Nonlinear Dynamics, Psychology, and Life Sciences, Vol. 17, No. 2, pp. 223-232.

Is Chaos Good for Learning?

J. C. Sprott[1], *University of Wisconsin - Madison*

Abstract: *This paper demonstrates that an artificial neural network training on time-series data from the logistic map at the onset of chaos trains more effectively when it is weakly chaotic. This suggests that a modest amount of chaos in the brain in addition to the ever present random noise might be beneficial for learning. In such a case, human subjects might exhibit an increased Lyapunov exponent in their EEG recordings during the performance of creative tasks, suggesting a possible line of future research.*

***Key Words*:** chaos, neural networks, training, learning, logistic map

INTRODUCTION

The idea that chaos is beneficial in natural systems is now widely accepted. For example, it was once thought that a healthy heart had a regular sinus rhythm, but it was discovered that perfect regularity is a pathological condition (Kleiger, Miller Bigger & Moss, 1987; Sará et al., 2008) and that a healthy heart is arguably weakly chaotic (Denton, Diamond, Helfant, Kahn, & Karagueuzian, 1990). Periodicity implies that the heart is not responding to its surroundings (Goldberger, 1996), and anesthetized subjects exhibit a decrease in heart rate variability (Esmaeili, Shamsollahi, Arefian, & Assareh, 2007). On the other hand, a fibrillating heart appears strongly chaotic and presages death.

Weak chaos allows the exploration of a wider range of conditions while still retaining a degree of memory and predictability. A prey in the wild has a better chance of eluding a predator if its fleeing motion is chaotic, but it must retain some memory of what movements it has previously made and how successful they were (Dawkins, 1995). In this case, evolution selects for weak chaos since it increases the chance of survival and hence reproduction (Humphries & Driver, 1967). Similarly, unpredictable behavior is often desirable in games as a way to hide one's intentions (Maynard Smith, 1982).

Since the human brain is a large network of nonlinearly interacting neurons, it is likely to be chaotic, and such chaos is potentially beneficial. Presumably, too much chaos in the brain is undesirable, and such people end up in asylums or perhaps prison. They behave unpredictably (Lynam & Widiger, 2001) and have impaired memory that inhibits learning. It has been shown that

[1]Correspondence should be addressed to: J. C. Sprott, Department of Physics, University of Wisconsin Madison, 1150 University Ave., Madison, WI, 53706-1390. E-mail: sprott@prhysics.wisc.edu

"mental noise" as measured by variations in reaction time correlates positively with neuroticism (Robinson & Tamir, 2005) and that the power spectrum associated with reaction time can be replicated by chaos from the logistic map (Clayton & Frey, 1997).

However, weak chaos might aid thinking or even be indispensable to it since it allows creativity and prevents one from getting stuck in a rut, endlessly regurgitating the same old ideas (Rolls, 2008). There is a relation between insanity and genius, with both types tending toward unique responses on word association tests, although originality is only one component of creativity (Eysenck, 1993).

One might argue that computers, which mimic some features of the brain, are not chaotic, and thus chaos is not essential for thinking. But a computer exhibits no creativity and only performs repetitive tasks that it is programmed to do. Furthermore, it is following a program supplied by a human, and the program usually involves nonlinearities in the form of branching statements (if-then-else) or explicit mathematical nonlinearities such as x^2 and sin x. As such, it is capable of modeling chaotic and even self-organizing systems, but not true creativity.

One indication of chaos in the brain comes from electroencephalograms. As with heartbeats, EEG signals are nearly periodic and are usually called "waves," although they do not necessarily propagate in space. The fluctuations are not perfectly periodic, however, and the resulting plots appear weakly chaotic (Babloyantz & Salazar, 1985). The oscillations are more nearly periodic in anaesthetized and comatose patients (Sará & Pistoia, 2010; Wu et al., 2011). There have even been attempts to estimate the fractal dimension of the apparent strange attractors and to use the results for diagnostic purposes (Jeong et al., 1998).

There have been numerous studies of the correlation of EEG and other signals from the brain with cognitive tasks. Dietrich and Kanso (2010) review 72 such experiments on creativity and insight in 63 articles, and the results are at best unpersuasive and at worst contradictory. Some workers reported an increase in amplitude and spatial coherence during the performance of creative tasks while others reported the opposite. Significantly, none of these studies appear to have recorded standard measures of chaoticity such as Lyapunov exponent and fractal dimension. Typically, they only consider changes in the amplitude and sometimes frequency and spatial correlation over a narrow range of frequencies without regard to deviations from periodicity. In fact, noise is usually suppressed or ignored in EEG recordings.

Prior to his death in 2005, the eminent medical physicist John R. Cameron often gave a popular lecture entitled "The Physics of Imagination and Creativity" in which he postulated that noise in the brain is responsible for one's ability to imagine new things and that a combination of imagination and knowledge lead to creativity (Cameron, 1988). He thought there are individual differences in the level of such noise and its effects, causing some people to be more imaginative than others, and that highly imaginative individuals tend to have bad memories and vice versa. Cameron points out that medical schools select individuals with good memories, which benefits the patient since no one wants a doctor with a great imagination and a terrible memory. He further proposed experiments in which subjects would be exposed to randomly fluctuating magnetic fields while being tested for short-term memory.

The idea that individuals have an inherent level of creativity is an old one. Spearman (1923) assumed that creativity is equivalent to intelligence, but he stimulated the work of Hargeaves (1927) who began studies of what was later called "divergent" and "convergent" thinking (Guilford, 1950, 1967), a terminology that nicely coincides with the divergent orbits in a chaotic system and the convergent orbits in a stable periodic cycle. Cattell (1971) proposed a related terminology of "fluid" and "crystallized" intelligence, although the later is a misnomer since both kinds of intelligence change over one's lifespan, with fluid intelligence peaking around age 30 or 40, while crystallized intelligence continues to grow until the onset of senility. Guastello, Guastello, and Hanson (2004) argue that creativity is a stable process that skilled individuals can turn on and off, and Logie (2011) claims that working memory is a measure of fluid intelligence. Furthermore, the many sources of noise in the brain have been identified (Faisal, Selen, & Wolpert, 2008) and the beneficial role of such noise is now widely appreciated (Rolls & Deco, 2010).

Although Cameron did not speak much about chaos, he would probably have been as willing to attribute imagination and creativity to chaos as to noise. In fact, it is very difficult to distinguish the high-dimensional chaos that occurs in complex dynamical systems from noise. Some people think that most noise is really chaos since it is produced by classical deterministic processes. Only in the atomic realm does quantum mechanics postulate true randomness, and even that has its detractors including Einstein who insisted that "God doesn't play dice with the world" (Hermanns, 1983, p. 58).

This paper addresses the related but more modest and easily testable hypothesis that weak chaos is beneficial for learning in an artificial neural network and thus presumably also in the human brain. It has been shown that large artificial neural networks tend to be chaotic (Dechert, Sprott, & Albers, 1999), and the usefulness of noise in attractor networks for retrieving an imprinted pattern has been demonstrated (Bar-Yam, 1997). However, the example given here of the benefits of chaos for training an artificial neural network is apparently new, simple, and especially elegant.

TRAINING SET

The task was to present the neural network with a single time series of 512 points taken from successive iterates of the discrete-time logistic map

$$x_{n+1} = Ax_n(1 - x_n)$$

at the accumulation point of $A = 3.5699456718\ldots$ where chaos onsets after any initial transient has decayed, and to train the network to predict each point in the time series from a single previous point. A representative sequence of 32 points from the time series is shown in Fig. 1. In essence, the neural network is approximating the logistic parabola.

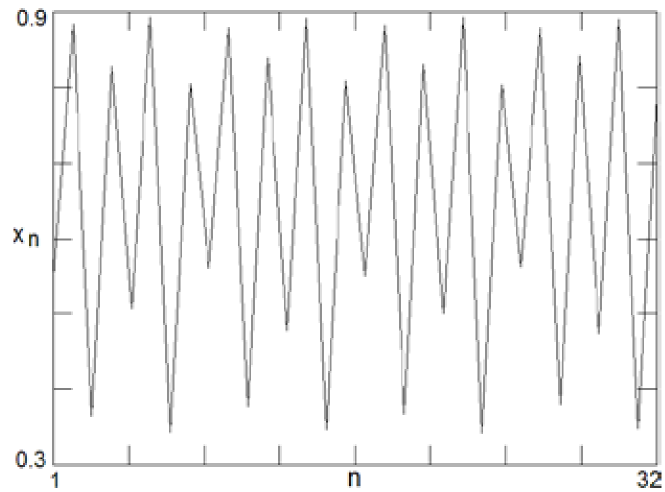

Fig. 1. A representative sequence of 32 points from the training set at the onset of chaos.

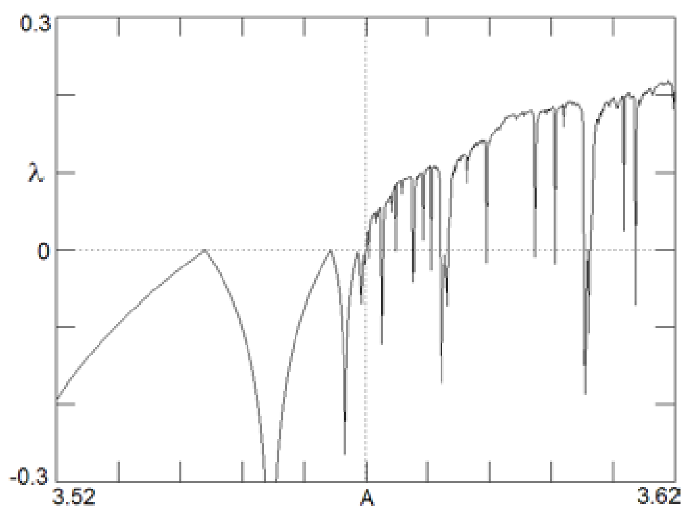

Fig. 2. Lyapunov exponent as a function of *A* for the logistic map, showing the accumulation point at A = 3.5699456718... where chaos onsets.

For this special value of *A*, the attractor is a Cantor set on the parabola with a fractal dimension of about 0.538 (Grassberger, 1981) and a Lyapunov exponent of exactly zero (Sprott, 2003). This is an example of a strange nonchaotic attractor (Feudel, et al., 2006). The reason for this choice is that in the process of training, the network should spend similar amounts of time in the nonchaotic (negative Lyapunov exponent) and chaotic (positive Lyapunov exponent) regions and visit regions where the Lyapunov exponent takes on a continuum of values in the vicinity of zero. Figure 2 shows how the Lyapunov exponent varies with *A* in the vicinity of the chosen value, and these regions are representative of the conditions encountered by the network as it approaches the solution.

THE NEURAL NETWORK

There are numerous architectures and learning rules for artificial neural networks (Hanson & Burr, 1990), but the single hidden layer, feedforward, discrete-time network chosen here is perhaps the simplest case that can exhibit chaotic dynamics. The network approximates each successive value x_n in the time series by a value y_n in terms of the previous value x_{n-1} using the formula

$$y_n = \sum_{i=1}^{N} b_i \tanh(a_0 + a_i x_{n-1})$$

where *a* and *b* are vector connection strengths whose values are adjusted to minimize the mean square error given by

$$e = \frac{1}{511} \sum_{n=2}^{512} (y_n - x_n)^2$$

The cases considered here used $N = 8$ neurons with the program LagSpace (Maus & Sprott, 2011).

Of particular importance is the stochastic training method in which the program explores a gradually shrinking Gaussian neighborhood in *ab*-space centered on the value that gives the smallest current error in the spirit of simulated annealing (Metropolis, Rosenbluth, Rosenbluth, Teller, & Teller, 1953). The training method uses randomness, but the states through which the network passes during the training are purely deterministic and are thus capable of exhibiting chaos. The program tries 10^5 combinations of *a* and *b* after which the error is typically $e \sim 10^{-5}$ and repeats for thousands of instances of the training. Figure 3 shows three typical training instances.

For each instance, the learning rate *R* given by

$$R = -\frac{d \log e}{d \log t}$$

(the negative slope of the curve in Fig. 3) and the Lyapunov exponent λ given by

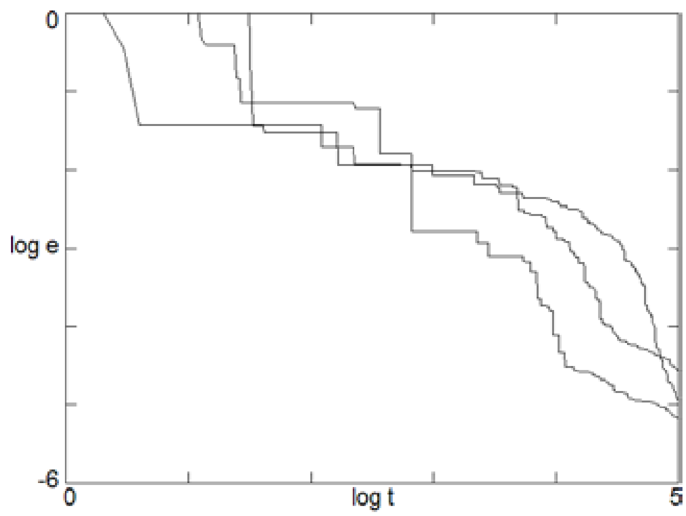

Fig. 3. Three typical instances of the training showing how the error *e* decreases with training trial.

$$\lambda = \frac{1}{512}\sum_{n=1}^{512} \log\left|\sum_{i=1}^{N} a_i b_i \operatorname{sech}^2(a_0 + a_i x_n)\right|$$

are calculated at every 1000 trials where t is the number of the trial, which is proportional to the time spent training. After many weeks of training, approximately 2.8×10^6 pairs of (R, λ) values were collected, a representative one hundred of which during one typical instance of the training are shown in Fig. 4. From the figure, it is evident that the Lyapunov exponent fluctuates with positive and negative values and converges toward zero as the error decreases.

As evidence that a neural network with only eight neurons is capable of training accurately, the smallest mean square error obtained was 4.4×10^{-9}, for which the calculated Lyapunov exponent was 2.0×10^{-3}. The slightly positive Lyapunov exponent is a result of its value having been calculated from a time series of only 512 points.

RESULTS

The main result of this paper is Fig. 5 which shows how the average learning rate varies with Lyapunov exponent in the vicinity of the exact solution whose Lyapunov exponent is zero. This figure was obtained by averaging the learning rate in each of 128 equally sized bins over the range of Lyapunov exponents from −0.05 to +0.05. The learning rate is about 50% greater when the artificial neural network is weakly chaotic ($\lambda \sim 0.02$) than when it is weakly periodic ($\lambda \sim -0.02$).

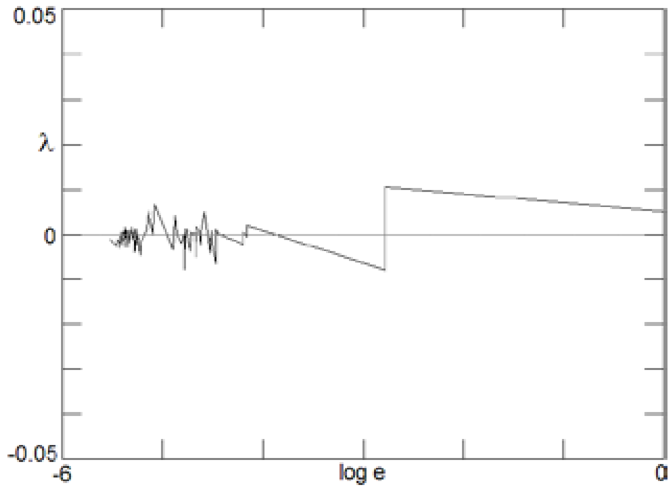

Fig. 4. Typical variation of the Lyapunov exponent during one instance of the training as the error decreases, showing how positive and negative regions are visited.

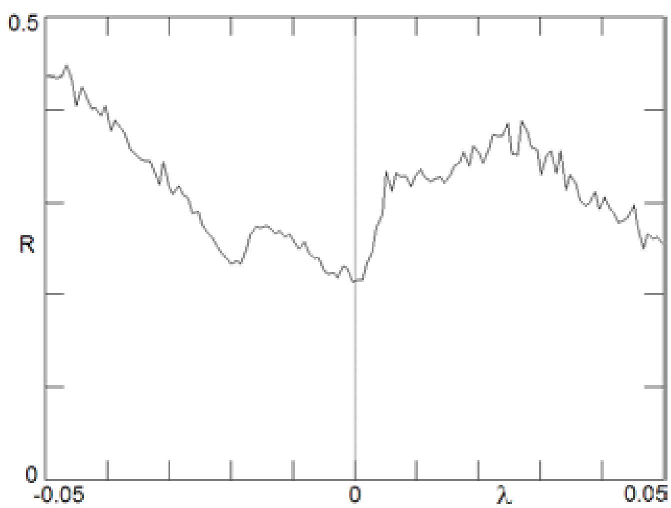

Fig. 5. Average learning rate as a function of Lyapunov exponent in the vicinity of the solution at λ = 0 showing that weak chaos (positive λ) is beneficial for learning in this artificial neural network.

CONCLUSIONS

While it is admittedly a stretch to conclude that this simple computer experiment implies that chaos is beneficial for human learning, it suggests a number of avenues of future research. First, it would be useful to test the robustness of the result by repeating the calculation with other chaotic systems, especially ones chosen to model neural dynamics such as the Hodgkin-Huxley model (Hodgkin & Huxley, 1952), models proposed by Rulkov (2002), and more complex models with many additional variables (Rulkov, Timofeev, & Bazhenov, 2004). In fact, artificial neural networks of this type are capable of generating high-dimensional chaotic dynamics (Albers, Sprott, & Dechert, 1998), and thus one such network could be used to produce a time series at the onset of chaos, which would then be used to train a second identical network.

A second line of research would use human subjects performing memory and creative tasks while their EEG signals are collected for later estimation of the Lyapunov exponent. The hypothesis is that the Lyapunov exponent would be somewhat larger during performance of the creative task, which could be something like drawing a picture or composing an essay than during the memory task, which could be something like recalling a sequence of numbers they had been previously shown and asked to remember. One could also test for individual differences in brain noise or chaos for subjects who score differently on standard tests such as the Torrance Test of Creative Thinking (Torrance, 1974), the Remote Associates Test (Mednick, 1962), or the Alternative Uses Test (Guilford, 1967).

A third more invasive line of research would follow Cameron's suggestion of exposing subjects to fluctuating magnetic fields while testing their memory and creativity. The technology for such tests already exists in the form of transcranial magnetic stimulation (TMS) therapy (Rossini & Rossi, 2007), which is regarded as safe and is used to treat neurological and psychiatric disorders (Slotema, Blom, Hoek, & Sommer, 2010), but typically employs only periodic magnetic pulses. If any effect was observed, one could then see how that effect depends on the character of the fluctuation, in particular its complexity (static, periodic, chaotic, or random) and chaoticity (Lyapunov exponent). A positive result could spawn an industry producing augmented cognition headwear marketed as "learning enhancement caps" programmed either to improve memory or to stimulate creativity, much as the pacemaker regulates the action of the heart. The placebo effect alone should ensure commercial success.

REFERENCES

Albers, D. J., Sprott, J. C., & Dechert, W. D. (1998). Routes to chaos in neural networks with random weights. *International Journal of Bifurcation and Chaos, 8*, 1463-1478.

Babloyantz, J. M. & Salazar, C. N. (1985). Evidence of chaotic dynamics of brain activity during the sleep cycle. *Physics Letters A, 111*, 152-156.

Bar-Yam, Y. (1997). *Dynamics of complex systems.* Reading, MA: Perseus Books.

Cameron, J. (1988). A proposed model for imagination and creativity. *Wisconsin Academy Review, June*, 33-36.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action.* Boston: Houghton Mifflin.

Clayton, K. & Frey, B. B. (1997). Studies of mental "noise." *Nonlinear Dynamics, Psychology, and Life Sciences, 1*, 173-180.

Dechert, W. D., Sprott, J. C., & Albers, D. J. (1999). On the probability of chaos in large dynamical systems: A Monte Carlo study. *Journal of Economic Dynamics and Control, 23*, 1197-1206.

Denton, T. A., Diamond, G. A., Helfant, R. H., Kahn, S., & Karagueuzian, H. (1990). Fascinating rhythm: A primer on chaos theory and its application to cardiology. *American Heart Journal, 120*, 1419-1440.

Dawkins, M. S. (1995). *Understanding animal behavior* (2nd edition.) Harlow: Longman.

Dietrich, A., & Kanso, R. (2010). A review of EEG, ERP, and neuroimaging studies of creativity and insight. *Psychological Bulletin, 136*, 822-848.

Esmaeili, V., Shamsollahi, M. B., Arefian, N. M., & Assareh, A. (2007). Classifying depth of anesthesia using EEG features, a comparison. *Conference Proceedings IEEE Engineering in Medicine & Biology Society, 2007* (pp. 4106-4109). Lyon: IEEE.

Eysenck, H. J. (1993). Creativity and personality: Suggestions for a theory. *Psychological Inquiry, 4*, 147-178.

Faisal, A. A., Selen, L. P. J., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience, 9*, 292-303.

Feudel, U., Kuznetsov, S., & Pikovsky, A. (2006). *Strange nonchaotic attractors.* Singapore: World Scientific.

Goldberger, A. L. (1996). Non-linear dynamics for clinicians: Chaos theory, fractals, and complexity at the bedside. *Lancet, 347*, 1312-1314.

Grassberger, P. (1981). On the Hausdorff dimension of fractal attractors. *Journal of Statistical Physics, 26*, 173-179.

Guastello, S. J., Guastello, D. D., & Hanson, C. A. (2004). Creativity, mood disorders and emotional intelligence. *Journal of Creative Behavior, 38*, 260-281.

Guilford, J. P. (1950). Creativity. *American Psychologist, 5*, 444-454.

Guilford, J. P. (1967). *The nature of human intelligence.* New York: McGraw-Hill.

Hanson, S. J. & Burr, D. J. (1990). What connectionist models learn: Learning and representation in connectionist networks. *Behavioral and Brain Sciences, 13*, 471-518.

Hargreaves, H. L. (1927). The "faculty" of imagination. *British Journal of Psychology* (Monograph Suppl. 10).

Hermanns, W. (1983). *Einstein and the poet: in search of the cosmic man.* Brookline Village, MA: Branden Press.

Hodgkin, A., & Huxley, A. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology, 117*, 500-544.

Humphries, D. A., & Driver, P. M. (1967). Erratic display as a device against predators. *Science, 198*, 1767-1768.

Jeong, J., Kim, D., Chae, J.,Kim, S. Y., Ko, H., & Paik, I. (1998). Nonlinear analysis of the EEG of schizophrenics with optimal embedding dimension. *Medical Engineering & Physics, 20*, 669-676.

Kleiger, R. E., Miller, J. P., Bigger, J. T., & Moss, A. R. (1987). Decreased heart rate variability and its association with increased mortality after acute myocardial infarction. *The American Journal of Cardiology, 59*, 256-262.

Lynam, D. R. & Widiger, T. A. (2001). Using the five-factor model to represent the *DSM-IV* personality disorders: An expert consensus approach. *Journal of Abnormal Psychology, 110,* 401-412.

Maynard Smith, J. (1982). *Evolution and the theory of games.* Cambridge: Cambridge University Press.

Maus, A., & Sprott, J. C. (2011). Neural network method for determining embedding dimension of a time series. *Communications in Nonlinear Science and Numerical Simulation, 16,* 3294-3302.

Mednick, S. (1962). The associative basis of the creative process. *Psychological Review, 69*, 220-232.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics, 21*, 1087-1092.

Robinson, M. D. & Tamir, M. (2005). Neuroticism as mental noise: a relation between neuroticism and reaction time standard deviations. *Journal of Personality and Social Psychology, 89*, 107-114.

Rolls, E. T. (2008). *Memory, attention, and decision-making: A unifying computational neuroscience approach.* Oxford: Oxford University Press.

Rolls, E. T. & Deco, G. (2010). *The noisy brain.* Oxford: Oxford University Press.

Rossini, P., & Rossi, S. (2007). Transcranial magnetic stimulation: diagnostic, therapeutic, and research potential. *Neurology, 68*, 484-488.

Rulkov, N. F. (2002). Modeling of spiking-bursting neural behavior using two-dimensional map. *Physical Review E, 65*, 041922.

Rulkov, N. F., Timofeev, I., & Bazhenov, M. (2004). Oscillations in large-scale cortical networks: map-based model. *Journal of Computational Neuroscience, 17*, 203-223.

Sará, M. & Pistoia, F. (2010). Complexity loss in physiological time series of patients in a vegetative state. *Nonlinear Dynamics, Psychology, and Life Sciences, 14*, 1-14.

Sará, M., Sebastiano, F., Sacco, S., Pistoia, F., Onorati, P., Albertini, A., & Carolei, A. (2008). Heart rate nonlinear dynamics in patients with persistent vegetative state: A preliminary report. *Brain Injury, 22*, 33-37.

Slotema, C. W., Blom, J. D., Hoek, H. W., & Sommer, I. E. C. (2010). Should we expand the toolbox of psychiatric treatment methods to include repetitive transcranial magnetic stimulation (rTMS)?" *The Journal of Clinical Psychiatry, 71*, 873-884.

Spearman, C. (1923). *The nature of "intelligence" and the principles of cognition.* New York: Macmillan.

Sprott, J. C. (2003). *Chaos and time-series analysis.* Oxford: Oxford University Press.

Torrance, E. P. (1974). *Torrance test of creative thinking.* Lexington, MA: Personal Press.

Wu, D., Cai, G., Yuan, Y., Liu, L, Li, G., Song, W., & Wang, M. (2011). Application of nonlinear dynamics analysis in assessing unconsciousness: A preliminary study. *Clinical Neurophysiology, 122*, 490-498.