High Enthusiasm and Low R-Squared

David Houle

Evolution, Volume 52, Issue 6 (Dec., 1998), 1872-1876.

Stable URL:
http://links.jstor.org/sici?sici=0014-3820%28199812%2952%3A6%3C1872%3AHEALR%3E2.0.CO%3B2-Z

BOOK REVIEWS

Evolution, 52(6), 1998, pp. 1872–1876

HIGH ENTHUSIASM AND LOW *R*-SQUARED[1]

DAVID HOULE
Department of Zoology, University of Toronto, Toronto, Ontario M5S 3G5, Canada
E-mail: dhoule@zoo.toronto.edu

Received October 20, 1998.

Asymmetry, Developmental Stability, and Evolution, by Anders P. Møller and John P. Swaddle, is an enthusiast's guide to the idea that subtle deviations from perfect symmetry can indicate the quality of individuals, or of the environments they inhabit. Møller and Swaddle argue that the value of these deviations as indicators of fitness is considerable, despite a body of evidence that the relationship is usually weak when it occurs, and is often undetectable. Population biology is periodically swept by such enthusiasms, by which I mean an idea which captures our imagination because it seems to explain a great deal, although it is supported by little confirmatory evidence. This jump to conclusions would not be a bad thing if our enthusiasms panned out more often, the way that, for example, the enthusiasm which greeted Watson and Crick has. However, in our field, yesterday's enthusiasm is more often today's cautionary tale about the complexity of the phenomena we try to explain.

My own field of evolutionary genetics is unfortunately an example of this process. The central problem of what maintains the genetic variation essential to all evolutionary change is unsolved, despite more than century of work. Enthusiasms have arisen for mutation as the major source, then for genetic drift, then for overdominant selection, then for coadapted gene complexes, then for mutation-drift balance, for marginal overdominance, for mutation-selection balance, and well, you get the idea. Although this list is peculiar to my field, I trust that its recitation may bring a shudder of recognition to those in other fields.

All this meandering is not without its good points. Enthusiasm can inspire critical tests of the new idea, or informative theorizing, increasing our understanding of the options. However, too many scientists are prepared to assume the validity of the latest idea rather than seek such critical tests. The result is that enthusiasm can have a distorting effect on a whole field, as it frequently leads to shoddy thinking and ill-motivated research. There are always more people ready to don skinny ties or baggy pants when others do so than to take a hard look at their fit.

My experience with scientific enthusiasm dates to my first days in graduate school, where I found my advisor in the grip of an enthusiasm for marginal overdominance as an explanation for allozyme polymorphisms. The data from some of the most active labs in the field was consistently revealing significant correlations between heterozygosity and whatever correlates of fitness fell to hand. The apparent ubiquity of these correlations suggested that some fundamental and general phenomenon was being revealed; my advisor was convinced that this phenomenon was a general form of heterozygote advantage. True, heterozygosity never did explain much of the variance in the fitness correlates, but theory clearly shows that only a small heterozygote advantage is necessary to maintain genetic variance indefinitely. I was both fascinated and skeptical, as I simultaneously soaked in the enthusiasm, learned the fundamentals of population genetics, and noted the glazed expressions of nonenthusiasts.

My experiments began with a study of the effect of allozyme genotype on development time in fruit flies. How well I remember my excitement as the bands swam up through the agar overlays, and how I agonized over the fuzzy lanes which were less clear than their neighbors. I was so eager for the results that I initially ran only the fastest and slowest developing flies, and then checked to see if they differed in heterozygosity. They did! In the end, I explained about 5% of the variance in development time, which satisfied me as a typical value for this sort of experiment.

The problem came when I tried to repeat the result. I did more experiments, increasing the sample size each time, but was never able to find any significant heterosis in my flies again. Thinking back to the first experiment, I remember the enthusiasm, the longing for the result, the fact that I knew the development time of each batch of flies I ran, the fuzzy lanes which might have been fuzzy because they were heterozygotes—or for some other reason. Mistrusting myself, in each subsequent experiment I had someone else code the samples in such a way that I did not know the phenotypes of the flies on the gels. I do not really know why similar experiments gave me different results, but I do know that the number of hopeful misscorings necessary to give an R^2 of 5% is soberingly tiny.

The parallels of the current enthusiasm for asymmetry studies with that for heterosis are striking. In each case, much of the appeal rests in the supposition that a few simple measurements of predictor variables (genic heterozygosity or morphological asymmetry) can reveal something important about the fitness of the organism. In each case, the proportion of phenotypic variance explained by the predictors is usually small. In each case, a substantial body of evidence does not show the predicted relationship. In each case, the ideas are promoted by a few scientists who are adept at observing the phenomenon on a regular basis.

Møller is one of the chief adepts of asymmetry studies and has produced both empirical studies and quantitative reviews in a remarkably prolific manner. Therefore, we know that this

[1] *Asymmetry, Developmental Stability, and Evolution.* Anders Pape Møller and John P. Swaddle. 1998. Oxford University Press, Oxford, U.K. xi plus 291 pp. HB $90.00, ISBN 0-19-854895-8; PB $39.95, ISBN 0-19-854894-X.

book will promote the importance of asymmetry to biology. What then should we look for in an enthusiast's manifesto? I look for an effective exposition of the basic hypotheses, a clear description of the evidence favoring the hypotheses, and hope for a revealing attempt to deal with competing ideas, even if colored by enthusiasm for the favored alternative. I must fault *Asymmetry, Developmental Stability, and Evolution* in the first two of these areas. Møller and Swaddle seem to have little grasp of the theory and consistently distort the evidence relevant to their ideas. Their failure to adequately consider alternative hypotheses is, however, very revealing.

The ruling premise of Møller and Swaddle's book is that the ability to develop in a predictable manner to an optimum phenotype is a good indicator of fitness. This property is defined as developmental stability. In most cases, however, we do not know what the optimum phenotype is, particularly as the optimum can vary from situation to situation. An important set of exceptions are morphological structures repeated on either side of an axis of symmetry. We can usually assume that, whatever the optimum is, it is the same on both sides of a symmetrical body. Precise measurements of such repeated structures often reveal slight deviations from perfect symmetry, usually referred to as fluctuating asymmetry (FA). Fluctuating asymmetry may therefore provide an indicator of developmental stability.

Møller and Swaddle go one step further, and effectively treat FA and developmental instability as the same thing, as in this passage: "As fluctuating asymmetry represents the ability of an individual to resist genetic and environmental factors during growth and development, then it could be viewed as a general health certificate of an individual" (pg. 111). This glosses over the crucial question of whether asymmetry does reflect anything general about development. Møller and Swaddle's statement would be justified if they had earlier reviewed the evidence for this claim and found it to be supported. However, quite the opposite is true. Møller and Swaddle's review of the evidence fairly demolishes any strong basis for the connection between fitness and asymmetry.

This review is found on pages 53–55, where Møller and Swaddle take up the seemingly arcane point of whether there is a correlation between asymmetries of different traits, that is whether there is an "asymmetry parameter" which reflects an individual's overall asymmetry. If asymmetry is a predictor of fitness, the proximal cause of the fitness differences can be the asymmetries themselves, if, for example, performance or mating success is directly affected by asymmetry. Alternatively, the asymmetries may be indicators of some underlying variable which directly affects both fitness and developmental instability. If asymmetry indicates either some property of the developmental environment, or the quality of the individual, we would expect that the degree of asymmetry of different traits would be positively correlated. Such correlations would indicate the existence of an asymmetry parameter, which might reasonably reflect an individual's developmental stability and, therefore, its fitness.

Now the majority of asymmetry studies have utilized traits chosen for their ease of measurement, rather than for any particular biological significance. A good example of such arbitrary traits are sternopleural bristles in *Drosophila melanogaster,* which are insignificant tufts of hairs on either side of the thorax. It is very difficult to imagine that bristle asymmetry affects fitness directly, although it may reflect some deeper property of the developmental environment or genome of the individual. If it does, then we would expect to be able to choose other arbitrary paired structures and observe that the asymmetries of these traits are more likely to be high when sternopleural bristles are asymmetrical. However, if there is no correlation, and the asymmetry of one trait does not even predict asymmetry of other traits, it would be difficult to see how asymmetry can predict anything else. On page 53, Møller and Swaddle show that the majority of studies find no correlations of asymmetries among individuals, and that even when significant correlations are found they are still very small. The situation is a bit more promising at the level of populations, as here the majority of cited studies do show positive correlations of average asymmetries of different traits across populations, although even these correlations are usually small.

Møller and Swaddle go on to list the sensible reasons that an asymmetry parameter might not exist. First, the potential perturbations of the developmental system may affect different traits in different ways, due to differences in the timing of development or to independence of the developmental processes underlying different traits. Second, they dismiss the possibility of general buffering processes to deal with these perturbations with the categorical assertion that "It is inconceivable that a single mechanism could effectively buffer a large number of potentially different developmental processes against a wide variety of environmental or genetic perturbations." Third, they note that, in the accepted model for the development of asymmetry, all individuals are more likely to be symmetrical than to have any particular degree of asymmetry, including those individuals with the lowest developmental stability. Any one measurement of asymmetry will be a poor indicator of the developmental stability which may underlie it (Whitlock 1996; Houle 1997). Even where an asymmetry parameter exists it will be difficult to detect. Despite these convincing empirical and conceptual conclusions, throughout the rest of the book we are exhorted to the completely contradictory view that FA "render[s] very sensitive measures of developmental stability" (p. 36).

Other less disastrous inconsistencies abound. For example, the issue of whether all forms of asymmetry are likely to reflect developmental stability is currently being debated. Palmer (e.g., 1996) has claimed that asymmetry where the ordered difference (i.e., right side–left side) between paired structures has a mean different from zero (directional asymmetry), or where the difference has two modes on either side of perfect symmetry (antisymmetry), does not reflect developmental stability. The reasoning is that cases of striking directional or antisymmetry, such as the asymmetric claws of some crustaceans, are probably adaptations, rather than flaws. In chapter 1, Møller and Swaddle take this position, stating that Palmer and Strobeck "refute" a connection between other types of asymmetry and developmental stability (p. 18), and that it is "imperative" that different forms of asymmetry are distinguished statistically before analysis (p. 21). Yet on page 51, Møller and Swaddle repeat the arguments of Graham et al. (1993) that all three kinds of asymmetry

may be informative about developmental stability. Møller and Swaddle seem to adopt the approving language of each source for their own ideas, with little attempt to relate or compare these viewpoints to one another.

Another sign of this uncritical use of language is that Møller and Swaddle frequently do not even try to define their terms. For example, in chapter 2, they launch into the full terminology of nonlinear dynamics, throwing around such terms as "phase-lock," "fractal dimension," and "attractor." In chapter 7, we suffer "vector sum of flow," "turning moments," "vortices," and "angle of attack." Although all these terms no doubt have their place, their undefined usage in a wide-ranging book is reminiscent of a certain class of undergraduate paper where jargon substitutes for understanding.

This suspicion that Møller and Swaddle do not really understand much of the material they discuss is amply confirmed in my own area of expertise. Chapters 3 through 5 essentially promote the view that genetic and phenotypic variance are finely tuned adaptations, rather than the more usual view of variances as epiphenomena caused by evolutionary forces acting primarily on means or norms of reaction. Recently, I suggested that there are seven distinct evolutionary processes which can affect genetic variance (Houle 1998). Møller and Swaddle consider ideas related to only two of these, the evolution of modifiers of variance and the evolution of mutation rates. Modifiers are genes which affect the properties of how other genes are expressed (or how often they mutate). Although modifier theory is a respected branch of population genetics, most population geneticists would regard evolution at loci which directly affect a trait as more important determinants of variance through such events as selective sweeps, fixation of alleles, or balanced polymorphisms. All of these processes hardly rate a mention from Møller and Swaddle.

This narrow approach to the evolution of variance would not be so damaging if Møller and Swaddle could give a convincing case for the importance of modifier evolution. Instead they show little evidence that they even understand their favorite hypothesis. For example, the early part of chapter 3 is devoted to the impact of selection on modifiers of variance, with the conclusion being that stabilizing selection will favor modifiers that decrease the phenotypic variance (or canalize the phenotype), and directional selection will favor modifiers that increase variance and decanalize the phenotype. Møller and Swaddle do not seem to realize that both of these expectations are not general. Directional selection alone is not sufficient to favor decanalizing modifiers, which are favored only when fitness rises faster than linearly with trait value (Lande 1980). This seems unlikely to be generally true even for directionally selected traits (Rowe and Houle 1996). For example, a 1% increase in survivorship to adulthood is likely to lead, at most, to a 1% increase in absolute fitness, regardless of the mean survivorship. This is a linear increase in fitness with trait value. If one takes potential tradeoffs between survival and other traits into account, for example if high survival comes at the expense of slow growth, it is even possible that the processes affecting survival are not under directional selection at all, even when survival is.

A similar problem for Møller and Swaddle's simplistic view of modifier evolution arises under stabilizing selection, where they expect that a modifier will be favored if it reduces phenotypic deviations from the optimum. However, Wagner et al. (1997) have recently shown that when canalization extends to genetic perturbations, deleterious alleles at loci that directly influence the trait are subjected to less intense selection, allowing those alleles to increase in frequency, thus increasing the phenotypic variance and lowering fitness. Therefore, stabilizing selection can favor either modifiers which increase the variance or those which decrease it. Although Møller and Swaddle briefly cite Wagner et al. (p. 81), they do not seem to realize the importance of their results.

Other astonishing assertions are sprinkled throughout the conceptual sections of the book. In chapter 2, we are led to believe that "chaos theory" explains the regulation of development, a fact that has escaped the attention of the developmental biologists I know. We learn that "directional selection in almost all studies has resulted in an increase in variance components" (p. 76), which will come as welcome news to animal and plant breeders in particular. The conclusion that "mutation rates increase under a variety of environmental conditions that deviate from the most commonly encountered environments" (p. 94) is considerably more understandable when we realize that these environmental conditions include presence of mutagens like ultraviolet light and "infection with parasitic DNA." The statement that "selection arises as an effect of a discrepancy between environmental conditions experienced by an individual and the conditions optimal for maintenance, growth, mating, reproduction, and survival" (pg. 100) would have surprised both Darwin and Fisher, who always thought that it was the variance in fitness which caused selection, rather than the mean.

A telling example of Møller and Swaddle's blindness to alternative hypotheses is the discussion of FA in the cheetah (pp. 118–119). The cheetah, which has become famous for its lack of genetic diversity, has a high level of FA for a cat species. This is consistent with the idea that inbreeding is a form of "stress" which, like environmental stress, often increases FA. However, a study of the two cat species most closely related to cheetahs has shown that they have levels of FA very similar to that of the cheetah, despite the fact that they are more variable genetically. Møller and Swaddle review this evidence, but then conclude that the cheetah must have been purged of the deleterious alleles which would have plagued the species when it first became inbred, restoring it to a normal level of developmental stability. The alternative hypothesis that genetic diversity and developmental stability are not related in cats is not mentioned. Dr. Pangloss would be pleased.

Even with all these major conceptual failings, this book does offer a useful series of reviews of the empirical literature. Many of the citations for these reviews are contained in a set of pages maintained at Oxford University Press' web site http://www.oup.co.uk/MS-asymmetry. The literature covered is extraordinarily wide ranging and is likely to include some material unfamiliar even to afficionados of these areas. On the other hand, some of these reviews do not adequately cover the published literature, such as the review of

the effect of directional selection on phenotypic variance mentioned above.

Although the breadth of these reviews is sometimes impressive, unfortunately the quality of the interpretations is shockingly biased. Studies are quoted as confirming a role for asymmetry that contain no relevant data, contradictory results within a paper are ignored, and results are misinterpreted. Random examination of just a handful of the cited studies is likely to reveal a case of such bias. For example, I examined just four papers cited in the web table concerning the existence of a population asymmetry parameter, before finding an example. Kat (1982) studied two species, and both are listed by Møller and Swaddle as confirming the existence of an asymmetry parameter. However, for one species, Kat only measured the asymmetry of a single trait, precluding any test for an asymmetry parameter. In the other species, two characters were measured, but their asymmetries were not significantly correlated.

Such seemingly minor lapses would ordinarily not be of concern when the good faith of the reviewer can be assumed. However, Møller has a track record of bias in a series of recently published reviews (Møller 1997; Møller and Thornhill 1997, 1998) which also form the basis for substantial material in this book. Each of these reviews has been subjected to devastating critiques (J. Evol. Biol. 10(1), 1997; Clarke 1998; Palmer, in press). What the critiques make clear is that each of these reviews is rife with the sort of errors made in citing the Kat (1982) study, and that virtually every error tends to confirm the importance of asymmetry.

Møller and Swaddle are unfazed by the criticisms they were aware of when their book went to press. They repeat the original conclusions of Møller and Thornhill's (1997) meta-analysis of the heritability of asymmetry down to the wildly inflated estimate of average heritability (pp. 112–114). Although they do address some of the criticisms of others, these are, in effect, dismissed as technical points that do not affect the overall conclusions. For those inclined to accept this version of events, I particularly recommend Whitlock and Fowler's (1997) critique, which points out that at least five of the cited studies provide no estimate of the correlations of FA among relatives at all, the majority of the remaining studies do not distinguish additive variance from other causes of correlations, and that the "effect size" measure chosen is not a measure of effect size at all. I suspect that the chief value of the literature reviews in Møller and Swaddle will be for students to learn the many misuses for statistics, much as Huff (1954) was used to introduce good statistical practice in the past.

The web page also contains a revealing clue to how this book was written. A note on the opening page states that "due to an error by Anders Pape Møller a section of the book . . . on pg. 54, second paragraph . . . was not attributed to its proper source, Dr. G. M. Clarke . . . ." The note goes on to tender an apology from Møller, and to state that "Dr. Clarke is cited as the sole source of this idea." I contacted Dr. Clarke, who made clear that this apology, however sincere, does not convey the whole story. The material in question is from an unpublished manuscript, which Clarke reports that Møller reviewed. Clarke furnished me with a copy of his manuscript, and it shows that the passage which appeared in Møller and Swaddle, more than 200 words in all, is virtually identical, word for word, to the corresponding passage in Clarke's manuscript.[2] A few phrases in this passage of Clarke's manuscript did not make an appearance in Møller and Swaddle, including one where Clarke cites a paper of his own. While Møller apologizes for the fairly innocuous sin of failing to cite the proper source of an idea, the real story is far more troubling. Tellingly, it turns out that the quote I included above, that "it is inconceivable" that a single mechanism could buffer perturbations to many traits, is one of those phrases also found in Clarke's manuscript. With this knowledge, the peculiarly contradictory nature of this book on crucial points becomes much easier to understand. One does not have to look far in the book to find analogous instances: the initial definition of developmental stability and canalization on pages 3–4 bears a striking resemblance to that of Zakharov (e.g., 1992), who is not cited.

So is there anything to this latest enthusiasm for asymmetry? It is clear that we should not turn to Møller and Swaddle to help us find the answer. Shoddy thinking, an unwillingness to confront the consequences of evidence, and biased interpretations are enough to tarnish even the shiniest new idea, if we let them. Others are more circumspect concerning asymmetry, more trustworthy, and perhaps almost as enthusiastic (e.g., Palmer 1996). The problem in determining the importance of asymmetry is that the effects of many of the phenomena are, if present, so small that observational and experimental results may easily be influenced one way or the other by seemingly trivial details of experiment or observation. Detailed studies of the genetic or developmental basis for asymmetry and any selection on it are more promising as they hold out the possibility of more repeatable and interpretable results, but they are also more technically demanding and time consuming.

Negative reviews often give a frisson of pleasure to the reader. I suspect that part of the thrill is in the shared assurance that the reader and the writer would never have fallen into the sorts of errors criticized. Before we feel too self righteous over any shortcomings of the asymmetry enthusiasm, let us remember our own heritage of enthusiasms for heterosis, for the broken stick model, for the macroevolutionary consequences of the breeder's equation, or for speciation by genetic revolutions. While I feel confident that the seamier manipulations of the asymmetry enthusiasm are indeed beyond most of us, our history of succumbing to enthusiasms is a deep-seated feature of our scientific culture. Few of us entirely resist; it would behoove us to remember this at all times. We have little choice but to seek inspiration from gurus of the newest ideas; sometimes they turn out to be partially right. However, we should never believe them without a struggle. If an idea seems too good to be true, it is probably not true. At least it has never been true in my experience.

Literature Cited

Clarke, G. M. 1998. Developmental stability and fitness: the evidence is not quite so clear. Am. Nat. 152:764–768.

[2] For those with Møller and Swaddle in front of them, the passage is on page 54 and runs from the phrase "Developmental processes" to the second citation of Mather.

GRAHAM, J. H., D. C. FREEMAN, AND J. M. EMLEN. 1993. Antisymmetry, directional asymmetry, and chaotic morphogenesis. Genetica 89:121–137.

HOULE, D. 1997. Comment on "a meta-analysis of the heritability of developmental stability" by Møller and Thornhill. J. Evol. Biol. 10:17–20.

———. 1998. How should we explain variance in the genetic variance of traits? Genetica 102/103:241–253.

HUFF, D. 1954. How to lie with statistics. W. W. Norton, New York.

KAT, P. W. 1982. The relationship between heterozygosity for enzyme loci and developmental homeostasis in peripheral populations of aquatic bivalves (Unionidae). Am. Nat. 119:824–832.

LANDE, R. 1980. Genetic variation and phenotypic evolution during allopatric speciation. Am. Nat. 116:463–479.

MØLLER, A. P. 1997. Developmental stability and fitness: a review. Am. Nat. 149:916–932.

MØLLER, A. P., AND R. THORNHILL. 1997. A meta-analysis of the heritability of developmental stability. J. Evol. Biol. 10:1–16.

———. 1998. Bilateral symmetry and sexual selection: a meta-analysis. Am. Nat. 151:174–192.

PALMER, A. R. 1996. Waltzing with asymmetry. BioScience 46: 518–532.

———. *In press.* Detecting publication bias in meta-analysis: a case study of fluctuating asymmetry and sexual selection. Am. Nat.

ROWE, L., AND D. HOULE. 1996. The lek paradox and the capture of genetic variance by condition dependent traits. Proc. Roy. Soc. London, Ser. B 263:1415–1421.

WAGNER, G., G. BOOTH, AND H. BAGHERI-CHAICHIAN. 1997. A population genetic theory of canalization. Evolution 51:329–347.

WHITLOCK, M. 1996. The heritability of fluctuating asymmetry and the genetic control of developmental stability. Proc. Roy. Soc. London, Ser. B. 263:849–854.

WHITLOCK, M., AND K. FOWLER. 1997. The instability of studies of instability. J. Evol. Biol. 10:63–67.

ZAKHAROV, V. M. 1992. Population phenogenetics: analysis of developmental stability in natural populations. Acta Zoologica Fennica 191:7–30.

Book Review Editor: J. Coyne