Event-Related Potentials Allow for Optotype-Based Objective Acuity Estimation

Sven P. Heinrich, Isabel Lüth, and Michael Bach

Visual acuity is routinely measured in countless patients every day. However, some cases are challenging due to the inability or unwillingness of the patient to reliably perform a psychophysical acuity test. There are examination techniques that may help reveal a lack of patient cooperation, some of which are easily available to the ophthalmologist or optometrist,\(^1\) while other techniques require specialized versions of psychophysical acuity tests.\(^5\) It would be clearly desirable to be able to rely on an objective physiological method of estimating acuity. Such a method would not only be useful in cases of suspected malingering, but also when the patient lacks the competence required for the psychophysical test.

Event-related potentials (ERPs) are an important tool in the investigation of information-processing ability in the absence of explicit behavior.\(^8\) Visual evoked potential recordings, often implemented in the form of sweep VEPs, can demonstrate that signals from the eyes reach the cortex and have successfully been used to estimate visual acuities.\(^9\)–\(^17\) However, they do not provide evidence for proper visual processing at higher stages of the cortical pathways\(^18\) and have even been reported to be increased in psychogenic visual loss in a group of children and young adults.\(^19\) In such a situation, ERP responses related to higher-level processing could be a decisive tool. In a few cases, the P500 of the event-related potential has been used to assess stimulus processing in cases of suspected functional loss or malingering.\(^21\)–\(^23\) We have recently demonstrated the feasibility of a cognitive acuity test with grating stimuli based on the P300 component of the event-related potential.\(^24\)

The P300 is regarded as a marker of high-level cognitive processing.\(^25\) It consists of several subcomponents, including the P3a (novelty P300) and the most widely known, P3b.\(^26\) The present study used the latter subcomponent, which is usually largest at parietal and central electrode locations with a peak time of approximately 300 to 600 ms after stimulus onset.\(^27\)\(^,\)\(^28\) Most commonly, it is recorded with a two-stimulus oddball sequence, where a series of frequent stimuli are interspersed with infrequent stimuli.\(^29\)–\(^31\) Subjective probability of the infrequent stimuli is a main determinant of the P300 amplitude.\(^32\) Other factors include objective probability\(^33\) and possibly environmental rareness.\(^34\) Interestingly, at least with cooperative subjects, randomness of the stimulus sequence is not an important factor.\(^35\) The functional significance of the P3b is not fully understood. The processes suggested to contribute to the P3b include the update of the working memory to reflect prevailing stimulus conditions, inhibition of ongoing activity that would facilitate information transmission, and postdecisional inhibition marking a closure of a stage of processing.\(^36\)
As the P300 is linked to conscious stimulus perception on a trial-by-trial basis, \(^ {57,58}\) P300-based acuity estimates should be closely linked to subjective acuity. In the present study, we provide a detailed quantitative account, going beyond the preliminary assessment in our previous study. \(^ {24}\) We specifically assess the feasibility of using optotypes, rather than gratings or checkerboard patterns, as stimuli. So far, methods for VEP/ERP-based objective acuity estimation have been based on gratings or checkerboard stimuli. Subjective acuity tests, on the other hand, typically use optotypes, such as the Landolt C, which is defined as the standard optotype by the International Organization for Standardization. \(^ {39}\) Acuity as measured with such optotypes represents so-called recognition acuity, \(^ {40}\) and it is known from psychophysical studies that optotype acuity and grating acuity differ in healthy subjects. \(^ {41,42}\) They are furthermore differentially affected by certain types of disorders of the visual system, such as amblyopia. \(^ {43}\) Such differences may also underlie discrepancies between VEP-based acuity estimates, which rely on gratings or checkerboards, and standard acuity measurements, which are performed with optotypes. For instance, Wenner et al. \(^ {44}\) have found that acuity estimates obtained with checkerboard VEPs are typically higher than those obtained psychophysically with optotypes. Because VEP recordings rely on the massive electrophysiologic response of primary visual cortex to gratings or similar stimuli, they are not suitable to be used with Landolt C optotypes, where the critical detail of the stimulus is rather subtle, namely a small gap in a ring. The P300 response of the ERP, however, is much less dependent on the low-level stimulus properties. This opens up the possibility for using optotypes for both objective and subjective acuity estimates, making a direct comparison possible.

**Methods**

Two experiments were performed, one with optotype stimuli in six different sizes and one with grating stimuli with six degrees of coarseness. In each experiment, the P300 was recorded from 12 subjects at different levels of dioptric blur. Psychophysical estimates of acuity were obtained with both optotypes and gratings in all subjects, irrespective of the type of ERP stimulus used. Because four subjects participated in both experiments, this resulted in 20 sets of psychophysical acuity data.

**Participants**

In each of the two experiments, 12 subjects participated after providing written informed consent. Among these were four subjects that participated in both experiments. All participants had no known disorders of the visual or neural system. The study was approved by the local review board and followed the tenets of the Declaration of Helsinki.

All tests and recordings were performed with one eye only. Subjects wore a trial frame with their habitual correction in addition to the degrading lenses. If subjects did not reach a decimal acuity of at least 1.0 with their habitual correction, they were refracted, and their habitual correction was replaced accordingly, which in all cases resulted in the acuity criterion being met.

**Stimuli and Procedure for P300 Recordings**

Stimuli were presented on a Philips GD403 CRT monitor at a distance of 3.6 m for 493 ms with an interstimulus interval of 493 ms. In one experiment, Landolt C optotypes in six different sizes with geometrically spaced gap widths of 0.78 to 22.1 min arc were used as infrequent stimuli among closed annuli as frequent stimuli (Fig. 1). Optotypes were presented in black on a 6.4\(^{\circ}\) \times 6.4\(^{\circ}\) white background. Weber contrast of the optotypes was 100%. In the other experiment, stimuli consisted of sine wave gratings of six geometrically spaced spatial frequencies in the range of 38.5 cycles per degree (cyc/deg) down to 7.35 cyc/deg, corresponding to stripe widths of 0.78 to 4.8 min arc. The gratings served as infrequent stimuli among frequent homogenous gray stimuli that had the same average luminance of 75 cd/m\(^2\) as the gratings (which was higher than the 43 cd/m\(^2\) of the dark gray background that filled the stimulus area). Michelson contrast of the gratings was chosen to be 40%, following the technical and physiological considerations by Bach et al. \(^ {15}\) The total grating size was 5.4\(^{\circ}\) \times 5.4\(^{\circ}\), surrounded by adumbrated crosshairs that remained on the screen during the interstimulus intervals to help subjects locate the center for fixation. For both Landolt C and grating stimuli, the screen outside the stimulus area was black.

Pilot experiments had shown that different ranges of the size of the stimulus's critical detail (Landolt C gap size or stripe width, respectively) would be necessary to cover the expected performance range of the subjects.

The rationale behind the stimulation scheme was as follows. Depending on visual acuity, infrequent stimuli below a certain Landolt C gap size or grating coarseness will appear identical to the corresponding frequent stimuli (closed annuli of the same size or homogenous gray fields). They will thus not elicit a P300 response. The threshold in size or coarseness between infrequent stimuli that elicit a P300 and those that do not may serve as estimate of visual acuity.

Typically, a total of approximately 400 to 500 stimuli were presented per condition, depending on the number of artifacts. Around one sixth of these were of the infrequent type, which is in the range that can be expected to yield a good signal-to-noise ratio. \(^ {45}\) Stimulus sizes were mixed within each block. In case of the Landolt C experiment, the frequent stimuli (closed circles) also occurred in all different sizes in order to avoid any bias toward a certain size. Infrequent stimuli, both gratings and optotypes, came in two orientations at 45\( ^{\circ}\) relative to vertical. The subjects were instructed to silently count one of the two orientations.

The acuity of the participants was degraded using lenses of +1, +2, and +4 diopters (D) in separate recording runs. The order of dioptric values was randomly altered between subjects. Two measurement runs were performed for each dioptric value, only separated by a set of subjective acuity tests that were also performed with the respective dioptric value and encompassed two test runs with Landolt Cs and two with gratings as described below.
ERP Recording and Analysis

The electroencephalogram (EEG) was recorded with a 32-channel EEG system with actiCAP active electrodes (Brain Products, Gilching, Germany). The FCz site was used as reference during recording, and a 70-Hz low pass was applied. Artifacts were rejected based on a 100-mV threshold criterion.

The data were analyzed offline using interactive software (Igor Pro; Wavemetrics, Lake Oswego, OR, USA). After rereferencing to the average of TP9 and TP10, trials were sorted according to stimulus type, pooled across recording runs with the same dioptic defocus, low-passed at 25 Hz, and averaged. Among the trials with infrequent stimuli, no distinction was made between the target and nontarget orientations.

The response difference between frequent and infrequent stimuli was computed to extract the oddball response, and the detailed analysis of the P300 focused on the Pz electrode. For each combination of defocus (three values) and stimulus size (six values), P300 magnitude was estimated from the maximum ERP amplitude in the 200- to 700-ms interval of the response difference. For each value of defocus, we fitted a sigmoid function (specifically, we used a logistic function) to describe the P300 amplitude as a function of the size of the critical detail. The threshold (size value at the inflection point) was taken as an estimate of the resolution limit. To facilitate reliable fitting even in situations of low signal-to-noise ratio, constraints were applied to the fit parameters, which prescribed the general orientation of the sigmoid function and prevented the function from falling below zero.

To circumvent the problem that in some subjects the maximum P300 amplitude was not reached with large values of defocus even with the largest optotype or coarsest grating, we first fitted the sigmoid to the data obtained with the lowest defocus. We then took the maximum and minimum values of this fit as fixed values for the fits to the higher-defocus data, leaving only threshold and slope variable.

Measurements of Subjective Acuity

Subjective acuity of the patients was measured using a psychometric test modeled after the Freiburg Visual Acuity Test and implemented in Igor Pro. Each participant completed two runs with Landolt C optotypes and two runs with sinusoidal gratings. Stimuli were presented on the same monitor at the same distance as the ERP stimuli. The same screen at the same distance as in the ERP experiment was used. Landolt C optotypes could have eight possible orientations, while gratings were oblique at either of the two 45° orientations. Contrast was identical to that in the P300 recordings. The optotype size or the stripe width of the grating, respectively, was determined through an adaptive staircase procedure.

While the ERPs for a given subject were recorded only to either grating or optotype stimuli, subjective acuity tests were performed with both types of stimuli in all subjects. The data of the two measurement runs with each stimulus type were pooled and submitted to a maximum likelihood fit, which determined the parameters of the corresponding cumulative Gaussian psychometric function. The participant’s acuity was computed from the threshold of this function (i.e., the inflection point). All further analysis was done on the logarithm of the threshold values. The two values obtained from repeat testing with the same stimulus type were averaged before further processing.
The fits were performed with the trust-region Levenberg-Marquardt least orthogonal distance method to account for the fact that there are measurement errors not only in the P300 data, but also in the psychophysical data. Note that the confidence intervals may be somewhat underestimated as they do not account for the partial correlation that may exist between the three data points belonging to a single subject.

In order to facilitate easy conversion between P300-based thresholds and psychophysical thresholds through a simple multiplicative factor (additive in the case of logarithmic data), we also fitted a line with the slope fixed to one to the logarithmic data. This resulted in ordinate intercepts of 0.36 (Landolt C; 95% CI, ±0.09) and 0.18 (grating; 95% CI, ±0.04) for the logarithmic data, corresponding to conversion factors of 2.3 (95% CI, 1.8...2.8) and 1.5 (95% CI, 1.4...1.7), respectively. Using these conversion factors, and assuming a normal distribution (the number of data points was too small to pursue a distribution-free approach), we computed approximate prediction intervals. In the case of Landolt C stimuli, in 95% of the cases the psychophysical acuity estimates are within 0.54 log units of the converted P300-based estimate. For gratings, this value is 0.26 log units.

**DISCUSSION**

The present study demonstrated that single-optotype acuity can be estimated objectively using ERPs. This is not possible...
with VEP recordings, which are not reliably elicited by subtle alteration of single-optotype stimuli.

In order to appraise the usability of the approach, we need to consider its relationship to psychophysical acuity estimates. The present psychophysical thresholds for optotype stimuli are consistent with previous studies that related dioptric blur to acuity. Subjects performed better with gratings than with optotypes, especially with strong blur. The most likely reason for this is spurious resolution, which enables the subjects to still correctly guess the orientation of a grating with a certain stripe width when the Landolt C with a gap size equal to the stripe width cannot be recognized anymore. It should be noted that the higher guessing probability with only two alternatives in the case of the grating orientation task is accounted for by using the psychophysical standard definition of threshold (i.e., the stimulus value for which the percentage of correct responses is in the middle between random guessing and 100%, barring lapses, is taken as threshold). In addition, signal detection theory predicts a further effect of the number of response alternative, unrelated to random guessing. This effect is far smaller than the difference between grating and optotype stimuli in the present study.

There is a close association between P300-based acuity estimates and the respective psychophysical estimates that were obtained with the same type of stimulus, which allows for deducing psychophysical acuity from P300-based acuity.
estimates when a conversion factor (an additive constant in case of logarithmic values) is applied. This suggests that the P300-based acuity estimates are governed by the same phenomena, namely spurious resolutions, as the psychophysical estimates. This precludes a straightforward conversion between, for instance, P300-based grating acuity estimates and psychophysical optotype acuity in the case of dioptric blur.

P300 thresholds were consistently higher than psychophysical thresholds. This is not surprising. The generation of the P300 relies on the percept that is generated immediately after the onset of the stimulus. Psychophysical thresholds were obtained with standard procedures that allow for prolonged inspection of the stimulus. At least for Landolt C stimuli, it is known that such a difference can have a sizable effect on the test outcome. Furthermore, the present oddball paradigm was based on the Landolt C gap or the grating stripes being present in infrequent trials and being absent in frequent trials, as opposed to the situation in a psychophysical acuity test in which the gap is always present, but the stimulus orientation has to be determined. Resembling a detection task, the P500-based acuity may thus be reduced. It should be noted that we took a very different approach to estimating thresholds from ERP amplitudes than that taken by, for instance, Norcia et al. and Bach et al. in studies on sweep-VEP-based acuity estimates. Simply speaking, in those studies the largest stimulus coarseness that would just fail to evoke a response was estimated through extrapolation. Fitting a sigmoid function to P300 amplitudes, as in the present study, is more analogous to fitting a psychometric function to behavioral data and only possible because the P300 amplitude does not depend much on physical stimulus parameters and moderate stimulus degradation as long as frequent and infrequent stimuli are easily distinguishable. This is similar to subjects showing nearly 100% performance in a psychophysical task as long as the stimuli are not too close to threshold. Near threshold, as with the binary outcome options of correct versus false in psychophysical tasks (barring randomly correct guesses), the P300 has been suggested to represent an all-or-nothing response. This means that intermediate amplitude values are the result of averaging full-amplitude responses with absent responses. There is thus considerable evidence that the P300 approach has a much more natural link to psychophysical acuity estimation than the VEP-based approach.

Spurious resolution is likely to be less relevant when visual impairments other than defocus are the underlying cause of reduced acuity, as the effect depends on the shape of the point-spread function. Importantly, discrepancies between grating acuity and optotype acuity do not occur only because of...
spurious resolution; the literature is inconsistent as to the magnitude and even as to the direction of the difference.42,56–59 One factor playing a role is the type of visual impairment. For instance, as mentioned above, grading acuity tends to be less affected by amblyopia than optotype acuity when estimated psychophysically or through visual evoked potentials.44 There is, furthermore, preliminary evidence for such a discrepancy being present in patients with central visual field loss and eccentric fixation, although the respective study compared a detection task with gratings to a forced-choice task with optotypes.60 In these groups of visual disorders, obtaining optotype-based objective acuity estimates, as piloted in the present study, might be particularly beneficial as one can expect a better agreement with standard subjective measures of acuity.

The prediction intervals appear relatively large. This is especially true for the Landolt C stimuli with a range of ±0.54 log units. Taking into account that the total range of measured values across blur levels was much larger with optotypes than with gratings, the relative size of the respective prediction intervals is not much different, and the prediction interval obtained for gratings (±0.26 log units) matches that reported by Bach et al.15 for VEP-based acuity estimates. One factor adversely affecting the prediction interval is the relatively low number of trials per condition, resulting from the large number of conditions that were to be compared. This reduced the signal-to-noise ratio.

In a clinical application, where only one type of stimuli and no artificial visual degradation would be tested, the number of trials per condition could be increased. Substantial further improvement may also be achieved by incorporating patient-specific advance knowledge as to the acuity range of interest. This is a particularly relevant factor because the amplitudes at only two or three size or width values determine the precise threshold value, while the other stimuli are too far away from threshold to contribute. However, it might be advantageous to include one stimulus size that is far above threshold as a control for task compliance and attention. When estimating the recording time in a clinical setting, it is important to consider whether the examination needs to be performed with both eyes or only with one eye. Clearly, comparing data from both eyes may be revealing in some cases. However, in many cases either the eye of interest is known (e.g., in amblyopia) or the eye is most likely irrelevant (e.g., in higher-level cortical disorders or psychogenic impairment of vision). In these cases, one could justifying performing the test for only one eye. Considering these aspects, the proposed technique for optotype-based objective acuity estimation would not typically be the first choice for a screening test. The use of individual prior knowledge, however, can ensure efficient testing in patients where other approaches may not yield reliable results.

In summary, the present study is, to our knowledge, the first to demonstrate the feasibility of obtaining objective acuity estimates with optotype stimuli. P300-based estimates of acuity show the same discrepancy between optotype and grating stimuli as do the psychophysical measurements, providing a perspective for meaningful objective acuity estimation in visual disorders that differentially affect grating and optotype acuity.

Acknowledgments

Supported by the Deutsche Forschungsgemeinschaft, Bonn, Germany (BA 877/18).

Disclosure: S.P. Heinrich, None; I. Lüth, None; M. Bach, None

References


54. Bonala B, Boutros NN, Jansen BH. Target probability affects the likelihood that a P300 will be generated in response to a target stimulus, but not its amplitude. Psychophysiology. 2008;45:93–99.


