INTELLIGENCE **16**, 361–381 (1992)

Brain Evoked Potentials and Intelligence: The Hendrickson Paradigm

P.T. BARRETT
H.J. EYSENCK
Institute of Psychiatry, Bethlem Royal Hospital

An attempt to replicate the Hendrickson (Hendrickson & Hendrickson, 1980, 1982) paradigm and results was undertaken on 40 subjects. The Hendrickson "rules for replication were also empirically examined with regard to their basis in fact rather than presumption. In addition to computing the specific Hendrickson measures of string, variability, and composite, averaged evoked potential integrated waveform amplitude and component latencies were computed. Results indicated a replication of the variability parameter and possibly component latency relationships only. The string correlations were significantly reversed in direction. The composite measure was conceptually and statistically nonsignificant. The Hendrickson paradigm was viewed as no more than a well-controlled auditory evoked potential paradigm. The nine rules suggested by Hendrickson and Hendrickson were shown to be largely irrelevant to obtaining evoked potential correlates with IQ.

INTRODUCTION

The relationship between averaged evoked potentials (AEPs) and psychometric intelligence has been explored since the mid 1960s, beginning with the work of Ertl reported in Chalke and Ertl (1965). Short visual AEP component peak latencies were found to correlate with high IQ in a group of 48 subjects. This work was replicated and extended further by Ertl and Schafer (1969), Ertl (1971, 1973), Bennett (1968), and Shucard and Horn (1972) among others. The average correlation found across these various studies was about −.30. Two recent reviews of the area of EEG correlates and IQ have comprehensively detailed the various studies, methodologies, and results found to date (Barrett & Eysenck, 1991; Deary & Caryl, in press).

This article reports the results from an attempted replication of what has been called the "Hendrickson paradigm." Based upon a novel model of synaptic structure, function, and nerve transmission, Hendrickson and Hendrickson (1980, 1982) derived two measures that could be extracted from an AEP. A complexity measure (otherwise known as the *string* measure) was assessed by computing the contour perimeter of the AEP waveform; the larger this value, the higher an individual's IQ. The second measure, the *variance*, was computed by

Correspondence and requests for reprints should be sent to P.T. Barrett, Institute of Psychiatry, Biosignal Lab, Bethlem Royal Hospital, Monks Orchard Road, Beckingham, Kent BR3 3BX, England.

taking the average variability of each sample point on an AEP over a number of epochs. The greater the variance, the lower an individual's IQ. Essentially they proposed that the neural transmission characteristics of high-IQ individuals is such that fewer propagation–transmission errors are made than is the case with low-IQ individuals. Consequently, within the AEP, high-IQ individuals will tend to have more complex AEPs, the individual component traces being less variable from trial to trial, thus preserving more of the detail of the single evoked potential response. In contrast, the low-IQ individuals will produce a more varied evoked response from trial to trial, yielding (when averaged) a smoother, less complex AEP. Thus, the high-IQ AEP should yield a longer string measure than the low-IQ AEP, and the variability measure should yield a lower value than that for low-IQ AEPs.

The empirical evidence regarding the complexity and/or variance of the AEP is drawn from two studies. The first, reported by Blinkhorn and Hendrickson (1982), correlated the complexity (string) measure with performance on Raven's Advanced Progressive Matrices (APM) and a variety of verbal ability tests. Specifically, data were collected from 17 male and 16 female students. Auditory AEPs were generated, the stimulus consisting of a 1000-Hz sine wave tone, of 30-ms duration, switched at the zero crossing point and presented binaurally through headphones at 85 dB (SPL). The interstimulus interval was quasi-random between 1 and 8 s. The recording derivation was bipolar with montage C_z–A_1 (10-20 System) and epoch time was 512 ms with a 0.5-ms sample speed (1,024 sample points for each of the 90 epochs). Various correlations between the string measure (computed from AEPs generated over 90, 64, and 32 epochs) and the APM yielded a midrange correlation of approximately .45. The verbal test scores did not correlate significantly with the string measure. However, because the range of the scores on the APM was restricted, Blinkhorn and Hendrickson corrected this value for a full range of IQ. The correlation was thus boosted to a maximum of .84. This value is reasonably close to one obtained by Hendrickson and Hendrickson (1980) in an analysis of some published data of Ertl's for which they obtained a correlation of .77 between WISC-IQs and the string scores from the AEPs.

In the second study (Hendrickson & Hendrickson, 1982) a sample of 219 schoolchildren (121 boys, 98 girls) was used. The WAIS was used to assess IQ, scores being generated for the 11 separate subscales, Performance, Verbal, and Full-Scale IQ. In addition to the complexity and variance measures defined before, D.E. Hendrickson defined a new composite measure. This measure was given simply as the variance score minus the string score. The stimulus presentation, data acquisition, epoch length, EEG derivation, and montage were the same as those reported in the first study, effectively the paradigm methodology. The correlations among the WAIS-IQ and string, variance, and composite AEP measures were .72, −.72, and −.83, respectively. The correlations between the WAIS performance total and the string, variance, and composite measures were

.53, −.53, and .60, respectively. The correlations between the WAIS-Verbal total and the string, variance, and composite measures were .68, −.69 and −.78, respectively.

Hendrickson and Hendrickson (1982, p. 210) presented nine points summarizing what should be done in order to replicate their results. These points or rules form the basis for the Hendrickson paradigm.

1. *Stimulus Choice:* The stimulus must be constant from trial to trial with replicable signal characteristics of amplitude, duration, and frequency, that is, no switching transients and no click stimuli.
2. *Stimulus Presentation:* The interstimulus interval should be varied on a pseudorandom basis to prevent habituation effects. The same random sequence of intervals should be used for all subjects.
3. *Electrodes:* Hendrickson and Hendrickson recommended Ag-AgCl electrodes.
4. *Amplifier:* They recommend that investigators use their amplifier, the circuit of which was given in Appendix N. This amplifier had no special upper frequency filtering and no notch filters. However, Hendrickson and Hendrickson failed to provide any bandwidth information. It would be reasonable to assume that its bandwidth extended up to at least 10 kHz.
5. *Calibration Signal:* A constant calibration signal must be fed into each data epoch for each subject. Hendrickson and Hendrickson assumed that the amplification used within a study would change from second to second, at a magnitude sufficient to invalidate any data so recorded.
6. *Recording Medium:* They recommended online digitization of data in contrast to storing analogue voltages on magnetic tape (due to tape stretch, variations in tape speed playback, etc.).
7. *A/D Conversion Sampling Rate:* 1000 Hz minimum.
8. *Epoch of Analysis Period:* 250 ms. They stated that longer epochs would invalidate the assumptions underlying their measures.
9. *Editing of Individual Records–Epochs:* The final number of records obtained for the purposes of averaging must be constant for each subject. The first few records in each session should be rejected for each subject as they tend to show muscle artifact.

The purpose of this article is to attempt to replicate the Hendrickson and Hendrickson results following as closely as possible their nine-point "rules." However, because Rules 4 and 5 are just assertions based upon incomplete specifications and opinion, we chose to use professional laboratory amplification with known (and published) physical standards and properties. We also ignored Rule 2. Because no explanation is offered by Hendrickson and Hendrickson for using the same random sequence for each subject, we decided to use a conventional approach of randomizing stimulus presentation uniquely for each subject.

Finally, we chose to generate a more comprehensive series of measures from the AEPS such that we could test some of the consequences of not following the nine points.

METHOD

Subjects

A total of 26 male and 14 female subjects took part in the experiment. They were volunteers from the local government unemployment bureau and the Institute of Psychiatry. The age range of the men was from 19 to 36 years (M = 23.92, SD = 4.15). The age range of the women was from 18 to 39 years (M = 26.29, SD = 8.30). Each volunteer was paid £5 for their assistance.

Apparatus

Experiment control, stimulus presentation, and data acquisition were controlled by an ACT SIRIUS 1 microcomputer, communicating with a BIODATA Microlink IEEE bus device incorporating 12-bit A/D and 8-channel multiplexer unit. EEG AC signal amplification was via BIODATA PA400 preamplifier and main amplifier units. The tone stimulus was presented by a MEDELEC ST10 stimulator unit, triggered by program instruction from the SIRIUS computer. The tones were delivered binaurally via TDH 39 audiometric, electromagnetic headphones. EEG electrodes were Ag-AgCl 9-mm disc electrodes, fixed to the scalp via collodion, using standard NEPTIC electrode gel as the surface contact medium.

Stimulus Characteristics

A 1000-Hz digitally synthesized sine wave was generated by the ST10 stimulator. The amplitude of the tone was 85 dB, the total duration was 30 ms. The tone envelope was shaped with a rise and fall time of 3 ms, yielding a plateau of 24 ms at maximum amplitude. Signal onset and offset were always at 0 V, there were no switching transients. The interstimulus intervals were randomized between the range of 3 to 8 s. Unlike Hendrickson and Hendrickson, who insisted on using the same random sequence for each subject, our sequence was uniquely randomized for each subject. The subjective perception of the stimulus was as a "soft" tone.

EEG Montage and Channel Identification

Data were acquired from eight channels, at a sampling speed of 1000 Hz, for an epoch duration of 512 ms. The EEG electrodes were located on the scalp according to the 10–20 System; electrode impedance was always less than 5 kΩ. Five channels of bipolar EEG were acquired:

Channel 1: C_z–A_1 (mastoid process)
Channel 2: C_z–A_2 (mastoid process)

Channel 3: O_z–C_4
Channel 4: O_z–C_3
Channel 5: O_1–O_2
Channels 6 and 7: Monitored both horizontal and vertical eye movements (EOG). For vertical movement, two electrodes were attached immediately above and below the midpoint of the left eye. For horizontal movement detection, two electrodes were placed on either side of the scalp, in line with both eyes. The purpose of the EOG monitoring was not to examine EOG potentials per se but rather to use the activity on either channel as an indicator of possible EEG artifact.
Channel 8: Monitored electromyographic (EMG) activity of the suprahyoid muscle group using a recording electrode placed just below the jaw near the right ear. The reference electrode was the left eye EOG electrode from Channel 6. Once again, EMG activity was used purely as an indicator of possible artifact within the EEG.

The subject was grounded using an electrode placed on the tip of the nose.

Psychometric Tests

Each subject completed the Eysenck Personality Questionnaire (Eysenck & Eysenck, 1975) and the I_7 Questionnaire (Eysenck, Pearson, Easting, & Allsopp, 1985), assessing Impulsivity, Venturesome, and Empathy. Both personality tests were administered by the SIRIUS PC. In addition, the full Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) was administered. The subjects also took part in a reaction time (RT) task reported in Barrett, Eysenck, and Lucking (1986).

Acquisition Details

Prior to each subject taking part in the experiment, the amplification channels were calibrated using an SLE battery-powered oscillator generating a 200 μV peak to peak sine wave. The signal was continuously sampled and displayed by the SIRIUS so that the Microlink offset could be centralized manually in order to yield a balanced signal on each channel; 100 epochs of 512 ms duration were acquired via online 12-bit A/D. Amplification range was ±100 μV (200 μV peak to peak) yielding a measurement resolution of 0.05 μV. The BIODATA PA400 filters were set to yield a frequency bandwidth of 0.8 to 300 Hz per channel (3 dB attenuation at these frequencies with a 20 dB/decade rolloff). For each epoch, the tone was sounded via the ST10 and acquisition was started simultaneously, the data being stored in RAM before being transferred to hard disc as a binary file. Due to the inability of the SIRIUS Microlink implementation to cope with direct memory access, we had to proceed from RAM to disc. This added a 2.25-s delay between each 512-ms epoch. Thus, our interstimulus interval could not match Hendrickson and Hendrickson's exactly. Our range was between 3 and 8 s; Hendrickson and Hendrickson's was between 1 and 8 s.

Procedure

All subjects took part in the experiment at 10 a.m. each morning. The EEG procedure was always implemented first. The RT and psychometric tests were counterbalanced in order of presentation after the EEG acquisition phase. EEG acquisition took place with the subject sitting in a darkened room; the tester sat in an adjoining room and monitored the data acquisition, display oscilloscopes, and ST10 stimulator. There was intercom communication between the subject and tester at all times. The subjects were asked to relax, keep their eyes closed, move as little as possible, and listen to some tones that would be presented through the headphones. A few tones were presented manually in order to familiarize the subject with their characteristics and amplitude. The tones were then presented to the subject. After the experiment had concluded for each subject, the data were transmitted via a secure serial line to a PRIME 2250 for offline mass storage. At the end of the experiment, we took delivery of a MASSCOMP 6600 system. All data were subsequently transferred from the PRIME to the MASSCOMP for parameter computation, further digital signal processing, and statistical analysis.

Parameter Computation

Prior to parameter computation, each epoch was passed through an amplitude artifact analysis program. This procedure converted all sampling values (0–4,095) to their microvolt equivalent, then examined each epoch in turn for any values that were not within the voltage range of $\pm 75\ \mu V$. If an epoch contained one or more such values, then that epoch was rejected entirely from any further averaging analysis. Note that all further analysis was based on microvolt value data points.

Having passed through the amplitude artifact analysis, the remaining epochs were submitted to the averaging and parameter computation program. Each epoch was initially detrended by subtracting the mean voltage for the epoch from each sample value. The mean detrended epochs were then averaged, and parameters extracted from this procedure. Ten parameters were extracted for further analysis.

The Hendrickson String Measure.

$$\text{STRING} = \frac{\sum_{i=2}^{N} (V_{i-1} - V_i)^2}{N - 1}$$

where V = the array of sample points defining the AEP; and
N = the number of sample points over which to make the calculation.

Two versions of this parameter were computed: $N = 256$ (STR256) and $N = 512$ (STR 512).

The Hendrickson Variability Measure.

$$\text{VARIABILITY} = \sum_{i=1}^{N} \left[\frac{\sum_{j=1}^{K} (x_j - \bar{x}_j)^2}{K} \right]$$

where X = the array of values for one sample point, taken across epochs;
K = the number of epochs over which the sum is computed; and
N = the number of sample points over which to make the calculation.

Two versions of this parameter were computed: $N = 256$ (VAR256) and $N = 512$ (VAR512).

Average Absolute Amplitude of the AEP.

$$\text{AMPLITUDE} = \frac{\sum_{i=1}^{N} |V_i|}{N}$$

where V = the array of sample points defining the AEP; and
N = the number of sample points over which to make the calculation.

Two versions of this parameter were computed: $N = 256$ (VAR256) and $N = 512$ (VAR512).

Amplitude Regression Intercept. This parameter was the intercept from a linear regression of mean absolute epoch amplitude against epoch sequence number. That is:

linear regression $y = A + b\mathbf{X_K}$

where

$$\mathbf{X}_K = \frac{\sum_{i=1}^{N} |X_i|}{N};$$

N = the number of sample points over which to make the calculation; and
X_K = the vector containing the mean absolute epoch amplitude values. The X_K vector elements are regressed against the epoch sequence numbers $j = 1$ to K.

Two versions of this parameter were computed: $N = 256$ (VAR256) and $N = 512$ (VAR512).

Latency Parameters. Two parameters were computed: the latency of the largest negative observed voltage within the range 80 to 140 ms, and the latency of the largest positive observed voltage within the range from the most negative voltage to 220 ms. Out of convention, these parameters were assigned the names N100L and P180L with the expectation of observing their mean values as 100 ms and 180 ms, respectively.

Finally, the number of epochs retained for averaging, after the amplitude artifact process had been implemented, was recorded as an 11th variable (NUMEPOK). Hendrickson and Hendrickson had previously maintained that identical numbers of retained epochs are required in order to replicate the IQ by AEP correlations. In addition, they stated that the string measure was very sensitive to the number of epochs used in its calculation. A final comment indicated that just from inspection of the data, there was some correlation between this number and the overall IQ of the subject. However, no values or directional information were presented.

Analyses Undertaken

Analysis 1 correlated the AEP variables with the WAIS subscale scores and IQ variables using Pearson correlations. Outlier analysis was undertaken on most significant correlations via SYSTAT/SYGRAPH v.5.0 (Wilkinson, 1991) influence plot analysis. The AEP variables were computed over an uneven number of epochs.

Analysis 2 tested the Hendrickson and Hendrickson assertion that high-frequency waveforms are crucial to the replication of their results or at least the replication of IQ by AEP parameter relationships. This was implemented by reanalyzing the entire data set of retained epochs (i.e., those that had not been rejected by the amplitude artifact program). Prior to the mean detrending and parameter analysis across the epochs, each epoch was passed through a digital finite impulse response (FIR) filter via the ILS (1988) digital signal-processing software). One of the properties of the FIR filter is that it preserves the phase characteristics of the input signal (no lag effects). The filter was designed on the MASSCOMP as a 36th-order low-pass filter with a passband from 0 to 40 Hz and a stopband from 60 to 500 Hz. At 30 Hz, there was about 10 dB signal attenuation, at 40 Hz there was 21 dB attenuation, at 60 Hz there was 53 dB attenuation. This filter had the effect of removing *all* high frequency "jitter" in an epoch. If the Hendrickson paradigm relied on high-frequency EEG to facilitate the IQ by AEP correlations, the use of the excessively filtered data should diminish substantially the Hendrickson string and variability parameter correlations with IQ. The filtered data were submitted to the same form of analysis as in Analysis 1.

Analysis 3 tested Hendrickson and Hendrickson's assertion that the number of epochs used to calculate the AEP parameters can substantially affect any correlational results. In addition, the proposed (but not computed) relationship between the number of epochs retained and IQ was also empirically examined. This analysis proceeded by noting the minimum number of epochs retained across all subjects via the amplitude artifact analysis. Having established this value, all other subjects' retained epoch counts were reduced accordingly by removing extraneous epochs from the first 5 or last 5 epochs in the 100-epoch set. Where this was not possible (as they had already been rejected), the program began removal at 6+ epochs. The statistical analysis took place as in Analysis 1.

RESULTS

Prior to Analyses 1, 2, and 3, the data were examined for EOG and EMG artifacts, using the amplitude artifacts detected on Channels 6, 7, and 8. These "artifacts" corresponded to an eyeblink or gross movement (such as swallowing or other jaw movement). The AEP for each subject was generated either by ignoring such artifacts on Channels 6–8, or by rejecting any epoch where an amplitude artifact had been detected on Channels 6–8. Comparative analysis of both datasets yielded very little difference in the IQ by AEP correlation patterns between EOG and EMG artifacted and nonartifacted EEG. Of note was the fact that few out-of-range epochs on Channels 6–8 were identified as amplitude artifacts on the remaining channels. Because removing these epochs had little effect on the AEPs generated, it was decided to proceed with all further analyses ignoring activity on the EOG and EMG channels.

With regard to Channel 5, O_1–O_2, no identifiable AEP was apparent in any subject's data. It was clear that there was no recognizable brain response using this bipolar montage and number of stimulus repetitions. Thus, all analyses were restricted to Channels 1–4.

Analyses

In order to give some feel for the measurement range of the variables used, Table 1 (p. 370) presents the means, standard deviations, and minimum and maximum values for the main variables used for all analyses. (The complete variable matrices containing the RT, Personality, AEP, and IQ data for each channel are available from the first author as either EXCEL v.3.0 spreadsheet or SYSTAT v.5.0 analysis files.)

In comparison with the published Hendrickson and Hendrickson (1982) data, our mean string and variability parameters are of a completely different magnitude. Personal communication with D.E. Hendrickson and A.E. Hendrickson on this matter indicated that they used a scaling procedure on the variability calculation in order to reduce the size of the parameter values. With regard to the string measure, they computed their measure on the raw digitized data points (range −2048 to +2048); we used the microvolt equivalents (range −100 to +100). As can be seen from the WAIS-IQ values in Table 1, the range is quite wide. However, only 1 subject had an IQ < 72.

Analysis 1. Tables 2, 3, and 4 (pp. 370–371) present the WAIS IQ by AEP parameter correlations. With regard to significance levels for the correlations, a

TABLE 1
Means, Standard Deviations, and Minimum and Maximum Values for RT, WAIS, and Channel 1 AEP Variables

Variable	*M*	(*SD*)	Min.	Max.
WAIS Verbal IQ	104.43	(18.22)	61.00	142.00
WAIS Performance IQ	104.35	(18.50)	62.00	144.00
WAIS Full-Scale IQ	105.23	(18.99)	61.00	134.00
Hendrickson String Measure (256-ms epoch) STR256	0.25	(0.18)	0.08	0.83
Hendrickson String Measure (512-ms epoch) STR512	0.21	(0.18)	0.06	0.82
Hendrickson Variability Measure (256-ms epoch) VAR256	43025.65	(15626.92)	13228.94	78282.92
Hendrickson Variable Measure (512-ms epoch) VAR512	83589.61	(29357.43)	25419.01	141093.17
Average Absolute Amplitude (256-ms epoch) AMP256	6.66	(2.12)	3.02	10.99
Average Absolute Amplitude (512-ms epoch) AMP512	4.87	(1.44)	2786.00	8.02
Epoch Amplitude Regression intercept (256-ms epoch) A256	12.70	(2.62)	7.37	17.49
Epoch Amplitude Regression intercept (512-ms epoch) A512	11.66	(2.27)	6.99	15.64
N100 latency (in ms) N100L	94.28	(11.11)	80.00	122.00
P180 latency (in ms) P180L	184.12	(19.28)	155.00	220.00
No. of epochs (excluding amplitude artifact) NUMEPOK	98.38	(1.98)	92.00	100.00

Note. N = 40.

TABLE 2
Pearson Correlations Between WAIS Verbal IQ and AEP Variables

WAIS Verbal IQ	Channel 1 $C_Z–A_1$	Channel 2 $C_Z–A_2$	Channel 3 $O_Z–A_4$	Channel 4 $O_Z–A_3$
STR256	−.37	−.28	−.17	−.14
STR512	−.33	−.26	−.15	−.13
VAR256	−.46	−.40	−.40	−.42
VAR 512	−.44	−.37	−.41	−.43
AMP256	−.38	−.31	.04	−.04
AMP512	−.49	−.43	−.04	−.10
A256	−.52	−.39	−.39	−.41
A512	−.53	−.40	−.40	−.41
N100L	−.19	−.07	−.01	−.16
P180L	−.18	−.09	−.27	−.25
NUMEPOK	.17	.09	.10	.12

Note. N = 40.

TABLE 3
Pearson Correlations Between WAIS Performance IQ and AEP Variables

WAIS Performance IQ	Channel 1 C_Z–A_1	Channel 2 C_Z–A_2	Channel 3 O_Z–A_4	Channel 4 O_Z–A_3
STR256	−.49	−.42	−.28	−.28
STR512	−.45	−.38	−.26	−.26
VAR256	−.40	−.38	−.35	−.38
VAR 512	−.37	−.34	−.35	−.38
AMP256	−.43	−.36	.04	−.05
AMP512	−.49	−.41	.01	−.07
A256	−.51	−.43	−.39	−.39
A512	−.52	−.42	−.40	−.40
N100L	−.18	−.11	.01	−.16
P180L	−.12	−.14	−.17	−.15
NUMEPOK	.24	.11	.15	.17

Note. $N = 40$.

value of .32 or greater would be significant at $p < .05$, two-tailed, a value of .40 or greater would be significant at $p < .01$, two-tailed. The 95% confidence limits on a correlation of .32 are between .01 and .57. For a correlation of .40, the limits are .10 and .63. For a correlation of .5, the limits are between .22 and .70. Given the number of correlations computed, it is probably more sensible to look for consistent patterning across variables rather than at significance values per se.

With regard to outlier observation analysis, there was 1 subject (Subject 3) whose data were a potential problem; this was the subject with a Full-Scale

TABLE 4
Pearson Correlations Between Full-Scale WAIS IQ and AEP Variables

Full-Scale WAIS IQ	Channel 1 C_Z–A_1	Channel 2 C_Z–A_2	Channel 3 O_Z–A_4	Channel 4 O_Z–A_3
STR256	−.44	−.35	−.23	−.21
STR512	−.40	−.32	−.21	−.20
VAR256	−.45	−.40	−.38	−.42
VAR 512	−.42	−.37	−.38	−.43
AMP256	−.42	−.34	.05	−.05
AMP512	−.51	−.44	−.01	−.10
A256	−.54	−.42	−.40	−.42
A512	−.55	−.43	−.41	−.43
N100L	−.21	−.11	−.02	−.18
P180L	−.16	−.12	−.22	−.20
NUMEPOK	.19	.10	.12	.15

Note. $N = 40$.

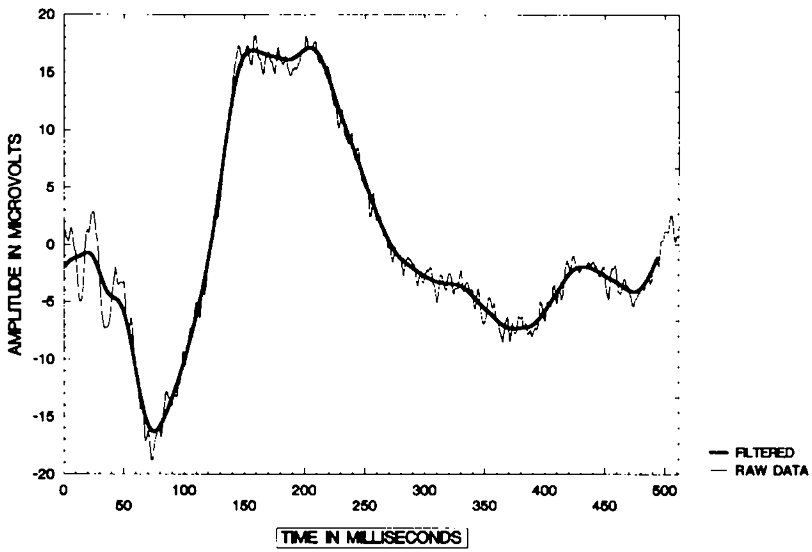

Figure 1. Subject 3's Full-Scale WAIS-IQ = 61 (Filtered vs. Unfiltered AEPs)

WAIS-IQ of 61. The largest effect observed was over the latency correlations N100L and P180L. Figure 1 shows the unfiltered and filtered AEPs for this individual. As can be seen from this figure, an early N100 trough is followed by a distorted P180 peak. Removal of this subject increased the negative latency by IQ correlations by a maximum of .10.

Although it is tempting to remove this subject from the analysis, it must be borne in mind that the effect of removal also decreased the magnitude of some correlations, but only by about .03–.05. In these other relationships, this individual's parameter values were not considered outliers either visually or analytically. Thus, it was decided to retain this subject within all parameter analyses rather than make subjective judgements about which parameter values to use and which to reject. It must be remembered that the parameters were all extracted from the same AEP. It would be rather irregular to use some parameters where it boosted expected values and reject others where it reduced expected values. The AEP is either valid or is not. We have chosen the former option, and now examine each parameter set in turn.

String. The string measure correlations are reversed in sign to those reported by Hendrickson and Hendrickson. Our data seem to indicate that AEP contour complexity, as defined by the string measure, decreases as IQ increases.

Variability. The variability correlations are in the same direction as the Hendrickson and Hendrickson correlations. Although their size is less than the −.72 reported by Hendrickson and Hendrickson, they are conceptually significant nevertheless.

Amplitude. The AEP amplitude measures are in the reverse direction than would be expected from previous research (Dustman & Beck, 1972; Haier, Robinson, Braden, & Krengel, 1984; Haier, Robinson, Braden, & Williams, 1983; Rhodes, Dustman, & Beck, 1969; Shagass, Roemer, Straunanis, & Josiassen, 1981). Noticeably, the amplitude measures fail to yield this relationship on Channels 3 and 4. String and amplitude measures correlated maximally at .34 on Channel 1. On Channels 2, 3, and 4, the correlations dropped to around .1. (over 256- or 512-ms epochs). Figure 2 gives a graphic illustration of the amplitude difference between the mean AEPs of the 5 highest and 5 lowest IQ subjects. The IQ range for the low-IQ group is between 61 and 82, the range for the high-IQ group is between 130 and 134.

Intercept Measure. This measure showed the highest relationships with IQ. It is related to the amplitude measure with a Pearson correlation of about .5 to .6, taken across channels and 256-ms and 512-ms epoch parameter values. It is related also to the string measure, across all channels, between .6 to .7.

N100L and P180L Latencies. All latency correlations were in the expected direction, but at a statistically and conceptually nonsignificant level. As Barrett and Eysenck (1991) indicated in their review of this area, replicability of these relationships is tenuous because they were only ever convincingly demonstrated on polarized IQ groups (no middle range).

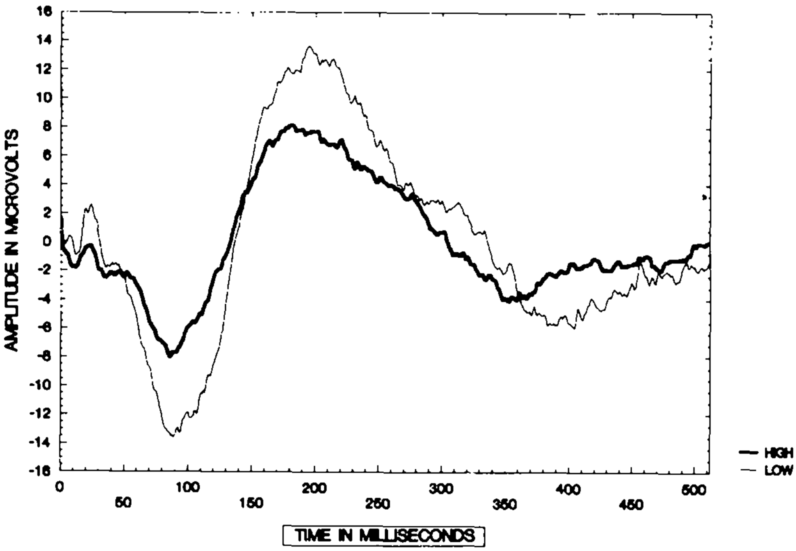

Figure 2. Low versus high WAIS-IQ groups (low IQ, M = 73; high IQ, M = 132)

The Number of Epochs. There was no relationship between the number of epochs and IQ. The directions are all positive but of such little magnitude that the relationship may be considered negligible.

Looking at Tables 2, 3, and 4, it is apparent that there is no significant difference between the 256-ms and the 512-ms epoch parameter relationships. In addition, the parameter relationships appear to be of equal size with either Verbal or Performance IQ.

Analysis 2. The results from this analysis yielded surprisingly little change to the overall correlational pattern. Only one parameter was substantially affected by the filtering operation, the string variable. This makes sense in that this parameter is computed directly from the waveform envelope. Table 5 presents the results for Channel 1 across the three derived WAIS-IQ variables. It is apparent from this table that the variability parameters, VAR256 and VAR512, are more related to Verbal IQ than Performance IQ, in line with the Hendrickson and Hendrickson finding. Channel 2, 3, and 4 data showed the same effect. As in Analysis 1, the amplitude parameter correlations disappeared for Channels 3 and 4. It is clear from Table 5 that there is no difference between filtered or unfiltered data with regard to the reported parameter relationships barring the string measure. Looking at the latency parameters, N100L and P180L, it might be argued that filtering has improved the level of relationships overall (with the potential outlier observation removed, these latency correlations increased on average by about −.06 to −.11, putting them in the range of the expected value of around −.3).

TABLE 5
Pearson Correlations Between WAIS Full-Scale, Verbal, and Performance IQs and FIR Filtered Data AEP Variables

FIR Filtered Data Channel 1 C_Z–A_1	WAIS-Full	WAIS-VIQ	WAIS-PIQ
STR256	−.19	−.16	−.22
STR512	−.23	−.19	−.25
VAR256	−.42	−.49	−.29
VAR 512	−.39	−.47	−.24
AMP256	−.42	−.38	−.42
AMP512	−.51	−.49	−.48
A256	−.51	−.53	−.43
A512	−.54	−.58	−.43
N100L	−.26	−.23	−.22
P180L	−.28	−.30	−.25
NUMEPOK	.19	.17	.24

Note. $N = 40$.

Analysis 3. This analysis yielded no conceptually significant differences at all between the correlations computed here and those computed in Analysis 1. On average, the differences between correlations was about ±.02. The largest observed difference value across all five channels was .04. Within this data set, and given the numbers of epochs rejected overall, it can be stated unequivocally that computing AEPs over unequal or equal numbers of epochs makes no difference to the resulting parameter relationships.

From these results, one parameter might be said to have been missing from all the analyses, the Hendrickson and Hendrickson "composite" variance-string score. However, this parameter appears to be no more than an ad hoc transformation of the two variables with no apparent rationale (a whole class of such arbitrary transformations could be applied in this manner). But, Hendrickson and Hendrickson did obtain their highest correlational values with IQ using this parameter. So, we computed this parameter across the four-channel data set from Analysis 1. Because the means and standard deviations of our string and variance measures were absolutely disparate, we first normalized each parameter data set, expressing each observation as a deviation score within a unit normal distribution. Then, by simple parameter subtraction, we formed the composite measure. This measure failed to correlate above .2 with Verbal, Performance, or Full-Scale IQ on Channels 1 and 2. The highest correlation of .2 was with Performance IQ on Channel 2. On Channels 3 and 4, the composite measure yielded negative correlations with IQ, contrary to Hendrickson and Hendrickson's results. The highest correlation was −.28 with Verbal IQ on Channel 4. All correlations were statistically *nonsignificant* ($p > .05$).

DISCUSSION

In light of the preceding results and recent work on AEP by IQ relationships, we can now assess the nine points presented by Hendrickson and Hendrickson as the optimal methodology to follow in replicating their results.

1. *Stimulus Choice:* Using a tightly controlled stimulus as required by Hendrickson and Hendrickson, we were unable to replicate the string measure relationships; however, we were able to replicate the variability parameter relationships with IQ. Caryl and Fraser (1985), using a standard laboratory oscillator and not controlling for onset or offset transients, replicated the size and direction of the Hendrickson and Hendrickson string parameter relationships. Haier et al. (1983), using a light flash stimulus (not recommended by Hendrickson and Hendrickson), also replicated the string parameter relationships. From this, we conclude that the presence of stimulus switching transients are irrelevant to replication of the Hendrickson and Hendrickson results.

2. *Stimulus Presentation:* There is no rationale to insist on the same random presentation to each subject within an experiment. To assume otherwise indicates that some random sequences are better than others at eliciting AEP by IQ parameter relationships. If this assumption holds, all investigators would have to use the Hendrickson and Hendrickson random sequence in order to maintain replicability. However, the random sequence is no longer "random" as such: It has now been imbued with unique, unknown properties.
3. *Electrodes:* Hendrickson and Hendrickson's preference for Ag-AgCl electrodes rather than gold or tin as in the Electrocap™ is just that, a preference.
4. *Amplification:* The number of studies that have and have not replicated the string parameter relationships including our own, used a wide variety of commercial amplifiers. There is no consistent pattern among studies to indicate that the use of the Hendrickson and Hendrickson amplifier is any better than the use of a commercial standard amplifier. With regard to the bandwidth recommendations, Analysis 2 demonstrated that whether EEG is filtered between 0.8 to 300 Hz or between 0.8 to about 30 Hz makes no significant difference to correlational results. We also note that Haier et al. (1983) recorded their EEG with a low-pass filter cutting in at 40 Hz, and were still able to replicate the Hendrickson and Hendrickson data. Stough, Nettlebeck, and Cooper (1990) were also able to replicate the string by IQ relationship using EEG filtered between 2 and 125 Hz. We conclude that the amplifier bandwidth is irrelevant for replicating Hendrickson and Hendrickson's results.
5. *Calibration Signal:* No successful replication of the Hendrickson and Hendrickson string measure relationships has adopted their recommendation for calibration.
6. *Recording Medium:* Their suggestion of online digitization is sound advice. All studies implemented to date have used online methodology.
7. *A/D Conversion Sampling Rate:* Hendrickson and Hendrickson recommended a minimum sampling rate of 1000 Hz. However, Haier et al. (1983) and Stough, Nettelbeck, and Cooper (1990) both used sampling rates of 256 Hz, and still maintained replicability of the Hendrickson and Hendrickson string measure. Although we adopted a 1000-Hz sampling rate, we were unable to achieve replication of this parameter. We consider that sampling rate is an irrelevant parameter to obtaining string by IQ parameter relationships down to about 100 Hz. Below 100 Hz the EEG must be filtered such that no frequencies above 50 Hz remain in the waveform. Frequency aliasing (high frequencies being sampled incorrectly can appear as low-frequency waveforms in the digitized record) into low frequencies caused by sampling above the Nyquist limit might well cause problems given the restricted bandwidth of the signal. In addition, peak identification will be tenuous as a 10-ms "peak" process would only be identified by one point.

The sampling jitter across the epochs for that peak would be likely to smear the "true" peak value.

8. *Epoch of Analysis Period:* Hendrickson and Hendrickson stated that a 250-ms epoch should be used. Our own data demonstrated little difference in IQ relationships between a 256- or 512-ms parameter epoch. Haier et al. (1983) demonstrated that a 508-ms epoch string measure correlated with IQ to a greater extent than did a 252-ms epoch measure. Stough et al (1990) suggested that an epoch length of 100 ms selected from within the total epoch might be more optimal. They found that selecting an AEP segment from 100 ms to 200 ms from within their 300-ms epoch yielded maximum string by IQ correlations. We conclude that epoch length is quite arbitrary to observing optimal AEP parameter by IQ relationships within the range of 100 to 512 ms. Each study appears to support different epoch lengths.
9. *Editing of Individual Records:* Analysis 3 demonstrated that, within the range of rejected epochs for our sample (from 92 to 100), there was no difference between computing AEP parameters across a fixed number of epochs or computing them over a variable number. However, although our range of rejection was limited, it did show that the global assertion of Hendrickson and Hendrickson on this point was unwarranted. It is still unknown whether variable numbers of epochs would significantly change results. Given the improvement of signal to noise is roughly equivalent to $\sqrt{N}$, where N = the number of epochs, the improvement in signal to noise ratio from 81 to 100 epochs is 9:1 to 10:1. Although we would recommend that this might be the lowest optimal bound, this is a subjective recommendation only, based on the small improvement in signal to noise ratio from 81 to 100 epochs.

Given we have generated AEPs to random stimuli, it is of relevance to question the existence or otherwise of a P300 component within our data set. Both Figures 1 and 2 demonstrate graphically that no such component could be identified unambiguously in the great majority of AEPs. Because the paradigm requires no stimulus-based decisions or responses to be made by a subject, it is perhaps not surprising that there is no identifiable P300 component in the waveform.

From the preceding critical assessment of Hendrickson and Hendrickson's points, we conclude that the Hendrickson paradigm as such is not a paradigm at all. Rather, any well-controlled attempt to generate auditory or visual evoked potentials will possibly yield AEP parameter by IQ correlations. However, the number of nonreplications of some or all parameters in studies that use high-specification methodology (such as our own) suggests that this area is still in a state of flux. It would appear that the methodological excellence or otherwise of all studies to date plays little part in determining parameter relationships. This is exemplified by our finding of negative correlation between the string measure and IQ. This is in the opposite direction to that expected from the original Hendrickson and Hendrickson work. Looking at the available evidence to date concerning this particular AEP parameter, we find that Shagass et al. (1981) observed statistically significant negative correlations between their string measure and Raven IQ within the range of −.22 to −.31 (N = 80). Haier et al. (1983) observed significant positive correlations between their string measure and Raven IQ within the range of .43 to .50 (N = 22). Caryl and Fraser (1985) observed a statistically significant positive correlation of .78 (N = 10) between the Hendrickson and Hendrickson string measure and AH4 IQ. Vetterli and Furedy (1985), in their recalculations using a subsample of Ertl and Schafer's (1969) data, observed a significant positive correlation of .80 (N = 20). However, when they carried out the same recalculations on a set of data from Weinberg (1969), they observed a statistically nonsignificant correlation of −.34 (12 subjects). Mackintosh (1986) reported a statistically nonsignificant correlation between the Hendrickson and Hendrickson string measure and Raven IQ of −.34 (N = 18). Vogel, Kruger, Schalt, Schnobel, and Hassling (1987) reported near-zero correlations between their version and the Hendrickson and Hendrickson version of the string measure and Raven-IQ (N = 236). Finally, Stough et al. (1990) observed statistically significant positive correlations within the .43 to .71 range between the Hendrickson and Hendrickson string measure and WAIS-IQ (N = 20). Noticeably, Stough et al. failed to observe any statistically significant correlations between Raven-IQ and the same string measure.

These conflicting results appear to show no pattern with regard to methodology quality or waveform amplitude parameter relationships. However, the sample sizes in most of the studies are so low that it is difficult to assess what the "true correlation" might be, let alone its direction. The Vogel et al. (1987) study is probably *the* most important attempted replication of Hendrickson and Hendrickson's work in that it used over 200 subjects and very exacting methodology. The problem with the study results is that, from confidence limits analysis of the Hendrickson and Hendrickson results, they should be considered a statistical anomaly. However, because the argument can be reversed and the Hendrickson and Hendrickson results likewise dismissed, it is apparent that recourse to sterile statistical inferences leads nowhere. Frankly, there is no easily discernable explanation for this disparate range of results. Rather, it is suggestive that some other factor (or factors) is operating on the paradigm, a factor powerful enough to affect bivariate relationships to a significant degree. However, with all the studies mentioned previously, so few other parameters have been computed and/or reported that critical evaluation of other features of the data sets is simply not yet possible. Barrett and Eysenck (1991) and Deary and Caryl (in press) both discussed this issue further, but both sets of reviewers concluded with the general proposal that more results are needed from more carefully crafted experiments before any substantive statements can begin to be made about the exact conditions under which significant AEP parameter by IQ correlations might be found.

Unfortunately, as with the initial beginnings of the inspection time experiments, the area is dogged by small sample research (*N*s = 10–20 per study), insufficient numbers of parameters computed (sometimes just the string measure alone), little detailed reporting of experiment methodology, and too little appreciation of the actual state of play of research in this area (limited, barely replicable results). However, with the right computer hardware and software, appropriate electrode technology (using the Electrocap™ rather than manual application of electrodes via paste or collodion), the typical AEP experiment can be implemented in a routine and very quick fashion. It is to be hoped that future investigators will cease paying attention only to those AEP parameters that fit their particular viewpoint and begin computing (as standard) a more complete parameter set based around waveform envelope complexity (string), epoch variability, waveform amplitude, peak component latencies, and component amplitude transients (trough to peak voltage difference and latency, etc.). These measures will at least enable investigators to look for common features among studies, to find whether relationships exist or not, and perhaps begin to determine the most consistent parameter under a wide variety of conditions. The fact that we are the only investigators to examine the variability parameter since Hendrickson and Hendrickson's (1982) study is barely credible, especially because Hendrickson and Hendrickson recommended this as the optimal AEP by IQ correlational parameter (1982, p. 208). In addition, it is this parameter that is the most important from a theoretical viewpoint. Evidence from the areas of simple and complex RT (Barrett et al., 1986; Frearson & Eysenck, 1986; Jensen, 1982) and from human sensory nerve conduction (Barrett, Daum, & Eysenck, 1990) has indicated that response variability correlates negatively with psychometric IQ. The size of this relationship is generally within the −.3 to −.6 range. Although the Hendrickson and Hendrickson biochemical model may not be an accurate explanatory process of the causation of this variability, the simple proposition of "noisy" neural transmission systems and their effects on cognitive performance is a pervasive and powerful concept.

Finally, although we have been somewhat critical of Hendrickson and Hendrickson's rules for replication, it must be remembered that it was Hendrickson and Hendrickson who gave this area the impetus it required. Without their seminal work, the status of the EEG by IQ area of investigation would be as it was, defunct.

In conclusion, this study has confirmed the Hendrickson and Hendrickson results indicating a negative correlation between AEP variability and IQ. The level at which this correlation is confirmed (about −.43 on average) across four scalp locations, using both filtered and unfiltered EEG, is conceptually and statistically significant.

REFERENCES

Barrett, P.T., Daum, I., & Eysenck, H.J. (1990). Sensory nerve conduction and intelligence. *Journal of Psychophysiology, 4*, 1–13.

Barrett, P.T., & Eysenck, H.J. (1991). Brain electrical potential and intelligence. In A. Gale & M.W. Eysenck (Eds.), *The handbook of individual differences: Biological perspectives.* New York: Wiley.

Barrett, P.T., Eysenck, H.J., & Lucking, S. (1986). Reaction time and intelligence: A replicated study. *Intelligence, 10,* 9–40.

Bennett, W.F. (1968). Human perception: A network theory approach. *Nature, 220,* 1147–1148.

Blinkhorn, S.F., & Hendrickson, D.E. (1982). Averaged evoked responses and psychometric intelligence. *Nature, 295,* 596–597.

Caryl, P.G., & Fraser, I.J. (1985, September). *The Hendrickson "string length" measure and intelligence—A replication.* Paper presented at Psychophysiology Society Scottish Conference.

Chalke, F., & Ertl, J. (1965). Evoked potentials and intelligence. *Life Sciences, 4,* 1319–1322.

Deary, I., & Caryl, P. (in press). Intelligence, EEG, and evoked potentials. In P.A. Vernon (Ed.), *Biological approaches to human intelligence.* Norwood, NJ: Ablex.

Dustman, R.E., & Beck, E.C. (1972). Relationship of intelligence to visually evoked responses. *Electroencephalography and Clinical Neurophysiology, 33,* 254.

Ertl, J. (1971). Fourier analysis of evoked potentials and human intelligence. *Nature, 230,* 525–526.

Ertl, J. (1973). Evoked responses and Fourier analysis. *Nature, 241,* 209–210.

Ertl, J., & Schafer, E. (1969). Brain response correlates of psychometric intelligence. *Nature, 223,* 421–422.

Eysenck, H.J., & Eysenck, S.B.G. (1975). *Manual to the Eysenck Personality Questionnaire.* London: Hodder & Stoughton.

Eysenck, S.B.G., Pearson, P.R., Easting, G., & Allsopp, J. (1985). Age norms for impulsiveness, venturesomeness, and empathy in adults. *Personality and Individual Differences, 6,* 613–620.

Frearson, W., & Eysenck, H.J. (1986). Intelligence, reaction time, and a new odd-man-out paradigm. *Personality and Individual Differences, 7,* 807–817.

Haier, R.J., Robinson, D., Braden, W., & Krengel, M. (1984). Psychometric intelligence and visual evoked potentials: A replication. *Personality and Individual Differences, 5,* 487–489.

Haier, R.J., Robinson, D.L., Braden, W., & Williams, D. (1983). Electrical potentials of the cerebral cortex and psychometric intelligence. *Personality and Individual Differences, 5,* 591–599.

Hendrickson, D.E., & Hendrickson, A.E. (1980). The biological basis of individual differences in intelligence. *Personality and Individual Differences, 1,* 3–33.

Hendrickson, D.E., & Hendrickson, A.E. (1982). The biological basis of intelligence. In H.J. Eysenck (Ed.), *A model for intelligence.* New York: Springer-Verlag.

ILS: Interactive Laboratory System Software. v.6.0 (1988). Encina, CA: Signal Technology Ltd.

Jensen, A.R. (1982). Reaction time and psychometric *g.* In H.J. Eysenck (Ed.), *A model for intelligence.* New York: Springer-Verlag.

Mackintosh, N.J. (1986). The biology of intelligence? *British Journal of Psychology, 77,* 1–18.

Rhodes, L.E., Dustman, R.E., & Beck, E.C. (1969). The visual evoked response: A comparison of bright and dull children. *Electroencephalography and Clinical Neurophysiology, 27,* 364–372.

Shagass, C., Roemer, R.A., Straunanis, J.J., & Josiassen, R.C. (1981). Intelligence as a factor in evoked potential studies in psychopathology: 1. Comparison of low- and high-IQ subjects. *Biological Psychiatry, 11,* 1007–1029.

Shucard, D., & Horn, J.L. (1972). Evoked cortical potentials and measurement of human abilities. *Journal of Comparative Physiological Psychology, 78,* 59–68.

Stough, C.K.K., Nettelbeck, T., & Cooper, C.J. (1990). Evoked brain potentials, string length, and intelligence. *Personality and Individual Differences, 11,* 401–406.

Vetterli, C.F., & Furedy, J.J. (1985). Evoked potential correlates of intelligence: Some problems with

Hendrickson's string measure of evoked potential complexity and error theory of intelligence. *International Journal of Psychophysiology, 3,* 1–3.

Vogel, F., Kruger, J., Schalt, E., Schnobel, R., & Hassling, L. (1987). No consistent relationships between oscillations and latencies of visual and auditory evoked EEG potentials and measures of mental performance. *Human Neurobiology, 6,* 173–182.

Wechsler, D. (1981). *WAIS-R Manual: Wechsler Adult Intelligence Scale-Revised.* New York: Psychological Corporation.

Weinberg, H. (1969). Correlation of frequency spectra of averaged visual evoked potentials with verbal intelligence. *Nature, 224,* 813–814.

Wilkinson, L. (1991). *SYSTAT/SYGRAPH v.5.0.* Evanston, IL.