Human Cognition and 1/f Scaling

Guy C. Van Orden
Arizona State University and the National Science Foundation

John G. Holden
California State University, Northridge

Michael T. Turvey
University of Connecticut and Haskins Laboratories

Ubiquitous 1/f scaling in human cognition and physiology suggests a mind-body interaction that contradicts commonly held assumptions. The intrinsic dynamics of psychological phenomena are interaction dominant (rather than component dominant), and the origin of purposive behavior lies with a general principle of self-organization (rather than a special neurocognitive mechanism). E. J. Wagenmakers, S. Farrell, and R. Ratcliff (2005) raised concerns about the kinds of data and analyses that support generic 1/f scaling. This reply is a defense that furthermore questions the model that Wagenmakers and colleagues endorse and their strategy for addressing complexity.

As science turns to complexity one must realize that complexity demands attitudes quite different from those heretofore common ... each complex system is different; apparently there are no general laws for complexity. Instead, one must reach for “lessons” that might, with insight and understanding, be learned in one system and applied to another. (Goldfeld & Kadanoff, 1999, p. 89)

In a previous article (Van Orden, Holden, & Turvey, 2003), we asserted that background noise in repeated measurements of cognitive performance includes 1/f scaling (also called 1/f noise, fractal time, or pink noise), which may be expected if the behavior of living beings is self-organizing. Wagenmakers, Farrell, and Ratcliff (2005, this issue) commented on various aspects of Van Orden et al. (2003), but we limit our reply to three concerns. First, they claimed that the data series of Van Orden et al. (2003) mostly comprised transient short-range correlations, not 1/f scaling. Second, and related to the first, is their claim that pink noise does not strictly imply self-organization but could be generated by componential models, in particular by the model of Granger (1980). Third, they worried that self-organized criticality is underspecified.

Can We Rule Out Transient Correlations?

The backbone of the commentary of Wagenmakers et al. (2005) is whether transient short-range correlations adequately characterize the data from Van Orden et al. (2003). The hypothesis of transient short-range correlations, however, carries the burden of proof because it asserts something readily observable, a short-range upper bound to visibly scale-free behavior. Wagenmakers et al. explained this fact: “The difference between a persistent 1/f noise process and a transient ... process” is that “a transient process flattens at the lower frequencies” (pp. 110). Although one expects an eventual breakdown in any natural, finite, scaling relation, transient processes are very short lived. 1/f-like patterns may recur across dozens of trials but not hundreds or thousands.

For a transient process, the relation between power (i.e., amplitude) and frequency of variation must break down at some long time scale where random variation appears. For a transient process, the spectral slope must level off to become the zero slope of white noise past some low frequency, a flat plateau of random variation. When plateaus of random variation reliably dominate low frequencies, they are quite obvious in spectral plots (e.g., see Figure 1). This is why visual inspection of spectral plots is such a common and important step in the procedures used to classify time series. When spectral and fractal methods are used in tandem, as in the Van Orden et al. (2003) method, it is especially difficult to mistake transient correlations for scaling relations because the scaling pattern will plateau in the plot of one or the other analysis (Rangarajan & Ding, 2000).

Another way to check for plateaus is simply to collect more data (Baryshev & Teerikorpi, 2002; Mandelbrot & Wallis, 1968/2002). Longer and longer data series reach into more and lower frequency ranges, which provide more and better opportunities to observe plateaus. For example, Figure 2 presents spectral plots of three trial series that appear in Figure 3. The boxes bound the frequencies visible in a spectral plot of 1,024 trials (middle panel) and 1,024 and 4,096 trials (bottom panel). In the examples, more data extend the scaling relation to more and lower frequency scales. The 8,192-trial series establishes the scaling relation over almost four decades of spectral frequencies: Event times of about 200 ms show
In all four panels, the square symbols represent the average power of each spectral coefficient, taken across participants, from a 511-frequency power spectrum analysis of the Van Orden et al. (2003; VOHT) trial series in the double-logarithmic domain. The simple reaction time coefficients were averaged across 10 participants; the naming coefficients were averaged across 20 participants. The error bars represent the standard error of each coefficient's mean power across participants. The two left panels depict the outcome of the analyses without detrending; the two right panels depict exactly the same analyses carried out after using the Van Orden et al. (2003) detrending procedures. Note that the detrending results in a dramatic loss of power for the coefficient corresponding to the lowest resolvable frequency (~3 on the x-axex) in each of the plots depicting the detrended data. Nevertheless, the remainder of the scaling relation stays essentially intact.

Figure 1. In all four panels, the square symbols represent the average power of each spectral coefficient, taken across participants, from a 511-frequency power spectrum analysis of the Van Orden et al. (2003; VOHT) trial series in the double-logarithmic domain. The simple reaction time coefficients were averaged across 10 participants; the naming coefficients were averaged across 20 participants. The error bars represent the standard error of each coefficient's mean power across participants. The two left panels depict the outcome of the analyses without detrending; the two right panels depict exactly the same analyses carried out after using the Van Orden et al. (2003) detrending procedures. Note that the detrending results in a dramatic loss of power for the coefficient corresponding to the lowest resolvable frequency (~3 on the x-axex) in each of the plots depicting the detrended data. Nevertheless, the remainder of the scaling relation stays essentially intact.

long-range dependence across a time span of 10.8 million ms (3 hr). We see comparable outcomes each time we collect longer data series.

So now a discrepancy exists between our conclusions and those of Wagenmakers et al. (2005) based on analyses of data from Van Orden et al. (2003). Wagenmakers et al. found plateaus; we did not. Why? Wagenmakers et al.'s model contrasts were conducted on detrended data, and here is where we see the problem. Detrending can be useful depending on the goals of an analysis (and we get back to this point further on); however, detrending can create false plateaus that do not exist in the actual data. Figure 1 illustrates the problem using the data from Van Orden et al. (2003). Detrending eliminates fluctuations at the lowest frequency and artificially truncates the scaling relation. Detrending flattens power in the lowest frequencies, which creates false plateaus. The false plateaus, in turn, mislead Wagenmakers et al.'s model-testing procedure. The model-testing procedure hinges on the $d$ parameter of the autoregressive fractionally integrated moving average, or ARFIMA, model. But the $d$ parameter is tuned to the region of false plateaus, the region wherein the scaling relation has been artificially flattened. Effectively, a positive test must detect variation that was explicitly removed by detrending. Consequently, detrending sets up Wagenmakers et al.'s analyses to misclassify scaling relations as transient short-range correlations.

The difficulty with our claim is that we also reported analyses in Van Orden et al. (2003) using detrended data. Although we had previously conducted analyses without detrending, we only reported outcomes for detrended data in Van Orden et al. (2003). We did not mention inspection for plateaus in analyses without detrending in part because we have yet to see plateaus in any
comparable data series. We and others have collected many data sets with outcomes like those shown in Figures 1 and 2 and found no reliable plateaus without detrending (see especially Thornton & Gilden, 2004). Consequently, we did not discuss the transient process hypothesis, although Wagenmakers et al. (2005) made plain that we should have clarified the issue.

Our argument hinged on the fractal dimension derived using relative dispersion analysis. One must understand what is at stake in a dispersion analysis to understand why detrending is so prominent in this analysis. Dispersion analysis is a renormalization group procedure used to examine how uncertainty in the estimate of a population mean changes with changes in sample size. In Van Orden et al. (2003), we explained how a dispersion analysis estimates fractal dimension and thereby evaluates the rate at which uncertainty decreases as more data are collected (within the limits of finite data sets). When its assumptions are met, it is among the more robust, unbiased, and efficient tools one can use to estimate fractal dimension (Caccia, Percival, Cannon, Raymond, & Bassingthwaighte, 1997; Eke, Hermán, Kocsis, & Kozak, 2002).

What role does detrending play? To get a sense of this, first take note of the apparent linear trend in the short, 1,024-trial data set in Figure 3. Next notice how other apparent ∼1,000-trial linear trends are simply part of larger oscillations in the 4,096- and 8,192-trial data sets collected from the same participant on different occasions. If data are fractal, then a data set of a given length is sampled from a fractal structure of greater length. Thus the nonstationary linear trend in the short data set is an abbreviated sample of longer range oscillatory trends. One must always bear this fact in mind in fractal analyses of finite samples. Statistical tools are limited in scope to the overt properties of a sample, so assumptions about the origins of samples must be finessed in the analysis. Otherwise, for instance, the powerful nonstationarity in the mean that is so characteristic of fractal processes may dominate and thus distort the outcome.

Nonstationarity in a sample actually contributes uncertainty and creates a bias in favor of the fractal account. By detrending we reduce this bias. Detrending artificially truncates the lowest frequency fluctuations to finesse the limits of statistical tools. It spectral coefficients that span the full range of resolvable spectral coefficients for a 1,024-trial series (cf. Gilden, Thornton, & Mallon, 1995). The slope of a regression line, fit to the lowest 25% of the coefficients, is −.95 (cf. Eke et al., 2002). The middle panel depicts a power spectrum for the 4,096-trial series in Figure 3. The first 511 coefficients were estimated with the Van Orden et al. (2003) window-averaging method, and the three lowest frequencies were estimated from a single pass of a 2,047-frequency analysis, for a total of 514 coefficients. The slope of a regression line fit to the lowest 25% of the coefficients is −.75. The inset box depicts the range of frequencies that would come from a 1,024-trial series. The bottom panel depicts a power spectrum for the 8,192-trial series. The first 1,023 coefficients were estimated with the Van Orden et al. (2003) window-averaging method, and the three lowest frequencies were estimated from a single pass of a 4,095-frequency analysis, for a total of 1,026 coefficients. The slope of a regression line fit to the lowest 25% of the coefficients is −.94. The inset boxes bound the ranges of the frequencies from 1,024- and 4,096-trial series, respectively. None of these series were detrended prior to the spectral analysis because the goal is to determine the full extent of the scaling relation.

Figure 2. The three plots depict spectral analyses for three data series in Figure 3. The top panel depicts a spectral analysis of the 1,024-trial series in Figure 3. The first 127 frequencies were estimated with the same window-averaging method described in Van Orden et al. (2003). The three lowest coefficients were derived from a single pass of a 511-frequency analysis and appended to the 127-frequency spectrum for a total of 130
Figure 3. The top panel depicts the vocal-response 1,024-trial series from Van Orden et al.'s (2003) Simple Reaction Time Participant 9, without detrending. The middle panel depicts a 4,096-trial button-press-response series for the same participant. The bottom panel depicts an 8,192-trial button-press series for this participant. The latter series illustrates a scaling relation close to the limits of what is reasonable for a continuous session. For instance, doubling the sample size to a 16,384-trial series would add just one additional lower frequency coefficient to the spectral plot depicted in Figure 2, but it would require about 6 hr to complete, which would be difficult to achieve without interruption.

imposes a characteristic scale—a constant, definite mean—at the largest scales of fluctuation. Thus detrending guarantees that the analyses are not dominated by a nonstationary linear trend. In turn, the Van Orden et al. (2003) analysis explicitly accommodated detrending by restricting the scales (i.e., frequencies) to one fourth the length of the original trial series, which excludes the false plateau (p. 340, footnote 2; see also Holden, in press). In other words, variation at the lowest frequencies is sacrificed to ensure a more reliable estimate of fractal dimension. Detrending is simply one part of a conservative procedure to estimate fractal dimension and evaluate uncertainty in estimates of the mean.

Does Pink Noise Imply Self-Organization?

As explained above, we have yet to find evidence of transient short-range correlations in our data. Apparently, we confront 1/f scaling. Granting 1/f scaling, one possible explanation would be found in a model with very many separate components. For instance, scaling relations that subtend thousands of trials could be mimicked by a process with a sufficiently large number of independent components, perhaps an infinite number of components. Such a model was outlined by Granger (1980) and is favored by Wagenmakers et al. (2005).

Wagenmakers et al. (2005) speculated that pink noise could be encapsulated in the “behavior of many independent groups of neurons, each with their own different autoregressive decay parameter” (p. 113). However, Wagenmakers et al.’s intuitions about independent neural activity, although plausible, do not mesh with the actual fractal character of the nervous system, so one must look elsewhere for independent component processes. Fractal fluctuations characterize behavior at individual synapses and collections
of synapses (Lowen, Cash, Poo, & Teich, 1997), and action potentials themselves reveal 1/f scaling (Matveev & Wang, 2000). Fractal fluctuations also appear in the event times of intermittent synchrony of electrical activity in the human brain (Gong, Nikolaev, & van Leeuwen, 2003), and long-range correlations appear in functional magnetic resonance imaging time series (Friston et al., 1995). In fact, 1/f scaling characterizes in vivo, excitable-tissue recordings “from the microscopic to the macroscopic” (Lowen et al., 1997, p. 5673; Bassingthwaighte, Liebovitch, & West, 1994; West & Deering, 1995).

In this light, we worry that Wagenmakers et al.’s (2005) newly proposed encapsulated mechanisms lack physical or psychological motivation before the fact and require that one posits implausible coincidences after the fact (see also Hausdorff & Peng, 1996; Thornton & Gilden, 2004). Each arbitrarily chosen frequency is equated with a special psychological or physical mechanism changing on that specific frequency of variation—infinitesimal frequencies equal infinite mechanisms. The amplitude of variation for each mechanism must line up with the scaling relation between power and frequency, which is treated effectively as a collective succession of coincidences. But the actual scaling relation is not tied to any specific frequencies. It provides no motivation for the specific frequencies and amplitudes of variation of models: It is scale free.

A scaling relation is legitimately a phenomenon in and of itself, a fact that Wagenmakers et al. (2005) did not consider. Criticality, as in self-organized criticality, predicts emergent 1/f scaling or pink noise, which should be widely observed in human performance. Pink noise is widely observed in human performance. Does the inverse of this deduction hold true? Does the presence of pink noise strictly imply self-organization? No. As stated in our original article, “Ubiquitous pink noise is not sufficient evidence for self-organized criticality; it is simply a necessary consequence” (Van Orden et al., 2003, p. 343).

Is Self-Organized Criticality Underspecified?

The third issue of this reply is Wagenmakers et al.’s (2005) concern that a hypothesis of self-organized criticality is underspecified. If we correctly understand the commentary, they would see any hypothesis as underspecified that did not eventually yield a conventionally reductive, mechanistic model of cognition.

We concede that the divide between Wagenmakers et al. (2005) and Van Orden et al. (2003) on this point may equal the distance between paradigms. Thus we must defend a point of view and a modeling strategy that they seem to reject out of hand. To begin with, it is legitimate to propose hypotheses that refer to emergent properties (Anderson, 1972), and it is reasonable to speculate that criticality emerges spontaneously in living systems (Bak, 1996; Kauffman, 1993, 1995). It is also credible that 1/f scaling in behavioral measures refers to an emergent property of a human being, in body and mind. Now how might we formulate this as a modeling problem? At what level do we model emergent behavior of the human organism; what are the building blocks? There is presently no workable entry level below the level of the emergent phenomena themselves.

For instance, 1/f scaling appears to be a universal feature of human behavior. Human universals are routinely attributed to causal bases in the human genome, a rationale that appears in evolutionary psychology and cognitive science. Thus the entry level for a fully specified causal model would be at the human genome. But complexity immediately overwhelms this entry point, far below the level of organismic behavior. There are too many interacting gene products, many of which are enzymes, receptors, members of signaling sequences, and other functional parts of metabolism. Deriving features of human behavior from “a system with this many interacting components . . . is clearly out of the question” (Whitesides & Isgamilov, 1999, p. 91).

We are equally frustrated as we move up in scale: “Studies of metabolic cycles and signaling pathways . . . demonstrate the current difficulty in rationalizing even the behavior of these relatively simple systems, much less the emergent properties of organisms” (Whitesides & Isgamilov, 1999, p. 89).

And what about the conventional scale of cognitive mechanisms? In Van Orden et al. (2003), we explained why emergent macrolevel behavior is antithetical to the conventional reductive pursuit of cognitive mechanisms. Aside from that, not one cognitive mechanism exists on which cognitive scientists can agree about its boundaries, its empirical shape, or details about its function. This criticism has been raised and elaborated on by many contemporary commentators (e.g., Harley, 2004; Stanovich, 2004; Thelen & Smith, 1994; Uttal, 1990, 2001; Watkins, 1990; Weldon, 1999).

So where is the entry level to model an emergent property? In Van Orden et al. (2003), we began with a question about 1/f scaling based on Juárez’s conjecture. Juárez (1999) had argued persuasively that conventional assumptions about intentional behavior are nonstarters and to date, to our knowledge, no one has successfully refuted her critique. She went on to speculate that workable assumptions might be found in self-organizing phenomena. She conjectured that intentional behavior originates in embodied states of self-organized criticality. She made this bold conjecture without apparent knowledge of widespread findings of pink noise in human performance or physiology, but her conjecture anticipates widely observable pink noise.

The empirical question that we posed in Van Orden et al. (2003) was whether 1/f scaling or pink noise is widely present in human performance. The phenomenological model of Gilden (2001) and colleagues addresses this question at the level it is posed (Thornton & Gilden, 2004). In their fractal model, spectral data comprise a 1/f scaling relation plus white noise, exactly the level of the emergent phenomenon. This is the modeling strategy common to studies of complex systems. “One should most often use a more phenomenological and aggregated description, aimed specifically at the higher level” (Goldenfeld & Kadanoff, 1999, p. 88). The modeling strategy works because patterns at the macroscale of behavior are independent of the details of microscale motion. Systems vastly different in their structural details may display the same recognizable patterns in their behavior—patterns that may include but are by no means limited to 1/f scaling.

Only the most transparent “nice pile” complex systems lend themselves to the kind of solutions that Wagenmakers et al. (2005) espoused. For instance, “the study of complexity in [chemical] systems of reactions . . . that can be described by nonlinear equations has been limited to the few that are both complex enough to be interesting and simple enough to be tractable” (Whitesides & Isgamilov, 1999, p. 89). These model systems are studied for abstract universal principles that may characterize more complex
opaque systems such as human activity. Model systems like sand piles are informative about interaction-dominant dynamics and self-organization, for example.

A fractal account does not necessarily require details of internal mechanisms to answer the theoretical questions that are posed. In the case of opaque complex systems, it is their behavior that motivates the hypothesis of self-organization, not the details of interacting primitives. Wagenmakers et al. (2005) did not consider these implications of modern complexity science. Not every scientific problem yields analytic solutions or succumbs to conventional reductive analysis (Camazine et al., 2001; Kelso, 1995; Rosen, 1991, 2000; Turvey, 2004; West & Deering, 1995).

Is self-organized criticality underspecified? From Wagenmakers et al.’s (2005) perspective, the answer appears to be yes, because a mechanistic model of emergence in a human organism is not forthcoming. From our perspective, the answer has to be no, because we make adequate progress as the assumptions and principles of complex systems, including self-organized criticality, are elaborated. The latter approach has worked famously for movement coordination (Amazeen, Amazeen, & Turvey, 1998; Kelso, 1995), and it may also suffice to characterize situated cognitive activity (e.g., Hock, Balz, & Smollon, 1998; Hock, Schönner, & Voss, 1997; Thelen, Schönner, Scheier, & Smith, 2001; Tuller, Case, Ding, & Kelso, 1994; Van Orden, Holden, Podgornik, & Aitchison, 1999). Moreover, this approach makes a place for intentionality, which is conspicuous by its absence in mainstream accounts (Van Orden & Holden, 2002).

Last, we remark on Wagenmakers et al.’s (2005) concern that Van Orden et al. (2003) advanced no predictions beyond $1/f$. Our primary target in Van Orden et al. (2003) was the question for cognitive science What kind of system do we study? The answer, a system with $1/f$ scaling, sets the stage for novel, heterodox expectations. In Van Orden et al. (2003), we discussed two immediate predictions: (a) cognition—whatever its nature—does not divide into statistically independent processes, and (b) the same processes govern cognitive performance in very short and very long time frames. These predictions remind us of, for instance, repeated calls to investigate in earnest the possibility of a single memory system and a unitary theory of forgetting (McGeoch, 1932; Melton, 1963; Nairne, 2002; Neath & Surprenant, 2003).

Beyond immediate expectations, surprising uninformed predictions await. Cursory looks at contemporary physics and biology are reassuring on this count. Early enterprises identified scaling and fractal phenomena, which laid the foundation for contemporary work: broad empirical investigations of nature’s mudler but more generic aspects guided by predictions derived through fractal equations (e.g., Kinzig & Harte, 2000; Shapir, Raychaudhuri, Foster, & Jorne, 2000). We anticipate similar developments in cognitive science.

References


Received March 26, 2004
Revision received October 26, 2004
Accepted October 27, 2004

Wanted: Old APA Journals!

APA is continuing its efforts to digitize older journal issues for the PsycARTICLES database. Thanks to many generous donors, we have made great strides, but we still need many issues, particularly those published in the 1950s and earlier.

If you have a collection of older journals and are interested in making a donation, please e-mail journals@apa.org or visit http://www.apa.org/journals/donations.html for an up-to-date list of the issues we are seeking.