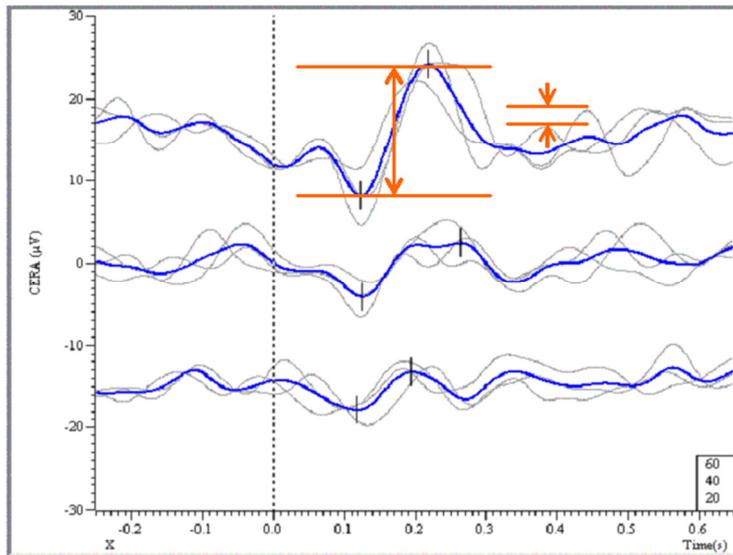


## Objective detection of the N1-P2 response

CAEP is often referred to as "objective audiometry", but CAEP analysis is usually subjective and accuracy depends on the skill of the operator. Machine scoring of waveforms, ideally resulting in a statistical measure of response confidence, is the ideal and has obvious attractions when the test is used in medico-legal cases where a claimant's compensation is based on the CAEP results or there is potential for dispute

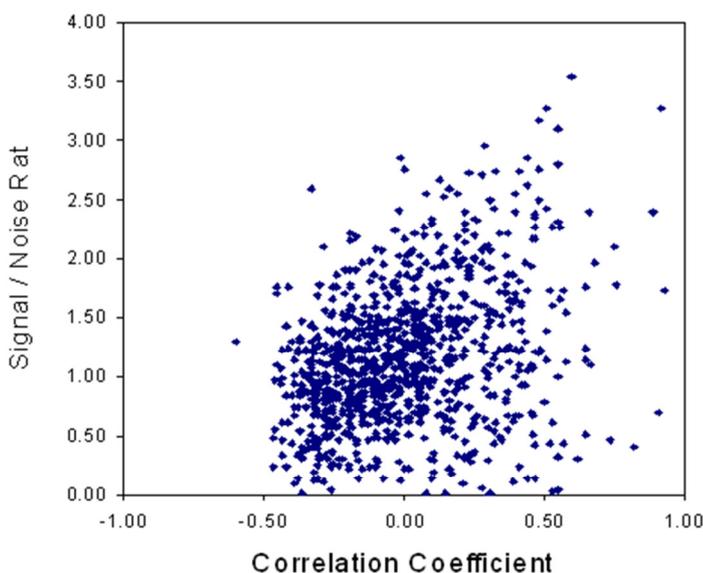


of CAEP test interpretation. We have recently extended our system to do this. A candidate N1-P2 response is automatically identified using a simple cursor placing algorithm: N1 is defined as the most negative point of the grand average in the latency range (from stimulus onset) 50ms to 250ms and P2 as the most positive point in the range N1 to 400ms. The user may move the cursors if not correctly placed by the algorithm. The N1-P2 amplitude is calculated (the "signal"). Noise is calculated as the point by point difference between a pair of sub-averages, averaged across the entire recording epoch (we use -250ms to +650ms). Since our system uses three sub-averages there are three possible combinations of sub-averages so the

three noise figures are averaged. We are therefore able to measure the signal to noise ratio of the response, SNR (this is not true SNR since the signal is peak-peak and the noise is RMS).

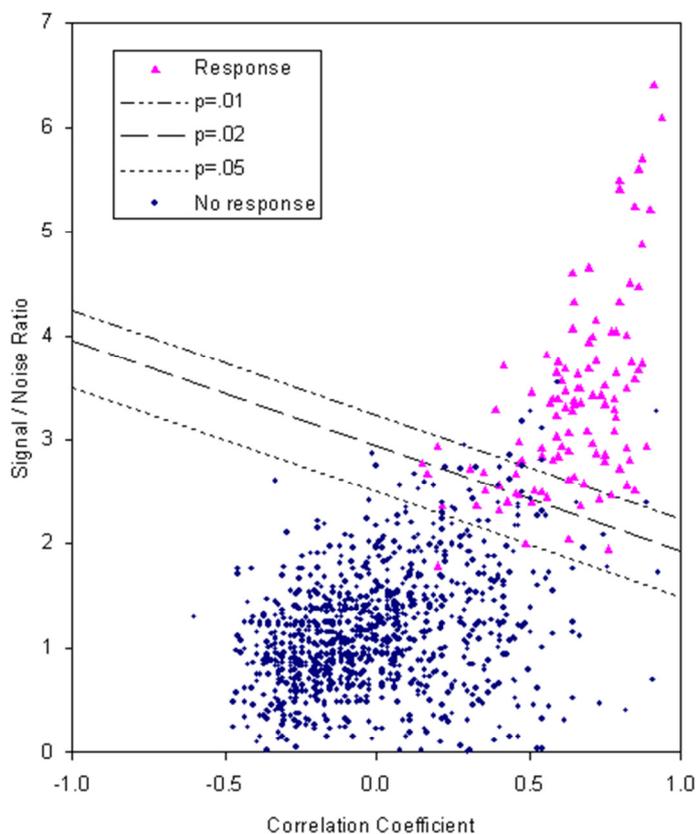
In order to use the SNR to calculate the chance that the identified "response" is simply random noise containing no genuine electrophysiological response (in other words to create a p-value) one must know the number of degrees of freedom in the recording then use F-tables. An alternative is to establish the

distribution of SNR values in a no-response population. This is the option we took, recording 1000 averages from volunteers tested without a stimulus and noting the resulting SNR. We also recorded the correlation coefficient (CC) of the sub-averages in the region around the potential response as identified by the algorithm.



The relationship between SNR and CC in this no-stimulus population shows very little correlation and there may (yet to be established!) be an advantage in using both SNR and CC in the calculation of the p-value. This makes sense: both SNR and CC will be high when there is a clear response and small (and random) when there is no response.

A waveform's p-value is calculated (from its SNR, CC or a combined variable) by calculating the proportion of no-stimulus cases whose value is equal to or greater than that measured from the patient. This method is attractive since it makes no assumptions about the shape of the reference distribution and is derived from real data using the same test paradigm and parameters.



This figure is as above but also shows the SNR and CC values for genuine N1-P2 responses. At high test levels one would record values in the top-right of the figure and as the stimulus level is reduced towards threshold the values approach, and are lost in, the area of uncertainty populated by the no-response population.

We have decided to use the simple combined variable (SNR + CC) when calculating p-values. The  $p=0.01$ ,  $p=0.02$  and  $p=0.05$  lines for this variable are shown in the figure. They are approximately orthogonal to the trajectory of real responses as test level is reduced, confirming that SNR + CC is not unreasonable as a parameter to separate response from no-response cases.

Our CAEP system now computes and displays the p-value upon completion of each average together with SNR, CC, N1-P2 amplitude etc. We have found the availability of p-values very helpful in the clinic, and feel that our clinical practice is now improved. In particular, p-values can identify circumstances where further averaging is needed to resolve "possible" responses. We have adopted  $p < 0.02$  as our criterion for response acceptance. Further work is needed to validate the method against conventional scoring.

This work was presented at the XXI Biennial Symposium of the IERASG in Brazil, June 2009.

### Accuracy of the CAEP threshold estimate in adults

Previous studies have used conventional stimulus presentation and data acquisition / manipulation methods. We know that our method is a good deal faster than conventional methods, mainly because we automate most of the predictable manual tasks. What we needed to demonstrate was the accuracy of the threshold estimate. Of course, this has been done before but not using our random pseudo-binaural stimulus and not at high frequencies. We employed 24 volunteers (mostly hospital staff) whose pure tone audiogram (PTA) was recorded by experimenter 1. Their CAEP was then conducted by experimenter 2, blind to the PTA results. Test frequencies of 1, 3 & 8 kHz (balanced order) were chosen because most hearing disability schemes use the frequencies of 1, 2 & 3 kHz. Conventional wisdom suggests that the CAEP amplitude is lower at high frequencies so we included 8 kHz to test this. Though not used in disability calculations, high frequencies are often helpful in matters of causation - demonstrating an audiometric notch associated with noise trauma.

**Results:** The mean error in the N1-P2 threshold estimate was 6.5 dB, with no significant effect of frequency. After correcting for this bias, 94% of individual threshold estimates were within 15 dB of the

behavioural threshold and 80% were within 10 dB. Establishing the 6 threshold estimates (3 frequencies, 2 ears) took on average 20.6 minutes.

### **Effectiveness of certain stimulus presentation features in increasing the N1-P2 amplitude**

We developed our "Optimised" CAEP test paradigm from the findings of the available literature (for details, see the page on this). However, we have taken much of this on trust and we certainly did not know whether there is any interaction between the effects of the parameters we have chosen. This study therefore addressed this issue by looking at them in isolation and in combination. Again, 24 volunteer staff are being used but only one ear is under scrutiny, at one frequency (3 kHz), at an intensity close to threshold (25 dB sensation level). In this study we hoped to identify any effect on CAEP amplitude of:

- varying the inter-stimulus interval of a monaurally presented stimulus
- inserting a 10s stimulus-free interval half-way through the averaging process to allow an adapted response to recover
- presenting the stimuli to one or other ear in a random fashion (at equal sensation level)

**Results:** There appeared to be no effect of any of the above on N1-P2 response amplitude. These findings were somewhat of a surprise and disappointing - apparently suggesting that these novel stimulus presentation features that we have developed and used for many years actually bring no advantage. Still, that's the point of undertaking the research!

In fact, the results must be viewed with an important fact in mind: the nature of the experimental design was such that subjects were exposed to ever-changing stimuli over a period of about 20 minutes. Thus, there appears to be no significant short-term effect of these features. What we have not addressed in our study is whether these features offer any advantage over conventional stimulation - i.e. monotonous monaural stimulation lasting up to an hour. We suspect we would see an advantage but that's another study!

We have produced a paper on this study, Lightfoot & Kennedy (2006); refer to the CAEP Literature page for details.

### **The effect of caffeine on the accuracy of the cortical response**

This study has looked at the effect of caffeine on the N1-P2 response and results were presented at the XXIII Biennial Symposium of the IERASG, New Orleans, USA, June 2013.

The study was a double-blind placebo controlled crossover design. We wanted to investigate whether the size of the response (and therefore the accuracy of the threshold estimate) could be enhanced by the use of caffeine since the literature suggests that the subject's general arousal level was a factor.

**Results:** In our sample and using a dose of 175 mg of caffeine there was no statistically significant increase in CAEP amplitude.