Long-range interactions in sequences of human behavior

Martin P. Paulus*
Laboratory of Biological Dynamics and Theoretical Medicine and Department of Psychiatry, School of Medicine, University of California, San Diego, La Jolla, California 92093-0603
(Received 26 June 1996; revised manuscript received 16 October 1996)

Measures derived from the thermodynamic formalism for dynamical systems are applied to human behavior to determine the degree of interaction within sequences of choices. Sequences of 500 binary choices generated by 22 human subjects are analyzed using autocorrelation and mutual information functions, as well as the fluctuation spectrum of local dynamical entropies or local spatial scaling exponents. The main results are as follows: (1) choices generated by the subjects are interdependent on short, midrange, and large scales relative to the 500 choices generated in the paradigm; (2) human behavior is characterized by large fluctuations of dynamical and geometrical choice characteristics; and (3) these geometric and dynamical characteristics can be related to a few competing principles. [S1063-651X(97)01203-8]

PACS number(s): 87.10.+e

The observation that complex temporal patterns can be generated by simple underlying systems [1,2] has drawn increasing interest to determine whether biological time series are generated by nonlinear dynamical systems. Further exploration of dynamical systems has shown that low-dimensional, chaotic dynamics is only one type of complex system's behavior [3–5]. In particular, spatially extended systems can generate time series that exhibit exponential divergence of initially nearby observations, characteristic for chaotic dynamical systems, but also exhibit time series that show a logarithmic divergence of nearby conditions [4]. Finally, ''apparently stochastic'' behavior can be generated by deterministic, spatially extended systems [6].

Ergodic theory has been used to describe both ''stochastic'' and deterministic systems and relies on topological and probabilistic invariant measures to characterize dynamical systems [7]. Both deterministic and stochastic systems can generate time series with indistinguishable topological and metric dynamical characteristics [7]. However, several investigators have recognized that a single measure does not adequately describe the complexity of the underlying dynamical system [8–10]. For example, the average uncertainty of a new observation does not capture information about subsequences that may be highly predictable or highly unpredictable. Moreover, similar average invariant measures can be generated by completely different distributions. Thus, in order to assess the contributions of subsets of observations with different dynamical properties the ergodic theory of dynamical systems was embedded in a thermodynamic formalism [8].

Here the thermodynamic formalism for dynamical systems is used to quantify the degree of interaction in sequences of human behavior to determine whether complex patterns of behavior are generated by simple underlying systems. Individual behavioral elements or subsequences of observations are considered analogous to individual particle or spin configurations within a macroscopic system [11]. Quantifying the interactions between individual behavioral elements provides measures to calculate macroscopic quantities of the system in analogy to the spin-spin interactions in statistical mechanics.

Several investigators have attempted to quantify the degree of interdependence found in sequences generated by biological systems [12–16]. In particular, long-range correlations have been described in a variety of systems ranging from eukaryote or prokaryote DNA [13,15] to human heartbeat intervals [16]. Four different methods have been suggested for the analysis of sequential data. First, the *power spectrum* and/or *autocorrelation* analysis attempts to determine the degree of linear relationship among elements of the biological sequence separated by k elements. Second, calculating a distance metric between observations separated by k elements uses the concept of a *fractal dimension* [17,18] to quantify the correlation in a sequence. Third, *Kolmogorov* or *metric entropy* measures [19] can be computed based on the divergence rate of similar sequences. In this case, the average branching rate or number of distinct observations weighed by the probability of their occurrence yields an estimate of the metric entropy as defined in the context of dynamical systems [20]. Finally, the *mutual information* can be computed from the joint probabilities of individual observations separated by k elements [13]. In particular, the latter two methods not only quantify linear interactions, but also assess higher-order dependences.

In many biological systems, individual observations are not organized on a particular temporal or spatial scale but show measurable relationships across a range of these scales [14,16,21]. Here the interdependence between elements of human behavior, presumably the most complex biological system, was assessed using a simple experimental setup to address the following questions. First, the nature and extent of the interactions between individual behavioral elements was determined using the methods described above. Based on the theoretical results from complex interacting systems, both logarithmic and exponential decays of information were hypothesized to coexist. To determine the nature of the interdependences on different scales, four different methods were used, i.e., autocorrelation functions, mutual information, spatial scaling exponents, and dynamical entropy analy-

*Fax: (619)-543-2493. Electronic address: martin@rat.ucsd.edu

sis. Second, the degree of fluctuation was examined to determine whether average measures sufficiently describe the interaction between individual behavioral elements. Based on the statistics of multiplicative processes, it was hypothesized that the degree of interdependence between human behavioral element fluctuates significantly. The fluctuation spectrum of local dynamical entropies and local spatial scaling exponents was computed to address this question.

The organization of human behavioral sequences is assumed to be analogous to the generation of patterns observed in coordinated hand movements [22]. In both cases, macroscopic variables are used to assess the self-organized state of the subject presumably due to the organization of responses on different levels in the central nervous system. Moreover, as behavioral elements are organized many possible combinations of elements are eliminated. This phenomenon is ubiquitous in complex systems and has been described by the Slaving principle [23] center manifold theorem [24], or adiabatic elimination [10].

The application of ergodic theory and the thermodynamic formalism to biological data is associated with a complex set of problems due to the statistical properties of these data sets [25]. Specifically, time scales are frequently not separable in biological data leading to nonstationary distributions. Moreover, statistical error measures of short sequences are typically large and numerical estimates are typically biased due to the poor converging properties of multiplicative processes. It has been proposed that the measures obtained should be tested against the least complex alternative generator providing the null hypothesis. Randomized data sets are used to test the null hypothesis that sequences result from an identically distributed random process. Therefore, to assess the sequential properties of human behavior the estimated dynamical properties behavioral sequences are compared to randomized data throughout this investigation.

It is reported here that human behavioral elements interact differentially on small and relatively large scales. While small-scale interactions appear to be determined primarily by exponential decay, the relatively-large-scale decay is characterized by logarithmic or even no noticeable decay of information. Thus both logarithmic and exponential decay of information occur and may result from the large fluctuation of local information decay indices. Therefore, it appears that multiple processes determine the organization of human choice behavior.

For the analysis of interdependences between human behavioral elements, consider a choice s_i from m different choice alternatives and $\{s_1 \cdots s_m\}$ in case of the binary choice task $m=2$. The experimental trial generating N choices is represented via $\{s\}^N$. For the thermodynamic formulation, consider subsequences of choices of length l starting at the jth choice $\{s\}_j^l$. Within the thermodynamic formalism, macroscopic quantities are derived to describe the behavior of the system based on $\{s\}_j^l$. In the case of choice behavior, a local interdependence measures its frequency, and the length of the subsequences considered corresponds to thermodynamic conceptualizations of energy, entropy, and volume, respectively.

The autocorrelation function $C(k)$ and the mutual information $I(k)$ have been used to assess the interdependence between individual elements in a sequence of experimental observations. Following the derivation of Herzel and Grosse [13], given the probability of observing an individual element i or element j, p_i and p_j, respectively, as well as the joint probability of element i and j separated by k observations, $p_{ij}(k)$, the mutual information is defined via

$$I(k)=\sum_{i,j=1}^{M} p_{i,j}(k)\ln\left(\frac{p_{i,j}(k)}{p_i p_j}\right).$$

The mutual information, when expressed in $\log_2$, quantifies the average bits of information contained in one observation that will specify the other observation. In analogy, assigning real numbers $\{a_i,\ i=1,\ldots,M\}$ to the partitioned elements in phase space and assuming ergodicity, the autocorrelation function is defined by

$$C_a(k)=\sum_{i,j=1}^{M} [p_{i,j}(k)-p_i p_j]a_i a_j .$$

For binary sequences, both functions yield similar information and have been used to determine long range correlations between individual experimental observations as detailed in [13].

The autocorrelation and mutual information functions quantify the statistical dependence of sets of symbols; however, they do not assess the deviation of the statistical dependence for individual symbol sequences. Following the development of the bilinear form for both the autocorrelation function and the mutual information as shown in [13], a matrix $\hat{D}(k)$ with entries $D_{ij}(k)=p_{ij}(k)-p_i p_j$ can be used to express both the autocorrelation function

$$C_a(k)=a\hat{D}(k)a_T$$

and the mutual information by a Taylor expansion

$$I(k)=\frac{1}{2\ln 2}\sum_{i,j=1}^{M}\frac{D_{i,j}(k)^2}{p_i p_j}+O(D_{i,j}(k)^3).$$

In both cases, the sum for each row and each column of $\hat{D}(k)$ has to vanish. However, the deviation of $D_{ij}(k)$ is not considered. Therefore, individual sequences may differ significantly from the stochastic independence, while the system as a whole may still behave as an independent system. Therefore, a large-fluctuation statistical approach may provide additional information about the statistical mechanics governing the individual interdependences between sequence elements.

The average dynamical entropy h quantifies the average degree of association between choices within the entire sequence of experimental observations. The assessment of the fluctuations of dynamical entropies provides important information about the rule that generates the sequence and are assessed within the framework of the large-fluctuation statistics [9] based on the interdependence between elements in subsequences of the observations. Specifically, local dynamical entropies are defined for the subsequences based on the subsequence occurrence probability (p_i) and the subsequence length (l_i), $h_\iota=-\ln(p_i)/l_i$. This measure is analogous to a ''local spread of information'' or the rate at which a particular subsequence becomes unique. A technique

analogous to nearest-neighbor methods has been suggested to calculate efficiently the average dynamical entropy, but can be easily extended to calculate these local dynamical entropies [26]. This method determines the uniqueness of a subsequence and calculates the local dynamical entropy via

$$h_i = \frac{\ln(N)}{l_i},$$

where N corresponds to the length of the entire sequence and l_i is the length of the subsequences that identifies the subsequence as unique. The development of the thermodynamic formalism for the dynamical entropies h_i based on this method is derived from the definition of the mass exponent via

$$1 = \sum_{i=1}^{N} \frac{e^{q l_i h_i}}{e^{l_i \tau(q)}}.$$

Typically, the mass exponent is considered a function of q. However, in this case, rearranging the above summation given the definition of h_i yields

$$1 = e^{q \ln(N)} \sum_{i=1}^{N} e^{-l_i \tau(q)},$$

thus giving q as an explicit function of $\tau(q)$,

$$q = \frac{\ln\left(\sum_{i=1}^{N} e^{-l_i \tau(q)} \right)}{\ln(N)}.$$

Subsequently, taking the inverse of the partial derivative of $q(\tau)$ to derive an explicit expression for $h(q)$ yields

$$\frac{1}{\partial q / \partial \tau} = \ln(N) \frac{\sum_{i=1}^{N} e^{-l_i \tau(q)}}{\sum_{i=1}^{N} l_i e^{-l_i \tau(q)}}.$$

Specifically, a τ weighted average of the unique subsequence lengths provide the estimation of the q-dependent entropies via

$$h(q) = \frac{\ln(N)}{\langle l \rangle_\tau},$$

where $\langle l \rangle_\tau$ stands for the ensemble average weighted by the mass exponent. The summation over all subsequences as given above can also be used to derive the expression for the fluctuation spectrum of local dynamical entropies $S(h)$[36]. The number of local entropies h_i, $Nb(h_i)$, is assumed to be described by

$$Nb(h_i) = e^{l_i S(h_i)}.$$

Consequently, rearranging the above sum over all observations to a sum over all local entropies within the range h, δh, gives

$$1 = \sum_{h=h_{\min}}^{h_{\max}} Nb(h) e^{qlh - l\tau(q)}.$$

Inserting the assumed distribution yields

$$1 = \sum_{h=h_{\min}}^{h_{\max}} e^{-l[-S(h) - qh + \tau(q)]}.$$

In this summation the individual terms will contribute most significantly if the term within square brackets approaches zero given the varying q parameter; thus

$$\min_q [-S(h) - qh + \tau(q)]$$

implies

$$S(h) = \tau(q) - qh,$$

which corresponds to a Legendre transform in from variable set $(q, \tau(q))$ to $(h, S(h))$.

The preceding was based on the assumption of an exponential spread of information. In contrast, here the investigation of the choice sequences is set up on a logarithmic scale. To analyze the sequences of choices, the data were transformed as

$$s(i) = \begin{cases} +1 & \text{if the } i\text{th choice} = R \\ -1 & \text{if the } i\text{th choice} = L. \end{cases}$$

Subsequently, a walk transformation function was constructed for the ith choice as

$$W(i) = \sum_{j=1}^{i} s(j).$$

This transformation is analogous to the path observed for a one-dimensional random walker [27]. It is well established that a random walker generates a self-similar trajectory and diffuses like t^h, $h = \frac{1}{2}$, corresponding to a fractal dimension of $2 - h = 1.5$. Moreover, anamolous diffusion has been defined via exponents $h > \frac{1}{2}$ or $h < \frac{1}{2}$ describing anomolously fast or slow diffusion, respectively. In the case of the binary choice task, anamolous fast diffusion corresponds to sequences of similar choices, i.e., simple repetitions of one or the other choice alternative. In contrast, anomolous slow diffusion is characterized by alternating choice sequences, i.e., anticorrelated adjacent steps. Finally, the autocorrelation function is described by a δ function for $h = \frac{1}{2}$, but retains residual correlations for $h \lessgtr \frac{1}{2}$. Anomolous diffusion is thus characterized by the interdependences on all levels of resolution.

Here the distance L between positions of W that are k steps apart is assumed to scale as k^h. Extending the method described in [16] in order to assess the fluctuations of h for

different subsequences of the W starting at the ith choice, a local spatial scaling exponent $d=2-h$ was defined by

$$L(i,k)\approx k^{2-d(i)}.$$

The exponent $d(i)$ quantifies the geometric properties of a walk segment. Specifically, a high exponent indicates that the distance does not increase significantly with an increasing number of steps, corresponding to anomolously slow diffusion or an anticorrelated behavioral sequence. On the contrary, a low exponent signifies a large increase of distance upon increasing the number of consecutive steps or fast diffusion via a correlated behavioral sequence.

To quantify the contribution of the occurrence of a particular local scaling exponent $d(i)$ to the ensemble average of all subsequences for the walk transformation, it has been proposed that the probability density $\rho(d)$ of finding a segment on scale k that exhibits a particular local spatial scaling exponent d can be described via

$$\rho(d)\approx k^{-[d-f(d)]}.$$

That is, the frequency of occurrence depends on the scale and a function that explicitly depends only on the local spatial scaling exponent, i.e., $f(d)$. In the special case where $f(d)=d$, we obtain the probability density to be 1 and nonvanishing. This d value characterizes the overall geometric scaling property of the entire walk transformation. This scaling function has been called a *spectrum of singularities* [28] or a *spectrum of local scaling exponents* [21].

Analogous to the fluctuation spectrum of local dynamical entropies, the spectrum of local spatial scaling exponents can be estimated using a canonical approach. Specifically, an undetermined multiplier q, is introduced and generates a maximum likelihood estimation of the probability density for the different levels or resolution. The mass exponent $\tau(q)$, defined via

$$k^{-\tau(q)}\approx\sum_{i=0}^{N-k} L(i,k)^{q-1},$$

consists of a q-weighted summation over all local lengths of path segments. Rearranging the summation to sum over all different local distances L with sampling width δL, i.e.,

$$k^{-\tau(q)}\approx\sum_{L} \rho(L(i,k))\delta L L(i,k)^{q-1},$$

using the scaling assumption $L(i,k)$, and eliminating prefactors yields

$$k^{-\tau(q)}\approx\sum_{L} \rho(L(i,k))\delta L k^{-(q-1)d(n)}.$$

Introducing the $f(d)$ function, one may replace the summation over the different lengths by a summation over the various local scaling exponents and their densities

$$k^{-\tau(q)}\approx\sum_{d} k^{-[d-f(d)]}k^{-(q-1)d}.$$

This summation will give appreciable contributions for various q values only if the difference between qd and $f(d)$ is minimal. Thus the mass exponent $\tau(q)$ can be used to obtain the $f(d)$ value using

$$f(d)\approx qd-\tau(q).$$

For the explicit calculation of $f(d)$, the q-dependent spatial exponent $d(q)$ that minimizes the difference above is given by

$$d(q)=\frac{\partial\ \ln[\tau(q)]}{\partial q}=\frac{1}{\ln(k)}\frac{\sum_{i=0}^{N} L(i,k)^{q-1}\ \ln[L(i,k)]}{\sum_{i=0}^{N} L(i,k)^{q-1}}.$$

The evaluation of these summations for various q values yields q-dependent averages of distances that correspond to different local scaling exponents. Specifically, for $q\rightarrow\infty$ only long distances will contribute significantly to this summation, thus evaluating diffusing, correlated sequences corresponding to low local scaling exponents. In contrast, for $q\rightarrow-\infty$, only short distances will contribute to the evaluation, yielding slowly diffusing, anticorrelated sequences corresponding to high local scaling exponents. Using least-squares fits to obtain the scale-independent functions, this method also provides error estimates of all quantities.

For the experiment, human subjects sat in front of a computer equipped with a mouse. Instructions were shown on the computer screen and additional questions regarding the experiment were answered by the experimenter. The computer screen displayed a house in the center flanked on either side with a sidewalk and a ditch. Across the ditches were two roads. The following scenario was described. In the house are several people waiting to catch a ride with a car that is driving on the road on the left- or the right-hand side. Only two people can wait at a time, one on each side, but neither can cross the ditch between the sidewalk and the road. The subject's task was to push the left or the right button in order to build a temporary bridge for one of the people, who may then cross the ditch to reach the car. The experimenter indicated that the car could arrive on either side without a particular preference or order in appearance. A trial was defined by the response of the subject to the situation described above. The paradigm consisted of 500 presentations of the situation, i.e., 500 trials. Each trial was initiated without delay by the button push of the previous response. Initially, two schematic figures appeared on the screen until the subject responds. Thereafter, the car was shown briefly on the computer screen for 250 msec on the far right- or left-hand side as determined by a random number generator. At the end of the paradigm, the computer screen was cleared and the subject was informed that the procedure had been completed.

Twenty-two subjects were recruited from an undergraduate course at San Diego State University screened for psychological abnormalities and were asked to participate in a larger research project of which the choice task experiment was a small component. The mean age of the subjects was 27.4 ± 2.8 yr and included both males and females.

The autocorrelation function was calculated for each subject for $0\leqslant k\leqslant32$ and the autocorrelation functions were av-

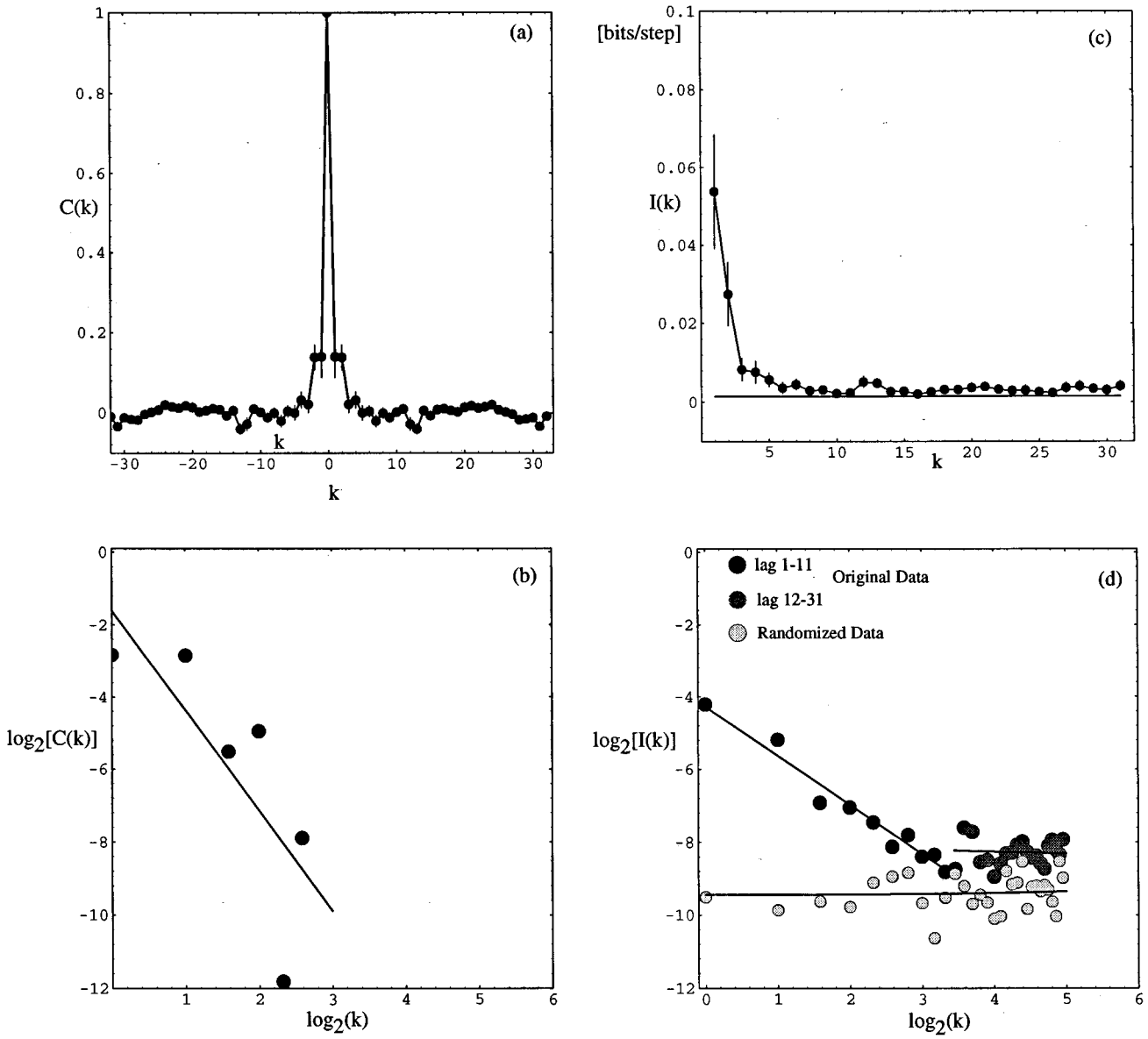

FIG. 1. (a) Average correlation function $C(k)$ and (b) average mutual information function $I(k)$ obtained by averaging the individual functions across all subjects. These functions are fitted logarithmically (for parameters see Table I) and are shown in (b) and (d), respectively. The line in (b) and (d) reflects the fitted function for $C(k)$ and $l(k)$, respectively.

eraged across the 22 subjects to obtain the group autocorrelation function [Fig. 1(a)]. The correlation between successive choices decays rapidly and approaches zero for a lag of 5. The doubly logarithmic fit of the autocorrelation function [Fig. 1(b)] indicates a power-law decay $C(k) \approx t^{-\gamma}$, with $\gamma = 2.76$. However, the regression coefficients exhibit large standard errors, indicating a poor fit (Table I).

Similar to the autocorrelation function, the mutual information function was computed for each subject for $0 \leqslant k < 32$ and the individual functions were averaged across all sub-

TABLE I. Parameters for the least-squares fit of the form Ak^B for the autocorrelation function $C(k)$ and the mutual information function $I(k)$. Two distinct regimes for $I(k)$ were observed ($k<12$ and $k \geqslant 12$) and fitted separately. Listed are the parameters (constant and exponent), the standard error (SE), the T statistic, the tail probability (p value), and the explained variance r^2.

	Logarithmic fit for autocorrelation and mutual information								
Parameters	A	SE	T statistic	p value	B	SE	T statistic	p value	r^2
$C(k)$	−1.61	2.06	−0.78	0.47	−2.76	1.14	−2.41	0.073	0.59
$I(k)$, $k<12$	−4.27	0.21	−20.04	0.00	−1.35	0.08	−15.89	0.00	0.97
$I(k)$, $k \geqslant 12$	−8.04	0.82	−9.71	0.00	−0.05	0.18	−0.27	0.79	0.00

jects to yield the group mutual information function [Fig. 1(c)]. The values of the group function are compared with the systematic overestimation according to Herzel and Grosse [13], given by $1/[2500-k\ln(2)]$, graphed in Fig. 1(d). The mutual information function exhibits a power-law decay $I(k)\approx t^{-\gamma}$, with $\gamma=1.348$. In this case, the fitted parameters are highly significant with small standard errors (Table I). Thus, compared to the autocorrelation function, the mutual information provides a more statistically robust measure of the information decay. Interestingly, the mutual information function exhibits a different behavior for $k\geqslant 12$. For these delays, the mutual information does not decay significantly as indicated by the fitting function (Table I). In addition, the group mutual information function exhibits consistently higher values than those obtained from the randomized data sets (light gray points) and those predicted by the systematic overestimation (line). In summary, the mutual information function for the subjects indicates two separate regimes: first, a significant decay of mutual information for consecutive choices with small lags ($k<12$), and second, a small but significant interdependence between choices that are separated by a larger lag ($k>12$). Therefore, even choices that are separated by many choices in between are not completely statistically independent.

The group fluctuation spectrum of local spatial scaling exponents $f(d)$ was obtained by averaging the fluctuation spectra of each subject. The scales considered varied from $k=1$ to 100 averaged over an ensemble size of $n=150$ consecutive choices. Each function was evaluated at a range of predetermined q values and averages for both d and $f(d)$ were obtained. Figure 2(a) shows $f(d)$ for 22 subjects (black points) and for the randomized data sets (gray points). The units on both axes correspond to ln[(length)]/ln[(scale)]. The t-score difference function is shown in Fig. 2(b). First, the average spatial scaling exponent $d=d(1)$ does not differ significantly between subjects ($d=1.4740\pm0.0262$) and the randomized data sets ($d=1.4690\pm0.0311$). Consequently, the average scaling behavior of choice sequences does not deviate from a random walk within the investigated region of scaling. Second, subjects exhibit a significantly increased contribution of highly anticorrelated choice sequence with high local spatial scaling exponents $d>1.66$ as indicated by $f_t(d)<-2.0$. Third, there is a slight and marginally significant increase of highly correlated choice sequences with $d<1.3$ with $f_t(d)<-2.0$. To summarize, subjects' choice sequences are characterized, on the average, by a random-walk pattern, but exhibit significant deviations of both highly correlated and anticorrelated choice subsequences.

Similar to the $f(d)$ function, the group fluctuation spectrum of local dynamical entropies $S(h)$ was obtained by q-dependent averaging of the fluctuation spectra for the individual subjects. Figure 3(a) shows the $S(h)$ for all subjects (black points) and for all randomized data sets (gray points). The units on both axes correspond to (bits)/(step). To statistically evaluate the difference between these two function, the t-score difference between the averaged fluctuation spectra is shown in Fig. 3(b). Similar to the $S(h)$ functions shown for the individual subject, the group functions reveal three distinct characteristics. First, the average or metric entropy for the subjects ($h_{\mathrm{met}}=0.7800\pm0.0178$) is significantly lower than h_{met} for the randomized data sets

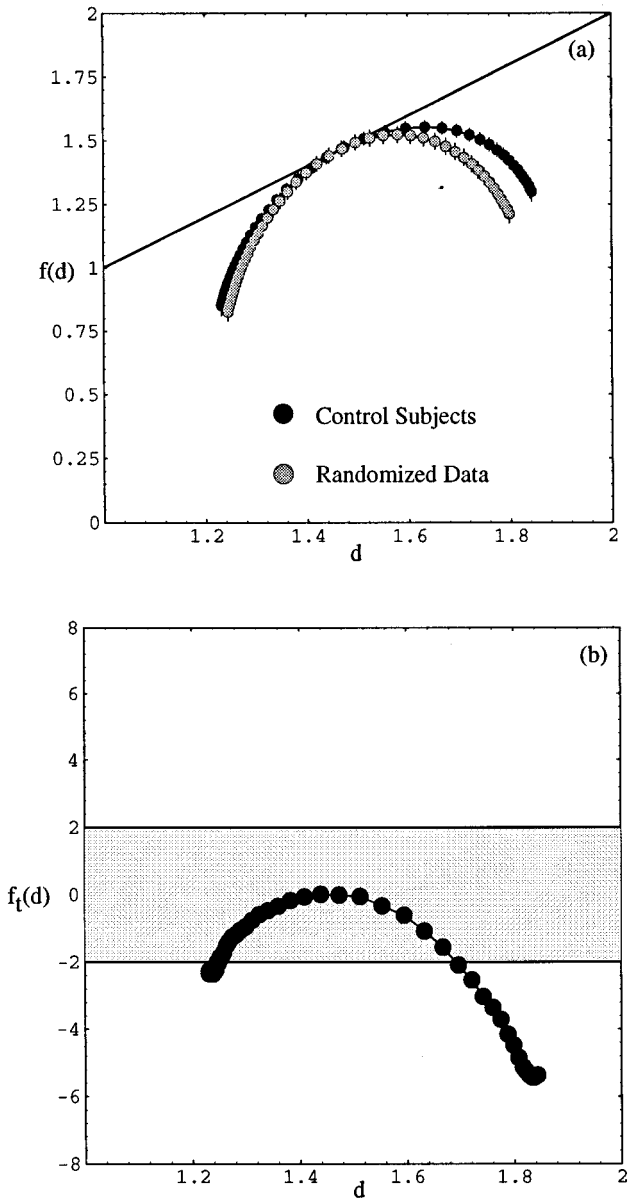

FIG. 2. (a) Average fluctuation spectrum of local spatial scaling exponents $f(d)$ and (b) t-score differences between original and randomized data $f_t(d)$. The gray area indicates insignificant differences between the original data and the randomized data set for different local spatial scaling exponents.

($\overline{h}_{\mathrm{met}}=0.8610\pm0.0061$). Second, the randomized data set is characterized by a lower topological entropy with a larger contribution of the topological entropy $[(h_{\mathrm{top}},S(h))=1.0210\pm0.0029,\ 0.9470\pm0.0018]$ as indicated by the higher maximum of the group fluctuation spectrum compared to the subject's maximum of the fluctuation spectrum. $[(h_{\mathrm{top}},S(h))=1.0670\pm0.0147,\ 0.9300\pm0.0031]$. Third, $S(h)$ for the subjects exhibits a significantly increased contribution of highly predictable subsequences and highly unpredictable subsequences as indicated by the t-score difference function. This function shows that the randomized data sets contain significantly fewer low ($h<0.75$) and high ($h>1.2$) dynamical entropy subsequences. To summarize,

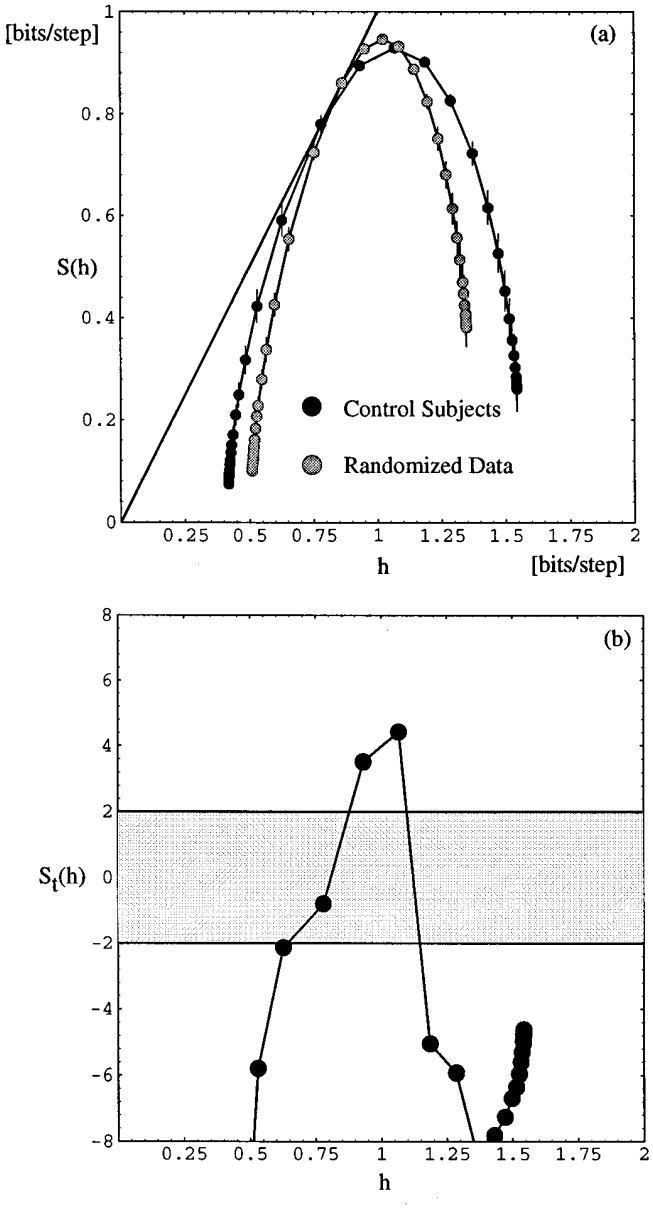

FIG. 3. (a) Average fluctuation spectrum of local dynamical entropies $S(h)$ and (b) t-score differences between original and randomized data $S_t(h)$. The gray area indicates insignificant differences between the original data and the randomized data set for different local dynamical entropies.

the group fluctuation spectrum of local dynamical entropies reveals a greater nonuniformity for the subjects relative to the randomized data sets.

There are three main results of this investigation. First, average measures of interdependence on a short scale ($1<k<10$) and a midrange scale ($5<k<22$) indicate that subsequent choices in a simple binary choice task are not random. Second, the subject's choice sequences are characterized by an increased variability in interdependence compared to the randomized data sets on all scales. Third, there are two scales ($k<12$ and $k>12$) with different characteristics of average information decay.

Conceptually, human choice behavior is located at the interface between external or internal constraints due to environmental or intrapsychic demands and observable actions of human beings. Various aspects of choice behavior have been extensively discussed in the psychological literature [29–34]. The results of these investigations suggest that decision making in the presence of uncertainty involves complex central processes that are sensitively influenced by a variety of experimental factors.

Here the mutual information function indicates that the interdependence between choices decays significantly within the adjacent ten choices. This would suggest that the average interdependence is best captured in statistical-mechanical models that incorporate explicitly these scales. However, the interdependence between choices that are separated by more than ten choices does not vanish. Instead, there is a significant long-range dependence between choices. The detailed assessment of the long-range behavior using the fluctuation spectrum of local spatial scaling exponents indicates that the subjects generate significantly more anticorrelated or alternating choice sequences than would be expected by chance. This finding had been anecdotally described previously for short sequences [35]. This investigation extends these findings for long sequences. Put simply, human subjects exhibit a ''balance'' between choice alternatives that extend across long ranges of choice sequences.

However, the subjects also generated an increased contribution of correlated or simple repetitive sequences when compared to the randomized data set. These findings had been described previously within the context of induction of sequence repetition [29]. Therefore, it appears that human choice sequences result from the balance of two competing principles, i.e., balancing choice alternatives and generating simple repetitive choice sequences. The competition between these two principles results in an increased variability of choice subsequences. This variability also extends to the dynamic description of the choice sequences. Specifically, both highly predictable and highly unpredictable choice subsequences contribute significantly more frequently to the average interdependence between choices than expected by chance. Within the entropic realm, there appears to be a competition between two dynamic principles: first, to generate novel, unpredictable behavioral sequences, and second, to repetitively use a familiar set of choice sequences. Thus the increased range of local dynamical entropies may result from the competition between uncertainty or novelty and familiarity or establishment of a characteristic set of rules generating the choice sequences.

In summary, the application of the fluctuation spectrum analysis to a simple behavioral paradigm is able to extract a few guiding principles for the generation of choice sequences. These principles can be used in a statistical-mechanical model to determine explicitly the influence of these principles on choice behaviors in analogy to the determination of order parameters in coordinated hand movements [22,23]. Moreover, the mutual information function can be used to determine the scale of a model for the obtained data sets based on the average interdependence. These findings further support the notion that complex behavior ranging from a molecular biological level [13,14,16] to a socioeconomic level [36] is based on long-range interactions between individual elements.

[1] R. M. May, Nature (London) **261**, 459 (1976).
[2] T. Elbert, W. J. Ray, Z. J. Kowalik, J. E. Skinner, K. E. Graf, and N. Birbaumer, Physiol. Rev. **74**, 1 (1994).
[3] J. A. C. Gallas, P. Grassberger, H. J. Hermann, and P. Ueberholz, Physica A **180**, 19 (1992).
[4] G. R. McNamara, Europhys. Lett. B **12**, 329 (1990).
[5] J. P. Crutchfield and J. E. Hanson, Physica D **69**, 279 (1993).
[6] Y. Braiman, J. F. Lindner, and W. L. Ditto, Nature **378**, 465 (1995).
[7] D. S. Omstein, Science **243**, 182 (1989).
[8] D. Ruelle, *Statistical Mechanics, Thermodynamic Formalism* (Addison-Wesley, Reading, MA, 1978).
[9] H. Fujisaka and M. Inoue, Phys. Rev. A **41**, 5302 (1990).
[10] Y. Oono, Prog. Theor. Phys., Supp. **99**, 165 (1989).
[11] D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, New York, 1987).
[12] W. Li, J. Stat. Phys. **60**, 823 (1990).
[13] H. P. Herzel and I. Grosse, Physica A **216**, 518 (1995).
[14] C.-K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin, F. Sciortino, M. Simons, and H. E. Stanley, Nature **356**, 168 (1992).
[15] R. F. Voss, Phys. Rev. Lett. **68**, 3805 (1992).
[16] C.-K. Peng, J. Mietus, J. M. Hausdorff, S. Havlin, H. E. Stanley, and A. L. Goldberger, Phys. Rev. Lett. **70**, 1343 (1993).
[17] J. Feder, *Fractals* (Plenum, New York, 1988).
[18] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, Phys. Rev. E **49**, 1685 (1994).
[19] R. Badii, Riv. Nuova Cimento **12** (3), 1 (1989).
[20] M. P. Paulus, M. A. Geyer, and D. L. Braff, Prog. Neuro-Psychopharm. Biol. Psychiatr. **18**, 1169 (1994).
[21] P. Ch. Ivanov, M. G. Rosenblum, C.-K. Peng, J. Mietus, S. Havlin, H. E. Stanley, and A. L. Goldberger, Nature **383**, 323 (1996).
[22] G. Schoner and J. A. Kelso, Science **239**, 1513 (1988).
[23] H. Haken, *Synergetics: An Introduction* (Springer, Berlin, 1983).
[24] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields* (Springer, New York, 1986).
[25] P. E. Rapp, A. M. Albano, I. D. Zimmerman, and M. A. Jimenez-Montano, Phys. Lett. A **192**, 27 (1994).
[26] P. Grassberger, IEEE Trans Inf. Theory **35**, 669 (1989).
[27] W. Feller, *An Introduction to Probability Theory and its Applications*, 3rd ed. (Wiley, New York, 1968).
[28] A. B. Chhabra, C. Meneveau, R. V. Jensen, and K. R. Sreenivasan, Phys. Rev. A **40**, 5284 (1989).
[29] C. D. Frith and D. J. Done, Psychol. Med. **13**, 779 (1983).
[30] S. Grossberg and W. E. Gutowski, Psychol. Rev. **94**, 300 (1987).
[31] M. J. Machina, Science **236**, 537 (1987).
[32] H. Rachlin, A. W. Logue, J. Gibbon, and M. Frankel, Psychol. Rev. **93**, 33 (1986).
[33] A. Tversky and D. Kahneman, Science **211**, 453 (1981).
[34] H. Rachlin, Psychol. Sci. **1**, 294 (1990).
[35] P. Bakan, Am. J. Psychol. **73**, 127 (1960).
[36] T. Bohr and D. Rand, Physica D **25**, 387 (1987)J.
[37] M. H. R. Stanley, L. A. N. Amaral, S. V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M. A. Salinger, and H. E. Stanley, Nature **379**, 804 (1996).