by either committee.

➤ The PlosBiol publication presents data from online experiments. In the published and therefore publicly accessible data set, however, there is no reference to online experiments or the results of online experiments are not available. These missing data were also not provided in response to a request from the journal PLoSBiol.

We uploaded all sessions of the online data analysis (feedback and open questions). In the beginning of the study we only obtained summary data from the yes-no trials, which were computed by the analysis program and are time stamped. We reported them together with the single trial data of the later sessions although these early sessions yielded lower percentage correct trials than the later single session data. Eliminating the summary statistics sessions would have biased the data in our favour, which we wanted to avoid and would not have been good scientific practice. In addition, we provided all data as requested by PLoS Biology.

➤ In the PLoSBiol publication the information regarding the number of online sessions is inconsistent and not comprehensible. Table 1 of the publication shows two online sessions with open questions for patient F, Fig. 2 shows 3 online sessions with open questions for the same patient F.

There was a mistake in Table 1, which stated 2 rather than 3 sessions for patient F, but an error is not falsification of data. These data were also uploaded with 3 sessions as noted in Figure 2.

➤ Since the whistleblower has access to the complete data sets of this work through the cooperation with the Birbaumer working group, he can prove that for many online experiments there are no data available and that there is a high probability that these experiments were not carried out. The online experiments for which data are available do not correspond to those in the publication.

The whistleblower had access to ALL data of the Birbaumer group after he expressed an interest in cooperation in September 2017. Since the group is continuously running experiments with the patients, he may have included data that were not subject of the specific study in PLoS Biology, which ran from April to October 2014. He obtained data that were collected until 2017.

Further activities of the whistleblower to clarify the data conflict with the authors of the PLoSBiol study

➤ Even before making contact with the MFT ombudspersons, the whistleblower made several attempts with Prof. Birbaumer and the first author of the PlosBiol publication to discuss and clarify the facts of the case on the basis of his arguments listed above (see passage: Justification of the accusations). In various discussions and via e-mail correspondence, the first author Dr. Chaudhary and the scientific head of the working group, Prof. Birbaumer, who is responsible for all scientific aspects, were informed about the mistakes and misrepresentations in the Plos Biol publication and the authors were be persuaded to correct the publication. According to the whistleblower, Dr.' Chaudhary, however, was completely unresponsive, while Prof. Birbaumer understood and understood the arguments of the whistleblower, but showed no interest in correcting the corresponding mistakes in the work (see Appendix 1A-1L). A former member of the Birbaumer working group, Mr Guilherme Auguste, who was involved in these discussions (at least via e-mail), but who was present at a meeting, also supported the view of the whistleblower.

es ablehnten, die Patienten, die getestet wurden, zu sehen und beide hatten keine Erfahrung mit BCIs bei ALS-Patienten. Wir zitieren aus dem Bericht des Hinweisgebers

As noted in our responses to the University commission and the DFG committee, the whistleblower did not participate in any data collection and did not familiarize himself with the processing of data collected by near infrared spectroscopy (NIRS). This made it difficult for the accused persons to collaborate with the whistleblower as originally planned, since he and the student he got to work with him refused to see the patients that were tested and neither had experience with BCIs in ALS patients. We cite from the report of the whistleblower to the accused of October 2017, which clearly demonstrates faulty handling of the fNIRS data, which he refused to change after being informed about it: "I performed a reanalysis of the raw data. I applied the Beer-Lambert transform to obtain Oxy/Deoxy-values and bandpass filtered the data from 0.01 Hz to 0.3 Hz. As the raw data contains data for all sensor/emitter combinations, this results in $8 \times 8 = 64$ channels with 2 values each (oxy/deoxy) resulting in a total of 128 values per timepoint." Had the ombudspersons given this to an expert, the faulty analysis of the data by the whistleblower would immediately have been known. This was, however, not done.

- Indicates that the fNIRS-BCI quotes of the PLoSBiol publication are incorrect (see Appendix 1 E). The whistleblower then submitted a commentary manuscript to the PLoSBiol publication, which he had incriminated, for publication in the journal PLoSBiol on 7 November 2017 (see Annex 2A1, A2: No evidence for communication in CLIS? A report about NIRS-based BCI communication in CLIS). On December 13, 2017, the whistleblower received a letter from the editor, Gabriel Gasque, of the journal PLoSBiol with a note/disclaimer and revision suggestions of the comment manuscript submitted by the whistleblower on the PLoSBiol study (see Annex 28). Following the review process by reviewers (one of the reviewers was Prof. Birbaumer), the journal issued a statement on March 12, 2018, regarding the rejection of the whistleblower's commentary manuscript. On 23 March 2018, the whistleblower filed a statement against this decision with the journal PLoSBiol and lodged a complaint about the review process and the bias of a reviewer and editor (see Annex 3). On 27 April 2018 the whistleblower received an e-mail from PLoSBiol with the information that a new reviewer had been consulted regarding his comment and that the comment of the whistleblower would be published after a revision together with a response comment by Professor Birbaumer (see Annex 4).

On August 1, 2018, the whistleblower received a request from PLoSBiol to review Prof. Birbaumer's comment as a reviewer. On 9 August 2018 PLoSBiol provided information that a "major revision" of the response commentary of the Birbaumer group had to take place, i.e. missing data should be made available; according to the whistleblower, the decision-letter sent by PLoSBiol to the whistleblower showed that he was probably the only reviewer of the Birbaumer response commentary. In this context, the whistleblower also points out that the second version of the response commentary of the Birbaumer Group was only slightly changed compared to the first version. Additional data on the still criticized data of the online-database was missing and open-also question sessions. In this regard, the whistleblower also states that the newly added data do not comply with scientific standards and that these results cannot be reproduced either.

As noted in our comments to the University committee both the author of the original article and the author of a comment review each other's statements, in addition to regular reviewers. This is a normal procedure of most journals we know and there is nothing wrong with it. Reasons for rejection of the commentary were that one of the reviewers felt that this scientific issue should be resolved within the research group and the insufficient description of the methods used by the whistleblower, especially with respect to handling of the fNIRS data. Neither the University nor the DFG committee found anything wrong with this procedure. In addition, in the 2019 commentary the data were recomputed by an independent person, Dr. Sudhir Pathak from the University of Pittsburgh, who obtained even better results than the accused.

In this context, the statements of the Birbaumer working group in its response commentary to the PLoSBiol commentary of the whistleblower are also confusing. Chaudhary and Birbaumer present the statement of the whistleblower in his PLoSBiol commentary that in the PLoSBiol publication different channels had been averaged, as a false statement. However, in the

PLoS Biol Publication the authors Chaudhary et al. several times write themselves that the mean was calculated across all channels (e.g. in Table 1 (row D) „number of channels averaged", also on page 18 and 1.-9 of the publication regarding online data analysis: „The mean of the relative change in. O2Hb across the channels was used"). In addition, Birbaumer has stated in his response comment regarding the first PLoSBiol commentary of the whistleblower that he (the whistleblower) in his commentary has not used averages with the words "Averaging signals is necessary and statistical/y justified, even the comparison between averaged signals is statistically correct. (see also Annex 4). Accordingly, the Birbaumer working group on this point itself, in its first Plos-Biol-Response commentary indicates that an averaging is needed but that they state in their second PLoSBiol response commentary, however, that this was wrong and that they never averaged over channels. Even more confusing is the attempt to try to support this point by quoting the PLOS Biol publication where they clearly state that averaging was done across all channels „*The mean (...) across all the channels was used*".

As we noted in our statements to the University and the DFG it is wrong to average across channels as done by the whistle blower. The channels are too far apart for averaging across them. Rather, the results have to be computed per channel and can later be averaged, which is exactly what we did and which was not criticized by any expert. This can also be seen in the commentary where we separately displayed the channels. Again, the lack of training of the whistle blower in this type of data analysis is evident.

In the PLoSBiol Response Comment from Chaudhary/Birbaumer to the PLoSBiol Comment of the whistleblower, the authors present new results of a t-test statistic performed for a patient (Pt. F) for each session and channel separately. In the corresponding Fig. 4 of the commentary, the authors give a significance value of <0.05 . However, the whistleblower's attempt to reproduce these data provided results with no significant differences.

We have noted in several statements to the University and the DFG that t-tests for yes and no responses have to be computed from the time series in the raw data which we provided on the zenodo server but cannot be computed from the data presented in Table S2 in the original publication, which served a different purpose (comparison of fNIRS and EEG data). The whistle blower simply used a wrong set of data for his analyses. His later made allegation that the axes of the figure in the commentary were manipulated, was also refuted.

CONCLUSION of the Ombudspersons

From the point of view of the ombudspersons of the Medical Faculty of the University of Tübingen, the arguments and statements made by the whistleblower appear plausible in the context of the submitted documents regarding the suspicion of scientific misconduct by the authors of the publication PLoSBiol 15(1) e1002593_. For an exact and detailed clarification of the complex facts which led to the accusations of the whistleblower we recommend that the University Commission for Scientific Misconduct obtains appropriate examination and clarification also using subject-specific independent expertise and review of the research protocols as well as the data records on the incriminated PLoSBiol publication.

Independent subject-specific expertise (BCI, fNIRS, ALS) was never obtained. The commission consulted an expert from the institute of the whistleblower who is only experienced in mathematical modelling. There is a potential conflict of interest due to the close proximity to the whistleblower and the fact that the expert was pa member of the university committee before he was called upon as an expert and was therefore not independent. .

Xxxx

(Ombudspersons)

