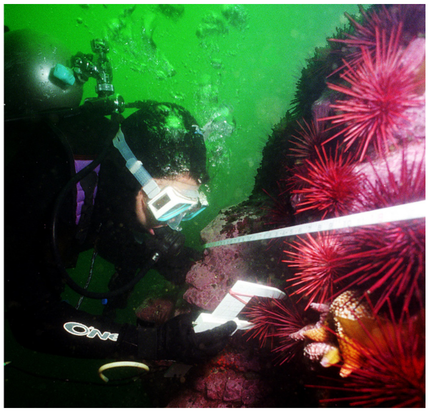

CONTRIBUTIONS

Commentary

Interpreting the Results from Multiple Regression and Structural Equation Models

The coefficients that are associated with pathways in multiple regression, as well as more advanced methods based on regression, such as structural equation models, are central to the interpretations made by researchers. The complex of factors that influence these coefficients make interpretations tricky and nonintuitive at times. Very often, inappropriate inferences are made for a variety of reasons. In this paper we discuss several important issues that relate to the interpretation of regression and path coefficients. We begin with a consideration of multiple regression. Here we discuss the different types of coefficients that can be obtained and their interpretations, with our focus on the contrast between unstandardized and standardized coefficients. Structural equation modeling is used to show how models that better match the theoretical relations among variables can enhance interpretability and lead to quite different conclusions. Here we again emphasize often-ignored aspects of the use of standardized coefficients. An alternative means of standardization based on the "relevant ranges" of variables is discussed as a means of standardization that can enhance interpretability.

Biologists have long used multiple regression in its various forms to examine relationships among explanatory and response variables. Over the past decade and a half, there has been a steady increase in the use of path analysis by biologists to serve the same purpose, but in the context of a more interpretive structure. Most recently, there has developed a considerable amount of interest in the more comprehensive capabilities of structural equation modeling (SEM) for understanding natural systems, again with the purpose of enhancing our interpretation of results. These methodologies have in common that they are based on the fundamental principles of regression and share many of the same issues when it comes to interpretation.

Researchers may not be aware that there has been a long history of discussion among quantitative social scientists and statisticians about the interpretation of results from both multiple regression and path analysis applications. The topic is sufficiently subtle and important that the central theme of Pedhazur's (1997) book on regression is the pitfalls of interpreting results. Among the many things he concludes is that results are frequently misinterpreted, particularly as they relate to the meaning of path coefficients. Many of these same issues apply to SEM. This discussion has involved a consideration of many topics, including the types of coefficients that can be calculated, the kinds of interpretations that can be supported using different coefficient types, and the importance of theory to interpretation. Here we illustrate some of these issues and discuss problems with the use of standardized coefficients, as well as a possible remedy.

An illustrative example

To illustrate the points being made in this paper we consider an example dealing with the response of shrublands to wildfire in Southern California (J. B. Grace and J. E. Keeley, *unpublished manuscript*). The data presented here represent a small subset of the variables in the complete study. In addition, the relationships among variables have been modified somewhat to meet the needs of the current paper. In this example, 90 sites were located in areas burned by a series of fires that occurred during a 2-week period in the fall of 1993 (Keeley et al., *in press*). Plots were established in all 90 sites and sampling began in spring of the first postfire year and continued for 4 more years, though only the data from the first sampling following fire are discussed here. At each site, the variables included (1) herbaceous cover (as a percentage of ground surface), (2) fire severity (based on skeletal remains of shrubs, specifically the average diameter of the smallest twigs remaining), (3) prefire stand age (in years), estimated from ring counts of stem samples, and (4) the elevation above sea level of the site. The data used in this analysis are summarized in Table 1. Again, the data presented are a subset of the original, and some relations in the data have been modified to make the example more applicable to our purposes.

Issues related to multiple regression

A multiple regression represents a particular model of relationships in which all potential explanatory variables (predictors) are treated as coequal and their interrelations are unanalyzed. As we shall see, the ability to obtain interpretable results from such models depends on the degree to which their structure matches the true relations among variables. Fig. 1 presents diagrammatic representations of a multiple regression model in which fire severity, stand age, and elevation are related to vegetation cover. Parameter estimates were obtained using the software Mplus (Muthén and Muthén 2005) under maximum likelihood estimation. Several types of coefficients were obtained from the analyses and are presented in Fig. 1, with each subfigure presenting a different view of the relations among variables.

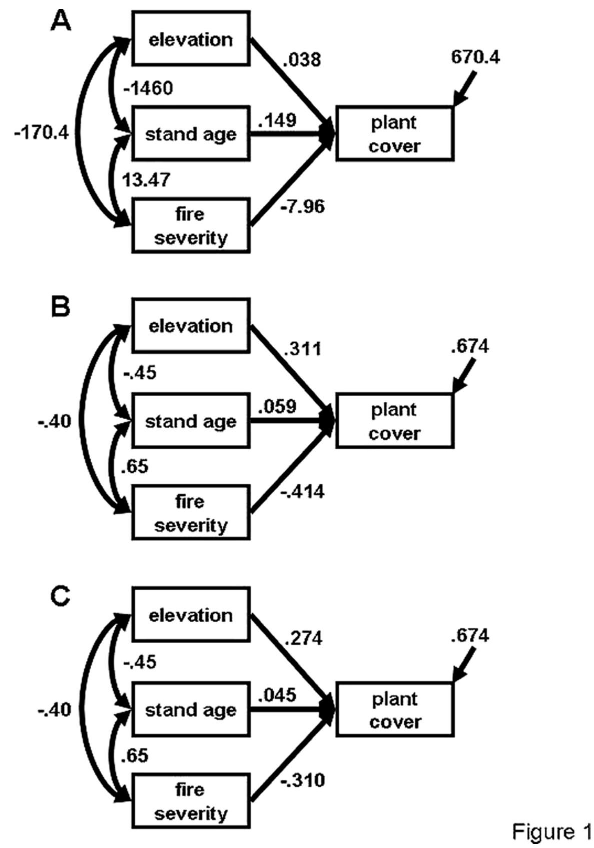

Fig. 1. Multiple regression results based on analysis of the data in Table 1. (A) Unstandardized parameters. (B) Standardized parameters. (C) Semipartial coefficients for the directional pathways.

Unstandardized coefficients

Fig. 1A presents the unstandardized path coefficients associated with the regression of plant cover on elevation, stand age, and fire severity. While the unstandardized coefficients are the most primary parameters obtained from a multiple regression, often they are not presented by investigators. In fact, typically the significance tests associated with regression are tests of the unstandardized parameters, and the standardized parameters are simply derived from the unstandardized coefficients and not directly tested. Characteristic of unstandardized parameters, they are expressed in the original units of the explanatory and dependent variables. With reference to a simple linear regression, unstandardized coefficients associated with directed paths represent the slope of the relationship. The same is true in multiple regression, although the slope is in *n*-dimensional space.

As we begin to interpret the results in Fig. 1A, note that the undirected relationships (double-headed arrows) represent the covariances among exogenous variables (predictors) in a model. In contrast, the coefficients associated with directed paths are partial regression coefficients. It is important for the discussion that follows to understand when the principles of partial regression apply. Simply put, partial regression represents a method of statistical control that removes the effect of correlated influences. Pathways that involve partial regression can be recognized by the following: (1) they involve a directed relationship (single-headed arrow), (2) the response variable (variable receiving the arrow) also receives other arrows, and (3) the multiple predictors affecting the response variable are correlated. As we can see from these criteria, all directed paths in multiple regression will involve partial regression as long as there are significant correlations among predictors. The question then is how are we to interpret such coefficients.

The literal definition of a partial regression coefficient is *the expected change in the dependent variable associated with a unit change in a given predictor while controlling for the correlated effects of other predictors*. There are actually several different ways we can look at partial regression coefficients. The most direct is to view them as parameters of an equation such as

$$\text{cover} = 0.038(\text{elevation}) + 0.149(\text{age}) - 7.96(\text{severity}) \qquad (1)$$

when variables are in their raw units. If we were able to plot a four-dimensional graph of cover against elevation, age, and severity, the unstandardized regression coefficients would be the slopes of the relationship in the plot. From this perspective, it should be clear that the coefficients estimate the mean influences of predictors on the response variable and the variation around the mean is ignored. Deviations from the mean in this case relate to the estimation of the probabilities that coefficients' values are zero. Thus, one interpretation of the unstandardized coefficients is that they are prediction coefficients. They also are descriptive coefficients in that they describe the association between cover and a one-unit change in the other variables. Hypothetically, these coefficients might also be viewed as explanatory. However, for such an interpretation to be valid, we must depend on the structure of the model to match the true dependencies among the predictors. As Pedhazur (1997:8) states, "Explanation implies, first and foremost, a theoretical formulation about the nature of the relationships among the variables under study." This point will be illustrated later in the paper when we discuss the structural equation model results for these data.

Referring back to our example, if we were to keep elevation constant for a set of plots, and the stands being burned were of a fixed age, a one-unit difference in the fire severity is associated with an average difference in cover of –7.96 cover units (i.e., the cover of the postfire community would differ by 7.96%). Similarly, if we were able to apply a fire of fixed severity while also holding stand age constant, a difference in elevation of 1000 m is associated with an expected difference of 38% in the postfire cover.

Standardized coefficients

Looking at Fig. 1A, we see that it is difficult to compare unstandardized coefficients among different pathways because the raw units are various. Cover varies in percentage points, elevation varies in meters, age varies in years, and fire severity varies in the units of an index based on the diameter of remaining twigs following fire. So, is a value of 0.038 (the coefficient for elevation effects on cover) large or small relative to the effect of another factor? The standardization of the coefficients based on the standard deviations of the variables is the approach typically used to make coefficients comparable. In essence, this puts variables in standard deviation units, and in that sense the expected impact of a standard deviation difference in one variable (say elevation) can be compared to a standard deviation difference in another variable (say fire severity). Though a convenient transformation, standardized regression coefficients are frequently misinterpreted, for reasons we will discuss next.

The most common misinterpretation of standardized coefficients is to interpret them as if they represent a partitioning of explained variance in the model. The fact that standardized coefficients are in standard deviation units contributes to the tendency to make this mistake. For example, the formula for standardized partial regression coefficients can be expressed in terms of the correlations among variables. In the case of two predictors, x_1 and x_2, and one response, y_1, this formula is

$$\gamma_{11} = \frac{r_{x1y1} - r_{x1x2} \times r_{x2y1}}{1 - r^2_{x1x2}} \tag{2}$$

where γ_{11} refers to the standardized partial regression coefficient representing the response of y_1 to x_1, and the r values represent the bivariate correlations among variables. This formula can be readily extrapolated to the case of more than two predictor variables (Pedhazur 1997).

Another relationship that applies to standardized coefficients is that the sum of all simple and compound associations between two variables equals the bivariate correlation between those two variables. For example, the bivariate correlation between elevation and cover is 0.45 (Table 1). With reference to Fig. 1B where standardized coefficients are presented, we find that the coefficients are those that satisfy the formula (allowing cover to be y_1, and elevation, stand age, and severity being $x_1 - x_3$)

$$r_{x1y1} = \gamma_{11} + r_{x1x2} \cdot \gamma_{12} + r_{x1x3} \cdot \gamma_{13}, \tag{3}$$

where γ_{11} is the response of y_1 to x_1, γ_{12} is the response of y_1 to x_2, γ_{13} is the response of y_1 to x_3, and r's refer to correlations.

A third property of standardized coefficients is that they can be related to the explained variance in our response variable using the equation

$$R^2 = r_{x1y1} \cdot \gamma_{11} + r_{x2y1} \cdot \gamma_{12} + r_{x3y1} \cdot \gamma_{12} \tag{4}$$

Pedhazur (1997). For our example presented in Fig. 1B, we find that the expression in Eq. 4 yields an R^2 of 0.326 (note the standardized error variance shown in Fig. 1B equals 1 minus the R^2).

Now, the properties of standardized coefficients give the impression that they solve a number of problems. Most obviously, they put all the coefficients in what seem to be the same units. However, they are only the "same" if we are willing to say that a standard deviation for one variable in one metric is *interpretationally equivalent* to a standard deviation of another variable that was measured in a different metric. This is an implicit assumption of using standardized coefficients and it is not obvious that this assumption is suitable other than in the fact that each is a standard deviation.

More seductive than that, however, is that standardized coefficients are expressed in terms of correlations, *which represent the variation associated with the relationships.* In the case of simple regression (involving one predictor variable), we know that the unstandardized coefficient represents the slope, while the standardized coefficient represents the square root of the variance explained in the response variable. Eq. 4 may give the false impression that this relationship between standardized coefficients and variance explained can be generalized to the case of multiple correlated predictors. However, it cannot be so generalized. To see why more readily, we now turn to the concept of semipartial coefficients and unique variance explanation.

Semipartial coefficients and the concept of shared variance explanation

The semipartial coefficient, when expressed in standardized form, represents a measure of the *unique ability of a predictor variable to explain variation in a response variable that cannot be explained by any other predictor variable in the model.* We can understand this in contrast to stepwise regression, which measures the sequential abilities of variables to explain residual variance. In sequential variance explanation, there is a pervading influence on the results by the logic used to determine the order of variables included. Here, the semipartials represent a measure of the minimum effect of a variable regardless of logical order. In the example in Fig. 1C, the coefficients associated with directed paths are semipartial coefficients, while the coefficients associated with undirected paths remain correlations. The unique variance explanation abilities of our three predictors (elevation, age, and severity) are 0.075, 0.002, and 0.096, the squares of the semipartial coefficients. Collectively, the three variables provide unique variance explanation of 0.173. Since the total variance explained by the full model is 0.326, we must conclude that 0.153 (roughly half) of the explained variance is shared among predictors.

The concept of *shared variance explanation* makes sense when we have predictor variables that are correlated for some unknown or unspecified reason. How are we to apportion the correlated explanatory power among predictors in a multiple regression? Since our relations among predictors are unanalyzed or not understood, we have no means to accomplish this. The implications of these relations can be seen if we compare the coefficients in Figs. 1B and C. It is to be expected that the partial regression coefficients are greater than the semipartial coefficients, with the degree of difference directly related to the strength of the correlations among predictors. It should be clear from the above discussion that as predictors become more highly correlated, their unique variance explanation ability decreases. It should also be clear from our presentation that the standardized partial regression coefficients (Fig. 1B) do NOT represent measures of variance explanation ability. Rather, *the standardized partial regression coefficients represent expected changes in* y *as a result of manipulations in* x *in standard deviation units while controlling for the correlated effects of other predictors.* The reason these coefficients cannot be used to represent variance explanation is simple; it is because we cannot guess how to apportion the variance explanation shared among predictors. In sum, the total variance explained in a multiple regression can only be attributed to the collection of predictors. The truth of this is most evident in nonlinear regression where individual predictors (e.g., x and x^2) may explain no variance by themselves, yet together they can explain substantial variance in some y.

Conclusions about the interpretability of multiple regression

While investigators commonly ask, "What is the relative importance of a set of causes controlling some observed phenomenon?" we must conclude that when predictor variables are correlated for unknown reasons, standardized partial regression coefficients do not provide an answer to this question. It is true that when correlations are not excessive, path coefficients can provide important insights. Multiple regression, which is inherently designed to ignore the causes behind the correlations among a set of predictors, makes for a particularly poor approach to understanding, however. This fundamental problem has been long recognized and is the central theme in Pedhazur's (1997) book on multiple regression. While Pedhazur discusses the problem from many different angles, his main conclusion is that without a theory to guide the analysis, a meaningful answer to the question of relative importance of factors is usually precluded in a multiple regression analysis. As we have seen, standardized regression coefficients do not equate to variance explanation. At the same time, measures of unique and shared variance explanation, which can be obtained using semipartial analysis, really don't address explanatory questions either, but instead, relate more to their unique roles as predictor variables.

Structural equation modeling

Since the interpretability of multiple regression results is typically limited by an insufficiently developed theoretical framework, we should consider what problems are solved using a theory-oriented method such as SEM. For those not familiar with SEM, it involves the use of a generalized multiequation framework that enables the analyst to represent a broad range of multivariate hypotheses about interdependencies (Bollen 1989). Path analysis, which is now familiar to most ecologists, is best known in analyses that only consider relations among observed variables. Modern SEM allows for the inclusion of unmeasured (latent) effects, as well as the specification of a wide range of model types. Importantly, SEM allows for evaluations of model fit that serve to permit overall testing of the model as a hypothesis. While SEM is most commonly based on maximum likelihood estimation, many model types can be solved using various least squares procedures. While we do not present latent variable examples in this paper, the issues discussed apply equally to such models.

We should begin by stating that SEM does not solve all problems associated with interpreting multivariate relations. Both inadequate data and insufficient theory can block substantial progress. Additionally, while SEM permits the implications of a causally structured theory to be expressed, the analysis itself does not contribute to the establishment of causality. This must come from other information. Nonetheless, the use of theory to guide our analysis within an SEM framework has the potential to remove many obstacles to interpretation. The example presented here is meant to illustrate that potential, but not to imply that the application of SEM automatically leads to a superior analysis.

Returning to the example of fire response by California shrublands, we now ask, "What do we know of the relations among our explanatory variables?" In this case, the authors of the original study felt they knew some important things, but we were unable to incorporate this information into the multiple regression performed in the previous section. First, substantial experience (Keeley 1991) indicates that postfire recovery by the plant community may be affected by fire severity because of impacts on seed survival. It is also possible that impacts to soil properties could contribute as well (Davis et al. 1989). The point is that fire severity is reasonably modeled as having a direct impact on plant cover. Stand age can be expected to have an effect on fire severity because older stands tend to have more fuel. A simple thought experiment illustrates the point. If we were to vary stand age (say, allow a stand to get older and accumulate more fuel), we might reasonably expect that it would burn hotter (though this would not be guaranteed). However, if we were to manipulate fire severity in a plot, that would certainly not affect the age of the stand. This logic and the experience upon which it is based encourages us to represent the relationship between stand age and fire severity as a directional one rather than a simple correlation. By a similar logic, we can see that the relationship between elevation and stand age should be represented as directional. If shrub stands tend to be younger as we go higher in elevation, which the data indicate, (e.g., if there were a reduced incidence of fire suppression at higher elevations), then picking a spot lower on the mountain will likely result in finding an older stand. On the other hand, if we were to allow a stand of shrubs to get older, we would not find that there was an associated change in elevation. Again, the use of thought experiments, which tap into our body of prior knowledge, suggest directional relationships among variables.

Some researchers may be uncomfortable with the logic used above to indicate directional relationships in causal models. This subject is beyond the scope of our discussion in this paper and we refer the reader to more in-depth treatments of the subject (e.g., Bollen 1989, Pearl 2000, Shipley 2000). For now, we accept such a procedure as reasonable and illustrate its consequences in Fig. 2. The path model represented in Fig. 2 illustrates the logic of the dependencies described above. In addition, it represents the possibility that there may be influences of elevation on cover that are unrelated to associated variations in stand age and fire severity. Because this model is not saturated (i.e., not all paths are specified), our model represents a testable hypothesis. Inherent in SEM practice is the evaluation of fit between model expectations and observed relations in the data. Our point here is not to elaborate on this point, but only to note this feature of SEM practice and then continue with our discussion of interpretation. The patterns of covariances specified in Table 1 in fact fit the model presented in Fig. 2 reasonably

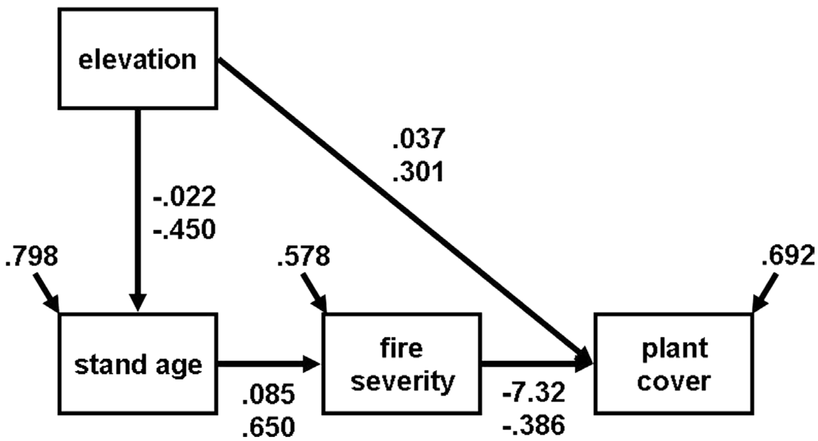

Fig. 2. Path model results based on analysis of data in Table 1. For path coefficients, upper numbers are the unstandardized coefficients, while lower numbers are standardized path coefficients. Error variables are standardized values. Model chi-square = 2.535 with 2 df, $P = 0.278$.

well (chi-square = 2.535 with 2 df and $P = 0.278$; note that a nonsignificant P value indicates the absence of significance deviations between data and model). This does not, of course, prove that the model is the correct one, only that it is consistent with the data.

The first thing we should do when interpreting the results in Fig. 2 is to consider which of our paths involve partial regression and which involve simple regression. Recall that response variables receiving two or more directed arrows will involve partial regression if the predictors involved are correlated. As stand age and fire severity only receive single directed arrows, their incoming pathways represent simple regression relations. We can see in fact that the correlations in Table 1 match the standardized path coefficients in Fig. 2 for these two pathways. Cover, on the other hand, has multiple influences and thus, the coefficients from elevation to cover and fire severity to cover are partial coefficients. What this means is that when we examine the relationship between elevation and stand age or between age and severity, there are no influences from other variables in the model to control for. On the other hand, the relationship between severity and cover is controlled for the covarying effects of elevation on cover. Similarly, the direct path from elevation to cover represents the effect once the influence of severity is removed.

Considering the unstandardized path coefficients in Fig. 2, we can see that the covariance between elevation and stand age can be understood as an expectation that age will decline on average by 2.2 years with an increase of 100 m. The covariance observed between stand age and fire severity can be understood as an expectation that severity will increase by 0.085 units with each year older a stand gets. Thus, we can understand the covariance between elevation and fire severity as the product of these two described relationships. Further, there is no indication of any other effect of elevation on fire severity except that mediated by stand age (because there is no direct path from elevation to severity to indicate some other effect).

The interpretation of unstandardized coefficients connecting severity and elevation to cover is somewhat different from those associated with a simple regression coefficient. We would draw the interpretation from Fig. 2 that increasing fire severity by one unit while holding all other conditions constant would cause a decrease in cover of 7.32%. The effect of elevation on cover is somewhat more interesting because of the presence of both direct and indirect effects on cover implied by the model. The direct path from elevation to cover predicts that if one were to choose a site 100 m higher than the mean and yet have an average severity fire, postfire cover would be 3.7% higher

Table 1. Covariances and correlations† among four variables relating vegetation regrowth in response to wildfire and the standard deviations of each variable (n = 90).‡ Matrix diagonals are the variances for the four variables.

Variables	Vegetation cover (% cover)	Fire severity (index values)	Prefire stand age (yr)	Elevation (m)
Cover	1,006.2	–26.2	–139.4	3686.3
Severity	–0.50	2.722	13.47	–170.4
Age	–0.35	0.65	157.8	–1459.6
Elevation	0.45	–0.40	–0.45	66,693
Standard deviations	31.72	1.65	12.56	258.25

†Note that the variance/covariance matrix can be reconstituted from the correlations and standard deviations presented. All analyses presented are based on the analysis of covariances.

‡The correlations among variables have been modified from the original to make the example more useful for the purposes of this paper. However, the standard deviations are as found by Keeley and Grace (submitted), thus the original scales for variables are preserved.

than the mean. On the other hand, the total effect of elevation on cover is 0.050, which indicates that if one moved upslope 100 m and allowed stand age and severity to vary as it naturally would (i.e., we are not holding them constant), there would be a net increase in cover of 5.0%. For the total effect of varying elevation, part of the increase in cover (1.3%) would result from the fact that stands would be younger (on average), 100 m higher, and associated fires would be expected to be less severe.

Consideration of standardized coefficients (Fig. 2) provides for an understanding of relationships expressed in terms of standard deviations. Such coefficients are both more easily compared (assuming different standard deviations can be thought of as equivalent) and somewhat more abstract. In these units, we see that if severity were increased by one standard deviation while elevation was held constant, cover would be expected to decrease by 0.386 standard deviations. On the other hand, if elevation was increased by one standard deviation, while holding severity constant, cover would increase by 0.301 standard deviations. Based on an estimated total effect of elevation on cover of 0.414, we can see that if elevation was increased one standard deviation without holding age and severity constant, then cover would increase 0.414 standard deviations. Thus, in terms of standardized units, the direct effect of elevation on cover is less (sign ignored) than the effect of severity (0.301 vs. 0.386), though the total effect of elevation on cover is greater (0.414).

So, how does all this relate to the question of the relative importance of different factors in affecting cover? If we accept standardization in terms of standard deviations as a reasonable basis for comparing coefficients (which is questioned below), it can be seen that the total influence of elevation on cover is greater than that of fire severity, with the total effect of stand age (–0.251) being least. The question we must now address is what it means to say that a pathway represents expected change in terms of standard deviation units.

Criticisms of standardization

While the above discussion appears to provide a suitable resolution of the question of how we may evaluate the importance of explanatory variables, we were forced to accept the caveat that standardization based on standard deviations was reasonable. Many metricians actually recommend that researchers avoid using standardized coefficients and focus on the unstandardized coefficients when seeking to draw conclusions from regression models (Darlington 1990, Luskin 1991). The reason for this is tied to the substantive meaning of unstandardized coefficients and the conditional nature of standardized coefficients. If we presume that our sample is fairly representative of some larger world, our unstandardized estimates represent the slopes of the relationships (i.e., the mean responses). When we use standardized coefficients, we interject additional variables into the problem, that of the sample variances. As Pedhazur (1997:319) so eloquently put it, "The size of a [standardized coefficient] reflects not only the presumed effect of the variable with which it is associated but also the variances and the covariances of the variables in the model (including the dependent variable), as well as the variances of the variables not in the model and subsumed under the error term. In contrast, [the unstandardized coefficient] remains fairly stable despite differences in the variances and the covariances of the variables in different settings or populations."

These criticisms of standardization appear rather powerful. In many ecological studies, we know that our samples often represent a tiny fraction of the total samples possible. Also, restrictions on randomization, for example because of accessibility problems or other sampling limitations, mean that sampling is often neither purely random nor fully representative; thus, variances can easily vary from sample to sample. Additionally, comparisons among populations based on standardized coefficients depend on the variances being constant across populations, which may frequently not be the case. Unstandardized coefficients are generally much more readily estimated with accuracy and less sensitive to differences in the variances of the variables across samples. Comparisons across populations (or between paths) in unstandardized coefficients do not depend on equal sample variances, and as a result, are more generalizable parameters than are those based on standardization. Altogether, there are assumptions that go into the interpretation of standardized coefficients and these are typically ignored, representing unknown influences.

A possible resolution using an alternative standardization procedure

Despite the criticisms of standardization, researchers generally would prefer a means of expressing coefficients in a way that would permit direct comparisons across paths. The debate over this issue goes back to Wright (1921), who originally developed path analysis using standardized variables. It was Tukey (1954) and Turner and Stevens (1959) who first criticized the interpretability of standardized values in regression and path models, and many others have since joined in that criticism. However, Wright (1960) argued in defense of standardized coefficients, saying that they provide an alternative method of interpretation that can yield a deeper understanding of the phenomena studied. Later, Hargens (1976) argued that when the theoretical basis for evaluating variables is based on their relative degrees of variation, standardized coefficients are appropriate bases for inference. Therefore, there are circumstances where standardized coefficients would be desirable. As Pedhazur's recent assessment of this problem concludes, ". . . the ultimate solution lies in the development of measures that have meaningful units so that the unstandardized coefficients . . . can be meaningfully interpreted."

So, how might we standardize using measures that have meaningful units? We must start by considering what it means to say that if x is varied by one standard deviation, y will respond by some fraction of a standard deviation? For normally distributed variables, there is a proportionality between the standard deviation and the range such that six standard deviations are expected to include 99% of the range of values.

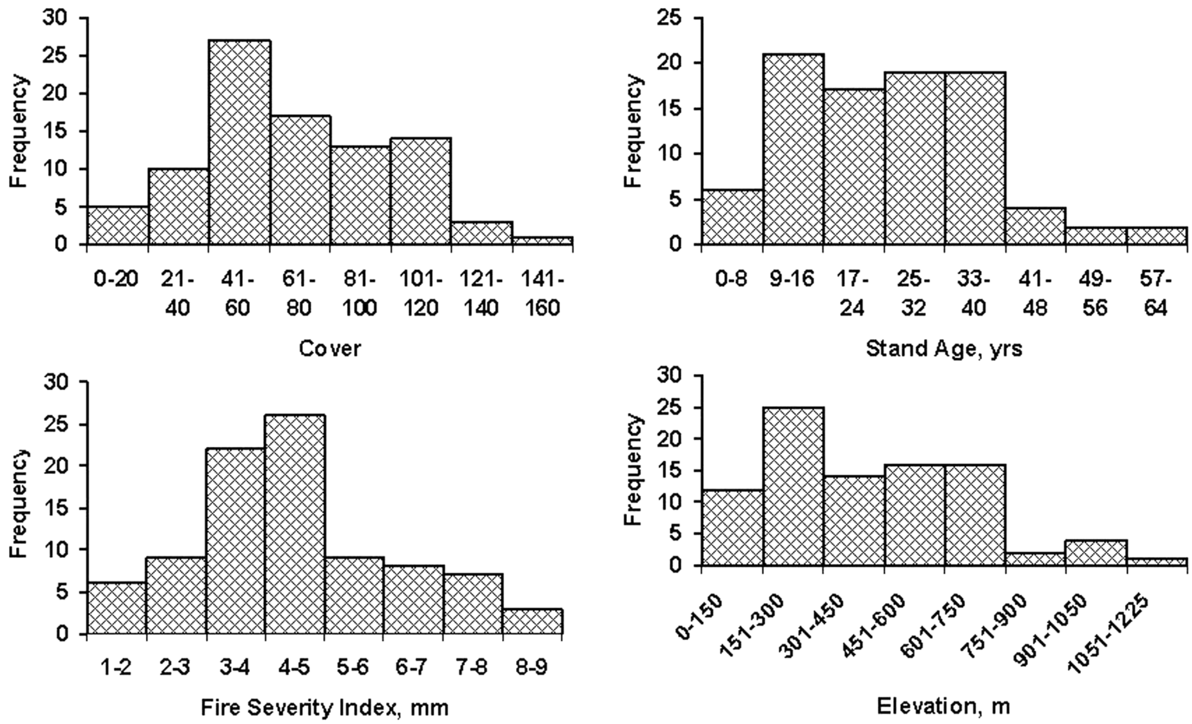

Fig. 3. Frequency diagrams for cover, stand age, fire severity, and elevation.

As discussed earlier, this may seem reasonable if (1) we have a large enough sample to estimate a consistent sample variance, (2) our variables are normally distributed, and (3) variances are equal across any samples we wish to compare. The reason why many metricians oppose standardized coefficients is because these three necessary conditions are not likely to hold generally. Of equal importance, rarely are these requirements explicitly considered in research publications and so we usually don't know how large violations of these requirements might be.

Fig. 3 presents frequency distributions for the four variables considered in our example. In the absence of further sampling, the repeatability of our sample variance estimate is unknown. This contributes to some uncertainty about the interpretability of coefficients standardized by the standard deviations. As for approximating a normal distribution, three of the four variables are truncated on the lower end of values. Cover can never be <0%, elevation likewise has a lower limit of expression relevant to terrestrial communities in this landscape, and stand age is also limited to a minimum value of between 0 and 1 year. None of these deviations are substantial enough to cause major problems with hypothesis tests (i.e., these variables are not wildly nonnormal); however, the deviations from idealized normality may very well impact the relationships between standard deviations and ranges. The observed range for cover was from 5% to 153% (overlapping canopies allow cover to exceed 100%), while six times the standard deviation yields an estimated range of 190%. The observed range for elevation was from 60 to 1225 m, while six times the standard deviation equals 1550 m. Stand age ranged from 3 to 60 years old, with six times the standard deviation equaling 75 years. Finally, fire severity index values ranged from 1.2 to 8.2 mm, while six times the standard deviation equals 9.9 mm. Thus, observed ranges are consistently less than would be estimated based

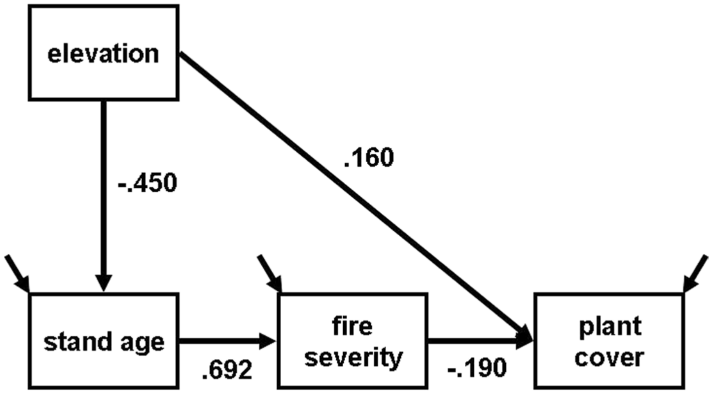

Fig. 4. Path analysis result showing coefficients standardized by the relevant ranges.

on standard deviations and the degree to which this is the case is slightly inconsistent (ratios of observed to predicted ranges for cover, elevation, age, and severity equal 0.78, 0.75, 0.76, and 0.71).

It is possible that in some cases information about the ranges of values likely to be encountered or of conceptual interest can provide a more meaningful basis for standardizing coefficients than can the sample standard deviations. We refer to such a range as the "relevant range." For example, if we have a variable whose values are constrained to fall between 0 and 100, it would not seem reasonable for the relevant range chosen by the researcher to exceed this value regardless of what six times the standard deviation equals. On the other hand, it may be that the researcher has no basis other than the observed data for selecting a relevant range. Even in such a case, we can choose to standardize samples that we wish to compare by some common range so as to clarify meaning across those samples. Whatever the basis for standardization, researchers should report both the unstandardized coefficients and the metrics used for standardization.

For the variables in our example, we specify the relevant range for cover to be from 0% to 270%. Obviously values cannot fall below 0%, but why chose an upper limit of 270%? Examination of cover values for all plots across the five years of the study show that values this high were observed in years 2 and 4 of the study. By using a relevant range of from 0% to 270%, we permit comparisons across years standardized on a common basis. Of course, this implies that the slopes measured will extrapolate to that full range, which is an assumption that should be evaluated closely. For elevation, the relevant range we choose is the observed range, from 60 to 1225 m. This span of 1165 m is chosen because we do not wish to extrapolate to lower or higher elevations, in case relationships to other variables are not robust at those elevations. For stand age, we specify the relevant range to be 60 years for basically the same reason. Finally, the fire index range chosen was also the observed range, which was 7.0 mm. It is clear that values could be obtained beyond this range in another fire. It is not known, however, whether the relationship between remaining twig diameter and herbaceous cover would remain linear outside the observed range.

Based on these determinations, we can generate path coefficients standardized on the relevant ranges. These coefficients are shown in Fig. 4. The biggest numeric differences between these values and those standardized using standard deviations (Fig. 2) is that the absolute values of the coefficients leading to cover are lower because of the large relevant range for this variable. The coefficient for the effect of age on severity is slightly higher, while that for the effect of elevation on age is unchanged. Using these coefficients now allows us to describe the importance of variables using their relevant ranges as the explicit context. These interpretations are *only* valid for relative comparisons within the *n*-dimensional parameter space defined by the relevant ranges. As fire severity increases across its relevant range, cover would be expected to decline by 19% of its relevant range. As elevation increases across its relevant range, the total change in cover from both direct and indirect causes would be an increase of 21.9% (the total effect). We now conclude from this analysis that the sensitivities of cover to fire severity and elevation (19% vs. 21.9%) are roughly equivalent in this study, though of opposing sign. It is possible to test whether these two estimates are reliable differences, which in this case, they are not.

Conclusions

It is important to recognize that the analysis of data has both an analytical element and a research element. By analytical element, we refer to the purely mathematical and statistical properties of the analytical methods. By research element, we refer to the fine art of applying analysis methods in the most meaningful ways. Formal training in statistics often emphasizes the analytical element and provides limited prescriptions for research applications that do not include a great deal of subjective judgment. What experienced statisticians have long known, however, is that for the application of statistical methods to be successful, strong guidance from the research perspective is required. Structural equation modeling is powerful specifically because it allows researchers to incorporate their accumulated knowledge into the analysis. Our advice regarding the interpretation of path coefficients is in that same vein. Rather than automatically allow sample standard deviations to represent the authoritative basis for standardizing coefficients, it is possible to insert our knowledge of the subject into the standardization process by explicitly considering the relevant ranges over which variables are to be considered. This procedure of standardizing based on substantive considerations acts to facilitate comparisons while avoiding problems associated with the sample-specific nature of standard deviations.

As with many new approaches, initial gains from defining and using the relevant range for standardization may be modest. Often the sample range will provide the best estimate available. However, as we accumulate additional information and focus on the ranges that are relevant to the inferences we wish to draw, much can be gained. Again, we recommend that unstandardized coefficients always be presented, regardless of the use of standardized coefficients of any sort. By also including either the sample standard deviations or the relevant ranges, which provide the bases for standardization, researchers can begin to compare both standardized and unstandardized values across studies. At present, there is a widespread and careless misapplication of standardized coefficients by researchers, both in the use of multiple regression and in the use of SEM/path analysis. Alternative means of comparing standardized coefficients may prove useful in drawing meaningful conclusions from analyses.

James B. Grace
US Geological Survey
and
Kenneth A. Bollen
University of North Carolina

Literature cited

Bollen, K.A. 1989. Structural equations with latent variables. John Wiley and Sons, New York, New York, USA.

Darlington, R. B. 1990. Regression and linear models. McGraw-Hill, New York, New York, USA..

Davis, F. W., M. I. Borchert, and D. C. Odion, 1989. Establishment of microscale vegetation pattern in maritime chaparral after fire. Vegetatio **84**:53–67.

Hargens, L. L. 1976. A note on standardized coefficients as structural parameters. Sociological Methods and Regression **5:**247–256.

Keeley, J. E. 1991. Seed germination and life history syndromes in the California chaparral. Botanical Review **57**:81–116.

Keeley, J. E., C. .J. Fotheringham, and M. B. Keeley. *In press.* Determinants of postfire recovery and succession in mediterranean-climate shrublands of

California. Ecological Applications.

Luskin, R. C. 1991. Abusus non tollit usum: standardized coefficients, correlations, and R^2s. American Journal of Political Science **35:**1032–1046.

Muthén, L. K., and B. O. Muthén. 2005. Mplus user's guide. Third edition. Muthén and Muthén, Los Angeles, California, USA.

Pearl, J. 2000. Causality: models, reasoning, and inference. Cambridge University Press, Cambridge, UK.

Pedhazur, E. J. 1997. Multiple regression in behavioral research. Third edition. Wadsworth, Belmont, California, USA.

Shipley, B. 2000. Cause and correlation in biology. Cambridge University Press, Cambridge, UK.

Tukey, J. W. 1954. Causation, regression and path analysis. Pages 35–66 *in* O. Kempthorne, T. A. Bancroft, J. W. Gowen, and J. L. Lush, editors. Statistics and mathematics in biology. Iowa State College Press, Ames, Iowa, USA.

Turner, M. E., and C. D. Stevens. 1959. The regression analysis of causal paths. Biometrics **15**:236–258.

Wright, S. 1968. Evolution and the genetics of populations. Volume 1. Genetic and biometric foundations. University of Chicago Press, Chicago, Illinois, USA.

Commentary

An Ecologist's Perspective of Ecohydrology

When my hydrological colleagues first brought up the term "ecohydrology" several years ago, I was simultaneously enthused, wary, and territorial. I still am. Enthused because the interface between ecology and hydrology still seems largely unmined, despite its key importance in ecosystems ecology—particularly in the water-limited systems that have been the focus of most of my work. Wary because although this interface does seem simultaneously unmined and important, the first response tends to be, "Well it's not news to ecologists that water is important in driving ecological processes and dynamics, and it is certainly not news to hydrologists that vegetation influences the water budget." And territorial because after feeling awash and striving to get my groundings in the ever-growing field of ecology, I was uneasy labeling any new collaborative endeavor—and particularly labeling myself—with a term ending in something other than "-ecology" or "-ecologist." A few years later, these points merit reflection and updating given the rapid growth in this area, which has affected me personally, as well as, I believe, a growing number of ecologists and hydrologists.

Most researchers have been cautious about labeling ecohydrology as a new field (Baird 1999, Bond 2002, Van Dijk 2004, Wilcox and Newman 2005). Rather, it is often referred to with respect to an increase in the interaction between ecology and hydrology. The terms "ecohydrology" and "hydroecology" have both been tossed around and have not been used consistently (Hannah et al. 2004). In general, "hydroecology" seems to be used more in association with aquatic ecology and riparian systems, whereas ecohydrology seems to be used more in association with terrestrial ecology, particularly for drylands. Most generally, there seems to be agreement that ecohydrology focuses on the interactions and interrelationships between hydrological processes and the pattern and dynamics of vegetation.

Debate remains about the relative newness and importance of ecohydrology (Hannah et al. 2004). Most colleagues I have spoken with who come from a hydrological background are particularly enthused about this growing area (see also Rodríguez-Iturbe 2000). Ecohydrology seems to have captured the interest of a subset of ecologists as well, although I have the sense there is not as much widespread enthusiasm as there appears to be in hydrology (see also Bond 2003). Many ecologists see it as just the next step in developing a new interface in ecology, similar to previous advances in plant ecophysiology or biogeochemistry. Some in natural resources believe that a wheel is being reinvented that is ignoring previous interdisciplinary contributions of watershed science and management. Although the latter perspective merits weight, I do believe that recent efforts in ecohydrology indeed represent a new level of interdisciplinary integration between current ecology and hydrology. Both fields have had substantial intellectual and membership growth over the past several decades since watershed resource management became established in an academic context. Some of the difference in perspective and level of enthusiasm for ecohydrology between the ecological and hydrological communities may reflect differences in their roots. Hydrologists have more direct roots in engineering relative to ecologists (see also Baird 1999), who like to view ourselves as being fundamentally rooted in multidisciplinary science. (It should be noted, however, that hydrologists often seem to be able to run circles around ecologists when it come to predicting relevant properties from the two respective disciplines.) Hence, ecologists and hydrologists may be viewing ecohydrology from different perspectives along the engineering–science continuum. I personally see the most evidence of the importance of greater integration between ecology and hydrology in a recent set of papers published in *Ecology* that resulted from an American Geophysical Union Chapman Conference on "Ecohydrology of Semi-arid Landscapes" (Wilcox and Newman 2005). This was one of the most exciting and stimulating workshops I have participated in, and the resulting papers (disclosure no. 1: I am a coauthor on two of the resulting papers) represent what I believe are novel syntheses that would be extremely unlikely to have been developed from either the ecological or hydrological communities alone.

Disclosure no. 2: I am still a bit uneasy with labeling myself as an "ecohydrologist," because it sounds like a specialty or "sub"-discipline in hydrology. But a close ecology colleague and friend pointed out, "Hey, Dave, you are an ecohydrologist—almost everything you study is related very tightly to the water budget, plant water use, and vegetation patterns and dynamics." I now use the term when it seems appropriate, but I also try to clarify that I am an ecologist, not a hydrologist, and I often refer to the area as "ecohydrology and vegetation dynamics." My active involvement at this interface was a major factor associated with my recent move to the University of Arizona, where I am working to build strong ties among related programs spread across three colleges: Hydrology and Water Resources in the College of Engineering; Ecology and Evolutionary Biology in the College of Science; and the Watershed Program within the School of Natural Resources in the College of Agriculture and Life Sciences. These academic units comprise a representative microcosm of much of the ecohydrology community at large. We now have a training grant in ecohydrology from the USDA and this fall I will teach a new course on "Dryland Ecohydrology and Vegetation Dynamics," so my near-future fate is somewhat coupled with upcoming development in ecohydrology. Developing this ecohydrology interface remains challenging, as does any interdisciplinary endeavor, but there currently is a great deal of interest and enthusiasm about it.

So what are these important, unmined areas in ecohydrology? Most generally, there is an important shift in emphasis between ecohydrology and the traditional focus of either ecology or hydrology. Ecohydrology, as noted above, focuses on the interactions and interrelationships between hydrological processes and vegetation pattern and dynamics. Traditionally, hydrology has focused in large part on issues of water yield and, as I perceive it (as a perhaps somewhat ignorant outsider), has invested much less effort in processes that are of particular interest in understanding ecological dynamics and the associated feedbacks between hydrology and vegetation dynamics. Most notably, I believe that a major challenge in ecohydrology is to develop much more predictive and well-tested relationships for the partitioning among the subcomponents of evapotranspiration (Loik et al. 2004, Huxman et al. 2005). Evapotranspiration represents the vast majority of the water budget—more than 95% of the total in most arid and semiarid ecosystems (Wilcox et al. 2003*b*). There is great ecological relevance in how this vast majority of the water budget is partitioned among major components, which include at least three: intercepted water that is assumed to evaporate back to the atmosphere, soil evaporation, and plant transpiration. Many models generate predictions about the partitioning among these three components, yet few field studies have rigorously estimated the various components, at least for arid and semiarid ecosystems (Reynolds et al. 2000, Huxman et al. 2005). Those few studies vary in ecosystem type, methods applied, and in time scale of measurements. Hence, we need to improve our ability to predict these components of the water budget and how they vary with vegetation patterns and dynamics. Indeed, some of the most important differences between nondegraded and degraded dryland ecosystems may be evident in the ratio of transpiration to total evapotranspiration (Huxman et al. 2005).

Ecologists have not done much better than hydrologists in tackling the evapotranspiration partitioning (but see Yepa et al. 2003 as an example exception). But perhaps the most important shift for ecologists in moving toward an ecohydrological emphasis is moving away from use of precipitation alone and toward a more comprehensive understanding of the water budget (Loik et al. 2004). In particular, we would like to have a more comprehensive understanding and quantitative ability to predict the amount of "plant-available water" at a site. (This is, of course, interrelated with partitioning components of evapotranspiration.) Precipitation has served as a powerful predictor of plant productivity and other ecological attributes in many systems. When coupled with other climatic variables, it also serves as the underlying driver for biogeography and biogeochemistry models. Yet vegetation dynamics might arguably be much more tightly related to soil moisture, and soil moisture dynamics can differ markedly from patterns of precipitation alone. There are many data sets that have one, two, or even three years of soil moisture data (the old familiar correlation with grant length), and there are several emerging data sets that are five-or-so years in length, thanks to advances in automated data collection for soil moisture and longer term studies such as those associated with the Long Term Ecological Research Network, but there remain few data sets spanning up to a decade or more (e.g., Scott et al. 2000). Arid and semiarid systems characteristically exhibit great interannual variability in precipitation input. We are learning more about how longer climate patterns can persist, and this insight highlights how critical it is to obtain longer-term soil moisture time series. Soil moisture may be much more heterogeneous than we have previously appreciated, varying substantially under trees and shrubs vs. between them, or at a smaller scale, with respect to the presence or absence of biological soil crusts (Loik et al. 2004). Similarly, soil water potential gradients may be affected by vegetation type, and can, surprisingly, draw upward as well as downward (Seyfried et al. 2005). Recent insights about hydraulic lift of water by plants add whole new levels of complexity to understanding ecohydrological processes (e.g., Zou et al. 2005). These factors all require a more detailed and ecologically relevant reassessment of the water budget at a site. As one colleague frequently reminds me, data collection is usually a humbling process.

Unraveling the feedbacks between ecology and hydrology remains challenging and will surely require both modeling and field-based approaches. Continued integration is needed between these two general approaches to avoid the "Do they ever even go out in the field?" vs. "Could they even model their way out of a paper bag?" schism, which is an ongoing challenge in most areas of environmental science. Progress in modeling feedbacks is highlighted in two recent books on ecohydrology: Eagelson's (2002) *Ecohydrology* and Rodríguez-Iturbe and Poporato's (2004) *Ecohydrology of Water-controlled Ecosystems* (2004). These texts both articulate the importance of vegetation in hydrology and the role of feedbacks, with the latter particularly emphasizing the importance of soil moisture. Their strength lies in their attempts to build toward generality, an approach that I applaud. Modeling approaches such as these will be critical to improving our understanding of feedbacks between components of the water budget and vegetation dynamics. It remains critical, however, for such approaches to remain well grounded in ecological processes. Eagelson's seminal papers of the 1970s and 1980s (see Eagelson 2002 and references therein), which laid the groundwork for his recent book, intrigued me when I first read them and continue to stimulate my thinking. Yet Kerkhoff et al. (2004), in a recent publication stemming from the senior author's dissertation (disclosure no. 3: I served on his graduate committee) documents how three fundamental assumptions in the proposed framework are all ecologically flawed. (The three are related to canopy stress minimization, successional stress minimization, and maximum soil productivity.) This example simply highlights one of many areas where further collaboration among ecology and hydrology and further integration of modeling and field-based approaches seems warranted.

Perhaps the clearest success story to date for ecohydrology is the unraveling of the dynamics of ecosystems with banded vegetation, in which the redistribution of runoff alters the spatial distribution of soil moisture and drives vegetation change, which in turn feeds back to runoff patterns (Ludwig et al. 1997, Tongway et al. 2001). In this case, the effects of vegetation on runoff have been clearly documented, as has been the response of vegetation to soil water inputs from runon. Hence the feedback mechanism in this case is nicely demonstrated. Importantly, a clear plan for improving land management has been developed as a result of the new insights for these systems (Ludwig et al. 1997). Similar processes appear to be relevant not only for systems with banded vegetation, but also to some degree for a diverse set of arid and semiarid ecosystems (Wilcox et al. 2003*a*, Ludwig et al. 2005). We need to tackle other areas of ecohydrology with a similar approach, capturing the vegetation effect on hydrological processes, the hydrological effect on vegetation, the resultant feedback dynamics, and the implications and applications for management.

Where is ecohydrology headed? Well, certainly there is a need to fully partition the water budget and to better quantify feedbacks, as discussed above. Other recent interdisciplinary endeavors in ecology such as plant physiological ecology have helped dramatically to reveal underlying mechanisms and to increase predictive capability. Recent progress in ecohydrology offers similar promise. In addition, ecologists are making great progress in explicitly clarifying the ways in which ecosystems provide goods and services to society, something that the hydrologists have had down since the inception of hydrology as a discipline. (You've got to have water.) This is perhaps most clearly highlighted in the new Millenium Ecosystem Assessment (2005). There are many ecohydrological challenges imbedded within the issues raised by the Millenium Ecosystem Assessment, with desertification being among the prominent issues raised. So in addition to improving our ability to partition the water budget and quantify feedbacks, another major issue for ecohydrology is to improve our understanding and ability to predict and manage how ecosystem dynamics affect ecosystem goods and services. I look forward to the challenges ahead with both my ecology and hydrology colleagues, and will enthusiastically embrace the emerging "ecohydrology" emphasis in the hope that we will be able to improve science and serve society through this framework.

Acknowledgments

I thank the following colleagues for their thoughts regarding this commentary and on ecohydrology in general; Brad Wilcox for introducing me to "ecohydrology" as an emerging area and for reminding me that "ecohydrology has been good to me, so I should be good to it"; Craig Allen for identifying my "inner ecohydrologist"; Chris Zou for helping with ecohydrological flow, Travis Huxman for bridging from ecohydrology to ecophysiology; and organizers and contributors to the previous Chapman Conference, ESA symposium, and special AGU sessions on ecohydrology.

Literature cited

Baird, A. J. 1999. Introduction. Pages 1–10 *in* A. J. Baird and R. L. Wilby, editors. Eco-hydrology: plants and water in terrestrial and aquatic environments. Routledge, New York, New York, USA.

Bond, B. 2003. Hydrology and ecology meet—and the meeting is good. Hydrological Processes **17**:2087–2089.

Eagelson, P. S. 2002. Ecohydrology: Darwinian expression of vegetation form and function. Cambridge University Press, New York, New York, USA.

Hannah, D. W., P. J. Wood, and J. P. Sadler. 2004. Ecohydrology and hydroecology: a "new paradigm"? Hydrological Processes **18**:3439–3445.

Huxman, T. E., B. P. Wilcox, D. D. Breshears, R. Scott, K. Snyder, E. A. Small, K. Hultine, W. Pockman, and R. B. Jackson. 2005. Woody plant encroachment and the water cycle: an ecohydrological framework. Special Feature on Ecohydrology. Ecology **86**:308–319.

Kerkhoff, A. J., S. N. Martens, and B. T. Milne. 2004. An ecological evaluation of Eagelson's optimality hypothesis. Functional Ecology **18**:404–413.

Loik, M. E., D. D. Breshears, W. K. Lauenroth, and J. Belnap. 2004. A multi-scale perspective of water pulses in dryland ecosystems: climatology and ecohydrology of the western USA. Special section on Precipitation Pulses in Arid Ecosystems. Oecologia **141**:269–281.

Ludwig, J., D. Tongway, D. Freudenberger, J. Noble, and K. Hodgkinson, editors. 1997. Landscape ecology: function and management. Principles from Australia's rangelands. CSIRO, Collingwood, Australia.

Ludwig, J. A., B. P. Wilcox, D. D. Breshears, D. J. Tongway, and A. C. Imeson. 2005. Vegetation patches and runoff-erosion as interacting ecohydrological processes in semiarid landscapes. Special

Feature on Ecohydrology. Ecology **86**:308–319.

Millennium Ecosystem Assessment. 2005. Ecosystems and human well-being: synthesis. Island Press, Washington, D.C., USA.

Reynolds, J. F., P. R. Kemp, and J. D. Tenhunen. 2000. Effects of long-term rainfall variability on evapotranspiration and soil water distribution in the Chihuahuan Desert: a modeling analysis. Plant Ecology **150**:145–159.

Rodriguez-Iturbe, I.. 2000. Ecohydrology: a hydrologic perspective of climate–soil–vegetation dynamics. Water Resources Research **23**:349–357.

Scott, R. L., W. J. Shuttleworth, T. O. Keefer, and A. W. Warrick. 2000. Modeling multi-year observations of soil moisture recharge in the semiarid American Southwest. Water Resources Research **36**:2233–2247.

Seyfried, M. S., S. Schwinning, M. A. Walvoord, W. T. Pockman, B. D. Newman, R. B. Jackson, and F. M. Phillips. 2005. Ecohydrological control of deep drainage in arid and semiarid regions. Ecology **86**:277–287.

Tongway, D. L., C. Valentin, and J. Seghieri, edtors. 2001. Banded vegetation patterning in arid and semiarid environments: ecological processes and consequences for management. Springer, New York, New York, USA.

Van Dijk, A. 2004. Ecohydrology: it's all in the game? Hydrological Processes **18**:3683–3686.

Wilcox, B. P., D. D. Breshears, and C. D. Allen. 2003*a*. Ecohydrology of a resource-conserving semiarid woodland: effects of scale and disturbance. Ecological Monographs **73**:223–239.

Wilcox, B. P., D. D. Breshears, and M. S. Seyfried. 2003*b*. Water balance on rangelands. Pages 791–794 *in* B. A. Stewart and T. Howell, editors. Encyclopedia of water science. Marcel Dekker, New York, New York, USA.

Wilcox, B. P., and B. D. Newman. 2005. Ecohydrology of semiarid landscapes. Ecology **86**:275–276.

Yepez, E. A., D. G. Williams, R. L. Scott, and G. H. Lin. 2003. Partitioning overstory and understory evapotranspiration in a semiarid savanna woodland from the isotopic composition of water vapor. Agriculture and Forest Meteorology **119**:53–68.

Zou, C. B., P. W. Barnes, S. Archer, and C. R. McMurtry. 2005. Soil moisture redistribution as a mechanism of facilitation in savanna tree–shrub clusters. Oecologia, *in press.* (DOI: 10.1007/s00442-005-0110-8).

David D. Breshears
School of Natural Resources,
Institute for the Study of Planet Earth, and
Department of Ecology and Evolutionary Biology
University of Arizona
Tucson, AZ, 85721-0043 USA
(520) 621-7259
Fax: (520) 621-621-8801
E-mail: daveb@email.arizona.edu

Commentary

A History of the Ecological Sciences, Part 18: John Ray and His Associates Francis Willughby and William Derham

John Ray (1623–1705) was the greatest naturalist and natural theologian of his time. He was assisted early in his career by patron, student, and zoologist Francis Willughby (1635–1672), and late in his career by cleric, natural philosopher, and natural theologian William Derham (1657–1735), who became his literary executor. Ray had a number of other associates who also contributed to his work, especially Martin Lister, Tancred Robinson, and Hans Sloane, all of whose roles are described in Charles E. Raven's encyclopedic biography of Ray (1942). Ray was the first naturalist to emphasize that natural history must be founded on an ability to identify plant and animal species, yet systematics was never the goal of his studies. His interest in natural theology encouraged his investigation into how nature works. Although his adult life was something of a struggle, he was nevertheless a constantly productive naturalist who produced numerous publications (Keynes 1951). The cumulative impact of his work was a major contribution to the Scientific Revolution during the 1600s (Levine 1983).

Ray (spelled Wray until 1670) came from modest circumstances: his father was a blacksmith and his mother a herbal healer. He absorbed her love of plants and religion. Little is known of his childhood, but if he had not been an excellent student, he would never have been admitted to Cambridge University. Arriving in 1664, he prepared for the ministry but showed a strong interest in botany and zoology. Since there were no courses offered in natural history, he joined a group of scholars who dissected animals to study comparative anatomy of vertebrates, and he published the first county flora in England, using as a model Gaspard Bauhin's flora of Basle, Switzerland. Raven (1942:81) described Ray's *Catalogus plantarum circa Cantabrigiam nascentium* (1660) as

> *a small octavo volume suitable for the pocket, is certainly an unpretentious . . . work. Few books of such compass have contained so great a store of information and learning or exerted so great an influence upon the future; no book has so evidently initiated a new era in British botany.*

Fig. 1. John Ray (Ray 1717).

Ray studied Cambridgeshire plants for 6 years before beginning work on the book and then took 3 years to complete it. In deference to the assistance of three friends (named in Ray [1975:24] including Willughby; a letter in Thompson [1974:112] illustrates that assistance), he did not even put his name on the title page. In an age still burdened with polynomials, correlating Cambridgeshire plants with those described in books on British and Continental plants was a demanding task, yet he found and identified 558 species, listed alphabetically, only one of which, a sedge, is of uncertain identity today. Fortunately, Ray's herbarium survives and is in Britain's Natural History Museum (Walters 1981:6–14).

Ray's *Catalogus* is directly relevant to ecology in his accurate recording of places where each species are found—bogs, woods, meadows, riverbanks. More important, he includes biological observations and conclusions. Under ash tree (*Fraxinus excelsior*) he explained the correlation between growth rings seen in a tree stump and the age of the tree, a study at the interface between ecology and physiology (Ray 1660:55; translated by Ewen and Prime, Ray 1975: 64–65):

> *The rings which are seen in the trunks and boughs of trees when cut crossways show more openly in the wood of this tree than in others. These rings in trees growing in the tropics are equidistant all round and have the heart of the tree in the true centre as Gassendus* tom 2. *P.178. observed in the wood of the Brazilian acanthus. In other regions situated either to the south or to the north they are expanded towards the equator and are contracted in the regions facing the pole so that the hearts are always found to be eccentric . . . .*
>
> *1. The age of a tree or branch is disclosed by the number of rings, unless the tree has stopped growing, the number of rings equals the number of years. 2. Normally the inner rings are closer together owing to pressure, probably in trees of great girth and growing old, the outside rings may be narrower through lack of vigour. 3. The pith is compressed as the tree ages; this is evident in the Elder. 4. The wood is harder and darker in the inner rings than in the outer, certainly never lighter . . . . 5. The tops of the trees have fewer rings and the inner rings of the trunk can be seen drawing to a point as they rise; the pattern thus formed is called in English* [he wrote in Latin] *the "grain of the tree". 6…my opinion is that so long as the tree is alive, it adds a new ring, even if it is a narrow one, every year; the age of a tree cannot be determined because its inside decays and the external rings become too narrow to count.*

Ray wrote this 5 years before Hooke announced his discovery of plant cells in *Micrographia*, and since Ray clearly did not examine tree rings under a microscope, he could not explain exactly how the rings grew. Under hops (*Humulus lupulus*) he observed (Ray 1660: 91, translated by Ewen and Prime, Ray 1975:81) that "The hop and probably other twining plants follow the course of the sun, that is they twist from east through south to the west never in the reverse direction. . . ." Under elm tree (*Ulmus procera*) he recorded how the growth of trees in the open isinfluenced by prevailing winds (Ray 1660:180; translated by Ewen and Prime, Ray 1975:126):

> *From the shape of a tall tree growing in the open air it is possible to say from what quarter of the heavens the stronger and more prevalent winds are accustomed to blow in any particular locality. Thus trees growing near the sea point to the east because those parts of the country are particularly exposed to frequent gales.*

He also explained some animal uses of various species. Under hemlock (*Conium maculatum*) he reported (Ray 1660:34; translated by Ewen and Prime, Ray 1975:54) that he had dissected the crop of a bustard (*Otis tarda*) and "found it stuffed with hemlock seeds;

there were only four or five grains of corn mixed with them. So even at harvest the bird leaves corn for hemlock." If Ray hoped this observation on food preference might help save bustards from farmers' ire, it seems unsuccessful—the last bustard was killed in Britain in 1835. Under deadly nightshade (*Atropa belladonna*) he commented (Ray 1660:157–158; translated by Ewen and Prime, Ray 1975:114) that snails and slugs commonly eat it despite its poison. (He added that these animals are hermaphroditic.)

His longest note under any plant is not about the plant itself but about its habitual insect pest. The discussion is under rape (*Brassica rapa*) and wild turnip (*B. napus*), where he reported (Ray 1660:134; translated by Ewen and Prime, Ray 1975:102) that "Caterpillars born on brassica have taught us that a close relationship exists between these stocks as the leaves of rape are eaten no less greedily that those of brassica although they scorn many other plants that we have offered them as food." He raised 10 or so of these caterpillars in a wooden box at the end of August 1658 and inadvertently discovered insect parasitism, but without fully understanding it (Ray 1660:134–138; translated by Ewen and Prime, Ray 1975:103):

> *Seven of them proved to be viviparous or vermiparous; from their backs and sides very many, from thirty to sixty apiece wormlike animalcules broke out; they were white, glabrous, footless and under the microscope [perhaps only a magnifying glass] transparent. As soon as they were born, they began to spin silken cocoons, finished them in a couple of hours, and in early October came out as flies, black all over with reddish legs and long antennae, and about the size of a small ant. The three or four caterpillars which did not produce maggots changed into angular and humped chrysalids which came out in April as white butterflies.*

He also described a case of parasitism of *Rosa canina* by the rose bedeguar (*Rhodites rosae*), and commented on previous authors' observations on the subject (Ray 1660:139–140; translated by Ewen and Prime, Ray 1975:105):

> *Sometimes a smooth hairy lump grows on the stalks of…*[Rosa canina]. *If you cut open this gall, you will find it packed with small white maggots; this is on the evidence of* Bacon nat. hist. cent.6 exp.562. Spigel isag. lib. 1, cap.10. Moufet. Theat insect. lib.2, cap.20. . . . *Spiegel, Moufet and Aristotle* (Arist. Lib.5. hist. cap.19) *say that beetles are borne from these maggots….*[but] *the maggots which lie hidden in the gall during the winter come out in the month of May the following year in the form of flies; their shape and proportion are like those of winged ants; their size is a little smaller . . . . Some of these flies are armed with a sting or spike protruding from the tail but others altogether lack this, so this probably makes a distinction between the sexes.*

Raven (1942:102–103) points out that some of Ray's observations on insects published in this first book were extended in his last book, *Historia insectorum* (1710); for example, he expanded his observations on insect galls in it on pages 259–260.

After sending his catalogue of Cambridgeshire plants to the printers (in Cambridge and London), Ray and Willughby took their first extended field trip, to northern England and the Isle of Man, which is equidistant between northern England and northern Ireland. The friends decided to compile natural histories of British plants and animals, and since Willughby's stronger interest was in animals, he would do them and Ray would do the plants. Before returning to Cambridge, Ray visited Thomas Brown at Norwich in August and they botanized along the Norfolk coast. Ray and Willughby's collaboration was very productive, though Willughby never got beyond the note-taking stage before his death at age 37 in 1672. In 1658, 1661, and 1662 Ray went on field trips without Willughby into other parts of Britain.

Fig. 2. Francis Willughby (Allen 1951).

Ray had trained for the ministry and was ordained, and he had intended natural history to be only an avocation. However, in 1662, after the Restoration, a Royalist Parliament passed a law requiring all ministers to sign a loyalty oath, and Ray, a Puritan, felt it violated his religious belief. His refusal to sign ended his clerical career, and his avocation became his life's work. In 1663 he and Willughby left on a 3-year trip to Western Europe to collect observations, specimens, and illustrations and to visit professors at several universities and a few unaffiliated naturalists (Ray 1673, Raven 1942:112–140, Allen 1951:419–422). This experience enabled the partners to broaden the scope of their studies beyond Britain, first to western Europe and later to the rest of the world known to Europeans.

In 1660, the 25-year-old Willughby had become a founding Fellow of the Royal Society of London—which happened at that young age only because he came from the nobility. In 1667 Ray was elected a Fellow, and in 1669 Willughby and he sent in "Experiments concerning the Motion of Sap in Trees, Made this Spring by Mr. Willughby and Mr. Wray," which the Society published in its *Philosophical Transactions*. Willughby had returned from the Continental trip before Ray and had begun these experiments himself (Welch 1972:76). Their experiments were exploratory, without a hypothesis, in a Baconian manner. Although Raven (1942:188) admitted that they made no fundamental discovery, he thought that this was "the first systematic attempt to study the physiology of a living plant and thus opened up a new field of research and gave a new direction to botany." In claiming such priority for Ray, however, Raven failed to consider the studies before 1669, discussed in Part 14 (Egerton 2004:210), though Raven may be right about these experiments stimulating studies by others. Botanist and historian Agnes Arber (1943) cited other examples in which Raven slighted other botanists while praising Ray. More recently, however, Morton (1981:210) followed Raven's example in claiming Ray as "the founder of plant physiology, even though his original contributions were modest." He based his judgment largely on the discussion of plant physiology in Volume 1 of Ray's *Historia Plantarum* (three volumes, 1686–1704); this is generally considered Ray's greatest scientific treatise. Ray was the first naturalist to pay special attention to the distinction between species, and he wrote his first essay on the subject in 1672 (published in 1757 and reprinted in Ray 1928:77–83). His later expression of his species concept in *Historia Plantarum* was long standing (Ray 1686:Volume I; translation by E. Silk, in Mayr 1982:256–257):

> *After a long and considerable investigation, no surer criterion for determining species has occurred to me than the distinguishing features that perpetuate themselves in propagation from seed. Thus, no matter what varia-*

tions occur in the individuals or the species, if they spring from the seed of one and the same plant, they are accidental variations and not such as to distinguish a species . . . . Animals likewise that differ specifically preserve their distinct species permanently: one species never springs from the seed of another nor vice versa.

In a different context, Ray explained, "I reckon all Dogs to be of one Species, they mingling together in Generation, and the Breed of such Mixtures being prolifick" (Ray 1717:21). Ray made important contributions to the classification of plants (Stevenson 1947, Sloan 1972, Morton 1981:201–203, 228–229, Stearn 1985–1986:113–117), including drawing a distinction between herbaceous Monocotyledons and Dicotyledons in his Methodus Plantarum (1682). Ray is often credited with being first to make this distinction (Raven 1942:195, Morton 1981:203, 228–229), but Mayr (1982:163) cites four predecessors. Although Ray was able to obtain funds to publish illustrations in the treatises on ornithology (from Emma, Willughby's widow) and ichthyology (from the Royal Society), both of which carried Willughby's name as author, he was unable to obtain funds for plates of the different species for his own books on plants (Henrey 1975:127–134, 266–269).

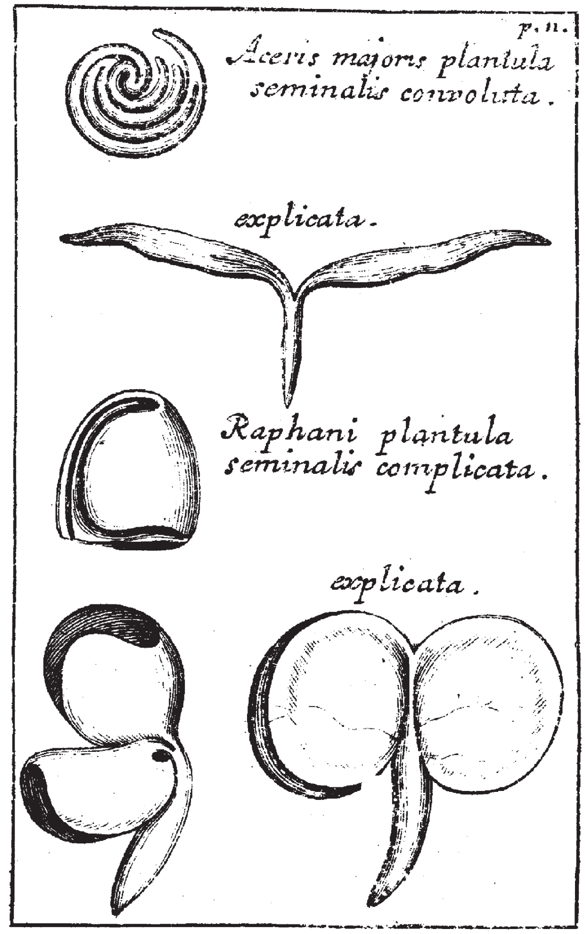

Fig. 3. This drawing of sycamore and radish seeds from Malpighi's *Anatomy of Plants* (1675) is reprinted in Ray's *Methodus Plantarum* (1682) and *Historia Plantarum* (1686).

Soon after Willughby's death in 1672, Ray turned to producing Willughby's *Ornithology*, which was a memorial to his patron and became the beginning of modern ornithology. Although he placed Willughby's name alone on the title page as author, Ray's contribution to the book was as much or more than Willughby's careful notes and collected illustrations (many from their European tour). This point is seen in an extract from Ray's letter on various birds to Martin Lister, 1 October 1667 (Ray 1928:113–115 [in Latin], Raven translation 1942:315):

Your observation of the Green Woodpecker corresponds with my own of the Black and both the Spotted Woodpeckers and the Wryneck. I once got out of the crops of these birds on dissection larvae as big as my small finger. The muscles and tendons by which they shoot out and retract their tongues deserve curious study.

Although Ray initially published the *Ornithology* in Latin (1676), 2 years later he published an enlarged English version. Two modern histories of ornithology (Stresemann 1975:43–45, Walters 2003:38–40) stress the importance of these authors' new classification of birds, and Ray performed the same service in 1693 for quadrupeds (Petit and Théodoridès 1962:317–320). Raven (1942:308–338) provides the most details on the *Ornithology*'s production and contents; Hall (1951:18–30) quotes the classification, human bird-hunting techniques, and the dodo; and Miall (1912:103–111) presents a briefer overview than Raven and more natural history extracts than Hall.

In the *Ornithology*, Chapter 3, "Of the generation of birds," our authors disagreed with William Harvey's belief (1651, exercise 29) that some hen eggs only come into existence after copulation. They thought (Willughby [and Ray] 1678:10–16) that hens are born with all the eggs in their ovaries that they ever lay. They cited five cases of longevity that seemed credible to them: a goose and a pelican had each been kept for 80 years; a pigeon had lived 22 years and had bred until its last 6 months; a linnet lived 14 years, and a goldfinch 23 years. When pigeons raise two young, Willughby wondered whether they were of opposite sexes; Ray replied that they usually are but sometimes are not.

Aristotle's *Historia Animalium* (600a15) claimed that swallows do not migrate in winter as other birds do, but hibernate, and naturalists revived this belief from the 1500s to the 1700s. Willughby and Ray (1678:212, quoted in Raven 1942:328) doubted this: "To us it seems more probable that they fly away into hot countries, viz. Egypt, Aethiopia etc." Their many natural history observations of ecological interest are illustrated in these six examples:

Lapwing (*Vanellus vanellus*) (Willughby [and Ray] 1678:308, quoted in Miall 1912:109–110):

It builds its nest on the ground, in the middle of some field or heath, open and exposed to view, laying only some few straws or bents under the eggs, that the nest be not seen. The eggs being so like in colour to the ground on which they lie, it is not easy to find them though they lie so open. The young, so soon as they are hatcht, instantly forsake the nest, running away, as the common tradition is, with the shell upon their heads, for they are covered with a thick down, and follow the old ones like chickens. They say that a lapwing, the further you are from her nest, the more clamorous she is, and the greater coil she keeps; the nearer you are to it, the quieter she is, and less concerned she seems, that she may draw you away from the true place, and induce you to think it is where it is not.

Fig. 4. Plate from Willughby [and Ray] (1678).

Blackbird (*Turdus merula*) (Willughby [and Ray] 1678:191, quoted in Miall 1912:111):

The blackbird builds her nest very artificially with outside of moss, slender twigs, bents and fibers of roots, cemented and joined together with clay instead of glue, daubing it also all over withinside with clay. Yet doth she not lay her eggs upon the bare clay, like the mavis, but lines it with a covering of small straws, bents, hair, or other soft matter, upon which she lays her eggs, both that they might be more secure and in less danger of breaking, and also that her young might lie softer and warmer.

Honey-Buzzard (*Pernis apivorus*) (Willughby [and Ray] 1678:72, quoted in Raven 1942:327):

It builds its nest of small twigs, laying upon them wool and upon the wool its eggs. We saw one that made use of an old Kite's nest to breed in, and that fed its young with the nymphae of wasps: for in the nest we found the combs of wasps' nests and in the stomachs of the young the limbs and fragments of wasp-maggots. There were in the nest only two young ones, covered with a white down, spotted with black. Their feet were of a pale yellow, their bills between the nostrils and the head white. Their craws large, in which were Lizards, Frogs etc….This bird runs very swiftly like a Hen. The female as in the rest of the Rapacious kind is in all dimensions greater than the male.

Dipper (*Cinclus cinclus*) (Willughby [and Ray] 1678:149, quoted in Raven 1942:327–328):

It frequents stony rivers and water-courses in the mountainous parts of Wales, Northumberland, Yorkshire etc. That I (J.R.) described was shot beside the river Rivelin near Sheffield in Yorkshire: that Mr Willughby described near Pentambeth in Denbighshire in North Wales. It is common in the Alps in Switzerland, where they call it Wasser-Anzeil. It feeds upon fish, yet refuseth not insects. Sitting on the banks of rivers it now and then flirts up its tail. Although it be not web-footed, yet will it sometimes dive or dart itself quite under water. It is a solitary bird, companying only with its mate in breeding time.

Cormorant (*Phalacrocorax carbo*) (Willughby [and Ray] 1678:330, quoted in Raven 1942:328):

. . . on the rocks of Prestholm Island near Bearmaris we saw a Cormorant's nest, and on the high trees near Sevenhays in Holland abundance . . . . besides this we have not known or heard of any whole-footed bird that is wont to sit upon trees, much less build its nest upon them.

Puffin (*Puffinus puffinus*) on the Isle of Man (Willughby and Ray 1678:333, quoted in Raven 1942:328):

The old ones early in the morning, at break of day, leave their nests and young and the island itself and spend the whole day in fishing in the sea…so that all day the island is so quiet and still from all noise as if there were not a bird about it. Whatever fish or other food they have gotten and swallowed in the day-time, by the innate heat or proper ferment of the stomach is (as they say) changed into a certain oily substance (or rather chyle) a good part whereof in the night-time they vomit up into the mouths of their young, which being therewith nourished grow extraordinarily fat.

The story of Willughby [and Ray]'s *Historia Piscum* (1686) is similar to that of the *Ornithology*: it was a joint effort, with editor Ray contributing more than Willughby. The latter had left fewer notes on fish than on birds, and Ray supplemented them by soliciting information from his naturalist colleagues (Raven 1942:339–370). The resulting volume contained many fewer natural history observations of ecological interest than the bird volume, no doubt because fish behavior is more difficult to observe than bird behavior. Miall (1912:112) pointed out that the fish volume depended heavily upon previous books by Rondelelt, Belon, Salviani, Gesner, and Marcgraf, and therefore "It cannot be said that this is a very important contribution to natural history." Even the observation from *Historia Piscium* that Miall mentioned, about sharks having the mouth on their bottom side as a provision of nature to ensure safety of other fish and to prevent sharks from dying from gluttony, is actually repeated from Aristotle's *De Partibus Animalium* (696b24–33, quoted in Egerton 1973:328–329). Nevertheless, two historians of ichthyology thought highly of this work. Cuvier in 1828 wrote (Simpson translation 1995:71):

> *Ray and Willughby had the honor of being the first to write an ichthyology in which the fishes were clearly described according to nature and classified based on characteristics drawn only from their structure, and in which their natural history was finally rid of all passages from ancient writings…*

More recently, Jordan wrote (1905:390) that "The basis of classification was first fairly recognized by" Ray and Willughby in *Historia Piscium*, which brought "order out of the confusion left by their predecessors." Their treatise described 180 species directly from nature and described 240 more from other authors' works. There was no later English edition.

In 1690 when Ray, age 63, began work on his *Historia Insectorum*, his health was already in decline. However, we saw above in notes to his Cambridgeshire catalogue of plants (1660) that he had an early interest in insects, the persistence of which is illustrated in this extract from his letter to Lister on 17 July 1670 (Ray 1718:69, quoted from Salmon 2000:252):

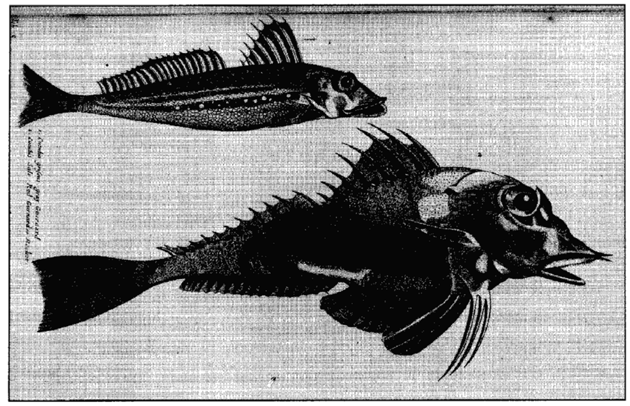

Fig. 5. Top: Flying gurnard (*Cephalacanthus volitans*); bottom: sea-robin (*Prionotus evolans*). Willughby [and Ray] 1686: Plate 13.

> *This summer we found here the same horned* Eruca [caterpillar], *which you observed about Montpelier, feeding on* Foeniculum [Seseli] tortuosum. *Here it was found on common Fennel: It has already undergone the first change into a* Chrysalis*, and we hope it will come out a* Butterfly *before winter.*

Ray also published a note on ants in 1671. Willughby's notes available to Ray were not limited to insects, but included worms and other invertebrates. As usual, Ray solicited and received help from other naturalists, and he used Lister's observations on spiders and beetles. For this project he was also aided by his wife, Margaret, and their four daughters—Margaret, Mary, Catharine, and Jane—who collected insects around their Black Notley home. In gratitude, he named several newly discovered butterflies and moths after his daughters. On 29 May 1693 his wife made an important discovery concerning a moth which Raven thinks was probably *Pachys betularia* (Ray 1710:177, Raven translation 1942:395):

> *It emerged out of a stick-shaped geometer caterpillar: it was a female and came out from its chrysalis shut up in my cage: the windows were open in the room or closet*

where it was kept, and two male moths flying round were caught by my wife who by a lucky chance were into the room in the night: they were attracted, as it seems to me, by the scent of the female and came in from outside.

Raven suggests that this was probably the first record of insect pheromones (though Raven did not use that term).

James Petiver (1663–1718) was Ray's only significant predecessor in naming British insects. He was a London apothecary (pharmacist) and nature collector who published on insects from 1695 until 1717 (Allen 2004, Stearns 1952, Salmon 2000:103). He and Ray were friends, not competitors, and he provided valuable assistance. In 1660 when Ray reported caterpillars producing flies instead of butterflies when chrysalises opened in the spring, he had been unsure how to interpret his observations. By the time he wrote *Historia Insectorum*, however, he understood (Ray 1710:114, translated by Raven 1942:104):

I think that the ichneumon wasps prick these caterpillars with the hollow tube of their ovipositor and insert eggs into their bodies: the maggots are hatched by the warmth of them, and feed there until they are full grown: then they gnaw through the skin, come out, and spin their cocoons.

There was no English edition of *Historia Insectorum* either. However, Bodenheimer (1928–1929:I, 486–494; II, 412–427) provided a German translation of extracts and also modern identifications of insects Ray discussed.

In addition to the ecological observations scattered through Ray's numerous natural histories, he also emphasized interactions among plants and animals in a more coherent way in his very influential book on natural theology, *The Wisdom of God Manifested in the Works of the Creation* (1691). The idea that one can learn about God by studying his creation arose among the ancient Greeks, and there were two basic arguments: (1) that the lives of plants and animals are designed to intertwine in ways to preserve the balance of nature, and (2) that the structure and function of the organs of the human body are designed to enable humans to flourish. The most famous discussion from antiquity of the former argument is in *De Natura Deorum* by Cicero, a first century BC Roman (Glacken 1967:54–61, Egerton 1973:30). Discussion of the latter argument is in several writings by the Greek physician Galen, in the 100s AD. In a very comprehensive survey of the history of natural theology, Neal C. Gillespie (1987) argues that there were few original contributions to the subject since Cicero and Galen until Ray, and that Ray made the most important contributions down to the time when the whole subject was challenged by Darwin's *Origin of Species* (1859). Glacken (1967:415–442), Raven (1942:452–478), and Zeitz (1994) essentially agree. Although I think Matthew Hale's *Primitive Origination of Mankind* (1677) was a more substantial contribution to the subject than Gillespie recognized (Egerton 2005), there is no doubt about the overwhelming importance of Ray's book on natural theology.

Ray had expressed his strong skepticism of spontaneous generation of animals in a note published in 1671, and that skepticism remained in his later writings. Arguments that he accumulated over the years were explained in *The Wisdom of God* (cited from the seventh edition, 1717:298–326, 1977). Another of Ray's concerns was the possible extinction of species. Since antiquity, it had been argued that all species are endowed with effective means of preservation (Egerton 1973). If any species actually became extinct, it could reflect against God's omnipotence or creative wisdom. Particularly worrisome were large fossil ammonites. Only the much smaller chambered nautilus had ever been found alive. Ray (1692:19–124) did not take a dogmatic stand, but pointed out that much of the world remained unexplored by European naturalists. In a posthumous essay, "Mr. Ray of the Number of Plants," (in Derham 1718:344–351), he also argued against the origin of new species or the extinction of previously existing species. Although he could not prove that species do not become extinct, he could emphasize their means of survival. This was another theme that went back to antiquity (Egerton 1973), and Ray (1717:110–146, 1977) marshaled the usual evidence along with a few new examples, including Lister's observation that swallows, like chickens, will continue laying eggs if previous eggs are removed from the nest daily (until 19 were laid), and Ray's own observation about woodpeckers' tongues being designed to extract insects from the trunks of trees or limbs.

A somewhat newer question, or at least newly answered (Ray 1717:368–373, 1977), was why there are multitudes of noxious insects. First, because it displays the riches of the power and wisdom of God. Second, because insects are eaten by other animals, many individuals are needed to prevent their extinction. Third, because insects are important food for birds, fishes, and various quadrupeds. Among his examples is an implicit food chain. Derham had, using a microscope, studied "those vastly small animalcula" (zooplankton), and found that they were food for small insects, which Ray had just said were eaten by fish, and of course he knew that people ate fish. Fourth, God can use noxious insects when he wishes to punish wicked persons or nations. Since it was known that insect pests are much worse in some places than in others, one may wonder why Ray did not conclude from his fourth point that wicked people are attracted to areas with many insect pests and virtuous people are not. That thought was "beyond the radar" of natural theologians, including Ray.

William Derham was a clergyman in Upminster, a town near London, which occupation left him with ample time to pursue his scientific studies (Atkinson 1952, Knight 1971, Smolenaars 2004). He became a Fellow of the Royal Society in 1702 and published 46 articles in its *Philosophical Transactions*, 1698–1735, many of them concerning the weather at Upminster. His justification for a clergyman engaging in scientific studies was that they provided material for his own two books on natural theology, *Physico-Theology* (1713) and *Astro-Theology* (1715), both of which were very popular and went through many editions and translations into other languages. Derham was bound to cover some of the same ground as Ray, but Derham also had new information and new perspectives (Glacken 1967:421–424). He provided a new synthesis of animal and human demography (Egerton 1967:135–144), and he had a larger store of knowledge of them than had Matthew Hale 36 years earlier (Egerton 2005). Derham seems to have first actually used the word "balance" in a discussion of the balance of nature (Derham 1716:171, 1977): "Thus the Balance of the Animal World is, throughout all Ages, kept even, and by a curious Harmony, and just Proportion between the increase of all Animals, and the length of their Lives, the world is through all Ages well, but not overstored." In discussing human demography, he drew upon the studies by John Graunt and later authors. He saw (1716:177, 1977) the tendency of births to be more numerous than deaths as

> *an admirable Provision for the extraordinary Emergencies and Occasions of the World; to supply unhealthful Places, where Death out-runs Life; to make up the Ravages of great Plagues, and Diseases, and the Depredations of War and the Seas; and to afford a sufficient number for Colonies in the unpeopled Parts of the Earth.*

He suggested that some of these calamities might be punishment for wickedness and also "wise Means to keep the Balance of Mankind ['s population] even. . . .

Ray had defended the wisdom of having mountains as providing a variety of abodes for a variety of species of plants and animals. Derham generalized this argument to explain that the diversity of soils and climates of the earth provide the needs for the large variety of existing species. In his chapter, "Of the Food of Animals," he further observed (1716:180–215, 1977) that animal species have special kinds of food, and also special anatomical features that enable them to obtain it, such as the long bills of woodcocks, snipes, and curlews, which they use to extract worms from the soil. It would have been difficult to make his argument for what we call ecological diversity had Derham chosen omnivorous species as examples; that is one limitation to his argument, and his neglect of competition between species is another. Perhaps a focus on the wisdom of creation diverted attention from these aspects of species interactions.

Thus, natural theology had limitations as a paradigm for understanding the living world. However, as a motivator for natural history studies, it played an important role in the thinking of European and American naturalists from the 1600s into the 1800s. John Ray and his associates, Francis Willughby and William Derham, provided the guidance and inspiration for many of these studies.

Literature cited

Allen, D. E. 2004. James Petiver (1663–1718), botanist and entomologist. Oxford Dictionary of National Biography, 3 pages, ‹oxforddnb.com›

Allen, E. G. 1951. The history of American ornithology before Audubon. American Philosophical Society Transactions **41**:385–591.

Arber, A. 1943. A seventeenth-century naturalist: John Ray. Isis **34**:319–324.

Atkinson, A. D. 1952. William Derham, F.R.S. (1657–1735). Annals of Science **8**:368–392.

Baldwin, S. A. 1986. John Ray (1627–1705): Essex naturalist. Baldwin Books, Witham, UK.

Bodenheimer, F. S. 1928–1929. Materialien zur Geschichte der Entomologie bis Linné. Two volumes. W. Junk, Berlin, Germany.

Cuvier, [J.L.N.F.] G. 1995. Historical portrait of the progress of ichthyology from its origins to our own time. Translated by A. J. Simpson. Johns Hopkins University Press, Baltimore, Maryland, USA.

Derham, W. 1716. Physico-Theology: or, a demonstration of the being and attributes of God, from his works of creation. Edition 4. W. Innys, London, UK.

Derham, W., editor. 1718. Philosophical letters between the late learned Mr. Ray and several of his ingenious correspondents, natives, and foreigners. W. and J. Innys, London, UK.

Derham, W. 1977. Physico-theology: or, a demonstration of the being and attributes of God, from his works of creation. [Reprint of 1716 edition.] Arno Press, New York, New York, USA.

Egerton, F. N. 1967. Observations and studies of animal populations before 1860: a survey concluding with Darwin's *Origin of Species*. Dissertation. University of Wisconsin, Madison, Wisconsin, USA.

Egerton, F. N. 1973. Changing concepts of the balance of nature. Quarterly Review of Biology **48**:322–350.

Egerton, F. N. 2004. A history of the ecological sciences, Part 14: plant growth studies during the 1600s. ESA Bulletin **85**:208–213.

Egerton, F. N. 2005. A history of the ecological sciences, Part 15: the precocious origins of human and animal demography and statistics in the 1600s. ESA Bulletin **86**:32–38.

Gillespie, N. C. 1987. Natural history, natural theology, and social order: John Ray and the Newtonian ideology. Journal of the History of Biology **20**:1–49.

Glacken, C. J. 1967. Traces on the Rhodian shore: nature and culture in western thought from ancient times to the end of the eighteenth century. University of California Press, Berkeley, California, USA.

Gunther, R. W. T. 1937. Early science in Cambridge. Oxford University Press, Oxford, UK.

Hall, T. S. 1951. A source book in animal biology. McGraw-Hill, New York, New York, USA.

Henrey, B. 1975. British botanical and horticultural literature before 1800. Volume 1. The sixteenth and seventeenth centuries, history and bibliography. Oxford University Press, Oxford, UK.

Jordan, D. S. 1905. A guide to the study of fishes. Volume 1. Henry Holt, New York, New York, USA.

Keynes, G. L. 1951. John Ray: a bibliography. Third edition. Faber and Faber, London, UK. 1976.

Knight, D. M. 1971. William Derham (1657–1735).

Dictionary of Scientific Biography **4**:40–41.
Levine, J. M. 1983. Natural history and the history of the scientific revolution. Clio **13**:57–73.
Mandelbrote, S. 2004. John Ray [*formerly* Wray] (1627–1705), naturalist and theologian. Oxford Dictionary of National Biography, 8 pages. ‹oxforddnb.com›
Mayr, E. 1982. The growth of biological thought: diversity, evolution, and inheritance. Harvard University Press, Cambridge, Massachusetts, USA.
Miall, L. C. 1912. The early naturalists: their lives and work (1530–1789). Macmillan, London, UK.
Mickel, C. E. 1973. John Ray: indefatigable student of nature. Annual Review of Entomology **18**:1–16.
Morton, A. G. 1981. History of botanical science: an account of the development of botany from ancient times to the present day. Academic Press, London, UK.
Petit, G., and J. Théodoridès. 1962. Historie de la zoologie des origines a Linné. Hermann, Paris, France.
Raven, C. E. 1942. John Ray, naturalist: his life and works. Cambridge University Press, Cambridge, UK.
[Ray, J.] 1660. Catalogus plantarum circa Cantabrigiam nascentium. Jeann. Field, Cambridge, UK.
Ray, J. 1671*a*. Concerning some un-common observations and experiments made with an acid juyce to be found in ants. Royal Society of London Philosophical Transactions **5**:2069–2077.
Ray, J. 1671*b*. Concerning spontaneous generation. Royal Society of London Philosophical Transactions **6**:2219–2220.
Ray, J. 1673. Observations topographical, moral, & physiological; made in a journey through part of the low-countries, Germany, Italy, and France: with a catalogue of plants not native of England. John Martyn for the Royal Society, London, UK.
Ray, J. 1686–1704. Historia plantarum species hactenus editas aliasque insuper multas noviter inventas & descriptas complectens. Three volumes. Mariae Clark, London, UK.
Ray, J. 1691. The wisdom of God manifested in the works of the creation. Samuel Smith, London, UK.
Ray, J. 1692. Miscellaneous discourses concerning the dissolution and changes of the world. Samuel Smith, London, UK. Reprinted 1968. Georg Olms, Hildesheim, Germany.
Ray, J. 1717. The wisdom of God manifested in the works of the creation. Edition 7. William Innys, London, UK.
Ray, J. 1928. Further correspondence. Edited by R.W.T. Gunther. Ray Society, London, UK.
Ray, J. 1975. Flora of Cambridgeshire. Translated by A. H. Ewen and C. T. Prime. Wheldon and Wesley, Hitchin, UK.
Ray, J. 1977. The wisdom of God manifested in the works of the creation. [Reprint of 1717 edition.] Arno Press, New York, New York, USA.
Salmon, M. 2000. The Aurelian legacy: British butterflies and their collectors. University of California Press, Berkeley, California, USA.
Sloan, P. R. 1972. John Locke, John Ray, and the problem of the natural system. Journal of the History of Biology **5**:1–53.
Smolenaars, M. 2004. William Derham (1657–1735), Church of England clergyman and natural philosopher. Oxford Dictionary of National Biography, 2 pages. ‹oxforddnb.com›
Stearn, W. T. 1985–1986. John Wilkins, John Ray, and Carl Linnaeus. Royal Society of London Notes and Records **40**:101–123.
Stearns, R. P. 1952. James Petiver, promoter of natural science. American Antiquarian Society Proceedings **62**:243–365.
Stevenson, I. P. 1947. John Ray and his contributions to plant and animal classification. Journal of the History of Medicine and Allied Sciences **2**:250–261.
Stresemann, E. 1975. Ornithology from Aristotle to the present. Translated by H. J. Epstein and C. Epstein. Harvard University Press, Cambridge, Massachusetts, USA.
Thompson, R. 1974. Some newly discovered letters of

John Ray. Society for the Bibliography of Natural History Journal **7**:111–123.

Walters, M. 2003. A concise history of ornithology. Yale University Press, New Haven, Connecticut, USA.

Walters, S. M. 1981. The shaping of Cambridge botany. Cambridge University Press, Cambridge, UK.

Webster, C. 1975. John Ray (1627–1705). Dictionary of Scientific Biography **11**:313–318.

Welch, M. A. 1972. Francis Willughby, F.R.S. (1635–72). Society for the Bibliography of Natural History Journal **6**:71–85.

Welch, M. A. 1976. Francis Willughby (1635–72). Dictionary of Scientific Biography **14**:412–414.

Willughby, F., and J. Wray. 1669. Experiments concerning the motion of sap in trees, made this spring. Royal Society of London Transactions **4**:963–965.

Willughby, F., [and Ray, J.]. 1678. The ornithology…in three books. Wherein all the birds hitherto known, being reduced into a method suitable to their natures, are accurately described…to which are added three considerable discourses. John Martyn for the Royal Society, London, UK.

Willughby, F., [and Ray, J.]. 1686. De historia piscium libri quatuor. Theatro Sheldoniano, Oxford, UK. Reprinted 1977. Arno Press, New York, New York, USA.

Zeitz, L. M. 1994. Natural theology, rhetoric, and revolution: John Ray's *Wisdom of God*, 1691–1704. Eighteenth Century Life **18**:120–133.

Acknowledgments

I thank David E. Allen, Wellcome Institute for the History of Medicine, Jean-Marc Drouin, Muséum National d'Histoire Naturelle, Paris, and Anne-Marie Drouin-Hans, Université de Bourgogne.

Frank N. Egerton
Department of History
University of Wisconsin-Parkside
Kenosha WI 53141
E-mail: frank.egerton@uwp.edu