Pediatric Exercise Science, 2011, 23, 17-22

Dynamic Factor Analysis and the Exercise Sciences

Christopher J. Hasson
Northeastern University

Kevin S. Heffernan
Tufts Medical Center

The genesis of human movement involves a complex series of events that depend on the integrated physiology of the neural, muscular, skeletal, metabolic, respiratory and cardiovascular systems. Since each of us has a unique set of physical attributes, such as muscle strength, aerobic capacity, and skeletal structure, our movement patterns serve as another set of "fingerprints". For instance, we can often tell that it is our friend in the distance by the characteristic way that they walk. Key to our ability to distinguish such movement patterns is the way in which the patterns evolve over time. If someone showed you a cine film depicting your friend's walking silhouette, but first chopped up the film, shuffled the pieces around, and glued the film back together, one would likely fail to identify the person in the video as their friend because crucial temporal relationships would be lost.

Detecting Altered Movement Patterns

Of interest to exercise scientists is that our movement patterns change when our physiology is modified due to training, injury, disease, or disability. Some changes are easy to detect, such as favoring one side while walking due to knee pain. Others are not so obvious, such as alterations in postural sway three days after a concussion (1). In the former case, traditional exercise science data analysis may be sufficient. Here, descriptive metrics such as mean or peak values (e.g., peak ground reaction force, joint torque, or joint angle) could allow an investigator to detect deviations from normalcy. However, in the latter concussion case, analysis is more challenging. For example, the peak postural sway amplitude of the injured person may not differ from uninjured individuals, but there may still be important changes in the underlying sway dynamics. As such, more sophisticated human movement analysis techniques are needed. More specifically, when attempting to differentiate physiology from pathophysiolgy, recognizing differences in *patterns* may be more telling than simply examining peak or mean values (2).

Hasson is with the Dept. of Biology, Northeastern University, Boston, MA. Heffernan is with the Molecular Cardiology Research Institute, Tufts Medical Center, Boston, MA.

Dynamic Factor Analysis

Dynamic factor analysis (DFA) is a technique designed specifically for time-series data with the goal of identifying underlying common patterns or trends using all available data, rather than a few isolated points. In general, this concept is not new, and has formed the basis of many nonlinear analysis techniques aimed at reducing the "dimensionality" of data by identifying common patterns (3). For instance, principal component analysis (PCA), developed in the 1930s (4), can distinguish between the walking patterns of normal adults and those with knee osteoarthritis (5). However, a key limitation of PCA is that the temporal ordering of the data are not taken into account. This means that the data can be shuffled in time (just like the cine film of your friend walking), but the result will be the same (6). The advantage of DFA is that the temporal ordering of data are considered, preserving valuable information about the system that produced the data.

Dynamic Factor Analysis in Practice

While DFA has been used successfully to study economic models (7), psychological phenomena (8), oceanographic data (6), hydrologic conditions and water quality (9,10) this powerful analysis technique has yet to be used in the movement sciences, until now. In the current issue of *Pediatric Exercise Science*, Bruno and associates use DFA to examine the power output of young swimmers and soccer players during a standard Wingate test.

When Bruno and colleagues examined lower body peak and mean power output during the Wingate test, they found that swimmers and soccer players attained comparable values. Based on these standard metrics, one would conclude that young soccer players and swimmers have similar power outputs during a Wingate test. However, DFA provided additional insight into the generation of power *over time* in these young athletes. Results from DFA revealed two common patterns or trends underlying the power output time series. One trend was related to a gradual increase in power output for about 10 s until peak power was achieved. This was sustained for close to 5 s, followed by a gradual drop in power (i.e., "late peak power and drop"). The second trend was characterized by an initial spike in power output, then a sharp decrease over roughly 15 s, followed by a gradual drop thereafter (i.e., "early peak power and drop"). In general, the power output of soccer players resembled the former pattern while swimmers were closer to the latter. The authors concluded that DFA improves assessment of power output in young athletes, enhancing overall discriminatory capabilities beyond that achieved by standard metrics alone.

At its core, DFA tries to model experimental data from a group of subjects as a combination of one or more common patterns or trends combined with noise, and can also include one or more explanatory variables (Figure 1). This model can be used to explore three questions: 1) what are the common patterns (trends) among subjects' data points? 2) can the subjects be grouped based on an affinity to one or more patterns (factor loadings)? 3) can these groupings be explained (explanatory variables)? Bruno and associates used DFA to address the first two questions, noting that there are indeed common Wingate power output patterns among swimmers and soccer players, and the athletes can be grouped based on these patterns. Unfortunately, the authors were only able to speculate on the physiologic mechanisms

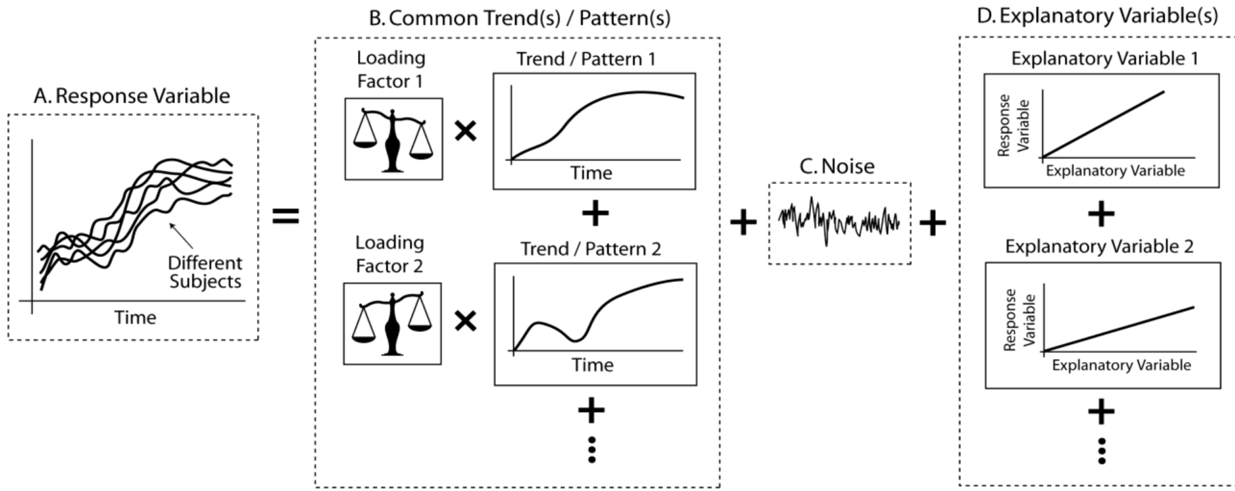

Figure 1 — Pictorial representation of dynamic factor analysis (DFA). Hypothetical response variable (e.g., change in power output over time) for a group of subjects (A). Each subject's data can be represented by a model consisting of several components (B-D). The first component is one or more trends or patterns (B). When more than one trend is included, each trend is weighted by a loading factor which characterizes the strength of the trend's relation to the data. The trends combine in a linear manner by addition. The second component is noise, which models the degree of random variation in the data (C). The third optional component is one or more explanatory variables (e.g., lactate accumulation, heart-rate, perceived exertion), which are used in a linear regression model to quantify the strength of relationship between the explanatory variable(s) and the response data (D).

that may explain these differences. One possibility put forth was related to differences in neuromuscular activation from sport-specific training adaptations. Had such data been collected (i.e., muscle activity recorded using electromyography), this information could have been entered into the DFA model as an explanatory variable to examine possible associations with the power output patterns. DFA can be looked at as a “super” regression model. Instead of examining correlates of a single summary metric (i.e., mean power output) DFA can provide information on correlates (explanatory variables) of *patterns* that emerge over time.

Significance of Dynamic Factor Analysis

The ability to distinguish groups and identify individuals at risk for injury, disease, or disability is of paramount importance to clinicians and researchers alike. Basing analyses on discrete metrics such as mean and peak values ignores a wealth of information embedded within the temporal dynamics of physiological data. For example, altered timing of oxygen uptake during the transition from rest to exercise differs with age (11) and has been witnessed in several disease states (e.g., heart failure, diabetes, peripheral artery disease, pulmonary hypertension; 12). This provides useful information on functional capacity, response to therapy, and long-term clinical outcome (12). Moreover, from a sports science perspective, oxygen uptake kinetics has been shown to be predictive of exercise performance (13). Thus, identification of temporal patterns in physiologic output is important; DFA may complement current techniques and further help clinically stratify high risk patients.

Applications to Human Movement

DFA may be of great value to future investigations in human movement and could have broad applications throughout all subdisciplines of exercise science. For example, DFA may be able to detect abnormal movement patterns among a group of individuals, and explain why individuals exhibit such patterns, possibly due to disease or injury. DFA could also be used to identify those at an increased risk of injury. For example, individuals who show increased knee instability may exhibit movement patterns that increase the likelihood of tearing their knee ligaments in future sports activities. Although Bruno and colleagues focused on differentiating between individuals, DFA could be used to examine how physiologic patterns change over time within individuals, such as with learning/training or dietary/drug interventions.

Predicting Future Trends

DFA can also be used to predict future trends. For example, Zuur and colleagues suggested that with 25–30 years of data on changes in fish populations, one might be able to predict future trends for the next 2–3 years using DFA (6). The predictive capabilities of DFA depend in part, on how well the model (trends + noise [+ explanatory variables]) characterizes the actual processes underlying the data. More baseline data may provide better estimates of these processes, improving future predictions. This capacity for prediction may allow researchers and clinicians to address very novel questions using DFA. Can the examination of patterns in heart rate and blood pressure before an exercise stress test predict development of an arrhythmia or severe hypotension following said exercise bout? Can the examination of respiration patterns over time predict occurrence of an exercise-induced asthmatic episode? More research will be needed to examine the utility of DFA for predicting physiologic outcomes in the exercise sciences.

Limitations to Dynamic Factor Analysis

Although DFA has potential as a useful analysis technique for exercise scientists, it is prudent to consider its limitations. A practical issue is that DFA is computationally intensive. Bruno and colleagues used custom DFA software (14), in which run-times (time taken to crunch the data) could exceed several hours as the number of subjects and common trends grows large (6). The results also tend to become inconsistent with such large data sets (6). Another limitation of DFA is that the common trends are combined in a linear fashion, and the explanatory variable regressions are linear as well. Therefore, nonlinear interactions between the components of the model are ignored. Finally, interpretation of DFA results may not be straightforward. The DFA model uses hypothetical latent variables that are deemed to be responsible for the observed patterns; however, no information is provided as to what these variables are. Adding explanatory variables to the model could help with interpretation, but this increases complexity and does not always improve the model. In general, one must keep in mind that when using advanced techniques such as DFA, extra care may be needed when interpreting results. This was perhaps best described by Spider-Man, who remarked, ". . . with great power there must also come – great responsibility!" (15).

Conclusion

As exercise scientists, we should carefully scrutinize our data and explore all that it has to offer. The time-varying patterns in physiological data reflect the properties of the system that produced that data—the human body. As such, analysis of physiologic signals with advanced techniques, such as DFA, may help us better understand movement patterns in health, disease and disability throughout the lifespan. At a minimum, these tools will allow us to recognize that it is indeed our friend walking in the distance.

References

1. Cavanaugh, J.T., K.M. Guskiewicz, C. Giuliani, S. Marshall, V.S. Mercer, and N. Stergiou. Recovery of postural control after cerebral concussion: new insights using approximate entropy. *J. Athl. Train.* 41:305–313, 2006.
2. Heffernan, K.S., J.J. Sosnoff, E. Ofori, et al. Complexity of force output during static exercise in individuals with Down syndrome. *J. Appl. Physiol.* 106:1227–1233, 2009.
3. Donoho, D. *High-dimensional data analysis: The curses and blessings of dimensionality*. AMS Math. Challenges Lecture, 2000.
4. Hotelling, H. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* 24:417–441, 1933.

5. Deluzio, K., and J. Astephen. Biomechanical features of gait waveform data associated with knee osteoarthritis: An application of principal component analysis. *Gait Posture.* 25:86–93, 2007.
6. Zuur, A., I. Tuck, and N. Bailey. Dynamic factor analysis to estimate common trends in fisheries time series. *Can. J. Fish. Aquat. Sci.* 60:542–552, 2003.
7. Sezgin, F., and B. Kinay. A dynamic factor model of the evaluation of the financial crisis in Turkey. *Bull. Soc. Sci. Med. Grand Duche Luxemb.* (Spec No 1):109–117, 2010.
8. Molenaar, P.C. The future of dynamic factor analysis in psychology and biomedicine. *Bull. Soc. Sci. Med. Grand Duche Luxemb.* (issue 2):201–213, 2006.
9. Munoz-Carpena, R., A. Ritter, and Y.C. Li. Dynamic factor analysis of groundwater quality trends in an agricultural area adjacent to Everglades National Park. *J. Contam. Hydrol.* 80:49–70, 2005.
10. Kuo, Y.M., and F.J. Chang. Dynamic factor analysis for estimating ground water arsenic trends. *J. Environ. Qual.* 39:176–184, 2009.
11. Armstrong, N., and A.R. Barker. Oxygen uptake kinetics in children and adolescents: a review. *Pediatr. Exerc. Sci.* 21:130–147, 2009.
12. Poole, D.C., T.J. Barstow, P. McDonough, and A.M. Jones. Control of oxygen uptake during exercise. *Med. Sci. Sports Exerc.* 40:462–474, 2008.
13. Jones, A.M., and M. Burnley. Oxygen uptake kinetics: an underappreciated determinant of exercise performance. *Int J Sports Physiol Perform.* 4:524–532, 2009.
14. Zuur, A., R. Fryer, I. Jolliffe, R. Dekker, and J. Beukema. Estimating common trends in multivariate time series using dynamic factor analysis. *Environmetrics.* 14:665–685, 2003.
15. Lee, S., and S. Ditko. *Amazing Fantasy #15*. New York: Marvel Comics, 1962.