

Forum

Behavioral syndromes: carryover effects, false discovery rates, and a priori hypotheses

Ned A. Dochtermann

Department of Biology, University of Nevada, Reno, NV 89557, USA

Behavioral syndromes and animal personalities are an increasingly active area of research within behavioral ecology (Bell 2007), and this interest stems in part from their potential evolutionary implications. In other areas of evolutionary ecology, trait covariances—like those observed in behavioral syndromes—have been demonstrated to have profound evolutionary implications (Lande and Arnold 1983). Trait covariances can dramatically affect how populations respond to selection by potentially constraining responses, generating trade-offs, or otherwise shaping evolutionary trajectories (Blows and Hoffmann 2005; Roff and Fairbairn 2007; Walsh 2007). Via these same evolutionary responses, behavioral syndrome research may well have similarly profound ramifications for our understanding of behavioral evolution. However, as with many novel areas of research, methodologies for the study of behavioral syndromes are still developing (e.g., Dingemanse et al. 2010).

While examining behavioral syndrome structure in hissing cockroaches (*Gromphadorhina portentosa*), Logue et al. (2009) discussed 2 key methodological issues for the study of behavioral syndromes: 1) How the order in which behaviors are assayed can affect subsequent behavioral responses and 2) How the evaluation of numerous behaviors results in a risk that correlations are statistical artifacts. Both of these aspects are of importance for the study of behavioral syndromes and warrant additional discussion as they may influence the inferences that researchers draw. Here, I focus on extending the discussion of order effects, their implications to behavioral syndrome research, and how to methodologically treat them. I also expand on some of the points made by Logue et al. (2009) regarding the evaluation of numerous behavioral correlations and discuss the role of confirmatory analyses.

ORDER EFFECTS

Order effects, the effects that the order in which behavioral assays are conducted have on subsequent behavioral responses, are an important issue for research in behavioral ecology (Díaz-Uriarte 2002). For behavioral syndrome research order effects can exaggerate or dampen the observed correlations, changing the inferences researchers draw (Logue et al. 2009). However, order effects are only one aspect of a group of methodological confounds more generally known as “carryover effects.” Carryover in general is the effect of past treatments, or behavioral assays, on subsequent behavior. Carryover effects are often, and inappropriately, ignored during the analysis of behavioral data (Díaz-Uriarte 2002). Carryover effects include the general effects of repeated testing (e.g., habituation to imposed stimuli) as well as specific effects of how behavioral stimuli are presented temporally (e.g., “winner” or “loser” effects). As a group, these effects can bias

results and can decrease the statistical power available to detect changes in behavioral responses (Díaz-Uriarte 2002).

Fortunately carryover effects can be dealt with in several ways. First, the presence of carryover effects can be explicitly evaluated using randomization testing (Logue et al. 2009). In this approach, individuals have their behavioral responses to a variety of conditions recorded in several different sequences of testing (e.g., for some individuals aggressive behavior would be recorded first and others would have foraging behavior recorded first). The difference in the magnitude of behavioral correlations between groups of individuals with specific testing sequences can then be estimated. This difference is then evaluated by shuffling individuals between groups at random to determine the null distribution of how large the difference between groups should be. Using this randomization testing approach, Logue et al. (2009) demonstrated that the observed correlations between latency to move and latency to forage as well as between courtship directed toward males and females in hissing cockroaches were not due to the order in which behavioral assays were conducted.

As an alternative to the randomization approach, carryover effects can be addressed methodologically by how behavioral assays are conducted. This alternative approach has the advantage of allowing researchers to statistically remove carryover effects from the behavioral correlations of interest. This removal of potential confounds may also allow a more precise estimate of effects and behavioral covariances.

Carryover effects can be controlled methodologically by assigning the order in which treatments are presented not randomly but rather by balancing the order in which treatments are presented (Díaz-Uriarte 2001). For example, if a researcher is interested in 4 behaviors that are often found to covary within a behavioral syndrome—like aggression toward conspecifics (AggC), aggression toward heterospecifics (e.g., boldness), routine formation, and activity—then 4 orders of treatment presentation (sequences) could be constructed that balance carryover effects:

Sequence	Testing period			
	1	2	3	4
1	AggC	Activity	Boldness	Routine formation
2	Boldness	AggC	Routine formation	Activity
3	Routine formation	Boldness	Activity	AggC
4	Activity	Routine formation	AggC	Boldness

This pattern of how behaviors are sampled, also called a Williams design (Díaz-Uriarte 2001), results in 4 sequences of how assays can be presented and has the characteristics that each assay is preceded and followed with each possible behavior (and the beginning or end of testing). Every behavioral assay is also conducted during each period of testing. Individuals can then be randomly assigned to one of the sequences in a balanced manner. By balancing the order of behavioral assays, Williams designs allow for the statistical control and testing of

the effects behavioral assays on the next assay conducted. However, incorporating sequence and period reduces residual degrees of freedom which may decrease statistical power, so proper attention should be paid to sample sizes. Using this sort of design, both sequence and period can then be statistically modeled using mixed-effects models, which will also accommodate data missing completely at random (Díaz-Uriarte 2001).

For behavioral syndrome research, individual behaviors can be treated as response variables of sequence and period. The residuals from the analysis (or the best linear unbiased predictors if random-effects models are used) can then be used to test for behavioral correlations with any potential order effects removed. Thus, the analysis of behavioral syndromes is not conflated with the methodological issue of what order behaviors were tested. Alternatively, a mixed model multivariate analysis of variance (MANOVA) could be used incorporating sequence, period, and proposed ecological causes as independent variables with the suite of observed behaviors acting as dependent variables. With this MANOVA approach, causal factors underlying variation in the suite of behaviors constituting a behavioral syndrome can be tested. Díaz-Uriarte (2001) provides further discussion of the statistical details for applying a MANOVA approach to data obtained from designs controlling for carryover effects. Unfortunately, to my knowledge, MANOVA methods have not yet been applied to the study of behavioral syndromes.

The methodological control of carryover effects also allows researchers to examine whether or not there are interactions between behaviors and the order in which they are measured (Díaz-Uriarte 2002). For example, if the same 4 behaviors were examined for males and females, the interaction between sex and sequence or period can also be tested. This might reveal differences in how sexes or other groups respond to changing conditions. Díaz-Uriarte (2001, 2002) presents an in-depth discussion of the treatment of carryover effects that is useful not just to behavioral syndrome research but all experimental approaches to the study of behavioral ecology.

THE CONSEQUENCES OF TESTING ALL THE POSSIBLE BEHAVIORAL CORRELATIONS

In addition to discussing order effects, Logue et al. (2009) identify another key issue with current methods of characterizing behavioral syndromes: researchers often cannot determine whether “statistically significant” correlations are statistical artifacts or representative of real trait correlations. For example, if a researcher is examining 7 behaviors there will be 21 bivariate correlation coefficients tested ($n = \frac{b(b-1)}{2}$; where n is the number of correlations and b the number of behaviors examined). For every 20 behavioral correlations tested one would be expected to be significant by chance alone (assuming $\alpha = 0.05$). Thus at least one correlation from our example will likely be identified as significant due simply to chance.

The traditional method of controlling for this problem is the Bonferroni correction (Quinn and Keough 2002). However, Bonferroni corrections often result in a general loss of power when controlling for spurious results (Nakagawa 2004). How then should researchers evaluate the correlations that constitute behavioral syndromes? This is a question that has also been faced in bioinformatics where hundreds to thousands of relationships are often tested.

Within bioinformatics one method by which this issue has been dealt with is by calculating the so-called “false discovery rate” (Storey and Tibshirani 2003). The false discovery rate is an estimation of how many observed significant relationships

are actually null effects. For behavioral syndrome research this would essentially be how often a behavioral correlation is spuriously identified as significant. This makes the false discovery rate a potentially powerful tool for behavioral ecologists as it may allow the more rigorous demonstration that behaviors within a behavioral syndrome are actually correlated. Consistent with this, Logue et al. (2009) controlled for the false discovery rate and demonstrated that behaviors of hissing cockroaches covaried as part of a behavioral syndrome.

When researchers evaluate all the possible correlations between measured behaviors they are essentially conducting an exploratory analysis (Bell 2007) and should follow the lead of Logue et al. (2009) and quantify and control for the false discovery rate. Analyses where all possible correlations are tested may also lead to a misestimating of the magnitude of correlations (Zhang 1992). However, testing all the possible behavioral correlations is not the only method by which researchers can investigate behavioral syndromes. Although the description of behavioral syndromes in an ever-increasing number of taxa is highly informative, there are now also numerous hypotheses regarding behavioral syndrome structure (what behaviors covary and whether positively or negatively). Thus researchers can test a priori hypotheses about behavioral syndrome structure.

The confirmatory testing of explicit and a priori hypotheses allows researchers to draw more general inferences than possible with exploratory analyses (Chatfield 1995) and require fewer statistical comparisons; reducing the risks identified by Logue et al. (2009) which necessitated the calculation of the false discovery rate. A focus on confirmatory tests of a priori hypotheses also allows concrete evolutionary questions about behavioral syndromes to be asked. For example, is the expression of behavioral syndrome structure under predation pressure a general phenomenon (Bell and Sih 2007; Dingemanse et al. 2007)? Are the behavioral covariances we observe in behavioral syndromes representative of underlying genetic correlations between behaviors (Dingemanse et al. 2009; Réale et al. 2009)? How and why do patterns of trait (behavioral) covariance differ between populations (Phillips and Arnold 1999)? It is the answer to these and similar questions that will provide key insights for our understanding of the evolution and evolutionary significance of behavioral syndromes in the future.

I thank David Logue and Tim Roth for helpful comments on an early version of this paper.

Key words: behavioral syndromes, carryover effects, family-wise error rates, personalities.

Address correspondence to N.A. Dochtermann. E-mail: ned.dochtermann@gmail.com.

Received 4 November 2009; revised 29 January 2010; accepted 31 January 2010.

REFERENCES

- Bell AM. 2007. Future directions in behavioural syndromes research. *Proc R Soc B Biol Sci.* 274:755–761.
- Bell AM, Sih A. 2007. Exposure to predation generates personality in threespined sticklebacks (*Gasterosteus aculeatus*). *Ecol Lett.* 10: 823–834.
- Blows MW, Hoffmann AA. 2005. A reassessment of genetic limits to evolutionary change. *Ecology.* 86:1371–1384.
- Chatfield C. 1995. Model uncertainty, data mining and statistical inference. *J R Stat Soc Ser A Stat Soc.* 158:419–466.

- Díaz-Urriarte R. 2001. The analysis of cross-over trials in animal behavior experiments: review and guide to the statistical literature [Internet]. Samizdat Press. Available from: <http://samizdat.mines.edu/>.
- Díaz-Urriarte R. 2002. Incorrect analysis of crossover trials in animal behaviour research. *Anim Behav.* 63:815–822.
- Dingemanse NJ, Dochtermann N, Wright J. 2010. A method for exploring the structure of behavioural syndromes to allow formal comparison within and between datasets. *Animal Behaviour.* 79: 439–450.
- Dingemanse NJ, Van der plas F, Wright J, Réale D, Schrama M, Roff D, Van der Zee E, Barber I. 2009. Individual experience and evolutionary history of predation affect expression of heritable variation in fish personality and morphology. *Proc R Soc B Biol Sci.* 276: 1285–1293.
- Dingemanse NJ, Wright J, Kazem AJN, Thomas DK, Hickling R, Dawnay N. 2007. Behavioural syndromes differ predictably between 12 populations of three-spined stickleback. *J Anim Ecol.* 76: 1128–1138.
- Lande R, Arnold SJ. 1983. The measurement of selection on correlated characters. *Evolution.* 37:1210–1226.
- Logue DM, Mishra S, McCaffrey D, Ball D, Cade WH. 2009. A behavioural syndrome linking courtship behavior toward males and females predicts reproductive success from a single mating in the hissing cockroach, *Gromphadorhina portentosa*. *Behav Ecol.* 20: 781–788.
- Nakagawa S. 2004. A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behav Ecol.* 15: 1044–1045.
- Phillips PC, Arnold SJ. 1999. Hierarchical comparison of genetic variance-covariance matrices. I. Using the Flury hierarchy. *Evolution.* 53:1506–1515.
- Quinn GP, Keough MJ. 2002. *Experimental design and data analysis for biologists.* Cambridge: Cambridge University Press.
- Roff DA, Fairbairn DJ. 2007. The evolution of trade-offs: where are we? *J Evol Biol.* 20:433–447.
- Réale D, Martin J, Coltman DW, Poissant J, Festa-Bianchet M. 2009. Male personality, life-history strategies and reproductive success in a promiscuous mammal. *J Evol Biol.* 22:1599–1607.
- Storey JD, Tibshirani R. 2003. Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A.* 100:9440–9445.
- Walsh B. 2007. Escape from flatland. *J Evol Biol.* 20:36–38.
- Zhang P. 1992. Inference after variable selection in linear regression models. *Biometrika.* 79:741–746.