Control Entropy: What Is It and What Does It Tell Us?

Stephen J. McGregor[1], Erik Bollt[2]

[1]Human Factors-Dynamical Systems Laboratory, Eastern Michigan University, Ypsilanti, MI and [2]Department of Mathematics & Computer Science, Clarkson University, Potsdam, NY

ABSTRACT

Complex tasks of motor control in humans, such as locomotion or postural control, exhibit patterns of variability that until recently have been indiscernible from random noise. Tools from the field of non-linear dynamical systems have been increasingly applied to measurements of these tasks and changes in these complex patterns have been identified. A particular tool, control entropy (CE), is a measure of the regularity, or conversely, the complexity of a signal and is used to infer the constraints present on a system. More importantly, CE can be used under non-stationary conditions, and can therefore identify changes in the complexity or constraints on a system under dynamic exercise conditions. In this review, we summarize the insight that has been gained from application of CE to signals from studies involving walking, running and postural control. We show that changing constraints can be identified during dynamic exercise and that these are reflected in changing CE. We also discuss how CE can identify increased complexity of tasks such as postural control in the fatigued state.

Key Words: complexity, locomotion, postural control, stationarity

INTRODUCTION

During measurement of motor tasks such as running or postural control, complex interactions of the various components of biological control systems often result in non-linear dynamical (time dependent) responses. Traditionally, only linear variability statistics have been available to examine the resulting data and underlying processes. Since the introduction of Approximate Entropy (AE) by Pincus in 1991 (16, 19) for the analysis of heart rate variability, there has been increasing interest in the use of non-linear regularity statistics such as AE, as means of non-invasively identifying clinically relevant diseased and/or perturbed physiological states. Specialized derivations of AE, such as Sample Entropy (SE) and Multiscale Entropy, to name a few, have increased the utility of regularity statistics, and hence their use in the literature (Figure 1).

Although AE and SE have found apparently wide application in various biomedical fields (4-7, 9, 15, 17, 18, 23, 25, 26), there are limitations with regard to the practical application of most of these currently available regularity statistics (20, 23). A well documented limitation to the application of AE, for example, is a bias resulting from self matching discretized samples (20). Another limitation inherent to all previously available approaches to entropy analysis is the requirement for stationarity. That is, the signal to be analyzed must be collected under steady state conditions, meaning the system must be stationary in a statistical sense (10). In fact, stationarity is required by many non-linear analysis techniques (23). This imposes serious limitations to successful clinical application because many diseased and perturbed physiological states (e.g. exercise) are monitored under non-stationary conditions. To address these limitations, we (1) recently developed a novel regularity statistic termed, Control Entropy (CE). Control entropy addresses the issue of stationarity inherent as a limitation in previous entropy approaches by calculating the entropy of the differences between neighborhood values, as opposed to the absolute values themselves.

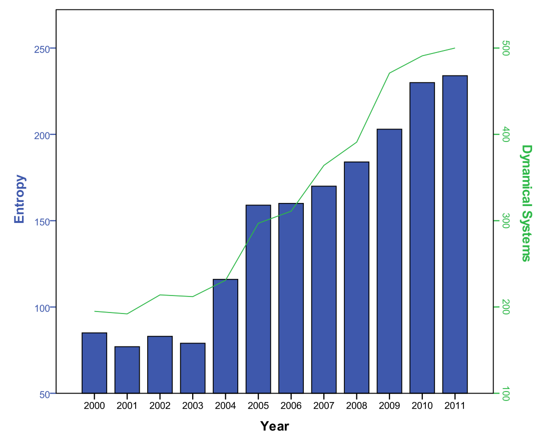


Figure 1. Yearly Pubmed citations from 2000-2011 using keyword search for “*”Entropy (e.g. Approximate entropy, Sample entropy, Shannon Entropy, Multi-scale entropy; (Blue bars) and Dynamical Systems (Green Line).

$$CE[x(t);w](t) = SampEn[\wp(\frac{dx}{dt}(t,t+w)],$$

Where CE is control entropy, x is sample from the time series, t is time, *SampEn* is sample entropy, *dx* is the change in *x* given *dt* the change in time, over window length *w* and partitions $\wp$. Notice especially that since the requirement for stationarity has been alleviated, it therefore is appropriate to measure this entropy as a function of time, and correspondingly, to display time series of CE(t) as a centerpiece in our investigations of changes in the underlying regularity of these complex systems This allows us to monitor and identify changes in the state of the system as it evolves over time.

Regularity statistics, such as CE, are tools that evaluate the variability of a measure, but in contrast to typical measures of variability (e.g. range, variance, standard deviation, etc.), they evaluate the variability in non-linear terms. This is important, because when assessing measures from *in vivo* complex systems such control of human postural stability and gait, there are components of the measurement that may appear to be random noise, but actually contain structured non-linear components that can provide insight into the underlying nature of the system. Systems such as human postural stability and gait are said to be "dynamical systems", that is they are systems that change and evolve in a time dependent fashion. So, in these systems, it is important to have tools that provide a window into non-linear evolution of the characteristics of the system over time. Therefore, the aforementioned issue of stationarity becomes important when applying non-linear techniques such as CE.

Despite the increasing prevalence of regularity statistics (aka entropy measures) in the literature, there is still a great deal of misunderstanding regarding the nature of the measures, their application and interpretation. Specifically regarding CE, an obvious question from the uninitiated might be, "what do changes in CE mean?" Generally, decreases in regularity statistics such as CE indicate increased regularity and reduced complexity (i.e. randomness) while increases in regularity statistics indicate decreased regularity and increased complexity. Here we take the term complexity to denote randomness rather than the notion, from information theory, of Komolgorov complexity, which signifies the algorithmic size of a computer code it takes to reproduce the data. When signal complexity is reduced, it can be inferred that the system is more constrained, and conversely, when signal complexity is increased, the system is more constrained. Further, in the case of CE, because of the unique characteristics of the algorithm, changes in CE are indicative of the controller's "effort" to maintain a given operating parameter. A simple representation of the difference in complexity and entropy of two waveforms is presented in Figure 2. The comparison of the sinewave versus the acceleration waveform can also be thought of in terms of predictability. Since the sinewave (Figure 2a) is perfectly regular, with 0 entropy, it is also perfectly predictable. At any point in time, the value of a point on the sinewave can be predicted with accuracy because the waveform is regular, and hence predictable. In other words, it is not complex. On the other hand, the acceleration waveform (2b) is more complex and less regular. It possesses a pattern that is somewhat regular and predictable, but the non-linear variance is such that the predictability of the waveform is not as certain as the sinewave. The level of unpredictability is another way to think of the complexity or lack of regularity of a system.

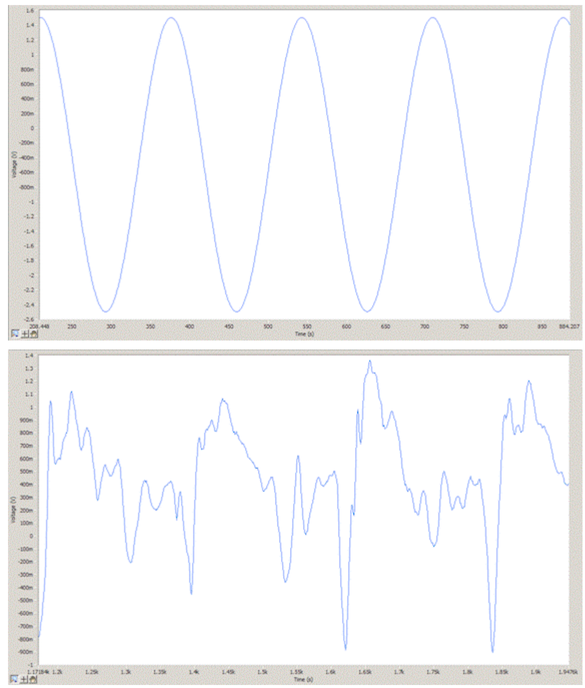


Figure 2. Waveform comparison between a) noiseless sinewave (mean ± S.D = -0.49±1.7) and b) accelerometer AP axis waveform collected from the approximate center of mass of a runner (mean ± S.D = -0.49±1.4). Despite the same mean and similar linear variability, the non-linear regularity (i.e. entropy) of the noiseless sinewave would be 0, while the accelerometer waveform is more complex and entropy would be higher (CE ~ 0.85).

Complexity of Running

To gain insight regarding the constraints of running, McGregor et al. (13) examined the CE of triaxial accelerometry signal obtained during incremental treadmill running in highly trained collegiate runners. Subjects started the trial by walking at 2kph and speed was increased 2 kph every 2 minutes until exhaustion. It was observed that CE increased in all 3 axes with increasing speed until the walk-run transition (8 kph). In the vertical axis, after the walk-run transition, the dominant K-L mode (Figure 3a) exhibited a decreasing CE until exhaustion.

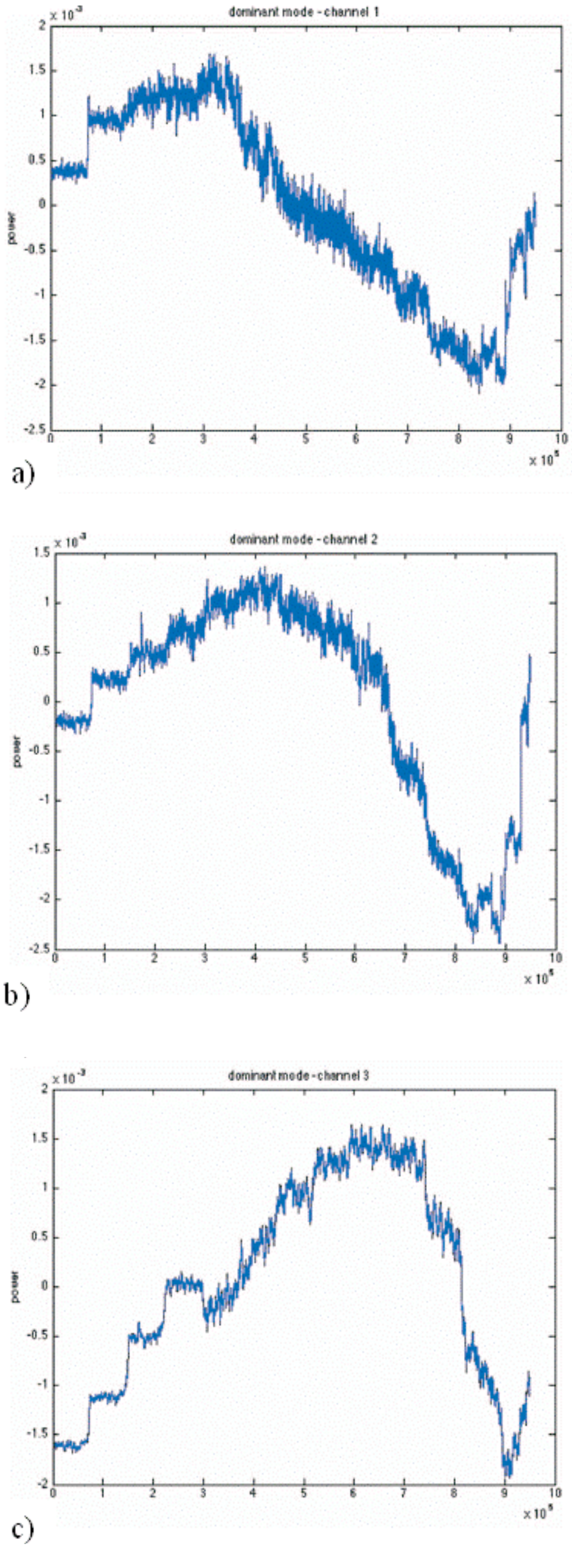


Figure 3. Karhunen-Loeve (KL) analysis: CE of accelerations in the a) VT b) ML c) AP axes. Subjects started at 2 kph after 2 min of quiet standing and speed was increased at 2 kph every 2 min until exhaustion. McGregor, S.J., et al., Chaos, 2009. **19**(2): p. 026109.

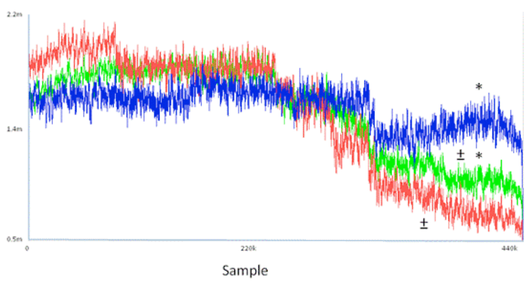
Further, in this axis, at exhaustion, CE was lower than both the initial standing baseline condition and the initial walk stage (2 kph). The "peak" CE response occurred sooner, during the walking phase, relative to the other two axes. The M/L and A/P axes both exhibited more consistent CE responses than the VERT axis, the dominant modes in these axes (Figures 3b and 3c). The CE increased in the A/P well into the run phase and generally peaked latest in this axis (Figure 3c).

This work was unique as no other investigators had previously examined changes in gait characteristics during running using non-linear regularity statistics, particularly in highly trained inter-collegiate runners. This allowed the investigators to examine changes in complexity of accelerometry over a wide range of running speeds, a range that would not be possible with untrained populations.

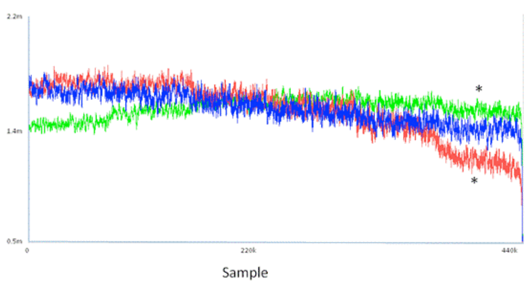
These results were interpreted to mean that with increasing CE, constraints are reduced, and peak CE would be associated with some least constrained parameters. Running in particular is constrained by an interaction between metabolic power generation, elastic and spring characteristics of the tissues and biomechanical coupling of joints involved in the task (21, 22). So, although the metabolic cost of running would be least constrained at 8 kph relative to higher speeds in this population, other specific parameters related to the A/P axis may still be relatively unconstrained as speed increases, because peak CE is not observed until later (18 kph) in the A/P axis (Fig 3c).

A particularly interesting aspect of this study was the observed CE response during the final stages of the test leading up to exhaustion. Because CE is a measure descriptive of system constraint (1) and reduced CE is indicative of increasing constraints, it was anticipated that as subjects approached fatigue, system constraints would be maximal, and this would be reflected in low, possibly minimal, CE. In keeping with this, in all axes, CE was lower at exhaustion relative to the first running stage (Figure 3). In particular, in both the VT and M/L axes, CE was lower at exhaustion than during standing where it might be expected constraints to accelerations of movement would be maximal. In the A/P axis, CE declined throughout the running stages to exhaustion, but CE at exhaustion was approximately equivalent to CE during standing.

To extend this work, we compared CE of triaxial accelerometry between highly trained inter-collegiate runners (HT) and untrained students (UT) (11). This required development of a new statistical approach designed specifically for group comparisons of CE under non-stationary circumstances. This was necessary because the evolution of signals produced by non-linear dynamical systems are time dependent and it is possible that simple comparison of mean CE values of signals that evolve differently over time may provide misleading insights regarding group differences. This new analysis was termed the R test and provided information regarding the "shape" of the evolution of the CE response over time. It was first observed differences in the shape of the CE response between axes (e.g. VT, ML, AP) within groups (Figure 4).



a)



b)

Figure 4. Dominant modes of control entropy responses by axis during incremental running on a treadmill. Control entropy (CE) of accelerations collected in high resolution at the approximate center of mass from a) untrained and b) highly trained runners during an incremental test. Karhunen-Loeve transformation was performed to generate a dominant mode for the CE response in each of three axes (VT = blue; ML = red, AP = green). Like symbols (*,±) indicate significantly different shapes of the dominant modes between axes. Li RDP, McGregor SJ, Busa MA, Skufca JD, Bollt E. *MBE.* Jan 1 2012;9(1):123-145.

These differences indicated that the complexity of movement in the individual axes evolved differently over time for both the HT and UT runners. When individual axes were compared between HT and UT though, the R test showed that shape of the evolution of CE was only different in the VT axis, but not ML or AP. Differences in the mean CE were found in each axis between groups. Because the shape of the CE responses was not different between HT and UT in the ML and AP axes, and the mean CE was significantly higher in both axes for HT compared to UT, we inferred that complexity of the running gait was greater in HT, and they were less constrained, in both axes, than UT. In the case of the VT axis, since the R test did show a significant difference in the evolution of the CE response between groups, the significantly higher mean CT in HT compared to UT should be interpreted with caution. It could be interpreted that the complexity of accelerations in the VT is higher in HT than UT, which would be expected, but results from other work from the same group, presented below, cloud this interpretation. So, it appears as though HT runners are less constrained than UT runners in the ML and AP axis, but it is difficult to say with confidence how their constraints compare in the VT axis.

Complexity of walking

A question that arises when thinking of constraints in terms of locomotion might be, does high fitness confer lower constraints at submaximal efforts? To address this question, we (14) examined the CE of high resolution accelerometry (HRA) while walking between collegiate runners (HT) who exhibited higher maximal aerobic fitness (VO2max) compared to untrained students (UT). A similar protocol to (13) was used for the walking stages, and the R test was used to compare the shapes of responses between groups. In this case, the R test showed that during walking, there were no differences between axes within the UT, but in the HT, both the VT and ML were significantly different than the AP. Between groups though, there were no significant difference in the shape of the CE responses by axis, and the mean CE was not different in the AP between groups. Surprisingly though, the mean CE was lower in both the VT and ML axes in the HT subjects compared to the UT. We interpreted this to mean that, despite their higher maximal cardiovascular fitness, HT runners were more constrained during walking than the UT runners. Therefore, it may be that runners with higher maximal capacity need to impose constraints at submaximal speeds because they are "optimized" to run at faster speeds. Alternatively, it may be that the CE response is indicative of preferred locomotion speed, and the HT are more comfortable at higher speeds, and hence less constrained. At the same time, they are less comfortable at lower speeds of walking, and hence more constrained. Further work will be necessary to answer this question.

Complexity of postural control

To examine the interaction of fatigue with complexity of balance, we (12) compared the CE of HRA at the center of mass while 10 college students maintained uni-pedal postural control following exhaustive exercise. The objective was to determine if fatiguing exercise (2 x Wingate anaerobic tests) would affect the complexity of postural control in the post-fatigue state compared to a post-rest control condition. Balance tests consisted of a series of five single-legged stances, separated by 30 s rest, performed while standing on the dominant leg for 15-s with the participant crossing the arms over the chest and flexing the non-dominant knee to 90 degrees.

R-test comparison of four conditions (pre-fatigue, PreFat; post-fatigue, PostFat; PreRest and PostRest) showed a significant shape difference for the CE response within conditions, between axes in all cases, except the PostFat state. This indicated the shape of the CE response was different between axes, except in the PostFat condition. Within axes though, R-test comparisons between conditions, showed differences were only present in PostFat for AP vs. PreFat ($p < .05$). Therefore, mean CE comparisons within axes, between conditions could be interpreted with confidence. A significantly higher overall CE was observed in the PostFat condition in VT and ML axes compared to PreFat and PostRest conditions. PostFat CE was also higher than PostRest in AP. These were unexpected results, as it was anticipated that fatigue would result in higher constraints of maintenance of postural control and consequently lower CE. Given previous literature regarding AE (2-4), as well as the aforementioned CE response in runners at exhaustion, it would be reasonable to expect a decrease in CE of postural control with fatigue. A higher CE though, could be interpreted outside of the context of constraints to mean that that the complexity of the control signal increased with fatigue in the VT and ML axes PostFat. It may be that postural control is a more complex task than running, or at least a task with more degrees of freedom, and therefore, the complexity of the task is increased with fatigue of some of the motor units involved in the task. The authors compared these observations to those observed by Donker (8) during measurements of postural control with a dual task. Donker argues that the attention is diverted during dual task postural control efforts, and in the case of fatigue, the circumstances may be similar (8). An alternative interpretation might be that fatigued muscles are less competent at executing the controller's commands to correct posture, and this requires exploration of alternative solutions to the postural control problem. This is in keeping with Vaillancourt (24-26) who has argued that changes or differences in complexity need to be viewed within the context of the specific task. The broad generalizations regarding the implications of low complexity, for example, indicating increased constraints or diseased states is too simplistic.

CONCLUSIONS

Here we have presented a summary of how non-linear regularity measures, particularly control entropy, can provide unique insight into the underlying constraints of complex biological tasks. Because the complex interactions of the various components of biological control systems often result in non-linear dynamical responses, tools such as control entropy that can be applied under non-stationary conditions are of great value. Further development of additional statistical tools, such as the R test, will enable the application of control entropy and other related tools under robust conditions that will provide novel insight that was previously unattainable. Using these tools it will be possible to investigate control strategies and related constraints during dynamic conditions such as maximal anaerobic exercise of short duration. Alternatively, it might be possible to sample signals, as opposed to time intervals (e.g. R-R intervals) from measures such as ECG or plethysmography during short duration intense exercise. These approaches are not possible using other currently available regularity statistical approaches. Much insight is yet to be gained from such approaches.

REFERENCES

1. Bollt EM, Skufca JD, and McGregor SJ. Control Entropy: A Complexity Measure for Nonstationary Signals. *Mathematical Biosciences and Engineering* 6: 1-25, 2009.
2. Cavanaugh JT, Guskiewicz KM, Giuliani C, Marshall S, Mercer V, and Stergiou N. Detecting altered postural control after cerebral concussion in athletes with normal postural stability. *British journal of sports medicine* 39: 805-811, 2005.
3. Cavanaugh JT, Guskiewicz KM, Giuliani C, Marshall S, Mercer VS, and Stergiou N. Recovery of postural control after cerebral concussion: new insights using approximate entropy. *Journal of athletic training* 41: 305-313, 2006.
4. Cavanaugh JT, Guskiewicz KM, and Stergiou N. A nonlinear dynamic approach for evaluating postural control: new directions for the management of sport-related cerebral concussion. *Sports Med* 35: 935-950, 2005.
5. Challis JH. Aging, regularity and variability in maximum isometric moments. *J Biomech* 39: 1543-1546, 2006.
6. Costa M, Goldberger AL, and Peng CK. Multiscale entropy analysis of biological signals. *Phys Rev E Stat Nonlin Soft Matter Phys* 71: 021906, 2005.
7. Costa M, Goldberger AL, and Peng CK. Multiscale entropy analysis of complex physiologic time series. *Phys Rev Lett* 89: 068102, 2002.
8. Donker SF, Roerdink M, Greven AJ, and Beek PJ. Regularity of center-of-pressure trajectories

depends on the amount of attention invested in postural control. *Exp Brain Res* 181: 1-11, 2007.

9. Ferrario M, Signorini MG, Magenes G, and Cerutti S. Comparison of entropy-based regularity estimators: application to the fetal heart rate signal for the identification of fetal distress. *IEEE Trans Biomed Eng* 53: 119-125, 2006.
10. Kantz AH, and Schreiber AT. *Nonlinear Time Series Analysis*. Cambridge University Press, 2003.
11. Li RDP, McGregor SJ, Busa MA, Skufca JD, and Bollt E. A statistical approach to the use of control entropy identifies differences in constraints of gait in highly trained versus untrained runners. *Mathematical biosciences and engineering : MBE* 9: 123-145, 2012.
12. McGregor SJ, Armstrong WJ, Yaggie JA, Bollt EM, Parshad R, Bailey JJ, Johnson SM, Goin AM, and Kelly SR. Lower extremity fatigue increases complexity of postural control during a single-legged stance. *J Neuroeng Rehabil* 8: 43, 2011.
13. McGregor SJ, Busa MA, Skufca J, Yaggie JA, and Bollt EM. Control entropy identifies differential changes in complexity of walking and running gait patterns with increasing speed in highly trained runners. *Chaos* 19: 026109, 2009.
14. McGregor SJ, Busa MA, Yaggie JA, Parshad R, and Bollt EM. Control entropy of gait: Does running fitness affect complexity of walking? . *Clinical Kinesiology* 65: 9-17, 2011.
15. Pincus SM. Approximate entropy as a measure of irregularity for psychiatric serial metrics. *Bipolar Disord* 8: 430-440, 2006.
16. Pincus SM. Approximate entropy as a measure of system complexity. *Proc Natl Acad Sci U S A* 88: 2297-2301, 1991.
17. Pincus SM. Assessing serial irregularity and its implications for health. *Ann N Y Acad Sci* 954: 245-267, 2001.
18. Pincus SM. Orderliness of hormone release. *Novartis Found Symp* 227: 82-96; discussion 96-104, 2000.
19. Pincus SM, Gladstone IM, and Ehrenkranz RA. A regularity statistic for medical data analysis. *J Clin Monit* 7: 335-345, 1991.
20. Richman JS, and Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol* 278: H2039-2049, 2000.
21. Saibene F, and Minetti AE. Biomechanical and physiological aspects of legged locomotion in humans. *Eur J Appl Physiol* 88: 297-316, 2003.
22. Sasaki K, and Neptune RR. Muscle mechanical work and elastic energy utilization during walking and running near the preferred gait transition speed. *Gait Posture* 23: 383-390, 2006.
23. Seely AJ, and Macklem PT. Complex systems and the technology of variability analysis. *Crit Care* 8: R367-384, 2004.
24. Vaillancourt DE, and Newell KM. Complexity in aging and disease: response to commentaries. *Neurobiol Aging* 23: 27-29, 2002.
25. Vaillancourt DE, Slifkin AB, and Newell KM. Regularity of force tremor in Parkinson's disease. *Clin Neurophysiol* 112: 1594-1603, 2001.
26. Vaillancourt DE, Sosnoff JJ, and Newell KM. Age-related changes in complexity depend on task dynamics. *J Appl Physiol* 97: 454-455, 2004.

AUTHOR CORRESPONDENCE:
Stephen J. McGregor, Ph.D.
Professor
School of Health Promotion and Human Performance
318 Porter Building
Eastern Michigan University
Ypsilanti, MI 48197
Ph: 734-487-0090
Fax: 734-487-2024