Published in: "Introduction to QEEG and Neurofeedback: Advanced Theory and Applications" Thomas Budzinsky, H. Budzinski, J. Evans and A. Abarbanel editors, Academic Press, San Diego, CA, 2008.

HISTORY OF THE SCIENTIFIC STANDARDS OF QEEG NORMATIVE DATABASES

Thatcher, R.W. [1] and Lubar, J.F.[2]

EEG and NeuroImaging Laboratory, Applied Neuroscience, Inc., St. Petersburg, Fl[1] , Brain Research and Neuropsychology Lab, University of Tennessee, Knoxville, TN[2].

Send Reprint Requests To:

Robert W. Thatcher, Ph.D.
NeuroImaging Laboratory
Applied Neuroscience, Inc.
St. Petersburg, Florida 33722
(727) 244-0240, rwthatcher@yahoo.com

Abstract

The nearly 40 year history of quantitative EEG (QEEG) normative databases are reviewed with special emphasis on the implementation of scientific and statistical standards. Differences between normative databases and standard control studies are discussed. The application of scientific and statistical standards such as peer reviewed publications, inclusion/exclusion criteria, number of subjects per age group, Gaussian tests for normality, cross-validation tests, amplifier matching, clinical correlations and FDA registration are presented in a historical context. A check list of "Gold Standards" for the evaluation of QEEG normative databases is presented in which the more checks then the higher the scientific and statistical standards for a given normative database. The goal of the paper is to provide an historical perspective and brief review of QEEG normative databases in order to encourage both users and authors of normative databases to strive for standardization.

Key words: QEEG normative database, cross-validation, Gaussian distributions

Introduction

Normative reference databases serve a vital and important function in modern clinical science and patient evaluation. There are numerous clinical normative databases that aid in the evaluation of a wide range of clinical disorders. For example, blood constituent normative databases, MRI, fMRI and Positron emission tomography (PET) normative databases, ocular and retinal normative databases, blood pressure normative databases, nerve conduction velocity normative databases, postural databases, bone density normative databases, ultra sound normative databases, genetic normative databases and motor development normative databases, to name a few. A comprehensive survey of existing clinical normative databases can be obtained by searching the National Library of Medicine's database using the search terms "Normative Databases" at: http://www.ncbi.nlm.nih.gov/sites/entrez.

All clinically applied normative databases share a common set of statistical and scientific standards that have evolved over the years. The standards include peer reviewed publications, disclosure of the inclusion/exclusion criteria, tests of statistical validity, tests of reliability, cross-validation tests, adequate sample sizes for different age groups, etc. Normative databases are distinct from non-clinical control groups in their scope and their sampling restriction to clinically normal or otherwise healthy individuals for the purpose of comparison. Another distinguishing characteristic of normative databases is the ability to compare a single individual to a population of "normal" individuals in order to identify the measures that are deviant from normal and the magnitude of deviation. Normative databases themselves do not diagnose a patient's clinical problem. Rather, a trained professional first evaluates the patient's clinical history and clinical symptoms and complaints and then uses the results of normative database comparisons in order to aid in the development of an accurate clinical diagnosis.

As mentioned previously the age range, the number of samples per age group, the mixture of gender and socioeconomic status, geographical distribution and thus a "representative" population are also distinguishing characteristics of a "normative" database because an individual is compared to a group of subjects comprising a reference normative database. In the case of QEEG, matching of amplifier frequency characteristics when a patient's EEG was acquired by a different amplifier than the database amplifier is also critical for normative databases but rarely important for standard "control group" studies. Cultural and ethnic factors and day-to-day variance and random environmental factors are typically factored into "normative" databases as "random control" factors, in contrast, a more limited sampling process is often used in non-clinical "control groups". The adequacy of the sample size of any database is related to the "effect size" and the statistical power and thus sample size varies depending on these factors (Cohen, 1977). In general, sample size is less important than careful calibration, elimination of artifact, accepted standards during the collection of data and accepted standards for the analysis of data and approximation to a Gaussian distribution. Peer reviewed publications are essential for all databases because high standards are required by anonymous reviewers and scientifically sub-standard databases will either not be published or if they are then the limitations are made public. To not publish a normative database in a peer reviewed journal is unacceptable and is a non-starter when a clinician considers the database that they are going to use to evaluate a patient. State licensing agencies and other authorities should be notified when sub-standard databases

are used to evaluate a clinical patient and certainly signed informed consent informing the patient that an unpublished and/or sub-standard database is being used to evaluate the patient is necessary to protect the public.

Definitions of Digital EEG and Quantitative EEG (QEEG)

Nuwer (1997) defined digital EEG as ". . . . the paperless acquisition and recording of the EEG via computer-based instrumentation, with waveform storage in a digital format on electronic media, and waveform display on an electronic monitor or other computer output device." The primary purposes of digital EEG is for efficiency of storage, the saving of paper and for the purposes of visual examination of the EEG tracings. An attempt was made to distinguish digital EEG from quantitative EEG by defining quantitative EEG (QEEG or qEEG) as "the mathematical processing of digitally recorded EEG in order to highlight specific waveform components, transform the EEG into a format or domain that elucidates relevant information, or associate numerical results with the EEG data for subsequent review or comparison." (Nuwer, 1997) (at p. 278). The reality is that there is no clear distinction between digital EEG and quantitative EEG because both involve mathematical transformations. For example, the process of analog-to-digital conversion involves transforms by analog and digital filtering as well as amplification and sample and hold of the electrical scalp potentials and re-montaging and reformatting the EEG. Clearly, digital EEG involves mathematical and transformational processing using a computer and therefore the distinction between quantitative EEG and digital EEG is weak and artificial.

Simultaneous Digital EEG Tracings and Quantitative EEG

Figure one illustrates a common modern quantitative EEG analysis where EEG traces are viewed and examined at the same time that quantitative analyses are displayed so as to facilitate and extend analytical power.

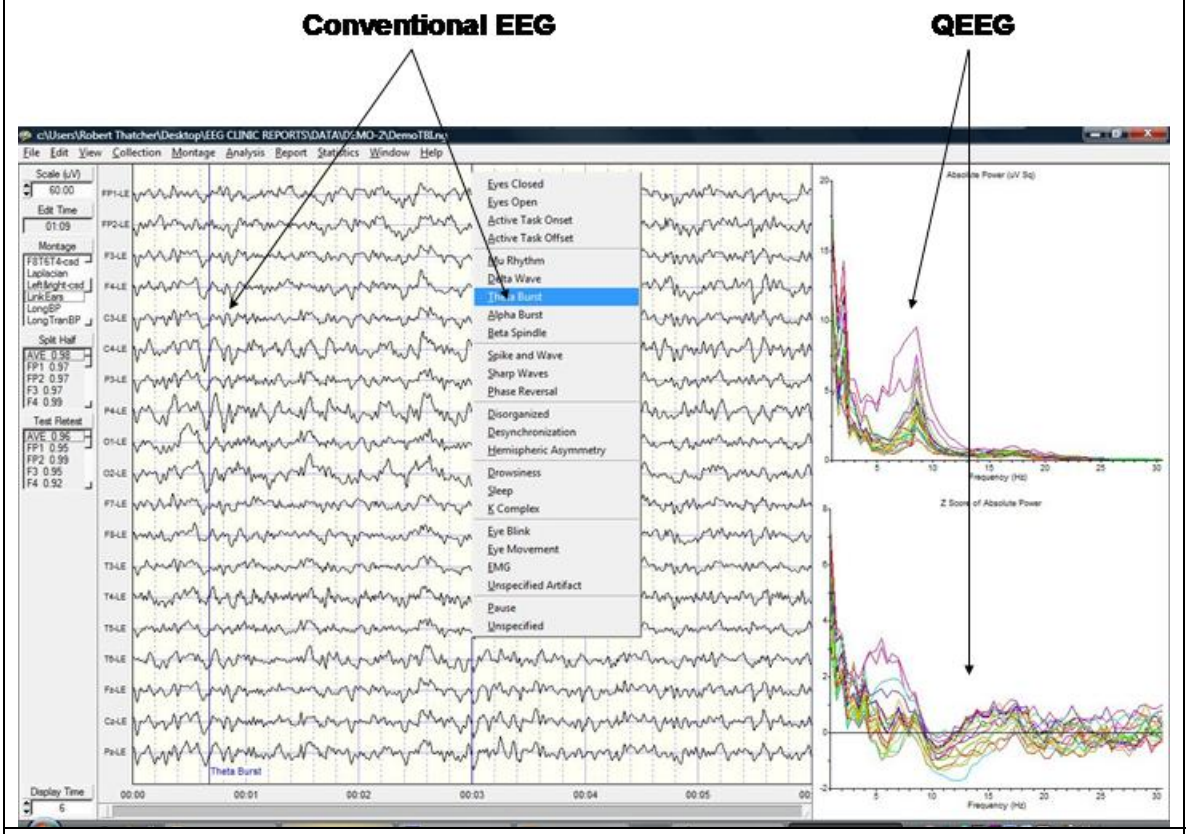

Fig. 1 – Example of conventional digital EEG (left) and QEEG (right) on the same screen at the same time. The conventional EEG includes examination and marking of EEG traces and events. The QEEG (right) includes the Fast Fourier Transform (Top right) and normative database Z scores (Bottom right).

Commonsense dictates that the digital EEG and QEEG when simultaneously available facilitates rapid and accurate and reliable evaluation of the electroencephalograpm. Since 1929 when the human EEG was first measured (Berger, 1929) modern science has learned an enormous amount about the current sources of the EEG and the manner in which ensembles of synaptic generators are synchronously organized. It is known that short distance local generators are connected by white matter axons to other local generators that can be many centimeters distant. The interplay and coordination of short distance local generators with the longer distant white matter connections has been mathematically modeled and shown to be essential for our understanding of the genesis of the EEG (Nunez, 1981; 1995; Thatcher and John, 1977; Thatcher et al, 1986).

The first QEEG study was by Hans Berger (1932; 1934) when he used the Fourier transform to spectrally analyze the EEG because Dr. Berger recognized the importance of quantification and objectivity in the evaluation of the electroencephalogram (EEG). The relevance of quantitative EEG (QEEG) to the diagnosis and prognosis of brain dysfunction stems directly from the quantitative EEG's ability to reliably and objectively evaluate the distribution of brain electrical energies and to compare different EEG measures to a normative database.

Test-Retest Reliability of QEEG

The clinical sensitivity and specificity of QEEG is directly related to the stability and reliability of QEEG upon repeat testing. The scientific literature shows that QEEG is highly reliable and reproducible (Hughes and John, 1999; Aruda et al, 1996; Burgess and Gruzelier, 1993; Corsi-Cabera et al, 1997; Gasser et al, 1985; Hamilton-Bruce et al, 1991; Harmony et al, 1993; Lund et al, 1995; Duffy et al, 1994; Salinsky et al, 1991; Pollock et al, 1991). The inherent stability and reliability of QEEG can even be demonstrated with quite small sample sizes. For example, Salinsky et al (1991) reported that repeated 20-second. samples of EEG were about 82% reliable, at 40 seconds the samples were about 90% reliable and at 60 seconds they were approximately 92% reliable. Gasser et al (1985) concluded that: "20 sec. of activity are sufficient to reduce adequately the variability inherent in the EEG" and Hamilton-Bruce et al, (1991) found statistically high reliability when the same EEG was independently analyzed by three different individuals. Although the QEEG is highly reliable even with relatively short sample sizes, it is the recommendation of most QEEG experts that larger samples sizes be used, for example, at least 60 seconds of artifact free EEG, and preferably 2 to 5 minutes, should be used in a clinical evaluation (Duffy et al, 1994; Hughes and John, 1999).

Although there are common purposes and applications of normative databases in clinical science, nonetheless, each type of normative database poses its own special requirements and details. In the sections to follow we focus exclusively on quantitative electroencephalographic (QEEG) normative databases. The goal of this paper is to present the history of the application of scientific standards as they apply to QEEG and to provide a practical guide for the understanding and evaluation of QEEG normative databases.

History of Standards of QEEG Normative Databases

The earliest quantitative EEG (QEEG) reference normative database was developed in the 1950s at UCLA as part of the NASA study and selection of astronaughts for purposes of space travel (Adey et al, 1961; 1964a; 1964b). The UCLA database involved several hundred carefully selected subjects who were candidates for the burgeoning NASA space exploration program as well as UCLA faculty and students. Careful clinical inclusion and exclusion criteria were not used because there was no intended clinical application of this early QEEG reference normative database. Instead, the essential quantitative foundations of QEEG normative databases were tested such as the calculation of means and standard deviations and measures of Gaussianity, complex demodulation, Fourier spectral analysis and basic statistical parameters necessary for any reference normative database.

Predictive accuracy and error rates depend on the data that make up a given EEG database as well as the statistical methods used to produce and compare QEEG normative databases. Historically, many of the statistical standards of normative databases were first applied by two Swedish Neurologist, Dr. Milos Matousek and Dr. Ingemar Petersen in 1973 in the first peer reviewed publication of a normative database (Matousek and Petersen, 1973a; 1973b). Matousek and Petersen set the standards of peer reviewed publications, clinical inclusion/exclusion criteria and parametric statistical standards for future QEEG normative databases. The cultural validity and reliability of the Matousek and Petersen 1973 database were established by E. Roy John and colleagues in 1975

when they successfully replicated, by independent cross-validation, the Matousek and Petersen Swedish database after collecting EEG from carefully screened 9 to 11 year old Harlem black children who were performing at grade level and had no history of neurological disorders (John, 1977; John et al, 1977; 1987).

History of Inclusion/Exclusion Criteria and "Representative Samples"

Matousek and Petersen (Matousek and Petersen, 1973a; 1973b) measured QEEG in 401 subjects (218 females) ranging in age from 2 months to 22 years and living in Stockholm, Sweden all without any negative clinical histories and performing at grade level. The sample sizes varied from 18 to 49 per one year age groupings. Similar inclusion/exclusion criteria were later used in the construction of the NYU normative database (John, 1977; John et al, 1977; 1987), the University of Maryland (UM) database (Thatcher, 1988; Thatcher et al, 1983; 1986; 1987; 2003; 2005a; 2005b) and Gordon and colleagues (2005) in the development of independent QEEG normative databases. Careful screening of the subjects that comprise a normative database is critical so that representative samples of healthy and otherwise normally functioning individuals are selected and individuals with a history of neurological problems, psychiatric problems, school failure and other deviant behaviors are excluded.

Representative sampling means a demographically balanced sample of different genders, different ethnic backgrounds, different socio-economic status and different ages. This is important in evaluating a QEEG normative database because the database is a "reference" in which many demographic factors must be included in order to minimize sampling bias.

History of Artifact Free Data and Reliability Measures

Sample adequacy in a QEEG normative database requires strict removal of artifact and measures of high test re-test reliability. Historically, multiple trained individuals visually examined the EEG samples from each and every subject that was to be included in the database. Removal of artifact by visual examination is necessary regardless of any digital signal processing methods that may be used to remove artifact. Split-half reliability and test re-test reliability measures with values > 0.9 is also important in order to provide a quantitative measure of the internal consistency and reliability of the normative database (John, 1977; John et al, 1987; Thatcher, 1998; Thatcher et al, 2003; Duffy, 1994).

Caution should be exercised when using reconstruction methods such as Independent Components Analysis (ICA) or Principal Component Analysis (PCA) to compute a QEEG normative database. In general, these methods should be avoided because they will invalidate the computation of coherence and phase differences because the regression and reconstruction effects the raw digital samples themselves and distorts coherence and phase. The best method of eliminating artifact is by making sure that high standards of recording are met and that the patient's EEG is monitored during recording so that artifact can be minimized. Elimination of artifact after recording should involve the deletion of the artifact from the analysis and not by regression and/or reconstruction using methods such as ICA or PCA.

History of Sample Size per Age Group

There is no absolute sample size that is best for a QEEG database because statistically sample size is related to the "effect size" and "power" (Hayes, 1973; Winer, 19971). The smaller the effect size then the larger the sample size necessary to detect that effect. The power of a statistical measure varies as a function of sample size and the effect size (Cohen, 1977). Another issue related to sample size is the degree to which a sample approximates a Gaussian distribution. As explained in the section below, increased sample size is often necessary in order to achieve closer approximations to Gaussian which in turn is related to the accuracy of cross-validation. Thus, the sample size is one of several inter-related issues in all normative databases and the sample size should not be singled out as being the most important factor in a QEEG normative database. It is best to refer to "adequate" sample size as measured by the extent to which the samples are Gaussian and the degree of cross-validation accuracy (John et al, 1987; Thatcher et al, 2003). The term "adequate" is related to the effect size, which in the case of human development is critical because different rates of maturation occur at different ages.

As mentioned previously, the Matousek and Petersen (1973a; 1973b) normative QEEG database had a total sample size of 401 in children ranging in age from 1 month to 22 years. It was known that there are rapid changes in EEG measures during early childhood and for this reason Matousek and Petersen (1973a) and Hagne et al (1973) emphasized using relatively large sample sizes during the period of time when the brain is changing most rapidly. For example, Hagne et al (1973) used a sample size of N = 29 for infants from three weeks of age to 1 year of age. In step with this fact were the subsequent QEEG normative databases at NYU (John et al, 1977; 1987) and UM (Thatcher, 1998; Thatcher et al, 1987; 2003) in which the preferential increase in sample size during early childhood was emphasized as well as during old age when potential rapid declines in neural function may occur.

History of Age Stratification vs. Age Regression

There are two general approaches that deal with the issue of sample size per age group: 1- Age stratification and, 2- Age Regression. Age stratification involves computing means and standard deviations of age groupings of the subjects (Matousek and Petersen, 1973a; John, 1977; Thatcher et al, 1987; 2003). The grouping of subjects and thus the number of subjects per age group depends on the age of the sample and the relative rate of maturation. Matousek and Petersen (1973a; 1973b) used one year age groupings, Thatcher et al (1987) (University of Maryland database) used one year age groupings as well as two and five year age groupings (Thatcher et al, 2003; 2005a; 2005b). A simple method to increase stability and sample size is to use "sliding" averages for the age stratification. For example, Thatcher et al (2003) used one year age groups with .75 year overlapping to produce a series of sliding averages and more recently used two year age groupings with .75 year overlapping. Which method is chosen depends on the accuracy of cross-validation and age resolution with careful examination of validation at different ages of the subjects.

The second method called "Age Regression" was first used by John et al (1977; 1980) in which a least squares regression was used to fit a straight line to the EEG data samples over the entire age range of the subjects. Once the intercepts and coefficients

are computed then one simply evaluates the polynomial equation using the age of the subject in order to produce the expected mean and standard deviation for that particular subject. A Z score is then computed by the standard method $Z = X - x/sd$. An important consideration when using an age regression method is the order of the polynomial and the amount of variance accounted for by a polynomial. If there are rapid maturational changes in the brain thus producing a "growth spurt" then a simple linear regression is likely to miss the growth spurt. A quadratic or cubic polynomial which will account for more of the variance over age will likely detect growth spurts better than a simple linear regression.

History of Gaussian Distribution Approximation and Cross-Validation

The statistics of replication and independent cross-validation of normative QEEG databases was first applied by E. Roy John and collaborators in 1974 to 1977 (John, 1977; John et al, 1977; 1987). As mentioned previously, the first independent cross-validation of a normative QEEG database was by John and colleagues in which the EEG from a sample of New York Harlem black children were compared to the Matousek and Petersen (5, 6) norms with correlations > 0.8 in many instances and statistically significant correlations for the majority of the measures (John, 1977). The importance of approximation to a Gaussian distribution was emphasized by both Dr. E. Roy John and Dr. Frank Duffy a Harvard Neurologist in the 1970s and 1980s. In 1994 the American EEG Association produced a position paper in which the statistical standards of replication, cross-validation, reliability and Gaussian approximation were iterated as acceptable basic standards to be met by any normative QEEG database (Duffy, 1994). The American EEG Society included the same standards. Dr. John and colleaques from 1980 to 1990s continued to evaluate and analyze the statistical properties of normative QEEG databases, including EEG samples obtained from different laboratories in non USA locations in the world. Gaussian approximations and reliability and cross-validation statistical standards for QEEG databases were applied to all of these databases by John and Colleagues (John et al, 1987; 1980; Prichep, 2005) and as well as by other QEEG normative databases, for example, Gasser et al (1988a; 1988b); Thatcher and colleagues (1983; 1986; 1987; 2003; 2005a; 2005b).

Figure 2 are examples of approximate Gaussian distributions and the sensitivity as calculated in figure 3. Table I is an example of a standard table of sensitivities for different age groups in the University of Maryland QEEG normative database (Thatcher et al, 2003).

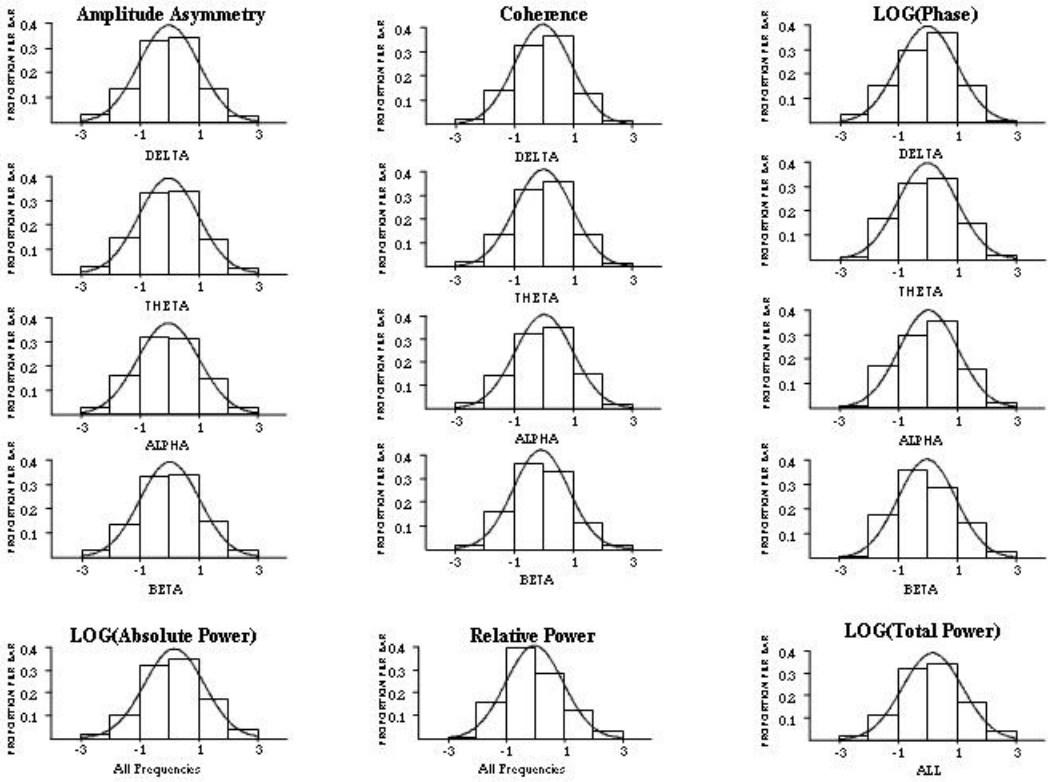

Fig. 2 - Histograms of the Z-Score Gaussian distributions and cross-validation for all ages (from Thatcher et al, 2003).

Figure 3 shows an example of Gaussian approximation and cross-validation of a QEEG normative database.

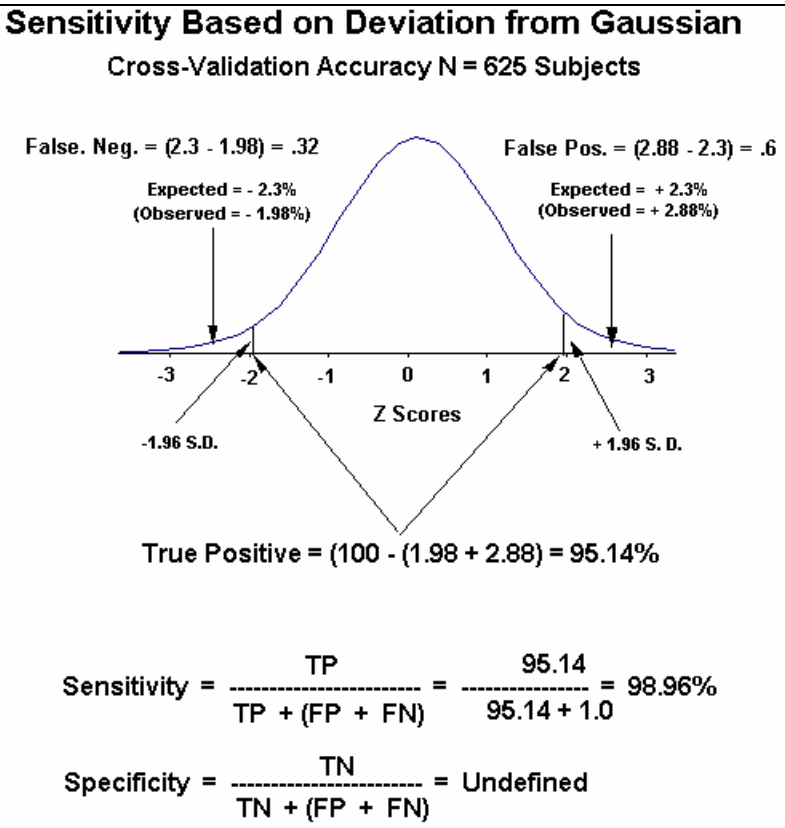

Fig.-3 An example of a normal or Gaussian curve showing values of Z (± 1.96), that includes the proportion which is .95 of the total area. The left and right tails of the distribution show probability values of .025 (one-tailed). The classification accuracy of any sample of subjects is based on the assumption of a normal distribution can be compared. The probability of finding an observed EEG value in a given range of any population can be determined and then the sensitivity of the sample can be tested by cross-validation (adapted from Thatcher et al, 2003).

Figure 3 is an illustrative bell-shaped curve showing the ideal Gaussian and the average cross-validation values of the database by which estimates of statistical sensitivity can be derived. True positives equal the percentage of Z-scores that lay within the tails of the Gaussian distribution. False negatives (FN) equal the percentage of Z-scores that fall outside of the tails of the Gaussian distribution. The error rates or the statistical sensitivity of a quantitative electroencephalogram (QEEG) normative database are directly related to the deviation from a Gaussian distribution. Figure 3 depicts a mathematical method of estimating the statistical sensitivity of a normative EEG database in terms of the deviation from Gaussian.

FFT Normative Database Sensitivities

2 STDEVs	CALC SENSITIVITY: FP=TP/(TP+FP) or FN=TP/(TP+FN)			
AGES	(+/- 2 SD)	(>= 2 SD)	(<= -2 SD)	
0-5.99	0.95448265	0.9771774	0.97730526	+/- **2 Std. Dev.**
6-9.99	0.95440363	0.9772031	0.97720054	
10-12.99	0.9543997	0.97724346	0.97715624	
13-15.99	0.95440512	0.97723601	0.97716911	
16-ADULT	0.9543945	0.97718143	0.97721307	
ALL	0.95442375	0.97720714	0.97721661	

3 STDEVs	CALC SENSITIVITY: FP=TP/(TP+FP) or FN=TP/(TP+FN)			
AGES	(+/- 3 SD)	(>= 3 SD)	(<= -3 SD)	
0-5.99	0.99743898	0.99871123	0.99872774	+/- **3 Std. Dev.**
6-9.99	0.99744112	0.99871611	0.99872501	
10-12.99	0.99744688	0.99873171	0.99871518	
13-15.99	0.99743186	0.99871951	0.99871234	
16-ADULT	0.99743835	0.99870216	0.99873619	
ALL	0.99744002	0.99871716	0.99872286	

Table I – Example of cross-validation and sensitivity tests of a normative database using the procedures described in Figure 3. (Adapted from Thatcher et al, 2003).

History of the use of the Z Score and QEEG Normative Databases

Matousek and Petersen (1973a; 1973b) computed means and standard deviations in one year age groups and were the first to use t-tests and Z scores to compare an individual to the normative database means and standard deviations. The T-Test is defined as the ratio of the difference between values divided by the standard deviation. The Z statistic is defined as the difference between the value from an individual and the mean of the population divided by the standard deviation of the population or

$$Z = \frac{x_i - \overline{X}}{SD}.$$

John and colleques (John, 1977; John et al, 1977; 1987) expanded on the use of the Z score for clinical evaluation including the use of multivariate measures such as the Mahalanobis distance metric (Cooley and Lohnes, 1971; John et al, 1987; John et al, 1988). A direct normalization of the Gaussian distribution using Z scores is useful in comparing individuals to a QEEG normative database (Thatcher, 1998; Thatcher et al, 2003). That is, the standard score form of the Gaussian is where the mean = 0 and standard deviation = 1 or, by substitution into the Gaussian equation for a bell shaped curve, then

$Y = \frac{1}{\sqrt{2\pi}} e^{-z2/2}$, where Y = Gaussian distribution and the Z score is a deviation in standard deviation units measured along the baseline of the Gaussian curve from a mean of 0 and a standard deviation = 1 and deviations to the right of the mean being positive and those to the left negative. By substituting different values of Z then different values of Y can be calculated. For example, when Z = 0, Y = 0.3989 or, in other words, the height of the curve at the mean of the normal distribution in standard-score form is given by the number 0.3989. For purposes of assessing deviation from normal, the values of Z above and below the mean, which include 95% of the area of the Gaussian is often used as a level of confidence necessary to minimize Type I and Type II errors (Hayes, 1973). The standard-score equation is also used to cross-validate a normative database which again emphasizes the importance of approximation to a Gaussian for any normative QEEG database.

Cross-Validations of Normative Databases: New York University and University of Maryland

As described previously, cross-validation is critical in determining the sensitivity and false positives and false negatives of a normative database. Due to the expense to acquire independent data, most cross-validations are computed using a leave-one-out cross-validation procedure (John et al, 1977; 1987; Thatcher et al, 2003; 2005a; 2005b). A completely independent cross-validation using different subjects is the best method of cross-validation although it is, as previously stated, more expensive and difficult and, accordingly, no independent cross-validations of two different normative databases have been conducted in the last 30 years, until recently. In 2007 an independent cross-validation of the New York University and the University of Maryland databases were conducted. The study was conducted because a company had collected raw digital EEG from several hundred clinical patients and had computed Z scores using the New York University (NYU) normative database (John, 1977; John et al, 1977; 1987; 1988). The question was: does the University of Maryland (UM) normative database produce similar of comparable Z scores as the NYU database using the same exact raw digital data? The correlation coefficients from the independent cross-validation between the NYU and UM normative databases is shown in Table II. The analysis included 332 psychiatric patients and an age range from 6.2 years to 84.9 years. Anterior includes electrodes Fp1/2, Fz, F3/4, F7/8, T3/4, C3/4 & Cz. Posterior includes electrodes O1/2, P3/4, T5/6 and Pz. The correlations ranged from 0.757 to 0.979. The high degree of cross-validation accuracy in this study is emphasized by the fact that at 331 degrees of freedom a correlation of 0.142 is significant at $P < .01$.

Table II
Correlation Coefficients from an Independent Cross-Validation of NYU vs UM Normative EEG Databases (Reprinted by permission of CNS Response, Inc.)

	Absolute Power	Absolute Power	Relative Power	Relative Power	Coherence	Coherence	Amp. Asym	Amp. Asym

	Anterior	Posterior	Anterior	Posterior	Anterior	Posterior	Anterior	Posterior
Delta	0.815	0.880	0.854	0.925	0.804	0.935	0.854	0.820
Theta	0.926	0.940	0.877	0.895	0.853	0.914	0.902	0.816
Alpha	0.951	0.958	0.901	0.887	0.873	0.046	0.899	0.979
Beta	0.820	0.882	0.757	0.784	0.848	0.900	0.846	0.876

Figure 4 are bar graphs of the correlation coefficients from the independent cross-validation comparison between the NYU and the UM Z scores. This study is important because it demonstrates a high degree of cross-correlation and cross-validation between two independent QEEG normative databases. Both the NYU and UM databases were constructed in medical centers with government grants and oversight and both have been clinically validated in peer reviewed publications (John et al, 1977; 1987; 1988l Thatcher et al, 1986; 1987; 2003; 2005b) as well has having FDA registration.

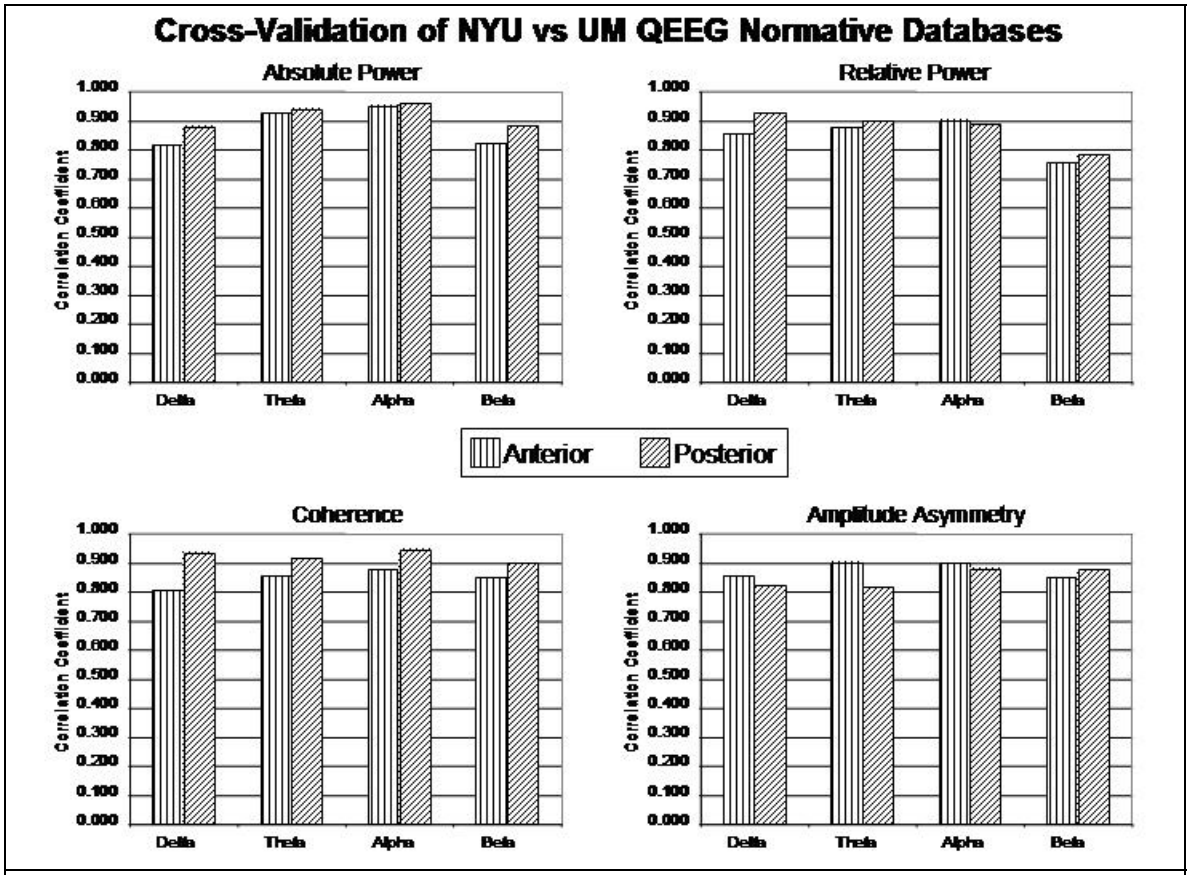

Fig. 4 – Results of an independent cross-validation comparison of Z scores from 332 psychiatric patients ranging in age from 6.2 years to 84.9 years using the NYU and UM normative databases. Anterior and posterior refers to the anterior and posterior location of electrodes. Highly significant independent cross-validation of was observed which shows the high degree of consistency between two peer reviewed and clinically validated QEEG normative databases. (Reprinted with permission from Brian McDonald, CNS Response, Inc.

History of Amplifier Matching and QEEG Normative Databases

Surprisingly, this particular standard was largely neglected during much of the history of QEEG normative databases. E. Roy John and colleagues (1982 to 1988)

formed a consortium of universities and medical schools that were using QEEG who met several times over a few years and was one of the supporters of the edited volume by John titled “Machinery of the Mind” (John, 1990). One of the important issues consistently raised at the consortium meetings was the need for “standardization”. In the 1980s it was technically difficult to match different EEG systems because of the infantile development of analysis software. This history forced most QEEG users to use relative power because absolute power was not comparable between different EEG machines. There was no frequency response standardization between different EEG machines and thus there was no cross-platform standardization of QEEG. It was not until the mid 1990s that computer speed and software development made amplifier matching and normative database amplifier equilibration a possibility. The first use of standardized matching of amplifiers was to the University of Maryland (UM) database. The procedure involved injecting micro volt calibration sign waves into the input of amplifiers of different EEG machines and then inject the same micro volt signals into the normative database amplifiers thus obtaining two frequency response curves (Thatcher et al, 2003). Equilibration of a normative QEEG database to different EEG machines is the ratio of the frequency response curves of the two amplifiers that are then used as amplitude scaling coefficients in the power spectral analysis. This was an important step because suddenly absolute power Z scores and normative database comparisons became possible. Relative power is a last resort type of measure to be used when there is no equilibration of absolute amplitude because relative power always distorts the spectrum and relative power depends on absolute power in order to interpret relative power. This is because relative power is a percentage of the whole and thus an increase in mid “beta”, e.g., 14 – 18 Hz will be seen as a decrease in “theta”, e.g., 4 – 7 Hz when in fact there is no change in theta and vice versa. The frequencies in absolute power are independent of each other and are not distorted. It is always best to use absolute values when ever possible and not relative values or even ratios. A ratio can change due to the denominator or the numerator and one can not determine which has changed without evaluating the absolute values used to compute the ratios.

As illustrated in Figure 5, a simple method of amplifier equilibration to exactly match the frequency characteristics of different amplifiers is to calibrate the amplifiers using micro-volt sine waves at discrete frequencies from 1 to 30 Hz (or whatever frequency range matches the normative database amplifiers) and then injecting the same calibrated sine waves into the inputs of the EEG amplifier to be compared to the normative database amplifiers. Then take the ratio of the micro-volt values at each frequency and use the ratios as gain or amplitude scalars in the FFT to exactly equate the spectral output values to the normative database amplifiers. This method creates a universal equilibration process so that micro-volts in a given amplifier are equal to micro-volts in all other amplifiers including the normative database amplifiers. By equilibrating amplifiers then direct comparisons between a given patient’s EEG and the normative database means and standard deviations is valid and meaningful. If amplifier matching is not accomplished then all normative database comparisons are potentially invalid and caution should be exercised not to use a normative database when amplifiers have not been equilibrated. We have found that amplifiers differ primarily from 0 to 2 Hz and in order to accurately match to the normative database amplifiers one can filter at 1 Hz, thus avoiding mismatches at less than 1 Hz. There are a wide variety of different

frequency response curves for different amplifiers and there is no one "gold standard" for EEG amplifiers. For older amplifiers that have a more limited frequency response, e.g., the NYU and University of Maryland amplifiers and Biologic, Grass and Cadwell, etc. then the match of frequencies is limited to the frequency range that is in common between the two amplifier systems. For example, Deymed has a nearly flat response from 0.5 Hz to 70 Hz and thus the match to the NYU & U of M amplifiers is only from 0.5 Hz to 30 Hz because the latter amplifiers used cut-off filters at approximately 30 Hz. Many amplifiers currently in use also have cut-off frequencies of around 30 Hz but there is still a lot of information in the EEG from 0.5 Hz to 30 Hz and equilibration is necessary to optimally use these amplifiers in a normative database comparison.

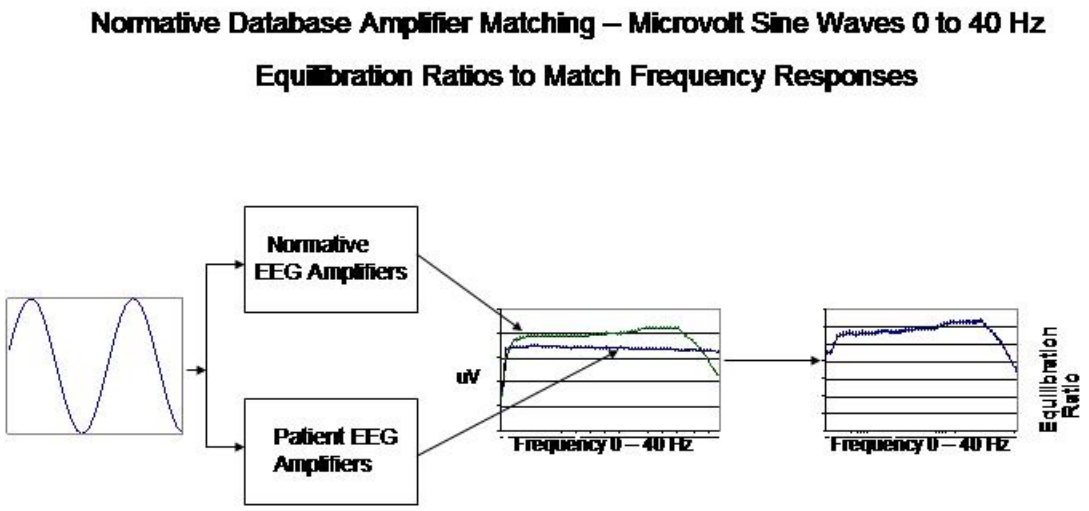

Fig. 5 – Flow chart of the amplifier standardization procedure. Microvolt (uV) sine waves are injected into the input of amplifiers and the frequency responses are calculated. The frequency response of the normative database amplifiers and the frequency response of the Deymed amplifier are in the middle graph. As shown in the right graph, EEG amplifier systems are then equated as the ratio of the two amplifier frequency response curves and the spectral analysis is adjusted based on the equilibration ratios or amplitude scalars (graph on the right) so that there is a standardized import and matching of amplifier systems with the common unit being micro volts (uV).

Figure 6 shows the frequency response curves of the University of Maryland and the NYU normative database amplifiers. It can be seen that the normative database amplifier is approximately 3 db down at 0.5 Hz and 27.5 Hz and approximately 35% attenuation at 30 Hz. Because of the sharp high and low frequency cut-offs of the normative database amplifiers the University of Maryland normative comparisons range

from 1.0 Hz to 30.5 Hz; a frequency range in which there is sufficient signal to insure accurate matching and equilibration to different amplifiers. For example, the NYU and University of Maryland amplifiers exhibit about 1 uV of peak-to-peak noise at 30 Hz and EEG beta frequency activity at 25 Hz to 30 Hz exhibits peak-to-peak amplitudes of about 6 uV to 12 uV. Thus, at 30 Hz there is adequate signal for amplifier matching.

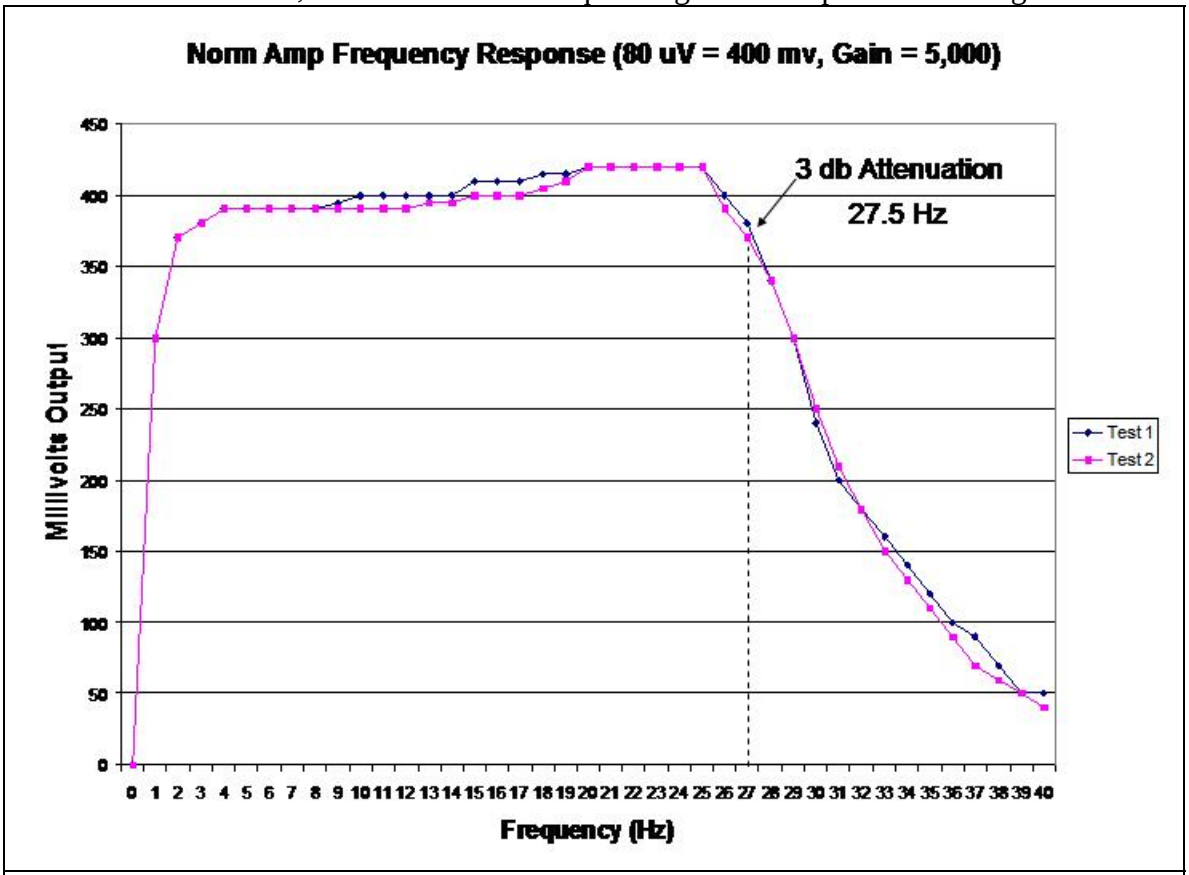

Fig. 6 – The frequency response of the University of Maryland normative database amplifiers. Sine waves at 80 uV peak-to-peak were used in the calibration and at a gain of 5,000 = 400 mV output. The 3 db high and low frequency cut-offs are approximately 1 Hz and 27.5 Hz. The amplifier filters were strong enough to insure no violation of the Nyquist sampling theorem at 100 Hz and 128 Hz sampling rates. Two independent calibration tests are shown demonstrating high accuracy.

Content Validity of QEEG Normative Databases: Neuropsychological Correlations

Content validity is defined by the extent to which an empirical measurement reflects a specific domain of content (Nunnally, 1978). For example, a test in arithmetic operations would not be content valid if the test problems focused only on addition, thus neglecting subtraction, multiplication and division. By the same token, a content-valid measure of cognitive decline following a stroke should include measures of memory capacity, attention and executive function, etc.

There are many examples of the clinical content validity of QEEG and normative databases in ADD, ADHD, schizophrenia, compulsive disorders, depression, epilepsy, TBI (Thatcher et al, 1998a; 1998b) and a wide number of clinical groupings of patients as reviewed by Hughes and John (1999). There are over 250 citations in the review by

Hughes and John and there are approximately twenty three citations to peer reviewed journal articles in which a normal reference database was used. Another recent review of QEEG normative databases and the clinical application of QEEG to psychiatric disorders cited 169 publications (Coburn et al, 2006). An internet search of the National Library of Medicine will give citations to more QEEG and content validity peer-reviewed studies using a reference normal group than were included in the Hughes and John review or the Coburn et al (2006) review. Finally, a recent review that emphasizes clinical correlations and clinical validation of a normative database is by Gordon et al (2005).

Figure 7 and Table III shows an example of the range of clinical correlations to full scale I.Q. in 373 normal individuals from 5 years to 55 years of age (2005c).

Table III
List of Correlations between Full Scale I.Q. and QEEG measures from 373 normal subjects age 5 years to 55 years (47).

QEEG Measure	Correlation Coefficient – QEEG and Full Scale I.Q. (Wisc-R)
Phase Difference	0.859
Coherence	0.842
Phase Reset per Second	0.785
Phase Reset Locking Interval Means	0.780
Amplitude Asymmetry	0.691
Phase Reset Duration Means	0.688
Burst Amplitude Means	0.574
Out-of-Phase Cross-Spectral Power	0.570
Cross Spectral Power	0.485
In-Phase Cross-Spectral Power	0.481
Absolute Power	0.443
Phase Reset Amplitude Means	0.372
Peak Frequency	0.218

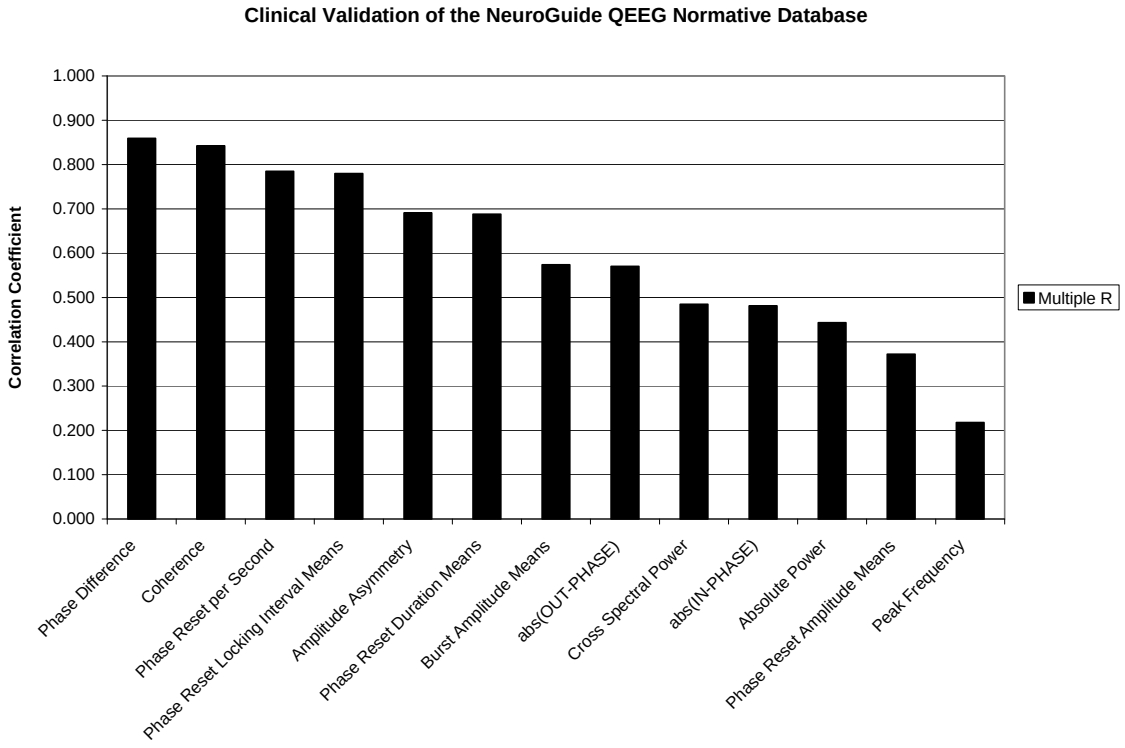

Fig. 7 – Correlations between QEEG measures and full scale I.Q. (WISC-R). N = 332 subjects from the University of Maryland QEEG normative database (see Table III). The highest correlations between QEEG and I.Q. are phase differences and coherence (47). The x-axis are different QEEG measures and the y-axis is the correlation coefficient in a multivariate regression analysis with full scale I.Q. as the dependent variable. Phase reset and Burst metrics are new measures which also exhibit high clinical correlations and clinical validation.

It can be seen in figure 7 that relative high correlations with I.Q. (0.859) are achievable when using a normative database and multiple regression of different variable types and that different QEEG measures exhibit different magnitudes of correlation. The multiple regression prediction of I.Q. is not intended to replace neuropsychological tests. However, an advantage of a QEEG normative database prediction of I.Q. is that it can be repeated without confounding by learning and it can be given to un-testable patients such as stroke, paralysis and uncooperative individuals. Also, QEEG predictions of intelligence provide an insight into which aspects of neural functioning such as location and connectivity contribute to the prediction of intelligence thus providing a deeper understanding of intelligence in an individual subject.

Content Validity of QEEG Normative Databases: Example for Traumatic Brain Injury

There are numerous peer reviewed journal articles showing high correlations between Z scores involving the UM and NYU and other normative databases over the last 20 years (see review by Hughes and John, 1999). It is beyond the scope of this chapter to attempt to review all of these studies. Instead, we will focus on one of the many clinical correlation sub-groups, namely, traumatic brain injury. The National Library of Medicine lists 1,672 peer reviewed journal articles on the subject of EEG and traumatic

brain injury. The vast majority of these studies involved quantitative analyses and, in general, the scientific literature presents a consistent and common quantitative EEG pattern correlated with TBI. Namely, reduced amplitude of the alpha and beta and gamma frequency bands of EEG (8 – 12 Hz and 13 – 25 Hz and 30 - 40 Hz) (Mas et al, 1993; von Bierbrauer et al, 1993; Ruijs et al, 1994; Korn et al, 2005; Hellstrom-Westas, 2005; Thompson et al, 2005; Tebano et al, 1988; Thatcher et al, 1998a; 2001a; Roche et al, 2004; Slewa-Younan, 2002; Slobounov et al, 2002) and changes in EEG coherence and phase delays in frontal and temporal relations (Thatcher et al, 1989; 1991; 1998b; 2001; Hoffman et al, 1995; 1996a; Trudeau et al, 1998). The reduced amplitude of EEG is believed to be due to a reduced number of synaptic generators and/or reduced integrity of the protein/lipid membranes of neurons (Thatcher et al, 1997; 1998a; 2001b). EEG coherence is a measure of the amount of shared electrical activity at a particular frequency and is analogous to a cross-correlation coefficient. EEG coherence is amplitude independent and reflects the amount of functional connectivity between distant EEG generators (Nunez, 1981; 1994; Thatcher et al, 1986). EEG phase delays between distant regions of the cortex are mediated in part by the conduction velocity of the cerebral white matter which is a likely reason that EEG phase delays are often distorted following a traumatic brain injury (Thatcher et al, 1989; 2001a). In general, the more severe the traumatic brain injury then the more deviant the QEEG measures (Thatcher et al, 2001a; 2001b).

Quantitative EEG studies of the diagnosis of TBI typically show quite high sensitivity and specificity, even for mild head injuries. For example, a study of 608 mild TBI patients and 103 age matched control subjects demonstrated discriminant sensitivity = 96.59%; Specificity = 89.15%, Positive Predictive Value (PPV) = 93.6% (Average of tables II, III, V) and Negative Predictive Value (NPV) = 97.4% (Average of tables III, IV, V) in four independent cross-validations. A similar sensitivity and specificity for QEEG diagnosis of TBI was published by Trudeau et al (1998) and Thatcher et al (2001a). All of these studies met most of the American Academy of Neurology's criteria for diagnostic medical tests of: 1- the "criteria for test abnormality was defined explicitly and clearly", 2- control groups were "different from those originally used to derive the test's normal limits", 3- "test-retest reliability was high", 4- the test was more sensitive than "routine EEG" or "neuroimaging tests" and, 5- the study occurred in an essentially "blinded" design (i.e., objectively and without ability to influence or bias the results).

History of 3-Dimensional Current Source Normative Databases

Parametric statistics that rely upon a Gaussian distribution have been successfully used in studies of Low Resolution Electromagnetic Tomography or LORETA (Thatcher et al, 2005a; 2005b; Huizenga et al, 2002; Hori and He, 2001; Waldorp et al, 2001; Bosch-Bayard et al, 2001; Machado et al, 2004). Bosch-Bayard et al (2001) created a Z score normative database that exhibited high sensitivity and specificity using a variation of LORETA called VARETA. A subsequent study by Machado et al (2004) extended these analyses again using VARETA. Thatcher et al (2005a) also showed that LORETA current values in wide frequency bands approximate a normal distribution after transforms with reasonable sensitivity. This same paper compared Z scores to non-parametric statistical procedures and showed that Z scores were more accurate than non-parametric statistics (2005a). Lubar et al (2003) used non-parametric statistics in an

experimental control study with similar levels of significance as reported by Thatcher et al (2005a). Figure 8 shows an example of how a log transform can move a non-gaussian distribution toward a better approximation to a Gaussian when using LORETA (Thatcher et al, 2005a; 2005b).

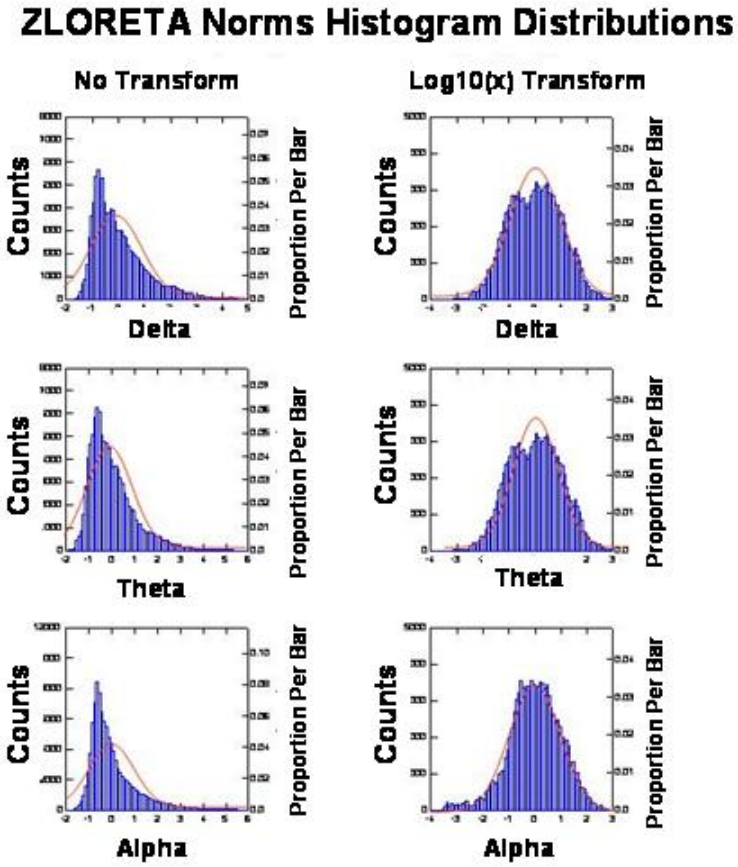

Fig. 8- Shows the distribution of current source densities before (left) and after (right) $\log_{10}$ transform for the delta, theta and alpha frequencies. It can be seen that reasonable approximation to Gaussian was achieved by the $\log_{10}$ transform. (From Thatcher et al, 2005a).

LORETA 3-dimensional current source normative databases have also been cross-validated and the sensitivity computed using the same methods as for the surface EEG (Thatcher et al, 2005b). Figure 9 shows an example of localization accuracy of a LORETA normative database in the evaluation of confirmed neural pathologies.

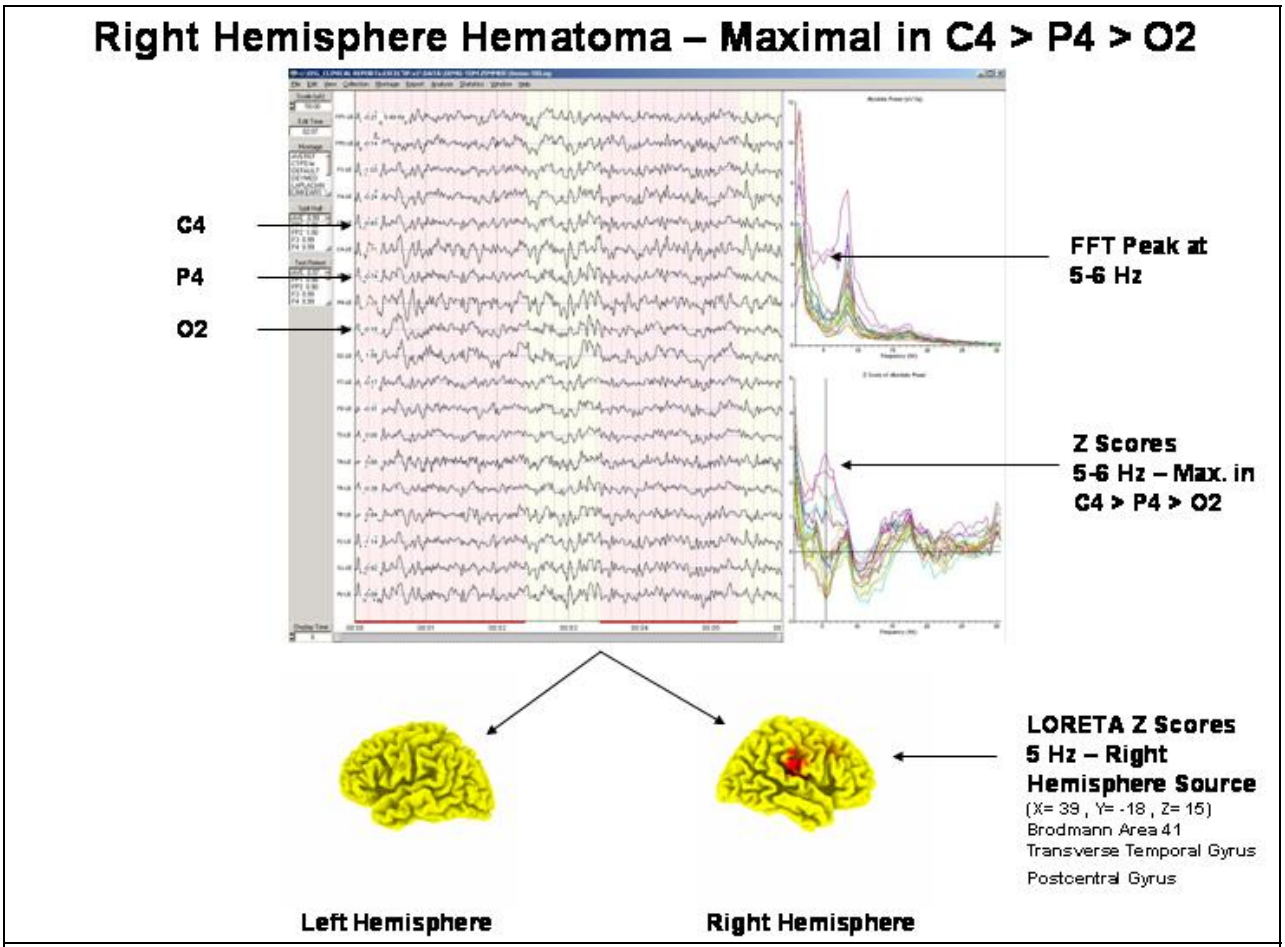

Fig. 9 - Top is the EEG from a patient with a right hemisphere hematoma where the maximum slows waves are present in C4, P4 and O2. The FFT power spectrum from 1 to 30 Hz and the corresponding Z scores of the surface EEG are shown in the right side of the EEG display. Bottom, are the left and right hemisphere displays of the maximal Z scores using LORETA. It can be seen that only the right hemisphere has statistically significant Z values. Planned comparisons and hypothesis testing based on the frequency and location of maximal deviation from normal on the surface EEG are confirmed by the LORETA Z score normative analysis. (From Thatcher et al, 2005b).

All of these studies demonstrated that when proper statistical standards are applied to EEG measures, whether they are surface EEG or 3-dimensional source localization, then high cross-validation accuracy can be achieved. Recently, Hoffman (2006) confirmed that high accuracy can be achieved using a LORETA Z score normative database to evaluate patient's with confirmed pathologies (e.g., left temporal lobe epilepsy and focal brain damage) using the University of Maryland normative database (Thatcher et al, 2003) and the University of Tennessee normative database (Lubar et al, 2003).

History of 3-Dimenional Source Correlation Normative Databases

Thatcher et al (1994), Thatcher (1995) and Hoechstetter et al (2004) used a multiple dipole source solution for scalp EEG electrical potentials and then used

coherence to compute the correlation between the 3-dimensional current sources and demonstrated changes in the correlation between current sources related to different tasks. Pascual-Marqui et al (2001) used low resolution electromagnetic tomography (LORETA) to compute current sources and then used a Pearson Product correlation coefficient to explore differences in source correlations between a normal control group and a group of schizophrenic patients. Recently, high statistical standards were applied to LORETA 3-dimenional source correlations in a QEEG normative database (Thatcher et al, 2007a). All of these studies revealed interesting and reproducible relations between current sources and network connectivity that provide a deeper understanding of the surface EEG dynamics.

The same statistical standards as enumerated previously were applied to the LORETA source correlation normative database, i.e., peer reviewed publication, gaussian approximation, removal of artifact, high reliability and cross-validation. The LORETA normative database studies prove that nearly any measure can be used in a normative database as long as the appropriate statistical and scientific standards are met.

History of Real-Time Z-Score Normative Databases

As mentioned above, many different normative databases can be constructed and validated as long as the basic scientific standards of gaussianity, cross-validation, amplifier matching and peer reviewed publications are met. A recent example of a new application of a normative database is the use of complex demodulation as a Joint-Time-Frequency-Analysis (JTFA) for the purposes of real-time biofeedback (Thatcher, 1998a; 1998b; 2000a; 2000b; Thatcher et al, 1987; 2003). This method has recently been implemented in EEG biofeedback systems and used to compute statistical Z scores in real-time. Complex demodulation is an analytic technique that multiples a time series by a sine wave and a cosine wave and then applies a low pass filter (Granger et al, 1964; Otnes and Enochson, 1977; Thatcher et al, 2007b). This results in mapping of the time series to the unit circle or "complex plane" whereby instantaneous power and instantaneous phase differences and coherence are computed. Unlike the Fourier transform which depends on windowing and integration over an interval of time, complex demodulation computes the instantaneous power and phase at each time point and thus an instantaneous Z score necessarily includes the within subject variance of instantaneous electrical activity as well as the between subject variance for subjects of a given age. The summation of instantaneous Z scores is Gaussian distributed and has high cross-validation (Thatcher et al, 1987; 2007), however, the individual time point by time point Z score is always smaller than the summation due to within subject variance. The use of within subject variance results in a more "conservative" estimate of deviation from normal solely for the purposes of instantaneous biofeedback methods. A standard FFT normative database analysis should first be computed in order to identify the electrode locations and EEG features that are most deviant from normal and that can be linked to the patient's symptoms and complaints. Linking a subjects symptoms and complaints, e.g., PTSD, Depression, Schizophrenia, TBI, etc. to functional localization of the brain is an important objective of those who use a normative database. Similar to a blood bank analysis, the list of deviant or normal measures are given to the clinician as one test among many that are used to help render a diagnosis. Linking de-regulation of neural activity in localized regions of the brain to known functional localization, for example,

left parietal lobe and dyslexia; right frontal and depression; cingulate gyrus and attention deficit; occipital lobes and vision problems, etc. are important to make by a trained clinician. Textbooks on functional localization in neurology and psychiatry are available to aid the clinician in learning about the link between a patient's symptoms and different brain regions (Heilman and Valenstein, 1993; Brazis et al, 2007). A link of the anatomical locations and patterns of a patients deviant Z scores is important in order to derive clinical meaning from the QEEG.

Once a QEEG normative database analysis is completed, then one can use a Z score biofeedback program to train patient's to move their instantaneous Z scores toward zero or the norm. The absolute value and range of the instantaneous Z scores while smaller than those obtained using the offline QEEG normative database are nonetheless valid and capable of being minimized toward zero. An advantage of a Z score biofeedback program is simplification by reducing diverse measures to a single metric, i.e., the metric of a Z score. Thus, there is greater standardization and less guess work about whether to reinforce or suppress coherence or phase differences or power, etc. at a particular location and particular frequency band.

Figure 10 shows the number of subjects per year in the normative EEG lifespan database, N = 625, that spans the age range from 2 months to 82 years of age. It can be seen that the largest number of subjects are in the younger ages (e.g., 1 to 14 years, N = 470) when the EEG is changing most rapidly. A proportionately smaller number of subjects represents the adult age range from 14 to 82 years (N = 155). In order to increase the time resolution of age, sliding averages were used for age stratification of the instantaneous Z scores for purposes of EEG biofeedback. Two year means were computed using a sliding average with 6 month overlap of subjects. This produced a more stable and higher age resolution normative database and a total of 21 different age groups. The 21 age groups and age ranges and number of subjects per age group is shown in the bar graph in figure 10.

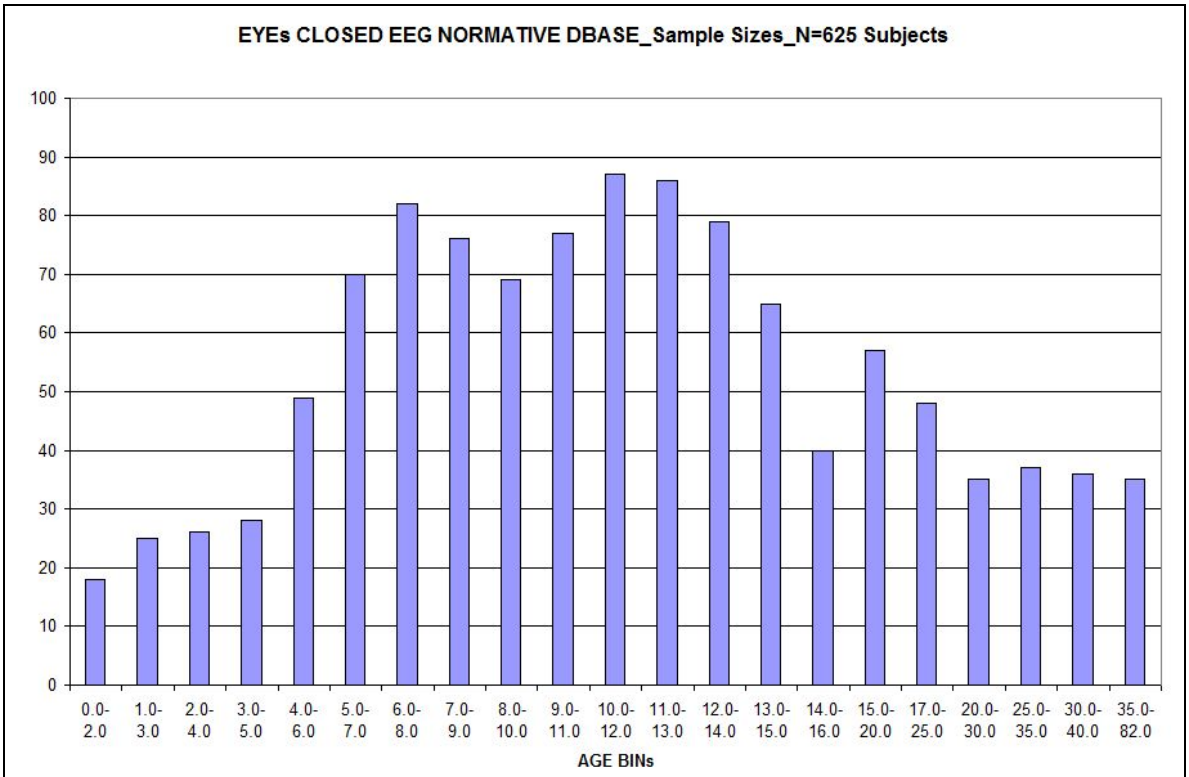

Fig. 10 - The number of subjects per age group in the Z score Lifespan EEG reference normative database. The database is a "life-span" database with the two months of age being the youngest subject and 82.3 years of age being the oldest subject. Two year means were computed using a sliding average with 6 month overlap of subjects. This produced a more stable and higher age resolution normative database and a total of 21 different age groups. The 21 age groups and age ranges and number of subjects per age group is shown in the bar graph (Adapted from Thatcher et al, 2003).

Active Tasks vs Eyes Closed and Eyes Open QEEG Databases

An active task refers to the recording of EEG and/or evoked potentials (EPs) while a subject performs some kind of perceptual or cognitive task. Many EEG and EP and event related potential (ERP) studies have reported reproducible changes in brain dynamics which are task dependent. Such studies are important for understanding normal and pathological brain processes responsible for perceptual and cognitive function. In contrast, an eyes closed or eyes open EEG state involves an alert subject simply sitting quietly and not moving. The eyes closed and/or eyes open conditions are commonly used as reference normative EEG databases because of the simplicity and relative uniformity of EEG recording conditions. Such databases can be compared across laboratories and populations with relatively high reliability. Active tasks, on the other hand, are dependent on the intensity of stimuli, the back ground noise of the room, the distance between the subject and the stimuli, the subject's understanding of the task instructions, the subject's motivation, etc. These are very difficult to control across experimenters or across clinics for the purposes of constructing a "reference" normative EEG database.

One of the most carefully constructed active task normative database is by Brain Resources, Inc. in Australia (Gordon et al, 2005). The BRC database does require replication of specific task conditions using a Neuroscan, Inc. EEG amplifier system.

The relative sensitivity and specificity of resting eyes open and eyes closed EEG versus an active task normative database has not been published to our knowledge. Another well constructed and tested active task normative database is the go no-go task developed by Russian scientists (Kropotov et al 2005) with medium to high sensitivity and accuracy in the evaluation of attention deficits and other disorders. We were unable to find any peer reviewed journal articles of EEG databases produced by Dr. Kropotov and therefore there is no information on the sensitivity, cross-validation, amplifier matching and other standards for EEG databases.

It should be kept in mind that the alert eyes closed EEG state is very much an active state, e.g., there is still about 20% glucose metabolism of the whole body occurring in the brain of an eyes closed subject (Herscovitch, 1994) (Raichle, 2002). During the eyes closed state, there is dynamic circulation of neural activity in connected cortical, reticular and thalamo-cortical loops (Thatcher and John, 1977; Nunez, 1995). The allocation of neural resource is simply different from when the subject is directing his/her attention to an experimentally controlled situation. Active tasks are very important because they reflect the switching and dynamic allocation of neural resource which also has clinical importance. However, a scientifically sound and stable resting EEG normative database can enhance and also facilitate the understanding of the underlying neural dynamics and clinical condition of a patient during an active task. For example, comparison to a resting baseline normative database during different active task conditions may help reveal anatomical localization of neural processes and network dynamics without the need for a comparison to an exactly matching active task.

Summary of Normative Database Validation and Sensitivity Tests

Figure 11 is a summary and overview of the procedures that are used to eliminate artifact, maximize reliability, approximate Gaussian distributions, cross-validate and compute clinical correlations as a standard set of procedures and sequences to make normative database creation more easy.

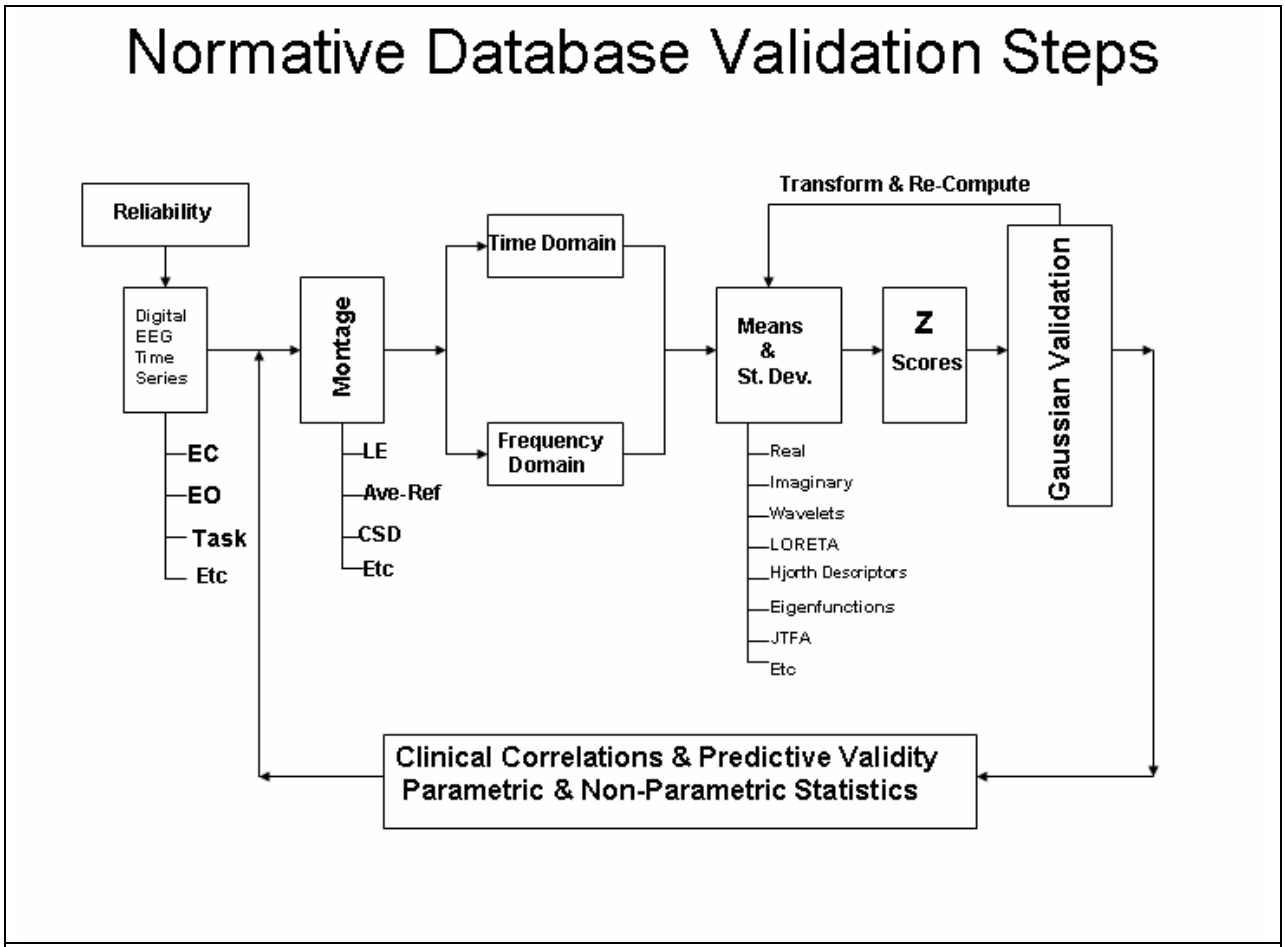

Fig. 11- Illustration of the steps involved in developing a normative QEEG database. The left is the start of the process with data acquisition, amplifier matching, artifact rejection and quality control. Approximation to a Gaussian is followed by cross-validation and then finally clinical correlations.

Figure 11 is an illustration of a step-by-step procedure by which any normative EEG database can be validated and sensitivities calculated. The left side of figure 11 is the edited and artifact clean and reliable digital EEG time series which may be re-referenced or re-montaged, which is then analyzed in either the time domain or the frequency domain.

The selected normal subjects are grouped by age with a sufficiently large sample size. The means and standard deviations of the EEG time series and/or frequency domain analyses are computed for each age group. Transforms are applied to approximate a Gaussian distribution of the EEG measures that comprise the means. Once approximation to Gaussian is completed, Z-scores are computed for each subject in the database and leave one out Gaussian cross-Validation is computed in order to arrive at optimum Gaussian cross-validation sensitivity. Finally the Gaussian validated norms are subjected to content and predictive validation procedures such as correlation with neuropsychological test scores and intelligence, etc. and also discriminant analyses and neural networks and outcome statistics, etc. The content validations are with respect to clinical measures such as intelligence, neuropsychological test scores, school

achievement and other clinical measures. The predictive validations are with respect to the discriminative, statistical or neural network clinical classification accuracy. Both parametric and non-parametric statistics are used to determine the content and predictive validity of a normative EEG database.

Gold Standard Check List for a Normative QEEG Database.

Table IV is a "Gold Standard" check list that summarizes the minimal standards of QEEG Normative databases that were discussed previously. Those clinicians interested in using a QEEG normative database are encouraged to enter a check for each of the standards that a given database has met. The more standards that are met then the better.

Table IV
List of "Gold Standards" by which to judge
QEEG Normative databases

	Standards	Yes	No
1	Amplifier Matching		
2	Peer reviewed publications		
3	Artifact Rejection		
4	Test Re-Test Reliability		
5	Inclusion/exclusion criteria		
6	Adequate Sample size per age group		
7	Approximation to a Gaussian		
8	Cross-Validation		
9	Clinical Correlation		
10	FDA Registered		

Problems in combining Sub-Standard QEEG Databases with scientifically acceptable databases

Often an EEG data sample from a patient is sent to a laboratory or QEEG service, and the data is compared to multiple databases including sub-standard databases. As expected, the results are often conflicting and contradictory and confusing. There is an assumption that somehow multiple comparisons to multiple databases is better than comparing a patient's EEG to a single well published database that has met high statistical and scientific standards. This assumption is wrong and potentially dangerous to unsuspecting patients and clinicians who are provided with multiple comparisons. If a patient or a clinician receives multiple database comparisons involving unmatched amplifier characteristics then they should ask the provider of the normative database for the methods of amplifier equilibration and for a list of the scientific standards of the normative databases. It is the responsibility of users of normative databases to know the scientific standards of the database that they are comparing their patient's to and to provide informed consent to patients in situations where the patient's EEG samples are compared to a non peer-reviewed database and/or unknown number of subjects per year database and/or unknown inclusion/exclusion criteria database and/or no statistical validation test database and/or a non-FDA registered database, etc. State law and the FDA and IRBs require wording in an informed consent form that is clear and

unambiguous in which the patient is informed that their EEG data will be compared to an unpublished or otherwise unknown QEEG normative database. Hopefully the "Gold Standards" check list in Table IV will help in this process.

Future Standardization of QEEG Normative Databases

The post Newtonian period of European history (1685 – 1850s) is marked by an emphasis on standards and rules as an outgrowth of Newtonian mathematics in the 1600s. It was recognized that standards were a prerequisite for the future industrial revolution involving mass production and efficient engineering and growth of new knowledge. A similar need for standardization of QEEG normative databases is present today. Amplifier equilibration and standardization has long been an elusive goal as mentioned previously. However, new technologies are available that provide for simple and inexpensive standardization of EEG amplifiers for purposes of comparison.

In the future the essential standard will be to equate the microvolt measurement of the electrical energies of the human brain recorded at different frequencies from different amplifiers using accepted statistical tests and standards of validation and verification as listed in rows 2 to 10 in Table IV.

References

Adey, W.R., Walter, D.O. and Hendrix, C.E. (1961). Computer techniques in correlation and spectral analyses of cerebral slow waves during discriminative behavior.Exp Neurol., 3:501-524

Adey, W.R. (1964a). Data acquisition and analysis techniques in a Brain Research Institute. Ann N Y Acad Sci., 31;115:844-866.

Adey, W.R. (1964b). Biological instrumentation, electrophysiological recording and analytic techniques. Physiologist., 72:65-68.

Arruda JE, Weiler MD, Valentino D, Willis WG, Rossi JS, Stern RA, Gold SM, Costa L. (1996). A guide for applying principal-components analysis and confirmatory factor analysis to quantitative electroencephalogram data. *Int J Psychophysiol* , 23(1-2), 63-81.

Berger, H. (1929). Uber das Electrenkephalogramm des Menschen. *Archiv. Fur. Psychiatrie und Neverkrankheiten*, 87, 527-570.

Berger, H. (1932). Uber das Electrenkephalogramm des Menschen. Vierte Mitteilungj. *Archiv. Fur. Psychiatrie und Neverkrankheiten*, 97: 6-26.

Berger, H. (1934). Uber das Electrenkephalogramm des Menschen. Neunte Mitteilungj. *Archiv. Fur. Psychiatrie und Neverkrankheiten*, 102: 538-557.

Brazis, P.W., Masdeu, J.C. and Biller, J. (2007). Localization in Clinical Neurology. Lippincott Williams and Wilkins, Philadelphia, PA.

Bosch-Bayard J, Valdes-Sosa P, Virues-Alba T, Aubert-Vazquez E, John ER, Harmony T, Riera-Diaz J, Trujillo-Barreto N. (2001). 3D statistical parametric mapping of EEG source spectra by means of variable resolution electromagnetic tomography (VARETA). Clin Electroencephalogr., 32(2):47-61.

Burgess A, and Gruzelier J. (1993). Individual reliability of amplitude distribution in topographical mapping of EEG. *Electroencephalogr Clin Neurophysiol* ., 86(4), 219-223.

Coburn, K.L., Lauterback, E.C., Boutros, N.N., Black, K.J., Arciniegas, D.B. and Coffey, C.E. (2006). The value of quantitative electroencephalography in clinical psychiatryu: A report by the committee on research of the American Neuropsychiatric Association. J. Neuropsychiat. and Clin. Neurosci. 18: 460-500.

Cohen, J. (1977). Statistical power analysis for the behavioral sciences. New York: Academic Press.

Cooley, W.W. and Lohnes, P.R. (1971). Multivariate Data Analysis. John Wiley & Sons, New York.

Corsi-Cabrera M, Solis-Ortiz S, Guevara MA. (1997). Stability of EEG inter- and intrahemispheric correlation in women. , *Electroencephalogr Clin Neurophysiol*, 102(3), 248-255.

Duffy, F., Hughes, J. R., Miranda, F., Bernad, P. & Cook, P. (1994). Status of quantitative EEG (QEEG) in clinical practice. *Clinical. Electroencephalography, 25* (4), VI - XXII.

Gasser, T., Verleger, R., Bacher, P., & Sroka, L. (1988a). Development of the EEG of school-age children and adolescents. I. Analysis of band power. *Electroencephalography and Clinical Neurophysiology, 69* (2), 91-99.

Gasser, T., Jennen-Steinmetz, C., Sroka, L., Verleger, R., & Mocks, J. (1988b). Development of the EEG of school-age children and adolescents. II: Topography. *Electroencephalography Clinical Neurophysiology, 69* (2), 100-109.

Gordon, E., Cooper, N., Rennie, C., Hermens, D. and Williams, L.M. (2005). Integrative neuroscience: The role of a standardized database. Clin. EEG and Neurosci., 36(2): 64-75.

Granger, C.W.J. and Hatanka, M. (1964). Spectral Analysis of Economic Time Series, Princeton University Press, New Jersey.

Hagne, I., Persson, J., Magnusson, R. and Petersen, I. (1973). Spectral analysis via fast Fourier transform of waking EEG in normal infants. In P. Kellaway & I. Petersen (Eds.), *Automation of clinical electroencephalography* (pp. 103-143). New York:

Raven Press.

Hamilton-Bruce MA, Boundy KL, Purdie GH. (1991). Interoperator variability in quantitative electroencephalography., *Clin Exp Neurol.*, 28,219-224, 1991.

Harmony T, Fernandez T, Rodriguez M, Reyes A, Marosi E, Bernal J. (1993). Test-retest reliability of EEG spectral parameters during cognitive tasks: II. Coherence. *Int J Neurosc*i , 68(3-4), 263-271.

Hayes, W. L. (1973). *Statistics for the social sciences* .New York: Holt, Rhinehart and Winston.

Heilman, K.M. and Valenstein, E. (1993). Clinical Neuropsychology (3rd ed.)., Oxford University Press, New York.

Hellstrom-Westas L, Rosen I. (2005). Electroencephalography and brain damage in preterm infants. Early Hum Dev. 81(3):255-61.

Herscovitch, P. (1994). Radiotracer techniques for functional neuroimaging with positron emission tomography. In: R.W. Thatcher, M. Halletr, T. Zeffro, E. R. John and M. Huerta, Functional Neuroimaging: Technical Foundations, Academic Press, San Diego.

Hoechstetter, K, Bornfleth H, Weckesser D, Ille N, Berg P, Scherg M (2004). BESA source coherence: A new method to study cortical oscillatory coupling. Brain Topography, 16: 233-238.

Hoffman DA, Stockdale S, Hicks L, et al: (1995). Diagnosis and treatment of head injury. *Journal of Neurotherapy*., 1(1), 14-21.

Hoffman DA, Stockdale S, Van Egren L, et al (1996). Symptom changes in the treatment of mild traumatic brain injury using EEG neurofeedback. Clinical Electroencephalography (Abstract). 27(3), 164.

Hoffman, D. (2006). LORETA: An attempt at a simple answer to a complex controversy. J. Neurotherapy, 10(1): 57-72.

Hori J. and He, B. (2001). Equivalent dipole source imaging of brain electric activity by means of parametric projection filter. Ann Biomed Eng. 2001, ;29(5):436-45.

Huizenga HM, de Munck JC, Waldorp LJ, Grasman RP. (2002). Spatiotemporal EEG/MEG source analysis based on a parametric noise covariance model. IEEE Trans Biomed Eng., 49(6):533-539.

Hughes, J. R. & John, E. R. (1999). Conventional and quantitative electroencephalography in psychiatry. *Neuropsychiatry, 11*, 190-208.

John, E.R. Functional Neuroscience, Vol. II: Neurometrics: Quantitative Electrophysiological Analyses. E.R. John and R.W. Thatcher, Editors. L. Erlbaum Assoc., N.J., l977.

John, E.R. Karmel, B., Corning, W. Easton, P., Brown, D., Ahn, H., John, M., Harmony, T., Prichep, L., Toro, A., Gerson, I., Bartlett, F., Thatcher, R., Kaye, H., Valdes, P., Schwartz, E. (1977). Neurometrics: Numerical taxonomy identifies different profiles of brain functions within groups of behaviorally similar people. Science, l96, :1393-1410.

John, E. R., Prichep, L. S. & Easton, P. (1987). Normative data banks and neurometrics: Basic concepts, methods and results of norm construction. In A. Remond (Ed.), *Handbook of electroencephalography and clinical neurophysiology: Vol. III. Computer analysis of the EEG and other neurophysiological signals* (pp. 449-495). Amsterdam: Elsevier.

John, E.R., Ahn, H., Prichep, L.S., Trepetin, M., Brown, D. and Kaye, H. (1980) Developmental equations for the electroencephalogram. Science, 210: 1255–1258.

John, E. R., Prichep, L. S., Fridman, J. & Easton, P. (1988). Neurometrics: Computer assisted differential diagnosis of brain dysfunctions. *Science, 293*, 162-169.

John, E.R. (1990). Machinery of the Mind: Data, theory, and speculations about higher brain function. Birkhauser, Boston.

Korn, A., Golan, H., Melamed, I., Pascual-Marqui, R., Friedman, A. (2005). Focal cortical dysfunction and blood-brain barrier disruption in patients with Postconcussion syndrome. *Journal of Clinical Neurophysiology, 22(1)*, 1-9.

Kropotov JD, Grin-Yatsenko VA, Ponomarev VA, Chutko LS, Yakovenko EA, Nikishena IS. (2005). ERPs correlates of EEG relative beta training in ADHD children. Int J Psychophysiol., 55(1):23-34.

Lubar, J. F., Congedo, M., & Askew, J. (2003). Low-resolution electromagnetic tomography (LORETA) of cerebral activity in chronic depressive disorder. International Journal of Psychophysiology, *49*, 175-185.

Lund TR, Sponheim SR, Iacono WG, Clementz BA. (1995). Internal consistency reliability of resting EEG power spectra in schizophrenic and normal subjects. *Psychophysiology*, 32(1), 66-71.

Machado, C., Cuspineda, E., Valdes, P., Virues, T., Llopis, F., Bosch, J., Aubert, E., Hernandez, E., Pando, A., Alvarez, M., A., Barroso, Es., Galan, L. and Avila, Y. (2004). Assessing acute middle cerebral artery ischemic stroke by quantitative electric tomography. Clin. EEG and Neurosci., 35(2), 116 – 124.

Mas, F., Prichep, L.S., Alper, K. (1993). Treatment resistant depression in a case of minor head injury: an electrophysiological hypothesis. *Clinical Electroencephalography, 24(3),* 118-22.

Matousek, M. & Petersen, I. (1973a). Automatic evaluation of background activity by means of agee-dependent EEG quotients. EEG & Clin. Neurophysiol., 35: 603-612.

Matousek, M. & Petersen, I. (1973b). Frequency analysis of the EEG background activity by means of age dependent EEG quotients. In P. Kellaway & I. Petersen (Eds.), *Automation of clinical electroencephalography* (pp. 75-102). New York: Raven Press.

Nunez, P. (1981). Electrical Fields of the Brain. Oxford Univ. Press, New York.

Nunez, P. (1995). Neocortical dynamics and human EEG rhythms. Oxford Univ. Press, New York.

Nunnally, J.C. (1978). Psychometric theory. New York: McGraw-Hill.

Otnes, R. K. & Enochson, L. (1972). *Digital time series analysis.* New York: John Wiley and Sons.

Pascual-Marqui, R.D., Koukkou, M., Lehmann, D. and Kochi, K. (2001). Functional localization and functional connectivity with LORETA comparison of normal controls and first episode drug naïve schizophrenics. J. of Neurotherapy, 4(4): 35-37.

Pollock VE, Schneider LS, Lyness SA. (1991). Reliability of topographic quantitative EEG amplitude in healthy late-middle-aged and elderly subjects. *Electroencephalogr Clin Neurophysiol.,* 79(1), 20-26.

Prichep, L.S. (2005). Use of normative databases and statistical methods in demonstrating clinical utility of QEEG: Importance and cautions. Clin. EEG and Neurosci., 36(2): 82 – 87.

Raichle, M. (2002). Appraising the brain's energy budget" PNAS vol. 99, no.16: 10237-10239.

Roche RA, Dockree PM, Garavan H, Foxe JJ, Robertson IH, O'Mara SM. (2004). EEG alpha power changes reflect response inhibition deficits after traumatic brain injury (TBI) in humans. Neurosci Lett., 13;362(1):1-5.

Ruijs, M.B., Gabreels, F.J., Thijssen, H.M. (1994). The utility of electroencephalography and cerebral computed tomography in children with mild and moderately severe closed head injuries. *Neuropediatrics, 25(2),* 73-7.

Salinsky MC, Oken BS, Morehead L. (1991). Test-retest reliability in EEG frequency analysis.. *Electroencephalogr Clin Neurophysiol.*, 79(5), 382-392.

Slewa-Younan S, Green AM, Baguley IJ, Felmingham KL, Haig AR, Gordon E. (2002). Is 'gamma' (40 Hz) synchronous activity disturbed in patients with traumatic brain injury? Clin Neurophysiol., 113(10):1640-1646.

Slobounov S, Sebastianelli W, Simon R. (2002). Neurophysiological and behavioral concomitants of mild brain injury in collegiate athletes. Clin Neurophysiol., 113(2):185-93.

Tebano, M.T., Cameroni, M., Gallozzi, G., Loizzo, A., Palazzino, G., Pezzino, G., Pezzini, G. and Ricci, G.F. (1988). EEG spectral analysis after minor head injury in man. *EEG and Clinical Neurophysiology, 70,* 185-189.

Thatcher, R.W. (2000a). EEG Operant Conditioning (Biofeedback) and Traumatic Brain Injury. . Clinical EEG, 31(1): 38-44.

Thatcher, R.W. (2000b) "An EEG Least Action Model of Biofeedback" 8th Annual ISNR conference, St. Paul, MN, September.

Thatcher, R. W., Walker, R. A. & Guidice, S. (1987). Human cerebral hemispheres develop at different rates and ages. *Science, 236,* 1110-1113.

Thatcher, R.W., Walker, R.A., Biver, C., North, D., Curtin, R., (2003). Quantitative EEG Normative databases: Validation and Clinical Correlation, J. Neurotherapy, 7 (No. ¾): 87 – 122.

Thatcher, R. W. (1998). EEG normative databases and EEG biofeedback. *Journal of Neurotherapy, 2* (4), 8-39.

Thatcher, R.W., North, D., and Biver, C. EEG inverse solutions and parametric vs. non-parametric statistics of Low Resolution Electromagnetic Tomography (LORETA). Clin. EEG and Neuroscience, Clin. EEG and Neuroscience, 36(1), 1 – 9, 2005a.

Thatcher, R.W., North, D., and Biver, C. Evaluation and Validity of a LORETA normative EEG database. Clin. EEG and Neuroscience, 2005b, 36(2): 116-122.

Thatcher, R.W., McAlaster, R., Lester, M.L., Horst, R.L. and Cantor, D.S. (1983).

Hemispheric EEG Asymmetries Related to Cognitive Functioning in Children. In: Cognitive Processing in the Right Hemisphere, A. Perecuman (Ed.), New York: Academic Press.

Thatcher, R.W., Krause, P and Hrybyk, M. (1986). Corticocortical Association Fibers and EEG Coherence: A Two Compartmental Model. Electroencephalog. Clinical Neurophysiol., 64: 123 - 143.

Thatcher, R.W. Tomographic EEG/MEG. (1995). Journal of Neuroimaging, 5, 35-45.

Thatcher, R., Wang, B., Toro, C. and Hallett, M. (1994). Human Neural Network Dynamics Using Multimodal Registration of EEG, PET and MRI. In: R. Thatcher, M. Hallett, T. Zeffiro, E. John and M. Huerta (Eds.), Functional Neuroimaging: Technical Foundations, Academic Press: New York.

Thatcher, R.W., Biver, C. J., and North, D. (2007). Spatial-temporal current source correlations and cortical connectivity, Clin. EEG and Neuroscience, 38(1): 35 – 48.

Thatcher, R.W. and Collura, T.F. (2006). Z-Score EEG Biofeedback. Int. Soc. for Neuronal Regulation, Atlanta, GA, Sept. 2006.

Thatcher, R. W., Biver, C., McAlaster, R & Salazar, A. M. (1998a). Biophysical linkage between MRI and EEG coherence in traumatic brain injury. *NeuroImage, 8* (4), 307-326.

Thatcher, R. W., Biver, C., Camacho, M., McAlaster, R. & Salazar, A. M. (1998b). Biophysical linkage between MRI and EEG amplitude in traumatic brain injury. *NeuroImage, 7*, 352-367.

Thatcher, R.W., North, D., and Biver, C. (2005c). EEG and Intelligence: Univariate and Multivariate Comparisons Between EEG Coherence, EEG Phase Delay and Power. Clinical Neurophysiology, 116(9):2129-2141.

Thatcher, R.W. and John, E.R. (1977). Functional Neuroscience, Vol. 1: Foundations of Cognitive Processes, E.R. John and R.W. Thatcher, Editors. L. Erlbaum Assoc., N.J.

Thompson, J., Sebastianelli, W., Slobounov, S. (2005). EEG and postural correlates of mild traumatic brain injury in athletes. *Neuroscience Letters, 4 377(3),* 158-163.

Thatcher, R.W., Cantor, D.S., McAlaster, R., Geisler, F. and Krause, P. Comprehensive predictions of outcome in closed head injury: The development of prognostic equations. (1991). Annals New York Academy of Sciences, 620: 82-104.

Thatcher, R.W., Walker, R.A., Gerson, I. and Geisler, F. (1989). EEG discriminant analyses of mild head trauma. EEG and Clin. Neurophysiol., 73: 93-106.

Thatcher, R.W., North, D., Curtin, R., Walker, R.A., Biver, C., J.F. Gomez M., and Salazar, A. (2001a). An EEG Severity Index of Traumatic Brain Injury, J. Neuropsychiatry and Clinical Neuroscience, 13(1): 77-87

Thatcher, R.W., Camacho, M., Salazar, A., Linden, C., Biver, C. and Clarke, L. (1997). Quantitative MRI of the gray-white matter distribution in traumatic brain injury. J. Neurotrauma, 14: 1-14.

.

Thatcher R.W., Biver, C.L., Gomez-Molina J.F., North, D., Curtin, R. and Walker, R.W., and Salazar, A. (2001b). Estimation of the EEG Power Spectrum by MRI T2 Relaxation Time in Traumatic Brain Injury. Clinical Neurophysiology, 112: 1729-1745.

Thatcher, R.W., North, D., and Biver, C. (2007b). Self-organized criticality and the development of EEG phase reset. Human Brain Mapping (In press, 2007).

Trudeau, D.L., Anderson, J., Hansen, L.M., Shagalov, D.N., Schmoller, J., Nugent, S. and Barton, S. Findings of mild traumatic brain injury in combat veterans with PTSD and a history of blast concussion", *J. Neuropsychiatry Clin Neurosci*., 10(3),308-313, 1998.

von Bierbrauer, A., Weissenborn, K., Hinrichs, H., Scholz, M., Kunkel, H. (1993). Automatic (computer-assisted) EEG analysis in comparison with visual EEG analysis in patients following minor cranio-cerebral trauma (a follow-up study). *EEG EMG Z Elektroenzephalogr Elektromyogr Verwandte Geb. 23(3),* 151-157.

Waldorp LJ, Huizenga HM, Dolan CV, Molenaar PC. (2001). Estimated generalized least squares electromagnetic source analysis based on a parametric noise covariance model. IEEE Trans Biomed Eng., 48(6):737-41.

Winer, B.J. (1971). Statistical Principles in Experimental Design. McGraw-Hill, New York.