



Evaluation of the “Freiburg Acuity VEP” on Commercial Equipment

Michael Bach · Jeffrey D. Farmer

Received: 24 May 2019 / Accepted: 17 September 2019 / Published online: 4 October 2019
© The Author(s) 2019

Abstract

Purpose To assess the implementation and outcome quality of the Freiburg Acuity VEP methodology (Bach et al. in Br J Ophthalmol 92:396–403, 2008) on the Diagnosys Espion Profile and E3 electrophysiological systems.

Methods We recorded visual evoked potentials (VEPs) from both eyes of 24 participants, where visual acuity (VA) was either full or reduced with scatter foils to approximately 0.5 and 0.8 LogMAR, resulting in a total of 144 recordings. Behavioral VA was measured in each case under the same conditions using the Freiburg Acuity Test (FrACT); VEP-based acuity was assessed with the “heuristic algorithm,” which automatically selects points for regression to zero amplitude.

Results Behavioral VA ranged from -0.2 to 1.0 LogMAR. The fully automatic heuristic VEP algorithm resulted in 8 of 144 recordings (6%) that were scored as “no result.” The other 136 recordings (94%) had an outcome of -0.20 to 1.3 LogMAR (which corresponds to a range of 20/12.5–20/400, or 6/3.8–6/

120, in Snellen ratios; or 1.6–0.1 in decimal acuity). The heuristic VEP algorithm agreed with the behavioral VA to within ± 0.31 LogMAR (95% limits of agreement), which is equivalent to approximately three lines on a VA chart.

Conclusions The successful implementation of the Freiburg Acuity VEP “heuristic algorithm” on a commercial system makes this capability available to a wider group of users. The limits of agreement of ± 0.31 LogMAR are close to the original implementation at the University of Freiburg and we believe are clinically acceptable for a fully automatic, largely objective assessment of visual acuity.

Keywords Visual acuity · Objective · VEP · Visual evoked potentials · Computer

Introduction

The objective assessment of visual acuity has become increasingly important over the past few years. One way to achieve this is based on visual evoked potentials (for reviews see [1, 2]), often termed “sweep VEP”, “stepwise sweep VEP” or “acuity VEP. Bach and coworkers [1] described a method that combined:

- Brief onset checkerboard presentation, yielding relatively high amplitudes,

M. Bach (✉)
Section for Functional Vision Research, Eye Center,
Medical Center – University of Freiburg, Faculty of
Medicine, University of Freiburg, Killianstr. 5,
79106 Freiburg, Germany
e-mail: michael.bach@uni-freiburg.de

J. D. Farmer
Diagnosys LLC, Lowell, USA

- Temporally in the steady-state region, allowing Fourier transform-based analysis,
- Laplace montage, yielding high noise rejection,
- Application of the Meigen/Bach statistic [3], yielding noise-corrected response and significance,
- An automated “heuristic algorithm” for regression point selection, enabling an acuity estimate, or the outcome of “no result” even when a “notch” [4] is present.

This approach was used in the Freiburg laboratory for a decade with high testability, (i.e., a “no result outcome” occurs in only 5–10% of the cases); problems in amblyopia have been described [5, 6], and the algorithm was successfully extended to very low acuities (≈ 2 LogMAR) [7]. The stimulation and recording system was also used in other laboratories and is available free of charge [8]. However, the platform hardware and software is outdated (e.g., MacOS 9), so a re-implementation was needed to enable a broader set of clinical and research users to operate the method in their clinics. Diagnosys expressed an interest in the method and implemented it following the method previously reported [1], and we herein report the outcome of that implementation.

Methods

Equipment and Stimuli

Steady-state VEPs were recorded using a Diagnosys Espion E3 System (Lowell, MA, USA). Checkerboard stimuli were presented in brief onset mode, two frames (33.3 ms) on at 100 cd/m² and six frames (100 ms) off, corresponding to 7.5 Hz with a stimulus distance of 180 cm and a contrast of 40%. Three sets of check sizes were used, one for the highest VA range (“Range A”), one for medium VA range (“Range B”) and one for the lowest VA range (“Range C”). For Range A, the check sizes were 0.37° to 0.05°, for Range B they were 1.19°–0.17° and for Range C they were 4.0°–0.57°. Six check sizes were used in each Range, and cumulatively across all three Ranges there were twelve unique check sizes (i.e., there is overlap of checks sizes between the Ranges). In this study for the three acuity conditions of each participant, Range A and Range B were always used, and depending on the adjusted acuity of the participant with the strongest

Bangerter foil, either Range B or Range C was used for the lowest acuity recording. The benefit of having three check size ranges available is that in clinical use this approach will usually keep the test time shorter for a patient. Typically, the clinic has a general understanding of the acuity range the patient is likely to fall within thereby enabling it to choose one of the three shorter protocols. In cases where the clinic does not have that knowledge, they can run the full set of twelve check sizes on a patient.

Freiburg Acuity Test (FrACT) [9] measurements were taken on a standard PC with a screen size of 58.5 cm (diagonal) using the same Bangerter foils used for the steady-state VEP recordings, also at a distance of 180 cm from the computer.

Recording

The VEP was recorded using gold cup electrodes at O_z, O₁ and O₂, referenced to F_z. In accordance with the ISCEV VEP standard [10], O_z was placed on the midline at 10% above theinion. O₁/O₂ were placed laterally to O_z at a distance of 10% the ½ head circumference on either side of the O_z electrode. Signals were amplified by a factor of 8, digitized at a rate of 1 kHz with 32-bit resolution and digitally filtered in the range of 5–50 Hz. Averaging was arranged to capture exactly eight on-/offset periods in 1066-ms epochs. forty sweeps were taken for each step within an artifact rejection window of $\pm 100 \mu\text{v}$. For each step, the Laplace transform is calculated from the signals obtained at the O_z, O₁ and O₂ electrode locations ($\text{VEP}_{\text{Laplace}} = 2\text{O}_z - (\text{O}_1 + \text{O}_2)$). The software then calculates the Fourier transform of each resulting signal at each step and plots the resulting six amplitudes by log spatial frequency (dominant of the check size [11]). Finally, the software calculates a visual acuity estimate based on methods described below. A simplified recording setup is depicted in Fig. 1. We define a set of six traces (along the chosen checkerboard set) to be one “recording.”

Participants

There were twenty-four participants in the study (14 male and 10 female), with an age range of 19 to 74 years old (mean age was 46.5), and in each case both eyes were tested. Participants were given the choice of either using their habitual eyeglasses during the test or

not, and that same condition was used for all tests for that participant. Under these conditions the participants' visual acuity ranged from approximately 0.6 to better than -0.15 LogMAR. Each participant completed one set of tests with full vision conditions (defined as participants with their chosen correction, and no Bangerter foil), which was typically 0.30 LogMAR or better. An additional set of tests were then conducted using a Bangerter foil that was intended to reduce the participants' acuity to approximately 0.4

LogMAR, and a final step with a foil intended to reduce acuity to approximately 1.0 LogMAR. Since every participant was recorded under several conditions, one might think that the "eyes or patients" problem might arise [12]. Given that we are using descriptive, not inferential statistics, this is not a problem.

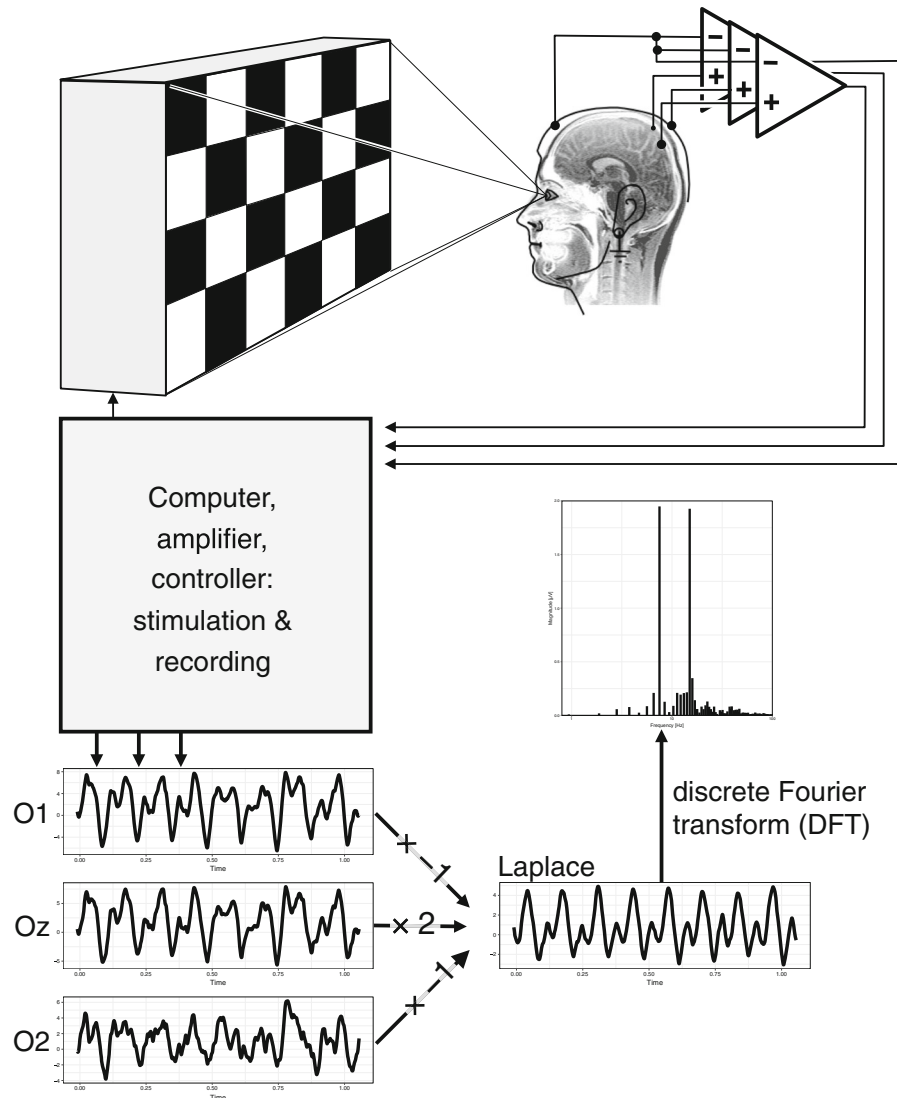


Fig. 1 Basic setup of the stepwise sweep VEP test methodology. The participant is positioned in front of a monitor that presents on-off checkerboard patterns. Signals are measured by the system from three active electrodes on the participant (O_z , O_1 , O_2), referenced to an electrode on the participant's forehead.

As the signals are recorded, the system first completes a Laplace transform on the data ($2O_z - O_1 - O_2$) and then a discrete Fourier transform to determine one VEP magnitude for each set of check sizes. This particular example also exhibits a strong 2nd harmonic

Analysis

Response traces were de-trended and subjected to a discrete Fourier transform (DFT). Because care was taken to choose the analysis interval (1066 ms) to be an integer multiple of the stimulation period, there is no overspill in the spectrum [13] and the noise can be estimated by averaging the magnitudes recorded at the two neighboring frequencies (6.5 and 8.5 Hz). The ‘true’ response magnitude at 7.5 Hz was calculated by non-linearly subtracting the noise from the magnitude measured at 7.5 Hz, and finally a significance for the response at 7.5 Hz was also calculated [3].

Responses are recorded over six check sizes. Ideally, the stimulation of the various check sizes would be interleaved, but this was not yet implemented on the system in this study. The six check sizes were selected from a range of 0.05° to 4.0° , as appropriate for the expected VA. The responses were processed as described above, resulting in 6 values for the response magnitude plus the associated significances. From these, the heuristic algorithm, starting at small check sizes, selects as many points as possible up to peak response and avoiding a notch [4] if present. The resulting points are regressed to zero magnitude on a log(spatial frequency) scale, resulting in the value SF_0 . SF_0 is divided by 17.6 (calibration factor, [1]), yielding a decimal acuity estimate $VA_{dec}(VEP)$. This is converted to LogMAR using the standard formula: $VA_{LogMAR} = -\log_{10}(VA_{dec})$. When insufficient points are found or other irregularities occur, a “no result” outcome is flagged, reducing “testability.”

The relationship between behavioral acuity and the VEP-based acuity estimate is quantified in terms of the Bland-Altman limits of agreement (LoA) [14]. Frequently for such a task the correlation coefficient is computed; that is, however, an inappropriate measure because it is normalized by range [12, 15].

Results

Altogether we recorded 864 traces (24 participants · 2 eyes · 3 VA conditions · 6 check sizes). In Fig. 2, one recording (representing one set of traces across the six selected check sizes) is depicted for the full vision condition (participant id 4, right eye). Its heuristic analysis outcome is seen in Fig. 3.

Across all 144 recordings, the heuristic algorithm reported “success” in 136 cases, and “no result” in 8 cases. This corresponds to a testability of 94%. In Fig. 4, the VEP-based acuity estimate of the 136 success cases is plotted versus their behavioral visual acuity. Since the unit “LogMAR” quantifies visual loss, not visual acuity, the LogMAR scale is inverted, showing good acuity at top right. The Bland-Altman limits of agreement were calculated to be ± 0.31 LogMAR and there was a tendency of the VEP-based acuity method to underestimate acuity in the lower acuity conditions, compared to the behavioral visual acuity measurements. One particular outlier disagrees by 0.5 LogMAR; inspecting its original data (and all others) showed no independent reason to exclude it.

Discussion

Using the Acuity VEP method from [1], we found a high testability (94%) and a reasonably close agreement of behavioral and VEP acuity estimates (95% limits of agreement of ≈ 3 lines). Behavioral test-retest LoA can be as low as 0.1 LogMAR [16, 17] but will be markedly higher in a clinical population. Thus, the LoA of ± 0.31 for the Acuity VEP seems acceptable and is very close to the one reported earlier [1]. The possibility of outliers should, of course, always be considered when analyzing patient data in clinic.

The current implementation of the Acuity VEP method steps through the six check sizes sequentially starting from the largest one, recording 40 sweeps at each check size. Each step takes approximately 40 s, resulting in a total of about 4 min of recording per eye. In principle, this allows the technician to stop adding steps with finer checks when the amplitude drops. However, this should never be done because insufficient data may be recorded thereby preventing a proper regression and the test could miss a notch region (e.g., the amplitude might rise again for smaller checks). In the present data set, two recordings had “no result” for this reason. In practice, the protocol is created in a stepwise fashion to allow the patient a brief time to rest at a few points during the test.

The VEP protocol had been set up with digital filtering in the range of 5 to 50 Hz. In hindsight, this seems unnecessarily narrow. However, this filtering does not affect the results because the heuristic

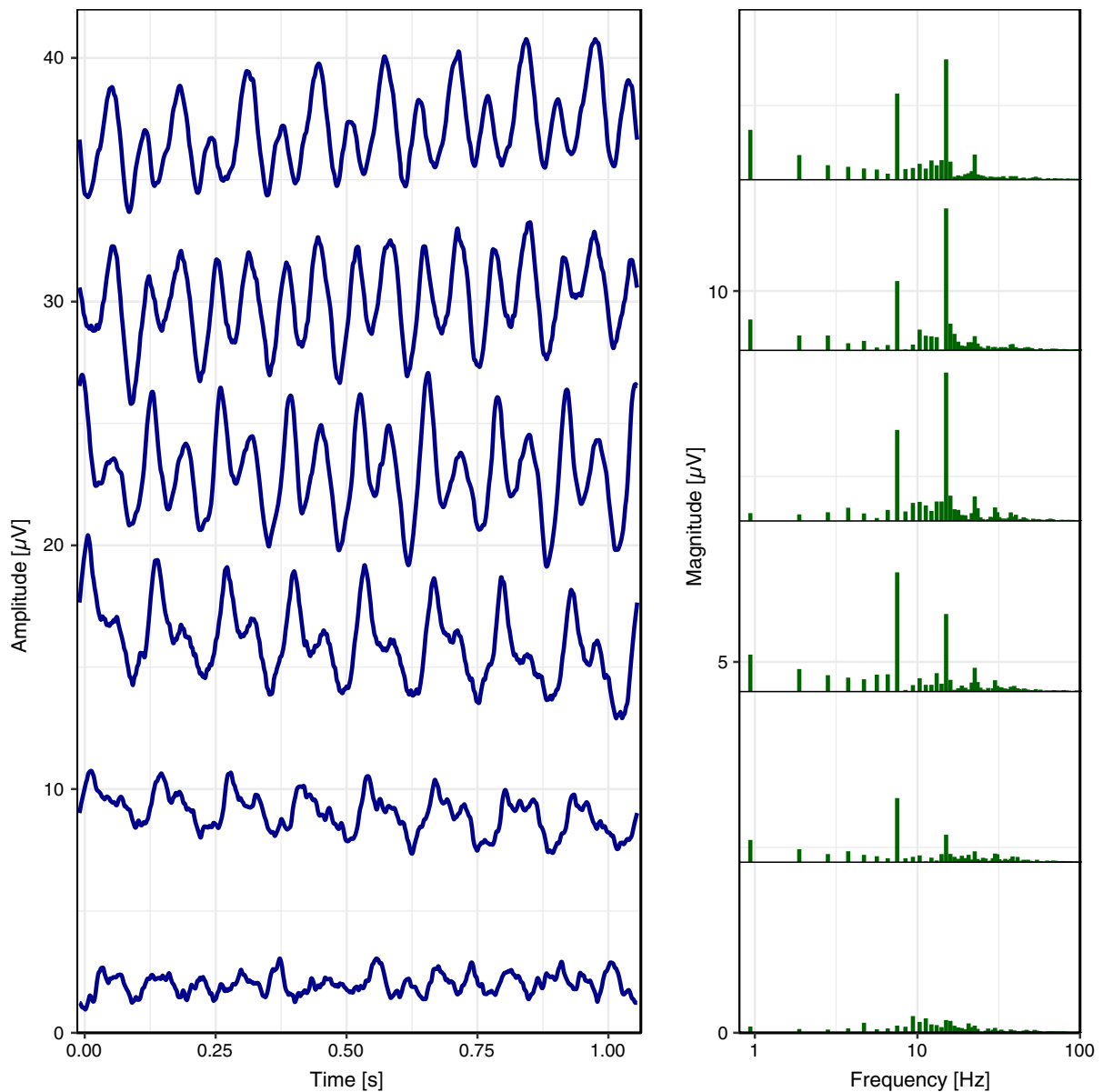


Fig. 2 Left: VEP traces (after calculating the Laplace transform) across a series of check sizes (0.05° bottom, 0.37° top). There are eight responses per sweep, with marked frequency doubling at intermediate check sizes. Right: Magnitude spectra of these traces after discrete Fourier transform (DFT). The first

harmonic response is at 7.5 Hz, a marked second harmonic is also obvious here (rarely as strong). No evidence of overspill is seen in the spectra. The magnitudes at the stimulus frequency (7.5 Hz) and their immediate neighbors form the basis for further analysis (Fig. 3)

algorithm is solely based on the 7.5 Hz spectral line (response) and its immediate neighbors (noise estimators). Even some mains intrusion (at 50 or 60 Hz, depending on locality) would not have a detrimental effect. In the future, the Acuity VEP protocol will use standard VEP filtering settings.

We have not addressed test-retest agreement here, since this is implicitly covered by the analysis of agreement between the behavioral visual acuity and the VEP acuity outcome. In future work, it may be of interest to specifically assess the test-retest agreement to analyze the relative variance contribution from interindividual vs. intraindividual sources.

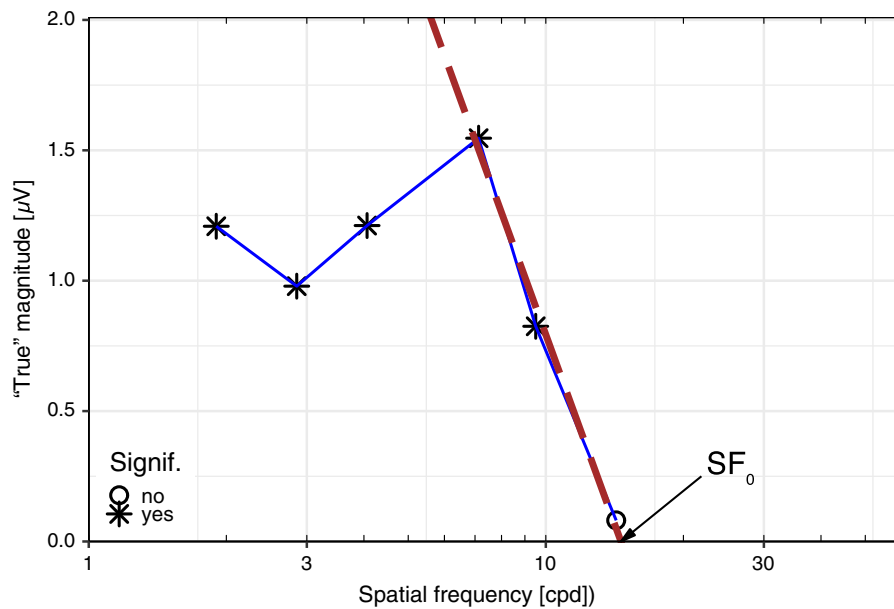


Fig. 3 Check size tuning curve based on Fig. 2 data. The abscissa depicts the dominant spatial frequency of the pertinent checkerboards. The ordinate depicts the spectral magnitude at 7.5 Hz after noise removal. The signal-to-noise ratio enables the calculation of individual significance, here indicated by stars versus circles. The dashed line represents the regression of the

three points selected by the heuristic algorithm and its extrapolation to zero magnitude. results in SF_0 , the spatial frequency for zero magnitude. In this example, SF_0 is 14.6 cpd, corresponding to a decimal acuity of $14.6/17.6 = 0.83$ or 0.081 LogMAR

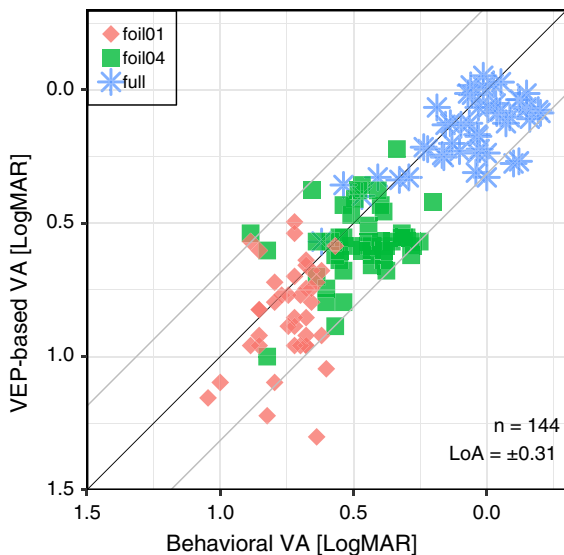


Fig. 4 Relation of the VEP-based visual acuity estimate (ordinate) and behavioral visual acuity (abscissa) using inverted LogMAR scales (good acuity top right). Blue stars indicate acuity in the “full vision” condition. Acuity was artificially reduced with foils to (nominally) decimal 0.4 (green squares) and decimal 0.1 (red diamonds). The Bland–Altman limits of agreement (LoA, grey lines parallel to the identity line) were calculated to be ± 0.31

The Freiburg Acuity VEP has previously been found to be of substantial aid in the management of patients with non-organic visual loss. The present study shows that the method has been implemented effectively in a commercial system, enabling its use in a broader set of clinical sites; the method is also being validated for pediatric applications. We would also welcome critical third party assessments without our own conflicts of interest. Finally, if the machine learning approach lives up to its promise [18], it can be applied to the Freiburg Acuity VEP method potentially extending its value to researchers and clinicians.

Acknowledgements We thank our participants for their kind cooperation.

Author contribution MB and JDF designed the study, JDF took the recordings, MB analysed the data and composed figures, both authors contributed equally to the manuscript.

Compliance with ethical standards

Conflict of interest Michael Bach served as consultant for Diagnosys. Jeff Farmer is CEO of Diagnosys LLC.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the Medical Center, Faculty of Medicine, University of Freiburg, Germany and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards [19].

Statement of human rights All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Statement on the welfare of animals This article does not contain any studies with animals performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Bach M, Maurer JP, Wolf ME (2008) Visual evoked potential-based acuity assessment in normal vision, artificially degraded vision, and in patients. *Br J Ophthalmol* 92:396–403. <https://doi.org/10.1136/bjo.2007.130245>
- Almoqbel F, Leat SJ, Irving E (2008) The technique, validity and clinical use of the sweep VEP. *Ophthalmic Physiol Opt* 28:393–403. <https://doi.org/10.1111/j.1475-1313.2008.00591.x>
- Meigen T, Bach M (1999) On the statistical significance of electrophysiological steady-state responses. *Doc Ophthalmol* 98:207–232
- Bach M, Joost W (1989) VEP vs spatial frequency at high contrast: subjects have either a bimodal or single-peaked response function. In: Kulikowski J, Dickinson C, Murray I (eds) *Seeing contour and colour*. Pergamon Press, Oxford, pp 478–484
- Wenner Y, Heinrich SP, Beisse C et al (2014) Visual evoked potential-based acuity assessment: overestimation in amblyopia. *Doc Ophthalmol* 128:191–200. <https://doi.org/10.1007/s10633-014-9432-3>
- Heinrich SP, Bock CM, Bach M (2016) Imitating the effect of amblyopia on VEP-based acuity estimates. *Doc Ophthalmol* 133:183–187. <https://doi.org/10.1007/s10633-016-9565-7>
- Hoffmann MB, Brands J, Behrens-Baumann W, Bach M (2017) VEP-based acuity assessment in low vision. *Doc Ophthalmol* 135:209–218. <https://doi.org/10.1007/s10633-017-9613-y>
- Bach M (2007) Freiburg evoked potentials. <http://www.michaelbach.de/ep2000.html>. Accessed 19 Aug 2013
- Bach M (1996) The Freiburg visual acuity test – automatic measurement of visual acuity. *Optom Vis Sci* 73:49–53
- Odom JV, Bach M, Brigell M, et al (2016) ISCEV standard for clinical visual evoked potentials: (2016 update). *Doc Ophthalmol* 133:1–9. <https://doi.org/10.1007/s10633-016-9553-y>
- Fahle M, Bach M (2006) Origin of the visual evoked potentials. In: Heckenlively J, Arden G (eds) *Principles and practice of clinical electrophysiology of vision*. MIT Press, Cambridge, pp 207–234
- Holopigian K, Bach M (2010) A primer on common statistical errors in clinical ophthalmology. *Doc Ophthalmol* 121:215–222. <https://doi.org/10.1007/s10633-010-9249-7>
- Bach M, Meigen T (1999) Do's and don'ts in Fourier analysis of steady-state potentials. *Doc Ophthalmol* 99:69–82
- Bland JM, Altman DG (1999) Measuring agreement in method comparison studies. *Stat Methods Med Res* 8:135–160. <https://doi.org/10.1177/096228029900800204>
- McAlinden C, Khadka J, Pesudovs K (2011) Statistical methods for conducting agreement (comparison of clinical tests) and precision (repeatability or reproducibility) studies in optometry and ophthalmology. *Ophthalmic Physiol Opt* 31:330–338. <https://doi.org/10.1111/j.1475-1313.2011.00851.x>
- Arditi A (2005) Improving the design of the letter contrast sensitivity test. *Invest Ophthalmol Vis Sci* 46:2225–2229. <https://doi.org/10.1167/iovs.04-1198>
- Bach M, Schäfer K (2016) Visual acuity testing: feedback affects neither outcome nor reproducibility, but leaves participants happier. *PLoS ONE* 11:e0147803. <https://doi.org/10.1371/journal.pone.0147803>
- Bach M, Heinrich SP (2019) Acuity VEP: improved with machine learning (submitted to DOOP)
- World Medical Association (2000) Declaration of Helsinki: ethical principles for medical research involving human subjects. *J Am Med Assoc* 284:3043–3045

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.