Serial correlations and 1/f power spectra in visual search reaction times

William McIlhagga

Vision Science Research Group, School of Life Sciences, University of Bradford, Bradford, UK

In a visual search experiment, the subject must find a target item hidden in a display of other items, and their performance is measured by their reaction time (RT). Here I look at how visual search reaction times are correlated with past reaction times. Target-absent RTs (i.e., RTs to displays that have no target) are strongly correlated with past target-absent RTs and, treated as a time series, have a 1/f power spectrum. Target-present RTs, on the other hand, are effectively uncorrelated with past RTs. A model for visual search is presented which generates search RTs with this pattern of correlations and power spectra. In the model, search is conducted by matching search items up with “categorizers,” which take a certain time to categorize each item as target or distractor; the RT is the sum of categorization times. The categorizers are drawn at random from a pool of active categorizers. After each search, some of the categorizers in the active pool are replaced with categorizers drawn from a larger population of unused categorizers. The categorizers that are not replaced are responsible for the RT correlations and the 1/f power spectrum.

Keywords: search, attention, reaction time, serial correlation, 1/f power spectrum, modeling


Introduction

Visual search is a laboratory abstraction of the common task of looking for something. In a visual search experiment, the subject looks for a particular target item hidden in a display of other items called distractors. In half of the displays, the target is absent. The subject’s task is to decide, as quickly and accurately as possible, whether the target is present or absent. Performance in visual search is usually measured by the speed of this decision i.e., the decision reaction time (RT). Search RTs depend on a number of factors, but the most important are the number of distractors, the presence of unique features in the target (Treisman & Gelade, 1980) and the similarity of target and distractor (Duncan & Humphreys, 1989).

During a visual search experiment, the subject’s RT varies from trial to trial. Some of this variation may be due to the randomness of the visual search stimuli—that is, extrinsic randomness—but some of it is undoubtedly due to intrinsic randomness in the visual processes involved in search. Because of this randomness, search experiments are repeated a large number of times and the RTs averaged to get a more accurate measure of performance.

The implicit assumption behind this averaging is that the actual sequence of RTs tells us nothing interesting about the processes of visual search. This is not the case. Chun and Wolfe (1996) noticed that RTs tended to be faster than average just prior to an error, and slower than average afterward. From this finding, they suggested that visual search is terminated by an adaptive process that speeds up search (and thus reduces accuracy) until a mistake is made, after which it slows search speed down to maintain an acceptable level of accuracy. Similar post-error slowing has been observed in other tasks (e.g., Brewer & Smith, 1989; Rabbitt, 1966; Rabbitt & Rodgers, 1977).

More interesting perhaps is the discovery that visual search RTs, when considered as a time series, have a 1/f power spectrum (Farrell, Wagenmakers, & Ratcliff, 2006; Gilden, 2001; Gilden, Thornton, & Mallon, 1995). The power spectrum of a time series is calculated by treating the sequence of RTs as an evenly spaced set of signal values, and taking the Fourier transform of them. The power spectrum of a time series is interesting for two reasons. First, 1/f power spectra are found in many diverse phenomena, such as the fluctuation of light intensity from quasars, current noise in resistors, semiconductors, and thermionic tubes, sea level fluctuations, music, intervals between heartbeats, and errors in time interval estimation, among others (Dutta & Horn, 1981; Milotti, 2002; Press, 1978; Voss & Clarke, 1975). The widespread occurrence of 1/f power spectra, and power spectra close to 1/f, suggests that there may be a universal mechanism behind all these phenomena. The second reason 1/f spectra are interesting is because, despite many attempts, no one has so far convincingly put forward any such universal mechanism.

The intriguing nature of 1/f power spectra can be appreciated by comparing 1/f power spectra to two other spectra which we do know the causes of. If visual search RTs were independent of each other, their power spectrum would be flat (i.e., “white noise”), and the process generating RTs would have no memory for past RTs. On
the other hand, if the RTs were generated by a random walk process, moving up or down by random amounts each trial, then they would show a $1/f^2$ power spectrum (“Brownian noise”). In this case, the process retains everything that happened to it in the past. A $1/f$ power spectrum is intermediate between the flat and the Brownian spectra, suggesting that some aspects of past RTs are retained for a very long time, while other aspects decay quickly.

The $1/f$ power spectrum discovered in visual search RTs was, however, found in a rather atypical search task: Gilden et al. (1995) used a complete report task where the subject had to say how many targets were present, from zero to four, in a display that had only four items. The $1/f$ power spectrum was found in the sequence of RT residuals, obtained by subtracting the mean RT for the particular search display (zero to four items) from the RTs. It is not clear whether the results obtained in this search experiment generalize to RTs obtained from ordinary search; in particular, whether they have a $1/f$ power spectrum.

In this paper I analyze the sequence of RTs obtained from visual search tasks with up to 100 items. Three types of search task were used. First, search in which the target is distinguished by a unique feature. In these “feature searches” the mean RT is independent of the number of features in the display. Second, search in which the target is distinguished by the absence of a specific feature, and in this case the RT increases linearly with the number of items. And third, a search in which the target is distinguished by the conjunction of features, where again the RT increases with the number of features.

In these standard visual search experiments, displays with a target and displays without a target are randomly interleaved, and the sequence of RTs obtained is therefore a random interleaving of a subsequence of RTs to target-present displays, and a subsequence of RTs to target-absent displays. These subsequences can be analyzed in two ways. First, one can compute the correlation between the current RT and past RTs. This measures, in a very direct way, the influence of past RTs on the present RT. Second, one can compute the power spectrum of the target-present RTs and target-absent RTs separately. This allows us to see if the power spectrum is $1/f$, or not.

In these experiments it was found that target-absent RTs (that is, RTs to displays where the target is missing) are strongly correlated with past RTs up to ten trials in the past. The correlation also increases with the mean RT. The target-absent RTs also have a $1/f$ power spectrum. On the other hand, the target-present RTs have little correlation with past RTs, and their power spectrum is nearly flat, so target-present RTs appear to be similar to a white-noise process.

One possible explanation for these results is that the target-absent searches are terminated by an adaptive deadline whereas the target-present searches end when the target is found. If the deadline changes over time with a $1/f$ power spectrum, this would account for the difference between target-present and target-absent power spectra. This deadline hypothesis was tested in a second experiment, in which searches of very different speed were interleaved, and it cannot account for the results obtained. Thus the correlations found between successive target-absent RTs cannot be explained by a deadline model of search.

What, then, does cause the correlations, and why should those correlations generate a $1/f$ power spectrum? The last part of this paper describes a model which answers these questions. In this model, display items are serially examined by a pool of categorizers, which categorize them as either target or distractor. Search ends when either the target item is found or no items remain (that is, Serial Self-Terminating Search, or SSTS). The categorizers in the pool have a distribution of speeds. After each search has ended, the pool of categorizers is refreshed by discarding the slowest one, and discarding the fastest one (which, by the speed-accuracy tradeoff, is the least accurate), and replacing them by selecting from a large population of dormant categorizers. Those categorizers which stay in the active pool for a long time are responsible for the correlations between RTs. Computer simulations show that the sequence of RTs generated by this model have a $1/f$ power spectrum, and the pattern of correlations between RTs closely parallels the pattern found in real visual search RTs. The exact reason why this model produces a $1/f$ power spectrum is not known, but may be because the refresh process produces a wide range of residence times for the categorizers in the active pool, with those categorizers whose speed is close to the median speed staying in the pool longer than those whose speed is far from it. Matlab code to generate $1/f$ power spectra via a simplified version of the model is included as Appendix A.

### Analysis of sequences of visual search RTs

Reaction times in a visual search experiment are collected in sequence, and once one has a sequence of data over time, it is reasonable to ask whether the sequence itself matters. Analysis of sequences of data is the domain of time series analysis (Chatfield, 1996). There are many ways of analyzing time series. Here, two methods will be used: serial correlations and spectral analysis. These methods, and the adjustments necessary to apply them to visual search RT sequences, will be described in this section.

Before beginning, it should be noted that the RT sequences collected here are not strictly speaking time series, since these are normally considered to be data collected at regular intervals. There is, however, no formal
difficulty in applying time series methods to merely ordered (rather than regularly spaced) sequences, and the methods will be valid if the data come from an event-based process (that is, a process that changes as events occur, rather than over continuous time), and approximately correct if the process is time-based, since visual search trials typically occur at approximately regular intervals.

Serial correlations

A standard method of analyzing a sequence of data is to compute its serial correlation, otherwise known as the autocorrelation. The autocorrelation of a sequence at lag $t$ is the correlation between all data in the sequence that are $t$ timesteps apart (Chatfield, 1996). The autocorrelation can be visualized by imagining the sequence of data duplicated, then shifted, and computing the correlation between all pairs of data where the two sequences overlap (see Figure 1a). Note that the shifted series is in the past relative to the unshifted series, because the subscript increases with time. If each item in the sequence is independent, the expected serial correlation is zero.

A zero serial correlation does not, however, imply independence from past events. The responses in the sequence could for example be dependent on the stimulus history, which is called a sequence effect (Kirby, 1980; Kornblum, 1973). With an appropriately balanced design, the reactions would end up being uncorrelated, because the stimulus presentations are uncorrelated. Sequence effects can be detected by conditioning mean responses on stimulus history rather than response history. Sequence effects, while an active area of research, are not the focus of this study, and were not analyzed.

Returning to serial correlations, the autocorrelation described above is designed to be applied to sequences where all the data are of the same type (e.g. temperature records, Dow-Jones indices, etc.). In visual search, however, responses are of two types, target-present and target-absent, which are randomly interleaved. For a sequence composed of interleaved target-present and target-absent RTs, the usual autocorrelation is inappropriate, and a slightly different definition of serial correlation is required.

The kind of serial correlation needed for a heterogeneous sequence is diagrammed in Figure 1b. The heterogeneous sequence, where $x_i$ represents one kind of response and $o_i$ a different kind, is copied and shifted, and all ordered pairs of the same type are identified; in this case where the first item in the pair is of type $x$ (unshifted) and the second item in the pair is of response type $o$, from the shifted series. The correlation is computed between the selected pairs.

Because visual search has two types of response, target-present and target-absent, there are four heterogeneous serial correlations:

1. between target-present and past target-present RTs,
2. between target-present and past target-absent RTs,
3. between target-absent and past target-present RTs,
4. between target-absent and past target-absent RTs.

The classification of responses into target-present and target-absent leaves out error responses, but the classification of error responses as either target-present or target-absent (or indeed leaving them out completely) was found to have little effect on the computed serial correlations.

All of these correlations can be computed at different lags (that is, different shifts of one sequence relative to the other), in order to see how long-lasting the correlations are.

Spectral analysis

Another method of analyzing time series is to estimate the power spectrum. Standard spectral analysis is most suited to series where all the data is of the same type, regularly spaced in time, so again some adjustments are needed to use spectral analysis on the heterogeneous sequence of target-present and target-absent RTs. The heterogeneous RT sequence can be thought of as two randomly interleaved subsequences, those due to target-absent responses and those due to target-present responses (Figure 2a). If either of these subsequences is considered in isolation (Figure 2b) it appears as a sequence with gaps in it. The ‘gaps’ can be thought of as data that was not measured, and the subsequence can be analyzed by methods designed to work with missing data or with data that is not regularly spaced in time.

Two methods of spectral analysis for sequences with missing data are either to fill in the gaps by interpolation, then do conventional spectral analysis, or preferably to use the Lomb–Scargle periodogram. The Lomb–Scargle periodogram is a least-squares estimate of the power of
sinusoidal components in the series. (For a good description of how it works see Press, Teukolsky, Vetterling, & Flannery, 1992.) The Lomb–Scargle periodogram was used in the spectral analyses presented here [using a Matlab implementation by Clifford (2004)]. It should be noted that the terminology surrounding spectral analysis can be ambiguous. In astronomy, where the Lomb–Scargle method was developed, the periodogram is taken to mean the power spectrum (Scargle, 1982), and that is the usage adopted here; this usage is different in other fields where the periodogram is often taken to mean the amplitude spectrum, which is the square root of the power spectrum.

Standard spectral analysis cannot tell us anything about the influence of one kind of response (e.g. target-present) on another (e.g. target-absent), unlike the serial correlations. There is a version of spectral analysis called the cross-spectrum which can, but it was not used because of the computational difficulty of computing the cross-spectrum between two sequences with missing data. Thus the spectral analysis presented here was only used on the target-present and target-absent RT subsequences, and misses any influence of one sequence on the other.

Experiment 1: Serial correlations

Methods

Visual search experiments were carried out to collect a set of RT sequences for analysis. Search was of three types.

1. Search for a cross among slanted distractors (Figure 3, left panel). This is a “feature search,” where the target has a feature that the distractors lack. Typically this kind of search is very fast, and the rate found here was 2.2 ms/item;
2. Search for a cross among slanted distractors, with a mixture of slants (Figure 3, middle panel). As is typical with such “within-dimension conjunction” searches (Wolfe et al., 1990), the search rate was fast: in this case 6.7 ms/item;
3. Search for a slanted line among cross shaped distractors (Figure 3, right panel), which yielded a slower search rate of 13.8 ms/item (Treisman & Gelade, 1980). This will be called “missing-feature” search because the target is missing a feature that the distractors have.

The experiment was programmed in Matlab using custom display software, and run on a workstation with

Figure 2. Spectral analysis of the RT time series. (a) Target-present RTs are marked as blue squares, target-absent as black circles. This sequence cannot be analyzed spectrally, as the RTs are of two different types, randomly interleaved. (b) Removing the target-present RTs gives a sequence of purely target-absent RTs, with gaps. Lomb–Scargle analysis can be used on this sequence to estimate the power spectrum.

Figure 3. Kinds of search task used. On the left is feature search, in the middle “within-dimension” conjunction search, and on the right missing-feature (sometimes called serial) search. Left and right searches show a classic search asymmetry for the slanted line feature, present in the target on the left but absent in the target on the right. The blobs on the items were intended to cover any potential intersection feature where the lines crossed.
an LCD display. Reaction times were measured using the Windows XP multimedia timer function `timeGetTime`, which has millisecond granularity when appropriately initialized with the Windows XP function `timeBeginPeriod`. Data was collected in blocks of 100 trials (Experiment 1) or 120 trials (Experiment 2, below). Each block consisted entirely of searches of the same number of items (20, 60, or 100 items), and search type (feature search, conjunction, or missing-feature search). Each combination of number of items and search type was repeated three times by every subject. Thus each block of 100 trials can be uniquely identified by the combination of subject \( \times \) search type \( \times \) set size \( \times \) repetition. In the Results section below, the data is averaged across blocks in various ways. For example, if a data point can be identified by the combination of search type \( \times \) set size, then this data point comes from an average across subject and repetition.

Trials were initiated by the computer. The search display was shown and remained on screen until the subject responded target-present or target-absent by clicking the left or right mouse button. The search display was removed when the subject responded and feedback was given only if the subject made an error. After a short delay (400 ms), the next trial started. Subjects were shown the target before each block. Five subjects (the author and four naïve paid volunteers) participated in the experiment. Experiments were approved by the Ethics Sub-Committee for Human Subjects at Bradford University. Consent was obtained from the volunteers.

Results

Serial correlations in RT

Serial correlations within and between target-present and target-absent RTs were calculated for each experimental block of 100 trials using the methods described in the previous section, then averaged across set size (20, 60, and 100 items), subject, and repetition. Separate averages were maintained for the different search types (feature, conjunction, or missing-feature). Error responses (typically 5% of responses, mostly misses) were classified according to the type of reaction, so that a miss is classed as a target-absent reaction. As stated earlier, the classification of errors has a minimal effect on the results.

Figure 4 shows the average serial correlations obtained for different search types at lags of up to 10 trials. Starred

![Figure 4](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933529/)
data points indicate significant correlations (5% level, Bonferroni corrected for 120 multiple comparisons using a critical value of 3.53). Correlations at lag 0 are not plotted, because they are either 1 (for target-absent:target-absent and target-present:target-present correlations) or undefined (for the other pairs, since it isn’t possible for a trial at lag 0 to be simultaneously target-absent and target-present).

Consider the target-absent RTs first (Figures 4a and 4b). There are strong and significant correlations between target-absent RTs and previous target-absent RTs up to 8 trials in the past (Figure 4a). The correlations are highest for missing-feature search, which has the longest RT, and lowest for feature search. The serial correlations between target-absent RTs and previous target-present RTs are rather lower (Figure 4b), but still sometimes significant. It is notable that none of the thirty-one nonsignificant correlations in Figures 4a and 4b between target-absent RTs and any past RTs are negative, implying that many more of them are significant than have been indicated by the highly conservative Bonferroni correction. These data show that target-absent RTs are strongly influenced by previous reaction times.

The picture with target-present RTs is somewhat different. Neither the serial correlations between target-present RTs and past target-present RTs (Figure 4c), nor those between target-present RTs and past target-absent RTs (Figure 4d) are particularly impressive and few are significant. Again, however, few of the correlations are negative, implying that there are more significant points than indicated by the Bonferroni corrected tests.

The pattern of correlations in Figure 4a suggests that the longer the RT, the stronger the serial correlation, since missing-feature search has the longest RT and the strongest correlation. This is confirmed by plotting the serial correlation at lag 1 between successive target-absent RTs against the mean RT (Figure 5). In this figure, mean RTs and lag 1 serial correlations are averaged over subject and repetition, and these block-averages were then plotted against each other. There is a clear trend in the averages, but the individual data show a wide range of correlations for the shorter RTs.

**Power spectra of RT time series**

The power spectra of the data was calculated as follows. Each sequence of RTs from a single experimental block was separated into two subsequences of target-present and target-absent RTs, as described previously. A Lomb–Scargle periodogram was calculated for the two subsequences separately. The resultant power spectra were averaged across all combinations of subject $\times$ search type $\times$ set size $\times$ repetition, except for feature searches, which were left out of the analysis because of their generally flatter spectra. A geometric average was chosen because there were large multiplicative differences between the spectra of different subjects. The spectra were averaged across search type because the differences between the power spectra of missing-feature and conjunction search were fairly small.

At the lowest frequencies, the target-absent RTs are similar to a $1/f$ power spectrum (Figure 6), but there is a lot of white noise which obscures this at frequencies...
above about 5. This transition from 1/f to white noise is similar to the spectra found by Thornton and Gilden (2005). The target-present RT sequence is more similar to white noise, consistent with the small serial correlations shown in Figures 4c and 4d.

The 1/f power spectrum shows that the target-absent RT series is being generated by a process with a long memory. However, the serial correlations (Figure 4a) appear to die out after about 10 trials. There is, however, no real contradiction here. Provided the serial correlations do not tend to zero too quickly, the amount of correlation between an RT and the aggregate of RTs some time in the past can be quite high, even if the correlation between an RT and a single RT some time in the past is not.

Individually, three subjects showed a clear 1/f power spectrum in the target-absent RTs. One of the subjects showed a spectra similar to 1/f^a, with a less than 1. The last subject showed a more Brownian-like spectrum (see supplementary material). The reason for the slightly different results from two of the subjects is unknown. They could be simply random deviations, since the entire power spectrum of these two subjects is highly variable. Possible structural reasons for the differences will be discussed in the section on modeling below.

### Experiment 2: Testing a deadline model

One way the above pattern of serial correlations might be produced is if search proceeds by examining items in turn until either the target is found, which yields a target-present response, or a deadline is passed, yielding a target-absent response. The deadline could be adaptively adjusted to task difficulty and this adaptive process could be responsible for the serial correlations.

The adaptive deadline hypothesis can be tested by interleaving searches of varying difficulty. Consider feature search (Figure 3, left panel), with a mean target-absent RT of 529 ms when there are 100 items, and conjunction search (Figure 3, middle panel) with a mean target-absent RT for of 825 ms for 100 items. Now consider what happens if these two kinds of search are randomly interleaved in a single block, so that each search trial is randomly either a conjunction search trial or a feature search trial. If target-absent RTs are caused by a single adaptive deadline, one would expect the mean target-absent RTs in the interleaved case to be the same regardless of whether the particular search trial was a conjunction search or a feature search, since it is controlled by the same deadline, and probably close to the average of 529 ms and 825 ms (i.e. 677 ms). The experiment in this section tests this hypothesis.

### Methods

RT data was collected in blocks of 120 search trials. A block of 120 search trials could be of three types: (a) pure conjunction searches (b) pure feature searches, or (c) a mixture of 60 conjunction and 60 feature searches, randomly interleaved. Displays always contained 100 items. In all of these searches, the target is the same (a cross), and whether the search is conjunction or feature-search is defined solely by the distractors (lines of a single slant for feature search, lines of mixed slant for conjunction search).

Each condition was repeated a number of times, and the sequence of RTs was recorded. Three subjects participated in these experiments, the author and two naïve observers. One of the naïve observers previously participated in Experiment 1. Experiments were approved by the Ethics Sub-Committee for Human Subjects at Bradford University. Consent was obtained from the volunteers. All other details were as for Experiment 1.

### Results

Mean target-present and target-absent RTs were calculated for feature and conjunction searches in both the pure (a) and (b) and mixed (c) conditions. These are shown in Figure 7, averaged across subjects (individual data similar). The mean target absent RT for the pure feature search was 529 ms (Figure 7, left hand green bar) and for

![Figure 7. Mean RTs for feature (feat) and conjunction (conj) searches. Error bars are 1 standard error. Red is mean target-absent RT, green is mean target-present RT, for searches with 100 items. "Pure feat" is the mean RT to feature searches. "Mix feat" is the mean RT to feature searches when they are intermixed with conjunction searches. "Mix conj" is the mean RT to conjunction searches when they are interleaved with feature searches. And "pure conj" is the mean RT to conjunction searches when they are not mixed with other searches.](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933529/)
the pure conjunction search was 825 ms (right hand green bar). The mean target-absent RTs for these searches in the interleaved blocks were 571 ms and 761 ms respectively (Figure 7, middle green bars), and are significantly different. However, if a single adaptive deadline controls the target-absent RTs, they should have been the same. These results are consistent with those of Chun and Wolfe (1996).

However, because the search type (feature or conjunction) could be identified by the kinds of distractors, subjects could, conceivably, switch between two separately maintained deadlines depending on the kind of search they are doing. If the serial correlations in target-absent RT were due to an adaptive deadline, and the mixed searches were controlled by two separate deadlines, we would expect serial correlations to be severely reduced in the mixed search conditions.

About 1/2 of successive pairs of target-absent RTs in the mixed condition both come from the same kind of search, either a feature search followed by another feature search, or conjunction followed by conjunction. It would be no surprise if these were similar to the serial correlations in the pure search condition. The most interesting serial correlation is where one of the target-absent RTs is from a conjunction search, and the preceding target-absent RT is from a feature search, or vice versa. These two RTs could be controlled by separate deadlines, and if this were the case we would not expect to see much correlation between them.

Two such mixed-search serial correlations were calculated. The first was between a target-absent RT in a conjunction trial and a previous target-absent RT in a feature search trial. This corresponds to the correlation at lag 1 in Figure 4a, but between two quite different kinds of search. The correlation, averaged over subjects and repetitions, was 0.33 ± 0.09. The complementary serial correlation, between a target-absent feature-search RT and a previous target-absent conjunction RT, was 0.18 ± 0.08. Both correlations are significant, and both are about the same value as the correlations in Figure 4a for their respective search types. Thus mixing the kind of search has little effect on the serial correlations, even though it greatly affects the RTs. This suggests that the serial correlations in target-absent RT are probably not due to one, or even several, adaptive deadlines.

Modeling of reaction times

There are many existing models for visual search; perhaps the most straightforward is serial self-terminating search (SSTS). In SSTS, display items are selected one at a time (without replacement) and categorized as target or distractor. If the currently selected item is a target, the search terminates. Otherwise, the search continues until no items remain to be examined. There are undeniably problems with SSTS (Horowitz & Wolfe, 1998), but even so it can account for a lot of the properties of visual search. In the SSTS model, the search RT is the sum of categorization times for all the items selected, plus a baseline time due to motor commands, retino-cortical transmission, and any other stimulus-independent delays. A variant of the purely serial one-at-a-time SSTS model is one with a limited parallelism (Houck & Hoffman, 1986; Pashler, 1987), in which a few items can be examined simultaneously. Provided the parallelism is far less than the number of items, limited parallel models function similarly to serial models.

If the SSTS model is used to explain search, where can the RT correlations come from? The RT is a sum of categorization times over all the items examined, so correlation may arise either from categorization times or from the number of items examined. However, in the SSTS model, the number of items examined is always uncorrelated from one trial to the next. When the target is absent the number of items examined is always the number of items in the display, which is constant, and hence has no correlation. When the target is present, the number of items examined is a random number between 1 and the number of items in the display, which does not depend on previous trials. Thus RT correlations can only arise from correlations between the categorization times in the SSTS model. Thus, to explain correlations, we must look at how categorizations might be accomplished.

A categorization is performed, naturally enough, by a categorizer, which might be thought of as some sort of “cell assembly” or “daemon.” The time needed for the categorizer to complete its work is a random variable, drawn from a probability distribution (e.g. exponential) with a given mean and variance. If there is only one kind of categorizer, and it does not change its characteristics over time, then categorization times are all independent, and the RTs are uncorrelated again.

Consider, then, what happens if there are multiple categorizers operating during visual search. Each time an item is selected, we might choose a categorizer at random from a pool of available categorizers and assign it to work on the selected item. When the categorization completes, another search item is chosen and another categorizer assigned to it, and the process continues until no items remain, or until the target is found. Depending on the number of items and the number of available categorizers, some of the categorizers may be reused many times during a search (However, none of the items is examined more than once). The time for the search to finish is again the sum of categorization times over the number of items examined. It is reasonable to assume that the categorizers in the pool all have different properties: some of them are fast and some are slow; some are unreliable and some are accurate. The properties of the categorizers in the pool determine the overall speed and accuracy of the visual search. Inevitably, some of the categorizers in the pool
will turn out to be unsuitable for the search task at hand, and unless they were removed from the pool would be detrimental to performance.

Suppose then that after every search we discard some categorizers from the active pool, and recruit some new ones. It makes sense to discard the slowest categorizer, because we want search to be fast. It also makes sense to discard the least accurate categorizer which, by the ubiquitous speed-accuracy tradeoff, is likely to be the fastest categorizer in the pool. The discarded categorizers are replaced by new ones, selected at random from a large dormant population of categorizers. They are selected at random because their suitability cannot be evaluated until they are actually used in a search.

This simple process of selecting categorizers is sufficient to generate correlations between RTs. Correlation arises from the similarity of the categorizer pool from one trial to the next. On one trial, the search time is a sum of a random sample of categorization times drawn from a particular categorizer pool. On the next trial, the search time is the sum of categorization times drawn from an almost identical pool (only two of the categorizers have changed). Thus the search times will tend to be similar. However, over larger time scales, the categorizer pool becomes less and less similar, and so the search times are also less similar. The pooling of categorizers thus produces a situation where, for a cluster of trials, the search times might be slightly longer than average because the pool is dominated by slow categorizers, while for another cluster of trials they might be shorter than average because of a preponderance of fast categorizers. This is sufficient to produce correlations between successive reaction times.

This process of selecting categorizers is also responsible for the 1/f power spectrum found in the RTs. This is demonstrated in computer simulations, described next. Intuitively, though, the cause of the 1/f power spectrum produced by this model can be identified. It is well-known that 1/f power spectra can be trivially produced by a system whose output is a sum of perturbations which have a wide variation in half-life or duration (Haldorff, 1968; Hausdorff & Peng, 1996). The problem with describing 1/f power spectra in this way is that there is no mechanism for the wide variation in duration of the events; it is simply assumed. In the pool of categorizers model proposed here, the mechanism for producing a wide range of durations is the selection process. Those categorizers whose categorization time is close to the pool median will tend to reside in the pool for a long time, whereas those categorizers whose categorization time is far from the median will tend to be removed rather quickly, if not immediately. Thus, the lifetime of the categorizer in the pool is determined by how close the categorizer speed is to the median speed in the pool.

The search model used in simulations (Figure 8) is specified by the following list of features. The reasoning behind each feature is, where warranted, given in square brackets.

1. There are \( N \) categorizers in the active pool, and an effectively infinite number in the dormant population. [The number of categorizers in the active pool has some effect on the power spectrum.]
2. Search is performed using a limited capacity parallel model. In this model, there are \( S \) memory ‘slots’, where \( S \) is a small number. If there is a free memory slot, an unexamined item is selected randomly from the display and placed in the free slot. Then an unused categorizer is selected from the active pool and assigned to categorizing the item. The slot is occupied until the categorizer finishes. The number of slots \( S \) must be smaller than the number of categorizers \( N \). [A pure serial model has \( S = 1 \). However, a serial model does not adequately model the coefficient of variation found in the RTs here. A limited capacity parallel model can. The number of slots \( S \) has little influence on the power spectrum.]
3. The RT is time from the start of search until the last categorizer finishes, plus a base reaction time \( B \).

Figure 8. Diagram of the search model. Items in the search display are categorized in a number of parallel processing slots by categorizers that are drawn from an active pool. Each processing slot marries up one search item with one categorizer, and the slot clears when the categorizer finishes. After each search, some categorizers in the active pool are returned to the dormant population, and others selected from it.
If one of the items is a target, the “last categorizer” is the one assigned to the target item, since once the target is found, ongoing categorizations in the other free slots can be ignored. [The base reaction time is the sum of fixed delays e.g. motor movements, and is responsible for the RT when there is only one item.]

4. The time for categorizer $i$ to complete its task is exponentially distributed with mean $m_i$. [The exponential distribution was chosen because it is skew and has been used successfully in other models of search, e.g. Bundesen (1990). This distribution has no effect on the $1/f$ power spectrum, but does affect the distribution of RTs; for example if the categorizer times instead had a Gaussian distribution, the overall RT would also be Gaussian.]

5. The mean categorizer times $m_i$ are also exponentially distributed with median $M$ and hence mean $M/\log(2)$. [The exponential distribution was used again to keep things simple. However, many positively valued distributions with the same median will work equally well. Some distribution of mean categorizer times is vital to generate $1/f$ power spectrum from the selection process.]

6. After each search, the pool of categorizers is refreshed in the following way: With probability $P$, the slowest categorizer (i.e. the one with the largest $m_i$) is replaced with one drawn randomly from the dormant pool. With probability $P$, the fastest categorizer is likewise replaced with one drawn randomly from the dormant pool. Note that this replacement process changes the distribution of categorizers in the pool, but the categorizer distribution will stabilize at some point. [The probability $P$ has an effect on the power spectrum.]

7. Search terminates when a categorizer finds the target, or all items have been examined.

The model thus has five parameters, $N, M, P, B$, and $S$. Parameters $M, B, S$ mostly affect the fit of the model to RT distributions, and parameters $N$ and $P$ mostly affect the RT correlations and the power spectrum. The model was only fitted to the “missing-feature” and “conjunction” experiments, since its essentially serial nature is a poor fit to the feature search experiments.

Fits to data were obtained by simulating visual search experiments using the model, then assessing how well the simulated data matched the real data. In a single block of 100 simulated trials, 50 were selected at random as target-present trials and the rest were target absent. In target-absent trials, the search continued until all the items were processed. In target-present trials, the number of items examined before the target was found was selected from a uniform distribution from 1 to the number of items. To fit both “missing-feature” and “conjunction” experiments with the same model, it was assumed that in a conjunction trial, the number of effective items was half the number of items in the search display. This might be because, say, half the distractors (those with one of the two slants present) can be filtered out preattentively (Egeth, Virzi, & Garbart, 1984).

Fitting the model to data is slow. The model is stochastic, and this means that the goodness of fit of the model, for a specific set of parameters, is a random variable. This precludes the use of canned optimization routines to find the best parameters. The model was therefore “fitted” to the data in Figures 4, 5, and 6 informally, where the parameters were varied by hand until a reasonable fit was obtained. The parameter values used were $N = 10, M = 56, P = 0.3, B = 460$, and $S = 4$.

Simulated results from this model have been shown in Figures 4, 5, and 6 above. In these figures, simulated data from the model is plotted alongside the real data. The solid lines in Figure 4 shows how closely the serial correlations from the simulated data match the measured correlations. The overall pattern of correlations is captured by the model. The model produces a strong correlation between target-absent RTs (Figure 4a) but a weak correlation between target-present RTs (Figure 4c). In addition, the model produces correlations between target-present and past target-absent RTs (Figure 4b). The difference between target-present and target-absent RT correlations (Figures 4a and 4c) is caused by the number of items examined in these two cases. In target-present searches, the number of items examined varies randomly from 1 to the number of items in the display. This variation makes it harder to see any correlation in categorizer processing times. Essentially, what is happening is that the target-present RT is produced by interrupting an exhaustive search of the display at some random time (i.e. when the target is found). The random interrupts remove almost all the correlations that would otherwise be apparent if search runs to completion, which occurs with the target-absent RTs.

Figure 5 shows how the simulated correlations scale with the mean RT. The lines in Figure 5 link point estimates of mean RT and correlation at three different set sizes: 20, 60, and 100. The model fits the trend in the data, but misses the fact that correlations are still quite high at short mean RTs. This suggests that there is an additional source of correlation not captured by the model.

Finally, the model power spectra are shown by thin solid lines in Figure 6. The model clearly produces a power spectrum which is close to $1/f$ at low frequencies, transitioning to white noise at higher frequencies. The transition to white noise is caused by the assumption that categorizer times are randomly distributed, since when categorizer times do not vary about their mean, a $1/f$ power spectrum holds to high frequencies (as in the code given in Appendix A). However, the transition to white noise at high frequencies could equally be caused by white noise in the base reaction time $B$ (Gilden, 2001).

It was noted earlier that not all subjects showed a $1/f$ power spectrum in the target-absent RTs, though all showed some correlation. The above model may be able to produce power spectra different from $1/f$. If for example...
the replacement probability \( P \) is reduced, the pool of categorizers changes less between trials and more correlation is observed, yielding a power spectrum steeper than \( 1/f \). If the number of categorizers \( N \) in the pool is reduced, a greater proportion get replaced after each trial, reducing the correlation between successive RTs, and whitening the power spectrum.

**Discussion**

I have shown here, as have others previously, that visual search RTs are dependent on past RTs. The dependency is clearly revealed in target-absent RT time series, but is not evident in target-present RTs because target-present RTs are cut short at a random moment (i.e. when the target is found). Target-absent RT searches on the other hand are allowed to continue to completion, so revealing the full extent of intertrial correlations. Not only are visual search RTs correlated with past RTs, but this correlation takes on a specific form: the target-absent RTs, considered as a time series, have a \( 1/f \) power spectrum.

A \( 1/f \) power spectrum is frequently accounted for by supposing that the system producing the \( 1/f \) power spectrum is the result of a sum of processes which have a wide range of correlation decay times (Halford, 1968; Hausdorff & Peng, 1996; Press, 1978). If the amplitude and lifetimes of the component processes are properly chosen, the power spectrum of the resultant system is \( 1/f \) over a wide range of frequencies. Explanations of this type have been invoked to account for \( 1/f \) spectra in human cognition (Hausdorff & Peng, 1996; Wagenmakers, Farrell, & Ratcliff, 2004). However, this kind of explanation just “transfers the mystery” (Press, 1978) of \( 1/f \) spectra to the hypothesized wide range of lifetimes of the component processes. If this kind of explanation is accepted, one must find a cause for the postulated wide range of lifetimes. Extremal dynamics (Miller, Miller, & McWhorter, 1992) is one method for generating processes or perturbations with a sufficiently wide range of lifetimes. Extremal dynamics is, however, rather abstract and it isn’t clear how it can be applied to the case of visual search. The categorizer replacement model suggested here is another way of generating a system that is a sum of processes with a wide range of lifetimes. In this model, the lifetimes of the categorizers are dependent on how similar the categorizer speed is to the median speed, because selection quickly replaces categorizers whose speed is far from the median.

There are alternative models for \( 1/f \) power spectra. Self-organized criticality (Bak, Tang, & Wiesenfeld, 1987) is an influential idea in statistical physics which was claimed to generate \( 1/f \) dynamics, though in fact it may not (Jensen, Christensen, & Fogedby, 1989). Van Orden, Holden, and Turvey (2003) have suggested a variety of processes that together might give rise to \( 1/f \) power spectra in RT series, but these suggestions are more descriptive than explanatory.

There are two notable omissions in the model. The first is the process of selecting categorizers from the dormant population. For simplicity, it was assumed that all the categorizers in the dormant population have the same distribution, so they can be selected at random, but this is unsatisfactory. Undoubtedly, most of the categorizers in the dormant pool will be totally irrelevant to the search, being instead designed to categorize grandmothers, VWs, and clouds, among other things, and the selection process needs some way of avoiding them if possible, and choosing categorizers that are more likely to be relevant to the task of distinguishing crosses from slanted lines.

The second omission is the complete lack of errors in the model. Even though some categorizers were assumed to be inaccurate, their resultant categorization was never actually used in the model. In general, we would expect a categorizer to return as its output the ratio of the likelihoods that an item is the target or distractor. If the likelihood ratio is too low, the item might need to be examined a second time by another categorizer, until the likelihood ratio reaches a level sufficient for a good decision. In addition, the likelihood ratio being reported by categorizers might occasionally be flat-out wrong. If the model is adjusted to create errors occasionally, what should happen when an error is made? Possibly, after an error it might be a good strategy to replace a larger number of categorizers in the active pool, as the error has indicated that the pool of categorizers contains too many unreliable units. If a large number of categorizers are replaced, this may lead to a sudden increase in RT, much as observed by Brewer and Smith (1989), Chun and Wolfe (1996), and Rabbitt (1966).

Finally, given that \( 1/f \) power spectra are observed in many different places, not just visual search RTs, what implications does the model presented here have for these other \( 1/f \) processes? For some other psychological processes, a similar model could be constructed. For example, to match time intervals (Gilden et al., 1995), one could imagine a pool of interval estimators, where the extreme estimators are replaced at regular intervals. The particular time interval produced by the subject would be a summary statistic (perhaps a sample mean) drawn from the estimator pool. This model, being functionally identical to the visual search model suggested here, would readily generate a \( 1/f \) power spectrum.

However, for \( 1/f \) power spectra in other domains, the model is less likely to be directly applicable. Nonetheless, one may discern a general principle which is embodied in the model. In the visual search model, the output is related to the sum of the speeds of the categorizers in the active pool. The lifetime of the categorizer in the pool is related to its speed: extreme speeds do not last as long as speeds close to the median. In a more general setting, the lifetime of a process could be related to its magnitude, with
processes of small magnitude lasting longer than those with large or extreme magnitudes.

**Appendix A**

The model given in the text is rather complex, and some of the complexities serve to obscure the way the model generates a $1/f$ power spectrum. Here, I give Matlab code for a simplified version of the model which produces a very clean $1/f$ power spectrum from the lowest to very high frequencies. For the given parameter set, $1/f$ power spectra result. The $1/f$ power spectrum is stable over a wide range of parameter values, but different spectra can result if the parameter values are changed markedly from those given. Small values of $n$ (around 5) or large values of trim (trim > $n/4$) generate spectra closer to white noise. To obtain a Brownian power spectrum, reduce the value of pr_replace to approx. 0.05 or less and/or increase $n$ to more than 100.

```matlab
function oneoverf
% oneoverf generates a 1/f spectrum. It may take some time on an average computer.
% Reduce reps to increase speed.

n = 25;  % number of categorizers in the active pool.
dur = 1024+300;  % duration of simulated block of trials + 300 step run in
reps = 500;  % number of repetitions to average over
nitems = 100;  % number of search items
m = 12;  % the median search speed for the categorizers
trim = 1;  % how many items to replace in the pool
pr_replace = 1;  % probability of replacing items

RT = zeros(reps,dur);  % stores the RT series

for r=1:reps
  % Initialize the active pool with categorizer speeds. These are drawn from
  % a uniform distribution, but other distributions will do too.
pool = 2*m*rand(1,n);
  for i=1:dur
    % idx says which categorizer (1..n) is applied to which item (1..nitems)
    % using a purely serial search process.
    idx = ceil(rand(1,nitems)*n);
    % The search time is the sum of assigned categorizer times.
    % Here all items are examined, which implies the target is always absent.
    % The model in the paper differs in that it generates random categorizer times
    % from a distribution with mean given by pool(idx).
    RT(r,i) = sum(pool(idx));
    % Replace the slowest & fastest categorizers with new ones from
    % the population of dormant categorizers
    if (rand<pr_replace)
      pool = sort(pool);
      pool([1:trim end-trim+1:end])=2*m*rand(1,2*trim);
    end
  end

  % work out the power spectra of each RT row, leaving out the 300 step run-in at the start
  p = abs(fft(RT(:,301:end),[],2)).^2;
  % average across rows
  p = mean(p,1);
  % trim the power spectra to have just positive frequencies
  n = length(p);
  p = p(2:floor(n/2));
  % plot the results
  n = length(p);
  loglog(1:n,p)
  hold on
  xlabel('Frequency')
  ylabel('Power')
  % plot the 1/f and 1/f^2 lines for reference
  plot(1:n,p(1)./(1:n).',':)
  plot(1:n,p(1)./(1:n).^2,:')
```

Downloaded From: https://jov.arvojournals.org/pdfaccess.ashx?url=data/journals/jov/933529/ on 01/14/2019
Acknowledgments

Commercial relationships: none.
Corresponding author: William McIlhagga.
Email: w.h.mcilhagga@bradford.ac.uk
Address: Vision Science Research Group, School of Life Sciences, University of Bradford, Bradford BD7 1DP, UK.

References


