Response time dynamics: Evidence for linear and low-dimensional nonlinear structure in human choice sequences

Alice Kelly[*], Andrew Heathcote, Richard Heath & Mitchell Longstaff

Department of Psychology, The University of Newcastle,

Callaghan, NSW, Australia 2308

[*] Please address all correspondence to Alice Kelly, School of Behavioural Sciences,

Building W, The University of Newcastle, Callaghan, NSW, 2308, AUSTRALIA;

email: akelly@psychology.newcastle.edu.au

Abstract

Response time (RT) is a commonly used measure of cognitive performance that is usually characterised as stochastic. However, useful information may be hidden in the apparently random fluctuations of RT. Dynamical systems analysis techniques allow an exploration of the alternative hypothesis that RT fluctuations are deterministic, albeit in a complex manner. We applied careful task construction and noise reduction and surrogate series tests to show that RT series from a forced-pace serial response time task have low dimensional chaotic characteristics. In Experiment 1, 80% of subjects' filtered RT series had low dimensionality, sensitive dependence on initial conditions, spectra close to $1/f$, and stable attractor geometry across sessions. In Experiment 2, we showed that the size of the inter-stimulus interval (ISI) determined the number of subjects with low dimensional chaotic series. A small ISI caused 100% of subjects to respond in the chaotic regime, whereas only 25% had a low dimensional chaotic RT component when the ISI was large. We argue that demanding task requirements cause a reduction in the dimensionality of the dynamics, producing RT fluctuations that may reflect a response strategy for controlling RT.

In recent years, many authors have proposed dynamical accounts of cognition as a replacement for symbol-processor models (Garson, 1996; Kelso, 1995; Lintern & Kugler, 1991; Luce, 1995; van Gelder, 1997). Accounts based on dynamical systems theory (van Gelder, 1997), complexity (Coveney & Highfield, 1995), and associated mathematical techniques, examine the processes involved in ongoing couplings between complex systems, such as humans, and their environments (Eliasmith, 1996).

One of the most interesting states that dynamical systems exhibit is low-dimensional chaos, which has the properties of stability and dynamic flexibility with few degrees of freedom (van Leeuwen, Steyvers & Nooter, 1997). A low dimensional chaotic model seems suited to describing the behaviour of a complex system such as the brain; it can incorporate the dynamics of the interaction between the organism and its environment, emergent self-organisation, multi-causal determinism, and the flexible manner in which the brain switches between stable cognitive states (Barton, 1994; Clark, 1997; Kelso, 1995; van Leeuwen et al, 1997; van Gelder, 1997). Luce (1995) argues that low dimensional chaotic models have the important advantage of parsimony because a realistically complex output can be generated by simple deterministic equations without requiring a separate random component to represent fluctuations around the pure signal. Hence, there are good reasons to attempt an explanation of cognitive and brain data using nonlinear dynamic analysis techniques.

However, it has proven difficult to detect low dimensional chaos in experimental data from psychology, limiting the applicability of dynamical models in the study of cognition and behaviour (Garson, 1996; Rapp, 1994). In the present paper, we show that if special care is taken with task construction and analysis, clear evidence for chaotic dynamics can be found in human response time (RT) data, and

that dynamical analyses, both linear and nonlinear, yield interesting new information about cognitive function.

Traditional approaches to response time data

Statistical analyses cleave RT measurements into two independent components, signal and error. The signal is modelled by smooth deterministic functions such as the mean, whereas the error accounts for inter-trial performance fluctuations. Further, RT models are often fit to averages over subjects, rather than to the behaviour of each individual. When behaviour is linear, the average accurately reflects individual performance, but may not when behaviour is nonlinear (e.g., Heathcote, Brown, & Mewhort, in press). Consequently, statistical analysis and model fitting often ignore two sources of information, dynamical structure in the inter-trial fluctuations and individual differences.

There are two lines of evidence for dynamical structure in performance fluctuations: short-term linear dependencies and longer-term rhythms. Linear dependencies between sequential responses have been found in vigilance tasks (Hollenbeck, et al., 1995; Makeig & Inlow, 1993), serial response tasks (Brewer & Smith, 1989; Fletcher & Rabbitt, 1978; Laming, 1968; Rabbitt, 1979; 1989; Remington, 1971; Smith & Brewer, 1995), and threshold perception tasks (Wertheimer, 1953). For example, Laming (1968) reported that the identity of the previous two stimuli had a strong influence on subjects' RT in a two-choice RT task. Luce (1986) showed that the influence of previous trials on RT extends to five trials into the past on two-choice tasks. Rabbitt and Rodgers (1977) showed that subjects respond more slowly following errors to avoid future errors. Luce (1995) argues that decision-making models that ignore dynamics will be incomplete and incorrect.

The passage of time can also lead to systematic dynamics in performance fluctuations. The sensitivity of attention to circadian rhythms has been well documented (Davies & Parasuraman, 1982). Some research also implicates "ultradian" rhythms, such as minute-range periodicities in vigilance performance. Makeig and Inlow (1993) presented subjects with ten auditory targets (noise bursts against a white noise background) every minute, a much higher rate than is typically used in vigilance tasks. They found that error rates peaked significantly at four-minute periods. Wertheimer (1953) found similar four-minute spectral peaks in a Fourier analysis of missed signals in a visual target discrimination task. Broadbent (1954) found that subjects were more likely to fail to respond to a stimulus at four-minute intervals in a forced-pace 5-choice SRT task.

A dynamic approach may be especially informative when the smooth functions examined by statistical analysis techniques, such as the mean, do not account for much of the variability in data. This appears to be a common situation, at least for individual RT data. For example, Gilden (1997) found that, for a variety of simple decision-making tasks, changes in the mean only accounted for around 10% of the variability in RT. Given the magnitude of fluctuations around the mean function commonly observed, and the evidence that a variety of cognitive tasks exhibit sequential and temporal structure, RT fluctuations should be investigated as a potential source of information. Alternative dynamical techniques have the potential to reveal any structure in cognitive performance fluctuations in RT.

Dynamical analysis techniques

Dynamical analysis examines performance data for evidence of dependence between its states at different times. Dependence implies that future states are a function of past states. The function governing the dependence can be either linear or

4

nonlinear. Linear dynamics produce relatively simple and regular outputs, so linear dynamical models usually include an added noise component to give a realistically complex output. Nonlinear dynamical systems, on the other hand, can give a complex output from simple recursive equations. The number of variables in the recursive equation is roughly analogous to the system's dimensionality. A low dimensional system can produce a complex output due to ongoing nonlinear interactions between a small number of variables. This complex behaviour is usually described as low-dimensional chaos. High dimensional nonlinear processes, on the other hand, have a large number of variables that interact in a nonlinear manner, and are referred to as high dimensional or hyper-chaos. Although the equations governing complex, dynamical behaviour are often unknown (Eliasmith, 1996), time series of the systems behaviour can be used to reconstruct the model (Takens, 1981).

The outputs of high dimensional nonlinear and linear stochastic processes are, in practice, very similar, and we will not differentiate them here. Rather, we will attempt to differentiate low dimensional chaotic structure from high dimensional deterministic and/or stochastic structure. A finding for low dimensional chaotic structure opens up the possibility that complex RT fluctuations can be modelled by relatively simple recursive equations with only a few parameters, although we will not pursue such explanatory modelling here. In the following sections, we present a more detailed account of linear, and low dimensional nonlinear, dynamical system characteristics, and an analysis technique associated with each characteristic.

Linear Dynamical Processes

We will use two complementary approaches to measure linear dynamical structure, time series analysis and Fourier analysis. The "Box-Jenkins Method" (Box & Jenkins, 1976) for time series analysis identifies autoregressive (AR) and moving

average (MA) linear dependencies between sequences of observations. In an MA(q)

process, the value of a stationary series at time t (X_t) is the weighted average of the

last q values of a sequence of independent, identically distributed random variables

(a_t, a_{t-1}, … a_{t-q+1}). In an AR(p) process, the stationary value X_t is a linear combination

of p past values (X_{t-1}, X_{t-2}, … X_{t-p}) and an independent random variable a_t.

The type of linear dependence, AR, MA, or some combination of AR and MA

processes, and their order (the values of p and q), can be determined by examining

autocorrelation and partial autocorrelation coefficients. These coefficients are

calculated between pairs of values at different lags, where lag is the number of time

steps between the values. The resulting lag functions have characteristic patterns that

differentiate AR and MA processes. The order of these processes can be determined

with the aid of significance tests on the coefficients (Cryer, 1986). For example, an

MA(q) process will have no significant autocorrelations beyond a lag of q and partial

autocorrelations that may be signific ant beyond q lags, but that decay exponentially to

zero as lag increases. An AR(p) process has a complementary pattern, with no

significant partial autocorrelations beyond p lags, and autocorrelations that may be

significant beyond p lags but that decay exponentially to zero with increasing lag.

Time series modelling has been used to study dependencies in data, for example, to

characterised error sequence effects (Brewer & Smith, 1989; Pressing, 1998).

Fourier analysis decomposes dynamic structure as a linear superposition of

component frequencies with different powers and phases. Fluctuations are

characterised as differently "coloured", depending on the linear slope of a log-log plot

of spectral power against frequency. White noise has a flat spectrum (zero slope) with

equal power at all frequencies, because each successive value is independently

distributed. Coloured noise, on the other hand, is classified as "pink" or "*1/f*" if the

slope of a log-log plot of the power spectrum is approximately -1. *1/f* noise, which characterises self-organising systems with a large number of degrees of freedom (Bak, Tang & Wiesenfeld, 1987; Bak & Creutz, 1994), has been found in a variety of biological processes such as heart rate (Kaplan & Glass, 1995) and DNA sequences (Buldyrev, et al, 1994). *1/f* processes are fractal in time, and have multiple self-similar time scales. Brown noise, in contrast, has a slope of -2 or greater, and has a high level of positive correlations.

Brown and *1/f* noise have been used to model human performance. Random walk models of decision processes are characterised by brown noise (Heath, 2000). Gilden (1997) found that approximately 25% of the trial-to-trial RT variability in simple decision tasks could be identified as *1/f* noise. Gilden proposed a model that represents the *1/f* noise as a decision-making component. The remaining variance was assumed to be white noise derived from other cognitive processes, motor fluctuations and measurement error. Pressing and Jolley-Rogers (1997) found that, when subjects were asked to generate a regular beat without feedback, low frequency fluctuations in performance over long series of trials exhibited *1/f* structure. They argued that the source of the *1/f* noise might be a multiple time scale attentional process, the temporal structure of which is fractal.

Linear time series analysis and Fourier analyses cannot, however, discriminate different types of dynamics. For example, many very different processes have a $1/f$ signature, including low dimensional chaos, high-dimensional stochastic systems with nonlinear dynamics (Dooley & Van de Ven, 1999) and superimposed linearly autocorrelated stochastic processes (Pressing, 2000). Consequently, nonlinear dynamical analysis techniques are required to distinguish low dimensional nonlinear dynamics from stochastic and high dimensional nonlinear dynamics.

7

<u>Nonlinear dynamical processes</u>

The distinction between low-dimensional chaos generated by a "strange attractor" and a stochastic process is fundamental, but difficult to make in practice due to the similarly irregular and complex appearance of their outputs. However, there are three fundamental characteristics of strange attractors that distinguish them from stochastic processes: they exhibit complex but well-defined geometrical structure, low, fractal dimensionality, and a decay in predicability over time due to sensitive dependence on initial conditions. The presence of each of these characteristics in the data can be assessed using a variety of nonlinear dynamical analysis tec hniques (for a review of these techniques see Heath, 2000). Three such techniques were selected to evaluate one of the fundamental aspects of low dimensional chaos, and to provide converging evidence for low dimensional chaos in RT sequences. The particular techniques were selected because they have been thoroughly investigated, so potential problems can be avoided or dealt with appropriately. The nonlinear analysis techniques are discussed in more detail below and summarised in Table 1.

INSERT TABLE 1 ABOUT HERE

Phase Portraits – assessment of visual structure in the series

Two dimensional phase space portraits reconstruct aspects of an attractor's geometry by plotting each data point in an observed time series $x(t)$ at time t against an estimate of its derivative at time t, such as half the difference between successive data points (Sprott, 1995). Deterministic processes such as limit cycles and chaos have structured phase portraits because their attractors limit the areas of phase space that they can visit. Stochastic processes, on the other hand, can visit all areas of the phase space within the data limits. Casdagli (1991) and Mpitsos (1994) argue that the

best evidence for low dimensional chaos is a complex but structured phase portrait.

For example, cyclical processes have highly structured simple shapes like closed

loops. Noise and stochastic processes appear complex but unstructured, like an

undifferentiated mass of jagged lines. Low dimensional chaos has a structured but

complex shape in phase space. Hence, phase portraits provide a useful initial check

for observable low dimensional structure in an experimental series.

Dimensionality Estimation – differentiating low and high dimensional chaos

A multi-dimensional attractor can be reconstructed from a one-dimensional time

series using an embedding space (Takens, 1981). The state of the system at any given

time is represented by a point in the embedding space with coordinates specified by

the last m values of the time series. When the dimensionality of a chaotic process is

smaller than m, its attractor will fill only a subset of the available dimensions of the

embedding space, whereas a stochastic system, with effectively infinite degrees of

freedom, will fill all available dimensions.

Dimensionality estimation techniques take advantage of these properties by

examining how the distribution of points in the embedding space changes with the

number of embedding dimensions. We will use the D2 dimensionality estimate

developed by Grassberger and Procaccia (1983). For stochastic systems with no linear

or nonlinear structure, D2 values increase as embedding dimension increases. For

chaotic systems, on the other hand, D2 values approach a finite asymptote. The

asymptote estimates the dimensionality of the system.

Nonlinear Prediction Analysis – detecting sensitive dependence on initial conditions

Possibly the best quantitative evidence for low dimensional chaos in

experimental series is an increase in prediction error over time, due to the primary

definitive characteristic of low dimensional chaos, sensitive dependence on initial conditions (Sugihara & May, 1990). Unlike other low dimensional processes, such as periods or fixed-states, the chaotic attractor is "fractal". This means that the attractor expands in a small number of directions at a local level while maintaining global stability. Sensitive dependence magnifies small prediction errors, so that the trajectory can only be predicted accurately in the short-term, despite the fact that it is deterministic. Prediction error thus increases as lag (time into the future) increases. Series composed entirely of noise are unpredictable on any time scale, so prediction error does not change with lag. High dimensional nonlinear processes also have high prediction error, due to extreme sensitive dependence. Extreme sensitive dependence means that future states are predictable only over very short time scales. Consequently, at longer time scales, high dimensional nonlinear processes appear stochastic. We can thus establish the presence low dimensional chaos through prediction errors that increase gradually as a function of prediction lag.

Problems associated with detecting low dimensional chaos

Tests for low dimensional chaos require long data series and are highly sensitive to confounds such as nonstationarity, linear autocorrelations in the data, noise, and artificial structure resulting from filtering the noise (Broomhead, Huke & Muldoon, 1992: Mayer-Kress, 1994; Rapp, 1994). Data collected in behavioural research is particularly sensitive to these confounds. Both adaptations to the task and fatigue can create nonstationarities, especially if the behaviour is sampled for an extended period of time. Linear autocorrelations due to sequence effects and noise due to measurement error and the effects of uncontrolled influential variables characterise cognitive data (Luce, 1986; 1995), and filtering must be used to attenuate the noise component to

detect chaos. However, both filtering and linear autocorrelations can then introduce dynamic structure that can be mistaken for chaos.

Because of these problems, in practice it is usually not possible to "prove" that a series is chaotic. We can, however, establish strong converging evidence. Using the multiple tests for chaos outlined above, we can provide both qualitative (phase plots) and quantitative (dimensionality estimation and nonlinear prediction) indices, which support the defining characteristics of low dimensional chaos (complex, but structured, attractor geometry, dimensionality and sensitive dependence on initial conditions). In addition, surrogate series comparisons (Theiler, Eubank, Longtin, Galdrikian & Farmer, 1992) can be used to discount other forms of structure, such as linear autocorrelation or filtering effects, which are known to confound these tests. Surrogate series are constructed from experimental series in a way that removes chaotic structure but maintains the confounding structure. If the surrogate acts like the experimental series in a test for chaos, the test is confounded. Alternately, if the test does not detect chaos in the surrogate, but does detect it in the experimental series, confounding can be discounted. Where confounding can be discounted and all three nonlinear dynamical measures indicate chaos, we can safely conclude that we have strong evidence for chaos in an experimental series.

Experiment One

We used RT as the dependent variable in the following studies. RT appears to be a good candidate for a dynamical analysis for a number of reasons. Firstly, RT is a non-invasive measure, and thus long series can be obtained under naturalistic conditions. Secondly, RT is known to be a rich source of information (Laming, 1968), and its statistical properties are well known (Luce, 1986). Finally, RT has been shown

to reflect complex, dynamical structure in other studies (e.g. Gilden, 1997), and thus may provide a good measurement of any cognitive dynamics.

The aim of Experiment One was to attempt to establish whether or not RT series contain nonlinear dynamical structure. To achieve this aim, we drew from past research to construct a task that is both amenable to nonlinear dynamical analysis and, as far as is possible, minimally subject to other confounding sources of structure such as nonstationarities. Specifically, we used a four choice serial response time (SRT) task with a compatible stimuli-response design. This design allowed us to collect long RT series, because the task was not too demanding for the subjects. The use of a randomised stimulus sequence had the dual effect of preventing learning, a source of nonstationarity, and maintaining attention to the task. Subjects were also given breaks every four minutes to minimise the effects of fatigue. Trial-onsets occurred at regular time intervals, so that we could sample the dynamics in real time, because subject-controlled pacing or randomising the stimulus-onset time may blur any (real-time) dynamics. Finally, a short inter-stimulus interval (ISI) was used to ensure that subjects were continuously engaged in the task and unable to take micro-rests between responses, as micro-rests may disrupt the RT dynamics.

Method

Subjects

Ten students, five female and five male, aged between 19 and 26 ($\underline{M} = 23$), from the University of Newcastle, Australia, provided informed consent to participate in the experiment. Subjects were paid $10 per session for their participation.

Apparatus

The stimulus-response apparatus consisted of a Psychology Software Tools Serial Response Box (Rogers, Schnieder, Pitcher & Zuccolotto, 1995) with a row of

12

five lights located 9mm apart, and a row of response keys located below each light. Four lights and response keys (excluding the middle pair) were used for stimulus presentation and response collection. The stimulus-response mapping was compatible, requiring subjects to press the response key positioned immediately below the active light. The Serial Response Box was located in front of a monitor attached to an IBM compatible computer running DOS. The apparatus measured RT to millisecond accuracy (Rogers, et al., 1995).

Procedure

Instructions to subjects were given on a sequence of computer screens, which the subject scrolled through at their own pace. The first screen instructed subjects to respond as quickly and as accurately as possible by pressing the key underneath the appropriate stimulus light. Subjects were also informed that they were expected to respond to stimuli about once every second. A second screen asked subjects to keep their response-fingers (the index and middle fingers of both hands) on the response keys for the duration of each block. Subjects completed three 1-hour sessions, each consisting of 11 blocks of 240 trials. During each block of trials, one of the four lights was turned on every second. Each light was turned on an equal number of times in each block, and the stimulus sequence was randomised without replacement.

The dependent variable, RT, was defined as the latency between the onset of the light and the full depression of a response button. On each trial, the stimulus remained illuminated until either a response was recorded or a maximum delay of 800 ms had expired, whichever occurred first. The time allowed for a response on each trial (800 ms) was determined from subjects' average RT in a small pilot study using a self-paced task. The stimulus for the next trial occurred 1000 ms after the onset of the previous stimulus regardless of if, or when, subjects responded on the previous

trial. Subjects took a break after every block for as long as they wished so that they did not feel fatigued, and were not making too many errors. Feedback to the subject was provided by two distinct computer-generated tones, one following incorrect responses and the other following missed signals.

<div align="center">Results</div>

A within-subjects ANOVA was conducted to compare mean correct and incorrect RT across the three experimental sessions. A decrease in mean correct RT from session one ($\underline{M}$ = 403.5 ms, $\underline{SD}$ = 27.7) to session three ($\underline{M}$ = 371.5 ms, $\underline{SD}$ = 34.0) just failed to reach significance, $\underline{F}(2,18)$ = 2.98, $\underline{p}$ = .08. The almost identical means in sessions two ($\underline{M}$ = 371.6 ms, $\underline{SD}$ = 26.0) and three ($\underline{M}$ = 371.5 ms, $\underline{SD}$ = 34.0) suggested that any improvement in performance due to subjects learning general task demands occurred in the first experimental session. There was, however, a significant difference in mean correct RT between sessions one and two, $\underline{t}(9)$ = 5.63, $\underline{p}$ < .001.

Overall, mean RT for incorrect responses ($\underline{M}$ = 322.5 ms, $\underline{SD}$ = 50.1) was faster than mean RT for correct responses, $\underline{F}(1,18)$ = 29.49, $\underline{p}$ < .0001. All subjects, except one subject in one session, produced less than 5% errors (missed and incorrect responses) in all sessions, with an average rate of 2.2% . Subject 5 committed 10% errors in her final session. A one-way repeated-measures ANOVA found no evidence for differences in mean percent errors between sessions, $\underline{F}$ < 1.

Linear dynamical analysis

The first block of 240 responses was deleted from each session to remove nonstationarity, due to "warm-up" effects, resulting in three series of 2400 responses for each subject. The series were concatenated across blocks within each session to provide sufficiently long data series. Longstaff and Heath (1999) argue that

concatenation is not a problem if the component series are sampled from the same attractor. Because dynamical analysis requires data to be sampled in real time, the erroneous responses were not deleted from the series.

Time Series Analysis

Autocorrelation and partial autocorrelation coefficients were calculated for all experimental series. The results were consistent with an AR(1) model for most subjects. The first partial autocorrelation coefficient was significant, with 95% confidence, for 22 of 30 series. Four series had two significant partial autocorrelation coefficients, and 4 had none. The size of the partial autocorrelation coefficient ($\underline{M}$ = 0.29; $\underline{SD}$ = 0.14) was not a function of session, $\underline{F}$ < 1, nor was the number of significant coefficients, $\underline{p}$ > .05. Autocorrelations were significant at many more lags than partial autocorrelations, the autocorrelation decreasing gradually with lag. A trend for the number of significant autocorrelations to decrease from 10.4 to 5.2 across experimental sessions one to three did not approach significance, $\underline{F}$ < 1.

Fourier Analysis

Thirty RT series, three per subject, were analysed separately. Fourier spectra were calculated for each series by the smoothed periodogram method with a modified Daniell smoother (span 3 to 5) with the mean removed. Five percent of each end of the series was tapered with a split cosine function (Bloomfield, 1976). A one-way ANOVA found that log power against log frequency slopes were very similar, and did not vary statistically, across sessions, $\underline{F}$ < 1. Differences between subjects dominated the slope data, explaining 82.5% of the variance. As can be seen from Table 2, all the spectral density slopes were negative, ranging from -0.097 to -0.765. Almost half the series had a spectral density slope in the 1/*f* range between -0.5 and -1.5, and the

15

average slope for all series was -0.44 ($\underline{SD}$ = 0.183). The slopes were not significantly different from -0.5, $\underline{t}(29) = 1.98$, $\underline{p} = 0.06$. These findings suggest that the RT series had a spectrum at the lower end of the 1/f range.

INSERT TABLE 2 ABOUT HERE

Nonlinear dynamical analysis

Series from the first session were excluded from the nonlinear dynamical analyses because of the significant decrease in mean RT between sessions one and two. Hence, each of the ten subjects contributed two experimental series to the analyses, with data from each session analysed separately. Stationarity within a session was checked with a t-test and Levene's test for homogeneity of variance comparing the mean and variance of the first and second half of each session. The results for all series were nonsignificant on both stationarity tests. In all cases, the split-half mean difference contributed less than 5% of the total RT variance.

Noise Reduction

The RT series were very noisy (see Figure 1a for an example of a RT series). We applied a noise reduction filter, local singular value decomposition, which employs local linear maps to approximate the underlying signal (Grassberger, Hegger, Kantz, Schaffrath & Schreiber, 1993; Kantz & Schreiber, 1997). The technique is analogous to principal component analyses carried out on local neighbourhoods of data points. Eigenvectors with smaller eigenvalues are expected to be primarily due to noise, and so are discarded. Eigenvectors with the largest eigenvectors, which are expected to be composed primarily of the signal, are used to generate the filtered series. Appendix A gives a detailed description of the filtering procedure. Figure 1b shows the filtered version of the series in Figure 1a.

The relative standard deviation, calculated as a ratio of the original and the filtered series standard deviations ($\underline{SD}_{filtered}/\underline{SD}_{original}$), provides a measure of the variability of the filtered series relative to the original series. We assessed the relative standard deviation for the series using dimension manifolds of one, two and three in the filtering process. The resulting series appeared equally structured with a dimension manifold of one or two. We used a dimension manifold of one for all series. The mean relative standard deviations for the experimental series ranged from 0.37 to 0.55, indicating that the structured component comprised, on average, 14% to 30% of the variance[1]. There was no difference between the relative standard deviations between sessions, $\underline{F} < 1$.

INSERT FIGURE 1 ABOUT HERE

Surrogate Series

We generated two surrogate series for each experimental series in order to test two independent null hypotheses. The first null hypothesis was that any low dimensional, chaotic structure detected in the experimental series resulted from the filtering process. The corresponding filtered sequence shuffled (FSS) surrogate was constructed in two parts. First, the observations in the experimental series were randomly reordered to disrupt any dynamical structure. Second, the reordered series was filtered to the same level as the experimental series by matching the standard deviations of the filtered experimental series with the FSS filtered series. Matching was necessary because the filter assumes that a low dimensional attractor generated the series. If there is an attractor, the standard deviation stabilises under repeated filtering. If this is not the case, the filter explodes "holes" in the data series (areas where no data visits), creating structure from statistical inconsistencies, and repeated

17

filtering reduces the standard deviation to an arbitrarily small value. The number of filtering iterations necessary for the FSS surrogates was generally two or three, around half of the five iterations required for the experimental series. Thus, the FSS surrogate series allowed us to compare the experimental series with a filtered random series that has the same statistical properties and a matched degree of filtering.

The second surrogate series, the Amplitude Adjusted Fourier Transform (AAFT) surrogate series, allowed us to test the null hypothesis that any evidence for low dimensional chaos is caused by linear dynamics (autocorrelations). The AAFT surrogate was developed by Hegger et al (1999) and is available through TISEAN 2.0 software[2]. To create the AAFT surrogates, the original series was rescaled to a Gaussian distribution, and then, following Fourier analysis, the phase was randomised. The series was then reconstructed using the inverse Fourier transformation. To prevent biases in the rescaling towards a Gaussian distribution, the surrogates were rescaled to the actual values and amplitudes of the original series. This surrogate provided an excellent comparison because it retains all the statistical and linear dynamic information including the mean, standard deviation and the linear dependencies. Thus, it disrupts only the nonlinear dynamics. It should be noted that the AAFT surrogates were created from the filtered experimental series so that all other structure, linear and statistical, in the filtered experimental series, was completely retained in the AAFT surrogate.

Phase Portraits

Eight of the ten subjects had phase portraits that were both complex and structured. Six of these eight subjects had complex shapes that were consistent across experimental sessions two and three. Figure 2 shows the phase portraits for five

18

subjects chosen because they provide the widest range of shapes. For example, the session two and three attractors for Subject One's series had a three-lobe structure. Each portrait for subject ten has an almost identical overall shape, down to the darker rings in the lower region. The phase portraits for Subjects Three and Seven exhibit very stable "Figure 8" shapes. Although the shape of Subject Two's portraits appear less defined than the portraits for the other subjects, the consistent structure is still maintained across the experimental sessions. The geometric structure shown in the phase portraits in Figure 2 is surprisingly consistent considering that the experimental sessions took place several days apart. Indeed, this consistency was also found in the initial practice session for many subjects.

INSERT FIGURE 2 ABOUT HERE

The phase portraits for Subjects Four and Six appeared to be relatively unstructured. An example phase portrait from one of these subjects' series is given in Figure 3a. The portraits of these two subjects were almost identical in all sessions. Although complex, the portraits have less defined shapes than do the portraits for the other subjects in Figure 2. Figure 3 also shows example phase portraits for the FSS and AAFT surrogate series. The phase portraits for all FSS surrogates were very similar, appearing complex but without defined shape. The straight lines in the upper region of the attractor reflect the effects of filtering, and represent outliers in the RT distribution. The portraits for Subjects Four and Six (eg. Figure 3a) were very similar to the portraits of the FSS surrogate (eg. Figure 3b), suggesting that the series for Subjects Four and Six are either filtered noise or high dimensional processes. The phase portraits for AAFT surrogates differed depending on the linear autocorrelation structure of the original series.

INSERT FIGURE 3 ABOUT HERE

Dimensionality Estimation

We used the Chaos Data Analyser package (Sprott, 1995) to calculate the

Grassberger and Procaccia (1983) Correlation Dimensionality Estimate ($\underline{D2}$)[3], which

measures the number of dimensions containing each series' attractor. A detailed

description of the procedure for calculating $\underline{D2}$ is given in Appendix A. $\underline{D2}$ estimates

for the experimental series were compared with the $\underline{D2}$ estimates for the AAFT and

the FSS surrogate series. As summarised in Table 1, if apparent low dimensional

structure in the experimental series was due to linear dependencies, the $\underline{D2}$ values for

the AAFT surrogates should be equivalent to those for the experimental series. On the

other hand, if the low dimensional structure was due to chaos, the $\underline{D2}$ estimates for the

experimental series should asymptote at a finite $\underline{D2}$, while those for the *AAFT*

surrogates increase with embedding dimension. Similarly, if apparent low

dimensional structure was due to filtering, the $\underline{D2}$ values should be equivalent for the

FSS series and the experimental series. If, alternatively, the low dimensional structure

was a characteristic of the experimental series, the D2 values for the experimental

series should asymptote at a finite dimensionality, whereas the FSS surrogate series

$\underline{D2}$ values should diverge from the experimental series and increase with the number

of embedding dimensions.

Figure 4(a) shows mean $\underline{D2}$ values for the 20 experimental series and their

corresponding AAFT and FSS surrogate series. The mean $\underline{D2}$ values are plotted using

one to ten embedding dimensions. The experimental series asymptoted at a finite

number of dimensions (around 2.8) by 10 embedding dimensions. The mean $\underline{D2}$

20

values for the AAFT surrogate series, and for the FSS surrogate series increased with

embedding dimension, and did not plateau within 10 embedding dimensions.

INSERT FIGURE 4 ABOUT HERE

$\underline{D2}$ values at 10 embedding dimensions are given in Table 3. A paired-sample

t-test showed there was no significant difference between $\underline{D2}$ values for the two

experimental sessions, $\underline{p} > .05$. A one-way ANOVA comparing the mean $\underline{D2}$ values at

10 embedding dimensions for the three types of series was significant, $\underline{F}(2,57) =$

75.96, $\underline{p} < .001$. The mean $\underline{D2}$ for the experimental series was significantly less than

the mean for the AAFT surrogates, $\underline{t}(19) = 8.96$, $\underline{p} < .001$, and the FSS surrogates,

$\underline{t}(19) = 19.96$, $\underline{p} < .001$. Hence, the low dimensionality reflected in the $\underline{D2}$ values for

the experimental series was not due to linear dynamic structure or to filtering the data.

INSERT TABLE 3 ABOUT HERE

Nonlinear Prediction Analysis

We used a prediction algorithm derived by Kantz and Schreiber (1997), and

made available through their TISEAN 2.0 Predict program. Appendix A describes

both the algorithm in detail and the parameter values chosen for this analysis. Percent

prediction error was defined as the absolute prediction error divided by the series

standard deviation. We analysed 20 experimental series, two from each of the ten

subjects, together with the corresponding FSS and AAFT surrogate series. Although

prediction error for linearly autocorrelated noise can increase over time, this increase

should not be as steep as that exhibited by a low dimensional chaotic series (Sugihara

& May, 1990). Thus the AAFT surrogate series, which retained all the linear dynamic

structure, provided a good comparison for the nonlinear prediction test. We also

compared prediction accuracy with the FSS surrogate series to check whether filtering introduced spurious sensitive dependence.

The results of the nonlinear prediction analysis are given in Figure 4(b). There was no effect of session number on the prediction error, $\underline{F} < 1$. The experimental series are more predictable than the AAFT series at all lags, and take twice as long as the AAFT series to approach 100% prediction, indicating that nonlinear dependencies are stronger, and have a larger span, than linear dependencies. The FFT surrogate acts like white noise at all lags, indicating that filtering did not introduce any sequential dependence.

Individual Differences

In general, the results reported so far support the hypothesis that the RT series have a low dimensional chaotic component. However, the phase portraits for subjects four and six appeared almost identical to the FSS surrogates, suggesting that their series are true noise that has been filtered. Given these findings, it would be expected that series for subjects four and six would also behave like the FSS surrogate series for the other analysis methods, with higher $\underline{D2}$ values and higher prediction errors. As can be seen in Table 3, the $\underline{D2}$ values for these subjects lie at the lower end of the $\underline{D2}$ range for all subjects, the opposite to what would be expected. However, if there is no attractor in the series, filtering continues to reduce the series dimensionality to a level that is determined by the parameter settings chosen for the filtering algorithm. Given that the filtering process was iterated five times on all experimental series, lower $\underline{D2}$ values for the series without attractors compared to series with attractors would indeed be predicted.

Overfiltering can be checked by the relative standard deviation. If there is no attractor in the series the standard deviation should decrease with successive iterations, and we should find that the relative standard deviations for subjects four and six are lower than the relative standard deviations for the remaining eight subjects. The average relative standard deviation for the three experimental series for subjects four and six were 0.270 and 0.313 respectively. The mean relative standard deviation for the remaining subjects was 0.386 ($\underline{SD} = 0.073$). Although these subjects' relative standard deviations were within two standard deviations of the mean, they were still comparatively low. When we re-filtered these series to equate their relative standard deviation with the mean for the other experimental series, we found that the mean $\underline{D2}$ estimates were indeed much higher (4.7 and 4.9 at 10 embedding dimensions for subjects four and six, respectively).

A large nonlinear prediction error supports the argument that a series is true noise. At lag one, the average nonlinear prediction errors for subjects four and six were 81.7%, and 48.5% respectively, more than five standard deviations higher than the mean prediction error for the remaining series ($\underline{M} = 16\%$, $\underline{SD} = 5.6\%$). This finding suggests that there was little linear or nonlinear structure available to the prediction algorithm to provide accurate predictions. Taken together, these findings suggest that the series for subjects four and six were either high dimensional or noise.

<div align="center">Discussion</div>

We have shown that RT from a SRT task exhibits not only linear dependencies but also nonlinear determinism, with a component of the RT fluctuations behaving like a low dimensional chaotic attractor for 80% of subjects. The component exhibited a stable, complex, and structured attractor shape in phase space, low-dimensionality and sensitive dependence on initial conditions. Surrogate

23

series comparisons showed that the evidence for low dimensional chaos was not due to either the effect of filtering or linear dependencies. Thus, the converging evidence from three tests for low dimensional chaos confirms our premise that trial-by-trial fluctuations in RT contain structure with low dimensional chaotic characteristics.

There were, however, several limitations with the methodology and results. Firstly, the dimensionality estimates are partly determined by the amount of filtering, and thus not being a true estimate of the dimensionality, should only be used comparatively. The nonlinear prediction error is more robust, and appears to depend more on the characteristics of the series than the effects of filtering. Thus, the nonlinear prediction error may be a more valid way of discriminating between series.

Secondly, the structured component of the RT comprised, on average, between 14% and 30% of the RT variance, as measured by the ratio of the filtered and original series standard deviations. This amount is relatively small, and suggests that the RT fluctuations contained predominantly noise. Gilden (1997) found a similar sized $1/f$ component in RT, around 25%, on a number of different decision tasks, while changes in the mean as a function of condition accounted for only 10% of the variability. Thus, even though we failed to account for all the variability in the series, a substantial component does behave like low dimensional chaos.

The relative standard deviation does not provide an accurate account of the absolute size of the structured component. Instead, its value appears more dependent on the parameters used in the filtering process. These parameters were chosen to facilitate reliable measurement of the low dimensional chaotic component. Inevitably, such filtering also removed some of the chaotic component as well as noise. Despite this caveat, it appears undeniable that our data did contain a substantial random

component. So, although nonlinear dynamical analysis techniques can potentially use all of the variance information, this goal was not achieved here.

Nonlinear dynamical analyses conducted on the RT data prior to filtering showed no evidence for low dimensional chaos. Hence, filtering was a necessary prerequisite for the detection of low dimensional chaos. As discussed in Appendix A, it is difficult to know the right filter settings without trial and error. However, the single set of parameter values reported there appeared to work well for most subjects. It should be noted that, although we tried a number of different settings, an attractor in phase space was never detected in the series for subjects four and six.

In addition to nonlinear dynamics, significant partial autocorrelations provide evidence for linear dynamics in the data. A possible source of the linear structure is repetitions in the stimulus sequence. Repetitions have been found to reliably cause up to fifth order autocorrelations in data from multiple-choice tasks (Luce, 1986). Although we accounted for the confounding effects of linear dynamics on tests for chaos with the AAFT surrogates, it would be informative to show that the nonlinear dynamical component occurs when the linear dynamics are minimised. This might not be achievable, however, because the linear and nonlinear dynamics are often dependent. For example, Nicolis (1991) showed that an AR(5) process could model a sample from the logistic nonlinear recursion, which is a low dimensional chaotic system. We will explore the dependence between the linear and nonlinear structure further in Experiment 2.

Our treatment of the data prior to analysis was somewhat different from the usual practice of removing outliers and incorrect responses. We did not remove any data points, so as to retain evenly spaced samples in time, an assumption of most dynamical analyses. To check the effect of this difference, we repeated the dynamical

25

analysis on series composed of only correct responses. We found that the attractor geometry of the intact and correct-only series differed substantially for only one of the thirty experimental series, series three from subject five, which had a high error rate of 10%. The correct only series also had slightly higher dimensionalities and slightly flatter spectral density slopes than the intact series. The small changes found were likely due to noise from discontinuities introduced by missing values. If erroneous responses were sampled from a different attractor than correct responses, their removal should have caused a decrease in dimensionality, rather than an increase.

The trend towards increased randomness when erroneous responses were removed was much more pronounced for subject five's session three series. This series behaved chaotically on all measures when intact. However, the correct-only series appeared random. This finding suggests that a large proportion of missing values in a series blurs the low dimensional dynamics to such an extent that they cannot be reconstructed. However, our analyses were robust against the effects of the small proportions of errors produced by most subjects in the SRT task. We conclude, therefore, that dynamical analysis should proceed on intact series, with care being taken to minimise the number of errors.

A puzzling question remains as to why eight subjects generated series with a low dimensional chaotic component in their RT fluctuations, while two subjects generated random fluctuations. In Experiment 1, subjects were forced, on average, to respond more quickly than they would have done under self-paced conditions, as the ISI was set at the mean RT from a self-paced pilot task. However, the ISI was invariant (1000 ms) over subjects, so the task may not have been equally demanding for all subjects. It is possible that RT series reflect a response strategy that only enters a low dimensional chaotic regime when the task is demanding. The idea that a brain

or cognitive system focuses resources by shedding degrees of freedom under demanding conditions, and thus moves from a high to low dimensional state, was proposed by Kauffman (1993), Kelso (1995) and Large and Jones (1999). For example, Kelso (1995) argued that a meaningful task is required before the brain can focus attentional resources, and focusing is achieved by reducing the total degrees of freedom to the most salient features of the task.

In Experiment 2 we manipulate ISI under forced-pace conditions to test the hypothesis that forcing subjects to respond rapidly generates a low dimensional chaotic RT series. We hypothesise that RT series from a fast forced-pace condition will exhibit the sort of chaotic behaviour observed in most subjects' series in Experiment 1. For a slower forced-pace condition, we expect less evidence for chaotic structure. To control the effects of individual differences, the ISI used in the slow and fast forced-pace conditions was set individually for each subject.

Experiment 2

In Experiment 2, each subject participated in one self-paced condition and two forced-pace conditions. The ISIs in the forced-pace conditions were set individually for each subject at values calculated from their performance in the self-paced condition. In the forced-pace fast (FPF) condition, the ISI was set to the average of the subject's RT under self-paced (SP) conditions. In the forced-pace slow (FPS) condition, the ISI was set at two standard deviations above the subject's average RT in the SP condition, giving subjects approximately the same amount of time to respond as they took under self-paced conditions. If fast pacing causes low dimensional chaos in RT series, the series for all subjects in the FPF conditions

27

should exhibit chaotic dynamics, whereas chaotic dynamics will be less likely in the FPS condition because subjects are not speeded beyond their natural responding rate.

Most research in psychology uses subject- or self- paced tasks, rather than forced- or fixed- pacing. Given that dynamics emerge in real time, it is possible that the forced-pacing used in our SRT task, which allows the underlying system to be sampled at regular intervals, is required for the dynamics to be measurable. Self-pacing means that the system is sampled at irregular intervals in real time, violating the assumptions of most dynamical analyses. Consequently, the use of self-paced tasks could explain why other researchers (e.g., Gilden, 1997) have failed to find evidence for low dimensional chaos in RT. We will compare the dynamics of self-paced and forced-pace tasks in Experiment 2 to examine this question. However, it should be noted that, because the SP condition was always run first in order to calibrate the primary focus of the experiment, the effect of ISI, comparison of self-paced and forced-pace performance is confounded with order of administration.

Several changes were made to the methods in Experiment 2 to address some of the methodological limitations of Experiment 1. The CPU timer updated every millisecond in Experiment 1, and the buttons used for responding on the Psychology Software Tools Button Box had a lot of travel and substantial depression resistance. Both of these factors may have affected the measurement of RT dynamics. In Experiment 2, we improved RT timing accuracy to 0.1 ms by using a dedicated IO card. Increased precision in the measurement of RT may be important for detecting the underlying chaotic dynamics because low dimensional chaos is highly sensitive to measurement error. Secondly, modified mouse buttons, with virtually no travel and little resistance, were used as more sensitive response keys.

The linear dynamical component in the RT from Experiment 1 was substantial, as shown by the significant first partial autocorrelation coefficients and the predicability of the AAFT surrogates. Stimulus repetitions are the predominant cause of linear dependencies in RT from multiple-choice tasks (Luce, 1986). To assess whether the nonlinear dynamics occur when linear dynamics are minimised, we removed repetitions from the stimulus sequences in Experiment 2. The sensitive response buttons used in Experiment 2 should also reduce dependencies in motor components of RT. Dependencies between successive responses can be caused by mechanical coupling between fingers, especially when the response required has a high amplitude and force.

<div align="center">Method</div>

Subjects

Twelve subjects, six males and six females, between the ages of 22 and 27 inclusive ($\underline{M} = 24$) participated in this study. Each subject was paid $10 for each one-hour experimental session. The subjects were volunteers from the University of Newcastle, Australia, who responded to an advertisement. Subjects provided informed consent.

Apparatus

A program written in Turbo Pascal 6.0 collected each subject's demographics and data, and presented both stimuli and instructions. The experiment was run on a Pentium computer running the DOS operating system to ensure proper timing control. The two outside buttons of two modified three-button electronic mouses were used as response keys. The computer's monitor presented the stimuli, and the computer's keyboard was used to initiate sessions and blocks of trials. Stimulus onset was synchronised with the video refresh signal (Heathcote, 1988).

29

Procedure

      The procedure was the same as that reported in Experiment 1 unless otherwise stated. Each session consisted of thirteen blocks of 240 trials. On each trial, one of four stimuli was illuminated. Stimulus locations, which remained on the computer screen throughout the block, were marked by four 27 mm by 54 mm empty grey rectangles with 5 mm borders on a black background. The stimuli were white rectangles that filled a stimulus location on each trial. The subject responded by pressing the mouse-key that mapped directly to the stimulus. Each subject completed six experimental sessions, two self-paced (SP) and four forced-pace (FP) sessions. Two of the forced-pace sessions had a slow (FPS) stimulus presentation rate and two had a fast (FPF) stimulus presentation rate. The first session in each condition was considered a "practice" session because the results from Experiment 1 suggested that learning effects were limited to the initial session. The ISI for the forced-pace conditions was calculated from the performance statistics of each subject in the second self-paced session, with an additional 200-ms pause inserted between the response window and the next stimulus. Feedback to the subject was provided by two distinct computer-generated tones, one following incorrect responses (all sessions) and the other following missed signals (forced-pace sessions).

      Equal numbers of stimuli of each type were presented in each block in a random order. The probability of repetitions was minimised by sampling the locations without replacement, unless the stimulus location was identical to the previous location. If it was identical, it was returned to the population, and a new location selected. This process could be repeated up to four times, after which the repetitive location was used. Repetitions were, therefore, only likely to occur on the last few trials of each block. All subjects completed the self-paced condition first. To

counter-balance the possible effects of session order and gender, every second female and every second male subject completed the slow forced-pace session before the fast forced-pace session. The remaining subjects completed the forced-pace sessions in the reverse order.

Results

Mean error rates (missed and incorrect responses) differed significantly between the first (practice) and the second (experimental) sessions for the fast forced-pace condition, $t(11) = 4.62$, $p < .001$, but not for the other conditions. Hence, subjects found the forced-pace condition more demanding than the other conditions, and required some learning during the practice session to fulfil its requirements, particularly avoiding missed responses. As high error rates may be problematic, and to maintain comparability between the different conditions, only data from the experimental sessions were subjected to further analyses, unless otherwise stated. The order of slow and fast forced-pace conditions did not have significant effects on error rates, miss rates, or mean correct or incorrect RT. Hence, the order of these conditions was discounted in the following analyses.

Preliminary Analysis

The average ISI for subjects in the fast force-pace condition was 449.2 ms, and the average ISI for the slow forced-pace condition was 578.1 ms. Subjects made significantly more incorrect responses in the FPF condition ($\underline{M} = 5.2\%$, $\underline{SD} = 3.0\%$) than the FPS condition ($\underline{M} = 3.1\%$, $\underline{SD} = 1.3\%$), $\underline{F}(1,22) = 4.77$, $\underline{p} < .05$. Subjects did not commit significantly fewer errors in the SP condition (3.3%, $SD = 2.8$) than either the FPF, $\underline{p} > .05$, or the FPS, $\underline{F} < 1$, conditions. Subjects missed more signals in the FPF ($\underline{M} = 6.9$, $\underline{SD} = 2.3$) than the FPS ($\underline{M} = 1.4$, $\underline{SD} = 1.5$) condition, $\underline{F}(1,22) = 47.04$,

$p < .0001$. Hence, even after extensive practice, subjects found the fast forced-pace condition more demanding than the slow forced-pace condition.

A 3x2 within-subjects ANOVA compared mean correct and incorrect RTs across experimental conditions. There was a significant main effect for condition, $F(2, 22) = 20.68$, $p < .001$, and the mean RT for incorrect responses was significantly less than for correct responses, $F(1, 11) = 29.26$, $p < .001$. Response type and condition did not interact, $F < 1$, indicating errors ($M = 185.13$, $SD = 32.88$) were faster than correct responses ($M = 207.27$, $SD = 26.69$) by a constant amount (about 40 ms) across conditions.

A one–way within-subjects ANOVA showed that the mean correct RT differed significantly across experimental conditions, $F(2, 22) = 18.92$, $p < .001$. FPF was faster than FPS, $F(1, 11) = 19.32$, $p < .001$, and FPS was faster than SP, $F(1,11) = 14.73$, $p < .001$. For mean RT on incorrect trials, the main effect of condition was significant, $F(2, 22) = 9.17$, $p < .001$, and FPF was faster than FPS, $F(1, 11) = 4.3$, $p < .001$. The average RT for the forced-pace conditions was less than for the self-paced condition, $F(1, 11) = 3.62$, $p < .01$, but the difference between the FPS and the SP conditions was not significant, $F < 1$. Thus, on correct trials, subjects responded faster under forced-pacing than self-pacing, and on correct and incorrect trials responded faster when the ISI was small. Therefore, subjects found forced-pacing more demanding than self-pacing, and a smaller ISI more demanding than a large ISI.

Linear Dynamical Analysis

Time Series Analysis

Autocorrelation and partial autocorrelation coefficients were calculated from the raw experimental series. As in Experiment 1, the results were consistent with an

32

AR(1) model for most subjects. For 23 of the 36 individual series, only the first partial autocorrelation was significant at the 0.05 level. For the remaining series, six had two, four had three and one had four significant partial autocorrelations (two had zero).

Average first partial autocorrelation in the FPF ($\underline{M}$ = 0.18, $\underline{SD}$ = 0.10) and FPS ($\underline{M}$ = 0.17, $\underline{SD}$ = 0.12) conditions was higher than in the SP ($\underline{M}$ = 0.11, $\underline{SD}$ = 0.05). However, a one-way ANOVA did not find a significant conditions effect, $\underline{F}$(2,33) = 2.36, $\underline{p}$ = .11. The average number of significant partial autocorrelations was identical for all conditions (1.5). The number of significant autocorrelations decreased from FPF ($\underline{M}$ = 3.25, $\underline{SD}$ = 4.75) to FPS ($\underline{M}$ = 3.58, $\underline{SD}$ = 3.73) to SP ($\underline{M}$ = 2.17, $\underline{SD}$ = 1.75), but the main effect of condition did not approach significance, $\underline{F}$ < 1. These findings indicate that the linear dependence varied little with experimental condition.

Fourier Analysis

The slope of log-log power versus frequency plots differed significantly across the experimental conditions, $\underline{F}$(2,33) = 4.56, $\underline{p}$ < .02, with the steepest slope occurring in the FPF condition ($\underline{M}$ = -0.31, $\underline{SD}$ = 0.12), followed by the FPS condition ($\underline{M}$ = -0.26, $\underline{SD}$ = 0.14) and then the SP condition ($\underline{M}$ = -0.16, $\underline{SD}$ = 0.08). Mean spectral density slope for the FPF condition was significantly less than -0.5, $\underline{t}$(11) = 4.95, $\underline{p}$ < .0001, suggesting that these series were not in the $1/f$ range. Although the spectral density slopes did not differ significantly between the two forced-pace conditions, $\underline{F}$ < 1, a comparison of the average slope for the forced-pace condition with the slope for the self-paced condition was significant, $\underline{F}$(1,22) = 8.11, $\underline{p}$ < .009. These findings suggest that the forced-pace series have a higher level of dynamical structure than do the series from the SP condition, which appear to be close to white noise. However, the mean SP slope was still significantly different from zero, $\underline{t}$(11) = 6.88, $\underline{p}$ < .0001.

33

<u>Nonlinear Dynamical Analysis</u>

In each of the following nonlinear dynamical analyses, we firstly compare the

series from the three experimental conditions, and, then, compare the experimental

series with their AAFT and FSS surrogate series.

Noise Reduction

The relative standard deviation ($\underline{SD}_{original}/\underline{SD}_{filtered}$) was kept constant between

conditions by controlling the number of iterations of the filtering process (see

Appendix A). The relative standard deviations, 0.34 in the FPF, and 0.37 in the FPS,

and 0.36 in the SP conditions, did not differ significantly according to a one-way

ANOVA, $\underline{F} < 1$. The average relative standard deviations for the FSS and AAFT

surrogates were identical to those of the experimental series in each condition.

```
-------------------------
INSERT FIGURE 5
-------------------------
```

Phase Portraits

All subjects generated complex but highly structured phase portraits in the

FPF condition. Phase portraits for a sample of six subjects, chosen because they

represented the most diverse range of attractor shapes, are shown in Figure 5.

Geometrical structure similar to that observed in the FPF phase portraits was evident

for only three subjects in the FPS condition. The portraits for the remaining nine

subjects in the FPS condition were similar to the portrait of an FSS surrogate series

(e.g. Figure 3b). Individual differences disappeared altogether in the SP phase plots,

which all appeared very similar to those of the FSS portraits. An example of a phase

portrait typical of 75% of the FPS series, and a phase portrait typical of all of the SP

series, is provided in Figure 6.

INSERT FIGURE 6 ABOUT HERE

Dimensionality Estimation

Figure 7(a) shows mean $\underline{D2}$ values for each experimental condition from one

to ten embedding dimensions. All experimental series asymptoted at different levels

by 5 embedding dimensions, although the FPS series continued to increase slightly.

The mean D2s at ten embedding dimensions are given in Table 4. Condition had a

significant main effect on mean $\underline{D2}$ values, $\underline{F}(2,33) = 7.66$, $\underline{p} < .002$. The mean $\underline{D2}$

value was significantly less for the FPF than the FPS condition, $\underline{t}(11) = 3.94$, $\underline{p} < .002$,

but $\underline{D2}$ was not significantly different between the FPS and the SP conditions, $\underline{t} < 1$.

INSERT TABLE 4 ABOUT HERE

Table 5 gives the results of inferential analyses comparing the $\underline{D2}$ values at 10

embedding dimensions for the experimental series from each condition with the two

types of surrogates. The results suggest that the low $\underline{D2}$ values for the FPF condition

are not a result of either the effects of filtering or linear structure in the data. The non-

significant difference between the experimental series and the FSS surrogate for the

SP condition suggests that the SP $\underline{D2}$ values are a result of the filtering process and

not a nonlinear dynamical property of the data. A similar conclusion is suggested by

the results for the FPS condition.

INSERT TABLE 5 ABOUT HERE

Nonlinear Prediction Analysis

Nonlinear prediction error was calculated for each series across ten trial lags.

The pattern of mean prediction error for the three experimental conditions is shown in

Figure 7(b). A within-subjects ANOVA was conducted on the percent prediction errors for the three experimental conditions and ten lags. Prediction error increased significantly as the lag increased, $F(9, 99) = 50.71$, $p < .0001$, and there was also a significant main effect of experimental condition, $\underline{F}(2, 22) = 6.56$, $\underline{p} < .001$. Figure 7(a) shows that the greatest differences in prediction error between the conditions occurred at lag one. The differences between the conditions tended to disappear as the prediction error approached 100%, resulting in a significant interaction between lag and condition, $\underline{F}(18,198) = 5.46$, $\underline{p} < .001$. The FPF series had the slowest rate of approach to 100% error as lag increased, suggesting that the RT series in the FPF condition had longer dependencies than the series in the FPS or the SP conditions.

INSERT FIGURE 7 ABOUT HERE

We compared the prediction error at lag one for each condition with their corresponding AAFT and FSS surrogates. For the FPF condition, there was a significant difference between the experimental ($\underline{M}$= 46.8%, $\underline{SD}$ = 18.7%) and FSS ($\underline{M}$ = 97.9%, $\underline{SD}$ = 11.1%) series, $\underline{F}(1,22) = 54.34$, $\underline{p} < .0001$, and between the experimental and AAFT ($\underline{M}$ = 61.4%, $\underline{SD}$ = 7.1%) series, $\underline{F}(1,22) = 4.72$, $\underline{p} < 0.04$. Hence, neither the effects of filtering nor the linear structure in the series can account for the nonlinear prediction accuracy of the experimental series in the FPF condition.

For the FPS condition, there was a significant difference between the experimental ($M = 62\%$, $SD = 27.7\%$) and the FSS ($M = 99\%$, $SD = 0.9\%$) series, $\underline{F}(1,22) = 15.99$, $\underline{p} < .001$, but the difference between the experimental and the AAFT ($\underline{M} = 72.9\%$, $\underline{SD} = 23.1\%$) series was not significant, $\underline{F}(1,22) = 1.31$, $\underline{p} = .264$. Hence, in the FPS condition, the linear structure in the AAFT surrogate can account for the predicability of the experimental series. In the SP condition, there was a significant

difference between the experimental ($\underline{M}$= 81.1%, $\underline{SD}$ = 11.1%) and FSS ($\underline{M}$ = 98.%,

$\underline{SD}$ = 1.1%) series, $\underline{F}(1,22)$ = 28.61, $\underline{p}$ < .0001, and between the experimental and

AAFT (M = 98.5%, SD = 6.6%) series, F(1,22) = 28.61, p < 0.04. While the latter

finding provides evidence for nonlinear structure in the SP condition, the large

magnitude of prediction error for the experimental and AAFT series at lag one

suggests that both the linear and nonlinear structure was weak.

Individual Differences

Individuals in the FPF and SP conditions were homogenous in their dynamics.

All the series in the FPF condition had a low dimensional chaotic component, because

they had low $\underline{D2}$ values, nonlinear prediction error patterns indicative of sensitive

dependence on initial conditions, and structured phase portraits. The series from the

SP condition appeared to be composed of noise, possibly with a weak chaotic

attractor, which was only detectable by the nonlinear prediction analysis.

The series from the FPS condition, on the other hand, appeared to be composed

of noise for 75% of subjects and contained a low dimensional chaotic component for

the remaining subjects. Subjects Two, Seven and Nines' series, with structured phase

portraits in the FPS condition, had higher relative standard deviation values ($\underline{M}$ =

0.442) than the remaining series, which did not have structured phase portraits ($\underline{M}$ =

0.357). This finding suggests that these series contained an underlying attractor. The

series with the structured phase plots also had a much lower prediction error at lag

one (33%) than did the series with unstructured phase plots (78%). Hence, the

difference between the two groups of series was consistent for these converging tests

of nonlinear dynamical structure.

INSERT TABLE 6 ABOUT HERE

Discussion

The results for Experiment 2, along with the results from Experiment 1 are summarised in Table 6. The results of Experiment 2 confirm the hypothesis that chaotic characteristics are more likely to emerge in RT series when the ISI is demanding. It was only when subjects were forced to make all responses faster than the average of their self-paced response time that chaotic characteristics were detected in all subjects' RT series. Only 25% of subjects had series with chaotic characteristics under fixed pace conditions that did not force them away from their usual self-paced response speed.

It is not clear from the analysis whether the subjects classed as "non-chaotic" had a low dimensional chaotic component that was too small to be detected with the observed level of noise, or whether these series simply did not have a low dimensional chaotic component. We applied a higher level of filtering to these series by iterating the filtering procedure up to five times. Phase portraits did not settle onto a visible attractor shape, and filtering continued to remove data until there was virtually no variation left in the signal. It appears, therefore, that these series did not contain a chaotic attractor. It is important to remember that the dimensionality estimates for the series with attractors reflect the nonlinear dynamics whereas the estimates for series without an attractor reflect the level of filtering. Thus, caution needs to be taken when comparing the dimensionality estimates of the series between conditions. We can conclude from the surrogate series analysis that the series in the self-paced and most of the series in the slow forced-paced had $D2$ values which were a result of the filtering and linear dynamics in the data, and that the series from the

forced-paced fast condition had $\underline{D2}$ values which were the result of nonlinear dynamical structure. The nonlinear prediction error is less dependent on the level of filtering and thus provides both a superior measure for contrasting the series dynamics between conditions, and confirms the results of the $\underline{D2}$ analysis.

The ISI used in Experiment 1 was 1000ms, whereas the ISIs in Experiment 2 were, on average, 578 ms and 449 ms for the slow and fast conditions respectively. However, the results of the dynamical analyses in the first experiment were more similar to the results for the forced-pace fast condition than the forced pace slow condition. Even with the longer ISI used in Experiment 1, most of the filtered RT series showed evidence for low dimensional chaos, whereas 75% of series from the slow-paced condition in Experiment 2, with a much shorter ISI, did not. The apparent inconsistency in the effect of different ISIs across experiments is probably due to the different apparatus used. The button box keys in Experiment 1 required more effort to depress and had more travel than did the mouse keys in Experiment 2. Hence, a longer ISI would be required in Experiment 1 for the same level of difficulty as in Experiment 2, purely due to mechanical effects.

All subjects were pushed into a low dimensional chaotic responding regime in the FPF condition in Experiment 2. In Experiment 1, on the other hand, some subjects, most notably subjects four and six, may have found this ISI too easy, and were not forced to respond in the chaotic regime. The individually set ISI in Experiment 2 was likely to be responsible for the increase to 100% of subjects responding in the chaotic regime.

There were, however, quantitative differences in the linear and nonlinear dynamics between the series in Experiment 1 and the force-paced fast condition of Experiment 2. Both the linear and the nonlinear dynamical components appeared to

39

be approximately twice as large in the Experiment 1 series. The average size of the first partial autocorrelation was 0.29 in Experiment 1 and 0.18 in the FPF condition in Experiment 2. The contribution of the nonlinear component to prediction accuracy was calculated as the difference between the AAFT and experimental series prediction errors at one lag. In Experiment 1, it was double (30.52%) the magnitude for the forced -pace fast condition in Experiment 2 (14.55%). These findings suggest that differences between the two experiments, such as the presence of repetitions and/or response button sensitivity, influenced the magnitude of linear and nonlinear dynamics. Determining the details of these effects of these factors is beyond the scope of the present analysis, however, Experiment 2 does clearly demonstrate that nonlinear dynamics are detectable even when their effects are minimised.

The linear and nonlinear components of dynamics appear at least partially independent. The linear structure was not statistically distinguishable between the fast and slow forced -pace conditions, as measured by both the autocorrelations in the RT series and the log-log spectral density slopes. The nonlinear dynamics of the fast and slow forced -pace conditions, in contrast, were easily distinguishable on all measures of low dimensional chaos. Conversely, linear dynamics were less evident in Experiment 2 than Experiment 1, but qualitatively similar low dimensional chaotic structure was observed for most subjects in Experiment 1 and all subjects in the fast forced-pace condition of Experiment 2. These dissociations between linear and nonlinear structure suggest that the nonlinear structure in RT, which emerges as a stable attractor with complex geometry, is a source of information in its own right.

The hypothesis that forced-pacing is required to detect a component with chaotic characteristics in RT sequences cannot be clearly arbitrated by the present results. There was little evidence for low dimensional chaos in the SP condition. SP

series had neither geometrical structure nor $\underline{D2}$ values that were different from the FSS surrogates, although the prediction analysis did suggest a weak sensitive dependence on initial conditions. There was stronger evidence for low dimensional chaos in the FPS condition, at least for a minority of subjects. However, the difficulty of these two conditions, while similar, was not exactly matched and comparison between these conditions is confounded with order of presentation of the conditions, as the SP condition always occurred first. Hence, it is not clear that the lack of chaotic characteristics in the SP condition was due to violations of the requirements of the dynamical analysis caused by the variable pacing, or due to reduced task difficulty or task order. However, self-paced conditions are used in most RT tasks in cognitive psychology, and so the question of whether or not nonlinear dynamics can be observed in self-paced RT is an important one. Clearly, more work needs to be done to determine if nonlinear dynamics can be observed in self-paced RT.

General Discussion

The results of Experiments 1 and 2 confirm that a portion of RT variance can behave like low dimensional chaos under forced-pace conditions, at least when the task is relatively difficult. However, other researchers have either found no evidence for chaotic structure in RT sequences (e.g. Gilden, 1997; Clayton & Frey, 1996), or the evidence has proven nonreplicable (e.g. Cooney & Troyer, 1994). We suggest that there may be several reasons for this discrepancy, both cognitive and technical. Firstly, we obtained long, stationary data sets and used sophisticated noise reduction techniques to clean our data. To our knowledge such techniques have not been applied previously to such long time series from cognitive processes. Another possible contributing factor is the simplicity of the SRT task. Each trial in the present

experiment was homogenous except for stimulus location. Gilden (1997), on the other hand, presented trials from a number of different conditions that required different amounts of cognitive processing, in a randomised sequence. This yielded a single RT series with a response occurring at multiple time scales that may have masked underlying low-dimensional dynamics.

To our knowledge, ours is the first attempt to provide a nonlinear dynamical analysis for forced-pace RT. Most previous experiments have used a self-paced paradigm. Hence, self-pacing may be responsible for previous failures to detect chaotic dynamics. Our finding that low dimensional chaotic structure reliably emerges in RT sequences under forced-pace but not self-paced conditions suggests sampling at regular intervals may be a general measurement requirement for detecting low dimensional chaos in RT. It is possible that low dimensional nonlinear dynamics underlie self-paced performance, but that the irregular measurement intervals make it difficult to detect. Alternately, the extra cognitive processes involved in self-paced responding may cause the underlying system to become higher dimensional. However, due to the confounding effects of task difficulty and task order in the present experiment, more work needs to be done to show definitively that chaotic characteristics do not occur in self-paced RT.

The effect of ISI on RT dynamics may be explained by subjects' response strategies to the increased task demands of the fast forced-paced condition. Any number of different psychological processes may be involved in the dynamics of the response strategy, and it is not apparent from the present results what the dynamics in RT reflect, nor is it within the scope of this paper to identify such processes. We make two speculations as to possible sources of the change in dynamics.

The dynamics in RT may represent motor coupling. Motor coupling between sequential responses will be stronger at shorter ISIs, where nonlinear dynamics are more evident. The difficult-to-press response buttons used in Experiment 1 may have increased the coupling between the movements involved in consecutive responding, as the time spent responding to each stimulus is increased, and so nonlinear dynamics are observed at longer ISIs. The stronger linear dynamics observed in Experiment 1 compared to Experiment 2 may also reflect this motor coupling.

The RT dynamics might also reflect a response strategy developed by each subject to satisfy the dual constraints of responding accurately yet rapidly. A response boundary on faster RT is imposed by cognitive and neurological limitations, and subjects cease to respond accurately if RT falls below this level (Luce, 1986). A second response boundary is imposed on slower RT when the ISI short. The two constraints trade-off, so responding too slowly leads to an increased likelihood of missing a signal, whereas responding too quickly leads to an increased likelihood of making an error.

There is evidence that subjects may use error feedback to control RT dynamically. Rabbitt and Rodgers (1977) showed that responses prior to an error increase in speed, and slow immediately following the error. They suggest that this "response programming" is employed by subjects in order to produce both fast and accurate responses. Brewer and Smith (1989) showed marked developmental differences in the ability to accurately control response programming, with younger subjects producing much larger than optimal corrections in response to an error. Under fast forced-pacing, subjects must respond within a very small time window because the ISI limit is close to the speed-accuracy limit, and so their control must be correspondingly accurate. It may be the control mechanisms that constrain

43

fluctuations in RT between the fast and slow response boundaries that are responsible for the nonlinear dynamics observed in the present experiments.

Conclusions

The first step to understanding a phenomenon is to establish conditions under which it can be reliably observed and measured. Nonlinear dynamical approaches to cognition have been limited because it has been difficult to find reliable evidence for low dimensional chaos. We have established that one set of conditions, fast forced pacing, results in detectable low dimensional chaotic dynamics in RT sequences for all subjects performing a serial response time task. We have also provided a methodology for detecting low dimensional chaos in behavioural data, including noise reduction and a converging set of linear and nonlinear tests.

Rapp (1994) has argued that chaos analysis is only valid if it uncovers new and useful information. We have shown that two phenomena often treated as "error", individual differences and inter-trial fluctuations, do contain useful information, specifically, low dimensional nonlinear dynamics that individually characterise each subject. From a theoretical standpoint, our results indicate that a component of complex inter-trial fluctuations in RT can be modelled by deterministic recursive equations with linear and nonlinear components. The successful development of such models would fulfil Luce's (1995) ideal of accounting for realistically complex and fluctuating behaviours more parsimoniously than can be done by linear stochastic models.

Acknowledgments

References

Bak, P., & Creutz, M. (1994). Fractals and self-organized criticality. In A. Bunde, & S. Havlin (Eds.) Fractals in science, pp. 26-47. Berlin: Springer-Verlag.

Bak, P., Tang, C., & Wiesenfeld, K. (1987). Self-organized criticality: An explanation of *1/f* noise. Physical Review Letters, 594, 381-384.

Barton, S. (1994). Chaos, self-organisation, and psychology. American Psychologist, 48, 5-14.

Bloomfield, P. (1976). Fourier analysis of time series: an introduction. NY: Wiley.

Box, G. E. P., & Jenkins, G. M. (1976). time series analysis: forecasting and control. San Francisco: Holden Day.

Brewer, N., & Smith, G.A. (1989). Developmental changes in processing speed: Influence of speed-accuracy regulation. Journal of Experimental Psychology, 118, 298-310.

Broadbent, D.E. (1954). Noise, paced performance and vigilance tasks. British Journal of Psychology, 44, 295-303.

Broomhead, D.S., Huke, J.P., & Muldoon, M.R. (1992). Linear filters and nonlinear filters. Journal of the Royal Statistical Society, Series B, 54, 373-382.

Buldyrev, S.V., Goldberger, A.L., Havlin, S., Peng, C.-K, & Stanley, H.E. (1994). Fractals in biology and medicine: From DNA to the heartbeat. In: A. Bunde, & S. Havlin (Eds.), Fractals in science, pp. 48-87. Berlin: Springer-Verlag.

Casdagli, M. (1991). Chaos and deterministic versus stochastic non-linear modelling. Journal of the Royal Statistical Society, Series B, 54, 303-328.

Clark, A. (1997). The dynamical challenge. Cognitive Science, 21, 461-481.

Clayton, K., & Frey, B. (1996). Inter- and intra- trial dynamics in memory and choice. In W. Sulis, & A. Combs, (Eds.), Nonlinear dynamics in human behaviour, pp. 90-106. Singapore: World Scientific.

Cooney, J. B., & Troyer, R. (1994). A dynamic model of reaction time in a short-term memory task. Journal of Experimental Child Psychology, 58, 200-226.

Coveney, P., & Highfield, R. (1996). Frontiers of complexity: the search for order in a chaotic world. NY: Ballantine.

Cryer, J. D. (1986). Time series analysis. Boston: PWS Kent.

Davies, D.R., & Parasuraman, R. (1982). The psychology of vigilance. London: Academic Press.

Dooley, K.J., & Van de Ven, A.H. (1999). Explaining complex organizational dynamics. Organizational Science 10, 358-372.

Eliasmith, C. (1996). The third contender: a critical examination of the dynamicist theory of cognition. Philosophical Psychology, 9, 441-463.

Fletcher, B., & Rabbitt, P.M.A. (1978). The changing pattern of perceptual analytic strategies and response selection with practice in a two-choice reaction time task. Human Factors, 20, 417-427.

Garson, J.W. (1996). Cognition poised at the edge of chaos: a complex alternative to a symbolic mind. Philosophical Psychology, 9, 301-322.

Gilden, D.L. (1997). Fluctuations in the time required for elementary decisions. Psychological Science, 8, 296-301.

Grassberger, P., & Procaccia, I. (1983). Characterization of strange attractors. Physical Review Letters, 50, 189-208.

Grassberger, P., Hegger, R., Kantz, H., Schaffrath, C., & Schreiber, T. (1993). On noise reduction methods for chaotic data. Chaos, 3, 127-141.

Heath, R.A. (2000). Nonlinear Dynamics: techniques and applications in psychology. Mahwah, NJ: Erlbaum.

Heathcote, A. (1988). Screen control and timing routines for the IBM microcomputer family using a high-level language. Behaviour Research Methods, Instruments & Computers, 20, 289-297.

Heathcote, A., Brown, S., & Mewhort, D. J. K. (in press). The Power Law repealed: The case for an Exponential Law of Practice. Psychonomic Bulletin & Review.

Hegger, R., Kantz, H., & Schreiber, T. (1999). Practical implementation of nonlinear time series methods: The TISEAN package. Chaos, 9, 413.

Hollenbeck, J.R., Ilgen, D.R., Tuttle, D.B., & Sego, D.J. (1995). Team performance on monitoring tasks: an examination of decision errors in contexts requiring sustained attention. Journal of Applied Psychology, 80, 685-696.

Kantz, H., & Schreiber, T. (1997). Nonlinear time series analysis. UK: Cambridge University Press.

Kaplan, D., & Glass, L. (1995). Understanding nonlinear dynamics. NY: Springer-Verlag.

Kauffman, S. (1995). At home in the universe. Oxford: Oxford Press.

Kelso, J.A.S. (1995). Dynamic patterns: the self-organisation of brain and behaviour. Cambridge: MIT.

Laming, D.R.J. (1968). Information theory of choice-reaction times. UK: Academic.

Large, E.W., & Jones, M.R. (1999). The dynamics of attending: How people track time-varying events. Psychological Review, 106, 119-159.

Lintern, G., & Kugler, P.N. (1991). Self-organization in connectionist models: Associative memory, dissipative structures, and thermodynamic law. Human Movement Science, 10, 447-483.

Longstaff, R.G., & Heath, R.A. (1999). A nonlinear analysis of the temporal characteristics of handwriting. Human Movement Science, 18, 485-524.

Luce, R.D. (1986). Response times: their role in inferring elementary mental organisation. UK: Oxford Press.

Luce, R.D. (1995). Why should mathematics play a role in psychology? Annual Review of Psychology, 46, 1-26.

Makeig, S., & Inlow, M. (1993). Lapses in alertness: coherence of fluctuations in performance and EEG spectrum. Electroencephalography and Clinical Neurophysiology, 86, 23-25.

Mayer-Kress, G. (1994). Localized measures for nonstationary time-series of physiological data. Integrative Physiological and Behavioural Science, 29, 205-210.

Mpitsos, G.J. (1994). The chaos user's tool kit. Integrative Physiological and Behavioural Science, 29, 307-310.

Nicolis, J.S. (1991). Chaos and information processing: a heuristic outline. Singapore: World Scientific.

Pressing, J. (2000). Sources for 1/f noise effects in human cognitive performance. In: Heath, R., Heathcote, A., Hayes, B., Hooker, C. (Eds.), Proceedings of the Fourth Australasian Cognitive Society Conference (CD-ROM). University of Newcastle: Australia.

Pressing, J. (1998). Error correction processes in temporal pattern preduction. Journal of Mathematical Psychology, 42, 63-101.

Pressing, J., & Jolley-Rogers, G. (1997). Spectral properties of human cognition and skill. Biological Cybernetics, 76, 339-347.

Rabbitt, P. (1979). How old and young subjects monitor and control responses for accuracy and speed. British Journal of Psychology, 70, 305-311.

Rabbitt, P. (1989). Sequential Reactions. In D.H. Holding (Ed.), Human skills, pp. 145-170. Chichester: Wylie.

Rabbitt, P., & Rodgers, B. (1977). What does a man do after he makes an error? An analysis of response programming. Quarterly Journal of Experimental Psychology, 77, 17-27.

Ramsey, J.B., & Yuan, H-J. (1990). The statistical properties of dimension calculations using small data sets. Nonlinearity, 3, 155-176.

Rapp, P. (1994). A guide to nonlinear dynamical analysis. Integrative Physiological and Behavioural Science, 29, 311-327.

Remington, R.J. (1971). Analysis of sequence effects for a four-choice serial reaction time experiment. Journal of Experimental Psychology, 77, 17-27.

Rogers, K.A., Schneider, W., Pitcher, E., & Zuccolotto, (1995). MEL professional: language reference guide. Pittsburgh: University of Pittsburgh.

Schiff, S.J., Sauer, T., & Chang, T. (1994). Discriminating deterministic versus stochastic dynamics in neuronal activity. Integrative Physiological and Behavioural Science, 29, 246-261.

Smith, G.A., & Brewer, N. (1995). Slowness and age: speed-accuracy mechanisms. Psychology and Aging. 10, 238-247.

Sprott, J.S. (1995). Chaos data analyser: user's manual. Raleigh, USA: Physics Academic Software.

Sugihara, G., & May, R. (1990). Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. Nature, 344, 734-739.

Takens, F. (1981). Detecting strange attractors in fluid turbulence. In: D. Rand, & L-S. Young (Eds.), Dynamical systems and turbulence, (pp. 365-381). USA: Springer.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., & Farmer, J.D. (1992). Testing for nonlinearity in time series: The method of surrogate data. Physica D, 58, 77-94.

van Gelder, T. (1997). Dynamics and cognition. In: J. Haugeland (Ed.) Mind design II, pp.421-451. Cambridge: MIT Press.

Van Leeuwan, C., Steyvers, M., & Nooter, M. (1997). Stability and intermittency in large-scale coupled oscillator models for perceptual segmentation. Journal of Mathematical Psychology, 41, 319-344.

Wertheimer, M. (1953). An investigation of the "randomness" of threshold measurements. Journal of Experimental Psychology, 45, 294-303.

Appendix A: Technical Details of Nonlinear Dynamical Analyses

The noise reduction procedure and the nonlinear prediction algorithm are described in detail in Hegger, Kantz and Schreiber (1999). The noise reduction algorithm and the prediction algorithm used are available through the TISEAN Project (Hegger, et al, 1999). The commands are **PROJECT** and **PREDICT** respectively. The dimensionality estimation algorithm was calculated using the Chaos Data Analyser (CDA) software (Sprott, 1995).

Data series used in the present paper are available from the authors' home pages at http://psychology.newcastle.edu.au/~akelly/QJEPdata/. These series have been made available to enable the reader to undertake a dynamical analysis on experimental data.

Locally Projective Nonlinear Noise Reduction

The algorithm assumes that the series is composed of a low-dimensional dynamical attractor and high-dimensional noise. According to the embedding theorems (e.g. Takens, 1983) we can reconstruct the system attractor if the embedding dimension (M) is high enough. However, the sub-manifold (q) on which the low-dimensional attractor lies cannot be contained within the embedding dimension if the attractor is corrupted by high dimensional noise. To overcome this problem, the locally projective filtering applies a correction $\Theta_n \in q$ with $\|\Theta_n\|$ small to each embedding vector $\mathbf{S}_n$, so that $\mathbf{S}_n - \Theta_n \in q$ and Θ_n is orthogonal on q. The first and last components of the delay vectors are fixed, and the rest are corrected. The algorithm also corrects for possible curvature introduced into the trajectory during the filtering process, by correcting the centre of mass for each neighbourhood, because this centre of mass has a bias towards the centre of curvature.

Parameter Choice

Some trial and error was required when choosing parameters for filtering, as the nonlinear dynamical properties of our data series were unknown. Over-embedding is necessary because of the high-dimensional noise in the unfiltered series. For this reason, we chose an embedding dimension of 10. Hegger et al. (1999) suggest a small delay is preferable unless the data is highly oversampled (in our case this was unlikely). Thus, because of the need to keep the parameters identical for each series, we used a delay of 1. For the dimension of the sub-manifold to which the data is being moved, we found that there was no real difference in the resulting series for 1, 2 and 3 manifold dimensions. Because the filtering only moves points in the direction of the manifold, and thus the dimension of this manifold does not reflect the dimension of the filtered series, Hegger et al. suggest that a too small manifold will not be harmful. This is especially the case for very noisy data such as ours. The neighbourhood size (100 data points) and the minimum radius (0.01% of the series) were chosen through trial and error because the amount of noise in the series was unknown, but was assumed to comprise a large portion of the variance. The parameter values chosen were fairly robust, as they were appropriate for all series with chaotic structure in the two experiments.

The number of iterations of the filtering process was chosen through trial and error, with the final values chosen because they gave the cleanest phase portraits of the series with chaotic structure, and because they provided a comparable amount of filtering for series without an attractor. Five iterations were used to filter the data from Experiment 1 and the FPF condition of Experiment 2, four for the FPS condition and three for the SP condition in Experiment 2.

<u>Grassberger-Procassia Dimensionality Estimation (D2)</u>

To calculate <u>D2</u>, the series, $x_{i, i = 1, N}$, has to be embedded in an m-dimensional

phase space. The trajectory was reconstructed in this space using Taken's method of

delays. Delay vectors, $\boldsymbol{x}_i$, of embedding dimension m and delay time t are defined by:

$\boldsymbol{x}_i = \{x_i, x_{i-t}, \ldots, x_{i-(m-1)t}\}, i = 1, N$. The correlation integral $C(R)$ is given by:

$$C(R) = \lim_{N \to \infty} \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \Theta(R - |\boldsymbol{x}_i - \boldsymbol{x}_j|)$$

Θ *(.)* is the Heaviside step function which maps positive arguments to 1 and non-

positive arguments on to 0. $C(R)$ is a measure of the probability that two arbitrary

points, $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, in the phase space will be separated by a distance R or less. D2 was

estimated from experimental data by the slope of the best fitting line when *ln C(R)*

was regressed on *ln R*. The correlation dimension was estimated as the average slope

of the cumulative curve over the middle one-quarter of the vertical scale (Sprott,

1995). <u>D2</u> was calculated for a range of embedding dimensions from 1 to 10. The

program used to calculate <u>D2</u>, Chaos Data Analyser (Sprott, 1995), provides standard

error bars for the <u>D2</u> estimates. For all individual series with attractors, the standard

error bars were very small, suggesting highly accurate measurement.

<u>Nonlinear Prediction</u>

The simple nonlinear prediction algorithm used in this paper approximates the

nonlinear component in the series (if it exists) in local regions (neighbourhoods) of

the embedding space, using a constant. As described by Hegger et al. (1999), all

neighbours of the embedding vector $\mathbf{s_n}$ in a delay embedding space are sought. To

predict the measurements at time $n + k$, the forecast is the average of the neighbours:

$$\hat{S}_{n+k} = \frac{1}{|U_n|} \sum_{s_j \in U_n} S_{j+k}$$

U is the local neighbourhood of data points and k is the number of steps ahead in time the prediction is made. The predicted values are compared with actual values k steps ahead.

Parameter Choice

The parameters required were the embedding parameters (dimension = 4, and delay = 1) and the neighbourhood size. The embeddings were chosen by trial and error, with the criterion being the smallest prediction error (Hegger, et al, 1999). The neighbourhood size was small (5) in the series from Experiment 1 because the noise was expected to be minimal in the filtered series. However, in Experiment 2, a neighbourhood of 50 was used because it reduced the prediction error for all series.

Tables

Table 1. Dynamical analysis techniques and the expected behaviour of different types

of series.

Series Type	Analysis Technique				
	AC (Auto-correlation)	Spectral Analysis	Phase Portrait	Dimensionality Estimation	Nonlinear Prediction
Stochastic	No significant AC coefficients	Zero Slope	Complex, but no clearly defined shape	High dimensionality	100% prediction error
Stochastic with Linear Dynamics	Significant AC coefficients	Negative Slope	Complex, but no clearly defined shape	High dimensionality	Prediction error may increase with lag due to correlations
Chaotic	Usually significant AC coefficients	Usually Negative Slope	Complex yet defined shape	Finite, fractal dimensionality	Prediction error increases with step size due to sensitive dependence

Table 2. Spectral density slopes from log power-log frequency plots for each subject

and experimental session in Experiment 1.

Subject	Session 1	Session 2	Session 3
1	-0.540	-0.762	-0.763
2	-0.534	-0.348	-0.501
3	-0.435	-0.636	-0.517
4	-0.156	-0.201	-0.097
5	-0.285	-0.245	-0.252
6	-0.298	-0.278	-0.117
7	-0.550	-0.374	-0.442
8	-0.430	-0.416	-0.394
9	-0.523	-0.541	-0.436
10	-0.706	-0.711	-0.527
Mean Slope	**-0.445**	**-0.451**	**-0.405**
SD	**0.161**	**0.200**	**0.202**

 Table 3. D2 values at 10 embedding dimensions for subjects in each session and their corresponding surrogate series in Experiment 1.

Subject	Experimental Series		FSS* Surrogate		AAFT** Surrogate	
	Session		Session		Session	
	1	2	1	2	1	2
1	2.53	1.95	4.85	5.30	2.69	2.91
2	2.78	2.53	5.42	5.20	5.74	4.75
3	3.36	3.38	5.42	5.41	4.94	4.54
4	2.71	2.10	5.60	5.61	5.84	6.10
5	3.30	3.10	5.00	5.41	5.75	5.20
6	2.70	2.25	4.70	5.04	5.55	5.73
7	3.10	3.53	5.25	5.27	4.93	4.80
8	2.32	2.87	5.60	5.83	4.27	5.70
9	3.79	3.41	5.89	5.46	5.46	4.80
10	2.20	3.15	5.02	5.00	3.52	3.76
Mean	**2.85**		**5.31**		**4.85**	
SD	**0.53**		**0.99**		**0.31**	

*FSS = Filtered Sequence Shuffled.
** AAFT = Amplitude Adjusted Fourier Transform.

Table 4. Dimensionality Estimates (D2) at 10 embedding dimensions for experimental series for each subject and condition of Experiment 2, and mean D2s for the experimental and surrogate series at 10 embedding dimensions with standard deviations in brackets.

Subject	Force-Paced Fast	Force-Paced Slow	Self-Paced
1	2.240	3.152	3.44
2	4.010	4.493	3.90
3	2.260	3.161	3.15
4	2.380	3.060	2.68
5	1.810	2.980	3.21
6	2.950	3.151	4.10
7	2.880	3.155	3.09
8	2.440	4.352	3.64
9	1.980	2.505	3.50
10	3.210	2.711	3.50
11	2.034	3.125	3.10
12	2.550	3.064	3.10
Mean	**2.56 (0.62)**	**3.24 (0.59)**	**3.37 (0.39)**
Mean FSS	**4.39 (0.53)**	**3.57 (0.68)**	**3.40 (0.46)**
Mean AAFT	**4.78 (1.01)**	**5.37 (0.76)**	**6.32 (0.22)**

Table 5. Results for inferential tests comparing the mean dimensionality ($\underline{D2}$)

estimates at 10 embedding dimensions of experimental series in each condition of

Experiment 2 with corresponding Filtered Sequence Shuffled (FSS) and Amplitude

Adjusted Fourier Transform (AAFT) surrogate series.

Condition	Surrogate	T	*p*
Force-paced fast (FPF)	FSS	6.67	0.0001
	AAFT	6.29	0.0001
Force-paced slow (FPS)	FSS	0.69	0.5060
	AAFT	8.95	0.0001
Self-paced (SP)	FSS	0.26	0.8010
	AAFT	24.13	0.0001

Table 6. Summary of results for series in each condition on different linear and

nonlinear analyses in Experiments 1 and 2.

	Linear Dynamical Analysis	**Phase Portrait Structure**	**Dimensional Estimation (D2)**	**Nonlinear Prediction Pattern**	**Conclusion**
Experiment 1	Mainly AR(1) Coloured noise, some series 1/*f*	80% of subjects complex, defined and individual 20% like FSS	Saturated at finite value. Series had Lower D2 than FSS* and AAFT** series	Initial low error increasing over lag	80% of subjects chaotic on all measures with consistent geometry in 2 sessions. 20% high dimensional noise with weak sensitive dependence detected only on the prediction analysis
Experiment 2					
Forced-Pace Fast (FPF)	Mainly AR(1) Coloured Noise, some series 1/*f*	100% subjects complex, defined and individual	Saturated lower than both FSS* and AAFT** mean D2	Initial low error increasing over lag	100% of subjects chaotic on all measures
Forced-Pace Slow (FPS)	Mainly AR(1) Coloured Noise	75% of subjects same as FSS*	25% subjects saturated lower than FSS* and AAFT. 75% same as FSS*	75% of subjects high initial error increasing slightly over lag. 25% low initial error increasing over lag.	25% of subjects chaotic on all measures. 75% high-dimensional noise.
Self-Paced (SP)	Mainly AR(1) White Noise range	100% subjects same as FSS*	Same as FSS*	High initial error increasing slightly over lag.	High-dimensional noise with a possible small, weak sensitive dependence detected only on the prediction analysis.

*FSS = Filtered sequence shuffled surrogate series.
**AAFT = Amplitude adjusted Fourier transform surrogate series

Figure Captions

Figure 1. A typical data series of RTs (a) prior to filtering, and (b) after filtering.

Figure 2. Phase portraits of series in experimental sessions two and three for five participants in Experiment 1.

Figure 3. A phase portrait for (a) subject four's second experimental session in Experiment 1, (b) an example of a filtered sequence-shuffled (FSS) surrogate series, and (c) an example of an AAFT surrogate series generated from a filtered experimental series.

Figure 4. (a) Mean $D2$ estimates with standard errors for filtered experimental RT series, and the AAFT and FSS surrogate series generated from each of the filtered experimental series. (b) Mean nonlinear prediction error with standard errors for experimental time series and the FSS and AAFT surrogate series at one to ten lags.

Figure 5. Phase portraits for six subjects' in the fast force-paced (FPF) condition of Experiment 2.

Figure 6. Examples of phase portraits for RT series from (a) the slow force-paced (SFP) condition and (b) the self-paced (SP) condition of Experiment 2.

Figure 7(a). Mean $D2$ estimates as a function of embedding dimension, with standard errors, for filtered experimental RT series in each condition of Experiment 2. (b) Mean nonlinear prediction error with standard errors for experimental series from the fast forced-pace (FPF), slow forced-pace (SFP), and self-paced (SP) conditions in Experiment 2.
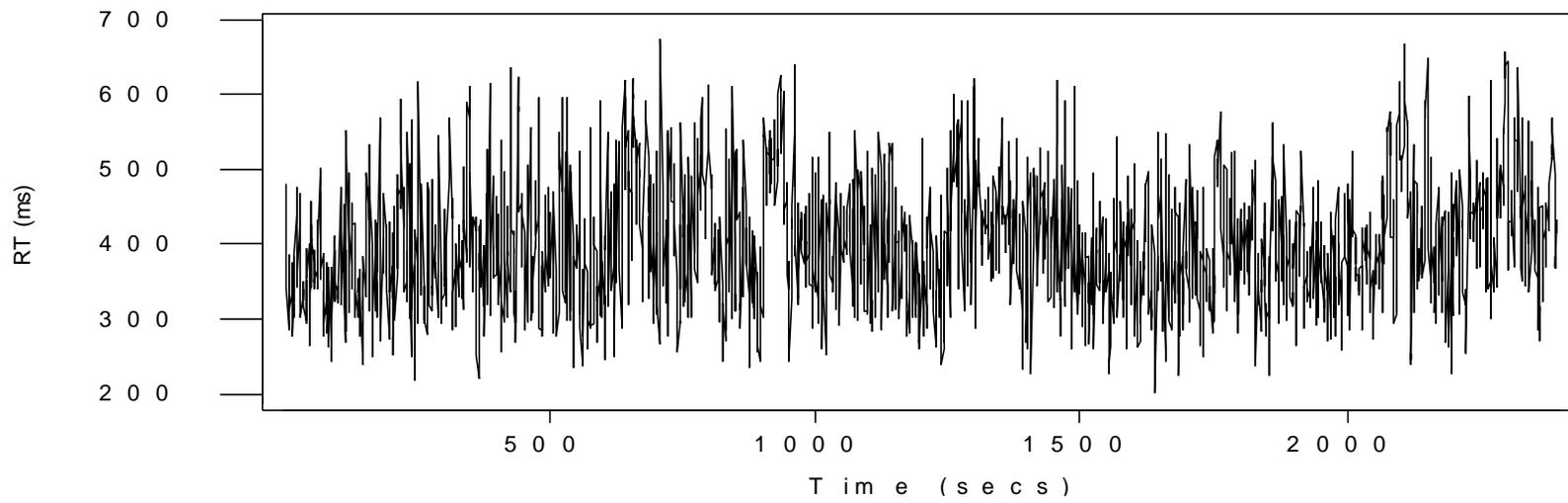
(a)
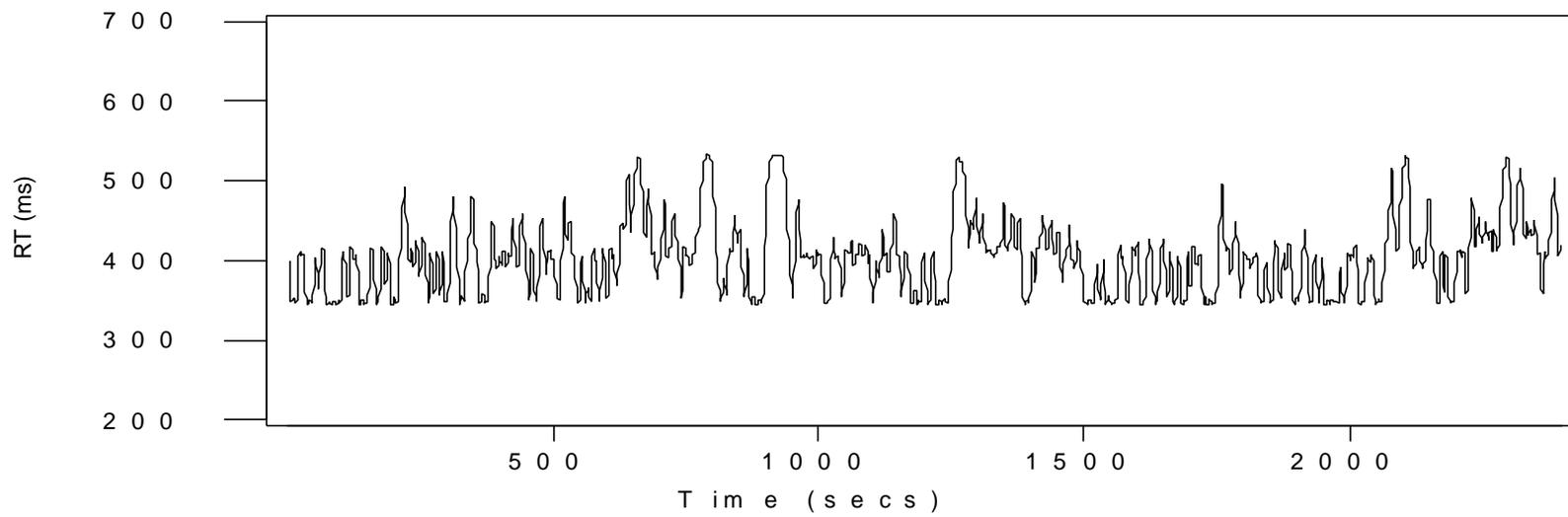
(b)
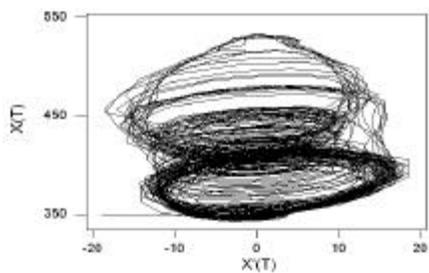
Subject 1, Session2
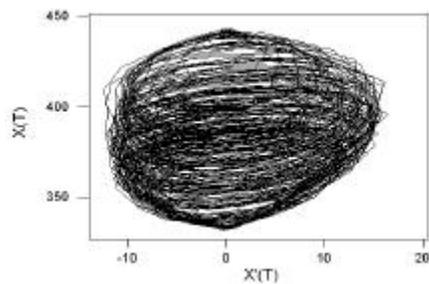
Subject 1, Session 3
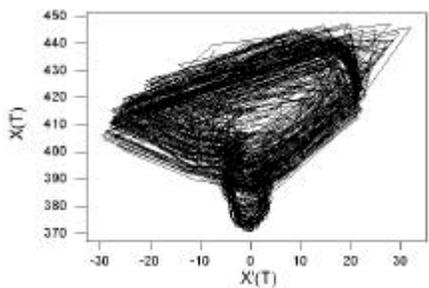
Subject 2, Session 2
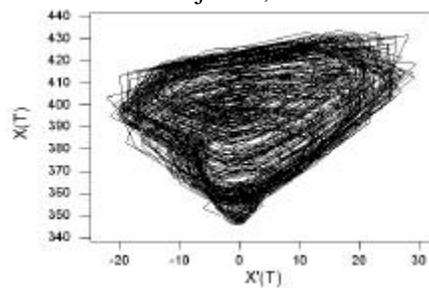
Subject 2, Session 3
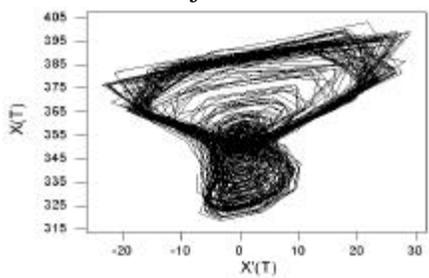
Subject 3, Session 2
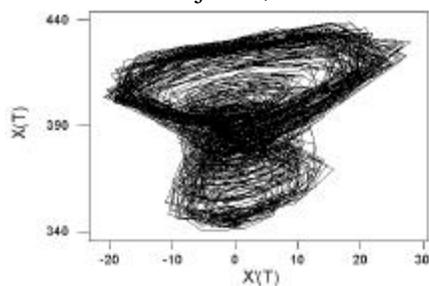
Subject 3, Session 3
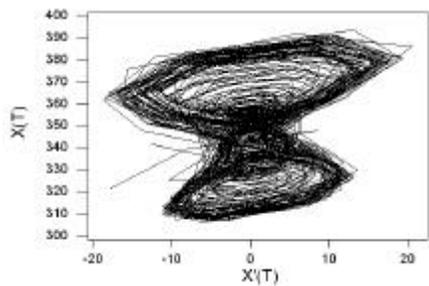
Subject 7, Session 2
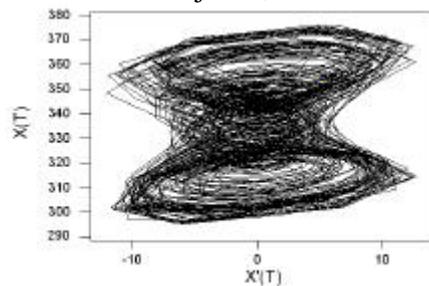
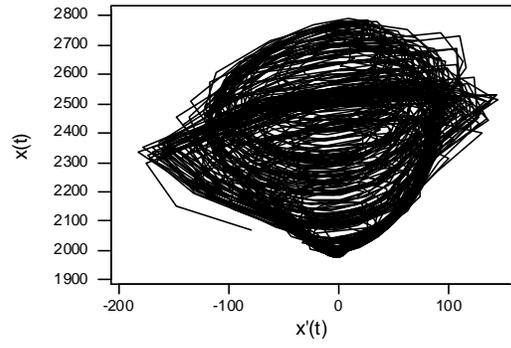Subject 7, Session 3
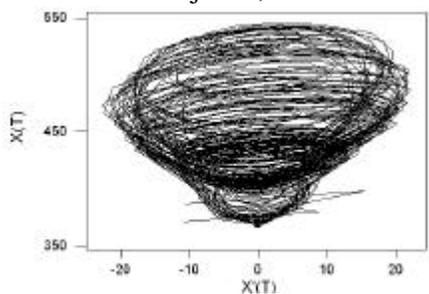
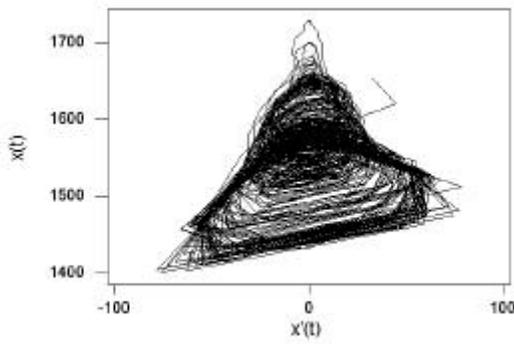Subject 10, Session 2

Subject 10, Session 3
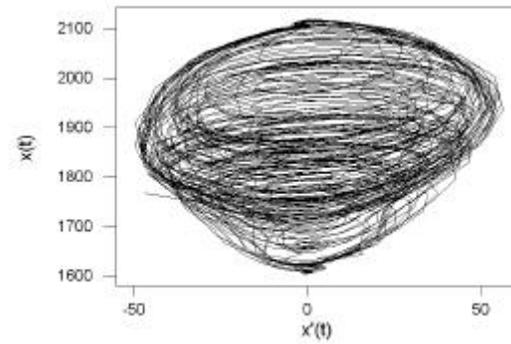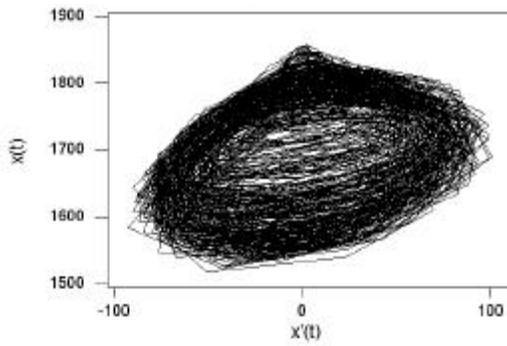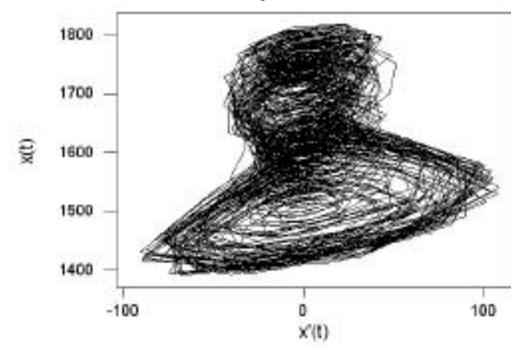
(a)

(b)

(c)
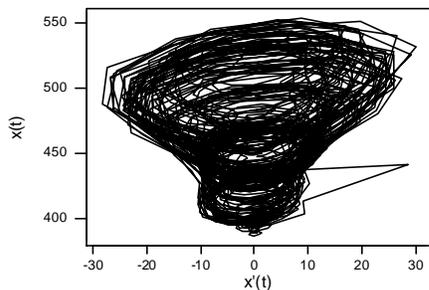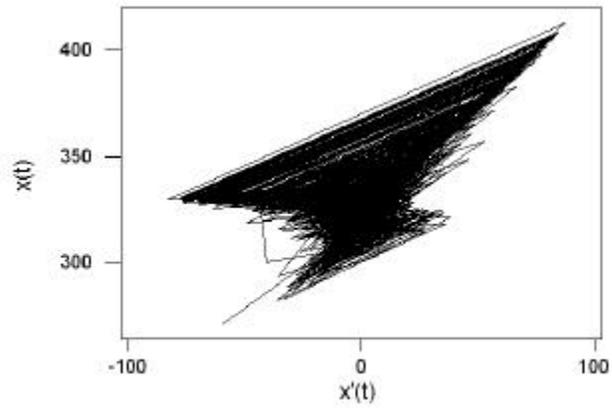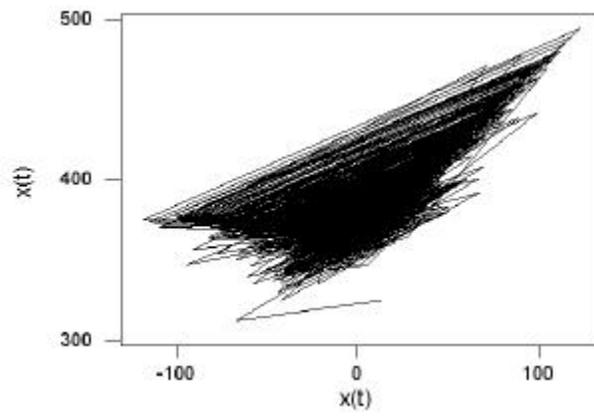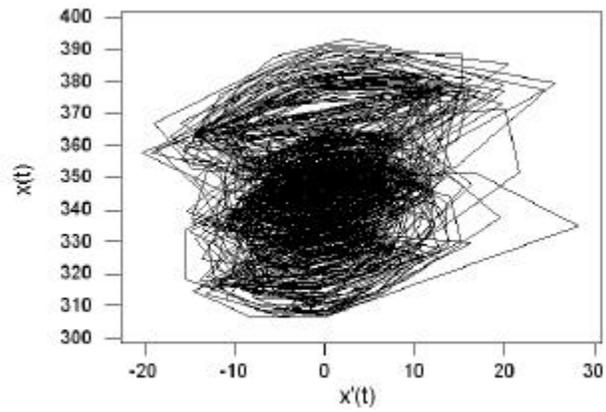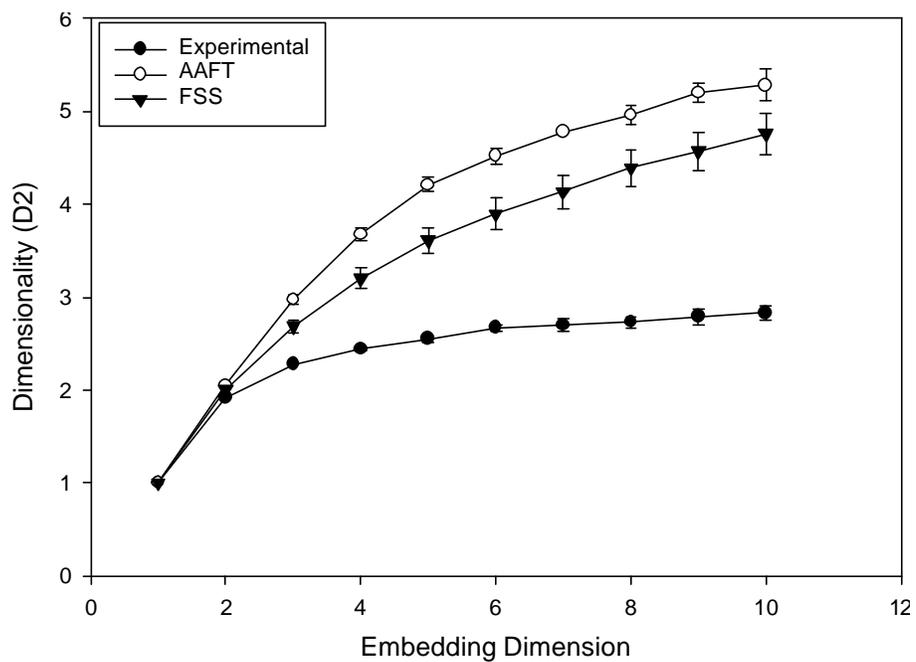
(a)
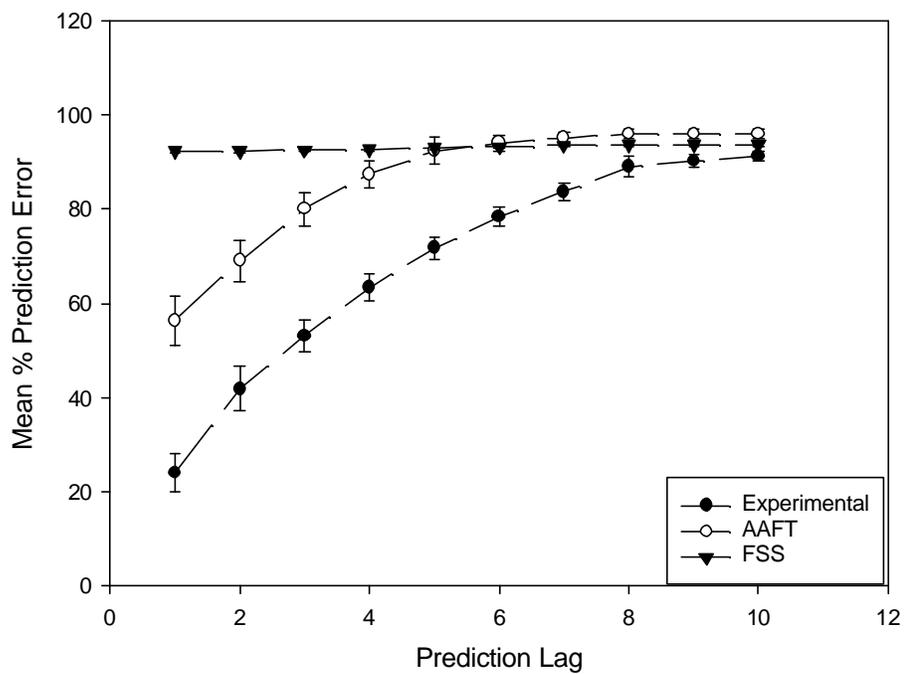
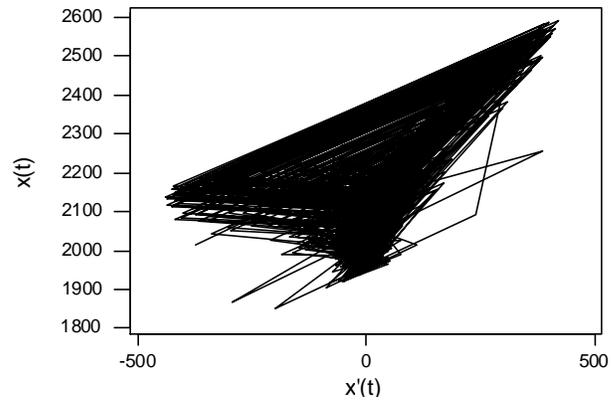

(b)

Response Time Dynamics
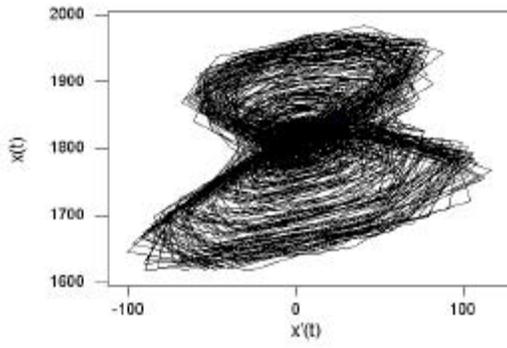
Subject 1

Subject 4

Subject 6

Subject 9

Subject 11

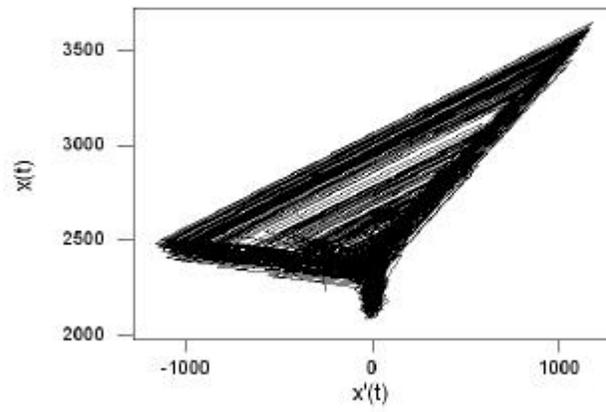Subject 12

3

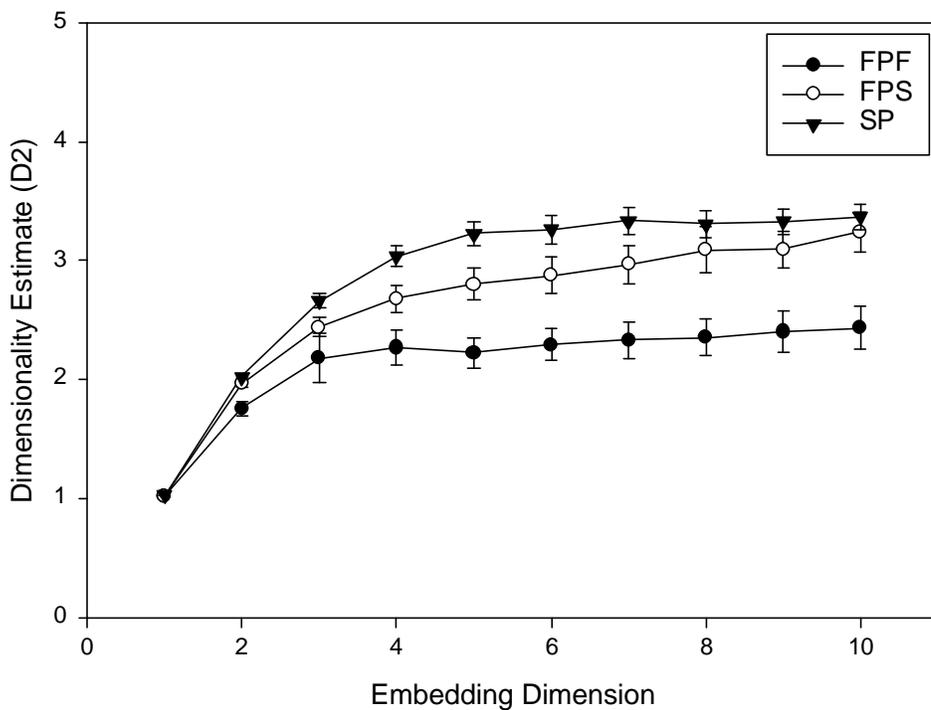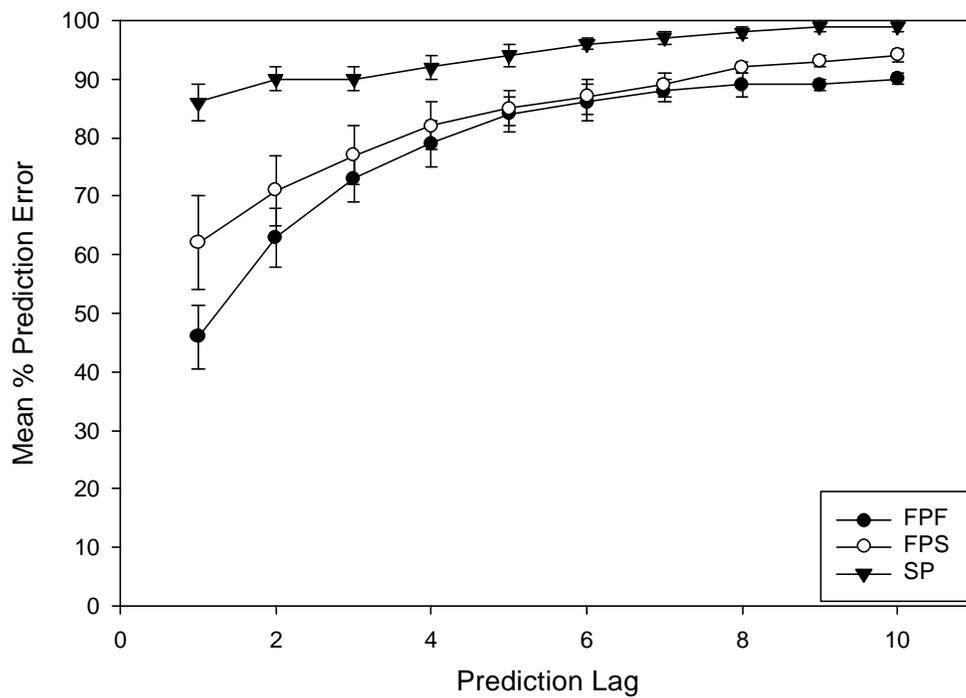(a)

(b)

(a)

(b)

Footnotes

[1] We found the relative standard deviation to be more dependent on the parameter values chosen for the filtering process than the absolute size of the structured component of the RT. Specifically, an increase in dimension manifold from 1 to 2 had a very large influence on the relative standard deviation, but almost no effect on the phase portrait structure of the series. Thus, the relative standard deviation were used comparatively to keep the level of filtering constant between series rather than being viewed as an absolute measure of the percentage of the variance accounted for by the series following the removal of high-dimensional noise.

[2] The TISEAN software is available from:

http://www.mpipks-dresden.mpg.de/~tisean/TISEAN_2.0/index.html

[3] For our relatively long series, the Grassberger and Procaccia D2 estimate produced good results. It gave essentially the same results as more recently developed algorithms for shorter series, such as *P*D2i (Schiff, Sauer & Chang, 1994).