TITLE: Vertex evoked potentials in a rating-scale detection task: relation to signal probability.¹

AUTHORS: Kenneth C. Squires, Nancy K. Squires, and Steven A. Hillyard
Department of Neurosciences, University of California, San Diego.

RUNNING HEAD: Threshold detection P5 and signal probability.

SEND PROOFS TO: Dr. Kenneth C. Squires
Department of Psychology
University of California, San Diego
La Jolla, California 92037

After July 10, 1974:
Department of Psychology
University of Illinois
Champaign, Illinois 61820
ABSTRACT

Vertex evoked potentials were recorded from human subjects performing in an auditory detection task with rating-scale responses. Three values of \textit{a priori} probability of signal presentation were tested. The amplitudes of the N1 and P3 components of the vertex potential associated with correct detections of the signal were found to be systematically related to the strictness of the response criterion and independent of variations in \textit{a priori} signal probability. No similar evoked potential components were found associated with signal-absent judgements (misses and correct rejections) regardless of the confidence level of the judgement or signal probability. These results strongly support the contention that the form of the vertex evoked response is closely correlated with the subject's psychophysical decision regarding the presence or absence of a threshold-level signal. The implications of these results for a general hypothesis for describing the sensitivity of the P3 component to psychological variables are discussed.

KEY WORDS: vertex potential, P3 component, N1 component, response criterion, Signal Detection Theory, signal probability, thresholds.
In the influential report of Sutton, Braren, Zubin, and John (1965), a late-positive (P3) component was shown to be introduced into the vertex evoked potential when a stimulus conveyed task-relevant feedback information (confirming or disconfirming a prior guess). Sutton et al. further demonstrated that the amplitude of P3 increased as the occurrence of the relevant stimulus became less probable. Many subsequent studies have verified that the amplitude of the P3 evoked by one of a set of clearly discriminable, task-relevant stimuli increases as a function of increasing improbability, unexpectedness or unpredictability (Tueting, Sutton, and Zubin, 1971; Squires, Hillyard, and Lindsay, 1973a; Donchin, Kubovy, Kutas, Johnson, and Herning, 1973; Wilkinson and Ashby, 1973; Friedman, Hakerem, Sutton, and Fleiss, 1973; Squires, Squires, and Hillyard, 1974). Furthermore, it has been firmly established that a P3 component can also be elicited by infrequent omissions of an expected stimulus from a repetitive sequence (Klinke, Fruhstorfer, and Finkenzeller, 1968; Barlow, 1969; Ruchkin and Sutton, 1973; Picton, Hillyard, and Galambos, 1974; Picton and Hillyard, 1974), thus suggesting that the relationship between P3 amplitude and event probability is similar whether the task-relevant event is stimulus presence or stimulus absence.

These conclusions are derived from situations where the stimuli were clearly recognizable and differentiable from one another. It appears that somewhat different rules apply when the alternative task-relevant stimuli are ambiguous, such as the presence or absence of a threshold-level auditory signal. In that case the P3 component is reportedly associated only with signal-present decisions (HITs), and its amplitude depends primarily upon the confidence level of those decisions (Hillyard, Squires, Bauer, and Lindsay, 1971; Paul and Sutton, 1972; Squires, Hillyard, and
Lindsay, 1973b). Hillyard et al. (1971) suggested that the amplitude of the P3 component elicited by signals that are difficult to detect or discriminate is governed by the interaction of two factors: P3 increases with increasing decision confidence but decreases with greater expectancy that the signal will occur. In a subsequent elaboration (Squires, Hillyard, and Lindsay, 1973a), it was postulated that internal neural models or "templates" were established for the purpose of recognizing each of the relevant stimulus alternatives; stimulus recognition then engenders a P3 wave that increases with the confidence of the recognition (i.e. the closeness of the "template match") but is reduced by the subject's prior expectation of that stimulus. This two factor hypothesis, however, does not readily explain the well-documented absence of a P3 component with correct signal-absent decisions (correct rejections) in the threshold detection paradigm (Hillyard, 1969; Hillyard et al., 1971; Paul and Sutton, 1972), particularly when those correct rejections are made with a high degree of confidence (Squires et al., 1973b) or are very improbable (Sutton and Paul, 1973). This absence of a P3 component for correct rejections in a threshold situation is especially puzzling in light of the aforementioned reports that P3 waves do accompany task-relevant omissions of suprathreshold signals (c.f. Sutton, Tucting, Zubin, and John, 1967).

The present study was designed to determine how the two main factors, confidence level of the decision and the probability of the decision, interact to determine the P3 amplitude for both signal-present and signal-absent decisions, with the aim of accounting for the relation of these types of decisions to the P3 component. The frequencies of occurrence of the various types of decisions were manipulated by varying the a priori probability of signal presentation in a threshold signal-
detection paradigm with decision confidence assessed on an eight-point rating scale. In particular, we wished to determine if the frequency of occurrence of the two kinds of decisions, over and above variations in the subject's response criterion, was a major determinant of P3 amplitude; if so, we anticipated that when signal absence was made an extremely rare event a P3 component might come to be associated with signal-absent decisions of high confidence. Thus, for all types of decision, the present study assessed the separate effects of decision confidence and decision probability in an attempt to arrive at the general principles that govern P3 amplitude in the signal-detection paradigm.
METHODS

Subjects

Four young adults with normal hearing who had previous experience in similar experiments served as subjects in a series of 6-10 two-hour experimental sessions over a period of two to three weeks. Two of the subjects (KS and NS) were experimenters.

Procedure

During testing the subject sat in a reclining chair in an acoustic chamber wearing TDH-39 earphones and fixating on a small neon bulb on the panel before him. His task was to decide on each trial whether or not a binaural 1000 Hz sinusoidal signal of 50 msec duration was presented against a background of wide-band white noise and to rate his confidence in that decision. The binaural noise background was continuously present at a level of 65 dB SPL. A signal intensity close to detection threshold (defined as 75% correct with signal probability of 0.5) was chosen for each subject and was used throughout the experiment.

Each trial began with a 200 msec flash of the neon bulb, which served as a warning signal. On the "signal-present" trials the offset of the warning signal was followed after 500 msec by the tonal signal; on the "signal-absent" trials no signal was presented. No additional stimulus served to mark the observation interval within which the signal might occur. The signal-present and signal-absent trials occurred randomly, but with a predetermined probability. The neon bulb was relit two sec after the warning light, thereby directing the subject to respond by pressing one of the eight numbered buttons on the panel before him. When highly confident that a signal had been presented the subject was instructed to press button number 1; ratings 2 and 3 indicated decreasing confidence that a
signal had been presented, and a "4" indicated a marginally confident
decision that a signal had been presented. Similarly, a rating of 5
indicated marginal confidence that the trial had been a signal-absent
trial, and ratings 6, 7, and 8 indicated increasing confidence that there
had been no signal presented during the observation interval. Immediately
after each button press the response light was turned off and one of two
remaining lights was illuminated for 750 msec, providing feedback as to
whether or not a signal had been presented on that trial. Inter-trial
times were randomized between four and six sec.

Each subject was given sufficient practice in the task to
stabilize his distribution of confidence ratings before data collection
began. Trials were presented in blocks of 75, with five or six blocks
per testing session.

Three values of a priori probability of signal presentation,
0.2, 0.5, and 0.8, were used in counter-balanced order. The subject was
informed of the signal probability prior to each block of trials.

**Evoked potential recording**

Evoked potentials were recorded from the vertex referred to
the right mastoid using Ag-AgCl electrodes (Beckman, non-polarizable)
and amplified with Grass model 7 polygraph amplifiers (bandpass flat from
0.5 to 120 Hz). Evoked potentials were sampled over an epoch of 500 msec
beginning at the onset of the observation interval (500 msec after the
offset of the warning light). Averaged waveforms were computed separately
for each of the sixteen stimulus-response outcomes, determined by the two
stimulus conditions (signal-present and signal-absent) and the eight confi-
dence-rating response categories. The vertical electro-oculogram was also
averaged concurrently with the evoked potentials to ensure the absence of eye-movement artifacts.

Stimulus timing, signal selection and on-line evoked-response averaging were under the control of a PDP-9 computer.
RESULTS

Psychophysical judgements

Since all of the subjects were experienced with rating-scale judgements, their response distributions rapidly stabilized and remained consistent across blocks of trials and across testing sessions. The average response distributions are shown in Table 1, where the frequency of occurrence of each signal condition and confidence-rating response is tabulated for the three levels of \textit{a priori} signal probability. Also presented is the percentage of correct choices for each confidence-rating category. Small numerical ratings signifying highly confident "signal-present" responses were associated with a high percentage of trials in which a signal was presented and, hence, a high percent correct. The percent correct diminished with decreasing confidence in the decision to a minimum for the mid-ratings and increased again for the higher numerical ratings which signified highly confident signal-absent decisions. Finally, the seven criterion cutoffs that define the eight separate confidence-rating categories are shown (see Green and Swets, 1966). The criterion cutoffs were derived from the response distribution data and are expressed in standard-deviation units above or below the mean of the assumed distribution of events resulting from signal-absent trials ($z_n$), since that distribution can reasonably be assumed to remain constant across variations in signal probability. It is evident that there was a systematic shift to a set of stricter criteria for making signal-present decisions as the \textit{a priori} probability of signal presentation decreased. For example, the value of $z_n$ for a rating of 1 increased from 2.15 to 2.23 to 2.60 as the signal probability decreased from 0.8 to 0.5 to 0.2. There was, however, no change in the overall detectability of the signal as the \textit{a priori} probability of its presentation was varied, as can be seen from the
receiver operating characteristic (ROC) curves shown in Figure 1. The mean values of the detectability measure, \( d_s \), were 1.04, 1.12, and 1.12 for signal probabilities of 0.2, 0.5, and 0.8 respectively. The parameter \( d_s \) was chosen as the appropriate detectability measure because of the evident asymmetry of the ROC curves which suggest that the variance of the signal-present and signal-absent response distributions were not equal (see Green and Swets, 1966, Chapter 4). In agreement with Schulman and Greenberg (1970), the slopes of the ROC curves plotted on normal-probability axes were found to systematically decrease with decreasing signal probability (slopes equalled 0.76, 0.68, and 0.57 for signal probabilities of 0.8, 0.5, and 0.2, respectively) consistent with an increase in the variability of the signal-present response distribution as signal presentation became less frequent.

Evoked potentials

The set of sixteen evoked-potential waveforms for one subject at a signal probability of 0.5 is shown in Figure 2. As shown previously (Squires et al., 1973b), the highest confidence HIT (a rating of 1, signal present) is characterized by a large negative component with a peak latency of about 165 msec (N1) followed by a large positive peak with a latency of about 330 msec (P3). For progressively less confident signal detections these components diminished in amplitude and increased in latency until
they became indiscernible at about rating-level 4. The evoked potential waveforms for the other two values of a priori signal probability (0.2 and 0.8) were similar in form to those in Figure 2, but differed in amplitude as shown in Figure 3. In Figure 3 the amplitude of P3 (expressed as the percent of the maximum P3 amplitude for each subject) is plotted as a function of confidence rating for both signal-present and signal-absent trials at the three values of signal probability. Component amplitudes were measured baseline to peak, with the baseline defined as the average voltage over the first 60 msec of the recording epoch. For signal-absent decisions (ratings 5-8, MISSES and CRs) no evoked components similar to the large N1 and P3 associated with HITs were observed regardless of the confidence of the decision or the a priori signal probability. In these cases the P3 amplitude was measured from the largest peak between 300-400 msec post-stimulus. The average amplitude of the waveforms at this latency for high-confidence MISSES and CRs was only 25% of that for high confidence HITs and was not affected by the signal or outcome probability. This result also held for an area measure of the waveforms which was determined as a check on the possibility that the P3 component for non-HIT trials might be poorly time locked to the averaging epoch, resulting in a small peak amplitude while encompassing a substantial positive area. The area function calculated for the interval between 250 and 450 msec post-stimulus onset, referred to the 60 msec baseline, was essentially the same as for the amplitude measure shown in Figure 3.

In Figure 3 the P3 amplitude decreases with decreasing confidence rating for HITs at all three signal probabilities. A given numerical rating
at a low signal probability, however, has a higher criterion cutoff (as defined by the \(z_n\) value) than the same rating with a high signal probability. The orderly relationship of P3 amplitude on HITs to criterion cutoff is shown for two subjects in Figure 4, where the evoked potentials are ordered by the criterion cutoff value regardless of the a priori signal probability and confidence rating. In Figure 5 the mean amplitudes over all subjects of P3 and N1 for HITs are plotted as a function of the mean value of the criterion cutoff at each numerical rating for all values of signal probability. Since it was not possible in some instances to make an accurate assessment of the N1 component amplitude for the lower confidence HITs, those data are not included. There is a clear decrease in the size of the P3 component with decreasing strictness of the criterion cutoff (correlations between P3 amplitude and \(z_n\) for individual subjects ranged from 0.55 to 0.86 with a mean of 0.70, \(p<0.001\)). Likewise, the amplitude of the N1 component can be seen to decrease with decreasing criterion cutoff (correlations for individual subjects ranged from 0.37 to 0.65 with a mean of 0.56, \(p<0.001\)). Most significantly, however, Figure 5 demonstrates no systematic influence of signal or outcome probability on the amplitude of P3 over and above that due to variations in criterion level at the different probabilities; in other words, P3 amplitudes as a function of criterion for all of the three probabilities fall along the same line.

In Figure 6 the amplitude of the vertex P3 accompanying the two highest confidence levels of HITs is plotted as a function of the mean
frequency of occurrence of those events at the different signal probabilities. There was a slight (non-significant) negative correlation between P3 amplitude and the relative frequency of a given event (dashed lines) which could be attributed to the variations in criterion cutoff. A positive correlation held, however, between the P3 amplitude and event probability for any fixed level of signal probability (solid lines). The P3 associated with the higher confidence decision was larger even though that event was more frequent, supporting the idea that P3 amplitude is determined by the criterion cutoff rather than the frequency of occurrence of the stimulus-response event.

---------------------------
INSERT FIGURE 6 HERE
---------------------------
DISCUSSION

In agreement with previous studies (Hillyard et al., 1971; Paul and Sutton, 1972; Squires et al., 1973b), the amplitude of the P3 component of the auditory evoked potential associated with correct detections of threshold-level signals (HITs) was found to be systematically related to the strictness of the response criterion. A highly confident, high-criterion HIT was associated with a large P3 component of relatively short latency, and for decreasingly confident detections the amplitude of that component decreased while its latency increased.

Seemingly at variance with the predictions of our previous proposals (Hillyard et al., 1971; Squires et al., 1973a, b), however, varying the a priori probability of signal presentation had no additional influence on the amplitude of the P3 component for HITs over and above that determined by the variations in criterion level. According to those previous formulations the amplitude of P3 was presumed to be directly related to the decision confidence and inversely related to the probability of making such a decision. Accordingly, it was expected that if decision confidence was held constant the a priori signal probability would determine the probabilities of decision outcomes and, in turn, the P3 amplitude. Using the objectively determined criterion cutoff as the measure of decision confidence, it was expected that the functions relating P3 amplitude to confidence rating would describe three separate curves corresponding to the three levels of signal probability. In fact, when plotted in this way, the three P3 amplitude versus criterion functions appear to lie along a single curve (Figure 5). Thus, while the P3 amplitude associated with a high-confidence HIT did decrease as the signal probability increased, this effect can be entirely accounted for by the shift to a less strict
criterion cutoff for that confidence rating and not by the variation in signal and decision probability.

In two previous studies (Tueting et al., 1971; Squires et al., 1973a), it has been demonstrated that for unambiguous feedback stimuli, which should be unaffected by perceptual factors such as decision confidence, there is a strong negative correlation between the amplitude of P3 and the a priori probability of stimulus occurrence. These results, along with those of Karlin and Martz (1973) showing a negative correlation between P3 amplitude and the probability of a signalled response, suggest that the amplitude of the P3 component elicited by readily discernible stimuli is largely determined by the probability of the task-relevant event (Tueting et al., 1971). The results of this study, however, indicate that the opposite relation holds at the two highest criterion level HITs at a fixed level of a priori signal probability: P3 amplitude and event probability are positively correlated under these circumstances. Thus event probability does not influence the P3 component elicited by ambiguous, threshold-level signals, which lie along a perceptual continuum, in the same way that it does for distinctive, suprathreshold events. For threshold-level signals the confidence factor evidently outweighs the event-probability factor, possibly because the multi-category rating events are not perceptually distinctive enough for the development of separate expectancies for each event.

The N1 amplitude versus response-criterion function was also largely uninfluenced by the a priori probability of signal presentation. This was to be expected if, as suggested previously (Squires et al., 1973b), the amplitude of N1 reflects the effective intensity of the stimulus. Since the identical signal intensity was used in all three probability conditions and yielded equal measures of detectability in all cases, the
trial-to-trial variations in stimulus effectiveness should be distributed equivalently, and variations in N1 amplitude should only reflect the differing selection of response criteria. Although N1 and P3 covary in the present study, their dissociability is evident even with threshold-level signals since N1 may be present when no decision is required of the subject, while P3 is not (Squires et al., 1973b).

Unlike for correct-detection trials (HITs), no evoked response components were found associated with correct rejections in any of the experimental conditions, thus verifying previous reports (Hillyard et al., 1971; Paul and Sutton, 1972; Squires et al., 1973b). In previous studies, however, the evoked potentials accompanying the correct rejections have received only a limited analysis. The results of this study, where both waveform amplitudes and areas were measured to compensate for variability in time locking, indicate that there is no variation in late positivity for the highest confidence CR over a wide range of response criteria ($z_n = -0.17$ to $z_n = -1.46$) corresponding to a frequency of occurrence ranging from 36% to 2% and a range of percent correct for that rating from 92% to 62%. Over a similar range of response criteria there is a profound change in the amplitude of the P3 component associated with HITs.

Within the theoretical framework of Squires et al. (1973a) there seem to be three possible explanations for this puzzling result. First, signal-absent decisions may never be made as confidently as signal-present decisions, particularly when the signal is near threshold and is embedded in noise. The manipulation of increasing the probability of signal presence, which resulted in a shift to a stricter criterion for high-confidence signal-absent decisions (1.5 standard deviations), also
produced a decrease in the percent correct for those decisions from 92% to 60% due to the increased probability of a MISS. Since the feedback signals made the subjects aware of their relatively low percent correct, there is some doubt as to the confidence with which these decisions were made, the numerical rating and criterion cutoff notwithstanding. In view of this, it seems unlikely that the probability manipulation, which was designed to increase the subject's confidence in signal-absent decisions to a level comparable with that for HITs, had the desired effect. In any case, there was no indication that the P3 amplitude for signal-absent decisions covaried with either the criterion cutoff (unlike the signal-present decisions) or with percent correct, over a wide range of values.

Secondly, the signal-absent event may have been consistently highly expected, regardless of the objective signal probability, since a clear signal-present decision was rare in all three experimental conditions. Accordingly, it may be impossible to produce a rare and unexpected stimulus omission using threshold-level signals. However, when stimuli are above threshold and signal-presence and signal-absence are distinctive events, the P3 appears to vary in a similar manner with the probability of occurrence for both (Ruchkin and Sutton, 1974; Squires, Squires, and Hillyard, in preparation).

Finally, it is possible that the subjects adopted a strategy whereby auditory information was evaluated only with respect to an internal template for the signal and that a P3 is associated only with an affirmative decision. One of the purposes of the probability manipulation was to induce the subject to modify such a strategy and to analyze inputs with reference to a template for signal absence, thus
reversing the standard association of P3s only with correct detections. If however, stimulus absence was an indistinct and highly expected event under all signal probabilities, the stimulus template and decision strategy would not be expected to change with the objective stimulus probabilities.

The results of this study confirm that the form of the evoked response associated with decisions in the threshold-detect paradigm are closely correlated with the subject's psychophysical response. The amplitude of the P3 and N1 components for HITs were directly related to the confidence level of the decisions, as measured by the objective criterion cutoff, over a wide range of probabilities of signal presentation. The precise relationship between the amplitude of P3 and response criterion reinforces the position of Sutton and colleagues (Donchin and Sutton, 1970; Paul and Sutton, 1973) in their continuing debate with Clark, Butler and Rosner (1969, 1970) on "the psychological significance of evoked potentials." We must emphasize the necessity for monitoring the subject's decision criterion and collecting evoked potentials according to finely graded categories of perceptual events if meaningful correlations of evoked potentials and psychophysical processes are to be obtained.

Furthermore, the P3 amplitude for a given criterion was found to be independent of the signal probability, the probability of making a particular decision, and the percentage correct (Sutton and Paul, 1973). There was no evidence that a P3 component was associated with any decisions of signal absence. While these results may be interpreted in line with our previous proposals for describing the behavior of the P3
component it is evident that the relationship predicted between P3 and signal or decision probability is complex and depends upon the discriminability of the signal alternatives.
REFERENCES


Results, and Evaluations," Washington, D. C., NASA.


FOOTNOTE

1. This research was supported by NASA grant NGR-05-009-35 to R. Galambos, by NIH grant NS 07454 to D. Norman and by a Sloan postdoctoral fellowship to N. Squires. Address reprint requests to: Dr. Kenneth C. Squires, Department of Psychology, University of Illinois, Champaign, Illinois 61820.
TABLE 1: Frequencies of occurrence of psychophysical responses at each a priori signal probability, plus the percent correct and criterion cutoff for each rating (mean of four subjects).

<table>
<thead>
<tr>
<th>RATING CATEGORY</th>
<th>( P(S) )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIGNAL</td>
<td></td>
<td>.047</td>
<td>.018</td>
<td>.015</td>
<td>.015</td>
<td>.020</td>
<td>.014</td>
<td>.020</td>
<td>.034</td>
</tr>
<tr>
<td>NOISE</td>
<td></td>
<td>.005</td>
<td>.012</td>
<td>.027</td>
<td>.042</td>
<td>.077</td>
<td>.122</td>
<td>.174</td>
<td>.358</td>
</tr>
<tr>
<td>%C</td>
<td></td>
<td>.91</td>
<td>.63</td>
<td>.37</td>
<td>.29</td>
<td>.61</td>
<td>.89</td>
<td>.90</td>
<td>.92</td>
</tr>
<tr>
<td>( z_n )</td>
<td></td>
<td>2.60</td>
<td>2.08</td>
<td>1.71</td>
<td>1.30</td>
<td>0.88</td>
<td>0.41</td>
<td>-0.17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIGNAL</td>
<td></td>
<td>.122</td>
<td>.055</td>
<td>.068</td>
<td>.075</td>
<td>.039</td>
<td>.041</td>
<td>.032</td>
<td>.042</td>
</tr>
<tr>
<td>NOISE</td>
<td></td>
<td>.008</td>
<td>.014</td>
<td>.033</td>
<td>.074</td>
<td>.079</td>
<td>.095</td>
<td>.100</td>
<td>.123</td>
</tr>
<tr>
<td>%C</td>
<td></td>
<td>.94</td>
<td>.78</td>
<td>.69</td>
<td>.50</td>
<td>.67</td>
<td>.69</td>
<td>.77</td>
<td>.77</td>
</tr>
<tr>
<td>( z_n )</td>
<td></td>
<td>2.23</td>
<td>1.78</td>
<td>1.30</td>
<td>0.70</td>
<td>0.27</td>
<td>-0.20</td>
<td>-0.74</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIGNAL</td>
<td></td>
<td>.252</td>
<td>.119</td>
<td>.118</td>
<td>.153</td>
<td>.056</td>
<td>.035</td>
<td>.028</td>
<td>.017</td>
</tr>
<tr>
<td>NOISE</td>
<td></td>
<td>.005</td>
<td>.014</td>
<td>.027</td>
<td>.067</td>
<td>.029</td>
<td>.031</td>
<td>.026</td>
<td>.023</td>
</tr>
<tr>
<td>%C</td>
<td></td>
<td>.98</td>
<td>.93</td>
<td>.82</td>
<td>.70</td>
<td>.55</td>
<td>.51</td>
<td>.47</td>
<td>.60</td>
</tr>
<tr>
<td>( z_n )</td>
<td></td>
<td>2.15</td>
<td>1.38</td>
<td>0.84</td>
<td>-0.03</td>
<td>-0.42</td>
<td>-0.90</td>
<td>-1.46</td>
<td></td>
</tr>
</tbody>
</table>
FIGURE LEGENDS

Figure 1: Receiver operating characteristic (ROC) curves for the four subjects at the three values of a priori signal probability. Axes are the probability of a signal-present decision when a signal was presented, $P(\text{HIT})$, and the probability of a signal-present decision when a signal was not presented, $P(\text{FA})$.

Figure 2: The evoked-potential waveforms for subject KS at an a priori signal probability of 0.5, averaged according to the sixteen combinations of two signal conditions (signal present and signal absent) and eight confidence-rating categories. Left column, signal present, and right column, signal absent. Confidence ratings 1 to 8 from top to bottom range from highly confident decisions that a signal was presented to highly confident that a signal was not presented. Numbers beside traces indicate number of trials in each averaged waveform. The waveform for the highest confidence FALSE ALARM was omitted due to an insufficient number of trials.

Figure 3: The average amplitude of the P3 component as a function of the confidence rating for signal-present and signal-absent trials at the three values of a priori signal probability. The amplitudes are normalized for each subject according to the maximum amplitude of P3 for that subject in all experimental conditions. All amplitudes are taken baseline to peak where the baseline is the average of the voltage over the first 60 msec of the recording epoch.
Figure 4: The HIT evoked potential waveforms of two subjects (NS and KS) for all three values of \textit{a priori} signal probability ordered according to the objective criterion cutoff. The criterion cutoff ($z_n$) corresponding to each waveform is listed as well as the confidence rating and \textit{a priori} signal probability.

Figure 5: Average amplitudes of the P3 and N1 components for HITs as a function of the criterion cutoff ($z_n$) for the three values of \textit{a priori} signal probability. Amplitudes calculated as in Figure 3.

Figure 6: Average amplitudes of the P3 components for the two highest confidence level HITs as a function of the frequency of occurrence of such decisions for the three values of \textit{a priori} signal probability. Amplitudes calculated as in Figure 3.