

Principles and Statistics of Individualized Live and Static Z-Scores

Thomas F. Collura^{1*} and Jeffrey Tarrant²

¹BrainMaster Technologies, Inc., and the Brain Enrichment Center, Bedford, Ohio, USA

²NeuroMeditation Institute, LLC, Eugene, Oregon, USA

Abstract

This report describes and briefly characterizes a method for computing quantitative EEG (qEEG) z-scores based on a modification of the typical methods used for qEEG reporting. In particular, it describes using a sample of EEG from a single individual, and creating a reference database from the individual sample, in contrast to using a population of individuals as the source data. The goal of this method is to quantify and localize within-subject changes that may arise due to time or various factors. We refer to this approach as “z-builder,” because the z-score reference is constructed or “built” on a per-subject basis in the office or laboratory and is not derived from a reference obtained from an outside source. It is confirmed that z-scores for EEG acquired during a test period can be calculated based on a single previously recorded reference sample from an individual, and that the resulting z-scores obey the expected statistical distribution. Reference data can be calculated using samples in the 1- to 5-minute range, and subsequent static or dynamic z-scores for a test sample can then be computed using this reference data in lieu of a population database. It is confirmed that, in the absence of systematic change in the EEG, z-scores generally fall well within the range of ± 1.0 , providing a sensitive indicator when changes do occur. It is shown that this method has value in assessing individual stability of EEG parameters and for quantifying changes that may occur due to time effects, aging, disorders, medications, or interventions.

Keywords: EEG; qEEG; statistics; database

Citation: Collura, T. F., & Tarrant, J. (2020). Principles and statistics of individualized live and static z-scores. *NeuroRegulation*, 7(1), 45–56. <https://doi.org/10.15540/nr.7.1.45>

***Address correspondence to:** Thomas F. Collura, 195 Willis Street, Suite #3, Bedford, Ohio 44146, USA. Email: tomc1@brainm.com

Copyright: © 2020. Collura and Tarrant. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (CC-BY).

Edited by:

Rex L. Cannon, PhD, SPESA Research Institute, Knoxville, Tennessee, USA

Reviewed by:

Jon A. Frederick, PhD, Lamar University, Beaumont, Texas, USA
Tanju Surmeli, MD, Living Health Center for Research and Education, Sisli, Istanbul, Turkey

Introduction

Z-builder is a method of producing z-scores based upon a reference that is computed from a single sample of EEG. The sample can be any length. The method operates in the same manner that would be used to estimate z-scores from a population of samples, except that it is based on a single sample of EEG from one individual, typically 1 to 3 minutes in length. The resulting norms consist of within-subject means and standard deviations for specified metrics, which are used in place of the typical “normative” samples arising from population-based databases (Collura, 2014).

In conventional normative databases, mean values for designated metrics are computed for each

individual, and then the individual mean values for each subject are combined to produce a population statistic, consisting of the population mean and the population standard deviation. This includes only one source of variance, that of the difference between individual mean values. The z-scores resulting from such an analysis are referred to as “population-static” z-scores.

Alternatively, the within-subject variation can be included in the analysis, providing a wider standard deviation for the z-score calculations. When instantaneous variation is introduced, the result is what is referred to as “population-dynamic” z-scores. The resulting z-scores are typically smaller in absolute value, for reasons explained by Thatcher (2008) and explained further below. To date, EEG

mapping for assessment has been typically done using population-static z-scores, and EEG neurofeedback using live z-scores has used population-dynamic z-scores. In both cases, population means and standard deviations are being used as the references. This raises fundamental concerns when it is recognized that individuals are unique and that using a population-based statistic has the undesirable result of causing every subject to be

compared to a group, raising concerns about the validity of these measurements.

When the distributions of individual and population metrics are compared, this aspect can be clarified. Figure 1 shows the relationship between static and dynamic metrics, and the resulting distributions and computed means and standard deviations.

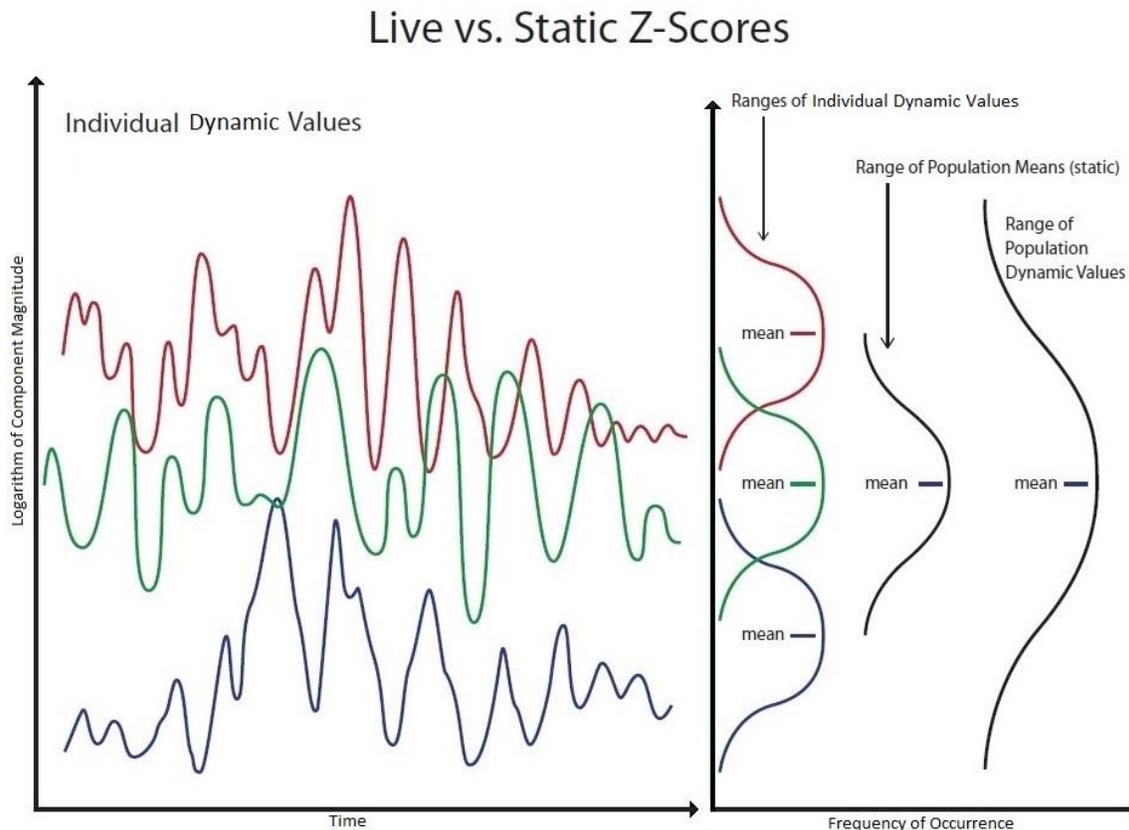


Figure 1. Live vs. Static Z-Scores. Example of dynamic values of a component metric (log magnitude) for three example individuals (left) and resulting statistical distribution ranges (right). Three individuals (red, green, blue) have unique means and standard deviations, which when combined, produce the population statistic. Population ranges can incorporate individual mean values (static), or individual variation as well (dynamic).

In existing live z-score methods, instantaneous z-scores are compared to a reference that is typically derived from a population. If the reference consists of static norms, then the z-scores will reflect how the instantaneous EEG compares to the mean value of a population, using the population variation as the standard deviation. This shows z-scores that, on average, will match the values shown in summary maps, also made from a static database. When a dynamic reference is used, the individual variation

within sessions is added to the standard deviation using an appropriate formula. In this case the instantaneous z-scores reflect the deviation from the full variation within the population, so that z-scores are smaller in absolute value. In both cases, however, the target mean values are the same. It is only the variability that differs. This means that “z = 0” has the same value in both cases.

Because the averaging process is a linear operation, it follows that if you take one mean value from each individual and compute a group mean, the result will be the same as if you were to individually include all the instantaneous data into one huge sample and average it. The resulting average value is the same for both approaches. As long as the averaging occurs before conversion to a z-score, this equivalence will be ensured. As a result, even when it is intended to use dynamic z-scores for training, the target values reflect population means. Thus, each subject is being rewarded for having an EEG more similar to one's peers, which is not a truly individualized approach. This realization has likely held back acceptance of qEEG and z-score neurofeedback for practitioners who object to having clients assessed and/or trained against a group statistic. If target means and standard deviations can be determined that more specifically reflect the individual's characteristics, then assessment and neurofeedback can be individualized to each client.

When using the z-builder approach, the reference means and standard deviations are derived entirely from one individual, and the variability is strictly across time, not across individuals. In other words, the approach described here uses exclusively the EEG data from the individual subject, and no across-subject data are used in the process. While this is conceptually different from using z-scores for population statistics, the mathematical formalism is the same. When applied to these measurements, the intent is not to make a decision related to some population. Rather, it is to determine the typical amount of variation in the repeated measures, to estimate noise and to test the null hypothesis, which in this case is that there is no change in the readings across time. This approach has the further benefit of directly answering the question "how stable are the data?" which is fundamental to the concept of the repeatability of qEEG-based measurements. As pointed out by Messick (1998), the validity of an approach does not depend on the properties of the measurement, but rather on the inferences that are made from the measurement. In this case, the inferences are whether a process is quantifiable and stable, what is the variability, and can we test the hypothesis that "something happened."

The key assumption for a z-score to be valid is that the reference sample and the computation methods ensure that the reference has a Gaussian (normal) distribution. The Gaussianity of single-subject statistics is demonstrated below. This method makes use of the concept of "repeated measurements,"

which has been used and characterized primarily in the field of analytic chemistry (Miller & Miller, 2016). In such applications, repeated measurements are used to reveal the presence of random errors as well as to quantify changes in time. Coming from the analog world, this method produces a number of samples from a theoretically infinite number of measurements we could make, and the set of all measurements is then considered to be the "population." In the present case, we both estimate the variability of the metrics of interest and also provide a statistical means of detecting statistically significant changes in the within-subject design. We shall see that if we take measurements rapidly, the mean and standard deviations of our measurements will converge to correct values, and that the concepts of sample independence and degrees of freedom are not applied to this model.

While it may be a useful assumption that the samples are independent, it is not relevant in a repeated-measurement design. What matters is simply how fast the parameters are changing, how fast we can measure them, and what is the standard error across time. Indeed, when constructing a dynamic norm, the samples that run across the session are not necessarily independent, since they come from repeated measurements from the same system. This time-dependent source of variability is used in z-builder to establish reference norms, and to compute z-scores for both assessment and for live neurofeedback purposes. Moreover, if there is a concern with regard to independence of successive samples when using z-builder, then that same concern would exist for any dynamic z-score reference that includes within-subject variation, including those used for many years. It is true, however, that in choosing the recording length and epoch size, attention must be paid to the choice of reasonable values. For example, the use of many small epochs does not necessarily increase the degrees of freedom, so that taking, for example, estimates 10 times per second and claiming 600 degrees of freedom in a 1-min sample would not be reasonable. We therefore dispense with the concepts of sample independence and degrees of freedom in this design. In order to help ameliorate this concern, we use a consistent sampling rate and computation rate of 256 per second, using the quadrature digital filters, throughout this work.

In order to justify the use of a digital filter (a form of real-time filtering) in lieu of the more common FFT, we compared the results of the digital filter outputs with the FFT amplitude computations, on 1-s

intervals, showing a very strong correlation. The method used is “quadrature filtering” also known as “synchronous demodulation,” which was developed originally for analog computers. In the digital version, we perform computations on every sample, at a rate of 256 per second. Because the filtering method used here allows estimation of signal amplitude and phase on every data sample (Collura, 1990), the resulting metrics are heavily oversampled. This results in an accurate estimation of means and standard deviations, which are known to converge when oversampling is used (Host-Madsen & Handel, 2000). See the Appendix for the basic equations confirming this result.

Figure 2 illustrates the quality of match between the static (FFT) calculation and quadrature digital filter outputs, which provide a dynamic (JTFA) computation. The degree of fit is 97%, once a correction is applied for the difference in the windowing techniques. This result is consistent with that reported by Kerson et al. (2019), which demonstrated a similar quality of fit across two different software and two different hardware platforms. This confirms that we may use digital filter amplitudes in this work, without any systematic disagreement with the results that would result from the conventional FFT method. When applying these calculations, we will restrict our analysis to z-scores, avoiding the issue of *t*-tests and relative degrees of freedom.

Comparison of FFT and JTFA (Quadrature Filter) Results 5 minutes, F3, EC

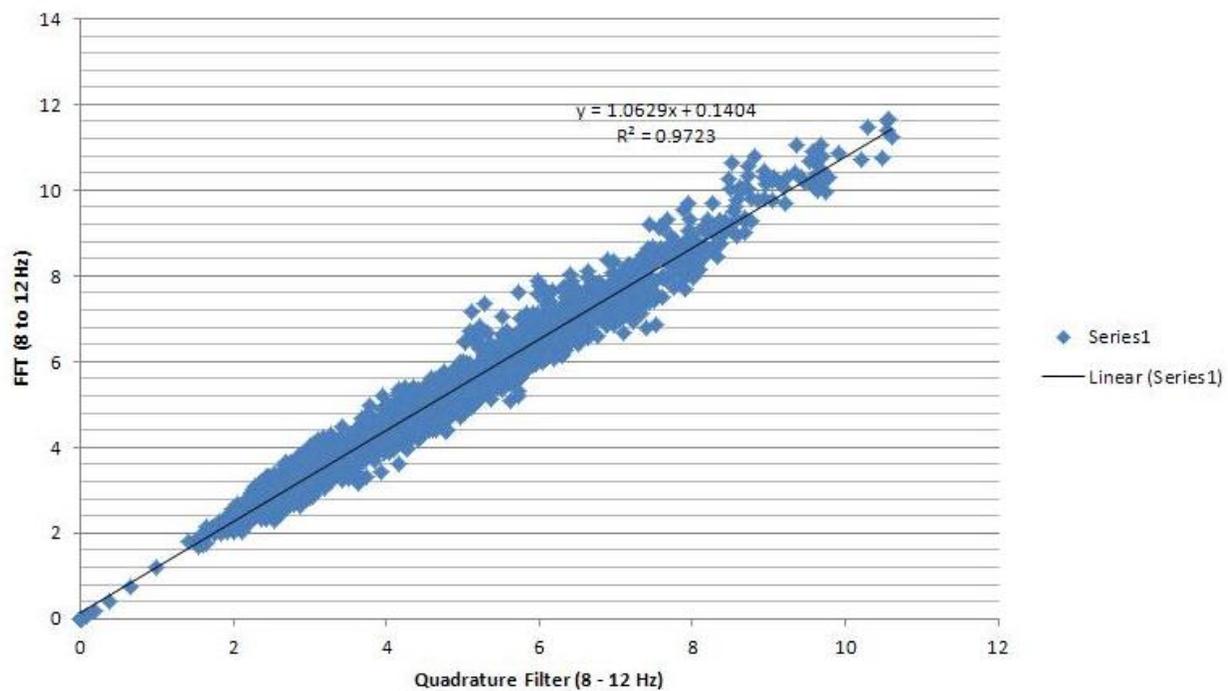


Figure 2. Illustration of the statistical agreement between FFT and JTFA computed amplitude values for the alpha band (8–12 Hz) from 5 minutes of EEG in one individual. When scatterplotted against each other, the results of an FFT analysis (y-axis) and a quadrature filter implementation of the JTFA (horizontal axis) demonstrate a statistical correlation of 97%.

Figure 3 illustrates the Gaussian distribution of the logarithm of the magnitude of one component (theta) for one individual. The goodness of fit is based upon a squared-error comparison of the actual data

with an idealized Gaussian with the same mean and standard deviation.

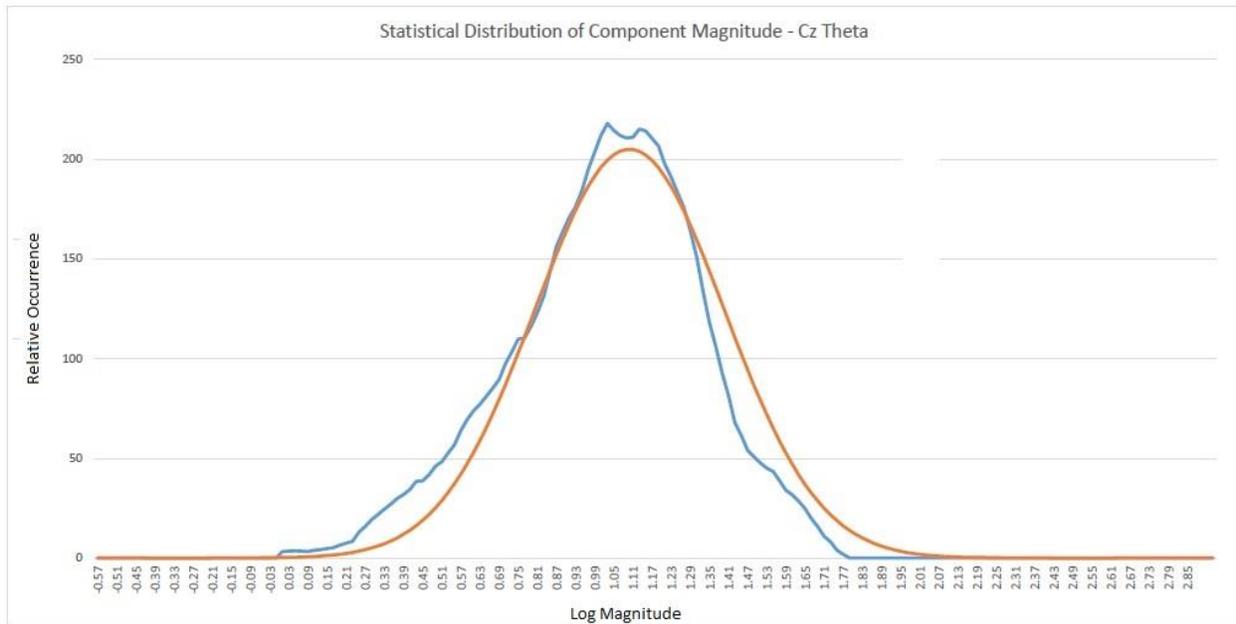


Figure 3. Demonstration of the Gaussianity of a sample of EEG processed for a single subject. Shown: log of Cz theta magnitude for a time-series of 5 minute, yielding $300 \times 256 = 76,800$ datapoints.

In order to determine the goodness of fit for all of the estimated magnitudes, a histogram was created with the 171 component estimates used in this study. Figure 4 illustrates the distribution of the Gaussianity estimates, confirming that Gaussianity is generally above 90%, and is centered at 94% Gaussianity. Figure 4 shows a histogram of the results of this

analysis. 93% of the components (159 out of 171) have a goodness of fit of 90% or above. Two distributions are evident, one centered at 0.94 Hz and a second centered at 0.89 Hz. The lower distribution in this example was found to contain reflect activity, particularly from frontotemporal leads.

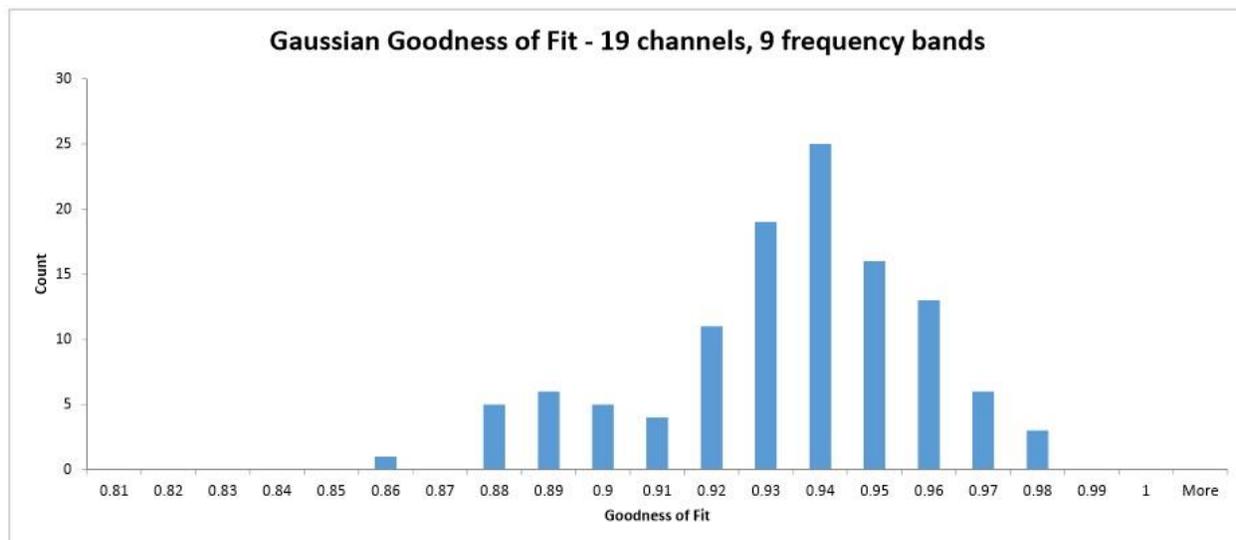


Figure 4. Histogram of the Gaussianity measurements from one individual, 19 channels and 9 frequency bands. This produces 171 component estimates. The Gaussianity of each component was computed for the same data.

Adapting Thatcher's (2008) notation, we denote the z-score based upon a population-static as Z_{PS} , corresponding to Thatcher's FFT, and his z-score based upon a population-dynamic and using the JTFA procedure as Z_{PD} . In order to compute either of these, the current sample is subtracted from the population mean, whether it is a static or an instantaneous calculation.

In both cases, the target value should be the same, after allowing for systematic differences such as windowing or other factors. This agreement in the raw values (and hence the mean) is shown in the example in Figure 2.

We use the following notation:

Z_{PS} for a population-static z-score

Z_{PD} for a population-dynamic z-score

Z_{IS} for an individual-static z-score

Z_{ID} for an individual-dynamic z-score

The sources of variation in a static z-score is solely SD_s which is the variation between subjects, when each subject contributes a mean to the statistic. The sources of variation in a dynamic z-score are SD_t due to the time-dependent activity, and SD_s . The two sources are combined in the average, so that (per Thatcher, 2008) $SD_D = (SD_t + SD_s) / 2$. This is the method that is used when applying what we call "population-dynamic" z-scores, which are based on population data. Combining population variation and time-based variation in this way elevates the issue of combining two different types of variation, arising from different mechanisms, in one measurement. It also introduces the question of whether each type of variation should be weighted equally as an average, or should they be weighted in a different manner.

The approach reported here avoids this concern, because we produce z-scores which are based entirely on the individual's instantaneous variation and no population statistics are introduced. We refer to these as "individual" dynamic z-scores. Oddly, this approach may be considered database-free, as it can be applied to any individual without requiring that a "normative" or "standardized" database be introduced. In the case of the static z-score, the standard deviation is SD_s . In the case of an individualized instantaneous z-score, the standard

deviation is simply that introduced by the subjects' EEG, variation across time, designated as SD_t .

When working with z-scores computed based on a normative sample and using a single individual as the measurements, there is a natural expectation that z-scores will follow the predicted distribution. That is, z-scores between ± 1 will occur approximately 65% of the time, and scores between ± 2 will occur approximately 95% of the time. This allows hypothesis testing, using these probabilities. Type 1 and type 2 error can be estimated using these distributions. The null hypothesis is that the person is entirely "average" and that no unusual z-scores will appear. For usual purposes, a range of ± 2.0 is used, and for medical determinations, a range of ± 2.5 or even ± 3.0 would be more common.

As stated previously, when working within an individual, the null hypothesis is not "this is from a normal individual," but rather that "nothing happened." That is, there is no change from the sample to the current measurements. Based on this consideration and the statistical principles described below, individualized z-scores occupy a tighter range than those from a normative analysis. It will be seen that a z-score outside ± 1 will be important in this case, and z-scores much outside of this range will be significant.

Static and Dynamic Z-Scores

We now look at the expected behavior of z-scores when using static or dynamic references, as well as population versus individual references. We can state in general that the expected value (mean target) for a population-dynamic z-scores is the same as that for a population-static z-scores.

$$E(X_{PD}) = E(X_{PS})$$

While this is desirable from some standpoints, such as uniformity when applying either type of analysis, the drawback is that the targets used for live z-score training for all individuals remain based on a population. That is, even when using current live z-score methods, the individual is still being compared to others, and his training targets are based on other people. Because of the additional variation included in the XJTFA calculation, we can state that, necessarily, z-scores from a population-dynamic process will be smaller in general than those from a population-static method:

$$|Z_{PD}| < |Z_{PS}|$$

That is, dynamically computed z-scores that incorporate both the across-subjects and the within-subjects variation will be smaller than conventional static z-scores. Moreover, we note that, generally, the expected value (mean target) of an individual-dynamic z-score is not the same as that of population-dynamic z-scores.

$$E(X_{ID}) \neq E(X_{PD})$$

Also, the expected value for an individual-dynamic z-score is not the same as that of population-static z-scores.

$$E(X_{ID}) \neq E(X_{PS})$$

In other words, when using an individual EEG as a reference, there is no reason to expect that the mean values will be the same as those from a population sample. Moreover, when an individualized approach is taken, and the included samples are from one individual only, then the sole source of variation is the time variation. Furthermore, the mean value for that individual will generally not be equal to the population average. Indeed, this may never happen. In general, the mean values for each individual will themselves follow a Gaussian distribution, which is in fact the mean data that is included in the static statistics.

Because this approach uses a different mean value (the subject's own mean value) and a different source of variation (equal to the *SD* for that individual), resulting z-scores will have a distribution that no longer reflects population statistics; it is solely a representation of that individual's mean values and variation across time.

It is reasonable to expect that the same transformations that produce Gaussianity in static and in dynamic z-scores should suffice to produce Gaussian distribution of instantaneous individual scores. This can be verified experimentally by applying a suitable test of Gaussian fit. Figure 3 demonstrates the Gaussianity of a 1-min sample of EEG transformed using the customary logarithmic equation used for static or dynamic statistics.

In order to estimate the significance of a z-score computed using this method, we can use basic statistical principles to determine how likely a given z-score would be, based on the expected results of the computations. Specifically, the references are based upon a specified sample of EEG, which includes the short-term variation, as the source of the standard deviation.

In order to compute z-scores, it is sufficient to demonstrate Gaussianity of the comparison data, as long as the current value is transformed in the same manner as the original samples, to follow that Gaussian distribution. The stability and usefulness of a short-term statistic is a separate issue, and must be addressed experimentally, in order to determine the realistic expected variation between samples from time to time. Indeed, repeatability studies of qEEG in general have confirmed that a single 1- to 2-min sample from an individual at rest indeed provides a useful set of estimates. The repeatability of that sample in a second recording minutes, hours, days, or even months later is a tacit assumption in the use of clinical qEEG, and one that has been evaluated. We therefore conclude that a single sample from an individual does provide a useful basis for computing expected means, as well as the expected standard deviation, which can be used for computation of z-scores at a future time.

In order to assign a probability, that is, a *p*-value to a given outcome, we examine the conditions used to produce the reference estimate as well as the details of how a particular z-score is being computed. In a case where we use, say, 1 minute to compute the mean and standard deviation of key variables, then use for example 10 seconds of live EEG to compute a semistatic z-score, we can estimate the likelihood of deviant z-scores appearing. As a simplifying assumption, we assume the subject is in stable and repeatable state; for example, eyes closed, not drowsy, etc. Additional factors such as change in conscious state or other EEG-related changes will of necessity produce more deviant z-scores. Therefore, this estimate provides a lower-bound to expected z-scores using this method.

We are now looking at the probability distributions of the z-scores themselves. It might seem intuitive that a z-score of 2, for example, indicates a deviation equal to 95% of the population, this will not always be the case. As shown by Thatcher (2008), for example, we know that dynamic z-scores will be smaller in value than static z-scores, when the same population is used for both estimates.

Whether the variation due the population is greater or less than the variation due to the intersubject variation is subject to measurement. It is clear that neither of them is insignificant; that is, neither the intersubject variation nor the intrasubject variation may be taken to be small. To calculate the significance of a particular individualized z-score, we make adjustments to the standard z-score probability

ranges, based upon the sampling details. As a first approximation, we can use the statistics of the t -test to provide an estimate. In a t -test, we compare two populations with different means and standard deviations, using the t value which is the difference in means divided by their joint standard deviation. This is similar to a z -score, which is the difference between two means, divided by the standard deviation of the reference population. When using individualized z -scores, if the standard deviation of a variable is assumed to remain constant as that variable varies in value, the z -score provides identical information as a t -score. When applied in this way, we are computing t -tests for “dependent means.” This is valid as long as (1) the data are normally distributed, (2) the scale of measurement is an interval or ratio, and (3) the measurements are matched in some way. In these circumstances, a t -test on dependent means can test a null hypothesis that there is no difference between the two means (Social Science Statistics, 2019).

Significance of t -scores is based on a computation that takes into account the number of samples (degrees of freedom) in the two samples. Similarly, when computing z -scores, the reference as well as the current value carry with them their respective degrees of freedom. When working with a time-series, successive samples are not independent in the same manner as samples from a population. As stated by Miller and Miller (2018), when taking successive samples from a process, the assumption of statistical independence is not made, and degrees of freedom are not comparable. Rather, successive samples serve to estimate the noise in the system, as well as the repeatability of measurements over time. Despite being based on a single individual, such a time-series is nonetheless a random variable and can be studied as such. In our analysis, we observe that the bandwidth of the filters is generally 4 Hz. Based on the inverse relationship of bandwidth and transient response (Collura, 2014), this corresponds to a transient rise time-constant on the order of 10 ms. Therefore, taking 256 calculations per second from each filter output should be more than adequate to capture the time behavior of the variables. Appendix I further shows that oversampling in this way does not compromise the estimate of the mean and standard deviation, as they converge when oversampled.

In general, if we compute an average and standard deviation from n samples of a random variable, the expected value of the mean is precisely the mean value, while the expected standard error is divided by n , so that the standard deviation is divided by the square root of n . For example, we take n samples of

EEG and use these to compute the expected mean (average) and the standard deviation of the EEG to produce the current estimates. When we take a subsequent sample to estimate the new mean, there will be averaging over the epoch chosen. The number of samples is not important, as shown in the Appendix, but the duration of the retest calculations does matter. The relationship between filter bandwidth and rise time-constant is given by

$$t = 1 / (2 * \pi * BW) = 0.35 / BW$$

So that, with a 4-Hz filter bandwidth, there is a time-constant of about 90 milliseconds.

The variation in the values will therefore occur with a maximum frequency on the order of 11 Hz, due to the filter bandwidth used. We propose that the ratio of the epoch chosen to the time-constant provides an estimate of how much damping will occur when averaging the filter outputs such as power, coherence, etc. That is,

$$\text{Effective reduction in variability} = (\text{approximated by}) \\ \text{epoch length} / \text{time-constant}$$

With a 10-s epoch and a 90-ms time-constant, the reduction is on the order of $10 / 0.09 = 110$. This is the reduction in the variance, so that this is the square of the reduction in standard deviation. Thus, using 10 s of subsequent EEG to compute a current value, the expected standard deviation becomes divided by the square root of 110, which is a scale factor of roughly 0.1. Thus, the value averaged over 10 s is expected to vary approximately one-tenth as much as an unaveraged estimate. Thus, with an optimally stable EEG, we would see 95% of z -scores within the range of approximately ± 0.3 . We conclude that, when taking a sufficient sample of reference EEG to obtain a convergence on means and standard deviations and then using a subsequent sample of EEG to calculate an updated mean value, the resulting z -score will have a strong tendency to be close to zero. In other words, our retest z -scores will typically be very low and close to zero. This is confirmed in the example data shown below.

Example Data

The data attached are z -builder comparisons base versus 10-min delay and base versus 30-min delay in the eyes-closed (EC) condition for three subjects. The results using 10 min of baseline EC and 10 s of test EEG, are shown in Figure 5.

Figure 5 shows three examples of a single-subject estimate of the variability in z-scores. Each histogram reflects the z-scores from 11 frequency bands and 93 sLORETA regions of interest. Each histogram therefore contains 1,023 z-scores. Reference sample of 10 min of EEG, test sample of 10 s of EEG.

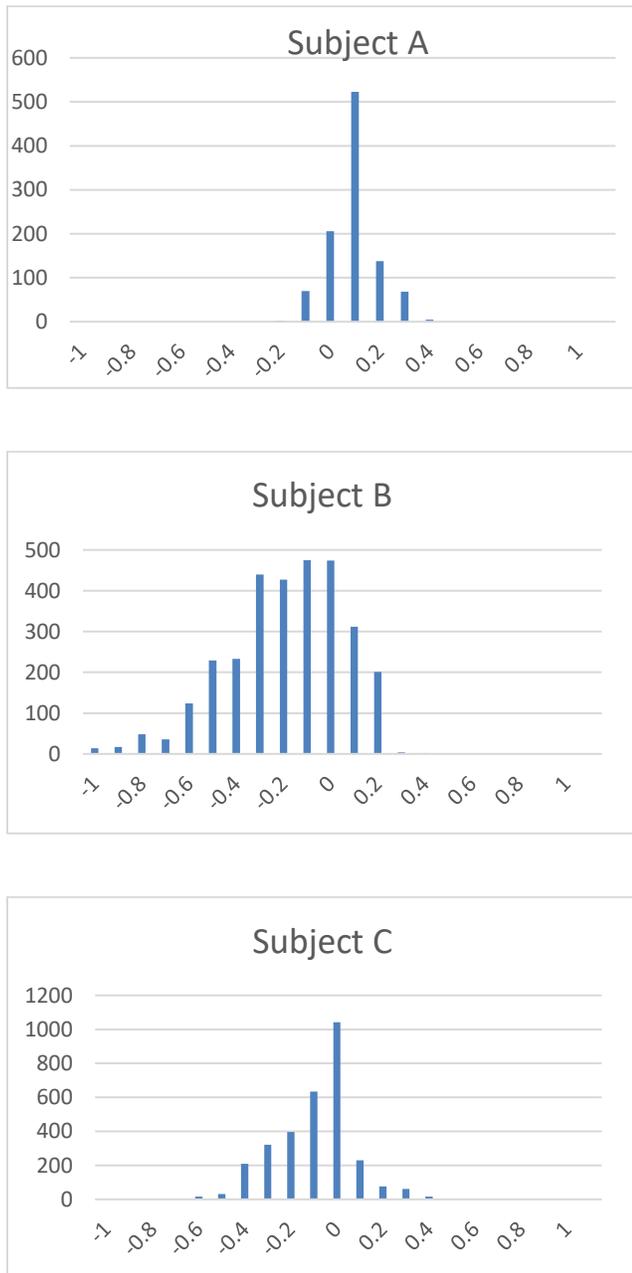


Figure 5. Three examples of a single-subject test–retest repeat measurements using 11 frequency bands and 93 sLORETA regions of interest.

A total of 12 such exercises was performed with the three subjects, with differences in the length of the reference sample (10 or 30 min) and the eyes condition (closed or open). The following table summarizes the results with these three subjects.

Table 1
Summary of z-score results for three subjects with 10-min and 30-min reference intervals.

	Min	Mode	Max	Width
Subject				
10 EC	-0.3	0.1	0.3	0.6
10 EO	-0.5	0.1	1.4	1.9
30 EC	-0.7	0	0.2	0.9
30 EO	-1.0	0	0.3	1.3
10 EC	-1.1	0	0.2	1.3
10 EO	-0.5	0.2	1.2	1.7
30 EC	-1.1	-0.3	0.1	1.2
30 EO	-0.5	0	0.7	1.2
10 EC	-0.7	0	0.4	1.1
10 EO	-0.6	0	-0.6	0
30 EC	-0.5	0	1.0	1.5
30 EO	-0.7	-0.2	0.3	1.0
Average	-0.6833	-0.0083	0.45833	1.14167

It is seen that, generally, such z-scores were within the range of ± 0.6 standard deviations, reflecting the stability of the EEG during the procedure. This demonstrates that it is possible to gather some minutes of EEG, then retest at some future point, and achieve a tight distribution of repeat measurements upon retest. Subjects A and C are quite stable and fall within the predicted range of ± 0.3 . In contrast, subject B is skewed, and appears to exhibit a systematic drop in z-scores with a shift to the left of 0.3 standard deviations, from test to retest. The appearance of z-scores less than -0.5 in this sample represents some type of (significant) change.

Estimating Changes

One benefit of using the z-builder approach is that it is free of the assumptions and implications of using a reference from a population that is purportedly “typical” or “normal.” Rather, this method recognizes the fact that everyone is different and has a unique set of EEG characteristics. Thus, this method can accurately determine the effects of different influences, without having to assume that the subject fits somehow into a wider population. When used for

neurofeedback (Collura, Thatcher, Smith, Lambos, & Stark, 2009), this approach would allow training to reflect any specified individual or state as the reference. This opens the possibility of individualized training that can aim to restore previous levels of brain activity or to train toward desired goals, such as reduction of specific characteristics.

A practical application of this method was reported by Siever and Collura (2017) who were able to produce sLORETA images of static z-scores for brain responses to repetitive photic, auditory, and magnetic stimulation. Examples are shown in Figures 6A–6C. Citing their work:

A reference data set was first constructed by taking 1 min from a 2-min at-rest baseline, and processed using BrainAvatar Z-Builder signal processing (Collura, 2012, 2013, 2014a, 2014b) to produce amplitude means and standard deviations for all frequency bands, for all scalp locations, sLORETA voxels, and sLORETA Regions of Interest (ROI). ROIs were computed for 97 different homologous regions including the Brodmann areas, the named lobes and regions, and for the hubs described by Hagmann. Once this data set was computed, it was possible to compute metrics for any other selected samples and, by comparison, convert all measurements into z-scores.

A 10-s sample was taken from the stimulus interval for each modality and analyzed using the subject's individual z-score database, producing z-score results. These z-scores show which qEEG components and locations have changed. These z-scores are not based on a normative reference database but are instead based on the subject's own initial EEG. Thus, z-scores reflect change from the initial state, and do not reflect "normality" or "abnormality" in any way. Because the use of z-scores in this manner involves multiple comparisons of many ROIs (97), Bonferroni correction was applied to the results shown here. (Siever and Collura, 2017, p. 82)

The resulting sLORETA images clearly show the sensory areas affected by the visual and auditory stimulation, as well as the fact that the pulsed EMF stimulation produced a frontal response, despite having been placed over the motor strip.

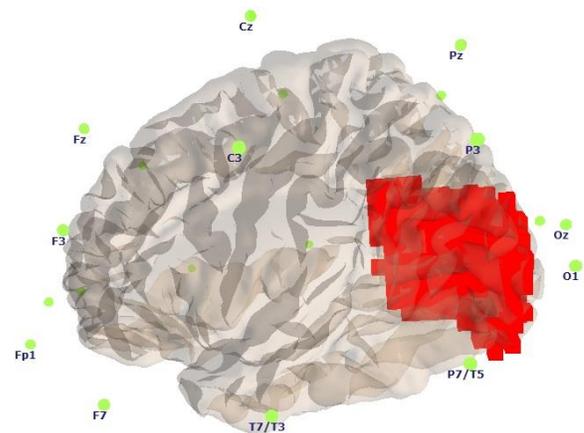


Figure 6A. Reduction in delta activity during binocular photic stimulation at 3.5 Hz. Left-side view. Affected areas include Brodmann 17 and 18 and the cuneus. These comprise the primary and secondary visual sensory areas.

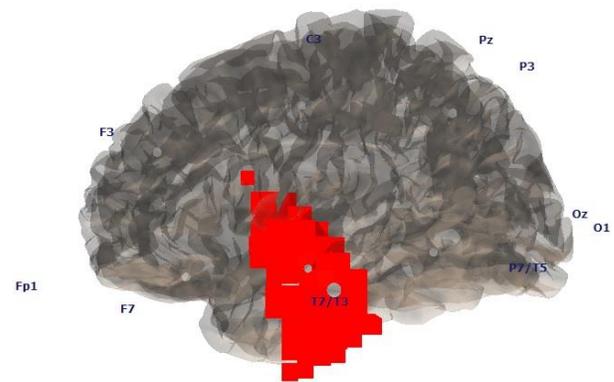


Figure 6B. Reduction in delta activity during auditory stimulation (clicks) at 3.5 Hz. Left-side view. Affected areas include Brodmann 13, 33, 24, 20, the Insula, and Sub Lobar areas. These areas are involved in auditory sensation, perception, awareness, and attentional control.

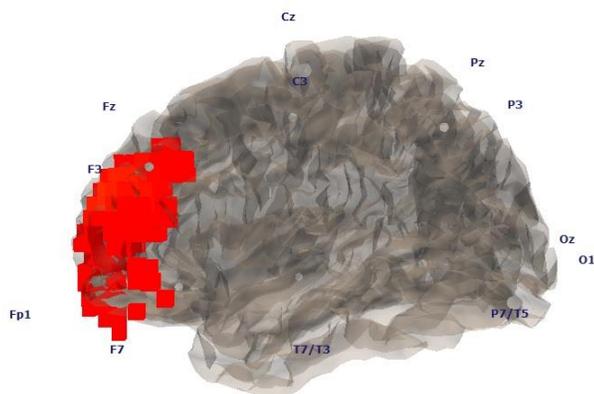


Figure 6C. Reduction in delta activity during pulsed electromagnetic stimulation at 3.5 Hz. Left-side view. Affected areas include Brodmann 45, 46, 10, and 44. These comprise the frontal lobe areas associated with executive function, attention, emotion, and decision-making.

Conclusion

These preliminary results verify that a single-subject approach to creating qEEG references and producing z-scores based upon time variation is feasible and statistically sound. Both surface and sLORETA z-scores can be computed in this manner. Individualized z-scores use a different theoretical foundation than population z-scores and are based upon time-effects, not population statistics. A continuous, repeated-measures approach that is based upon the analog world is applicable, in contrast to a sampling approach based upon population statistics. The population-based concepts of sample independence and degrees of freedom are not applicable in this situation; hence, they do not limit this approach. This method has shown the ability to quantify the stability of an individual's EEG and also to detect, quantify, and map changes that could arise due to the effects of time, state changes, disorders, or external influences.

Author Disclosure

Thomas Collura is the founder and an owner of BrainMaster Technologies, Inc. The authors have no grant support, financial interest, or conflicts of interest to report.

References

- Collura, T. F. (2014). *Technical Foundations of Neurofeedback*. New York, NY: Routledge.
- Collura, T. F. (2014, Spring). Specifying and developing references for live z-score neurofeedback. *NeuroConnections*, 26–39. Retrieved from https://docs.wixstatic.com/ugd/cba323_b824c922625941808b2d633bc63f3df7.pdf
- Collura, T. F. (1990). Real-time filtering for the estimation of steady-state visual evoked potentials. *IEEE Transactions on Biomedical Engineering*, 37(6), 650–652. Accessed from <https://www.ncbi.nlm.nih.gov/pubmed/2354847>
- Collura, T. F., Thatcher, R. W., Smith, M. L., Lambos, W. A., & C. R. Stark (2009). EEG Biofeedback training using live Z-scores and a normative database. In J. R. Evans, T. H. Budzynski, H. K. Budzynski., & A. Arbanal (Eds.), *Introduction to quantitative EEG and neurofeedback: Advanced theory and applications, Second Edition* (pp. 103–142). New York, NY: Elsevier.
- Host-Mandel, A., & Handel, P. (2000). Effects of sampling and quantization on single-tone frequency estimation. *IEEE Transactions on Signal Processing*, 48(3), 650–662. Accessed from <http://www-ee.eng.hawaii.edu/~madsen/papers/48sp03-hostmadsen.pdf>
- Kerson, C., deBeus, R., Lightstone, H., Arnold, L. E., Barterian, J., Pan, X., & Monastra, V. (2020). EEG theta/beta ratio calculations differ between various EEG neurofeedback and assessment software packages: Clinical interpretation. *Clinical EEG and Neuroscience*, 51(2), 114–120. <https://doi.org/10.1177/1550059419888320>
- Messick, S. (1998). Test validity: A matter of consequence. *Social Indicators Research*, 45(1–3), 35–44.
- Miller, J. N., & Miller, J. C. (2018). Statistics of repeated measures. In J. N. Miller & J. C. Miller (Eds.), *Statistics and chemometrics for analytical chemistry* (6th ed., pp. 17–26). Harlow, England: Pearson. Retrieved from <https://manualzz.com/doc/27792868/statistics-and-chemometrics-for-analytical-chemistry>
- Siever, D., & Collura, T. (2017). Chapter3 – Audio-visual entrainment: Physiological mechanisms and clinical outcomes. In J. R. Evans & R. P. Turner (Eds.), *Rhythmic stimulation in neuromodulation* (pp.51–95). New York, NY: Elsevier.
- Social Science Statistics (2019). T Test Calculator for 2 Dependent Means. Retrieved from <https://www.socscistatistics.com/tests/ttestdependent/default.aspx>
- Thatcher, R. W. (2008, April). Z-Score EEG biofeedback conceptual foundations. *NeuroConnections*, 9–11. Retrieved from https://docs.wixstatic.com/ugd/cba323_426c6449511f48968c49c4ee94fa0c7e.pdf

Received: September 30, 2019

Accepted: January 3, 2020

Published: March 25, 2020

Appendix

The following demonstrates that if you sample a signal to acquire real data such as amplitudes, temperatures, etc., as you sample the signal more often, the estimated mean and variability can only improve. It is shown that oversampling will not change the expected mean value, and it will only improve the estimate of the variability (variance). Given the following time series:

$$X = x_1, x_2, x_3, \dots x_n$$

The mean is defined as:

$$\mu = \frac{1}{n} \sum_{i=0}^n x_i$$

The variance in the data is defined as:

$$\sigma^2 = \frac{1}{n-1} \sum_{i=0}^n (x_i - \mu)^2$$

We can represent an oversampled version of this signal by replacing each x with a copy, thus doubling the sampling rate, under the assumption that there is negligible difference between samples. That is ensured by the fact that the signal is filtered and that we are sampling at a high enough rate. The oversampled data can be written as:

$$X_o = x_1, x_1, x_2, x_2, x_3, x_3 \dots x_n, x_n$$

The mean value of the oversampled signal does not change, as should be intuitively evident, as we are only adding more of the same data points. This can be shown as:

$$\mu_o = \frac{1}{2n} \sum_{i=0}^n 2x_i$$

So that when the 2's cancel, we find that $\mu_o = \mu$.

As for the variance estimate, if we designate the variance of the oversampled signal as σ_o^2 we have the following:

$$\sigma_o^2 = \frac{1}{2n-1} \sum_{i=0}^n 2(x_i - \mu_o)^2$$

Comparing this with the equation for the original variance, we find:

$$\sigma_o^2 = \frac{2n-2}{2n-1} \sigma^2$$

Which for large n is a ratio very close to 1.000. For example, if we incorporate 256 values into our estimate over 1 second, the values have converged within a factor of $(512-2) / (512-1)$, which is a ratio of 1.002.

Therefore, as the number of samples increases, the estimated variance converges strongly to a final value, which can be considered the "true" variance which for very high values of n is essentially independent of the sampling rate. This is superior to methods that use FFTs and sliding bins, because those approaches require compromises associated with epoch selection and sliding factors. The approach of oversampling the data from continuous quadrature filter data circumvents these limitations entirely and produces results equivalent to an analog system.