Person. individ. Diff. Vol. 16, No. 1, pp. 3–32, 1994
Printed in Great Britain

0191-8869/94 $6.00 + 0.00
Pergamon Press Ltd

THE RELATIONSHIP BETWEEN EVOKED POTENTIAL COMPONENT AMPLITUDE, LATENCY, CONTOUR LENGTH, VARIABILITY, ZERO-CROSSINGS, AND PSYCHOMETRIC INTELLIGENCE

P. T. BARRETT[1] and H. J. EYSENCK[2]

[1]University of Surrey, Department of Psychology, Guildford, Surrey GU2 5XH and
[2]Institute of Psychiatry, De Crespigny Park, Denmark Hill, London SE5 8AF, England

(Received 20 May 1993)

Summary—Two hundred subjects provided data from within three separate studies that attempted to replicate correlations between averaged evoked potential (AEP) indices and psychometric IQ. In addition, AEP zero-cross analysis was undertaken as a specific test of a proposition made within the Weiss quantum theory of intelligence. Measures of AEP variability, mean individual epoch amplitude, and P180 component latencies were found to correlate negatively with IQ at around −0.50 across the three studies. However, the consistency and size of relationship in the results was found to be a function of selecting subjects whose AEP P180 component amplitude was greater than some specified, sample dependent, target value. The zero-cross analysis, contrary to predictions, yielded no correlations with IQ. Robinson's cerebral arousability theory was noted as a possible explanatory framework for the results. In addition, it was noted that if Robinson is correct in his assertion of the complex analogue nature of the evoked response, conventional AEP analysis is no longer relevant.

INTRODUCTION

Studies of intelligence and cognitive functioning and intelligence have nearly always adopted a correlational or factor analytic paradigm (Carroll, 1993). An implicit condition for the use of such methods is the **homogeneity** of the samples studied; if males and females differed in **level** or **composition** of IQ, clearly they would have to be analysed separately. Yet there is evidence from several sources which suggest that there may be several causes of heterogeneity in samples typically used in such studies, and that such heterogeneity may confound the aims of the analyses normally carried out. Thus samples of children **high** on N (neuroticism) have been shown to have a very much simpler factor structure than a similar sample of children **low** on N (Lienert, 1968; Eysenck & White, 1969; but see Cohen & Williams, 1967). Low IQ groups show higher test intercorrelations than high IQ groups (Detterman & Daniel, 1989); similarly, lower ability *S*s show higher heritabilities and lower shared environmentality (Detterman, Thompson & Plomin, 1990). Even looking at items within a simple test, heterogeneity can be important (Van der Ven, 1992); application of a dichotomous Rasch model to the 'Cubes' test demonstrated the existence of several different strategies within which there was homogeneity, but the existence of which caused the Rasch model to fail badly. Another example is verbalization of problem solutions in non-verbal tasks (Franzen & Merz, 1976; Merz, 1969); this reduces interindividual variance in relevant IQ tests. In a similar manner, when Hick's law has been used to measure intelligence in reaction time (RT) experiments, a subsample of testees fails to adhere to Hick's law (Barrett, Eysenck & Lucking, 1986). For inspection time (IT) experiments, it has also been found that a subsample adopts a strategy relying on visual after-effects rather than adopting the majority strategy (Nettlebeck, 1982). Such sub-samples usually differ from the majority by showing different-size correlations between the variable measured (RT or IT) and IQ, suggesting that it is not very meaningful to assess overall correlations combining the two groups apparently differing in strategy, any such correlation being in part a function of the (accidental) mixture of numbers of the two such samples. The samples may of course differ in the respective contribution made by components of the total score. Thus RT can be divided into DT (decision time) and MT (movement time), and different sub-samples may show different correlations with IQ of these two components of RT (Whitley & Montano, 1992).

Little has been written on the problems presented by such heterogeneity, but recognition of the facts presented suggests certain important consequences. Theoretically, how can psychometric theory deal with the problem presented? If theory requires homogeneous groups for analysis, how can they be obtained? If different strategies create heterogeneity, should we not admit their study is an important part of work on individual differences? Most important, from the point of view of this article, is it possible to account for widely differing (non-replicable) results in terms of heterogenous samples? And if so, how should we proceed to obtain homogenous samples?

Studies of the correlation between IQ and different aspects of the averaged evoked potential (AEP) have often given dissimilar results. This may suggest the absence of any true relation, but there is too much meaningful agreement to make this hypothesis acceptable. We decided to look at the possibility that heterogeneity of samples might be responsible; in other words, for some people there might exist a **strong** correlation between IQ and certain aspects of the AEP, while for others such a correlation might be lacking. Such a hypothesis can only be tested if certain guidelines are adhered to. It is always possible to create two such groups artificially by selecting them on the basis of their position in the scatter diagram; retain those in the positive quadrants (+ + and − −), and contrast with those in the negative quadrants (+ − and − +). Such a procedure would of course succeed (inevitably), but provide no useful information.

What is needed is the discovery of an **objective** indicator **independent** of IQ which **repeatedly** succeeds in splitting the total sample into a high- and low-correlation group, and which gives a theoretically meaningful reason for that distinction. In addition, such an indicator should, if possible, be applicable in fields other than AEP research, e.g. it should discriminate between high IT–IQ correlation groups and low IT–IQ correlation groups. It is the major purpose of this article to search out such an indicator, and present evidence for its usefulness; as will be seen, the indicator suggested conforms to all the rules laid down, with the possible exception of the sound theoretical basis; we have suggested attentional variables, as have others for RT and IT work, but only direct testing of deductions from the theory can put it on a firmer footing or disprove it altogether.

Before detailing the methodologies adopted in the reported studies below, it is useful to briefly review the area within which they have been implemented. The relationship between AEPs and psychometric intelligence has been investigated 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 *S*s. This work was replicated and extended further by Ertl and Schafer (1969), Ertl (1971, 1973), Bennett (1968), and Shucard and Horn (1972) amongst others. The average correlation found across these various studies was about −0.30. During the 1980s, Hendrickson and Hendrickson (1980. 1982), and Blinkhorn and Hendrickson (1982) produced new evidence on the correlation of AEP and IQ measures. Based upon a novel model of synaptic structure, function, and nerve transmission, the Hendricksons 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 length of the AEP waveform; the larger this value, the higher an individual's IQ. The second measure, the *variance*, was computed by 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/variance of the AEP is drawn from two studies. The first, reported by Blinkhorn and Hendrickson (1982) using 33 *S*s and auditory stimulation, correlated the complexity (string) measure with performance on Raven's Advanced Progressive Matrices (RAPM: Raven, 1983) and a variety of verbal ability tests. Various correlations between the string measure (computed from AEPs generated over 90, 64 and 32 epochs) and the APM yielded a mid-range correlation of approx. 0.45. The verbal test scores did not correlate significantly

with the string measure. 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 sub-scales, performance, verbal, and overall IQ. The correlations among the WAIS IQ and string and variance AEP measures were 0.72 and −0.72, respectively.

Since these pioneering studies, there have been further attempts to replicate the results. Unfortunately, apart from Barrett and Eysenck (1992a), the only attempts at replication have focused on amplitude and/or string measures to the exclusion of component latencies and variability. The general consensus on the replicability of individual parameter correlations is that there is little consistency in the patterning of results. Correlations have even reversed sign across studies. 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, 1992b; Deary & Caryl, 1992). Since these reviews, some further evidence has been reported by Pelosi, Holly, Slade, Hayward, Barrett and Blumhardt (1992) indicating that in a Sternberg digit probe identification task, IQ and the peak-to-baseline amplitude of an N290 component were correlated at up to 0.67 with total WAIS IQ (19 *S*s). In addition, an analysis of the dimensional complexity of the EEG (using non-linear dynamics to determine the fractal dimensionality of spontaneous EEG recorded under various cognitive load conditions) reported by Lutzenberger, Birbaumer, Flor, Rockstroh and Elbert (1992), demonstrated that under resting conditions, high IQ *S*s have higher dimensional complexity than those with low IQ. No correlations were computed but the reported ANOVA results were statistically significant (34 *S*s). The significance of this result is based upon the concept of chaotic firing patterns of neuronal cell assemblies. The less focused activity required of a *S*, the more cell assemblies are assumed to fire in a competitive, chaotic, fashion. Using non-linear dynamical (chaos) methods of analysis, an estimate can be made of the amount of competitive firing taking place. Lutzenberger *et al.* hypothesized that higher IQ *S*s would demonstrate greater 'activity' in the EEG than would low IQ *S*s. Of further importance was the fact that conventional frequency analysis of the EEG did not yield any differences between the groups.

During the late 1980s, converging lines of evidence from other measurement paradigms are indicating that there is an enduring relationship between indices of human biological systems and structures and psychometric IQ. For example, nuclear imaging of brain structure and function has produced some preliminary results that suggest that this methodology has a major part to play in the eventual understanding of brain–behaviour relationships. Yeo, Turkheimer, Raz and Bigler (1987) used computerized axial tomography scans of 41 individuals in order to compute brain hemispheric volume. Their results indicated that total brain or total hemispheric volumes were not related to IQ test scores. However, a simple measure of hemispheric asymmetry (left-right hemisphere size) correlated 0.57 with an IQ difference score computed by subtracting WAIS performance IQ score from the verbal IQ score. This correlation indicates that the larger the verbal IQ in relation to performance IQ, the greater the size of the left brain hemisphere to the right hemisphere. Although the correlation between brain asymmetry and the IQ difference score was larger for the male than for the female *S*s the two coefficients were statistically equal in size. Willerman, Schultz, Rutledge and Bigler (1991), using magnetic resonance imaging (MRI) to assess brain size in 40 college students, found that brain size did correlate with psychometric IQ scores at around 0.35, after corrections for body size and a *deflationary* correction for the extreme measurement range of IQ in their sample. Hemispheric asymmetries were computed in a similar fashion to the Yeo *et al.* (1987) method, yielding a significantly different pattern of correlations. These asymmetry correlations (reported in Willerman, Schultz, Rutledge & Bigler, 1992) demonstrated that for males, hemispheric asymmetry correlated 0.44 with the verbal minus performance IQ difference score, replicating the Yeo *et al.* finding. However, for the female *S*s this correlation was −0.55, indicating a larger right hemisphere being associated with verbal performance. In fact, for the female group, it was found that the size of the left hemisphere better predicted non-verbal performance than verbal performance, a finding reversed within the males. Ankney (1992) in a re-examination of brain mass data initially collected from autopsy records by Ho, Roessmann, Straumfjord and Monroe (1980a, b) also demonstrated that brains from males are about 100 g heavier than female brains, correcting for body height and body surface area. These data are interpreted by Ankney in the framework of a general assumption that specific abilities, at which

males and females excel, are related to specific areas or quantities of brain mass. This assumption is given some credence in the review of work in this area by Kimura and Hampson (1992).

Finally, Andreasen, Flaum, Swayze, O'Leary, Alliger, Cohen, Ehrhardt and Yuh (1993), again using MRI imaging to measure the volume of various brain cavities and structures found that full-scale, verbal, and performance WAIS IQ correlated variously with certain brain structure volumes, and with cortical grey matter volume (but not with white matter volume). Correlations computed over 67 normal *S*s averaged about 0.40 between size/volume and IQ. It is interesting to note that the earlier work by Lynn (1989) and others has been validated, in part, by this more recent direct approach at assessing brain size.

Another approach at assessing brain system/structure and IQ relationships has been that using positron emission tomography (PET) to detect radioactive isotope residues within the brain (generally tagged onto glucose), measuring glucose metabolic uptake and location. The 'activation' form of experiment is one where the radioactive substance is given immediately prior to a *S* completing a cognitive or behavioural task. The radioactive isotope tagged-glucose is subsequently metabolized by the specific areas of the brain that are involved in the task, then imaged via the PET methodology. The 'passive' form of experiment is where cognitive or behavioural measures are completed external to the imaging process. Glucose metabolic uptake is effected during resting conditions where an individual is isolated from sound and light while the isotope doped glucose is administered and ultimately imaged. The cerebral spatial uptake of glucose is then correlated with the external variable test scores. De Leon *et al.* (1983) compared 15 young normal *S*s with 22 elderly normal *S*s and 24 Alzheimer patients on WAIS IQ and other tests of cognitive function in a passive format experiment. Correlations of up to 0.6 were found between glucose metabolic rate and IQ, higher cortical activity being associated with a higher metabolic rate. These correlations were computed over a combined normal elderly and Alzheimer patient group, with no significant difference between the elderly and young normal groups. Chase, Fedio, Foster, Brooks, Di Chiro and Mansi (1984) in a similar study essentially replicated these results. However, this form of passive experiment has generally failed to generate consistent brain–behaviour relationships (Duara, Grady, Haxby, Ingvar, Sokoloff, Margolin, Manning, Cutler & Rapoport, 1984; Haxby, Grady, Duara, Robertson-Tchabo, Kozarz, Cutler & Rapoport, 1986; Boivin, Giordani, Berent, Amato, Lehtinen, Koeppe, Buchtel, Foster & Kuhl, 1992). In contrast, the activation study of Haier, Siegal, Nuechterlein, Hazlett, Wu, Paek, Browning and Buchsbaum (1988), using 8 *S*s who completed RAPMs during the glucose uptake period, indicated that higher cortical glucose metabolic rate was related to lower performance on the Matrices. This result was confirmed by Parks, Loewenstein, Dodrill, Barker, Yoshii, Change, Emran, Apicella, Sheramata and Duara (1988) using a test of verbal fluency, and Berent, Giordani, Lehtinen, Markel, Penny, Buchtel, Starosta-Rubenstein, Hichwa and Young (1988) who showed negative correlations between WAIS memory scores and glucose metabolism rate. These activation studies indicate that high IQ *S*s appear to solve problems more 'efficiently' than do low IQ *S*s, requiring less energy to maintain performance at a higher level of accuracy. Haier, Siegel, MacLachlan, Soderling, Lottenberg and Buchsbaum (1992a) tested this hypothesis in an experiment that examined learning of a spatial game task within a group of 8 normal *S*s. It was hypothesized that learning should produce a decrement in cerebral glucose uptake. The *S*s were initially injected with the glucose and PET scanned while playing the game for the first time, they then practised the game constantly over a 2 month period and were then scanned again while playing the game. The results indicated significant widespread reductions in glucose uptake across several regions of the brain. Haier, Siegel, Tang, Abel and Buchsbaum (1992b) extended the analysis to examine whether higher IQ *S*s had greater reductions in glucose uptake than lower IQ *S*s. The pattern of correlations between RAPM, WAIS IQ scores, and metabolic reduction coefficients indicated significant negative relationships, thus confirming the hypothesis. Taking all the activation studies reported above, there are now indications of a consistent negative relationship between glucose uptake rate and cognitive abilities, even if the probable true size of such a relationship cannot yet be estimated.

In addition to the above, human peripheral nerve transmission characteristics have recently become the focus of attention to researchers interested in the biological foundations of intelligence. Two parameters have been examined in some detail, nerve conduction velocity (the speed at which a nerve impulse travels through axons and across synapses), and nerve conduction variability (the

variability in action potential response to a constant stimulus). The first reported work assessing the relationship between human peripheral nerve conduction characteristics and intelligence was that carried out by Vernon and Mori, reported initially at a conference (Vernon & Mori, 1989), and subsequently reported in full in Vernon and Mori (1992). Using data from 85 mixed sex university students, they computed three measures of conduction velocity from the right arm median nerve in the finger–wrist–arm segment. Electrical stimulation of the nerve was supramaximal. A first principal component was computed for the 10 sub-tests of the Multidimensional Aptitude Battery (Jackson, 1984) and separately for the three nerve conduction velocity estimates. Correlating *S*s' scores on each principal component yielded a correlation of 0.42, indicating higher conduction velocity in higher IQ *S*s. In addition, it was found that the wrist-to-finger conduction velocities correlated higher with IQ and reaction time than the velocities computed from the wrist-to-elbow or wrist-to-axilla. Test–retest coefficients based upon 15 of the *S*s over a period from 1 to 3 weeks, indicated high reliability of measurement: 0.86 (wrist-to-finger), 0.81 (wrist-to-elbow), and 0.95 (wrist-to-axilla). Vernon and Mori (1992) reported a replication of their initial experiment using a further 88 mixed sex undergraduate students, following the same measurement procedures. Conduction velocity was computed only between the finger–wrist segment of the median nerve. Once again, a first principal component composite IQ measure was used for correlational purposes. The correlation between conduction velocity and IQ was 0.48. Barrett, Daum and Eysenck (1990), in a study using 44 mixed sex *S*s, examined the relationships between several measures of median nerve transmission characteristics, RAPM score, personality (using the Eysenck Personality Questionnaire, Eysenck & Eysenck, 1975), and 3-bit choice reaction time. Both conduction velocities and action potential variabilities were computed from a segment of nerve between the finger and wrist, on both arms. No significant correlation was found between nerve conduction velocity and intelligence or choice reaction time, rather, it was variability in conduction that correlated -0.44 with IQ. However, a statistically significant correlation of 0.37 between velocity and EPQ Psychoticism was observed. Electrical stimulation of the nerve was between threshold and 2 mA above threshold; effectively 15–20 mA below supramaximal levels. Reed and Jensen (1991) reported a study carried out on 200 male college and university students, assessing nerve conduction velocity of the median nerve in the wrist–elbow segment of a *S*s' preferred hand. The velocities were correlated with scores on Ravens Standard and Advanced Progressive Matrices, and with 3 reaction time tasks (simple, 3-bit, and oddman). Electrical stimulation of the nerve was supramaximal. Arm conduction velocities showed no relationship with IQ and inconsistent relationship to the various measures of reaction time. Test–retest reliability of the conduction velocities was computed using 14 *S*s who were re-assessed from several days to several weeks after their initial test session. This computed reliability was 0.63. In addition to the wrist–elbow conduction velocities, Reed and Jensen computed an estimate of retina to visual cortex brain pathway nerve conduction velocities, computed by measuring latencies of components from visual AEPs. This brain conduction velocity correlated 0.26 with IQ. This particular result is reported in detail in Reed and Jensen (1992). Finally, Barrett and Eysenck (1993) partially replicated the results from their initial 1990 study using the same methodology as before. Only nerve conduction variability correlated significantly with IQ (after adjustments for IQ range compression). However, the correlation between EPQ Psychoticism and nerve conduction velocity was not replicated. Two important factors emerged from this recent study, the variability measure had very low test–retest reliability (0.34), the velocity measures demonstrated an average of 0.83 test–retest reliability.

Some recent work in the neurosciences has indicated that basilar dendrite length and segment count of supragranular pyramidal cells in Wernicke's area are related to age and educational attainment (Jacobs & Scheibel, 1993; Jacobs, Schall & Scheibel, 1993). Using tissue extracted from neuropathology autopsies of 20 individuals aged between 18 and 79 years, it was shown that age correlated negatively with dendritic length up to -0.69. In addition, educational attainment was associated with significantly greater dendritic length and segment counts. These results are very significant in that they are the first indications from human brain tissue that intellectual/environmental enrichment may have a neurohistological effect on such tissue. Since IQ and educational attainment are also correlated at about 0.5 (Kline, 1991), it is also intriguing to wonder if such a process is limited by a gene or gene complex that may underlie intra extracellular nerve conduction properties. In other words, neurohistological enrichment may happen in all our brains,

but one or more 'limiters' of central nervous system (CNS) and cerebral nerve conduction properties affects the efficiency of such nerve pathways. A recent article by Miller (submitted) suggests one such mechanism, the myelination of nerves. From an exhaustive review of the available evidence from histological through to psychological results, he concludes that lack of myelination leads to transmission errors and hence lowering of cognitive functioning as seen in individuals of low IQ. While no new data are offered in this review, the predictions following from it are both clear and testable. In addition, the explanatory power of his model is suggestive that it may be partially correct. Certainly, the work from the various nerve conduction studies reported above seems to lend weight to his arguments. Miller (1992) has also recently published a proposition that the reported correlations between myopia and intelligence are a result of a genetic mechanism that is responsible for brain and eye size.

The secondary purpose of the current analyses reported below is to comprehensively examine a large array of parameters that can be computed from an auditory AEP. As noted above, the studies that have been implemented in this area have only focused on one or two parameters, to the exclusion of the remaining possible indices. This has made it almost impossible to compare results from various studies as few share even one parameter in common. Thus, we have reported the results from an extensive cluster of parameter analyses, cross-replicating our results across three distinct investigations, utilizing the combined data from 200 *S*s. Further, in order to address one feature of the quantum model of intelligence as proposed by Weiss (1986, 1987, 1989, 1992), a zero-cross analysis of the AEPs was undertaken. Weiss has argued that the number of zero-crossings in an AEP, not only correlates with psychometric IQ but also specifies the memory span for any individual. While the quantum model background to support such an assertion is rather complex, Weiss demonstrated that the proposition could be validated quite simply in a re-analysis of Ertl and Schafer's (1969) data.

METHOD

Study 1

Subjects

Twenty nine female and 45 male adult *S*s took part in the study at the Institute of Psychiatry. They were recruited from the local unemployment bureau, advertisements in local newspapers and in the London Mensa newsletter. The age range of the males was from 17 to 56 years, with a mean age of 26.6 and SD of 8.6. The age range of the females was from 18 to 41 years, with a mean age of 26.3 and SD of 6.8.

Psychometric tests

Each *S* completed a computer administered Eysenck Personality Questionnaire (EPQ, Eysenck & Eysenck, 1975) assessing the personality traits of Psychoticism, Extraversion, Neuroticism, and Social Desirability, and a computer administered I_7 questionnaire (Eysenck, Pearson, Easting & Allsopp, 1985) assessing the personality traits of Impulsivity, Venturesomeness, and Empathy. With regard to the assessment of cognitive abilities, the RAPM was administered under speeded conditions, within a 20 min time constraint. Finally, each *S* was also administered the WAIS-R intelligence test (Wechsler, 1981).

Apparatus

Experiment control, stimulus presentation, and data acquisition was 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 and electrode montage

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 msec. The tone envelope was shaped with a rise and fall time of 3 msec, yielding a plateau of 24 msec 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 3 to 8 sec. EEG data was acquired from a single EEG channel, using an electrode placed at the scalp vertex position Cz, referenced against linked mastoid process electrodes. The *S*s was grounded using an electrode placed on the tip of the nose. Electrode impedance of all electrodes was always less than 5 kΩ.

Acquisition details

Prior to each *S* taking part in the experiment, the amplification channels were calibrated using an SLE battery powered oscillator generating a 200 μV peak-to-peak (p-p) sine wave. The signal was continuously sampled and displayed by the SIRIUS in order that the Microlink offset could be centralized manually in order to yield a balanced signal on each channel. 100 epochs of 512 msec duration were acquired via on-line 12-bit A/D. Sampling speed was 1000 Hz. Amplification range was ± 100 μV (200 μV p-p) yielding a measurement resolution of 0.05 μV. The BIODATA PA400 filters were set to yield a frequency bandwidth of 0.8 to 30 Hz per channel (30% signal attenuation at these 'cutoff' frequencies with 90% attenuation at 300 Hz). 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. All experiment data was subsequently transferred to a MASSCOMP M6600 computer for off-line signal processing and analysis.

Procedure

EEG data was acquired from each *S* on each of two consecutive 'working' days (about a fifth of all the *S*s were tested on a Friday and then on the following Monday). The *S*s were administered the WAIS-R on one day, and the RAPM, EPQ, and the I_7 questionnaire on the other day. The order of psychometric test completion was counterbalanced across days. Prior to EEG acquisition, the *S*s completed several reaction time tasks that have been reported elsewhere (Frearson, Barrett & Eysenck, 1988). In all, *S*s were engaged in approx. 2 hr of cognitive test and reaction time tasks prior to the EEG phase of the study. EEG acquisition took place with the *S* sitting in a darkened room, with the tester sitting in an adjoining room monitoring the data acquisition, display oscilloscope, and ST10 stimulator. There was intercom communication between the *S* and tester at all times. The *S*s 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 *S* with their characteristics and amplitude. The tones were then presented to the *S*.

Parameter computation

Prior to parameter computation, each epoch was passed through an amplitude artifact analysis program. This procedure converted all sampling values (0 to 4095) to their microvolt equivalent, then examined each epoch in turn for any values that were not within the voltage range of ± 75 μV. 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 datapoints. 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 the following parameters extracted from this procedure:

AEP component amplitude and latencies. The overall mean absolute amplitude of the AEP was computed initially from a 1–256 msec (*AMP*256) and a 1–512 msec (*AMP*512) epoch length AEP using the formula below:

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

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

Two other amplitude measures were extracted based upon locating the maximum negative waveform voltage between 80 and 140 msec (*N100A*) and the maximum positive waveform voltage observed within the range from the latency of the most negative voltage (*N100L*) and 220 msec epoch length. This positive voltage was labelled *P180A*, with a latency of *P180L*. The peak identification routine was computer implemented and based simply on identification of maximum amplitude values without regard to waveform jitter, noise, or shape. From these parameters, two others were generated:

$$DIFFAMP = P180A - N100A$$

which is the peak-to-peak voltage between the N100 and P180 components.

$$DIFFLAT = P180L - N100L$$

which is the time taken for the voltage to traverse from maximum negativity to maximum positivity, in milliseconds.

Finally from a regression of individual evoked potential mean absolute amplitude over epochs, an intercept and slope parameter was computed respectively for 1–256 msec epochs (*A256* and *B256*), and for 1–512 msec epochs (*A512* and *B512*). For each individual evoked potential epoch that composed the AEP, the mean absolute amplitude of that epoch was computed using the equation given above for AEP amplitude. Whereas that equation used the array of AEP sample voltages, here we use the single epoch voltages. Given K epochs, with a mean absolute amplitude X for each epoch, we regress X on K using the linear, least squares equation:

$$Y = a + bX_K$$

where X_K = the mean absolute amplitude for epoch k of K; a = the intercept parameter; and b = the slope parameter.

AEP contour length (string)

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

where: V = the array of sample voltages defining the AEP; and N = the number of sample points i, over which to make the calculation. Otherwise known as the Hendrickson String parameter, this is a measure of the length of the waveform envelope. The AEP length was measured between 1–256 msec duration (*STR256*) and between 1–512 msec duration (*STR512*).

AEP variability

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

where: v = the voltage at each sample point within an epoch j of K; $\bar{v}_i$ = the mean of the values for sample point i computed across epochs K; K = the number of epochs over which the variances are computed; and N = the number of sample points defining the AEP (256 or 512). Otherwise known as the Hendrickson Variability parameter, this is a measure of the variability of the AEP, assessed by computing the mean variance of each sample point in either a 1–256 msec epoch (*VAR256*) or a 1–512 (*VAR512*) epoch.

The number of epochs retained for averaging (after the amplitude artifact process had been implemented) was also noted. All parameters above were computed over day 1 and day 2 datasets, the differences between all 'occasion' parameters forming a third set of EEG variables.

On visual inspection of some of the AEPs generated across the two occasions, it was apparent that 50 Hz mains noise was contaminating some of the records. In order to remove this noise, and maintain consistency with our previous work in this area (Barrett & Eysenck, 1992a), we used the same linear phase digital FIR filter to smooth each epoch prior to the averaging process. It 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 (68% loss) signal attenuation, at 40 Hz there was 21 dB attenuation (91% loss), at 60 Hz there was 53 dB attenuation. This filter had the effect of removing **all** high frequency 'jitter' and mains noise from within the EEG epoch. Due to the filter process, the epoch length was reduced from 512 to 494 msec. All parameters computed above for the digitally unfiltered data were subsequently recomputed over the digital filtered data.

AEP zero-cross count. For the zero-cross analysis, only the AEPs composed from digitally filtered epochs were used. Secondly, an epoch length of 300 msec was specified in order to remain consistent with Weiss (1992) who argued that the epoch should be long enough to permit identification of a possible P300 component. Zero-cross counts were then computed over the AEP for each *S*, within the range of 1–300 msec. Also, a further set of counts was computed over a mean-detrended AEP. This latter set was essentially a conservative re-adjustment of any baseline DC shift of the AEP. Thus, the first procedure was based upon the crossing from +ve to −ve voltage (or vice-versa) of the AEP trace, not allowing for any DC offset in the AEP (the mean voltage of the AEP was not equal to 0 V). The second procedure removed any DC shift such that the AEP mean voltage was also 0 V. No attempt was made to derive peak components from the zero-cross intervals, rather a simple count was generated.

The possible difference between the results produced by the two procedures is extremely significant. The problem with zero-crossing parameters, especially in the manner in which Weiss is attempting to use them, is that any oscillatory noise within the AEP waveform around 0 V can significantly enhance the zero-cross count. It is to be noted that for his analyses, Weiss (1992) is using the figural data first reported in Ertl and Schafer (1969), based upon EEG acquired within a 3–50 Hz bandwidth. Also, in contrast to the studies reported here, these AEPs were generated using a visual light-flash stimulus. The quantum model of intelligence as proposed by Weiss makes no distinction between quantal state transitions evoked by different sensory events and their relationship to psychometric IQ.

Study 2

Subjects

Sixty one female and 25 male adult *S*s took part in the study at the Biosignal Laboratory, Institute of Psychiatry. They were recruited from the local area via advertisements in local newspapers. The age range of the males was from 19 to 49 years, with a mean age of 31.6 and SD of 9.8. The age range of the females was from 18 to 53 years, with a mean age of 36.2 and SD of 9.2.

Psychometric tests

All *S*s completed the EPQ Revised (EPQR, Eysenck, Eysenck & Barrett, 1985) which assesses the same traits as the EPQ. In addition, the I_7 questionnaire was also completed. Both administrations were via paper and pencil. Psychometric intelligence was assessed using the Jackson Multidimensional Aptitude Battery (MAB, Jackson, 1984), a group-administerable, speeded analogue of the WAIS-R.

Apparatus

EEG acquisition was effected using a MASSCOMP M6600 computer system, sampling across 7 channels at 1000 Hz, using sample and hold continuous-duty 12-bit A/D sampling (continuous, parallel-channel sampling, with no multiplexing error across channels). 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 an HP VECTRA PC. The tones were delivered binaurally via Hills SH-22 lightweight electromagnetic headphones. The headphone-delivered SPL was checked using a Brüel and Kjaer model 2231 sound level meter with microphone type 4155. The IT task was implemented via an HP VECTRA 286 PC controlling a custom-built stimulus and timing unit. The IT stimulus is composed of multiple

segment red bar LEDs arranged in the form of an inverted U. Each multiple segment LED display consists of 5 LED bars of length 30 mm and width 1 mm. The luminance rise time of the LEDs is ~100 nsec with an illuminance of 0.6 microcandela. The LEDs are flush mounted into a matt black panel that forms the front face of the stimulus box. All stimulus timing is accurate to within 1 msec, with all timing effected using hardware rather than software triggers. Two microswitch buttons on 3 m length cables are connected to the back of the stimulus box; these buttons are used by a *S* to indicate a response. The LED bar display can be energized in three standard ways, showing a longer red bar on the right side, the left side, and a mask that energizes all LED bars that are not illuminated as part of the stimulus.

The RT task was implemented by the same computer, using another custom-built stimulus and timing unit. Precise details of both sets of apparatus are given in Barrett and Eysenck (submitted). Skin conductance level (SCL) was recorded using two Ag–AgCl electrodes attached to the distal phalanges of the middle finger of the non-preferred hand. These electrodes were attached to an Electronic Developments skin conductance meter, calibrated in mSiemens, with a 0.5 V potential maintained across the electrode pair. Finally, EEG was recorded using an ElectroCap™ with embedded 9 mm Sn disc electrodes and specially formulated electrode paste for the tin electrodes.

Stimulus characteristics and electrode montage

A 1000 Hz digitally-synthesized sine wave was generated on demand by the ST10 stimulator. The amplitude of the tone was 85 dB, the total duration was 30 msec. The tone envelope was shaped with a rise and fall time of 3 msec, yielding a plateau of 24 msec at maximum amplitude. Signal onset and offset were always at 0 V. The randomized interstimulus interval was between 1 and 4 sec. The electrode montage is given in Fig. 1 below. The numbers in italics correspond to the channel numbers used occasionally in the reporting of the results. The reference electrodes were also 9 mm Sn discs, formed into an ear clip that fitted onto each earlobe. These earlobe electrodes were linked within the custom headbox. The earth/ground electrode was located midway between Fpz and Fz locations within the 10–20 system.

Acquisition details

Prior to each *S* taking part in the experiment, the 7 EEG amplification channels were calibrated using an HP 3325A signal generator providing a 300 μV p-p sine wave. This represented the maximum bandwidth of the EEG channels for recording purposes. Thus, the amplification range was $\pm 150\,\mu$V (300 μV p-p) yielding a measurement resolution of 0.07 μV. Sampling speed was 1000 Hz. The BIODATA PA400 filters were set to yield a frequency bandwidth of 0.2 to 100 Hz per channel (30% signal attenuation at these 'cutoff' frequencies with 90% attenuation at 1000 Hz). Anti-aliasing hardware filters were present on the MASSCOMP signal-conditioning A/D board.

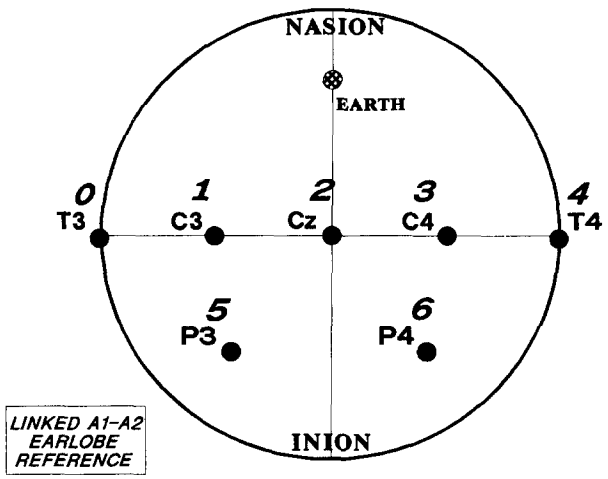

Fig. 1. Electrode montage for study 2.

Two further channels of data were recorded, one for SCL and one for the TTL pulse defining auditory stimulus onset from the ST10 signal generator. For convenience and precise stimulus locking, these channels were both sampled at 1000 Hz as part of the entire incoming data ensemble.

Procedure

All *S*s were initially administered the EPQR, I_7, and MAB IQ tests at group testing sessions several months prior to the laboratory-based tasks. On the day of the EEG acquisition, a *S* was initially administered 20 trials of a 3-bit, 8-light, choice RT task followed by a 20 trial Odd-Man-Out task (Frearson & Eysenck, 1986). The total task duration for these two RT strategies was approx. 8 min. Immediately after these tasks, the *S* completed an adaptive IT paradigm that lasted between 4 and 12 min (depending upon how quickly the algorithm resolved a *S*s' IT). The precise details of these tasks are given elsewhere (Barrett & Eysenck, submitted). After these chronometric tasks, the *S*s' hands were washed with warm water, but no soap. Then the *S* was seated in a comfortable chair in the EEG acquisition room, facing another IT stimulus array, at a distance of about 6 feet from the array. Two skin conductance electrodes were placed on the distal phalanges of the *S*s' middle finger of the non-preferred hand. The *S* was grounded to the skin conductance meter via a 9 mm Ag–AgCl electrode placed on the back of the hand to which the active electrodes were affixed. SCL was recorded within the range of $\pm 10\ \mu$S, with a measurement sensitivity of $0.005\ \mu$S. The ElectroCap™ was then placed over the *S*s' scalp. Skin preparation was confined to the use of blunt needle abrasion, effected through each of the EEG electrodes. The reference electrodes were then placed on the earlobes. All electrode impedances were maintained below 5 kΩ.

Following the application of all recording electrodes, the IT response buttons were placed onto the arms of the chair in which the *S* was sitting. The *S* was then given a set of verbal instructions:

> "You will close your eyes and relax in the chair. After a few minutes, you will hear some tones through your headphones. When these stop, we will ask you to open your eyes and begin the inspection time task. After this task has finished, we will tell you to close your eyes again and relax. Once again, after a few minutes you will hear some more tones. When these finish, the experiment is complete."

It was stressed that a *S*s' eyes must remain closed except for when they were engaged in the IT task. In addition, there must be no talking or excess movement except under exceptional circumstances or replying to the investigator's instructions. Since the *S* had already just completed the IT task in another test room (with the RT tasks), specific instructions or training were not required. Given acknowledgement of these basic instructions by the *S*, the investigator placed the headphones over the *S*'s ears, turned off the lights in the *S* room, and closed the door to the room in which the *S* now sat alone. The room was lit only by an indirect, reflected, light source that yielded an effective illuminance level of about 2.4 lx measured at the face of the IT stimulus box using a Minolta T-1 illuminance meter. From the acquisition control room, the *S* was monitored visually using a low-light infrared CCTV camera, and auditorily using a microphone placed near the *S*. The *S* was then asked by the investigator to confirm they were ready. On this confirmation, the microphone link from the investigator to the *S* was cut. The SCL level was balanced on $0\ \mu$S over a period of about 30 sec (in order to allow the SCL to stabilize a little). Then the experiment control program was started on the HP VECTRA PC and the MASSCOMP began continuously acquiring data from the 7 EEG channels, SCL, and tone pulse channel. Three minutes of silence was followed by the 100 tones, lasting a further $3\frac{1}{2}$ min or so (1 to 4 sec randomized interstimulus interval). When the tones were finished, the MASSCOMP ceased recording and the *S* was asked to open his/her eyes and place a hand over the respective left/right response buttons for the IT task. After a few seconds, the IT task was started on another remote PC. The MASSCOMP also began recording EEG activity again. This IT task was time-limited to 5 min duration, if the IT had not been resolved in this time, then the task was halted. On the termination of the task, the MASSCOMP stopped recording, and the *S* was asked to remove their hands from the response buttons, close their eyes, and relax again. After a few seconds, the experiment control program was re-initiated and the MASSCOMP once again started continuous recording. As before, 3 min of

silence was followed by another 100 tones, lasting a further $3\frac{1}{2}$ min. The MASSCOMP ceased recording 3 sec after the last tone.

Thus, we have two occasions on which AEPs were generated using tone stimuli, with a cognitive task intervening between occasions. For the purposes of analysis below, we shall refer to the two AEP sessions as occasion 1 and 2. All data on the MASSCOMP was digitally stored, yielding approx. 20 Mb of data for each *S*. All subsequent signal processing analyses were implemented off-line.

Parameter computation

The same parameters computed for Study 1 were also computed on this data. Since the evoked potentials were embedded in a continuous stream of EEG signals, the tone 'pulse' channel provided the stimulus onset signals from which the 100 512 msec epochs could be extracted from the 7 EEG channels. The detailed analyses of the spontaneous EEG, cortical muscle potentials and SCL data, are not reported here. Rather, this report confines itself to an examination of AEP parameters only.

The data were recorded with an effective bandwidth between 0.1 and 200 Hz. The Biodata PA400 hardware filters have very gradual pass-to-stopband transition functions, hence at a measured 2.6 dB cutoff at 100 Hz, there is only 26% attenuation of the signal, at 150 Hz, there is 52% loss, and at 200 Hz, there is 69% loss. Since high frequency (>60 Hz) cortical muscle potentials can have amplitudes as great at 100 μV p-p, even a 70% attenuation will leave a 30 μV p-p waveform in the predominantly low frequency EEG waveform. Thus it was decided that for AEP analysis, all epochs would again be digitally filtered prior to averaging. Instead of using the filter that was used in Study 1, it was decided to design a new filter that preserved more of the upper EEG frequencies while severely attenuating frequencies above about 50 Hz. The FIR filter was designed on the MASSCOMP as a 44th order low-pass filter with a near flat passband from 0 to 30 Hz and a stopband from 90 to 500 Hz. At 50 Hz, there was about 5 dB (44% loss) signal attenuation, at 60 Hz there was 10 dB attenuation (68% loss), at 80 Hz there was 20 dB attenuation (90% loss), and at 100 Hz, attenuation was 70 dB (virtually 100% loss). This filter had the effect of removing **all** high frequency 'jitter' and mains noise from within the EEG epoch. Due to the filter process, the epoch length was reduced from 512 to 490 msec. All parameters computed above for the digitally unfiltered data were subsequently recomputed over the digital filtered data.

RESULTS

All AEP parameters used in the reported analyses were drawn from the digitally filtered epochs. Correlational analyses for studies 1 and 2 indicated no conceptually or statistically significant relationships between the AEP variables and psychometric IQ, on either of the test occasions or, in the case of study 2, across all 7 measurement channels. This is in marked contrast to our recently published study (Barrett & Eysenck, 1992a) replicating the Hendrickson variability correlation with IQ. From the Introduction above, demonstrating that investigators had difficulty replicating the essential IQ $\times$ AEP parameter correlations, it now seemed that we could not even replicate our prior results.

However, during the acquisition of *S* data in study 2, conceptually significant correlations of up to about -0.30 between IQ and AEP variability, intercept, and latency parameters were observed within the first 30–40 *S*s' data. These correlations subsequently disappeared with the addition of another 10 or so *S*s. Looking closely at the first 40 *S*s' data for outlier effects that might be producing such correlations, it was apparent that there were no obvious outliers. What was apparent though was the number of AEPs that did not have a well-defined P180 positive component (the N100-P180 components of the tone or click elicited auditory AEP are effectively a signature of this particular waveform, Regan, 1989). That is, the component was not clearly identifiable as a single-peak or was of such low amplitude in comparison to the N100 component that it could not be *visually* identified as a peak at all. Given the use of objective, computer-based, maximum-value peak identification, the maximum value within a window will always be found. However, this is not to say that the maximum value can be said to thus identify a conceptually significant peak, such as the P180. If a peak is smeared or flattened or bipolar, only visual confirmation or some very sophisticated computer-based rules would show that the 'peakedness'

Table 1. The means, SDs, and ranges for the various samples and sub-samples used in the reported analyses

		Verbal			Perform			Total		
Sample	N*	Mean	SD	Range	Mean	SD	Range	Mean	SD	Range
B & E (1992)—Total (WAIS)	40	104.4	18.2	61–142	104.4	18.5	62–144	105.2	19.0	61–134
Study 1, Total (WAIS)	74	107.9	15.0	75–137	103.5	13.4	69–140	106.6	14.5	74–134
Study 2, Total (MAB)	86	110.8	13.0	80–137	109.6	13.9	70–137	110.4	12.8	78–132
Selected *Ss*, B & E (1992)	37	103.5	18.3	61–142	103.6	18.2	61–144	104.3	18.9	61–134
Selected *Ss*, study 1	49	109.3	14.5	75–137	103.8	12.9	74–134	107.6	13.9	81–134
Non-selected *Ss*, study 1	25	105.0	16.0	78–129	103.0	14.5	69–140	104.6	15.7	74–132
Selected *Ss*, study 2, channel Cz (2)	38	111.7	12.5	84–136	109.4	15.2	70–137	110.8	13.1	78–131
Non-selected *Ss*, study 2, channel Cz (2)	48	110.2	13.6	80–137	109.8	13.1	81–130	110.1	12.6	87–132
Common-subset, selected *Ss*, study 2	26	110.3	12.3	84-129	106.3	14.7	70–133	108.7	13.0	78–130
Common-subset, non-selected *Ss*, study 2	31	111.4	15.0	80–137	111.5	12.5	86–129	111.5	13.2	87–132

*N = number of *Ss*.

is of poor quality or abnormally low amplitude. On the basis of a simple hypothesis that there was something 'wrong' about AEP traces with a *subjectively* ambiguous P180, an *objective* screening procedure was implemented via computer, using the simple criterion of rejecting all *S*s' with a P180 amplitude less than the mean amplitude of the P180 component computed over the total *S* dataset i.e. having computed the mean P180 component amplitude over all *S*s a filter was created whereby any *S* with a P180 amplitude less than the mean value was rejected from further analysis. This screening procedure was carried out for all *S*s, across all channels, and across both test occasion datasets. In addition, the same procedure was carried out on the data from study 1, and the data from the earlier published work reported in Barrett and Eysenck (1992a). For these latter two studies, it was noted that the boundary limit could be set at 1 SD below the mean without substantially affecting the resultant parameter correlations. This reflected the subjective observation that the data from these two studies appeared to be of a better 'quality' with regard to the definition of the N140-P180 segment of the AEP. Finally, a secondary computer-based screening procedure then took place over a single AEP parameter. This was the 256 msec variability measure. Where the IQ × variability scatterplot for each dataset indicated possible 'out-of-range' outliers (as defined by the placement of a 95% confidence ellipse around the scatterplots) these *S*s were also removed from all parameter analyses. It is this parameter that is the most sensitive indicator of intraindividual epoch disparity. In all, 18 distinct datasets were analysed in the above specified manner. Each analysis yielded correlational results between the same AEP parameters and IQ.

Before reporting the correlations, it is of use to examine the concept of statistical significance under these conditions. Given each analysed reduced-data matrix consists of 3 IQ variables and 16 AEP parameters, a total of 864 correlations were computed (the total number computed using all pertinant variables is closer to 5760). Statistical significance in the accepted sense of the word is almost meaningless under these conditions. Rather, the pattern, structure, and replicability of correlations is of far greater significance than attempting to arbitrarily correct correlations for the number of coefficients computed. In addition, given these coefficients are being computed from *post-hoc* selected data, the whole concept of a significance test is questionable. With regard to the reduction in sample sizes due to the selection conditions, the size of correlation coefficients can fluctuate dramatically as the sample size is reduced. However, there is a somewhat mistaken belief that this fluctuation is directional in that the changes in size of the correlation coefficients is related directly to the sample size. This is incorrect. The reasons for such changes have more to do with the sampling of values on the two variables in question rather than any simple function of the number of observations. Of course, as the sample size is decreased, so the likelihood of capitalization on chance errors in sampling increases. However, these errors can cause over- or underestimation of the parameters, they are not specifically directional except by chance. So, to recap, the subset data is to be examined more for its replicability of coefficient direction and size than notional statistical significance.

Table 1 presents the main distribution parameters for standardized IQ within the total and subset samples used in all analyses below. The verbal and performance sub-tables represent the verbal and performance composite IQ scores as assessed by both the MAB and WAIS IQ tests. Both sets of test scores are standardized with means of 100 and SDs of 15. As can be seen from this table, only

Table 2. Correlations between Psychometric IQ and 18 sets of observations across scalp electrode and study, string/contour length

	256 msec epoch			490/494 msec epoch		
	Verbal IQ	Perf. IQ*	Fullscale IQ	Verbal IQ	Perf. IQ	Fullscale IQ
Vertex electrode						
B & E (1992), A1-Cz, $n = 37$	−0.10	−0.17	−0.14	−0.14	−0.21	−0.18
B & E (1992), A2-Cz, $n = 37$	−0.05	−0.21	−0.13	−0.08	−0.23	−0.15
Study 1, day 1, [A1 + A2]-Cz, $n = 49$	−0.11	0.21	0.02	−0.12	0.20	0.09
Study 1, day 2, [A1 + A2]-Cz, $n = 49$	−0.07	0.30	0.09	−0.09	0.31	0.09
Study, day 2, Occ 1, [A1 + A2]-Cz, $n = 38$	−0.12	−0.22	−0.18	−0.18	−0.25	−0.22
Study 2, Occ 2, [A1 + A2]-Cz, $n = 40$	0.09	−0.12	−0.02	0.02	−0.14	−0.06
Study 2, left hemisphere [A1 + A2]						
Temporal, Occ 1, T3, $n = 47$	−0.01	−0.04	−0.03	−0.07	−0.09	−0.09
Temporal, Occ 2, T3, $n = 37$	−0.27	−0.28	−0.29	−0.27	−0.28	−0.29
Mid-temporal/vertex, Occ 1, C3, $n = 39$	−0.11	−0.12	−0.11	−0.16	−0.13	−0.16
Mid-temporal/vertex, Occ 2, C3, $n = 39$	−0.00	−0.09	−0.05	−0.05	−0.07	−0.06
Parietal, Occ 1, P3, $n = 39$	−0.13	−0.12	−0.13	−0.21	−0.16	−0.19
Parietal, Occ 2, P3, $n = 39$	−0.04	−0.05	−0.05	−0.19	−0.10	−0.16
Study 2, right hemisphere [A1 + A2]						
Temporal, Occ 1, T4, $n = 41$	−0.33	−0.28	−0.33	−0.29	−0.23	−0.28
Temporal, Occ 2, T4, $n = 38$	0.03	−0.03	−0.01	−0.01	−0.05	−0.04
Mid-temporal/vertex, Occ 1, C4, $n = 42$	−0.07	−0.15	−0.12	−0.07	−0.14	−0.11
Mid-temporal/vertex, Occ 2, C4, $n = 38$	−0.10	−0.14	−0.12	−0.16	−0.19	−0.18
Parietal, Occ 1, P4, $n = 38$	−0.18	−0.21	−0.21	−0.19	−0.18	−0.20
Parietal, Occ 2, P4, $n = 37$	−0.17	−0.12	−0.11	−0.17	−0.16	−0.18

*Perf. IQ = performance IQ.

the B & E (1992)* data has excessively larger SDs than expected. This, as reported and discussed in the 1992 paper, is due mainly to the inclusion of the *S* with IQ 61. Removal of this *S* made little or no difference to any correlational relationships but did reduce the Total WAIS SD by 1.2 points. In all datasets, the mean IQ is higher than the expected average.

Table 2 presents the correlations of the AEP string/contour length parameter with psychometric IQ. The unambiguous result from this table of correlations is that no significant positive correlation between this parameter and IQ is present. The direction of most of the coefficients is negative—reflecting the overwhelming negative-direction correlations between AEP amplitude and IQ that is found within our various datasets; the string length parameter correlating consistently within the range of 0.60 and 0.85 with mean absolute AEP amplitude, computed over digital filtered EEG data. This positive correlation was initially demonstrated by Haier, Robinson, Braden and Williams (1983) and Haier, Robinson, Braden and Krengel (1984). An important feature of the analysis is also indirectly indexed in the above table, this feature casts doubt upon the recent hypothesis put forward by Bates and Eysenck (1993) regarding the relationship between the string measure, intelligence, and attention. They reported significant negative correlations (up to −0.61) between the string measure and IQ, based upon *normalized* bimodal stimulus AEPs acquired from *S*s taking part in an IT task. Their explanation for these negative correlations was based upon a proposition that cognitive load/attention mediated the correlations between the string measure and IQ such that under high load/attending conditions, correlations betueeh these parameters would be negative (indexing neural energy). Where cognitive load was low, the measures would be expected to correlate positively (indexing neural capacity). In addition to the data they presented in their paper, they also referenced the Barrett and Eysenck (1992a) paper that demonstrated a high negative correlation (−0.44) between total WAIS IQ and 256 msec epoch string length using *unfiltered* EEG. They claim that *S*s were 'attending' to the stimuli after being given the instructions to 'listen to the tones'. However, the important point made in this particular paper was that when using *filtered* EEG, the correlation dropped to −0.19. As can be seen from Table 2, removing 3 *S*s from this dataset yields a correlation of −0.14 for the Cz-A1 channel data. Noticeably, in the digitally unfiltered data, string and mean absolute waveform amplitude correlated 0.34. In the filtered dataset, this increased to 0.79. For study 2, Cz channel, digitally unfiltered data, the correlation between 256 msec epoch string and amplitude was 0.12 and 0.15 for occasion 1 and 2 datasets, respectively (86 *S*s). In the digitally filtered datasets, these, correlations increased to 0.77 and 0.71, respectively (86 *S*s). Within the selected study 2 datasets used above in Table 2, these correlations were 0.74 and 0.71, respectively (38 and 40 *S*s). The sensitivity of the string measure correlations

*B & E (1992) refers to the data from the Barrett and Eysenck paper published recently.

to the filtering operation is suggestive not of attentional effects but of high frequency (>45 Hz), low amplitude, 'noise' on the waveform envelope. The amplitude × IQ correlations as shown in Barrett and Eysenck (1992a) and in Table 4 are far more stable indicators of any EEG × IQ relationships. For example, in Barrett and Esenck (1992a), the filtered EEG 256 msec epoch, Cz-A1 channel, amplitude × IQ correlation was −0.42, for the unfiltered data it was also −0.42. Noting the sizeable correlations in study 2, channel T4, occasion 1, the filtered EEG string × IQ correlation is −0.33. For the unfiltered data from the same 'selected' *S*s* it is +0.28. For the same dataset, correlating mean absolute amplitude and IQ, the correlations are −0.46 and −0.33 respectively. The correlation between string and amplitude in the unfiltered data is −0.09, in the filtered data, 0.65. It would appear that the string measure is actually a confounded amplitude measure, the confounding being due to noise/jitter present on the envelope of the AEP. Of course, whether or not waveform amplitude is indexing attention is another point, one that is partly addressed in this paper and more specifically in Robinson (1993).

A further point is worthy of discussion here. Bates and Eysenck normalized the AEP on the basis that such normalization would yield string measures independent of AEP amplitude. As Barrett (1988) has previously pointed out, non-linear transformations of data can have unpredictable consequences on AEP parameters. In order to demonstrate the consequence of this transformation on our data, we computed the string and mean absolute AEP amplitude (MAmp) over normalized AEPs for occasion 1 data from studies 1 and 2, channel Cz (the total length AEP was normalized in each case, then we computed the parameters over a 256 msec and 490/494 msec epoch). For study 1, non-normalized AEP, 74*S*s, the 256 msec string vs MAmp correlation was 0.91, its normalized data counterpart was 0.34. For the 494 msec AEP, the non normalized data correlation was 0.82, its normalized counterpart was −0.46. In the non-normalized data, the 256 and 494 msec MAmp correlated 0.93, its normalized counterpart was −0.40. For study 2, the same procedure was followed, using all 86 *S*s. The non-normalized data, 256 msec string vs MAmp correlation was 0.77, its normalized data counterpart was −0.24. For the 490 msec AEP, the non-normalized data correlation was 0.69, its normalized counterpart was 0.15. In the non-normalized data, the 256 and 494 msec MAmp correlated 0.95, its normalized counterpart was −0.21. These correlations do not support the assumption that normalizing an AEP will automatically produce statistical independence between string and amplitude measures.

Since Bates and Eysenck do not explicitly state how they normalized the AEPs, there is the possibility that rather than standardize each AEP to its own mean and SD, they may have standardized each sample point on their total epoch across all *S*s. That is, for each sample point, the mean and SD was computed over the array of values given by all the *S*s. Then for every *S*'s AEP, each sample point was standardized using its particular mean and SD. This procedure was implemented accordingly on study 2 data, channel Cz, occasion 1. Given the non-normalized data correlations between MAmp and string of 0.77 and 0.69 for the 256 and 490 msec epochs, respectively, the cross-*S* normalized data yielded correlations of 0.44 and 0.45. Thus, even using this methodology, the string and amplitude measure are still related to a non-trivial degree, albeit this relationship is almost halved. This is not to say that Bates and Eysenck's parameters were not independent within their particular dataset, just that since no analysis was reported in their paper, the possibility remains that their string measure was still confounded with amplitude. However, as is evident from the correlations above, the normalization of the AEP by either method does seem to be producing unpredictable correlational effects amongst amplitude and string measures. Given the analysis above and the comprehensive analysis within Robinson (1993), it does seem unlikely that the string measure can ever be independent of the amplitude of the waveform envelope.

Table 3 presents the correlations of the AEP variability parameter with psychometric IQ. The pattern of correlations across study, channel, and occasion, suggests a consistent, significant correlation between variability and IQ. Of all the channels, it is only the T3 electrode channel in study 2 which seems to provide no relationship between the parameter and IQ. What is more surprising is that the hemispherically opposite electrode T4 does seem to contain a systematic relationship between these variables. Electrode positions T3 and T4 are located just above the ears,

*Selected and non-selected *S*s refers to those *S*s whose data passed or failed, respectively the two screening procedures. Within study 2, different *S* can compose the subset data for each channel and test occasion.

Table 3. Correlations between Psychometric IQ and 18 sets of observations across scalp electrode and study, variability

	256 msec epoch			490/494 msec epoch		
	Verbal IQ	Perf. IQ	Fullscale IQ	Verbal IQ	Perf. IQ	Fullscale IQ
Vertex electrode						
B & E (1992), A1-Cz, $n = 37$	−0.51	−0.33	−0.45	−0.50	−0.28	−0.42
B & E (1992), A2-Cz, $n = 37$	−0.39	−0.25	−0.34	−0.37	−0.20	−0.30
Study 1, day 1, [A1 + A2]-Cz, $n = 49$	−0.39	−0.38	−0.43	−0.44	−0.42	−0.48
Study 1, day 2, [A1 + A2]-Cz, $n = 49$	−0.31	−0.16	−0.28	−0.32	−0.20	−0.30
Study 2, Occ 1, [A1 + A2]-Cz, $n = 38$	−0.52	−0.36	−0.46	−0.52	−0.37	−0.46
Study 2, Occ 2, [A1 + A2]-Cz, $n = 40$	−0.29	−0.22	−0.27	−0.29	−0.23	−0.27
Study 2, left hemisphere [A1 + A2]						
Temporal, Occ 1, T3, $n = 47$	−0.08	0.01	−0.04	−0.08	0.01	−0.04
Temporal, Occ 2, T3, $n = 37$	−0.41	−0.21	−0.33	−0.41	−0.22	−0.34
Mid-temporal/vertex, Occ 1, C3, $n = 39$	−0.48	−0.33	−0.43	−0.47	−0.33	−0.42
Mid-temporal/vertex, Occ 2, C3, $n = 39$	−0.36	−0.25	−0.32	−0.35	−0.23	−0.31
Parietal, Occ 1, P3, $n = 39$	−0.60	−0.50	−0.57	−0.59	−0.50	−0.57
Parietal, Occ 2, P3, $n = 39$	−0.44	−0.31	−0.39	−0.42	−0.29	−0.38
Study 2, right hemisphere [A1 + A2]						
Temporal, Occ 1, T4, $n = 41$	−0.43	−0.40	−0.44	−0.43	−0.40	−0.45
Temporal, Occ 2, T4, $n = 38$	−0.19	−0.21	−0.21	−0.18	−0.21	−0.20
Mid-temporal/vertex, Occ 1, C4, $n = 42$	−0.38	−0.28	−0.35	−0.38	−0.29	−0.35
Mid-temporal/vertex, Occ 2, C4, $n = 38$	−0.41	−0.36	−0.41	−0.41	−0.38	−0.41
Parietal, Occ 1, P4, $n = 38$	−0.56	−0.36	−0.49	−0.55	−0.35	−0.48
Parietal, Occ 2, P4, $n = 37$	−0.27	−0.19	−0.24	−0.28	−0.22	−0.26

directly over the temporalis muscles which are implicated in jaw movement and teeth clenching. These two electrodes also recorded an apparently continuous stream of high frequency muscle potentials for most *S*s throughout the experiment in study 2. The modulation of this activity appears to be an individual difference parameter in its own right and is the subject for further analysis within the laboratory.

Table 4 presents the correlations of the AEP mean absolute amplitude parameter with psychometric IQ. As pointed out above, this table of correlations demonstrates the general negative relationship found between amplitude and IQ within our data. However, the small size of most of the correlations is not suggestive of a strong relationship with IQ. It is quite possible that the measurement of most of the epoch is not an efficient parameter estimate in that too much 'random' noise is being added into the parameter. Noticeably Haier *et al.* (1983, 1984) achieved their correlations on a subset of the AEP waveform epoch (from 140 to 200 msec) or using specific component amplitudes. In our studies, AEP component amplitudes N100 and P180 were generally uncorrelated with IQ, as was the peak-to-peak voltage parameter DIFFAMP referred to in the Method section above.

A further measure of AEP amplitude is the AEP amplitude regression intercept as defined above

Table 4. Correlations between Psychometric IQ and 18 sets of observations across scalp electrode and study, absolute amplitude of AEP

	256 msec epoch			490/494 msec epoch		
	Verbal IQ	Perf. IQ	Fullscale IQ	Verbal IQ	Perf. IQ	Fullscale IQ
Vertex electrode						
B & E (1992), A1-Cz, $n = 37$	−0.34	−0.41	−0.39	−0.47	−0.49	−0.50
B & E (1992), A2-Cz, $n = 37$	−0.24	−0.37	−0.31	−0.37	−0.43	−0.41
Study 1, day 1, [A1 + A2]-Cz, $n = 49$	0.00	0.27	0.13	0.08	0.32	0.20
Study 1, day 2, [A1 + A2]-Cz, $n = 49$	−0.01	0.33	0.15	−0.04	0.36	0.14
Study 2, Occ 1, [A1 + A2]-Cz, $n = 38$	−0.15	−0.24	−0.20	−0.16	−0.22	−0.20
Study 2, Occ 2, [A1 + A2]-Cz, $n = 40$	0.13	−0.10	0.02	0.06	−0.08	−0.00
Study 2, left hemisphere [A1 + A2]						
Temporal, Occ 1, T3, $n = 47$	0.18	0.10	0.14	0.23	0.12	0.19
Temporal, Occ 2, T3, $n = 37$	0.00	−0.16	−0.08	−0.09	−0.25	−0.18
Mid-temporal/vertex, Occ 1, C3, $n = 39$	−0.17	−0.27	−0.23	−0.19	−0.27	−0.24
Mid-temporal/vertex, Occ 2, C3, $n = 39$	0.10	−0.12	−0.01	0.10	−0.11	0.00
Parietal, Occ 1, P3, $n = 39$	−0.10	−0.13	−0.12	−0.06	−0.05	−0.06
Parietal, Occ 2, P3, $n = 39$	0.18	0.19	0.20	0.08	0.20	0.15
Study 2, right hemisphere [A1 + A2]						
Temporal, Occ 1, T4, $n = 41$	−0.41	−0.46	−0.46	−0.34	−0.41	−0.40
Temporal, Occ 2, T4, $n = 38$	−0.07	−0.24	−0.16	−0.08	−0.22	−0.15
Mid-temporal/vertex, Occ 1, C4, $n = 42$	−0.16	−0.22	−0.19	−0.11	−0.18	−0.15
Mid-temporal/vertex, Occ 2, C4, $n = 38$	−0.14	−0.16	−0.15	−0.18	−0.15	−0.17
Parietal, Occ 1, P4, $n = 38$	−0.22	−0.47	−0.37	−0.19	−0.39	−0.31
Parietal, Occ 2, P4, $n = 37$	−0.05	−0.23	−0.13	−0.17	−0.24	−0.21

Table 5. Correlations between Psychometric IQ and 18 sets of observations across scalp electrode and study, epoch ampl. regression intercept

	256 msec epoch			490/494 msec epoch		
	Verbal IQ	Perf. IQ	Fullscale IQ	Verbal IQ	Perf. IQ	Fullscale IQ
Vertex electrode						
B & E (1992), A1-Cz, $n = 37$	−0.51	−0.42	−0.49	−0.57	−0.43	−0.53
B & E (1992), A2-Cz, $n = 37$	−0.33	−0.34	−0.35	−0.41	−0.36	−0.40
Study 1, day 1, [A1 + A2]-Cz, $n = 49$	−0.18	0.07	−0.09	−0.25	−0.04	−0.19
Study 1, day 2, [A1 + A2]-Cz, $n = 49$	−0.23	0.14	−0.09	−0.31	−0.01	−0.21
Study 2, Occ 1, [A1 + A2]-Cz, $n = 38$	−0.53	−0.37	−0.47	−0.56	−0.41	−0.51
Study 2, Occ 2, [A1 + A2]-Cz, $n = 40$	−0.25	−0.29	−0.28	−0.27	−0.26	−0.28
Study 2, left hemisphere [A1 + A2]						
Temporal, Occ 1, T3, $n = 47$	0.01	0.08	0.04	−0.05	0.04	−0.01
Temporal, Occ 2, T3, $n = 37$	−0.31	−0.23	−0.28	−0.26	−0.18	−0.23
Mid-temporal/vertex, Occ 1, C3, $n = 39$	−0.46	−0.33	−0.41	−0.48	−0.34	−0.43
Mid-temporal/vertex, Occ 2, C3, $n = 39$	−0.25	−0.27	−0.27	−0.26	−0.22	−0.25
Parietal, Occ 1, P3, $n = 39$	−0.52	−0.35	−0.45	−0.52	−0.35	−0.45
Parietal, Occ 2, P3, $n = 39$	−0.38	−0.30	−0.36	−0.40	−0.27	−0.36
Study 2, right hemisphere [A1 + A2]						
Temporal, Occ 1, T4, $n = 41$	−0.58	−0.55	−0.60	−0.57	−0.52	−0.58
Temporal, Occ 2, T4, $n = 38$	−0.17	−0.30	−0.25	−0.11	−0.23	−0.18
Mid-temporal/vertex, Occ 1, C4, $n = 42$	−0.46	−0.33	−0.42	−0.47	−0.35	−0.44
Mid-temporal/vertex, Occ 2, C4, $n = 38$	−0.45	−0.46	−0.48	−0.45	−0.46	−0.47
Parietal, Occ 1, P4, $n = 38$	−0.58	−0.38	−0.51	−0.54	−0.31	−0.45
Parietal, Occ 2, P4, $n = 37$	−0.30	−0.29	−0.30	−0.33	−0.29	−0.32

in the Method section. Table 5 presents the correlations of this parameter with psychometric IQ. This table provides clear evidence for a negative relationship between this parameter and IQ. However, it must be recognized that this parameter correlates at about 0.80 with the variability measure. Thus, the patterning of correlations is seen to be very similar to that in Table 3 above. The slope parameter from the regression was quite independent of IQ with a mean parameter value of near zero across all channels, studies, and occasions. This confirms that the intercept parameter is primarily an estimate of the average of the mean absolute amplitudes for each epoch, computed over the individual epochs used to form the AEP.

Table 6 presents the correlations of the AEP P180 latency and DIFFLAT parameter (P180-N100 latency) with psychometric IQ. The results from this table demonstrate an enduring negative correlation between the P180 latency and DIFFLAT parameter and IQ. The relationship appears to be stronger with the P180 component than with the difference parameter. There also appears to be some hemispheric specificity in the pattern of relationships with the higher correlations clustered around the vertex locations.

In order to assess the reliability of the AEP parameters computed above, the test–retest reliability of each parameter was computed, with a 1 day interval and re-application of electrodes in the case

Table 6. Correlations between Psychometric IQ and 18 sets of observations across scalp electrode and study, P180 and P180-N100 latencies

	P180 Latency			P180-N100 Latency		
	Verbal IQ	Perf. IQ	Fullscale IQ	Verbal IQ	Perf. IQ	Fullscale IQ
Vertex electrode						
B & E (1992), A1-Cz, $n = 37$	−0.38	−0.34	−0.37	−0.28	−0.23	−0.25
B & E (1992), A2-Cz, $n = 37$	−0.20	−0.18	−0.20	−0.14	−0.11	−0.12
Study 1, day 1, [A1 + A2]-Cz, $n = 49$	−0.00	−0.36	−0.15	0.10	−0.21	−0.02
Study 1, day 2, [A1 + A2]-Cz, $n = 49$	0.14	−0.30	−0.04	0.19	−0.14	0.07
Study 2, Occ 1, [A1 + A2]-Cz, $n = 38$	−0.40	−0.38	−0.42	−0.31	−0.30	−0.33
Study 2, Occ 2, [A1 + A2]-Cz, $n = 40$	−0.26	−0.25	−0.28	−0.27	−0.25	−0.28
Study 2, left hemisphere [A1 + A2]						
Temporal, Occ 1, T3, $n = 47$	−0.21	−0.06	−0.14	0.06	0.09	0.08
Temporal, Occ 2, T3, $n = 37$	−0.03	0.12	0.04	0.20	0.03	0.13
Mid-temporal/vertex, Occ 1, C3, $n = 39$	−0.36	−0.30	−0.36	−0.17	−0.15	−0.18
Mid-temporal vertex, Occ 2, C3, $n = 39$	−0.19	−0.20	−0.21	−0.17	−0.12	−0.16
Parietal, Occ 1, P3, $n = 39$	−0.34	−0.44	−0.42	−0.28	−0.42	−0.38
Parietal, Occ 2, P3, $n = 39$	−0.22	−0.24	−0.24	−0.16	−0.14	−0.16
Study 2, right hemisphere [A1 + A2]						
Temporal, Occ 1, T4, $n = 41$	−0.23	−0.24	−0.25	−0.14	−0.27	−0.23
Temporal, Occ 2, T4, $n = 38$	−0.35	−0.28	−0.33	−0.17	−0.19	−0.19
Mid-temporal vertex, Occ 1, C4, $n = 42$	−0.32	−0.44	−0.41	−0.25	−0.39	−0.34
Mid-temporal/vertex, Occ 2, C4, $n = 38$	−0.17	−0.18	−0.18	−0.09	−0.18	−0.14
Parietal, Occ 1, P4, $n = 38$	−0.05	−0.27	−0.17	0.16	−0.19	−0.02
Parietal, Occ 2, P4, $n = 37$	−0.17	−0.32	−0.25	−0.07	−0.25	−0.17

Table 7. Test–retest reliability of six AEP parameters within Study 1 and 2

	String	Variability	Amplitude	Intercept	P180Lat	P180-N100
Vertex electrode						
Study 1, day 1 vs 2, [A1 + A2]-Cz, $n = 49$	0.73	0.78	0.80	0.75	0.82	0.82
Study 2, Occ 1 vs 2, [A1 + A2]-Cz, $n = 31$	0.82	0.85	0.79	0.68	0.65	0.54
Study 2, left hemisphere [A1 + A2]						
Temporal, Occ 1 vs 2, T3, $n = 31$	0.87	0.91	0.85	0.80	0.55	0.55
Mid-temporal/vertex, Occ 1 vs 2, C3, $n = 31$	0.83	0.81	0.81	0.76	0.43	0.44
Parietal, Occ 1 vs 2, P3, $n = 32$	0.70	0.89	0.81	0.72	0.60	0.37
Study 2, right hemisphere [A1 + A2]						
Temporal, Occ 1 vs 2, T4, $n = 31$	0.84	0.84	0.81	0.83	0.49	0.40
Mid-temporal/vertex, Occ 1 vs 22, C4, $n = 32$	0.85	0.91	0.81	0.79	0.54	0.47
Parietal, Occ 1 vs 2, P4, $n = 26$	0.84	0.91	0.83	0.85	0.72	0.57

Epoch length = 256 msec, selected *S* group.

Table 8. Test–retest reliability of six AEP parameters within Study 1 and 2

	String	Variability	Amplitude	Intercept
Vertex electrode				
Study 1, day 1 vs 2, [A1 + A2]-Cz, $n = 49$	0.71	0.77	0.73	0.71
Study 2, Occ 1 vs 2, [A1 + A2]-Cz, $n = 31$	0.78	0.87	0.75	0.79
Study 2, left hemisphere [A1 + A2]				
Temporal, Occ 1 vs 2, T3, $n = 31$	0.73	0.91	0.70	0.83
Mid-temporal/vertex, Occ 1 vs 2, C3, $n = 39$	0.76	0.84	0.78	0.77
Parietal, Occ 1 vs 2, P3, $n = 32$	0.73	0.90	0.73	0.75
Study 2, right hemisphere [A1 + A2]				
Temporal, Occ 1 vs 2, T4, $n = 31$	0.88	0.83	0.71	0.77
Mid-temporal/vertex, Occ 1 vs 22, C4, $n = 32$	0.83	0.91	0.82	0.84
Parietal, Occ 1 vs 2, P4, $n = 26$	0.84	0.93	0.86	0.87

Epoch length = 494/490 msec, selected *S* group.

of study 1, and a 5 min 'high cognitive activity' interval in study 2. Tables 7 and 8 provide the results for 256 msec and total epoch length parameters, respectively. These parameters were computed using the selected *S* groups only in order to maintain comparative coherence with the data reported above. It is readily apparent from these two tables that a remarkably high level of reliability is demonstrated within each dataset. The parameters with the lowest estimates of reliability appear to be those based upon the measurement of latency within study 2, the P180 and DIFFLAT (P180-N100) parameters. However, it must be noted that the sample size for each of the sub-groups is now reduced due to the fact that data from only those *S*s whose data appears in both test occasions for each channel was used. Test–retest reliabilities computed using the total dataset for the three studies were generally comparable with those reported above in Tables 7 and 8.

In contrast to the data above, it was decided to recompute the test–retest reliabilities using the *S*s who had been rejected from the analyses. These data are presented in Tables 9 and 10 below. In most instances, the results are comparable to those given above. The latency parameters are again the most unreliable of all the various measures and in fact show the largest decrement in size in comparison with the selected *S* reliabilities. However, for the remaining parameters, the reliability coefficients are remarkably high. This table indicates that the AEP rejection criteria are not selecting out data based upon random, destructive noise, but rather on some property (or lack of it) within the AEP. *The parameters computed from the AEPs within the non-selected subset are as reliable as those from within the selected subset.* We will return to this extremely significant finding in detail below.

Finally, in order to unambiguously demonstrate the effect of the sub-sampling of the data based upon the P180 amplitude criterion, two tables of data were created using the *S*s whose data appeared in every occasion and channel within study 1 and 2, within both the selected and non-selected subsets. For example, in study 2, only those *S*s whose data was accepted in every channel on both test occasions were used in this analysis. They were identified as the common-subset selected *S*s.* Likewise for the common-subset, non-selected *S*s. Tables 11 and 12 below present the results of the correlational analyses with the AEP parameters computed using the total epoch length.

*Common-subset selected and non-selected *S*s refers to those *S*s whose data passed or failed, respectively the two screening procedures within all 7 channels and on both test occasions within study 2.

Table 9. Test–retest reliability of six AEP parameters within Study 1 and 2

	String	Variability	Amplitude	Intercept	P180Lat	P180-N100
Vertex electrode						
Study 1, day 1 vs 2, [A1 + A2]-Cz, $n = 25$	0.89	0.49	0.87	0.59	0.64	0.62
Study 2, Occ 1 vs 2, [A1 + A2]-Cz, $n = 39$	0.70	0.94	0.86	0.87	0.58	0.49
Study 2, left hemisphere [A1 + A2]						
Temporal, Occ 1 vs 2, T3, $n = 33$	0.87	0.94	0.80	0.84	0.29	0.48
Mid-temporal/vertex, Occ 1 vs 2, C3, $n = 39$	0.71	0.93	0.87	0.88	0.40	0.47
Parietal, Occ 1 vs 2, P3, $n = 40$	0.70	0.94	0.60	0.87	0.10	0.14
Study 2, right hemisphere [A1 + A2]						
Temporal, Occ 1 vs 2, T4, $n = 38$	0.74	0.96	0.88	0.88	0.52	0.52
Mid-temporal vertex, Occ 1 vs 22, C4, $n = 42$		0.52 0.44	0.95	0.83	0.88	0.49
Parietal, Occ 1 vs 2, P4, $n = 37$	0.74	0.93	0.67	0.82	0.29	0.23

Epoch length = 256 msec, non-selected *S* group.

As can be seen from these tables, there is a clear difference between the size of correlations within the selected *S*s data in Table 11, and those from Table 12. The only correlation in Table 12 worthy of interest is that computed within study 1, on day 1, between the variability parameter and IQ. This correlation is +0.42 whereas in the selected sample dataset, it is −0.48. Thus, it is not surprising that in the overall sample for this dataset the observed correlation is near zero. Note also the reversal of sign of the correlation between the P180-N100 parameter and IQ in Table 11. As stated above, channel T3 was the least significant of all channels in the demonstration of a correlation between various AEP parameters and IQ. The artifactual correlation is evidence of the poor quality of AEP data from this channel.

In concluding the parameter analyses, the results are reported for the zero-cross analysis of the DC-uncorrected and corrected AEPs. For the sake of brevity, only those counts computed from the vertex electrodes in the three studies are reported. Table 13 provides the parameters from the DC-uncorrected AEPs. As can be seen from this table, there is little test–retest reliability for these parameters. In addition, some AEPs only yield 1 zero-cross due to the AEP waveform only crossing 0 V once throughout the entire 300 msec epoch. This can happen where the epoch begins at a negative voltage, rises to a positive voltage at around 140 msec duration and fails to cross 0 V before 300 msec. What is important from the analysis is that none of the parameters correlated conceptually or statistically significantly with IQ, even amongst the sub-tests. Correcting the AEPs for DC shifts also failed to yield any significant correlations between IQ and zero-cross count. Table 14 provides the parameters for this analysis.

Both tables provide clear evidence that the zero-cross hypothesis of Weiss is badly flawed in that it appears to be relevant only to AEPs evoked by visual flash stimuli. Even running the analyses with a minimum valid zero-cross count >2 yields no change in the neutral correlational patterns. Although memory span was not directly measured in this investigation, the WAIS-R did provide an estimate of digit span. This was not related at all to zero-cross count.

Summarizing the results above, it is apparent that by selecting a subset of data from the 3 studies, based solely upon an amplitude boundary on the P180 component and an adjustment for outlier variability, all three studies show a high degree of correlational concordance. Further, within study 1 and 2, the correlational relationships are mostly higher using the data from the first test occasion. What is not clear is why this apparently arbitrary criterion should produce such dramatic effects.

In order to show the selection effects visually, Figs 2 and 3 below provide the AEPs from 3

Table 10. Test–retest reliability of six AEP parameters within Study 1 and 2

	String	Variability	Amplitude	Intercept
Vertex electrode				
Study 1, day 1 vs 2, [A1 + A2]-Cz, $n = 25$	0.90	0.49	0.87	0.39
Study 2, Occ 1 vs 2, [A1 + A2]-Cz, $n = 39$	0.63	0.94	0.81	0.87
Study 2, left hemisphere [A1 + A2]				
Temporal, Occ 1 vs 2, T3, $n = 33$	0.80	0.93	0.77	0.83
Mid-temporal/vertex, Occ 1 vs 2, C3, $n = 39$	0.67	0.93	0.83	0.89
Parietal, Occ 1 vs 2, P3, $n = 40$	0.79	0.94	0.59	0.89
Study 2, right hemisphere [A1 + A2]				
Temporal, Occ 1 vs 2, T4, $n = 38$	0.64	0.96	0.85	0.84
Mid-temporal/vertex, Occ 1 vs 22, C4, $n = 42$	0.51	0.96	0.78	0.89
Parietal, Occ 1 vs 2, P4, $n = 37$	0.83	0.94	0.60	0.84

Epoch length = 494/490 msec, non-selected *S* group.

Table 11. Correlations between full-scale IQ score and AEP parameters, 490/494 msec epoch, occasion 1, common-subset selected *Ss* ($N = 26$)

	String	Variability	Amplitude	Intercept	P180	P180-N100
Vertex electrode						
Study 1, day 1, [A1 + A2]-Cz, $n = 49$	0.01	−0.48	0.20	−0.19	−0.15	−0.02
Study 2, Occ 1, [A1 + A2]-Cz	−0.10	−0.58	−0.13	−0.54	−0.54	−0.46
Study 2, left hemisphere [A1 + A2]						
Temporal, Occ 1, T3	−0.22	−0.32	−0.00	−0.25	0.08	0.39
Mid-temporal/vertex, Occ 1, C3	−0.07	−0.59	−0.14	−0.50	−0.51	−0.33
Parietal, Occ 1, P3, $n = 39$	−0.22	−0.63	−0.01	−0.54	−0.53	−0.35
Study 2, right hemisphere [A1 + A2]						
Temporal, Occ 1, T4	−0.28	−0.40	−0.28	−0.47	−0.21	−0.13
Mid-temporal/vertex, Occ 1, C4	−0.23	−0.60	−0.24	−0.66	−0.47	−0.42
Parietal, Occ 1, P4	−0.27	−0.57	−0.24	−0.61	−0.34	−0.10

adjacent channels, on test occasion 1, for 2 *Ss* of differing IQ drawn from the study 2 common-subset selected *S* group. Figures 4 and 5 provide the AEPs from 3 adjacent channels, on test occasion 1, for 2 *Ss* of differing IQ drawn from the common-subset non-selected *S* group.

These plots have not been chosen to represent the most marked differences between the two groups but rather to show a representative difference using *Ss* of similar IQ disparity. Three features of these plots are evident. Firstly, there is an obvious difference between the selected and non-selected *Ss* in the P180 region of interest. The amplitude and shape of the waveform in that area is quite different between the two groups. The N100 amplitude differs less between the two groups (mean N100 amplitude in the selected *Ss*, vertex electrode, is $-8.6\,\mu$V, for the non-selected group it is $-5.6\,\mu$V. For the P180 component amplitude the means are 8.6 and $3.5\,\mu$V, respectively). Secondly, the similarity of the channel waveforms, even in the non-selected group, is quite clear. It is important to remember here the high test–retest reliability of measures such as amplitude, variability, and contour length. Thirdly, there appears to be a higher degree of oscillatory activity on the AEP waveform in both *Ss*' drawn from the non-selected *S* group.

Moving further into the construction of these AEPS, it is highly illuminating to observe how the AEPs are built up from the individual epochs. Thus, for every *S*, AEPs were computed using blocks of sequential 25 epochs to form four sub-averages. These sub-average AEPs provide an easy visual guide as to the homogeneity of the individual epochs composing the main AEP. Figures 6 and 7 below show the vertex Cz sub-averages for Figs 2 and 3 above. Figures 8 and 9 show the corresponding sub-averages for Figs 4 and 5 above.

The figures clearly demonstrate the lack of homogeneity in the non-selected *Ss*' data in comparison to the selected *Ss*' data. However, in Fig. 6, there is a clear difference between the first and last 25 epochs and the middle 50 epochs. Note the peak at around 290 msec in Fig. 6, within the first block of trials. It is tempting to call this a P300 component except that there is no stimulus uncertainty. Further note how this peak is barely present within Fig. 2. The overriding result from looking at Figs 6–9 is that, in this particular paradigm, the use of averaging over all epochs is not an effective way of summarizing supposedly homogenous brain responses. There are *Ss* who do demonstrate remarkable consistency in their individual responses, but these *Ss* are the rarity. Finally, it is worthwhile to note just why the variability measure is correlating with IQ in the selected *S* group. The 256 msec variability estimate for *S* 23, with Total MAB IQ of 90, whose data is shown in Figs 4 and 8, is 16996.25. For *S* 80, with Total MAB IQ of 119, whose data is shown

Table 12. Correlations between full-scale IQ score and AEP parameters, 490/494 msec epoch, occasion 1, common-subset non-selected *Ss* ($N = 31$)

	String	Variability	Amplitude	Intercept	P180	P180-N100
Vertex electrode						
Study 1, day 1, [A1 + A2]-Cz, $n = 25$	0.05	0.42	0.19	0.21	−0.12	0.09
Study 2, Occ 1, [A1 + A2]-Cz	0.07	−0.01	−0.17	−0.03	−0.07	−0.15
Study 2, left hemisphere [A1 + A2]						
Temporal, Occ 1, T3	−0.08	−0.27	−0.07	−0.23	−0.01	0.05
Mid-temporal/vertex, Occ 1, C3	0.08	−0.00	−0.11	−0.03	−0.08	−0.17
Parietal, Occ 1, P3, $n = 39$	0.16	0.12	−0.00	0.11	−0.13	−0.08
Study 2, right hemisphere [A1 + A2]						
Temporal, Occ 1, T4	0.18	−0.07	0.09	−0.04	−0.13	−0.17
Mid-temporal/vertex, Occ 1, C4	0.09	−0.08	−0.07	−0.08	−0.03	−0.08
Parietal, Occ 1, P4	0.14	0.07	−0.10	0.07	−0.11	−0.26

Table 13. Zero-cross counts and test–retest reliabilities for the vertex AEPs from three studies, DC uncorrected AEPs

DC-uncorrected AEPs	Total sample					Selected sample				
Study	*N*	Mean	SD	Range	R-R*	*N*	Mean	SD	Range	R-R
B & E (1992), [Cz – A1]	40	2.9	1.1	1–6	0.36	37	2.8	1.1	1–6	0.40
Study 1, day 1, [Cz – A1 + A2]	74	3.5	1.1	2–6	0.47	49	3.5	1.1	2–6	0.58
Study 2, Occ 1, [Cz – A1 + A2]	86	3.9	1.4	1–9	0.39	38	3.8	1.2	2–8	−0.02

*R-R = test–retest reliability coefficient.

in Figs 5 and 9, the variability estimate is 13487.14. It is fairly obvious that what is causing the disparity in variability estimate is the greater lack of homogeneity in the individual epochs.

Finally, a rather interesting and possibly significant result is that from an analysis of the IT data within study 2. Using a new parameter (Barrett & Eysenck, submitted) that produced a monotonic rank estimate for IT (R_{it}) rather than just the actual time in milliseconds, it was found that R_{it} computed from the task undertaken in the EEG lab correlated with Total MAB IQ at 0.02 ($N = 31$) in the common-subset non-selected *S* group. For the corresponding selected group, this measure correlated at −0.55 $N = 26$, $P = 0.0036$ two-tail, 95% confidence is from −0.21 to −0.77). For the IT task implemented prior to that undertaken within the EEG lab, these correlations were −0.24 and −0.61 respectively (all correlation scatterplots were checked for outliers—there were some but their removal simply enhanced the large negative correlation for the selected *S*s). In order to confirm that this result is not simply a function of the small sample sizes of the two groups, the same analysis was carried out on the vertex Cz dataset for the selected and non-selected *S*s with sample sizes of 38 and 48 *S*s, respectively. This yielded correlations between R_{it} and IQ of −0.65 (non-EEG Lab IT, 95% confidence from −0.42 to −0.88) and −0.41 (EEG Lab IT) for the selected group and −0.30 and −0.02 for the non-selected group. Broadly, similar to the effect computed over the common subsets. Recognizing the fact that a core of 26 individuals who form the common-subset are in each selected group, it is of interest to perform another comparative analysis on the largest selection group generated for electrode position T3 (channel 0). The sample sizes for the selected and non-selected groups are 47 and 39, respectively. There are 21 extra *S*s in the selected group over and above the 26 common subset. The results of the same R_{it} analysis yielded correlations of −0.59 and −0.42 for the selected group and −0.24 and −0.01 for the non-selected group. Once again, the differential effect in the patterning of correlations is present. The stability of these results (which also extends across the other selected vs non-selected groups) suggests that this is not a chance or random effect but one that is quite systematic. For reference purposes, the total sample ($N = 86$) correlations between R_{it} and IQ were −0.45 and −0.21 for the non-EEG Lab and EEG Lab IT tasks, respectively. Test–retest across the two task conditions was 0.36. Within the common-subset *S* group, the same test-retest was 0.55. For the common-subset non-selected *S*s, this coefficient was 0.28. These R_{it} results are another indication that there are systematic individual differences operating within *S* samples. To ignore these differences by simply aggregating data is liable to produce results that will barely be replicable by other investigators unless they happen to 'get lucky' with their *S* mix. These results should provide some thought for those investigators working in the area of individual differences research and for those who interpret results from outside the area.

DISCUSSION

In reviewing the results and figures above, the simplicity of the sample selection criterion is apparent both in its execution and in the clarity of results thus produced. It might be argued that all that has been achieved by such selective sampling is the removal of artifactual AEPs from each

Table 14. Zero-cross counts and test–retest reliabilities for the vertex AEPs from three studies, DC corrected AEPs

DC-corrected AEPs	Total sample					Selected sample				
Study	*N*	Mean	SD	Range	R-R	*N*	Mean	SD	Range	R-R
B & E (1992), [Cz – A1]	40	2.5	1.1	1–6	0.72	37	2.4	1.1	1–6	0.70
Study 1, day 1, [Cz – A1 + A2]	74	3.3	1.2	1–6	0.45	49	3.3	1.2	1–6	0.56
Study 2, Occ 1, [Cz – A1 + A2]	86	4.1	1.4	1–8	0.24	38	3.8	1.3	1–7	0.04

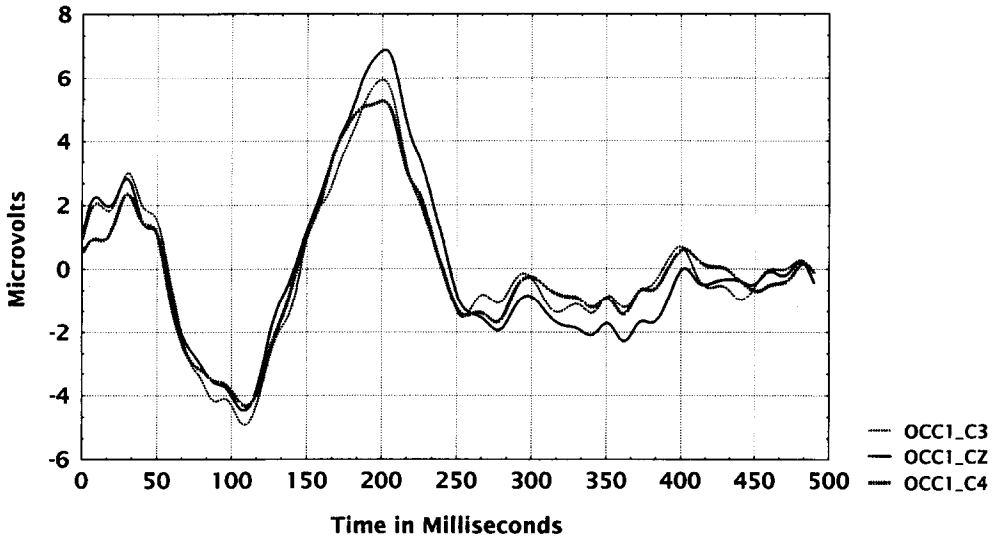

Fig. 2. AEPs from electrodes at scalp positions C3, Cz and C4. From selected *S* No. 23, occasion 1, full-scale IQ = 90 (OCC1-C3 = occasion 1, channel C3, OCC1-Cz = occasion 1, channel Cz, OCC1-C4 = occasion 1, channel C4).

particular dataset. However, Tables 7–10 suggest that although the AEPs may be artifactual, parameters computed from them are stable over days (study 1) or occasions (study 2). If random error were a cause of such artifact, test–retest reliability should be low or non-existent. Admittedly, parameter estimate comparisons are not the same as waveform envelope comparisons but the AEP contour length measure and to a lesser extent, the P180 component latency reliabilities are indexes of AEP similarity. For study 1, day 1 vs 2, common-subset non-selected *S* contour length reliability is 0.89, P180 latency test–retest is 0.64, P180 amplitude test–retest is 0.89, N100 latency test–retest is 0.57, N100 amplitude test–retest is 0.85. These indices do not support the proposition that the data is corrupted by random noise. If we take into account the R_{it} analyses on the selected and non-selected subset groups, it appears that the same 'effect' is evident in both the EEG and performance task. That is, within the same subset of *S*s' chosen on the basis of their P180 amplitude, enhanced correlations between both AEP parameters, R_{it} parameters, and psychometric IQ are found. This result adds further weight to the argument that rejection of AEPs is solely a function of some form of indirect artifact rejection. It begins to look more like that the *S* groups differ on the basis of some external criterion that affects performance assessment generally. This criterion is not a function of IQ, as Table 1 clearly demonstrates. One possibility is that 'attention' is the moderator of the AEP × IQ relationships. In the task, the maintenance of attention is crucial to performance on the task. If full attention is not paid to the stimulus, then it is impossible to complete the task. Work by Näätänen and others (Näätänen, 1990; Näätänen & Michie, 1979; Näätänen & Picton, 1987; Teder, Alho, Reinikainen & Näätänen, 1993) has demonstrated that the amplitude of the AEP components N100 (N1) and P180 (P2) can be moderated by the effects of attention to the stimuli, such that unattended stimuli produce AEPs of lower component amplitude than those from attended stimuli. We are suggesting that it is this effect, that of involuntary attentional processing of stimuli, that may be moderating the IQ × AEP correlations. If *S*s do not process the stimuli, the evoked potentials are of low amplitude and uncertain or ambiguous component structure. If they do process the stimuli, however involuntary, the resultant evoked potentials are better defined in their component structure and amplitude. From the R_{it} analyses, we might also hypothesize that these same *S*s have difficulty in attending to any task requiring attention to be focused and maintained. Unfortunately for this hypothesis, there were no statistical differences between R_{it} scores from either group. If attentional deficit was the explanation of the subset correlational differences, then it might have been expected to significantly affect performance in the IT task. However, its effect has been confined to removing the correlation between IQ and R_{it} within the non-selected sub-groups. Perhaps, as with the EEG, the effect simply produces more

variability in performance, leading to less accurate measures? Variability in the AEP paradigm was significantly lower in the study 1 selected group as compared to that in the non-selected group, $P < 0.001$, two-tail. For study 2, in either the common-subset or global selected groups, the difference was not significant at $P < 0.26$ and $P < 0.31$, respectively. Thus, the explanation for these sub-group correlational differences remains elusive although it is now apparent that the means of selection can be operationalized around a specific component amplitude parameter. Finally, there were no statistically significant personality score differences between the two groups in either study 1 or 2 datasets. However, EPQR Extraversion was consistently lower in the non-selected groups in study 1 and 2. The mean scores for the non-selected group in study 1 and 2 were 13.68 and 13.74, respectively, for the selected groups they were 15.29 and 16.38 respectively. Since the variances were statistically equal in both datasets, it would appear that given a greater number of *S*s producing the same range of scores, these means would have been statistically significant.

Although the hypothesized selection criterion is objective, it is not an efficient screening mechanism. This is likely to come from a new method of averaging evoked potentials that is based upon some form of pattern matching or similarity clustering, whether metric or non-metric. Figures 6–9 demonstrate that simple averaging of all 'non-artifactual' evoked potentials can be extremely error-prone. Unless the potentials are highly homogenous in terms of their envelope and amplitude, averaging 'smear' will be masking the effects of adding of two completely different waveforms. Preliminary work by us has indicated that the use of the entire epoch for these multivariate pattern-matching algorithms is not useful. Rather, a focused region-of-interest analysis appears to yield far greater differentiation of epochs, and subsequently increased homogeneity in any sub-group of potentials. This focusing on a region of the waveform (100–200 msec) has already been used successfully by Haier *et al.* (1983, 1984), Zhang, Caryl and Deary (1989a, 1989b), and Stough, Nettlebeck and Cooper (1990). One of the analyses currently being undertaken by us is the complete re-analysis of all our data from the three studies discussed above focusing all relevant parameter computations within various epochs within a 70–200 msec time window.

One other feature to arise from the analyses above is the lack of correlation between AEP zero-cross counts, contour length/string measures, and IQ. It is significant that both these measures were originally proposed and validated using Ertl and Schafer's (1969) published set of specimen potentials. From this study (and others referred to in the Introduction), it is now apparent that neither measure appears to be valid with the use of auditory stimulation. The basis of both these measures is derived from models that are attempting to explain long latency multiple peaks and troughs in high IQ *S* waveforms. Robinson (1993) has also elucidated in more depth the problem

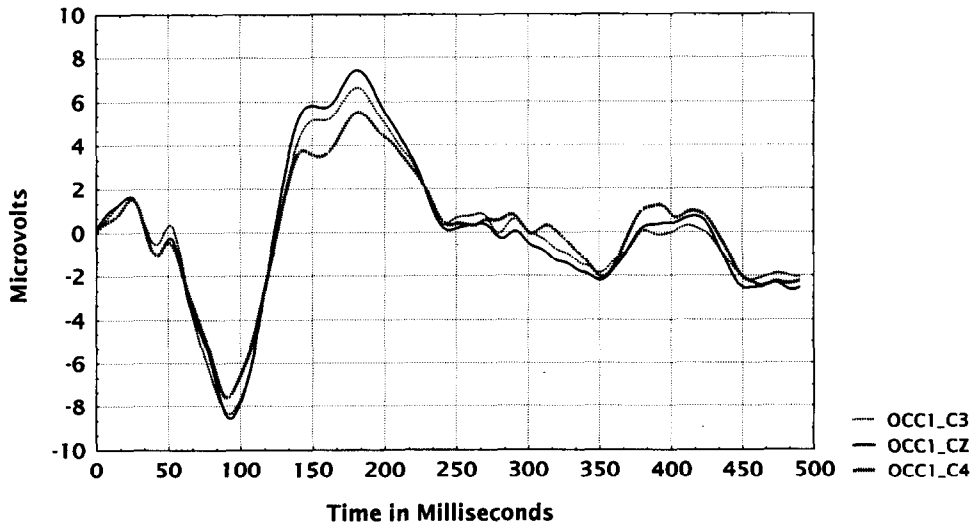

Fig. 3. AEPs from electrodes at scalp positions C3, Cz, and C4. From selected *S* No. 80, occasion 1, full-scale IQ = 119 (OCC1-C3 = occasion 1, channel C3, OCC1-Cz = occasion 1, channel Cz, OCC1-C4 = occasion 1, channel C4).

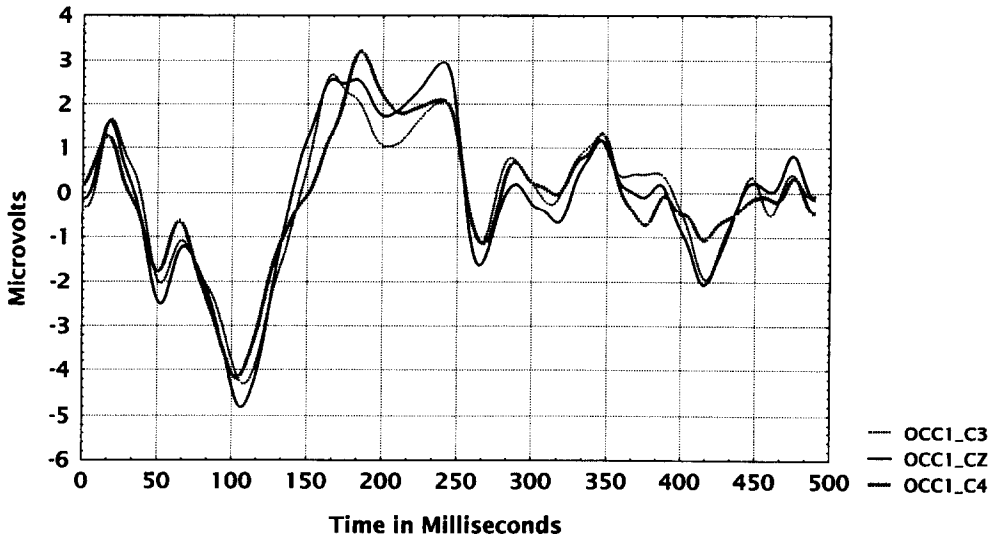

Fig. 4. AEPs from electrodes at scalp positions C3, Cz, and C4. From non-selected *S* No. 12, occasion 1, full-scale IQ = 92 (OCC1-C3 = occasion 1, channel C3, OCC1-Cz = occasion 1, channel Cz, OCC1-C4 = occasion 1, channel C4).

of using contour length measures of AEP waveform envelopes. In simple stimulus auditory evoked long latency potentials (from 40 msec or so onwards), there are no multiple peaks and troughs. The auditory stimulus waveform is triphasic. Any measure that is derived to produce measures based on the occurrence of multiple peaks and troughs is thus bound to fail. Where the string measure has been found to yield a positive correlation with IQ, this is probably more to do with the amplitude of the AEPs rather than any 'complexity' of the AEP waveform (see Robinson, 1993 for evidence of this assertion). In addition, examination of Figs 6–9 should provide a warning to anybody attempting contour length measures. The standard AEP average contour is likely to be heavily distorted by individual epoch fluctuations. Unless these non-homogenous epochs are removed prior to final averaging, it is unlikely that the string measure will correlate with anything other than AEP amplitude. Of course, this once again begs the question of what is a 'non-homogenous' epoch.

One alternative explanation of the results and procedures above is that we have been picking our way through evoked potentials whose properties are not defined solely by the stimulus characteristics and a fixed response output. In other words, if there is no long latency evoked potential component structure, and more than one possible output for a constant stimulus, then it will prove almost impossible to find a homogenous set of epochs within *S*s or AEPs across *S*s that can be said to share some common characteristics. Robinson's arousal theory (1993), albeit in not such strong terms, effectively suggests that this is the case. Part of Robinson's thesis is that the short latency brain response to any brief sensory stimuli is solely an analogue response. That is, there is no concept of sequential processing of stimuli through cognitive processing networks, each network stage producing some transient output to be subsequently labelled as a component of the waveform. In contrast to the accepted view of evoked potentials, Robinson argues (on the basis of some evidence) that the brain response is not the output from a 'stages of processing' serial processor, but the confounded output response of three oscillatory/reverberatory neural systems. Thus the evoked potential is seen as a function of the phase differences of three output waveforms, whose frequencies lie near 4, 7, and 11 Hz. The phase difference (and amplitude) of these three waveforms determines the resultant shape of the evoked potential envelope. The components N100, P180/200, and P300 are viewed as no more than products of waveform addition. They are not processing stages but just regions of particular phase characteristics. Robinson claims that the reliability of these 'components' is illusory in that there is no underlying 'component' being measured, just a region where the phase structure produces a region of maximum or minimum amplitude of the waveform envelope. Further, arousal influences the reverberatory activity

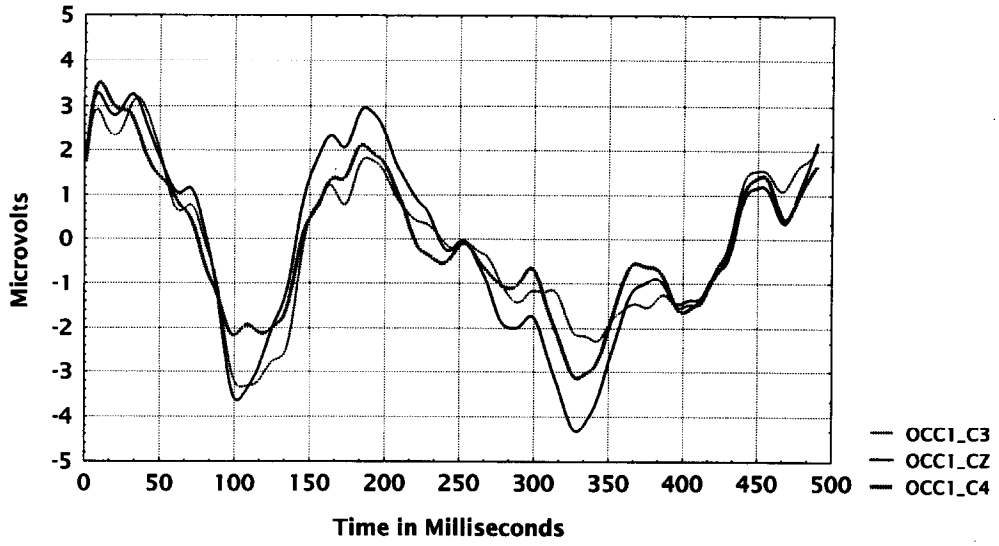

Fig. 5. AEPs from electrodes at scalp positions C3, Cz, and C4. From non-selected *S* No. 83, occasion 1, full-scale IQ = 120 (OCC1-C3 = occasion 1, channel C3, OCC1-Cz = occasion 1, channel Cz, OCC1-C4 = occasion 1, channel C4).

(assumed to be generated within the brain-stem, thalamus, and cerebral cortex), such that arousal will produce specific damping effects on the sinusoidal output waveforms. Thus, not only is an evoked potential seen as a product of not one but multiple waveforms, but the damping of the waveforms is also viewed as a function of the arousal 'state' of the *S*. Thus, any attempt to use measures of AEPs based upon the concept of a fixed output single waveform system is liable to meet with failure or be a rather unstable process.

The rather comprehensive analyses and seemingly puzzling selection issues noted above must give pause for thought to anybody who has read Robinson's previous work (1982, 1983, 1986, 1987, 1991, 1993) and judged it misguided. Although there are features of his theory that have not been replicated (Barrett & Eysenck, 1992b), this specific proposal on the nature of the brain response provides both an extremely powerful explanatory mechanism and a new way forward for AEP research. Our data and procedures above cannot in themselves test this aspect of his theory,

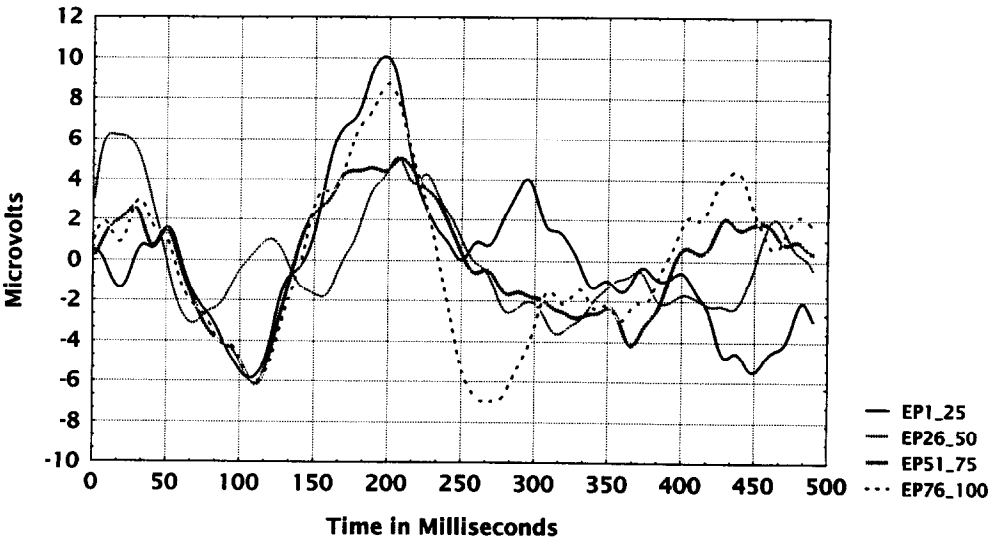

Fig. 6. AEPs computed from blocks of 25 sequential evoked potentials (EPs). From selected *S* No. 23, occasion 1, full-scale IQ = 90 (EP1-25 = EPs 1 to 25, EP26-50 = EPs 26 to 50. EP51-75 = EPs 51 to 75, EP76-100 = EPS 76 to 100).

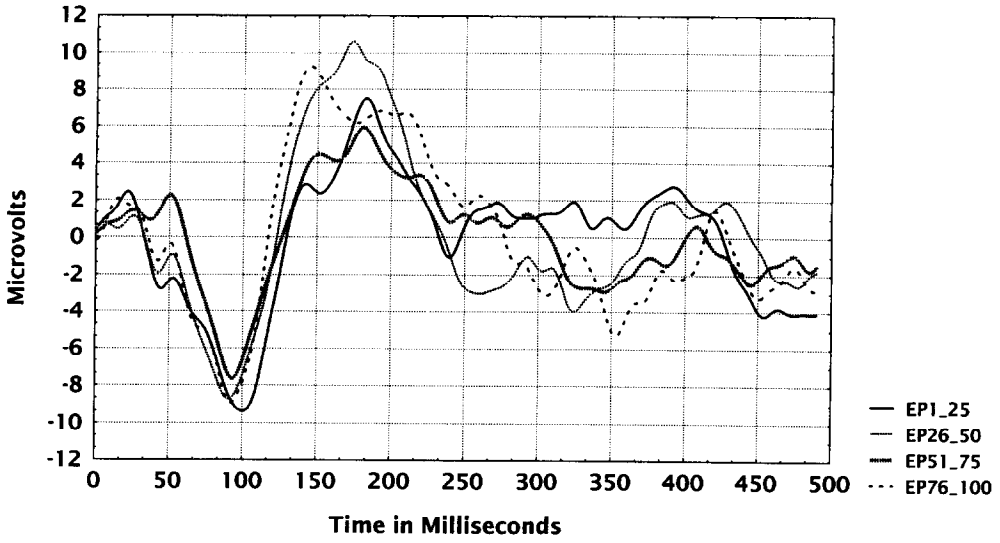

Fig. 7. AEPs computed from blocks of 25 sequential evoked potentials (EPs). From selected *S* No. 80, occasion 1, full-scale IQ = 119 (EPI-25 = EPs 1 to 25, EP26-50 = EPs 26 to 50. EP51-75 = EPs 51 to 75, EP76-100 = EPS 76 to 100).

but the difficulties we have experienced in the identification of a class or homogeneous potentials does lend weight to the argument that we have been using entirely the wrong methods of analysis to achieve the aim of testing a relationship between brain activity and IQ. It is quite possible that, in a rather haphazard manner, we have isolated those *S*s in our selected groups who happened to be in an intermediate state of arousal claimed by Robinson to be the optimal state to permit the assessment of AEP × IQ relationships.

Finally, if research is to continue within this area, the notion of individual differences in *S*s brain responses has to be addressed. The data above show that marked differences in results can exist within at least 2 sets of data, encompassing over 160 *S*s. Apparently neutral results were shown to be a product of two contrasting groups of *S*s within a supposed homogenous group. Also, if Robinson is correct in his assertion that the evoked response is a complex phase function of three damped sinusoids (he does provide some evidence for this assertion) then conventional analyses of AEPs are valueless. It now looks as though the Hendrickson paradigm of tone or click stimuli given in isolation, with no specific task involvement, is less than optimal for eliciting AEP × IQ relationships. It appears that a better route is to avoid the late auditory components altogether and instead record the sensory responses from the brainstem auditory pathways. Alternatively, Robinson's assertions about the composition of an AEP will have to be tested. If he is right, conventional AEP analysis is rendered obsolete and inaccurate, if he is wrong then enhanced multivariate techniques of analysis are required for evoked potential homogeneity analysis.

CONCLUSIONS

(1) Measures of AEP variability, mean individual epoch amplitude, and P180 component latencies were found to correlate negatively with IQ at around −0.50 across three separate studies, encompassing over 120 *S*s. However, the consistency of results was a function of selecting *S*s whose P180 component amplitude was greater than some specified, sample dependent, target value.

(2) It is suggested that conventional averaging of evoked potentials is no longer sufficient for future work in this area. Evidence from our data above indicates individual difference measurement information is being lost as well as distortions created in the AEP waveform envelope.

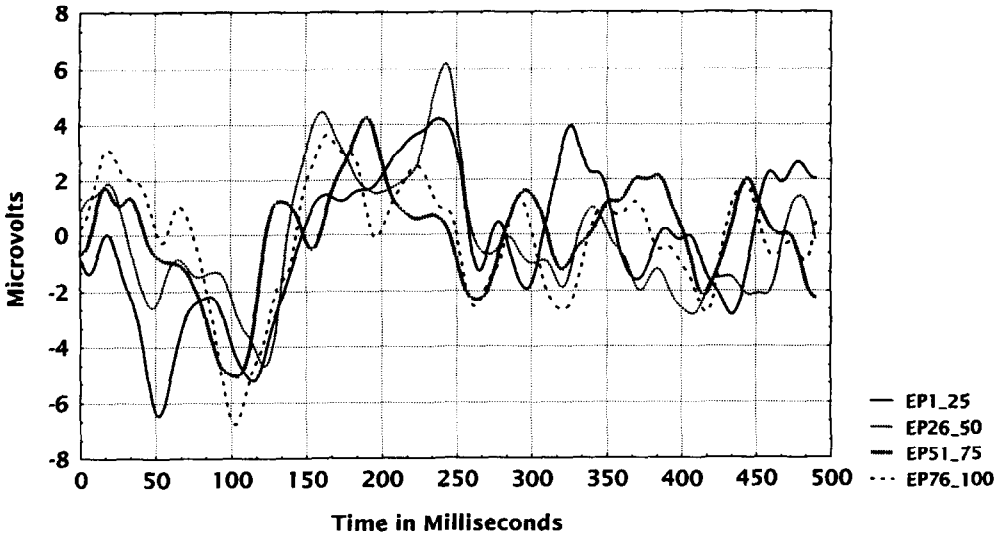

Fig. 8. AEPs computed from blocks of 25 sequential evoked potentials (EPs). From non-selected *S* No. 12, occasion 1, full-scale IQ = 92 (EPI-25 = EPs 1 to 25, EP26-50 = EPs 26 to 50. EP51-75 = EPs 51 to 75, EP76-100 = EPS 76 to 100).

(3) The proposition within Weiss' quantum theory of intelligence concerning the relationship of zero-crossings of the EP to IQ was not confirmed from a zero-cross analysis of data from 200 *S*s, drawn from three separate studies. No relationship between zero-cross count and IQ was found amongst any of the channel data or test occasions. Further it was stated that the proposition was based solely upon a set of data that were derived from visual stimuli, yielding evoked potentials whose characteristics are far removed from the typical auditory, triphasic response.

(4) Robinson's cerebral arousability theory was noted as a possible explanatory framework for our results. In addition, it was noted that if Robinson is correct in his assertion of the complex analogue nature of the evoked response, conventional AEP analysis is no longer relevant to the description of brain activity.

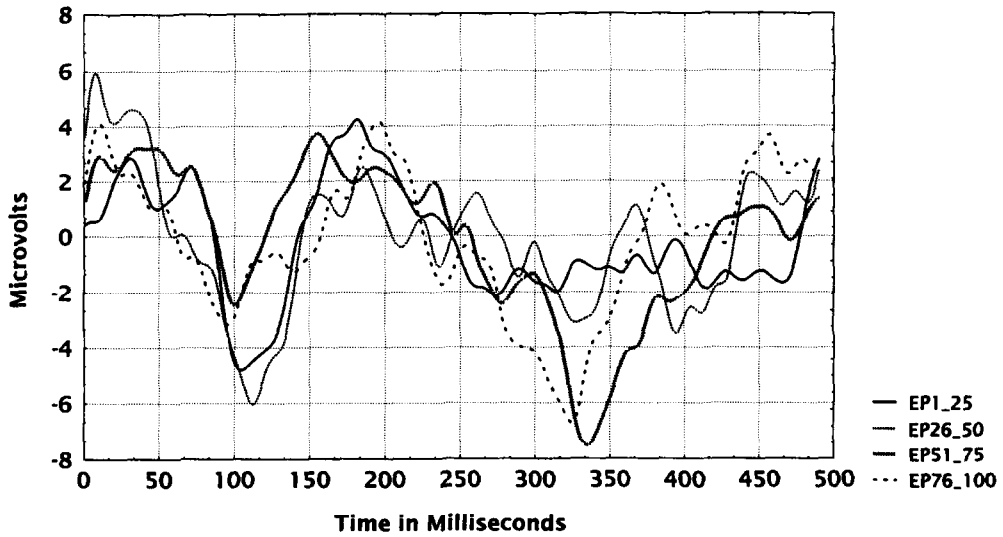

Fig. 9. AEPs computed from blocks of 25 sequential evoked potentials (EPs). From non-selected *S* No. 83, occasion 1, full-scale IQ = 120 (EPI-25 = EPs 1 to 25, EP26-50 = EPs 26 to 50. EP51-75 = EPs 51 to 75, EP76-100 = EPS 76 to 100).

Acknowledgements—We are indebted to the Pioneer Fund and to Philip Morris Inc. for grants which made this set of experiments possible. In addition, our thanks go to Mrs Jackie Marshall for all the testing of all the subjects within these experiments.

REFERENCES

Andreasen, N. C., Flaum, M., Swayze, V., O'Leary, D., Alliger, R., Cohen, G., Ehrhardt, J. & Yuh, W. (1993) *American Journal of Psychiatry, 150,* 130–134.

Ankney, C. D. (1992). Sex differences in relative brain size: the mismeasure of woman, too? *Intelligence, 16,* 329–336.

Barrett, P. T. (1988). Paul's Page: Normalising Transformations on EEG data. *Journal of Psychophysiology, 2,* 71–73.

Barrett, P. T. & Eysenck, H. J. (1992a). Brain evoked potentials and intelligence: The Hendrickson paradigm. *Intelligence,* 361–382.

Barrett, P. T. & Eysenck, H. J. (1992b). Brain electrical potentials and intelligence. In Gale, A. & Eysenck, M. W. (Eds), *Handbook of Individual Differences: Biological Perspectives.* New York: Wiley.

Barrett, P. T. & Eysenck, H. J. (1993). Sensory Nerve conduction and Intelligence: A replication. *Personality and Individual Differences, 15,* 249–260.

Barrett, P. T. & Eysenck, H. J. (submitted). The measurement of Inspection time and its relation to psychometric intelligence, reaction time, and personality. *Intelligence.*

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

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

Bates, T. & Eysenck, H. J. (1993). String length, attention, and intelligence. *Personality and Individual Differences, 15,* 363–371.

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

Berent, S., Giordani, B., Lehtinen, S., Markel, D., Penny, J. B., Buchtel, H. A., Starosta-Rubinstein, S., Hichwa, R. & Young, A. B. (1988). Positron emission tomographic scan investigations of Huntington's disease: cerebral metabolic correlates of cognitive function. *Annals of Neurology, 23,* 541–546.

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

Boivin, M. J., Giordani, B., Berent, S., Amato, D. A., Lehtinen, S., Koeppe, R. A., Buchtel, H. A., Foster, N. L. & Kuhl, D. E. (1992). Verbal fluency and positron emission tomographic mapping of regional cerebral glucose metabolism. *Cortex, 28,* 231–239.

Carroll, J. (1993). *Human cognitive abilities.* Cambridge: Cambridge University Press.

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

Chase, T. N., Fedio, P., Foster, N. L., Brooks, R., Di Chiro, G. & Mansi, L. (1984). Wechsler adult intelligence scale performance: cortical localisation by Fluorodeoxyglucose F18-Positron Emission Tomography. *Archives of Neurology, 41,* 1244–1247.

Cohen, R. & Williams, K. (1967). Eine Unteruschung zur Abhängigkeit der Intelligenz-struktur von Neurotizismus. *Zeitschrift fur experimentelle und angewandte Psychologie, 14,* 71–88.

De Leon, M. J., Ferris, H., George, A. E., Christman, D. R., Fowler, J. S., Gentes, C., Reisber, B., Gee, B., Emmerich, M., Yonekura, Y., Brodie, J., Kricheff, I. I. & Wolf, A. P. (1983). (1983). Positron emission tomographic studies of Aging and Alzheimer disease. *American Journal of Neuroradiology, 4,* 568–571.

Deary, I. & Caryl, P. (1992). Intelligence, EEG, and Evoked Potentials. In. Vernon, P. A. (Ed), *Biological approaches to the study of human Intelligence.* Norwood NJ: Ablex.

Detterman, D. K. & Daniel M. H. (1989), Correlations of mental tests with each other and with cognitive variables are highest for low IQ groups. *Intelligence, 13,* 349–359.

Detterman, K. Thompson, L. A. & Plomin, R. (1990), Differences in heritability across groups differing in ability. *Behavior Genetics, 20,* 369–384.

Duara, R., Grady, C., Haxby, J., Ingvar, D., Sokoloff, L., Margolin, R. A., Manning, R. G., Cutlet, N. R. & Rapoport, S. I. (1984). Human brain glucose utilization and cognitive function in relation to age. *Annals of Neurology, 16,* 702–713.

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 of the Eysenck Personality Questionnaire.* London: Hodder and Stoughton.

Eysenck, H. J. & White, P. O. (1969). Personality and the measurement of intelligence. *British Journal of Educational Psychology, 24,* 197–202.

Eysenck, S. B. G., Eysenck, H. J. & Barrett, P. T. (1985). A revised version of the Psychoticism scale. *Personality and Individual Differences, 6,* 21–29.

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.

Franzen, U. & Merz, F. (1976). Einfluss des Verbalisierens auf die Leistung bei Intelligenzpruefungen: Neue Unterusuchungen. *Zeitschrift fuer Entwicklungspsychologie und Paedagogische Psychologie, 8,* 117–134.

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

Frearson, W., Barrett, P. & Eysenck, H. J. (1988). Intelligence, Reaction Time, and the effects of smoking. *Personality and Individual Differences, 9,* 497–517.

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.

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., Siegel, B. V., Tang, C., Abel, L. Buchsbaum, M. S. (1992b). Intelligence and changes in regional cerebral glucose metabolic rate following learning. *Intelligence, 16,* 415–426.

Haier, R. J., Siegel, B. V., MacLachlan, A., Soderling, E., Lottenberg, S. & Buchsbaum, M. S. (1992a). Regional cerebral glucose metabolic changes after learning a complex visuospatial/motor task: a positron emission tomographic study. *Brain Research, 570*, 134–143.

Haier, R. J., Siegal, B. V., Nuechterlein, K. N., Hazlett, E., Wu, J. C., Paek, J., Browning, H. & Buchsbaum, M. S. (1988). Cortical glucose metabolic rate correlates of abstract reasoning and attention studied with positron emission tomography. *Intelligence, 12*, 199–217.

Haxby, J. V., Grady, C. L., Duara, R., Robertson-Tchabo, E., Kozarz, B., Cutler, N. R. & Rapoport, S.I. (1986). Relations among age, visual memory, and resting cerebral metabolism in 40 healthy men. *Brain and Cognition, 5*, 412–427.

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 Eysenck, H. J. (Ed.), *A model for intelligence*. New York: Springer-Verlag.

Ho, K. C., Roessmann, U., Straumfjord, J. V. & Monroe, G. (1980a). Analysis of brain weight: 1. adult brain weight in relation to sex, race, and age. *Archives of Pathology and Laboratory Medicine, 104*, 635–639.

Ho, K. C., Roessmann, U., Straumfjord, J. V. & Monroe, G. (1980b). Analysis of brain weight: 2. adult brain weight in relation to body height, weight, and surface area. *Archives of Pathology and Laboratory Medicine, 104*, 640–645.

Jackson, D. N. (1984). *The Multidimensional Aptitude Battery*. Port Huron, MI: Research Psychologists Press.

Jacobs, B. & Scheibel, A. B. (1993). A quantitative dendritic analysis of Wernicke's area in humans. 1. Lifespan changes. *Journal of Comparative Neurology, 327*, 83–96.

Jacobs, B., Schall, M. & Scheibel, A. B. (1993). A quantitative dendritic analysis of Wernicke's area in humans. 1. Gender, Hemispheric, and Environmental Factors. *Journal of Comparative Neurology, 327*, 97–111.

Kimura, D. & Hampson, E. (1992). Neural and hormonal mechanisms mediating sex differences in cognition. In Vernon, P.A. (Ed.), *Biological approaches to the study of human intelligence*. Norwood NJ: Ablex.

Kline, P. (1991). *A psychometric view of intelligence*. London: Routledge.

Lienert, G. A. (1968). Die Faktorenstruktur der Intelligenz als Funktion des Neurotizismus. *Zeitschrift fur experimentelle und angewandte Psychologie, 10*, 140–159.

Lutzenberger, W., Birbaumer, N., Flor, H., Rockstroh, B. & Elbert, T. (1992). Dimensional analysis of the human EEG and intelligence. *Neuroscience Letters, 143*, 10–14.

Lynn, R. (1989). A Nutrition theory of the secular increases in intelligence: positive correlations between height, head size, and IQ. *British Journal of Educational Psychology, 59*, 372–377.

Merz, F. (1969). Der Einfluss Verbalisierens auf die Leistung. *Zeitschrift fur experimentalle und angewandte Psychologie, 16*, 114–137.

Miller, E. M. (1992). On the correlation of Myopia and Intelligence. *Genetic, Social, and General Psychology Monographs, 88*, 363–383.

Miller, E. M. (submitted). Intelligence and Brain myelination: a hypothesis. *Behavioral and Brain Sciences*

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences, 13*, 201–287.

Näätänen, R. & Michie, P. T. (1979). Early selective attention effects on the evoked potential: a critical review and reinterpretation. *Biological Psychology, 8*, 81–136.

Näätänen, R. & Picton, T. W. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology, 24*, 375–425.

Nettlebeck, T. (1982). Inspection time: an index for intelligence? *Quarterly Journal of Experimental Psychology, 34a*, 299–312.

Parks, R. W., Loewenstein, D. A., Dodrill, K. L., Barker, W. W., Yoshii, F., Chang, J. Y., Emran, A., Apicella, A., Sheramata, W. & Duara, R. (1988). Cerebral metabolic effects of a verbal fluency test: A PET scan study. *Journal of Clinical and Experimental Neuropsychology, 10*, 565–575.

Pelosi, L., Holly, M., Slade, T., Hayward, M., Barrett, G. & Blumhardt, L. (1992). Event related potential correlates of performance of intelligence tests. *Electroencephalography and Clinical Neurophysiology, 84*, 515–520.

Raven, J. C. (1983). *Manual for Raven's Progressive Matrices*. London: H. K. Lewis.

Regan, D. (1989). *Human brain electrophysiology*. New York: Elsevier.

Reed, T. E. & Jensen, A. R. (1991). Arm nerve conduction velocity (NCV), Brain NCV, reaction time, and intelligence. *Intelligence, 15*, 33–47.

Reed, T. E. & Jensen, A. R. (1992). Conduction velocity in a brain nerve pathway of normal adults correlates with intelligence level. *Intelligence, 16*, 259–272.

Robinson, D. L. (1982). Properties of the diffuse thalamocortical system, human intelligence, and differentiated vs integrated modes of learning. *Personality and Individual Differences, 3*, 393–405.

Robinson, D. L. (1983). An analysis of human EEG responses in the alpha range of frequencies. *International Journal of Neuroscience, 22*, 81–98.

Robinson, D. L. (1986). The Wechsler Adult Intelligence Scale and Personality Assessment: Towards a biologically based theory of intelligence and cognition. *Personality and Individual Differences, 7*, 153–159.

Robinson, D. L. (1987). A Neuropsychological model of personality and individual differences. In Strelau, J. & Eysenck, H. J. (Eds), *Personality dimensions and arousal*. New York: Plenum Press.

Robinson, D. L. (1991). On the neurology of Intelligence and Intelligence factors. In Rowe, H. A. (Ed.), *Intelligence: Reconceptualisation and measurement*. Hillsdale, NJ: Lawrence Erlbaum.

Robinson, D. L. (1993). The EEG and Intelligence: an appraisal of methods and theories. *Personality and Individual Difference, 15*, 695–716.

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

Teder, W., Alho, K., Reinikainen, K. & Näätänen, R. (1993). Interstimulus interval and the selective attention effect on auditory ERPs: 'N1' enhancement vs processing negativity. *Psychophysiology, 30*, 71–81.

Van der Ven, H. G. S. (1992). Item homogeneity in a spatial test: a Rasch Analysis of Amthauer's cubes test. *European Journal of Psychological Assessment, 8*, 189–199.

Vernon, P. A. & Mori, M. (1989). Intelligence, reaction times, and nerve conduction velocity [Abstract]. *Behaviour Genetics, 19*, 779.

Vernon, P. A. & Mori, M. (1992). Intelligence, reaction times, and peripheral nerve conduction velocity. *Intelligence, 16*, 273–288.

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

Weiss, V. (1986). From memory span and mental speed toward the quantum mechanics of intelligence. *Personality and Individual Differences, 7*, 737–749.

Weiss, V. (1987). The quantum mechanics of EEG-brain dynamics and short term memory. *Biol. Zentralblatt, 106*, 401–408.

Weiss, V. (1989). From short term memory capacity toward the EEG resonance code. *Personality and Individual Differences, 10*, 501–508.

Weiss, V. (1992). The relationship between short-term memory capacity and EEG power spectral density. *Biological Cybernetics, 68*, 165–172.

Whitley, J. & Montano, I. L. (1992). Relation between reaction time and movement time in college wrestlers. *Perception and Motor Skills, 74*, 171–176.

Willerman, L., Schultz, R., Rutledge, A. N. & Bigler, E. D. (1991). *In vivo* brain size and intelligence. *Intelligence, 15*, 223–228.

Willerman, L., Schultz, R., Rutledge, A. N. & Bigler, E. D. (1992). Hemisphere size asymmetry predicts relative verbal and nonverbal intelligence differently in the sexes: an MRI study of structure-function relations. *Intelligence, 16*, 315–328.

Yeo, R. A., Turkheimer, E., Raz, N. & Bigler, E. D. (1987). Volumetric Asymmetries of the human brain: intellectual correlates. *Brain and Cognition, 193*, 15–23.

Zhang, Y., Caryl, P. G. & Deary, I. J. (1989a). Evoked potential correlates of inspection time. *Personality and Individual Differences, 10*, 379–384.

Zhang, Y., Caryl, P. G. & Deary, I. J. (1989b). Evoked potentials, inspection time, and intelligence. *Personality and Individual Differences, 10*, 1079–1094.