

## NEURAL NETWORK AND MULTIVARIATE ANALYSES: PATTERN RECOGNITION IN ACADEMIC AND SOCIAL RESEARCH

Chris Ninness,<sup>1</sup> Marilyn Rumph  
Logan Clary, David Lawson  
John-Thomas Lacy, Sarah Halle  
Ranleigh McAdams, Sonya Parker  
Diane Forney  
*Stephen F. Austin State University*

**ABSTRACT:** Neural networks are the modern tools that focus most heavily on the logical structure of measurement/assessment, as well as the actual results we attempt to identify by way of scientific inquiry. Employing the *Self-Organizing Map* (SOM) neural network, we reexamined a well-recognized and commonly employed dataset from a popular applied multivariate statistics text by Stevens (2009). Using this textbook dataset as an exemplar, we provide a preliminary guide to neural networking approaches to the analysis of behavioral outcomes. When employing conventional multivariate procedures only, the sample dataset demonstrated significant familywise error rates; however, these outcomes did not provide sufficient information for identifying the curvilinear patterns that existed within these records. When converted to natural logs and reanalyzed by the SOM, the exemplar dataset showed the actual best fit performance patterns exhibited by all members of the experimental and control groups. The SOM and related neural network algorithms appear to have unique potential in the recognition of nonlinear but unified data patterns frequently exhibited within academic and social outcomes. In particular, the SOM allows the researcher to conduct a "finer grain" analysis identifying critically important similarities and differences that can inform treatment well beyond the probability values derived from conventional statistical techniques.

**KEYWORDS:** neural network, self-organizing map, pattern recognition, probability, multivariate analysis

Psychology has been described as an ailing scientific discipline. To paraphrase Forness (1970), the field's academic affliction becomes particularly obvious with an examination of published research within the annals of our most prestigious professional literature. Here, psychologists often present with such symptoms as violations of the assumptions of normality and linearity, inflation of findings, and accumulations of jargon deposits within their conclusions. Discussing an analogous set of conditions Forness (1970) forwards the notion, "The syndrome might best be described as a paralysis of the analysis." (cited in Ninness & Glenn, 1988, p. 6).

---

<sup>1</sup> We would like to express our gratitude to Dr. David Jeffrey for his support and guidance during the original development of our research in the area of artificial intelligence and neural networking architectures. Correspondence concerning this article should be addressed to Chris Ninness, School Psychology PhD Program, Stephen F. Austin State University, P.O. Box 13019-SFA Station, Nacogdoches, TX 75962.  
E-mail: cninness@sfasu.edu

During the decades following Forness's comments, it has become increasingly obvious to many researchers that most traditional data analysis methodologies are insufficient to address the magnitude and diversity of data inundating their respective academic disciplines (e.g., Kline, 2004). In particular, the applied research within the behavioral and social sciences literature abounds with evidence obtained by way of traditional linear multivariate procedures as applied to outcomes that are inherently nonlinear. Although often employed when no logical alternatives seem to be available, univariate statistical tests simply are not capable of analyzing nonlinear data (see Table 1 regarding linearity issues). Even the more complex statistical techniques are severely limited by the core assumptions pertaining to traditional multivariate analysis (e.g., homogeneity of variance, normality, linearity, absence of multicollinearity, and random and independent sampling). Indeed, each univariate and multivariate statistical test comes with its own set of unique assumptions that must be satisfied in order to be employed with any approximation of accuracy.

Paradoxically, during the past four decades, new and substantially more robust forms of nonlinear neural network classification procedures have been developed (Heaton, 2008), but only in very recent years have these methodologies become widely available to researchers in the social and behavioral sciences (Haykin, 2009). Neural network methodologies appear to have particular value with respect to identifying and understanding an ever-growing number of important social/economic problems (e.g., Kaski & Kohonen, 1996) and an exponentially growing number of environmental issues (e.g., Allamehzadeh & Mokhtari, 2003; Cavazos, 2000). As per Ward and Houmanfar (2011), early behavior analysis focused on fundamental research and sought to make the design and analysis of behavior increasingly sophisticated. Modern behavior analysis seeks to take complex data and simplify it through research that occurs in natural and simulated environments, employing the most robust and scientific tools at its disposal. Neural networks stand as some of the most modern scientific tools available for conducting a comprehensive data analysis.

Artificial neural networks (ANNs) are nonlinear data driven algorithms that are particularly well suited to identifying inconspicuous patterns within an array of complex and highly diversified datasets. Among a vast range of related problem solving strategies, many neural networking algorithms are employed in pattern recognition and the development of graphical illustrations of nonlinear outcomes (Reusch, Alley, & Hewitson, 2005). Thus, neural network analyses may prove increasingly valuable in the understanding and development of nonlinear outcomes derived from the analysis of human behavior. Further, it is our position, that neural network analyses can significantly enhance and extend traditional statistical analysis (e.g., group comparisons using multivariate analysis) by providing a finer-grain examination of the similarities and differences of individual case scores within and between groups, beyond group mean comparisons.

Structurally and functionally, artificial neural networks are "neural," primarily in the sense that the network algorithms were initially inspired and developed around the "biological structure" of human neural physiology (e.g., Hebb, 1949); however, research in the development of artificial neural networking does not advance parallel research in human neural physiology, nor is human neural science informed by developments in artificial neural architectures. As per Burgos (2007), a direct relation between the neural network model, "...and the biological structure of the brain is not integral to the construction of the model. That is to say, the models are not informed by neuroscientific knowledge about the biological structure and functioning of

Table 1. *Special Terminology and Brief Definitions*

<b><u>Term</u></b>	<b><u>Definition</u></b>
Linear Data	Refers to measurement outcomes that may be represented on the coordinate axis such that they approximate a straight line passing through a swarm of data. Usually, such distributions do not align perfectly with a straight line; however, the scores are uniformly spread around a best fit line passing through the data.
Nonlinear Data	Refers to measurement outcomes that may be represented in accordance with any dispersion of measurements that form an irregular pattern. The nonlinear data points may be erratic and asymmetrical (such as those seen in stock market fluctuations). Other nonlinear patterns may be repetitive (such as those seen in the cumulative records as a function of fixed-interval schedules of reinforcement).
Neural Networks in the form of Self-Organizing Maps (SOM)	Software programs based on a sequence of mathematical functions are referred to as nodes. These nodes receive input values in the form of vectors (series of independent variables). Using the Euclidean distance formula, each node is scanned in an attempt to identify nodes that are most similar to the input values. The node that best approximates the input vector is identified and referred to as the “best matching unit” (Kohonen, 2001). The radius of the best matching unit is calculated, and the weight for each node is altered such that they become more closely aligned with the input values. This step is repeated until the SOM finds the best pattern approximating each set of input values/vectors. Over a series of iterations (cycles), common patterns emerge. These patterns are classified (categorized and given a specific numerical designation). The reader is referred to Heaton (2008) for a detailed discussion.
Euclidean Distance	(Alternatively Pythagorean distance) is the distance between two data points obtained as a function of calculating the root of squared differences between specific locations on the coordinate axis. The SOM algorithm locates data points sharing the least distance between them by implementation of the Euclidean distance formula within the algorithm.
Class or Classification of Data	Refers to a set or category of things having some property or attribute in common and differentiated from others. Neural network outcomes that are categorized or cluster together based on some common characteristic or attribute fall within a neural network class are named accordingly. Note, however, that neural network may provide numerical rather than nominal classifications.
Articulation	Refers to the musical performance technique that affects the transition or continuity on a single note, or between multiple notes or sounds.

brains” (p. 119). However, as forwarded by Burgos, there are several studies in which post hoc interpretations of neurological structures have been inferred subsequent to the development of a new artificial neural network design/algorithm (e.g., Myers, Gluck, Weinberger, & Mattay, 2005).

From a direct problem solving perspective, neural networks are congenial with (or complemented by) several statistical algorithms. In the areas of behavioral and social sciences, outcomes conventionally evaluated by factor analysis, principal components analysis, and a wide range of related multivariate procedures may now be evaluated (or at least supplemented) by way of neural network strategies (see Ninness, Lauter, Coffee, Clary, Kelly, Rumph, Rumph, Kyle, & Ninness, 2012, for a discussion). While a comprehensive description of neural networking methodology is well beyond the scope of this paper, we will provide a condensed overview of two commonly employed techniques with a special emphasis on one increasingly popular and exceptionally robust neural network model.

*Supervised neural networks* must undergo a series of training sessions with training data in conjunction with expected output response patterns, before they are capable of providing solutions to problems. When input values are at variance with anticipated outcomes, the weights are sequentially modified (i.e., the training sequence is adjusted and reinitiated) in order to gradually diminish the existing inaccuracies (see Heaton, 2008, for an advanced discussion).

The primary complication associated with employing supervised networks is that the researcher must have preexisting and detailed knowledge of accurate outcomes in order to identify correct response patterns generated by the systems. Training for ANNs may occur over an extremely large number of trials and must continue until the network establishes the best output patterns for a given array of inputs. Notwithstanding, ANNs have established their value in the mathematical formulation of many behavior analytic investigations (e.g., Burgos, 2001), and there are many applications for which these types of networks are particularly serviceable (e.g., Donahoe & Burgos, 1999). These networks are exceptionally well suited to the development of mathematical learning models since they can be trained to simulate specific response patterns and learning outcomes. For example, Donahoe and Burgos (2000) described a series of feedforward neural network manipulations utilized to assess simulated response strength to feedback in terms of decreasing and increasing operant behavior. Kemp and Eckerman (2001) advanced a variation on this theme by exposing a feedforward network to various schedules of reinforcement, and they obtained cumulative records showing response patterns reminiscent of nonverbal organisms on the same schedules. More recently, Burgos (2007) developed a series of four simulations using feedforward architectures to demonstrate autoshaping and automaintenance.

An entirely different type of neural networking system is frequently referred to as *unsupervised* (sometimes described as competitive learning networks), the most prominent of which is the Self-Organizing Map (SOM) Kohonen (1982; 1984; 2001). This type of architecture is capable of repeatedly exposing itself to the same data such that it acquires the “ability” to identify subtle data configurations within extremely diversified and nonlinear input distributions (see Table 1 regarding *nonlinear* data). A few examples of datasets in which the SOM has proved uniquely capable of recognizing nonlinear patterns have included financial analyses (Deboeck & Kohonen, 1998), complex socio-economic datasets (Kaski & Kohonen, 1996), and the forecasting of earthquake aftershocks (Allamehzadeh & Mokhtari, 2003). SOM pattern recognition systems have been used to construct technologies such as car navigation (Pomerleau,

1991), language processing with modular neural networks, and distributed lexicon language processing (Miikkulainen & Dyer, 1991), and mapping neural activity within the visual cortex (e.g., Miikkulainen, Bednar, Choe, & Sirosh, 2005). The SOM also has played an important role in the recognition and grouping of human blood plasma profiles (Kaartinen, Hiltunen, Kovanen, & Ala-Korpela, 1998), the investigation of irregular and potentially dangerous climate changes (Cavazos, 2000), the identification of insulin resistance disorder and related conditions (Valkonen, Kolehmainen, Lakka, & Salonen, 2002), as well as the recognition and classification of gene expression configurations (Tamayo, Slonim, Mesirov, Zhu, Kitareewan, Dmitrovsky, Lander, & Golub, 1999).

Presently, the Kohonen SOM algorithm is renowned as one of the most elegant competitive learning networks ever designed (Ultsch, 2007). Unlike supervised neural network systems, the Kohonen SOM “acquires the ability” to classify nonlinear datasets without repetitive inspection of each outcome by the researcher. As described in Ninness et al. (2012), rather than successively examining a sequence of training simulations with training data using correction procedures employed by the user, the SOM acts to “self-organize” and systematically minimize the Euclidean distance among input values, producing a constellation of *best fit* data patterns (see Table 1 regarding *Euclidean distance*). That is, given that a set of unambiguous patterns may exist within a given chaotic nonlinear dataset, the SOM algorithm isolates and numerically classifies nonlinear patterns that form a visually conspicuous cluster of output values. The SOM’s unique capability to capture underlying data patterns and provide illustrative displays of crucial dimensions and characteristics of a given dataset makes this algorithm a particularly important tool when conducting any form of data exploration/data mining and distinguishes it from classical multivariate statistical methodologies (Lagus, Honkela, Kaski & Kohonen, 1996). Detailed mathematical descriptions of neural networking computational procedures can be found in works by Haykin (2009) and Duda, Hart & Stork (2001). Several types of neural networking software architectures have been developed by SAS and are available at:

<http://www.sas.com/technologies/analytics/datamining/miner/neuralnet/index.html>. Likewise, IBM SPSS has developed a series of neural networking architectures including Kohonen networks. Specific details may be found at: [http://pic.dhe.ibm.com/infocenter/spssmodl/v15r0m0/index.jsp?topic=%2Fcom.ibm.spss.modeler.help%2Fkohonennode\\_general.htm](http://pic.dhe.ibm.com/infocenter/spssmodl/v15r0m0/index.jsp?topic=%2Fcom.ibm.spss.modeler.help%2Fkohonennode_general.htm)

From a broader perspective, neural networking procedures in the forms of ANNs and SOMs have been used to analyze a wide variety of data regarding the scientific classification and prediction of world events. Ekonomou (2010) used ANN methodology to predict Greek energy consumption. The network forecast generated a near point-to-point match with the actual data collected for the years 2005-2008. In a similar vein, ANN predictions were 95% accurate in the forecasting of electrical usage in Turkey (Oğcu, Demirel & Zaim, 2012). Clearly, forecasting energy availability and consumption will be a critical issue as countries face diminishing supplies and attempt to regulate the effect on the global climate and energy consumption. Neural networks have even identified nonlinear patterns in the analysis of conventional versus contemporary perspectives regarding Japanese social order and evolving modes of communication (Takane & Oshima-Takane, 2002).

Using SOM architecture, data from the Amsterdam-Amstelland police, family violence cases were correctly classified (Poelmans, Elzinga, Vianene, & Dedene, 2010), resulting in important advances in the identification, prediction, and management of a wide range of domestic problems. In a similar vein, the SOM was employed to develop a *choloropleth* (color

coded) map (Guo, Liao & Morgan, 2007). This project revealed previously unidentified temporal trends in types of terrorist targets across thirty-eight years of international terrorist incidents. The SOM revealed the specific characteristics of weapons (and resulting forms of fatalities), and generated/mapped locations pertaining to terrorist weapons.

During the past ten years, several behavioral software/neural network design laboratories have attempted to make the SOM more *flexible*, and the version described here is now capable of analyzing assorted inputs from a wide range of diversified metric scales concurrently. For example, Ninness, Rumph, McCuller, Harrison, Vasquez, Ford, Ninness, & Bradfield, A (2005) used an early version of this type of architecture to identify and remediate mathematical errors that occurred during computer-interactive instruction. In this study, the SOM identified trigonometric errors in which the participants exhibited common mathematical misconceptions (e.g., regarding order of operator precedence). Identifying student error patterns allowed us to develop a more efficient remediation strategy during the second experiment within the same paper. A more recent study conducted by the same behavioral software design team applied SOM pattern recognition principles to diversified nonlinear data distributions from behavioral and physiological research, extending previous findings that validate the predictive ability of SOM neural networks (Ninness et al., 2012). Three studies analyzing independent datasets were conducted. The first study examined data related to Congressional members' voting patterns. The original votes noted as *Yes*, *No*, and ? were converted to a numerical scale as dummy variables. Follow-up analysis of the SOM classifications by way of logistic regression showed that party affiliations and the probability of specific legislators voting for or against key pieces of legislation were accurately predicted. Using the type of votes cast on congressional legislation, the SOM was able to create classifications of voting patterns. Based on these voting patterns, predictions were made as to how votes were cast on other key pieces of legislation as well as the legislators' affiliations.

The second study analyzed single-subject changes of nine physiological dependent variables. The variables were recorded before, during, and after exposing the subject to a cold pressor challenge. Clearly defined classifications of the physiological variables emerged across the three conditions. Study 3 examined cell diagnostics related to distinguishing benign from malignant cells using digitized breast tissue images. A class of cells (29.7 % of total cases) was determined to be *undifferentiated* and therefore extracted from the post hoc analysis. The remaining cases were classified by SOM outcomes into three classes of benign or malignant cells. Following removal of data classified as undifferentiated, an overall level of accuracy of 99.51% in predicting cell type was obtained by way of follow-up multiple logistic regression and the SOM classifications.

As previously described, new computer technologies are inundating researchers with diversified nonlinear data (Kline, 2009), and very often, classical statistical methodologies are insufficient to address the diversity/nonlinearity of data saturating many academic disciplines. The behavioral and social science research literature abounds with evidence obtained by way of traditional linear multivariate procedures as applied to outcomes that are inherently nonlinear. This methodological complication is exacerbated by the fact that even the more sophisticated statistical techniques may be compromised when non-independent and nonlinear datasets are analyzed by way of conventional multivariate and univariate approaches. Below, we describe neural network techniques in conjunction with classical multivariate statistics, by reanalyzing and systematically contrasting an exemplar dataset from a popular applied multivariate statistics text (Stevens, 2009). In doing so, we provide a preliminary guide to neural networking

methodologies in the analysis of behavior, and we show the benefits of extending the analysis beyond the original/conventional multivariate procedures. As stated by Thuneberg and Hotulainen (2006), “In the psycho-educational field, SOM can show common and expected features among average subjects, but it can also reveal hidden characteristics of the sub-population...which do not share the same properties of the majority of the population” (p. 98). Thuneberg and Hotulainen forward the notion that, “It makes possible to examine both the common trends and the deviations: groups which as they were hypothesized to act, and groups which surprise us” (p. 98).

### Method

As indicated above, the current paper employs a dataset/exemplar from a popular multivariate text by Stevens (2009). In this study, Ambrose (1985) compared elementary school children in an experimental group who received clarinet lessons by way of programmed instruction to students in a control group who were trained by way of traditional music lessons (Stephens, 2009, p. 174). This particular dissertation dataset was selected as a SOM exemplar precisely because it is widely available, and it is employed pedagogically for the purposes of illustrating the utility of applied multivariate analysis within the social and behavioral sciences. For the purposes of comparison, these data were analyzed by way of conventional multivariate procedures and subsequently analyzed by way of the SOM neural network. SOM software was developed by the first author using Microsoft Visual Studio C# 2010 Professional. SOM neural network analyses were conducted on a Toshiba Satellite, i7-3517U CPU, 6 GB RAM, with a 256 GB solid state drive.

### Data Preprocessing

Even though the SOM is an unsupervised neural network developed for the purpose of classifying obscured data patterns within erratic, dissimilar, and nonlinear distributions, finding a common input metric (scaling system) has been a continuing challenge in the general application of this algorithm (see Table 1 regarding *classification*). This is particularly true when the input values are composed of highly diversified metrics with extremely divergent upper and lower end values (e.g., blood pressure, intelligence, and body-weight). Under these conditions, the SOM ability to identify extremely subtle curvilinear patterns has been less impressive (Arciniegas-Rueda, Daniel, & Embrechts, 2001).

In order to address this issue, many neural network systems employ some type of preprocessing of raw data previous to running data through a neural network algorithm. Preprocessing refers to standardizing, transforming, converting characters to numerical values, or identifying locations where omitted values might exist within datasets (Kotsiantis, Kanellopoulos, & Pintelas, 2007). Although several preprocessing strategies have been described in the literature (Vishwanathan & Murty, 2002), we have found that z-score transformations and logarithmic transformations produce excellent uniform distributions while maintaining the primary characteristics of the original input values.

It is important to note that z-score transformations produce normalized metrics using diversified inputs within the same input array (e.g., brain wave activity, blood pressure, academic skills) and thus provide a format for comparing input values around a mean of zero. Logarithmic transformations give diversified measurements a common metric and concurrently show the differences at the upper and lower ends of the log scales. In this study, we employed logarithmic

transformations previous to SOM analysis. Transformation of raw data to natural logarithms does not change the relative shape or positions of the data points across dependent measures. It does, however, compress all of the data points so as to allow the SOM architecture to more clearly discriminate and recognize cohesive patterns within the entire set of input values. It should be noted that our version of the SOM is capable of employing raw data, data transformed into z-scores, natural logs, or common logs, but it is entirely up to the individual researcher's best judgment as to which type of transformation (if any) is most suitable to analyze his/her available data.

## Analysis

As a post hoc measure of SOM grouping effectiveness/pattern recognition, we employed partial least squares regression (PLSR) (see Monecke & Leisch, 2012 for a discussion). This particular multiple regression technique is especially valuable when analyzing an array of “non-independent” predictor variables and when homogeneity of the covariance matrices is in question (Carrascal, Galva'n, & Gordo, 2009). Moreover, this regression procedure is comparatively robust with regard to violations of the assumption of linearity (critical to traditional multiple linear regression) and to the employment of small sample sizes even with a relatively large number of dependent variables (Garthwaite, 1994). We employed *articulation* as the output/predicted variable when running the partial least squares regression analysis (see Table 1 regarding *articulation*).

As previously stated, Stevens (2009) provides data from a study by Ambrose (1985) that compared elementary school children in an experimental group who received clarinet lessons by way of programmed instruction to students in the control group who were trained by way of traditional music lessons. Measurements were obtained by averaging the results of the ratings from two independent judges who directly observed each student's recital. Multiple dependent variables entailed the following musical components of a student's recital: interpretation, tone, rhythm, intonation, tempo, and articulation. As described by Stevens (2009), traditional multivariate procedures (Hotelling's  $T^2$ ) revealed that the familywise multivariate null hypothesis was rejected at the .05 level ( $F = 3.7488, p < .016$ ). Post-hoc procedures by way of the Bonferroni adjustment required each dependent variable to be assessed at a more stringent level (where  $\alpha = .05/6 = .0083$ ), and four of the six dependent variables were statistically reliable with  $p$ -values at or below .006. While Hotelling's  $T^2$  and post-hoc procedures are informative in determining levels of statistical probability, a series of probability values (standing alone) do not provide the reader with a great deal of perspective regarding how student performances show differentiated behavior/skill patterns within and between group. Subsequent to verifying the multivariate outcomes described above, the entire Ambrose (1985) dataset was transformed into natural log values, and a SOM analysis was conducted. It should be noted that our software system allows the process of log transformation and SOM data analysis to be conducted simultaneously. As shown in Figure 1, all raw scores are loaded into a CSV file where the Ambrose data included six columns and twenty-three rows. LL in Figure 1 indicates the lower limit (least number) of classifications anticipated by the researcher. Given these inputs, the researcher simply selects the type of SOM analysis to be conducted by clicking the appropriate button on the software system. For Windows operating systems, the interested researcher may obtain a sample SOM program by contacting the first author.



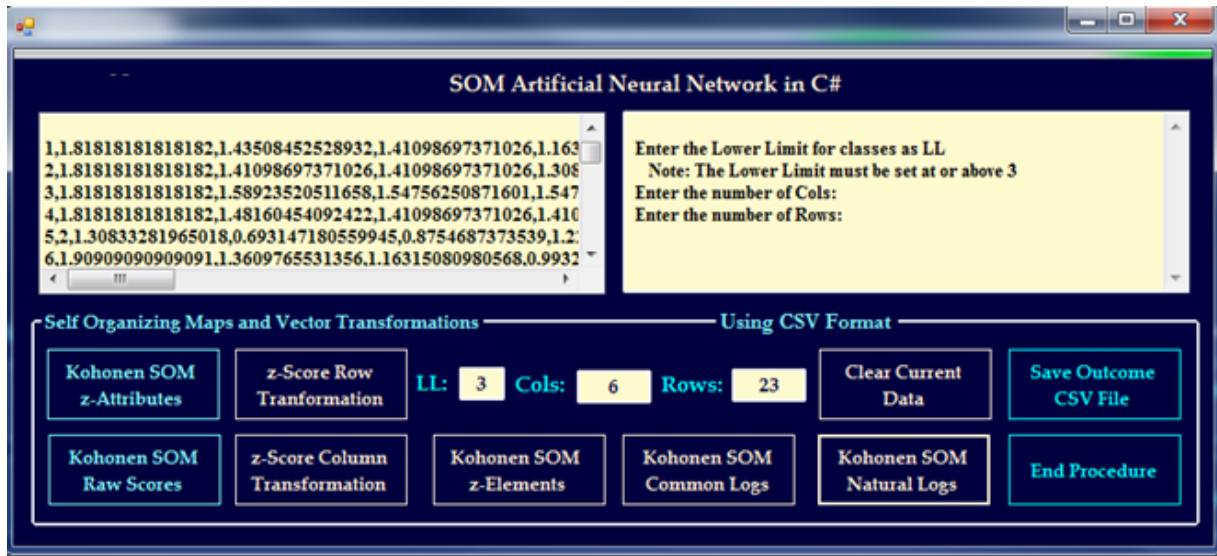


Figure 1. Interested researchers may contact the first author and download our most recent version of the Self-Organizing Map (illustrated above). Instructions for employing this software system are likewise freely available upon request.

## Results

Figure 2 shows the output patterns of the experimental and control groups subsequent to treatment in the form of logarithmic scale scores where PLSR analyses indicate moderate coefficients of determination for both group outcomes. Specifically, in predicting the output variable articulation, the unadjusted PLSR coefficients were calculated at 0.690 for participants in the experimental group and 0.766 for participants in the control group.

Based on the above statistical results, one might expect a more unified series of outcomes encapsulating the respective members of the control and experimental groups; however, two of the five patterns recognized by the SOM actually contained members of both groups. That is, the performance patterns of participants illustrated in Figure 3 show the actual best-fit student performance patterns. Importantly, when these data were classified by the SOM, ratings of student performances showed visually distinctive patterns and exceptionally high PLSR coefficients. Class 1 is composed of seven participants within the experimental group ( $R^2 = 0.919$ ). This class appears elevated, well-unified, and well differentiated from the other classes. Of critical importance, the SOM shows us that other members of the experimental group performed in such a different manner so as to be classified within other groups. It may be of interest to study these highly correlated/clustered groups as they relate to the treatment variables.

Class 2 includes six of the twelve control group participants and appears distinctive with respect to the grouping of students' ratings of lower musical skill levels ( $R^2 = 0.971$ ). This is particularly conspicuous for skill ratings located at 2 and 3 along the  $x$ -axis. Critically important is the fact that Class 3 includes two experimental group and two control group members showing *the same* well-integrated pattern across all six ratings of musical skills ( $R^2 = 0.988$ ). Just as important, Class 4 is composed of two experimental participants and one control group participant, all of whom were classified as having extraordinarily similar skill patterns ( $R^2 = 0.993$ ). Class 5 shows nearly perfect correspondance between two members of the control group;

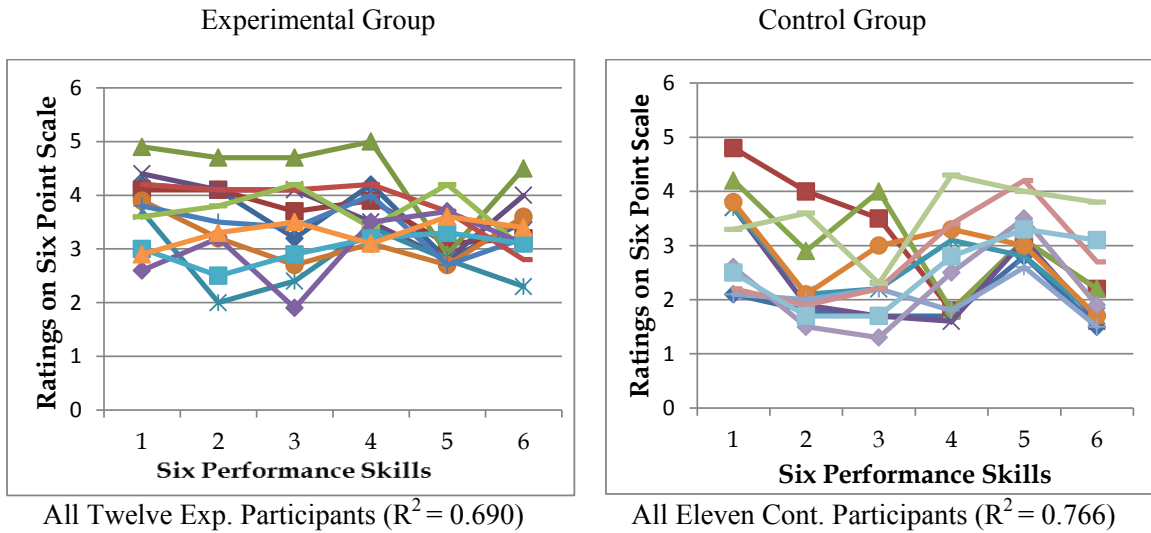


Figure 2. Panel 1 shows twelve members of the experimental group where the PLSR value = 0.690. Panel 2 shows eleven members of the control group where the PLSR value = 0.766. Rather than a single trend line depicting the *average* performance of all participants, each line represents the individual performance ratings across all six dependent variables.

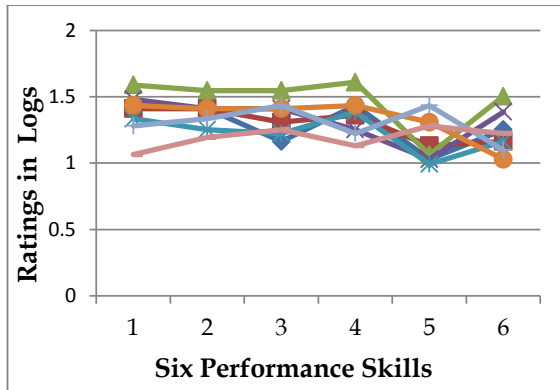
however, PLSR procedures cannot be conducted on classification with fewer than 3 participants in a class. Nevertheless, visual inspection of these data patterns within Class 5 shows the extent to which these outcome patterns overlap conspicuously. The above highly correlated and well differentiated skill patterns were obscured entirely when the analysis was confined to outcomes based on Hotelling's  $T^2$  and conventional multivariate post-hoc procedures.

### Discussion

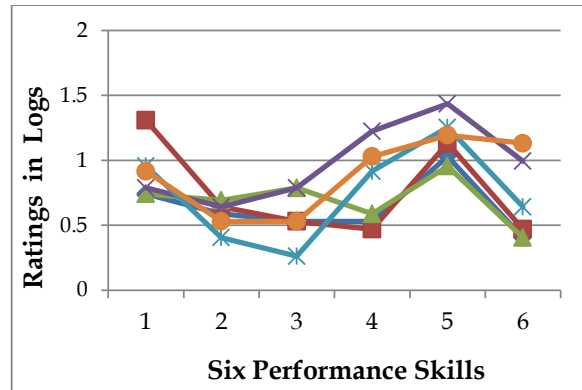
As previously mentioned, neural networks demonstrate increasing accuracy as the data employed become more reliable. Often, rating scales are problematic for neural networks since the values obtained are, by definition, subjective estimates of values that exist along an uneven number line. Nevertheless, we have found that data obtained by way of rating scales often can be “improved” (in the sense that they can be transformed and show more cohesive patterns) when converted to logarithmic scales. Indeed, logarithmic scales have been long employed by behavior analysts to convert the diversified range of human behaviors into a common metric (Lindsley, 1971), and we have found transforming raw scores to logarithmic scores to be an especially valuable strategy during SOM neural network data analysis.

Our immediate ambition was to compare and contrast our version of the SOM with conventional multivariate methodology in an effort to obtain output patterns that are not recognized when using traditional multivariate statistical techniques. Employing the archival dataset from Ambrose (1985), we reanalyzed variables associated with students' recitals. When student performance outcomes (in the form of rating scale data) were transformed into natural logs and analyzed by the SOM, the skill sets of all 23 students were disaggregated into five graphically distinct performance classes. That is, when converted to natural logs and classified by the SOM, measures of student performances showed distinctive and highly correlated pattern formations that were entirely absent when the analysis was confined to conventional multivariate

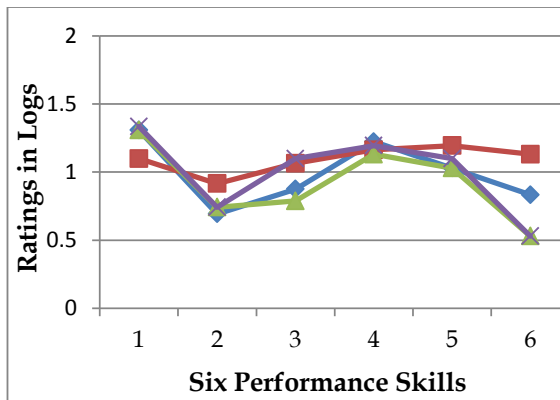
## NEURAL NETWORK AND MULTIVARIATE ANALYSES



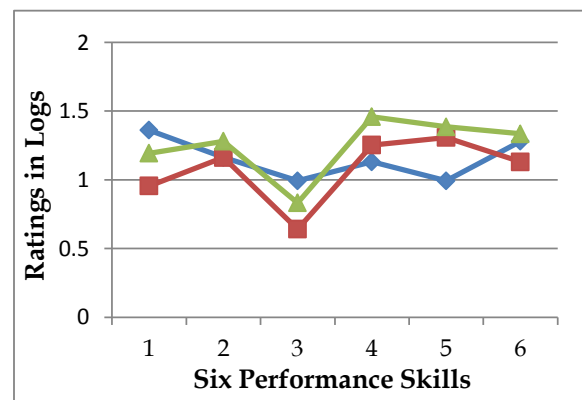
Class 1: Seven Exp. Participants ( $R^2 = 0.919$ )



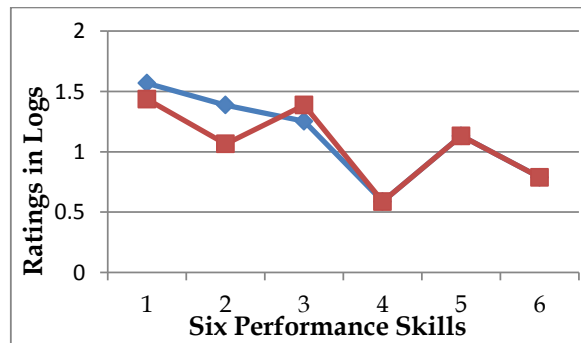
Class 2: Six Cont. Participants ( $R^2 = 0.971$ )



Class 3: Two Exp. and Two Cont. ( $R^2 = 0.988$ )



Class 4: Two Exp. and Two Cont. ( $R^2 = 0.993$ )



Class 5: Two Cont. Participants

Figure 3. Class 1 is composed of seven members of the experimental group. Class 2 is composed of six members of the control group. Class 3 is composed of two experimental and two control group members. Class 4 includes two experimental and one control group participant. Class 5 is composed of two members of the control group. Note that each line represents an individual participant's performance across the six dependent variables.

statistical procedures. These outcomes are consistent with previous investigations focusing on math performance (Ninness et al., 2005) as well as with findings pertaining to SOM analysis of Congressional members' voting patterns, neurophysiological changes, and diagnostics related to distinguishing benign from malignant cells (Ninness et al., 2012).

Ironically, the purpose of employing multivariate analysis and post hoc techniques is to allow researchers to disambiguate findings among and between participants across experimental arrangements. This effort is made to obtain a clearer understanding of the extent to which specified treatments influence the behaviors of participants identified as members of control and experimental groups. However, something very critical is left out of the conventional statistical analyses. Over half a century ago, Sidman (1960) noted within the arena of conventional psychological methodology, subject variability is viewed as a underlying source of experimental *error* while in behavior analytic methodology variability is considered a critical source of experimental *interest* (as described in Ninness et al., 2002, p. 64). The same paradox is apparent in the contemporary statistical analysis of group behaviors. Conventional multivariate statistics applied to the social and behavioral sciences remains devoid of the ability to identify behavior change that may be critically important for individuals as well as "groups within groups." In this sample "textbook study," unified and well-differentiated student performances were undetectable when the examination of available data was confined to conventional multivariate analysis and post hoc procedures.

From our perspective, SOM analyses have particular relevance in identifying clinically significant changes among and within groups and individuals. These procedures allow the researcher to conduct a "finer grain" analysis within comparison groups, identifying meaningful similarities and differences that can inform treatment for individual participants rather than based simply on group means. From a behavior analytic perspective, there may be a more pervasive issue that needs to be considered when looking at the outcomes generated in applied and basic research settings.

Behavior analysts have long been suspicious about outcomes that are predicated on the assumptions of conventional/classical statistics. Single-subject designs allow the researcher to visually inspect graphical illustrations of behavior and determine the extent to which functional control of the independent variable has been established. The determination of functional control does not rely on underlying theoretical assumptions such as homogeneity of variance, homoscedasticity, linearity, or random and independent sampling. In the world of single-subject design, such underlying assumptions make no sense.

It seems reasonable to question the extent to which the SOM architecture might be useful to behavior analysts. As mentioned at the beginning of this paper, we (Ninness et al., 2012) previously employed the SOM using data obtained by way of an A-B-A single-subject design to determine the potential coordinated effects of nine variables associated with sympathetic nervous system cardiac activation and various components of the inner ear during a series of cold pressor challenges. Subsequent to SOM analysis, several well-differentiated classifications were identified by the SOM architecture. In an earlier study (Ninness et al., 2005), we employed a single-subjects design to identify and remediate mathematical errors that occurred during computer-interactive instruction. From our perspective, the SOM has considerable potential to facilitate single-subject research strategies, particularly when the single-case design entails several dependent variables.

At this stage of modern technological development, the same might be said with regard to large group data analysis. As suggested earlier, neural networks are the modern tools that focus

most heavily on the logical structure of measurement and assessment, as well as the actual results we attempt to identify by way of scientific inquiry. While conventional multivariate statistical techniques remain valuable in identifying significant differences between and among groups, the pervasive and almost inevitable violations of basic assumptions underlying conventional statistics suggest the need for alternative (or at least supplemental) group comparison techniques (e.g., Anderson & Robinson, 2001; Edgington, 1995; Ninness et al., 2002; Manly, 2006). We suggest that contemporary neural networking strategies and particularly the SOM neural network are congenial with conventional multivariate statistics in the sense that these strategies provide additional details for conducting a finer grain analysis of diversified outcomes. This appears to be true for a quickly growing number of nonlinear measurements obtained from within the social and behavioral sciences.

### References

- Allamehzadeh, M., & Mokhtari, M. (2003). Prediction of aftershocks distribution using Self-Organizing Feature Maps (SOFM) and its application on the Birjand-Ghaen and Izmit earthquakes. *Journal of Seismology and Earthquake Engineering*, 5, 1–15.
- Ambrose, A. (1985). *The development and experimental application of programmed materials for teaching clarinet performance skills in college woodwind technique courses* (Unpublished doctoral dissertation). University of Cincinnati, Cincinnati.
- Anderson, M. J., & Robinson, J. (2001). Permutation tests for linear models. *Australian & New Zealand Journal of Statistics*, 43, 75–88.
- Arciniegas Rueda, I. E., Daniel, B., & Embrechts, M. (2001). Exploring financial crises data with self-organizing maps (SOM). In N. Allinson, L. Allinson, H. Yin & J. Slack (Eds.), *Advances in Self Organizing Maps* (pp. 30–39). London: Springer-Verlag.
- Burgos, J. E. (2001). A neural-network interpretation of selection in learning and behavior. *Behavioral and Brain Sciences*, 24, 531–532.
- Burgos, J. E. (2007). Autoshaaping and automaintenance: A neural-network approach. *Journal of the Experimental Analysis of Behavior*, 88, 115–130.
- Cavazos, T. (2000). Using self-organizing maps to investigate extreme climate events: an application to wintertime precipitation in the Balkans. *Journal of Climate*, 13, 1718–1732.
- Carrascal, L. M., Galva'n, I. & Gordo, O. (2009). Partial least squares regression as an alternative to current regression methods used in ecology. *Oikos*, 118, 681–690.
- Deboeck, G. (1998). *Visual explorations in finance with self-organizing maps*. London: Springer-Verlag.
- Donahoe, J. W., & Burgos, J. E. (1999). Timing without a timer. *Journal of the Experimental Analysis of Behavior*, 71, 257–263.
- Donahoe, J. W., & Burgos, J. E. (2000). Behavior analysis and revaluation. *Journal of the Experimental Analysis of Behavior*, 74, 331–346.
- Duda, R., Hart, P., & Stork, D. (2001). *Pattern classification*. New York: John Wiley & Sons.
- Edgington, E. S. (1995). *Randomization tests*. New York, Marcel Dekker.
- Ekonomou, L. (2010). Greek long-term energy consumption prediction using artificial neural networks. *Energy*, 35, 512–517.
- Forness, S. (1970). Educational prescription for the school psychologist. *Journal of School Psychology*, 8, 96–98.
- Garthwaite, P. H. (1994). An interpretation of partial least squares. *Journal of the American Statistical Association*, 89, 122–127.
- Guo, D., Liao, K., & Morgan, M. (2007). Visualizing patterns in a global terrorism incident database. *Environment and Planning B: Planning and Design*, 34, 767–784.
- Haykin, S. S. (2009). *Neural networks and learning machines* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.
- Heaton, J. (2008). *Introduction to neural networks for C#* (2nd ed.). St. Louis, MO: Heaton Research, Inc.
- Hebb, D. (1949). *The organization of behaviour*. New York, John Wiley.
- Kaartinen, J., Hiltunen, Y., Kovanen, P., & Ala-Korpela, M. (1998). Application of self-organizing maps for the detection and classification of human blood plasma lipoprotein lipid profiles on the basis of H NMR spectroscopy data. *NMR Biomed*, 11, 168–176.

- Kaski, S. & Kohonen, T. (1996). Exploratory data analysis by the Self-organizing map: Structures of welfare and poverty in the world. In A. P. N. Refenes, Y. Abu-Mostafa, J. Moody, J., & A. Weigend, A. (Eds.), *Neural Networks in Financial Engineering. Proceedings of the Third International Conference on Neural Networks in the Capital Markets* (pp. 498-507). World Scientific: Singapore.
- Kotsiantis, S., Kanellopoulos, D., Pintelas, P., (2007). Handling imbalanced datasets: A review, *GESTS International Transactions on Computer Science and Engineering*, 30, 25-36.
- Kemp, S. N., & Eckerman, D. A. (2001). Situational descriptions of behavioral procedures: The In Situ testbed. *Journal of the Experimental Analysis of Behavior*, 75, 135-164.
- Kline, R. B. (2004). *Beyond significance testing*. Washington, DC: American Psychological Association.
- Kline, R. B. (2009). *Becoming a behavioral science researcher: A guide to producing research that matters*. New York, NY: Guilford.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43, 59-69.
- Kohonen, T. (1984). *Self-organization and associative memory*. Berlin: Springer-Verlag.
- Kohonen, T. (2001). *Self-organizing maps* (3<sup>rd</sup> ed.). Berlin: Springer-Verlag.
- Lagus, K., Honkela, T., Kaski, S., & Kohonen, T. (1996). Self-organizing maps of document collections: A new approach to interactive exploration. In E. Simoudis, J. Han & U. Fayyad, (Eds.), *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining* (pp. 238-243) Menlo Park, California: AAAI Press.
- Lindsley, O. R. (1971). Precision teaching in perspective: An interview with Ogden R. Lindsley, Ann Duncan interviewer. *Teaching Exceptional Children*, 3, 114-119.
- Manly, B. F. J. (2006). *Randomization, bootstrap and Monte Carlo methods in biology*. London: Chapman and Hall.
- Miikkulainen, R. & M. G. Dyer. (1991). Natural Language Processing with Modular PDP Networks and Distributed Lexicon. *Cognitive Science*, 15, 343-399.
- Miikkulainen, R., Bednar, J. A., Choe, Y., & Sirosh, J. (2005). *Computational Maps in the Visual Cortex*. Berlin: Springer. networks and distributed lexicon. *Cognitive Science*, 15, 343-399.
- Monecke A, Leisch F (2012). semPLS: Structural Equation Modeling Using Partial Least Squares. *Journal of Statistical Software*, 48, 1–32.
- Myers, C., Gluck, M., Weinberger, D. & Mattay, V. (2005). Neural mechanisms underlying probabilistic category learning in normal aging. *Journal of Neuroscience*, 25, 11340-11348.
- Ninness, C., & Glenn, S. (1988). *Applied behavior analysis and school psychology: A research guide to principles and procedures*. Westport, CT: Greenwood Press.
- Ninness, C., Lauter, J. Coffee, M., Clary, L., Kelly, E., Rumph, M., Rumph, R., Kyle, R., & Ninness, S. (2012). Behavioral and Biological Neural Network Analyses: A Common Pathway toward Pattern Recognition and Prediction. *The Psychological Record*, 62, 579-598.
- Ninness, C., Rumph, R., McCuller, G., Harrison, C., Vasquez, E., Ford, A., Ninness, S., & Bradfield, A. (2005). A relational frame and artificial neural network approach to computer-interactive mathematics. *The Psychological Record*, 55, 561-570.
- Ninness, C., Rumph, R., Vasquez, E., & Bradfield, A. (2002). Multivariate randomization tests for small-*n* behavioral research. *Behavior and Social Issues*, 12, 64–74.
- Oğcu, G., Demirel, O. F. & Zaim, S. (2012). Forecasting Electrical Consumption with Neural Networks and Support Vector Regression. *Procedia – Social and Behavioral Sciences*, 58, 1576-1586.
- Poelmans, J., Elzinga, P., Viaene, S., & Dedene, G. (2010) Curbing domestic violence: Instantiating C-K theory with Formal Concept Analysis and Emergent Self Organizing Maps. *Intelligent Systems in Accounting, Finance and Management*, 17, 167-191.
- Pomerleau, D. (1991). Efficient training of artificial neural networks for autonomous navigation. *Neural Computation*, 3, 88-97.
- Reusch, D. B., Alley, R. B., & Hewitson, B. C. (2005). Relative performance of self-organizing maps and principal component analysis in pattern extraction from synthetic climatological data. *Polar Geography*, 29, 227–251.
- Sidman, M. (1960). *Tactics of scientific research*. New York: Basic Books.
- Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences*. Hillsdale, NJ: Erlbaum.

## NEURAL NETWORK AND MULTIVARIATE ANALYSES

- Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., Lander, E., & Golub, T. (1999). Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proceedings of the National Academy of Science*, 96, 2907-2912.
- Takane, Y., & Oshima-Takane, Y. (2002). *Nonlinear generalized canonical correlation analysis by neural network models*. In S. Nishisato, Y. Baba, H. Bozdogan, & K. Kanefuji (Eds.), *Measurement and multivariate analysis* (pp. 183–190). Tokyo: Springer-Verlag.
- Thuneberg, H & Hotulainen, R. (2006). Contributions of data mining for psycho-educational research: what self-organizing maps tell us about the well-being of gifted learners. *High Ability Studies*, 17, 87-100.
- Ultsch, A. (2007). Emergence in self-organizing feature maps. *Proceedings of the Workshop on Self-Organizing Maps*. Bielefeld, Germany.
- Valkonen, V., Kolehmainen, M., Lakka, H., & Salonen, J. (2002). Insulin resistance syndrome revisited: application of self-organizing maps. *International Journal of Epidemiology*, 31, 864-871.
- Vishwanathan, S. V. N. & Murty, M. N. (2002). Use of MPSVM for Data Set Reduction. In A. Abraham, L. Jain, & J. Kacprzyk (Eds.), *Recent Advances in Intelligent Paradigms and Applications*, Studies in Fuzziness and Soft Computing (pp. 1-18), Berlin: Springer Verlag.
- Ward, T., & Houmanfar, R. (2011). Human simulations in behavior analysis (1987-2010): Facilitating research in complex human behavior. *Behavior and Social Issues*, 20, 72-101.