THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2009, 62 (3), 585–604

Psychology Press
Taylor & Francis Group

Detrended windowed (lag one) autocorrelation: A new method for distinguishing between event-based and emergent timing

Loïc Lemoine and Didier Delignières
EA 2991, University Montpellier I, Montpellier, France

The aim of this study was to test different methods for distinguishing between two known timing processes involved in human rhythmic behaviours. We examined the implementation of two approaches used in the literature: the high-frequency slope of the power spectrum and the lag one value of the autocorrelation function, ACF(1). We developed another method based on the Wing and Kristofferson (1973a) model and the predicted negative ACF(1) for event-based series: the detrended windowed (lag one) autocorrelation (DWA). We compared the reliability and performance of these three methods on simulation and experimental series. DWA gave the best results, and guidelines are given for its appropriate use for identifying underlying timing processes.

Keywords: Timers; Autocorrelation; Spectral analysis; Event-based; Emergent.

Timing has long been considered as a general-purpose ability (Ivry & Hazeltine, 1995). This idea is supported by a number of correlation studies, showing that timing variability is significantly correlated across a range of different rhythmic tasks. Keele, Pokorny, and Ivry (1987) showed, for example, significant correlations in timing precision between tapping with the index finger and tapping with the forearm. Keele, Pokorny, Corcos, and Ivry (1985) observed significant correlations between time estimation and time production in a tapping task. These results suggested the presence of a common timing process, a fundamental timing ability that could be shared by a variety of tasks. The simplest task for studying timing processes is the index finger tapping task. Two different paradigms have been considered: the continuation and the synchronization tapping tasks. In the former participants have to tap as regularly as possible following an initially prescribed tempo, whereas in the synchronization task participants have to synchronize their tap with a metronome. The main variables studied are the intertap intervals in continuation and the series of asynchronies between the metronome and the associated tap in synchronization.

Wing and Kristofferson (1973a, 1973b) proposed a model (W-K model) to account for timing variability in continuation tapping tasks, which supposed that the production of each interval was based on two independent processes: an internal clock, providing a series of temporal

Correspondence should be addressed to Loïc Lemoine, EA 2991 Motor Efficiency and Deficiency, Faculty of Sports and Physical Education, University Montpellier I, 700, avenue du Pic Saint Loup, 34090 Montpellier, France. Email: lemoine.l@free.fr

http://www.psypress.com/qjep
DOI:10.1080/17470210802131896

intervals C_i, and a motor component, responsible for the execution of the tap at the expiration of the interval C_i. This motor component does not operate instantaneously, and all taps have an associated motor delay, M_i. In terms of these two components, the observed I_i interval is written as:

$$I_i = C_i + M_{i+1} - M_i \tag{1}$$

In the original model, C_i and M_i were assumed to be independent random variables, each identically distributed over the index i (i.e., stationary). In the frequency domain a stationary process gives rise to fluctuations with equal power across all frequencies and is termed white noise (Eke et al., 2000).

Based on the hypothesis that the clock (C) and motor (M) components are stationary independent random variables, Wing and Kristofferson (1973a) demonstrated that the clock and motor variances could be estimated as follows:

$$\sigma_D^2 = -\gamma_I(1) \tag{2}$$

$$\sigma_C^2 = \gamma_I(0) - 2\sigma_D^2 \tag{3}$$

where σ_C^2 and σ_D^2 represent clock and motor variances, respectively, and $\gamma_I(x)$ the lag(x) autocovariance of I series. However, several studies noted problems with these estimators with negative motor variance calculated in continuation tapping series (Collier & Ogden, 2004; Ivry & Corcos, 1993; Wing & Vorberg, 1996). Vorberg and Hambuch (1996) suggested that these anomalous estimates arose from drift in the series of intertap intervals and advocated the use of short series, combined with extended practice, to avoid such problems. In contrast, Collier and Ogden (2004) considered the drift to be of interest in its own right and recommended partitioning the total variance of continuation tapping series in three components: the clock and motor variances of the W-K model, plus the drift variance. They applied their model to simulation and experimental series. In addition to giving a separate estimate of the drift variance component, they concluded that even in the absence of drift, the use of long series (150 data points) rather than short series (30 data points) resulted in estimators for their model that were equal to or better than those for the W-K model.

Vorberg and Wing (1996) proposed, on the basis of the W-K model, another model to account for timing variability in asynchrony series collected during synchronization tapping tasks. An autoregressive term was added to the W-K model in order to account for feedback correction as following:

$$A_{i+1} = (1 - \alpha)A_i + I_i - \tau \tag{4}$$

where A represents asynchrony, α is the autoregressive parameter, I the time intervals produced via the W-K model, and τ a constant corresponding to time intervals prescribed by the metronome.

Pressing and Jolley-Rogers (1997) proposed a similar model but also a method based on the covariance in order to estimate the variance of the clock and motor delay components, as proposed by Wing and Kristofferson (1973a) for continuation tapping. In addition, they proposed using the power spectral properties of series (power spectra plotted in log−log coordinates) in order to estimate the variance of components of the model and to determine the model that best fits the data. Spectral analysis is then applied on event series of asynchronies resulting in spectra representing the squared amplitude of fluctuations over the length of these fluctuations (i.e., the frequency of these fluctuations) in log−log coordinates. The highest frequencies in spectra represent the fluctuations over the adjacent asynchronies, whereas the lowest represent the fluctuations over the half of the length of the series (see Figure 1). This spectral approach has the advantage of graphically separating estimation of the clock drift (low-frequency shape) and motor delay (high-frequency shape). This separation of model components allows determination of the exact nature of fluctuations of each (i.e., short-term correlations, long-term correlations, or uncorrelated fluctuations) and consequently leads to a better understanding of timing mechanisms.

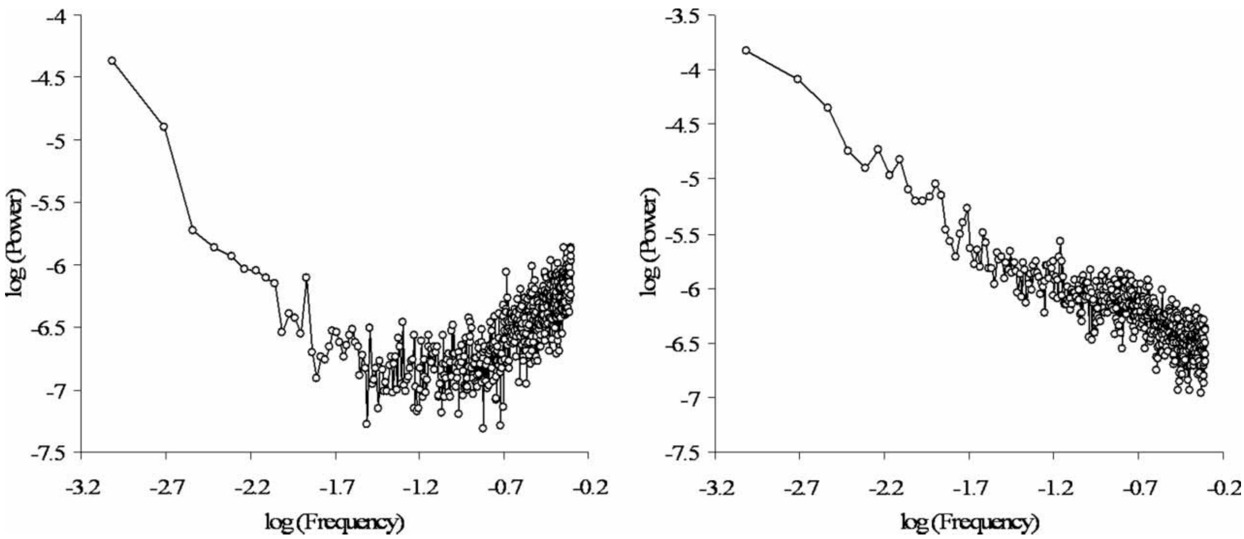

Figure 1. *Examples of averaged power spectra (in log–log coordinates), for the tapping task (left) and the oscillatory task (right) (as shown by Delignières, Lemoine, & Torre, 2004). As the series are event series and not time series, the frequencies in the abscissa axes refer to the spacing between data related to the series length. The highest frequency represents adjacent intervals, and the lowest frequency half the length of the series (Gilden et al., 1995).*

Recently, results have been reported that challenge the view that timing is a unitary general ability. For example, Robertson et al. (1999) analysed timing variability in finger tapping and in continuous circle drawing. They showed that individual performances in timing variability in the tapping task were not significantly correlated with individual performances in timing variability in continuous drawing. The authors suggested that these two tasks might involve distinct timing control processes. Finger tapping was considered as representative of a class of tasks where timing occurred through the operation of an "internal clock" in which the temporal goal was explicitly represented. This mode of timing control was then referred to as explicit. The key feature of this class of tasks could be related to their discrete nature (i.e., in contrast to a smooth movement the discreteness refers to a jerky movement)—for example, finger tapping appears as the concatenation of discrete movements, in response to cognitively generated time intervals. They contrasted this with another set of tasks that requires the movement to be made in a smooth, continuous manner. In this case, timing could be controlled on the basis of a different process, resulting from the operation of mechanically based timing—for instance, resonance of a mass-spring system whose period is controlled by nontemporal parameters such as muscle stiffness. The authors refer to this latter type of timing as implicit.

Schöner (2002) proposed on a more formal basis a similar distinction between event-based (explicit) and emergent (implicit) timers. The former are based on the hypothesis of "internal clocks" generating periodic timing events that trigger motor responses. Emergent timers, in contrast, exploit the limit cycle dynamics of effectors, considered as self-sustained oscillators. Within this limit cycle particular "anchoring" events (such as movement reversals) can be used to delimit time intervals. In order to establish the importance of the continuous/discontinuous distinction, Zelaznik, Spencer, and Ivry (2002) developed an intermittent circle drawing task, where participants were instructed to pause between each circling cycle. Despite the similarity in spatial trajectory of intermittent and continuous circle drawing, temporal variability on the intermittent drawing task correlated with finger tapping, and neither task correlated with continuous drawing. These results reinforce the distinction between two classes of timing processes: event-based timing used in discrete tasks and emergent timing used in continuous smooth movements.

However, the correlational approach used in the aforementioned studies (Robertson et al., 1999; Zelaznik et al., 2002) does not allow the exact nature of the underlying timing process for a particular experimental series to be determined. Ideally one would like to establish, by comparing two tasks, which of the two different modes of timing control is being used in a given situation. Delignières, Lemoine, and Torre (2004) proposed a method for distinguishing between event-based and emergent timing processes. They analysed the spectral properties of time interval series. As previously shown by Pressing and Jolley-Rogers (1997) a difference in the nature of fluctuations has repercussions for the shape of spectra. They applied the spectral analysis to time intervals series produced in continuation tapping or in rhythmic continuous oscillations of the forearm. The first situation was supposed to exploit an event-based timer and the second an emergent timer. They showed that each kind of timer was characterized by a specific signature in the log–log power spectrum (Figure 1). For tapping series, the log–log spectrum was characterized by a positive slope in the high-frequency region, whereas for oscillation series, the high-frequency slope remained negative. In both cases, the low-frequency slope presented a negative linear slope, close to −1, generally considered as revealing the presence of $1/f$ noise in the series. This $1/f$ noise represents a specific kind of variability, reflecting the presence of nonrandom variability and long-range correlations in the series. It is characterized by a linear decrease in variability with the decrease of time scales, contrary to white noise, which comprises random fluctuations uncorrelated in time with constant variability whatever the time scale of observation.

The log–log spectrum with positive slope at high frequencies for tapping has already been described by Gilden, Thornton, and Mallon (1995). These authors interpreted this result on the basis of the W-K model (Equation 1) According to Gilden et al. (1995), the positive slope observed in the high-frequency region of the log–log spectrum could arise from the presence of a differenced white noise process in the series ($M_{i+1} - M_i$). Their interpretation suggests that the clock component should be considered as a source of $1/f$ noise and not as a white noise process as originally hypothesized by Wing and Kristofferson (1973a).

The observation of a negative slope in the high-frequency region for rhythmic forearm oscillation was interpreted by Delignières et al. (2004) as the typical signature of an emergent timing process. According to the authors, the motor error in this kind of emergent timing process is distributed over the entire movement and directly affects time interval, and not the successive events that delimit the interval, as in the event W-K model. The macroscopic difference between both models is that the two motor delay components are replaced by a single motor error component. The simplest formulation of such an emergent model should read as follows:

$$I_i = D_i + \xi_i \tag{5}$$

where D_i represents the duration of cycle i, and ξ_i a Gaussian white noise term representing the motor error that affects the time interval. The presence of this white noise term is revealed by a simple flattening of the slope in the high-frequency region of the power spectrum. Consequently, the linear negative slope present in the low-frequency region gives an account of the temporal component (D) of the model and denotes the presence of $1/f$ fluctuations as with the temporal component of the W-K model (C).

Determining the nature of the timing process used in each task condition during the continuation phases is one approach (Delignières et al., 2004; Robertson et al., 1999; Zelaznik et al., 2002) that differed from previous work that determined the variance of each component in the timing models (Collier & Ogden, 2004; Pressing & Jolley-Rogers, 1997; Wing & Kristofferson, 1973a). However, the method developed by Delignières et al. (2004) allows the identification of the nature of the timing process underlying a particular series. This capacity to distinguish between timers on the basis of the statistical properties of experimental data is important, since some recent studies have shown that the exploitation of event-based timers in discrete tasks and emergent timers in continuous tasks is not systematic (Delignières et al., 2004; Lemoine, Torre, & Delignières, 2006). A better understanding of the factors that induce the exploitation of each kind of timer is necessary and requires further experimental effort. Therefore, a basis for distinguishing timing mechanisms underlying each series is required in developing an understanding of human timing behaviour.

A criticism that may be levelled against spectral analysis is the need for extended series because of the bias and high variability of estimators based on short series (Delignières et al., 2006). Power spectral analysis is known to give reliable results with a series of at least 2^9 or 2^{10} data points (Delignières et al., 2006; Delignières, Torre, & Lemoine, 2005; Eke et al., 2000). Obtaining such extended series in psychological experiments requires very prolonged trials (for example, 15 minutes in tapping for a frequency condition of 1.25 Hz), and it is easy to imagine that this may make it difficult for participants to sustain their motivation and attention at a constant level through the trial. We developed, on the basis of the timing models (see Equations 1 and 5), a new method, which will be presented here: the *detrended windowed (lag one) autocorrelation* (DWA), which appears to function well with shorter series. The aim of the present study was then to compare the performances of spectral analysis and DWA in detecting the exploitation of event-based or emergent timers from simulated or experimental series and their robustness with short series.

DETRENDED WINDOWED (LAG ONE) AUTOCORRELATION

DWA is based on a statistical expectation in event-based timing series: that lag(1) autocorrelation, ACF(1), which represents the short-term dependencies in the series, should exhibit a negative value between −.5 and 0 (Wing & Kristofferson, 1973a). According to these authors, this property results from the presence of two identical motor delay terms, but with opposite sign, in successive intervals (see Equation 1). The presence of such terms in successive intervals induces a negative ACF(1). By contrast, in accounting for timing variability in continuous tasks (Equation 5), the model proposed by Delignières et al. (2004) suggests that ACF(1) should be positive. Thus, the D component generates $1/f$ fluctuations, which are known to produce persistent positive autocorrelation over time, and the absence of the motor delay terms avoids introducing negative ACF(1). The presence of a white noise term in the model reduces time dependence in the series, but the autocorrelation function remains positive. The basic idea of the present paper was to distinguish between event-based and emergent timers on the basis of the sign of the ACF(1).

Some tapping experiments have failed to show the negative ACF(1) expected under the W-K model (Ivry & Corcos, 1993). Vorberg and Wing (1996) suggested that the presence of drift in time interval series that frequently occurs in such experiments could result in positive autocorrelation in series, which could counterbalance the negative dependence induced by the event-based production of intervals. The drift parameter introduced by Collier and Ogden (2004) into the W-K model added a new source of positive autocorrelations in series, which could explain this absence of negative lag(1) autocorrelation in continuation tapping series.

Our aim was to determine the best set of parameters for reducing drift, which masks the negative correlations arising from the motor component of the event-based model, without overly reducing estimates of the variance of the D component of the emergent timing model, which induces the positive ACF(1) in this model. As noted earlier, using short series, increasing the participants' practice or removing the linear trend from tapping data have all been suggested as means to reduce drift. We explore alternate methods for removing the linear trend of data series before calculating ACF(1). Indeed, this operation allows us to decrease the relative part of drift variance, which can occur in association with a tendency to tap at a preferred frequency (Ogden & Collier, 1999). Consequently, the nature of the short-term dependencies contained in event-based timing series is more efficiently detected through the ACF(1).

Another source of positive dependence lies in the assumed $1/f$ properties of the C and D components. Indeed, the drift and $1/f$ properties of temporal components (C and D) constitute two different variance sources. As noted by Madison (2004), the $1/f$ properties of tapping series are not affected by training, contrary to drift in series, which is reduced by training. The $1/f$ noise processes are considered as stationary signals. However, as previously noticed, they produce persistent positive autocorrelations that could also hide the expected negative short-term dependence in the series. Nevertheless, it is important to recall that one key property of fractal processes (here the working of the temporal components C and D) is that their variance increases with the length of the bin over which it is computed. Indeed, the $1/f$ noise describes the linear and proportional decrease of the amplitude of a frequency component with increase in its frequency. Consequently, the longer the time used to estimate the properties of the data series, the greater the tendency for the low-frequency components of the series to appear and increase the estimated variance. This phenomenon could partly explain the weaker estimations from the autocovariance method proposed by Wing and Kristofferson (1973a) on long series, assuming the clock component to be a white noise source when it is actually a $1/f$ source. As such, one could suppose that the negative ACF(1) would be expected in short series, but would disappear when the autocorrelation function is computed over long series, as predicted by Collier and Ogden (2004).

Thus, DWA aims to favour the expression of negative dependence by repeatedly computing ACF(1) over a narrow moving window of length n. In each window the series is detrended before computing the autocorrelation. Then the window is moved forward by one point, and the operation is repeated, and so on, until the window has been swept over the entire series. The mean ACF(1) $\bar{\gamma}(1)$ is then calculated by averaging all estimates. At this point two questions remain for a complete implementation of this algorithm: the length of the window (n), and the nature of the detrending (linear or polynomial order 2) applied within each window. A simulation study was performed in order to determine the optimal choices for these two parameters.

SIMULATIONS

In order to test the capacity of DWA, lag one autocorrelation, and spectral analysis to detect the exploitation of event-based or emergent timing process, we generated series according to Equations 1 and 5 simulating both timing behaviours. The difference between these two models is based on the noise term that is added to the main components C and D. In the case of event-based timer a differenced white noise term was added simulating the two motor delay terms, whereas a simple Gaussian white noise term was added in the case of emergent timing simulating the motor error term.

In order to simulate the $1/f$ properties of the C and D components of both models we used the algorithm designed by Davies and Harte (1987; for a detailed presentation, see Cannon, Percival, Caccia, Raymond, & Bassingthwaighte, 1997). This algorithm allows the generation of "exact" fractional Gaussian noise series with known Hurst exponent (H). We generated fractal series with H exponents equal to 0.9 (close to $1/f$ noise) and with 0 mean and unit standard deviation. We also generated series of Gaussian white noise and differenced white noise in order to simulate the noise components, with 0 mean and unit standard deviation.

Wing and Kristofferson (1973a) showed that the variance of the temporal component is proportional to the produced time intervals whereas the motor component variance remains constant whatever the time interval produced. Changing the target frequency in rhythmic tasks has consequences for the relative contribution of the noise variance component of the models. The relative strength of noise in the series (i.e., that proportion of the total variance attributable to noise) has consequences for the high-frequency slope observed in log−log power spectra (Delignières et al., 2004; Gilden et al., 1995; Lemoine et al., 2006) and ACF(1). Indeed, the variances of event-based and emergent series are composed of the variance of the main component C and D, plus the variance of the noise components M and ξ, respectively. As the noise variance becomes larger, the high-frequency slope in the spectra increases (Gilden et al., 1995), and the ACF(1) becomes more negative (Wing & Kristofferson, 1973a).

It is important to be able to determine the effects of noise level in event-based and emergent timing on identification of the underlying timing processes. We varied noise strength by applying a multiplicative coefficient (p for event-based timers and q for emergent timers) to the noise part of the series in order to increase or decrease its contribution to overall variability. The p values varied from 2 to 1 in steps of 0.5 and from 1 to 0.1 in steps of 0.1, and q from 0 to 1 in steps of 0.5 and from 1 to 1.9 in steps of 0.1. We used shorter steps (0.1) for event-based series when the relative strength of noise was weak (all methods exploiting the differenced nature of the noisy part of series) and conversely for emergent series when noise tended to extinguish autocorrelations in the series. These values of p and q represented a percentage of the noisy part in the total variance of 90%, 68%, 34%, 25%, 16%, 7%, and 2% for p values of 2, 1, 0.5, 0.4, 0.3, 0.2, and 0.1, respectively, and 1%, 21%, 52%, 72%, 74%, 79%, and 80% for q values of 0.1, 0.5, 1, 1.5, 1.6, 1.7, 1.8, and 1.9, respectively. Testing these different relative strengths of noise was also considered important to determine which methods allow reliable identification of timing processes in a wider range of frequency conditions, the relative strength of noise being related to the frequency of time intervals.

Another goal was to test the effect of series length on the performance of each method. We know that spectral analysis requires long series in order to detect the nature of correlations present in the series (Delignières et al., 2006). The DWA method should work better on long series because of the greater number of estimates involved. Our aim was obviously to determine the method that worked best on short series to avoid fatiguing participants, which might be expected to increase with longer trials. We tested series of 1,024, 512, 256, 128, and 64 data points. In each condition (event-based vs. emergent series, noise strength, and series length), 80 series were simulated.

In order to determine the optimal parameters for implementing DWA, we tested window lengths ranging from 20 to 50 data points, in steps of 10. We selected this range because it represents the most commonly used sequence lengths in studies of timing variability employing autocorrelation-based analyses (Vorberg & Wing, 1996; Wing & Kristofferson, 1973a, 1973b). We also compared the performances of two detrending strategies (linear vs. polynomial of order two).

Spectral analyses were performed using $^{low}PSD_{we}$ (see Appendix), as recommended by Eke et al. (2000) and Delignières et al. (2006), with some preprocessing operations applied before spectral analysis. We compared two alternative measures of the high-frequency slope of spectra: either the top half of spectra ($N/2$), or the top three quarters ($N/4$). Finally, we computed the lag one autocorrelation, ACF(1), over the entire series. For each method, we assessed the median percentage of misclassifications (event-based series identified as emergent series, and conversely), and the variability of estimates.

Results

We present in Table 1 the simulation-based estimates of $\bar{\gamma}(1)$, calculated over all series lengths, as a function of p and q values representing noise strength. As can be observed, DWA gave inaccurate results for event-based series, with positive $\bar{\gamma}(1)$ values, when noise strength was low. Conversely, DWA gave unexpected negative $\bar{\gamma}(1)$

Table 1. *Simulation-based estimates of mean $\bar{\gamma}(1)$ indices calculated over all series lengths*

		Order of detrending							
		Linear				*Order 2 polynomial*			
Series type	p/q coefficients	*20*[a]	*30*[a]	*40*[a]	*50*[a]	*20*[a]	*30*[a]	*40*[a]	*50*[a]
Event-based (p)	0.1	***0.22***	***0.30***	***0.35***	***0.38***	***0.12***	***0.22***	***0.28***	***0.31***
	0.2	***0.16***	***0.24***	***0.28***	***0.31***	***0.06***	***0.16***	***0.21***	***0.25***
	0.3	***0.07***	***0.14***	***0.18***	***0.21***	− 0.03	***0.06***	***0.11***	***0.15***
	0.4	− 0.02	***0.05***	***0.09***	***0.12***	− 0.11	− 0.02	***0.02***	***0.05***
	0.5	− 0.09	− 0.03	***0.01***	***0.03***	− 0.17	− 0.10	− 0.05	− 0.03
	0.6	− 0.16	− 0.11	− 0.07	− 0.05	− 0.23	− 0.17	− 0.13	− 0.11
	0.7	− 0.21	− 0.17	− 0.14	− 0.12	− 0.28	− 0.22	− 0.19	− 0.17
	0.8	− 0.25	− 0.21	− 0.18	− 0.17	− 0.30	− 0.26	− 0.23	− 0.21
	0.9	− 0.28	− 0.25	− 0.23	− 0.21	− 0.33	− 0.29	− 0.27	− 0.25
	1	− 0.32	− 0.29	− 0.27	− 0.25	− 0.36	− 0.32	− 0.30	− 0.29
	1.5	− 0.40	− 0.38	− 0.37	− 0.37	− 0.42	− 0.40	− 0.39	− 0.39
	2	− 0.42	− 0.42	− 0.41	− 0.41	− 0.44	− 0.43	− 0.43	− 0.42
Emergent (q)	0.1	0.24	0.32	0.37	0.39	0.14	0.24	0.30	0.33
	0.5	0.14	0.22	0.26	0.28	0.05	0.14	0.19	0.22
	1	0.03	0.09	0.13	0.15	***− 0.04***	0.03	0.07	0.10
	1.1	0.01	0.07	0.10	0.12	***− 0.06***	0.01	0.05	0.07
	1.2	0.00	0.05	0.08	0.10	***− 0.07***	***0.00***	0.04	0.06
	1.3	***− 0.01***	0.04	0.07	0.08	***− 0.08***	***− 0.01***	0.02	0.04
	1.4	***− 0.02***	0.02	0.05	0.07	***− 0.09***	***− 0.03***	0.01	0.03
	1.5	***− 0.03***	0.01	0.04	0.05	***− 0.10***	***− 0.04***	***− 0.01***	0.01
	1.6	***− 0.03***	0.02	0.05	0.06	***− 0.09***	***− 0.03***	0.00	0.02
	1.7	***− 0.04***	0.01	0.03	0.04	***− 0.10***	***− 0.04***	***− 0.01***	0.01
	1.8	***− 0.05***	***− 0.01***	0.02	0.03	***− 0.11***	***− 0.05***	***− 0.02***	***− 0.01***
	1.9	***− 0.05***	***− 0.01***	0.01	0.03	***− 0.11***	***− 0.05***	***− 0.02***	***− 0.01***
Mean *SD*		0.0688	0.0718	0.0749	0.0773	0.0670	0.0707	0.0737	0.0762
Median percentage of misclassifications		40	21	17	14	53	40	25	22

Note: The p (event-based timing simulations) and q (emergent timing simulations) coefficients represent the level of noise variance. Different combinations of parameters were tested. The first parameter was the order of detrending (linear vs. polynomial order 2) applied on the window before calculating the lag(1) autocorrelation, and the second was the length of the window ([a]window size from 20 to 50 by steps of 10). Errors of classification are shown in bold italics. An error of classification is determined by the sign of the value, which should be negative for event-based timing and positive for emergent timing. Mean standard deviation and median percentage of misclassifications are presented in the bottom two rows.

values for emergent series when noise strength was high. Nevertheless, it should be noted that in both cases these inaccurate estimates were obtained with noise levels that were rarely encountered in experimental data with time intervals under 1,000 ms (emergent timer: Delignières et al., 2004; event-based timer: Collier & Ogden, 2004; Ivry & Corcos, 1993)—that is, the motor variance represents less than 25% of the total variance with event-based simulations and more than 70% of the total variance with emergent simulations.

Table 1 suggests that the best combination for DWA should use a 30-point window and linear detrending. Polynomial detrending (order 2) yielded particularly bad results for emergent series, especially for short window lengths. Using long windows (40 and 50 points) gave bad results for event-based series, especially with linear detrending. Combining linear detrending with 30-point windows maintains a low variance in estimations (mean variance = 0.0718) and a low median percentage of misclassifications (21%). The polynomial detrending with 50-point windows gives approximately the same results; nevertheless the variance of estimations is higher (0.0762 vs. 0.0718). The 30-point window with linear detrending combination is used in the following analyses to compare performances of DWA and spectral analysis and to test DWA on experimental series.

We present in Table 2 the results of spectral analysis, with the high-frequency slopes obtained over all series lengths, for event-based and emergent series. The results are presented for the two proposed criteria for spectral analysis ($N/2$ and $N/4$). As can be seen, the mean variance of estimations, calculated over all series, was higher for the $N/2$ (0.9693) than for the $N/4$ criterion (0.4077). Additionally, the median percentage of misclassifications was lower for the $N/4$ criterion, so we decided to use it for the following analyses.

We also present in Table 2 the results of lag one autocorrelation, ACF(1), as a function of p and q values. As can be observed, in spite of a lower variability than DWA (0.0665) and a low median percentage of misclassifications (5%), ACF(1) gave erroneous results in the case of the event-based timer.

Figures 2 and 3 contrast the results of DWA (top), spectral analysis (middle), and ACF(1) (bottom), for event-based (left), and emergent (right) series. We present in Figure 2 the mean indices—$\bar{\gamma}(1)$, high-frequency slopes, and ACF(1)—obtained for each series length. As can be seen, series length did not have a noticeable effect except for ACF(1), with best results for short series in the case of event-based timers and conversely for emergent timers. Figure 3 presents the standard deviations of the three indices, according to series length and noise strength. As can be seen, all methods gave more variable results as series length decreased. Because of the different sizes of samples of estimations for the three indices of interest and in order to compare the variability of the three methods, we normalized the three samples of results to unit standard deviation. This operation was carried out per method and per timing simulation (event-based vs. emergent). First we calculated the global mean over all measured values, 80 (simulated series) $\times$ 12 (noise strengths, p or q coefficient) $\times$ 5 (series lengths: 1,024, 512, 256, 128, 64), as well as the global standard deviation. Thus, we subtracted from each measured value the global mean and then divided by the global standard deviation. Subsequently, we calculated the standard deviation of these new normalized values for each noise strength (p or q) in each series length: the normalized standard deviation. Finally, we averaged these normalized standard deviations over all values to obtain the mean normalized standard deviation. The mean normalized standard deviation was higher for high-frequency slopes (0.52 for event-based and 0.81 for emergent timer) than for $\bar{\gamma}(1)$ (0.30 and 0.58, respectively) and ACF(1) (0.22 and 0.45, respectively).

TEST ON EXPERIMENTAL DATA

In order to confirm the accuracy of methods for identifying the underlying timing process, we applied them on experimental data collected during previous studies on continuation tapping tasks thought to elicit the use of an event-based

Table 2. *Simulation-based estimates of mean high-frequency slopes and lag one autocorrelation calculated over all series lengths in function of* p *and* q *coefficients representing the level of noise variance*

	p/q coefficients	$N/2$	$N/4$	*ACF(1)*
Event-based (p)	0.1	***− 0.69***	***− 0.81***	***0.57***
	0.2	***− 0.46***	***− 0.58***	***0.51***
	0.3	***− 0.09***	***− 0.28***	***0.43***
	0.4	0.16	0.01	***0.33***
	0.5	0.39	0.22	***0.23***
	0.6	0.70	0.43	***0.14***
	0.7	0.64	0.57	***0.05***
	0.8	0.79	0.69	− 0.02
	0.9	0.80	0.81	− 0.08
	1	0.79	0.94	− 0.13
	1.5	0.95	1.16	− 0.30
	2	0.86	1.21	− 0.38
Emergent (q)	0.1	− 0.84	− 0.91	0.59
	0.5	− 0.54	− 0.59	0.47
	1	− 0.23	− 0.26	0.29
	1.1	− 0.16	− 0.24	0.27
	1.2	− 0.21	− 0.22	0.24
	1.3	− 0.16	− 0.19	0.22
	1.4	− 0.20	− 0.17	0.20
	1.5	− 0.14	− 0.13	0.18
	1.6	− 0.16	− 0.19	0.16
	1.7	− 0.05	− 0.14	0.15
	1.8	− 0.16	− 0.11	0.13
	1.9	− 0.04	− 0.13	0.12
Mean *SD*		0.9693	0.4077	0.0697
Median percentage of misclassifications		41	26	5

Note: ACF(1) = lag one autocorrelation; p = event-based simulations; q = emergent simulations. Criteria of two high-frequency slopes were tested: $N/2$ representing the high-frequency slope computed on the top half of the spectra, and $N/4$ representing the high-frequency slope computed on the top three quarters of the spectra. Errors of classification are in bold italics. An error of classification is determined by the sign of the value, which should be negative for event-based timing and positive for emergent timing for ACF(1) values and conversely for high-frequency slopes. Mean standard deviation and median percentage of misclassifications are presented in the bottom two rows.

timer and continuation oscillation tasks thought to elicit the use of an emergent timer.

Tests on event-based timers

Method

A total of 10 right-handed participants who were nonmusicians took part in this experiment (5 women and 5 men, mean age = 27.56 years, *SD* = 6.23). Participants performed in a synchronization-continuation tapping paradigm using the index finger of the dominant hand in three frequency conditions (2.7, 1.8, and 1.25 Hz corresponding to time intervals of 370, 556, and 800 ms, respectively). Participants placed their hand on a desk with a response key under the index finger of the dominant hand, the other digits being kept relaxed. They were instructed to synchronize one tap of the index finger with each beep of the metronome. After 25 cycles, the metronome stopped, and participants had to continue tapping as regularly as possible, following the initial tempo. The metronome was provided through headphones and was created with a

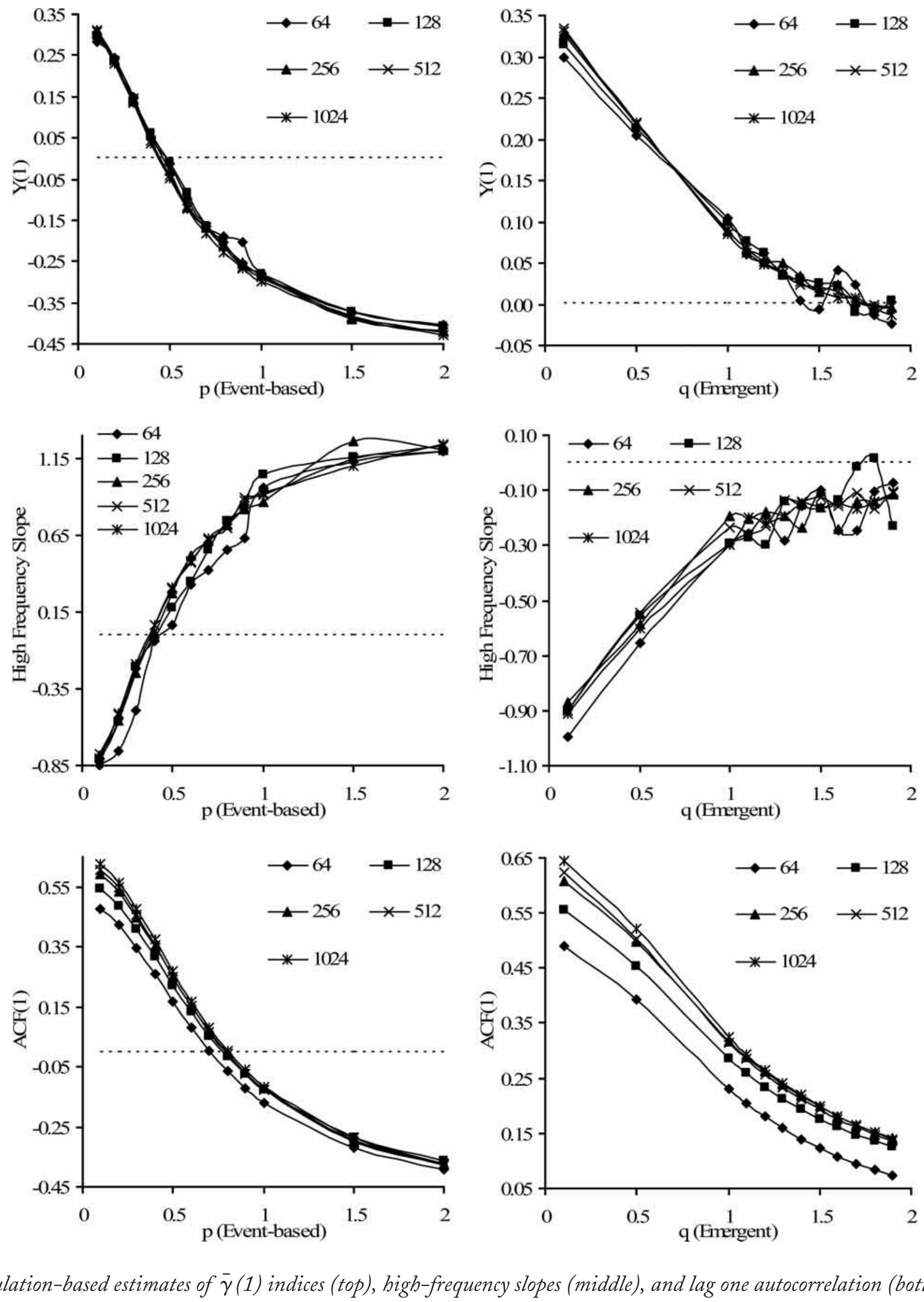

Figure 2. *Simulation-based estimates of $\bar{\gamma}(1)$ indices (top), high-frequency slopes (middle), and lag one autocorrelation (bottom) estimations for all series lengths. Figures on the left show the evolution with p coefficients (noise variance level in event-based simulations), and those on right show the evolution with q coefficients (noise variance level in emergent simulations). $\bar{\gamma}(1)$ indices and high-frequency slopes presented are those of the retained parameters in each method (30-data-point window with linear detrending for DWA and N/4 criteria for the ${}^{low}PSD_{we}$). DWA = detrended windowed (lag one) autocorrelation.*

Matlab script. The duration of trials was fixed in order to collect 1,200 time intervals. The task was performed twice per frequency condition in two identical morning sessions separated by 2 months. The order of frequency conditions was randomly determined at the time of the first session and was identical for the second.

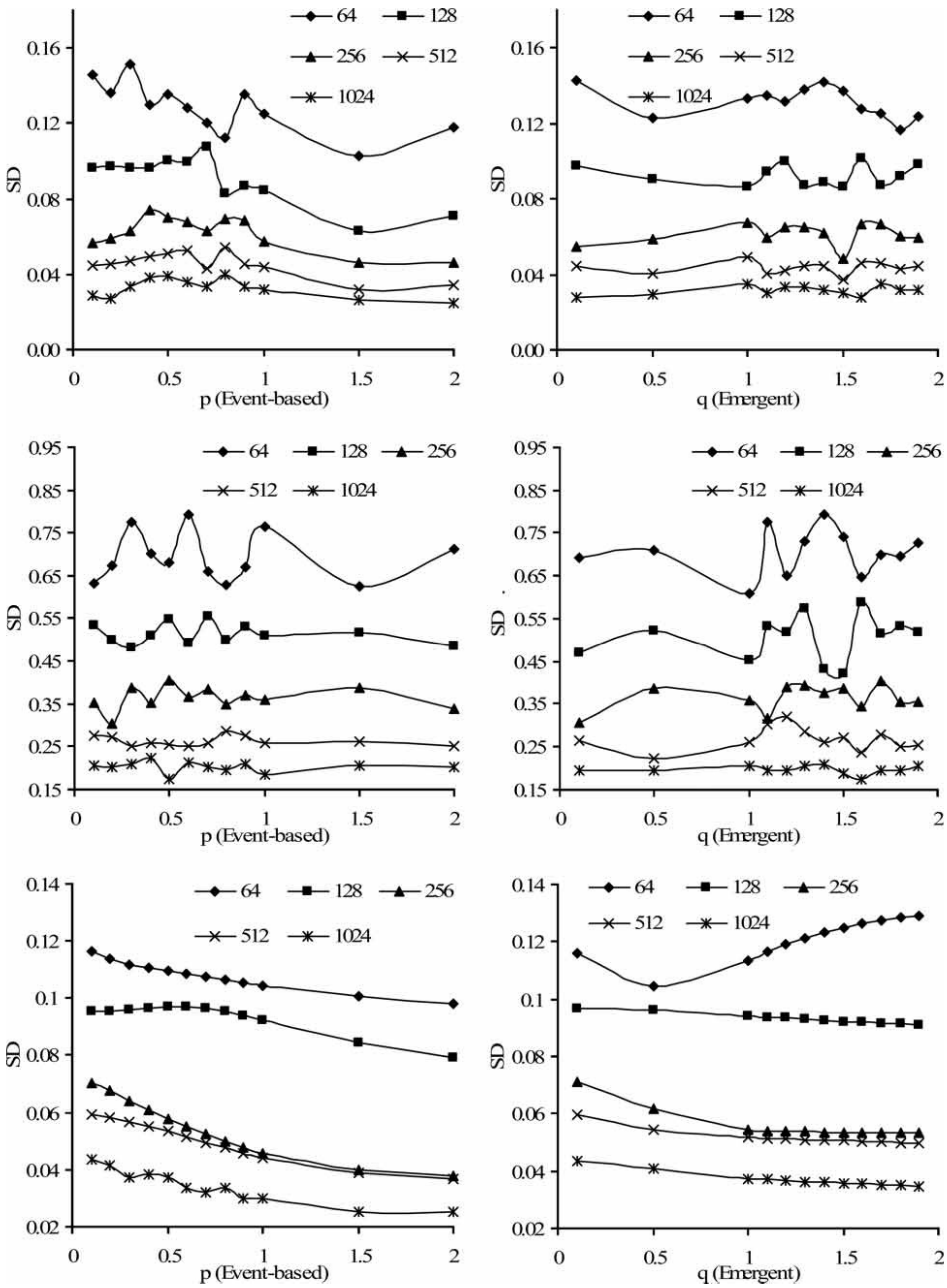

Figure 3. *Simulation-based estimates of standard deviation (SD) of $\bar{\gamma}(1)$ indices (top), high-frequency slopes (middle), and lag one autocorrelation (bottom) estimations with series length. Figures on the left show the evolution with p coefficients (noise variance level in event-based simulations), and those on the right show the evolution with q coefficients (noise variance level in emergent simulations). SDs presented are those of the retained parameters in each method (30-data-point window with linear detrending for DWA and N/4 criteria for the $^{low}PSD_{we}$). DWA = detrended windowed (lag one) autocorrelation.*

Results

The results of presenting the three methods of analysis, DWA, $^{low}PSD_{we}$, and ACF(1), are presented in Table 3. Tapping is supposed to be a discrete task favouring the use of an event-based timing process and should yield negative $\bar{\gamma}(1)$, positive high-frequency slopes, and negative ACF(1). As can be seen, all methods gave coherent

Table 3. *Mean indices and standard deviation of estimations of DWA, ACF(1), and $^{low}PSD_{we}$ obtained in tapping experimental series in the three frequency conditions*

		Frequency					
		2.7 Hz		*1.8 Hz*		*1.25 Hz*	
	Series length	*M*	*SD*	*M*	*SD*	*M*	*SD*
DWA	1,024	− 0.29	0.11	− 0.19	0.11	− 0.11	0.11
	128	− 0.17	0.17	− 0.20	0.14	− 0.18	0.19
ACF(1)	1,024	− 0.01	0.26	***0.12***	0.26	***0.31***	0.22
	128	***0.07***	0.30	− 0.02	0.26	0.00	0.32
$^{low}PSD_{we}$	1,024	0.82	0.52	0.33	0.36	0.22	0.42
	128	0.31	0.72	0.63	0.67	0.52	0.77

Note: DWA = detrended windowed (lag one) autocorrelation; ACF(1) = lag one autocorrelation. Two lengths of series were tested (1,024 and 128 data points). "Misclassifications" are in bold italics.

mean indices. Individual indices measured with DWA were negative in most cases, except for two series in the 1.25-Hz condition. The $^{low}PSD_{we}$ gave a higher number of misclassifications (1 in the 2.7-Hz condition, 3 in the 1.8-Hz condition, and 6 in the 1.25-Hz condition). The ACF(1) gave the highest number of misclassifications, with positive values in most cases (65% of misclassifications). Correlation coefficients were calculated between identical frequency conditions. We obtained significant correlations for DWA in the 2.7- and 1.8-Hz conditions ($r = .71$, $p < .05$; $r = .86$, $p < .01$, respectively), for high-frequency slopes in the 2.7- and 1.8-Hz conditions ($r = .67$, $p < .05$; $r = .72$, $p < .05$, respectively), and for ACF(1) only at 2.7 Hz ($r = .69$, $p < .05$).

In order to test the methods on short series, we applied them to the first 128 data points of series. The results were not significantly different from those with longer series for DWA, $F(1, 9) = 0.19$, $p > .05$, and high-frequency slopes, $F(1, 9) = 0.06$, $p > .05$, but differences were found for ACF(1), $F(1, 9) = 6.47$, $p < .05$, with a decrease in indices with the decrease in series length. However, we obtained a higher number of misclassifications for DWA and spectral methods and lower for ACF(1) indices. Nevertheless, the number of misclassifications with ACF(1) was always higher than that with high-frequency slope and always higher with high-frequency slope than with DWA: 33, 15, and 10 for the ACF(1), high-frequency slope, and DWA, respectively.

Working on event-based timing, we also calculated the relative contribution of the motor component to the total variance using the Wing and Kristofferson method (i.e., the covariance method, see Equations 3 and 4), with or without linear detrending of series, and using the Collier and Ogden (2004) method (the differentiated method) in order to estimate the relative strength of noise in the series. These estimations allow us to check the accuracy of tested methods in the process of timers identification, the three methods—DWA, high-frequency slope, and ACF(1)—giving inaccurate results for different noise strength level. These calculations were made on long and short series (1,024 vs. 128). The series where methods gave a negative variance for one of the components were considered as erroneous. We obtained 72% of erroneous series for the covariance method with long series without detrending, 55% with long series with detrending, 55% with short series without detrending, and 47% with short series with detrending. The relative part of the motor component in the total variance was 27%, 32%, 54%, and 57%, respectively. In the case of differentiated method we obtained 38% of erroneous series with long series and 57% with short series. In that case no detrending was applied because of the drift component included in the model. The relative part of the motor component in the total variance was 47% with long series and 28% with short series.

Tests on emergent timers

In a second step, we applied these methods on series collected during an oscillation task thought to elicit the use of an emergent timer.

Method

A total of 13 right-handed participants who were nonmusicians took part in this experiment (4 women and 9 men, mean age = 26.62 years, *SD* = 4.61). As in the tapping task, participants were tested in a synchronization-continuation paradigm, but the task involved oscillation of the forearm as in Delignières et al. (2004). Participants were seated in a chair and had to continuously oscillate a joystick with their dominant forearm. First, they had to synchronize the reversal point in the abduction position of their movement with the beeps of the metronome. When the metronome stopped, participants had to continue to oscillate with the joystick as regularly as possible. The main variable was the series of time intervals collected between each reversal point in the abduction position. The conditions were exactly the same as those in the tapping experiment, with three frequency conditions (2.7, 1.8, and 1.25 Hz) tested in two sessions separated by 2 months and at the same time of the day.

Results

Results are presented in Table 4. As can be seen, all methods gave coherent mean indices with positive values for DWA and ACF(1) and negative values for $^{low}PSD_{we}$. Nevertheless, the individual indices revealed some misclassifications: 13 (17%) for $^{low}PSD_{we}$, 7 (9%) for DWA, but none for ACF(1). The excellent results of the last method nevertheless appear suspect as this method seemed to yield positive values whatever the series analysed. The coefficients of correlation between the same frequency conditions were calculated. We obtained significant correlations only for the 1.8-Hz condition for the DWA and spectral indices ($r = .77$, $p < .01$; $r = .64$, $p < .05$, respectively).

Working on shorter series (first 128 data points), we obtained the same results with DWA and spectral analysis, $F(1, 12) = 0.17$, $p > .05$; $F(1, 12) = 0.01$, $p > .05$, respectively, but not for the ACF(1) with a decrease of autocorrelations with the decrease of series length, as with tapping series, $F(1, 12) = 85.59$, $p < .00001$. The number

Table 4. *Individuals indices of DWA, ACF(1), and* $^{low}PSD_{we}$ *obtained in oscillation experimental series in the three frequency conditions*

		Frequency					
		2.7 Hz		*1.8 Hz*		*1.25 Hz*	
	Series length	*M*	*SD*	*M*	*SD*	*M*	*SD*
DWA	1,024	0.10	0.06	0.09	0.09	0.14	0.08
	128	0.11	0.10	0.09	0.12	0.14	0.13
ACF(1)	1,024	0.50	0.20	0.48	0.23	0.66	0.41
	128	0.31	0.21	0.27	0.18	0.18	0.20
$^{low}PSD_{we}$	1,024	− 0.25	0.31	− 0.38	0.37	− 0.40	0.31
	128	− 0.25	0.58	− 0.31	0.50	− 0.49	0.64

Note: DWA = detrended windowed (lag one) autocorrelation; ACF(1) = lag one autocorrelation. Two lengths of series were tested (1,024 and 128 data points).

of misclassifications increased for all indices with 10, 19, and 4 for the DWA, spectral analysis, and ACF(1), respectively.

DISCUSSION

The aim of this study was to determine the most appropriate method for identifying the timing process underlying time intervals produced during the performance of rhythmic tasks. This constitutes a necessary step to determine factors eliciting the use of event-based and emergent timing processes, in order to develop a better understanding of the nature of motor timing. Indeed, some methods used in the literature—for example, the correlational method (Robertson et al., 1999; Zelaznik et al., 2002)—cannot determine the timing process used for individual participants, but assumes the use of a common timing process in a particular condition, which results in conclusions that cannot be generalized.

Three timing process identification methods were tested: two methods currently used in the literature and a third method that we developed on the basis of the W-K model (Wing & Kristofferson, 1973a, 1973b). These methods were applied first to simulated series in order to define the best parameters for each method and to test their performances and reliability. In order to confirm these results, we applied the methods to experimental series collected in tapping and oscillation tasks, which we assumed would elicit the use of an event-based timer and an emergent timer, respectively (Delignières et al., 2004). We also tested the effect of series length on timing process identification for determining the best method of working on short series in order to avoid overlong trials, likely to fatigue participants and which could require additional components in the models, such as the drift component of Collier and Ogden (2004).

The first conclusion that can be drawn is the failure of ACF(1) to determine the timing process used in the control of rhythmic activities. As observed in our first study on tapping experimental data, this method categorized correctly the timing process only in 35% of cases with long series and 45% in short series. The simulation study yielded similar results with an inability to detect the use of event-based timers with low levels of noise, which could be encountered in experimental data (motor variance less than 58% of the total variance). The absence of preprocessing operations, as used in DWA (detrending and windowing), results in a failure to distinguish between timers. This method, while previously recommended by Wing and Kristofferson (1973a, 1973b), should not be used in this form for the discrimination of timing processes under the conditions evaluated in the studies reported here. However, the ACF(1) method gives accurate results with shorter sequences and trained participants, an approach used for avoiding the characteristic drift occurring at the beginning of trials (Ogden & Collier, 1999), which hide the short term correlations and consequently the negative lag(1) autocorrelation with event-based timing. However, the effect of training on timing processes exploited was not studied and could have consequences for the nature of the underlying timing process. Thus it might be argued that the aim of an identification method should be to identify this process for every trial and not just after training.

The two methods that gave the best results were DWA and $^{low}PSD_{we}$. Simulation data allowed us to determine the best combination of parameters for both methods. In the case of DWA the variability of estimates increased with window lengths and was higher with linear than with polynomial detrending (Table 1). Nevertheless, the number of misclassifications increased as window length decreased and was higher with polynomial detrending (Table 1). The best compromises were associated with linear detrending and a 30-data-point window and with polynomial detrending (order 2) and a 50-data-point window. The former combination yielded lower variability of estimates (0.0718 vs. 0.0762), and we then decided to use these parameters for the DWA method. In the case of $^{low}PSD_{we}$ the choice of the better parameter as being $N/4$ was easier with larger variability with the $N/2$ than with the $N/4$ criteria and higher median percentage of misclassifications with the $N/2$ criteria.

Regarding the simulation results, DWA and spectral analysis seemed to give identical results with an advantage to spectral analysis with the detection, on average, of the correct timing process with lower noise level in the case of event-based timing simulations and higher noise level in the case of emergent timing simulations. However, the variance of estimates with the spectral analysis was generally higher than that with DWA. This tendency was confirmed by the mean normalized standard deviation, which was larger for spectral analysis than for DWA. The difference in mean normalized standard deviation between methods was 0.22 and 0.23 for event-based and emergent timing processes, respectively, representing 73% and 40% of the DWA mean normalized standard deviation, respectively. The higher variability of estimates with spectral analysis could explain the supposed better performances of this method with mean indices. Indeed, whatever p or q values the mean standard deviation was identical (Figure 3). As can be seen in Tables 1 and 2, the difference separating the mean indices measured with both methods and the value representing a misclassification decreased as noise level increased for event-based timing process simulations and conversely for emergent timing process simulations. The greater the difference between the mean standard deviation observed and the difference separating the mean index measured and the value representing a misclassification, the higher the probability of obtaining individual misclassifications. As can be seen in Tables 1 and 2, there were more p and q values where this difference was larger in the case of spectral analysis than for DWA. The spectral analysis allowed detection of the underlying timer when many series were collected. However, on isolated trials its high variability induced higher probability of misclassifications than did DWA. The required length of series for these methods did not allow testing of participants more than twice because of the duration of tasks, involving fatigue of participants, and number of conditions tested. It then seems important to determine the method giving the most robust estimations on isolated series in order to identify the correct timing process testing participants only once or twice per task and condition. The most suitable and reliable method is thus the DWA.

The analyses performed on experimental data confirmed this tendency with more reliable results for DWA. We obtained with this method a lower percentage of misclassifications assuming the two respective models for discrete (event-based) and continuous (emergent) timing, whatever the length of series. The higher variability of $^{low}PSD_{we}$ (Tables 3 and 4), evidenced also during the simulation study, could explain the higher rate of misclassifications observed with this method.

However, DWA gave erroneous results for series with a low level of noise for event-based timing (p coefficients) and a high level of noise for emergent timing process (q coefficients). This process could be increased for event-based timing simulations using fractal series with a higher exponent corresponding to higher positive correlations in the series, which would hide all the more the negative short-term fluctuations and consequently the negative lag(1) autocorrelation on DWA. In the case of emergent timing simulations the inverse is true, with an increase of misclassifications using fractal series with a lower fractal exponent involving a faster fall of correlations with the increase of noise level, which lowers positive correlations in series. However, the mean fractal exponent measured on experimental series was 0.88 in tapping series and 0.95 in oscillation series. Madison (2004) reported lower exponents in a tapping experiment with shorter series length (256 data points). Delignières et al. (2004) reported higher exponents for emergent timing processes than for event-based timers. The use of simulated series with a fractal exponent of 0.9 was then justified. Madison (2004) showed a significant increase of fractal exponents with time interval duration. As mentioned above for noise level, it seems important in a tapping experiment to work with frequency conditions greater than 1 Hz (i.e., time intervals shorter than 1 s) in order to avoid the appearance of critical level of positive correlations in series affecting the detection of event-based timing process. In addition, the use of frequency conditions above 3 Hz (i.e., time intervals shorter than 333 ms) has to be taken with caution for emergent timing, since the high level of noise in these conditions lowers the positive correlations in the series. However, the noise level is also likely to depend on the movement used in the rhythmic task and its amplitude.

Originally the use of ACF(1) was proposed with short sequences to explore an event-based timing model (Wing & Kristofferson, 1973a). More precisely, the lag(1) autocovariance was applied to measure the variance of the motor component. Nevertheless, the appropriateness of this method was questioned (Ivry & Corcos, 1993; Vorberg & Wing, 1996), and several other methods were proposed. Indeed, the ACF(1) method is adequate, especially when using shorter sequences and trained subjects yielding stationary data and for the purposes of estimating WK model parameters. Ivry and Corcos (1993) recommended measuring the variance of clock and motor components via the regression slope between the variance of tapping series and the duration of intervals and the intercept of the slope. Nevertheless, these methods were not used in order to detect the use of an event-based timer, but only to measure the variance of model components (Collier & Ogden, 2004; Ivry & Corcos, 1993; Pressing & Jolley-Rogers, 1997; Vorberg & Wing, 1996). Delignières et al. (2004) proposed discriminating between timers through spectral analysis. However, this method required very long series (Delignières et al., 2006). A single trial lasted on average 12 minutes, leading to evident problems of fatigue or concentration. DWA and ${}^{low}PSD_{we}$ methods were shown to give reliable results even with short series. Nevertheless, the variability of estimations increased as series length decreased. We recommend working on series of 128 data points, which is a good compromise between task duration and method reliability. Working on this length of series with a method able to identify accurately the timing process underlined is an advantage as different conditions could be tested without providing overlong trials resulting in fatigue of participants, which allows testing new hypotheses on the use of different timing processes in human motor control. In addition, this length of series could allow measuring of the fractal exponent of series via the maximum likelihood estimation (Deriche & Tewfik, 1993), which has been proven to work accurately with short sequences (Delignières et al., 2006) and has been used to compare and understand the functioning of timing processes (Delignières, Torre, & Lemoine, 2008). The tests on experimental short series reassured us in showing no differences between indices measured on long and short series even if there was an increase in misclassifications. This increase could be explained by the increase of the variance of indices' estimations with short series as observed with simulations. The standard deviation of $\bar{\gamma}(1)$ indices increased from 0.11 in average to 0.16 in tapping series and from 0.08 to 0.11 in oscillation series. The same phenomenon was observed with spectral analysis.

The decrease of correlations between indices measured in the same frequency conditions could also be interpreted in this way. Testing participants twice per frequency condition in each session should be a necessary precaution in order to measure a mean index closer to the exact index as recommended with fractal exponents (Delignières et al., 2006; Rangarajan & Ding, 2000). This precaution is possible with shorter sequences of 150 time intervals without generating trials that were too long and sessions with fatigue of participants and could allow significant correlations obtained with long series. The significant correlations obtained with DWA in the tapping experiment (at 2.7 Hz and 1.8 Hz) were found at frequencies close to the preferred frequencies reported in the literature. McAuley, Jones, Holub, Johnston, and Miller (2006) detected the spontaneous tempi in tapping between 300 ms and 600 ms, depending on the age of participants. These better correlations close to preferred frequency could be explained by a decrease of the relative part of clock variance close to preferred frequency due to lowering of drift in time intervals produced toward preferred frequencies and a higher efficiency of "clock" and consequently by an increase of the relative part of motor variance. Indeed, using the differentiated method (Collier & Ogden, 2004) on long tapping series, we expected to obtain according to their model an increase in the contribution of drift and clock variance and a decrease of motor variance with increasing interval duration (Wing & Kristofferson, 1973a). However, on average we obtained a higher motor variance in the 1.8-Hz condition, with a weaker clock variance than that in the 2.7-Hz condition. Close to the preferred frequency the "internal clock" seems to work more efficiently, which could be explained by the use of stable strategies by participants (Delignières et al., 2008), which could induce the better correlations. In the case of emergent timers the preferred frequency seems to be close to 1.8 Hz with significant correlation at this frequency. Nevertheless, these results have to be confirmed on short series using multiple measurements as recommended above.

CONCLUSION

On the basis of simulated and experimental data analyses, the DWA method was recommended to discriminate between event-based and emergent timekeeping. The analysis of experimental data showed the failure of lag(1) autocorrelation to discriminate event-based timers on long series. The parameters advised for DWA were a 30-data-point window length and linear detrending. This method allows working on short series with a minimal series length recommended at 128 data points.

Original manuscript received 19 January 2007
Accepted revision received 20 January 2008
First published online 12 June 2008

REFERENCES

Cannon, M. J., Percival, D. B., Caccia, D. C., Raymond, G. M., & Bassingthwaighte, J. B. (1997). Evaluated scaled windowed variance methods for estimating the Hurst coefficient of time series. *Physica A*, *241*, 606–626.

Collier, G. L., & Ogden, R. T. (2004). Adding drift to the decomposition of simple isochronous tapping: An extension of the Wing–Kristofferson model. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 853–872.

Davies, R. B., & Harte, D. S. (1987). Tests for Hurst effect. *Biometrika*, *74*, 95–101.

Delignières, D., Lemoine, L., & Torre, K. (2004). Time intervals production in tapping and oscillatory motion. *Human Movement Science*, *23*, 87–103.

Delignières, D., Ramdani, S., Lemoine, L., Torre, K., Fortes, M., & Ninot, G. (2006). Fractal analyses for "short" time series: A re-assessment of classical methods. *Journal of Mathematical Psychology*, *50*, 525–544.

Delignières, D., Torre, K., & Lemoine, L. (2005). Methodological issues in the application of monofractal analyses in psychological and behavioral research. *Nonlinear Dynamics in Psychology and Life Science*, *9*, 435–461.

Delignières, D., Torre, K., & Lemoine, L. (2008). Fractal models for event-based and dynamical timers. *Acta Psychologica*, *127*, 382–397.

Deriche, M., & Tewfik, A. H. (1993). Maximum likelihood estimation of the parameters of discrete fractionally differenced Gaussian noise process. *IEEE Transactions on Signal Processing*, *41*, 2977–2989.

Eke, A., Herman, P., Bassingthwaighte, J. B., Raymond, G. M., Percival, D. B., Cannon, M., et al. (2000). Physiological time series: Distinguishing fractal noises from motions. *Pflugers Archiv*, *439*, 403–415.

Fougère, P. F. (1985). On the accuracy of spectrum analysis of red noise processes using maximum entropy and periodogram methods: Simulation studies and application to geographical data. *Journal of Geographical Research*, *90*, 4355–4366.

Gilden, D. L., Thornton, T., & Mallon, M. W. (1995). $1/f$ noise in human cognition. *Science*, *267*, 1837–1839.

Ivry, R. B., & Corcos, D. M. (1993). Slicing the variability pie: Component analysis of coordination and motor dysfunction. In K. M. Newell & D. M. Corcos (Eds.), *Variability and motor control* (pp. 415–447). Urbana, IL: Human Kinetics.

Ivry, R. B., & Hazeltine, R. E. (1995). Perception and production of temporal intervals across a range of durations: Evidence for a common timing mechanism. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 3–18.

Keele, S. W., Ivry, R. I., & Pokorny, R. (1987). Force control and its relation to timing. *Journal of Motor Behavior*, *19*, 96–114.

Keele, S. W., Pokorny, R. A., Corcos, D. M., & Ivry, R. (1985). Do perception and motor production share common timing mechanisms: A correctional analysis. *Acta Psychologica (Amst)*, *60*, 173–191.

Lemoine, L., Torre, K., & Delignières, D. (2006). Testing for the presence of $1/f$ noise in continuation tapping data. *Canadian Journal of Experimental Psychology*, *60*, 247–257.

Madison, G. (2004). Fractal modeling of human isochronous serial interval production. *Biological Cybernetics*, *90*, 105–112.

McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M., & Miller, N. S. (2006). The time of our lives: Life span development of timing and event tracking. *Journal of Experimental Psychology: General*, *135*, 348–367.

Ogden, R. T., & Collier, G. L. (1999). On detecting and modeling deterministic drift in long run sequences of tapping data. *Communications in Statistics—Theory and Methods*, *28*, 977–987.

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

Rangarajan, G., & Ding, M. (2000). Integrated approach to the assessment of long range correlation in time series data. *Physical Review E*, *61*, 4991–5001.

Robertson, S. D., Zelaznik, H. N., Lantero, D. A., Bojczyk, K. G., Spencer, R. M., Doffin, J. G., et al. (1999). Correlations for timing consistency among tapping and drawing tasks: Evidence against a single timing process for motor control. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1316–1330.

Schöner, G. (2002). Timing, clocks, and dynamical systems. *Brain and Cognition*, *48*, 31–51.

Vorberg, D., & Wing, A. (1996). Modeling variability and dependence in timing. In H. Heuer & S. W. Keele (Eds.), *Handbook of perception and action* (Vol. 2, pp. 181–262). London: Academic Press.

Wing, A. M., & Kristofferson, A. B. (1973a). Response delays and the timing of discrete motor responses. *Perception and Psychophysics*, *14*, 5–12.

Wing, A. M., & Kristofferson, A. B. (1973b). The timing of interresponse intervals. *Perception and Psychophysics*, *13*, 455–460.

Zelaznik, H. N., Spencer, R. M., & Ivry, R. B. (2002). Dissociation of explicit and implicit timing in repetitive tapping and drawing movements. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 575–588.

APPENDIX

Spectral analysis

$^{low}PSD_{we}$

We used the $^{low}PSD_{we}$ method initially proposed by Fougère (1985) and modified by Eke et al. (2000), which includes some preprocessing operations before the application of the fast Fourier transform (FFT): First the mean of the series was subtracted from each value, and then a parabolic window was applied: Each value in the series was multiplied by the following function:

$$W(j) = 1 - \left(\frac{2j}{N+1} - 1\right)^2 \qquad \text{for } j = 1, 2, \ldots N. \tag{A1}$$

This transformation induces a tapering of the series and is supposed to reduce the leakage in the periodogram. Spectral leakage is the term used to describe the loss of power of a given frequency to other frequency bins in the FFT. There are edge effects arising from the discontinuity at the bounds that cause spectral leakage. It implies that windowing in the time domain corresponds to smoothing in the frequency domain. This smoothing reduces sidelobes associated with the window. Finally, a linear detrending was applied to the resulting series. The FFT algorithm was then applied on the obtained series.

A fractal series is characterized by the following power law:

$$S(f) \propto 1/f^{\beta} \tag{A2}$$

where β is the spectral exponent, f the frequency, and $S(f)$ the correspondent squared amplitude. β is estimated by calculating the negative slope $(-\beta)$ of the linear regression of log $S(f)$ against log f. β equals 0 for white noise, 2 for ordinary Brownian motion, and 1 for $1/f$ noise. As proposed by Eke et al. (2000) we excluded in the fitting of β the high-frequency power estimates ($f > 1/8$ of maximal frequency). This method was proven to provide more reliable estimates of the spectral exponent.