

## USE OF STRUCTURE COEFFICIENTS IN PUBLISHED MULTIPLE REGRESSION ARTICLES: $\beta$ IS NOT ENOUGH

TROY COURVILLE  
Sam Houston State University

BRUCE THOMPSON  
Texas A&M University

The importance of interpreting structure coefficients throughout the General Linear Model (GLM) is widely accepted. However, regression researchers too infrequently consult regression structure coefficients to augment their interpretations. The authors reviewed articles published in the *Journal of Applied Psychology* to determine how interpretations might have differed if standardized regression coefficients and structure coefficients (or else bivariate *rs* of predictors with the criterion) had been interpreted. Some dramatic misinterpretations or incomplete interpretations are summarized. It is suggested that beta weights and structure coefficients (or else bivariate *rs* of predictors with the criterion) ought to be interpreted when noteworthy regression results have been isolated.

Following the publication of Cohen's (1968) seminal article, "Multiple Regression as a General Data-Analytic System," social scientists began to take seriously the notion that regression is the univariate General Linear Model (GLM). Knapp (1978) subsequently extended this view by presenting canonical correlation analysis as the multivariate GLM. More recently, structural equation modeling has been represented as the most general case of the general linear model, when measurement modeling is incorporated simultaneously into substantive modeling (Bagozzi, Fornell, & Larcker, 1981). Even when only measurement modeling is conducted, the same basic variance partitioning methods used in substantive modeling are applied, albeit for a different purpose (Dawson, 1999).

---

Xitao Fan served as Action Editor for this manuscript.

Educational and Psychological Measurement, Vol. 61 No. 2, April 2001 229-248  
© 2001 Sage Publications, Inc.

The general linear model has helped researchers come to understand that all parametric analyses (a) are correlational, (b) apply some system of weights to measured/observed variables to estimate scores on composite or synthetic variables that then become the analytic focus, and (c) yield effect size analogs of  $r^2$  values (e.g., Fan, 1996, 1997; Thompson, 1991, 2000). This last realization ultimately helped lead to encouraging effect size reporting in all research (American Psychological Association [APA], 1994, p. 18) and, more recently, to the repeated suggestion by the APA Task Force on Statistical Inference that effect size reporting should occur in all research (Wilkinson & APA Task Force on Statistical Inference, 1999).

The Task Force emphasized, “*Always* [italics added] provide some effect-size estimate when reporting a  $p$  value” (Wilkinson & APA Task Force on Statistical Inference, 1999, p. 599). Later, the Task Force wrote,

*Always* [italics added] present effect sizes for primary outcomes. . . . It helps to add brief comments that place these effect sizes in a practical and theoretical context. . . . We must stress again that reporting and interpreting effect sizes in the context of previously reported effects is *essential* [italics added] to good research. (p. 599)

In construct or predictive validity or in substantive regression studies, once statistically significant and/or noteworthy effects are detected, the origins of detected effects must then (and only then) be explored. Of course, many researchers interpret standardized regression (sometimes called  $\beta$  or beta) weights for this purpose. Many such researchers deem unimportant any predictor variables with near-zero beta weights.

However, the flaws of interpreting only beta weights have been noted. Thompson and Borrello (1985) argued that structure coefficients are just as important in regression as they are in other GLM methods, such as descriptive discriminant analysis and factor analysis. Indeed, the editorial board members of *Educational and Psychological Measurement* subsequently cited that article as one of the most important measurement-related publications of the past 50 years (Thompson & Daniel, 1996).

The purpose of the present study was to characterize what regression researchers are actually doing in published regression research as regards the interpretation of regression results using  $\beta$  weights and/or structure coefficients. In particular, we wanted to summarize the interpretations offered by authors in published regression research, and where possible conduct our own interpretations by supplementary analyses that then led to independent interpretations on our part. We wanted to determine how dramatic the differences might be for actual versus alternative interpretations in which  $\beta$  weights and structure coefficients (or else bivariate  $r$ s of predictors with the criterion) are interpreted as part of a more complete system of results.

### Regression Interpretation Strategies

On a superficial first-glance basis, a regression interpretation focusing solely on  $\beta$  weights erroneously seems reasonable, given one formula for the multiple  $R^2$  effect size (e.g., Thompson, 1995):

$$R^2 = \beta_1 (r_{YX_1}) + \beta_2 (r_{YX_2}) + \dots + \beta_p (r_{YX_p}). \quad (1)$$

A superficial examination of the formula erroneously intimates that a predictor variable with a near-zero beta weight does not add to the predictive efficacy of the model.

Of course, such a view falls apart beyond a cursory examination. For example, a predictor (e.g.,  $X_p$ ) may have a large absolute correlation with  $Y$  but have a zero  $\beta$  weight, if one or more other correlated predictors are assigned credit for that predictor's shared explanatory ability. Indeed, in some cases a predictor with near-zero beta weight may be a very good predictor or even the single best predictor (e.g., Thompson & Borrello, 1985).

Furthermore, many researchers making such misinterpretations also fail to recognize that this  $\beta$ -weight-focused interpretation strategy is *context dependent* on having an *exactly* correctly specified model, because adding or deleting a single predictor could radically alter all the weights and thus all the interpretations resulting from them (e.g., Thompson, 1999b). In the words of Dunlap and Landis (1998), "The size of the regression weight depends on the other predictor variables included in the equation and is, therefore, prone to change across situations involving different combinations of predictors" (p. 398).

But as Pedhazur (1982) has noted, "The rub, however, is that the true model is seldom, if ever, known" (p. 229). And as Duncan (1975) has noted, "Indeed it would require no elaborate sophistry to show that we will never have the 'right' model in any absolute sense" (p. 101).

For these and other reasons, some researchers have suggested that structure coefficients (or alternatively  $r_s$  of predictors with  $Y$ ; see Pedhazur, 1997, pp. 899-900) must be interpreted in conjunction with the standardized weights when predictors are correlated (see Cooley & Lohnes, 1971, p. 55; Darlington, 1968; Thompson, 1997b; Thompson & Borrello, 1985; Thorndike, 1978, pp. 170-172). In the regression case, a structure coefficient ( $r_s$ ) is the bivariate correlation between a given predictor variable and the synthetic variable, predicted  $Y$  or  $\hat{Y}$ .

#### *Two Coefficients Should Be Interpreted*

However, interpreting only structure coefficients would be just as erroneous as interpreting only beta weights. A superficial examination of Equation 1

might incorrectly intimate that predictor  $X_p$  with a zero correlation with  $Y$  cannot contribute to the  $R^2$  effect size, because for this variable the product term in Equation 1,  $\beta_p (r_{YX_p})$ , equals zero regardless of what the predictor's  $\beta$  weight is.

Yet a variable may have a zero correlation with  $Y$  but a sizeable non-zero  $\beta$  weight. This does affect  $R^2$  by allowing the  $\beta$  weights for other predictors to deviate further from zero than the boundary of their respective zero-order correlations with  $Y$ , thus making some of the other product terms within the equation larger, and thus making  $R^2$  larger. This is exactly what happens in the classic "suppressor variable" case described by Horst (1966) based on pilot training data from World War II (see also Henard, 1998; Lancaster, 1999; Stevens, 1996, pp. 106-107; Woolley, 1997).

In this classic example, notwithstanding the fact that verbal ability was uncorrelated with pilot ability, using verbal ability scores in the regression equation to predict pilot ability actually served to remove the contaminating influence of verbal ability from the other predictors, which effectively increased the  $R^2$  value from what it would have been if only mechanical and spatial abilities were used as predictors. As Horst (1966) noted,

To include the verbal score with a negative weight served to suppress or subtract irrelevant ability, and to discount the scores of those who did well on the test simply because of their verbal ability rather than because of abilities required for success in pilot training. (p. 355)

### *Regression Purposes and Three Heuristic Examples*

Stevens (1996) and others were careful to distinguish two distinct applications of multiple regression: prediction versus explanation (or theory testing). Huberty and Petoskey (1999) discussed the distinctions at some length. In a pure prediction case we usually have (a) a group of people for whom data on a criterion variable and several predictors are available and (b) a group of people for whom the same predictors are available, but the criterion variable is not known or has not yet occurred. We can derive a prediction equation (rule or factor) from the first group and—if (and only if) (a) the people in the two groups are reasonably similar and (b) the regression equation works fairly well in the first group—then we can reasonably apply the weights from the first group's equation in the second group to make predictions of the absent criterion scores within the second group.

When our research application is purely predictive, in a sense interpretation may be irrelevant. We may very much desire an accurate prediction, but we may not care why our predictive rule works. For example, parents may care very much to know that they can obtain a very accurate prediction of the adult height of their 2-year-old children ( $\hat{Y}_i$  as a prediction of  $Y_i$  for each of the  $i$ th children) using the rule  $\hat{Y}_i = 0.0 + (2 \times X_i)$ , where  $X_i$  is the height of the chil-

dren at age 2. As parents, we may not care why this rule or equation works so well, as long as it works. Thus, interpretation may be less relevant in regression prediction applications.

But in theory testing or explanation applications, interpretation is very relevant. We want to know how useful the variables are. We may have theory suggesting that some variables should be important in one or more senses in the model, and theory that other variables should not be useful in any sense in the model. Here we may then need to examine regression beta weights, but, it will be argued, we will not (generally) want to interpret only regression beta weights.

Four cases will be distinguished here. Case 1 involves a single predictor variable. In this case, the multiplicative weight for predicting  $Z_Y$  using the scores on  $Z_{X_1}$  is  $r_{X_1Y}$  (i.e.,  $\hat{Y} = \beta(Z_{X_1})$ , and  $\beta_{X_1} = r_{X_1Y}$ ). Case 1 is actually a special case of Case 2—uncorrelated predictors, because with only one predictor there can be no nonzero correlations among the predictors. We will therefore briefly illustrate Case 1 using Case 2, and illustrate Cases 3 and 4 as well, all using the example of multiple regression using two predictor variables.

*Case 2: Uncorrelated predictors.* Table 1 presents the various calculations associated with Case 2. In Case 2, the predictors are perfectly uncorrelated with each other, and the denominator in each beta calculation equals 1.0 because  $1.0 - r_{X_1X_2}^2 = 1.0 - 0.0$ . Furthermore, the numerator always simplifies to equal the correlation of  $Y$  with each predictor because the right side of the numerator equals 0.0 (e.g.,  $r_{YX_2} [r_{X_1X_2}] = 0.7071[0.0]$ ). Finally, in Case 2,  $R^2$  equals the sum of the  $r^2$  values of each predictor with  $Y$ , because

$$R^2 = \beta_1 (r_{YX_1}) + \beta_2 (r_{YX_2}) + \dots \beta_p (r_{YX_p}),$$

which in this example equals

$$(r_{YX_1})(r_{YX_1}) + (r_{YX_2})(r_{YX_2}),$$

or

$$R^2 = (r_{YX_1})^2 + (r_{YX_2})^2.$$

Thus, when predictors are uncorrelated, each predictor's so-called standardized regression coefficient equals each predictor's correlation with  $Y$ . Furthermore, because the structure coefficient for a given predictor equals the correlation of the predictor with  $Y$  divided by the multiple correlation  $R$ , in this case the beta weights, the structure coefficients for the predictors and the correlations of the predictors with  $Y$  will all rank order the predictors identically, except that the structure coefficients will be scaled in a different metric (unless  $R = 1.0$ ).

Table 1  
*Regression Analyses for Three Cases*

Case 2				Case 3				Case 4			
Correlation matrices											
	$X_2$	$Y$	$\hat{Y}$		$X_2$	$Y$	$\hat{Y}$		$X_2$	$Y$	$\hat{Y}$
$X_1$	.0000	.7071	.7071	$X_1$	.9990	.7071	.9998	$X_1$	-.7071	.7071	.7071
$X_2$		.7071	.7071	$X_2$		.7060	.9983	$X_2$		.0000	.0000
$Y$			1.0000	$Y$			.7072	$Y$			1.0000
$\beta$ weight calculations											
$b_1 = \frac{r_{YX_1} \pm (r_{YX_2})(r_{X_1X_2})}{1.0 \pm (r_{X_1X_2})^2}$											
$= \frac{.7071 \pm (.7071)(.0000)}{1.0 \pm .0000^2}$				$= \frac{.7071 \pm (.7060)(.9990)}{1.0 \pm .9990^2}$				$= \frac{.7071 \pm (.0000)(\pm .7071)}{1.0 \pm (-.7071)^2}$			
$= \frac{.7071 \pm .00007071 \pm .7053}{1.0 \pm .0000 \quad 1.0 \pm .9980}$				$= \frac{.7071 \pm .0000}{1.0 \pm .5000}$				$= \frac{.7071}{.5000}$			
$= \frac{.7071}{1.0}$				$= \frac{.0018}{.0020}$				$= 1.4142$			
$= .7071$				$= .9068$							
$b_2 = \frac{r_{YX_2} \pm (r_{YX_1})(r_{X_1X_2})}{1.0 \pm (r_{X_1X_2})^2}$											

$$\begin{aligned}
&= \frac{.7071 \pm (.7071)(.0000)}{1.0 \pm 0.0000^2} \\
&= \frac{.7071 \pm 0.0000}{1.0 \pm 0.0000} \\
&= \frac{.7071}{1.0} \\
&= .7071
\end{aligned}$$

$$\begin{aligned}
&= \frac{.7060 \pm (.7071)(.9990)}{1.0 \pm 0.9990^2} \\
&= \frac{.7060 \pm 0.7064}{1.0 \pm 0.9980} \\
&= \frac{-.0004}{.0020} \\
&= -.1999
\end{aligned}$$

$$\begin{aligned}
&= \frac{.0000 \pm (.7071)(\pm 0.7071)}{1.0 \pm (-.7071)^2} \\
&= \frac{.0000 \pm 0.5000}{1.0 \pm 0.5000} \\
&= \frac{.5000}{.5000} \\
&= 1.0000
\end{aligned}$$

$R^2$  calculations

$$R^2 = \beta_1(r_{YX_1}) + \beta_2(r_{YX_2})$$

$$\begin{aligned}
&= .7071(.7071) + .7071(.7071) \\
&= .5000 + .5000 \\
&= 1.0000
\end{aligned}$$

$$\begin{aligned}
&= .9068(.7071) + -.1999(.7060) \\
&= .6412 - .1412 \\
&= .5001
\end{aligned}$$

$$\begin{aligned}
&= 1.4142(.7071) + 1.0000(.0000) \\
&= 1.0000 + .0000 \\
&= 1.0000
\end{aligned}$$

---

*Note.* Calculations were computed to six decimal places, but are rounded here to four decimal places. The bivariate correlation between a predictor variable and the composite variable,  $\hat{Y}$  (where  $\hat{Y} = [\beta_1 Z_{X_1}] + [\beta_2 Z_{X_2}]$ ), is the structure coefficient for that predictor variable. The bivariate correlation between  $Y$  and  $\hat{Y}$  is also  $R_{(YX_1X_2)}$ .

*Case 3: Correlated predictors.* When the predictor variables are correlated (i.e., are collinear or multicollinear), the beta weight for a given predictor no longer equals the correlation of that predictor with  $Y$ . Instead, the beta weight computations take into account all the pairwise correlations of the observed variables with each other.

This is done so that portions of  $Y$  variance that are redundantly explained by two or more predictors will not be multiply counted as explained. Thus, as Table 1 illustrates for Case 2, although both predictors in the heuristic example explain about half the variance in the  $Y$  scores, because the two predictors were almost perfectly correlated with each other ( $r = .999$ ), together the two predictors still explain little more than half ( $R^2 = .5001$ ) of the variance in the  $Y$  scores.

The example also illustrates the folly of interpreting beta weights as if they are correlation coefficients; this misinterpretation is unfortunately all too common. For example,  $X_2$  in the example has a beta weight of  $-.1999$ , even though  $r_{YX_2} = 0.7060$ .

*Case 4: Suppressor variables.* Table 1 illustrates a dramatic example of suppressor effects. Here,  $X_2$  has a zero correlation with  $Y$ ,  $X_1$  explains only 50% ( $-.7071^2$ ) of the variance in  $Y$ , and yet together the two predictors explain 100% ( $R^2 = 1.00$ ) of the variance in  $Y$ . Again, the folly of interpreting beta weights as correlation coefficients is demonstrated, in that (a)  $\beta_1$  is greater than the maximum value of  $r$  and (b)  $\beta_2 = 1.00$  when  $r_{YX_2} = 0.0$ .

### Structure Coefficients Within the GLM

Throughout the general linear model (GLM), structure coefficients are bivariate correlation coefficients between a given measured/observer variable and a latent/synthetic variable. For example, in multiple regression, the structure coefficient for predictor  $X_1$  is the correlation between the scores of  $n$  people on  $X_1$  with the same  $n$  people's scores on the predicted outcome variable,  $\hat{Y}$  (Cooley & Lohnes, 1971).

Similarly, in either exploratory or confirmatory factor analysis, the structure coefficient for measured variable  $X_1$  on Factor I is the correlation between the scores of  $n$  people on  $X_1$  with the same  $n$  people's factor scores on Factor I (Wells, 1999). And in canonical correlation analysis, for example, the structure coefficient of measured criterion variable  $X_1$  on Function I is the correlation between the scores of  $n$  people on  $X_1$  with the same  $n$  people's criterion-variable composite scores on Function I (Thompson, 1984, 2000).



*Emphasis Elsewhere in the General Linear Model*

Huberty (1994) has noted that

if a researcher is convinced that the use of structure  $rs$  makes sense in, say, a canonical correlation context, he or she would also advocate the use of structure  $rs$  in the contexts of multiple correlation, common factor analysis, and descriptive discriminant analysis. (p. 263)

For example, principal components analysis is actually an implicit part of canonical correlation analysis (CCA) and all the parametric methods subsumed by CCA (Thompson, 1984, pp. 11-16). Regarding exploratory component and factor analysis, Gorsuch (1983) emphasized that a “*basic* [italics added] matrix for interpreting the factors is the factor structure” (p. 207).

Regarding confirmatory factor analysis, Thompson (1997b) and others (e.g., Bentler & Yuan, 2000) have emphasized that when factors are correlated, the measured variables have nonzero structure coefficients even with the factors on which pattern coefficients have been fixed to be zero, and that these structure coefficients *must* be consulted to arrive at correct interpretations. Similarly, as regards descriptive discriminant analysis, Huberty (1994) noted that “construct definition and structure dimension [and not hit rates] constitute the *focus* [italics added] of a descriptive discriminant analysis” (p. 206).

Again, most researchers agree that the interpretation of structure coefficients is essential to understanding canonical results. As Meredith (1964) suggested, “If the variables within each set are moderately intercorrelated the possibility of interpreting the canonical variates by inspection of the appropriate regression weights [function coefficients] is practically nil” (p. 55). Levine (1977) was even more emphatic:

I specifically say that one *has* to do this [interpret structure coefficients] since I firmly believe as long as one wants information about the nature of the canonical correlation relationship, not merely the computation of the [synthetic function] scores, one must have the structure matrix. (p. 20)

And Cohen and Cohen (1983) observed that “interpretation of a given canonical variate is best undertaken by means of the structure coefficients, which are simply the (zero-order) correlations of that variate with its constituent variables (as was  $r_{\hat{Y}_i}$  in MRC)” (p. 456).

*Computation of Regression Structure Coefficients*

Regression structure coefficients for the regression case can be computed by estimating the  $\hat{Y}$  scores and then requesting the correlations of the predic-

tor variables with these scores. However, logistically it may be easier to compute regression coefficients simply by dividing a given  $r$  between predictor scores and the  $Y$  scores by the multiple correlation coefficient,  $R$ . For example, the structure coefficient for predictor  $X_1$  can be computed as:

$$r_s = r_{YX_1} / R.$$

### *An Interpretation Alternative*

Pedhazur (1997) noted that “such coefficients are simply zero-order correlations of independent variables with the dependent variable divided by a constant, namely, the multiple correlation coefficient. Hence, the zero-order correlations provide the same information” (p. 899). But Thompson and Borrello (1985) argued, “However, it must be noted that interpretation of only the bivariate correlations seems counterintuitive. It appears inconsistent to first declare interest in an omnibus system of variables taken only two at a time” (p. 208).

In a recent American Educational Research Association (AERA) invited address, Thompson (1999a) stated, “The reason that structure coefficients are called ‘structure’ coefficients is that these coefficients provide insight regarding what is the nature or structure of the underlying synthetic variables of the actual research focus” (p. 15). This view of structure coefficients certainly suggests the interpretive importance of these coefficients.

However, Pedhazur (1997) has argued that “because one may obtain large structure coefficients even when results [i.e.,  $R^2$ ] are meaningless, their use in such instances may lead to misinterpretations” (p. 899). He then presented a hypothetical data set involving an  $R^2$  of .00041, for which the  $r_s$  for the first predictor variable was .988. Pedhazur then said, “These are impressive coefficients, particularly the first one. . . . But what is not apparent from an examination of these coefficients is that they were obtained from meaningless results” (p. 899).

This objection seems unusual. As Thompson (1997a) explained,

All analyses are part of one general linear model. . . . When interpreting results in the context of this model, researchers should generally approach the analysis hierarchically, by asking two questions:

—Do I have anything? (Researchers decide this question by looking at some combination of statistical significance tests, effect sizes . . . and replicability evidence.)

—If I have something, where do my effects originate? (Researchers often consult both the standardized weights implicit in all analyses and structure coefficients to decide this question.) (p. 31)

As Pedhazur himself acknowledged (1997, p. 899) regarding other GLM analyses, such as descriptive discriminant and canonical correlation analy-

ses, one would only bother to examine the structure coefficients after one has determined that the results are noteworthy.

So, given this hierarchical contingency-based approach to the interpretation of all GLM results, including regression, this criticism of Pedhazur (1997) seems irrelevant. In short, in a regression in which noteworthy effects have been isolated, and only then, one ought to interpret either the standardized weights and the correlations of the predictors with *Y* or the standardized weights and structure coefficients.

Note that neither perspective (i.e., weights vs. structure coefficients) is inherently superior or correct. Only the use of both sets of coefficients presents the full dynamics of the data when predictors are correlated, as is commonly expected in behavioral research. For example, a near-zero weight with a large squared structure coefficient indicates that a predictor might have been useful in a prediction, but that the shared predictive power of that predictor was arbitrarily (i.e., not wrongly, just arbitrarily) assigned to another predictor. Conversely, when a predictor has a large absolute beta weight but a near-zero structure coefficient, a suppressor effect is indicated, as discussed previously (e.g., Horst, 1966).

### Sample

The studies we examined were collected from the *Journal of Applied Psychology* from 1987 (Volume 77) to 1998 (Volume 84). This is an APA I/O psychology journal in which regression is used with some frequency. To be considered, the articles had to use multiple linear regression to analyze the data. The articles also had to present a correlation matrix that included all the variables used in each analysis, so that reanalyses were readily possible. Finally, the article authors had to have deemed their results sufficiently noteworthy that they interpreted their effects, so that these interpretations could be contrasted with alternative interpretations that also consulted structure coefficients. Thirty-one articles met these criteria.

### Results

In the 31 articles that met the study's three inclusion criteria, there were 110 regression analyses performed. In all of these analyses, the authors only interpreted standardized weights as opposed to either (a) standardized weights and structure coefficients or (b) standardized weights and correlations between the predictors and the outcomes. Of course, the authors reported the correlation matrices, to be included in our examination, but these authors did not consult either these coefficients or the structure coefficients when evaluating the import of the predictors.

Of the 110 analyses, 103 (94%) contained at least one discrepancy between the standardized weights and structure coefficients as regards the

rank orderings of the predictive powers of the predictor variables. The standardized weights can be interpreted to evaluate the importance of the predictors in the context-specific setting, in which it is presumed that the model is exactly correctly specified; the weights tolerate no redundancy in credit for shared predictive power when the predictors are correlated. Structure coefficients, on the other hand, evaluate which predictors do or could produce the predicted outcome scores. When predictors are correlated, different interpretations may arise from the two perspectives, both of which have interpretative value.

In 37 (34%) of the 110 analyses, the beta weights failed to identify the single-best predictor variable. In 90 (82%) analyses, either the first, second, or third single-best predictor variable was not identified. Finally, in 77 (70%) of the analyses, the interpreted beta weights did not identify the single-worst predictor, which is extremely important if a researcher uses the beta weights to reduce the number of predictor variables applied in a model.

Of course, the interpretive pictures painted by the weights are not intrinsically wrong, any more than the picture painted by the structure coefficients is intrinsically correct. However, beta weights are affected by the presence or the absence of any other predictor in the model, and the interpretations arising from these weights are context specific and presume that the model is exactly correctly specified. And we will never distinguish between suppressor effects and the direct predictive power of a predictor if we limit our interpretations solely to the examination of standardized weights. We turn now to some illustrative differences in interpretations arising when a fuller set of results is considered.

### *Illustrative Results*

Space restrictions preclude complete presentation of all examples here. Suffice it to say that some gross misinterpretations or incomplete interpretations of regression results occurred in the articles that we studied.

Some examples may convey the general tenor of these problems. Because as the number of predictor variables increases, the likelihood of multicollinearity increases and so then does the opportunity for discrepancies in interpretations arising from beta weights as against structure coefficients, the examples are categorized by the number of predictor variables. However, prior to turning to these examples, a common misinterpretation is briefly noted.

*Misinterpretation of beta weights as measuring relationship.* It must be remembered that the correlations between predictors and the  $Y$  scores and between predictors and the  $\hat{Y}$  scores (i.e.,  $r_s$ ) are correlation coefficients. That is, the results are bounded by  $-1$  to  $+1$ , measure relationship, and have signs reflecting the pattern (direct or inverse) of the relationship.

Beta weights, on the other hand, do not measure relationship. They do not have universal statistical boundaries (i.e.,  $-1$  to  $+1$ ). The weights, for example, can be positive when the predictor's relationship with the criterion is negative, as occurred in the Case 2 example presented in Table 1. Or, the weights can be large and nonzero when the correlations of the predictors with the  $Y$  scores are zero, as occurred in the Case 3 example in Table 1.

Clearly, beta weights should not be interpreted as measuring relationship! A beta weight evaluates, given one unit of change (e.g., increase) in  $Z_{x_1}$ , how much  $\hat{Y}$  will change. For example, if  $X_1$  (and thus  $Z_{x_1}$ ) is perfectly uncorrelated with  $Y$  and has a beta weight of  $-2.0$ , this means that if a person's score was higher on  $Z_{x_1}$  by one unit,  $\hat{Y}$  would be lower by two units. This is important to evaluate as part of interpretation, but does not evaluate the statistical issue of correlation! Indeed, a predictor may have a zero correlation with  $Y$  but have the largest absolute value of  $\beta$  for that model.

*Five or fewer predictors.* In a study of organizational citizenship behavior, Podsakoff, Ahearne, and MacKenzie (1997) noted that the

data reported in this table indicate that both sportsmanship (standardized  $b = .393, p < .05$ ) and helping behavior (standardized  $b = .397, p < .05$ ) had significant positive relationships [*sic*] with the quantity of output and accounted for about a quarter of the variance (25.7%) in this criterion variable. The data also indicated that helping behavior was negatively related [*sic*] (standardized  $b = -.424, p < .05$ ) to the percentage of paper produced that was rejected. . . . Civic virtue was not found to be related [*sic*] to either the quantity or quality of output, and sportsmanship was not related [*sic*] to the quality of output. (p. 266)

However, in reanalysis, the structure coefficients indicated that helping behavior and civic virtue were the best predictors of quantity of output, with both having positive relationships with quality, as opposed to negative and no relationship, respectively.

Maslyn and Fedor (1998) examined the relevance of measuring different foci in politics. The authors reported that

LMX and participant age were positively related to organizational commitment. In contrast, group-focused politics were negatively associated with organizational commitment.

Turnover intentions also were significantly predicted by the set of control variables, accounting for 33% of the variance. In this case, LMX and participant age were both negatively related [*sic*] to turnover intentions, whereas the group-focused perceptions of politics were not predictive of turnover intentions. (pp. 650-651)

The structure coefficients in this reanalysis indicated that LMX was the best predictor ( $r_s = -.829$ ). Although group focus had the most near-zero beta weight in predicting turnover intentions (reported as  $-.00$ ), which the authors

noted, the structure coefficients ( $r_s = .542$ ) indicated that this variable indeed had sizeable predictive ability.

*Six to 10 predictors.* In an article on executive recognition and the write-off of problem loans, Staw, Barsade, and Koput (1997) reported that

regression analyses again demonstrated that the relative turnover of top managers at  $T - 1$  significantly predicted both adjusted provision for loan loss ( $B = .003, p < .01$ ) and adjusted net loan loss at Time  $T$  ( $B = .0046, p < .001$ ). The relative turnover of other senior managers showed similar effects ( $B = .0041, p < .001$ , for provisions; and  $B = .0023, p < .005$ , for write-offs). However, once again, turnover of outside board members did not predict either provision for loan loss or net loan loss ( $B = -.0023, ns$ , and  $B = -.0024, ns$ , respectively). (p. 137)

The structure coefficients suggest a different picture. Contrary to the authors' findings, for adjusted provision for loan loss and adjusted loan loss, the single-most important predictor was the relative turnover of outside board members (i.e., bank outside directors), and the relative turnover of top managers (i.e., bank presidents, chief executive officers, and chairs) was the least important predictor of all three.

Furthermore, Staw et al. (1997) observed that "turnover in banks' operating management was significantly associated with the way banks dealt with problem loans" (p. 138). However, the importance of structure coefficients is demonstrated by the finding that, from this alternate perspective, it was actually the turnover of outside management, not the turnover of the banks dealing with problem loans, that was the single-best predictor.

Vliert, Euwema, and Huismans (1995) found that "as superiors, the sergeants treated their subordinate more effectively if they removed forcing ( $\beta = -.58, p < .01$ ) or added precess controlling ( $\beta = .46, p < .01$ ) or accommodating ( $\beta = .21, p < .05$ )" (p. 276). However, the unreported structure coefficients indicated that the sergeants treated their subordinates more effectively if they added problem solving ( $r_s^2 = .743$ ) first and foremost, whereas the third-most effective technique was to remove avoiding ( $r_s^2 = .411$ ).

Similarly, Wanberg, Watt, and Rumsey (1996) maintained that

our multiple regression analyses found conscientiousness and job-seeking support to be significant, positive predictors of job-seeking frequency and job-seeking intention. One additional variable, gender, was also a significant predictor of job-seeking intention, with women being more likely than men to have future intentions of looking hard for work. (p. 83)

This result was reported to be at odds with previous results. However, looking at the structure coefficients, the rank of the structure coefficient for gender was seventh, not second, as the beta weights suggested.

In their findings on perceived equity, motivation, and final-offer arbitration in major league baseball, Bertz and Thomas (1992) stated that

multiple ordinary least squares regression indicated that arbitration outcome significantly predicted subsequent performance. The pre-arbitration performance variables were the most significant predictors of subsequent performance. As hypothesized, the coefficient on lost arbitration was negative and significant, suggesting that losing arbitration had detrimental effect on subsequent performance. (p. 283)

After computing the structure coefficients for this model, it was found that losing arbitration was found to have no effect on the player performance ( $r_s = -.086$ ). That is, losing arbitration was an unnoticed suppressor variable.

In their study of the influence of structural features of local unions on members' unions commitment, Mellor, Mathieu, and Swim (1994) reported that "only age ( $\beta = -.14, p < .05$ ) and union-family conflict ( $\beta = -.42, p < .01$ ) evidenced significant linear effects" (p. 206). However, the structure coefficients indicated something different. Union-family conflict was the single-most noteworthy predictor of union commitment. Furthermore, age was actually only the ninth-best single predictor of union commitment.

*Eleven to 15 predictor variables.* In studying the roles of job enrichment and other organization interventions on self-efficacy, Parker (1998) noted that "decision-making influence (the second measure of job enrichment) did not make a significant independent contribution to the regression equation" (p. 842). However, the structure coefficients indicated that, of the organizational variables, decision-making influence was the second-best predictor of role breadth self-efficacy.

Parker (1998) also indicated that "it is relevant to observe that both self-esteem and proactive personality were significant predictors of RBSE ( $\beta = .11, p < .01$ , and  $\beta = .24, p < .001$ , respectively), suggesting that these personality factors are associated with self-efficacy" (p. 842). However, the structure coefficients indicated that although proactivity was the second-best predictor of RBSE, self-esteem was only the sixth best.

*Sixteen or more predictor variables.* In explaining the results of their study of substance abuse and on-the-job behaviors, Lehman and Simpson (1992) reported that on the Psychological Withdrawal Behaviors scale, "Alcohol use, lifetime drug use, and substance use at work each had significant  $b$  weights in the full regression. Other important predictors of psychological withdrawal behaviors included age, education, self-esteem, depression, tenure with city, job involvement, job satisfaction, organizational commitment, and power" (pp. 315-316). However, surveying the structure coefficients, the ordering of the variables predicting Psychological With-

drawal Behaviors was different from the ordering associated with the beta weights. The eight most important predictor variables, according to the structure coefficients, were organizational commitment ( $r_s = -.740$ ), job satisfaction ( $r_s = -.606$ ), job involvement ( $r_s = -.564$ ), self-esteem ( $r_s = -.428$ ), drug use ( $r_s = .428$ ), faith in management ( $r_s = -.402$ ), loyalty ( $r_s = -.351$ ), and age ( $r_s = -.337$ ).

For Physical Withdrawal Behaviors, Lehman and Simpson (1992) noted that "the strongest individual predictors were positive affect, pay level, power, and lifetime drug use. Substance use at work and alcohol use also had significant  $b$  weights" (p. 316). However, looking at the structure coefficients, the strongest single predictor variables were recent drug use ( $r_s = .551$ ), drug use ( $r_s = .416$ ), gender ( $r_s = -.409$ ), substance use at work ( $r_s = .409$ ), self-esteem ( $r_s = .372$ ), and age ( $r_s = .326$ ).

For Antagonistic Work Behaviors, Lehman and Simpson (1992) noted that "the strongest individual predictors were being White, faith in management, job involvement, job satisfaction, and loyalty. Substance use at work was the only substance use variable to have a significant  $b$  weight" (p. 316). However, the structure coefficients suggested that the most important predictor variables were faith in management ( $r_s = -.668$ ), job satisfaction ( $r_s = -.524$ ), loyalty ( $r_s = -.479$ ), organizational commitment ( $r_s = -.325$ ), and substance abuse at work ( $r_s = .322$ ).

Finally, in Tannenbaum, Mathieu, Salas, and Cannon-Bowers's (1991) article on the influence of training fulfillment on the development of commitment, self-efficacy, and motivation, they indicated that training fulfillment was positively related to training motivation, but the structure coefficients indicated that training fulfillment was actually negatively related to training motivations. Similarly, they also mentioned that inspections were negatively related to physical self-efficacy, but the structure coefficients indicated that there was actually a positive relationship between these two variables.

## Discussion

In most cases, regression researchers ought to interpret  $\beta$  weights and structure coefficients (or else bivariate correlations of predictors with the criterion) once a noteworthy omnibus effect is detected. Our study demonstrates that this is not merely some pedantic statistical concern. Researchers may misinterpret or incompletely interpret their regression results by consulting only selected aspects of their analyses.

The finding that so few regression researchers in the articles we studied consulted structure coefficients (i.e., none) is not atypical as regards the practices in other journals (Burdenski, in press). For example, Bowling (1993) reported that the *Journal of Counseling Psychology* published 20 articles that used multiple regression analysis between January 1990 and April 1993, but



that authors of only 3 studies reported structure coefficients in their results and only a few provided a correlation matrix that would allow an ambitious researcher to derive them post facto. In this vein, Dunlap and Landis (1998) noted that although structure coefficients “are invariably computed for canonical correlations by modern statistical software, they are never reported for multiple regression analysis” (p. 398).

We emphasize again that we do not advocate ignoring the regression beta weights. However, we must remember that these weights and the interpretations arising from them are context specific. The confidence we vest in interpretations of the weights hinges on our certainty that our model is exactly correctly specified. The weights can *all* change *dramatically* with the addition or the deletion of a *single* predictor.

It can be useful to also consult regression structure coefficients or the correlations of the predictors with  $Y$  to obtain another perspective on dynamics within our data. This consultation may yield the insight that a predictor with a near-zero beta weight actually was the single-best predictor. Or we may discover that the predictor is a suppressor that improves the model  $R^2$  not by directly predicting  $Y$  but indirectly doing so by removing extraneous variance from other predictors.

Of course, when predictors are perfectly uncorrelated, both sets of coefficients will yield identical interpretations, because  $\beta$  in this case for a predictor will equal  $r_{YX}$ , and because  $r_s$  equals  $r_{YX} / R$ , so the two sets of coefficients will merely be scaled differently. But, as our review showed, in many articles predictors are often correlated, just as they often are in the reality being studied.

When interpreting regression results, once noteworthy effects have been detected it may be best to consult the full system of results, just as we routinely would in applications of other members of the general linear model analytic family. The two sets of coefficients— $\beta$  weights and structure coefficients—provide us with a more insightful stereoscopic view of dynamics within our data. Interpreting only beta weights, once noteworthy omnibus effects have been isolated, usually will not yield sufficient understanding of all the relevant dynamics within our data. Other results may also augment interpretation (e.g., Johnson, 2000).

As Cohen and Cohen (1983) argued, “It is also important to keep in mind the zero-order correlations of  $X_i$  with  $Y$  (and hence with  $\hat{Y}$ )” (p. 113). Of course, Ezekiel’s (1930) old admonition remains prescient:

Except insofar as the effort to reduce the variables to specific numerical statement, definitely related, forces the investigator to think more clearly and definitely about his problem, statistical analysis is not a substitute for logical analysis, clear-cut thinking, and full knowledge of the problem. (p. 351)

## References

- American Psychological Association. (1994). *Publication manual of the American Psychological Association* (4th ed.). Washington, DC: Author.
- Bagozzi, R. P., Fornell, C., & Larcker, D. F. (1981). Canonical correlation analysis as a special case of a structural relations model. *Multivariate Behavioral Research*, 16, 437-454.
- Bentler, P. M., & Yuan, K.-H. (2000). On adding a mean structure to a covariance structure model. *Educational and Psychological Measurement*, 60, 326-339.
- Bertz, R., Jr., & Thomas, S. (1992). Perceived equity, motivation, and final-offer arbitration in major league baseball. *Journal of Applied Psychology*, 77, 280-287.
- Bowling, J. (1993, November). *The importance of structure coefficients as against beta weights: Comments with examples from the counseling psychology literature*. Paper presented at the annual meeting of the Mid-South Education Research Association, New Orleans. (ERIC Document Reproduction Service No. ED 364 606)
- Burdenski, T. K., Jr. (in press). The importance of structure coefficients in multiple regression: A review with examples from published literature. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 6). Stamford, CT: JAI.
- Cohen, J. (1968). Multiple regression as a general data-analytic system. *Psychological Bulletin*, 70, 426-443.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Cooley, W. W., & Lohnes, P. R. (1971). *Multivariate data analysis*. New York: Wiley.
- Darlington, R. B. (1968). Multiple regression in psychological research and practice. *Psychological Bulletin*, 69, 161-182.
- Dawson, T. E. (1999). Relating variance partitioning in measurement analyses to the exact same process in substantive analyses. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 5, pp. 101-110). Stamford, CT: JAI.
- Duncan, O. D. (1975). *Introduction to structural equation models*. New York: Academic Press.
- Dunlap, W. P., & Landis, R. S. (1998). Interpretations of multiple regression borrowed from factor analysis and canonical correlation. *Journal of General Psychology*, 125, 397-407.
- Ezekiel, M. (1930). *Methods of correlational analysis*. New York: Wiley.
- Fan, X. (1996). Canonical correlation analysis as a general analytic model. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 4, pp. 71-94). Greenwich, CT: JAI.
- Fan, X. (1997). Canonical correlation analysis and structural equation modeling: What do they have in common? *Structural Equation Modeling*, 4, 65-79.
- Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Henard, D. H. (1998, January). *Suppressor variable effects: Toward understanding an elusive data dynamic*. Paper presented at the annual meeting of the Southwest Educational Research Association, Houston, TX. (ERIC Document Reproduction Service No. ED 416 215)
- Horst, P. (1966). *Psychological measurement and prediction*. Belmont, CA: Wadsworth.
- Huberty, C. J. (1994). *Applied discriminant analysis*. New York: Wiley and Sons.
- Huberty, C. J., & Petoskey, M. D. (1999). Use of multiple correlation analysis and multiple regression analysis. *Journal of Vocational Education Research*, 24(1), 15-43.
- Johnson, J. W. (2000). A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate Behavioral Research*, 35, 1-19.
- Knapp, T. R. (1978). Canonical correlation analysis: A general parametric significance testing system. *Psychological Bulletin*, 85, 410-416.
- Lancaster, B. P. (1999). Defining and interpreting suppressor effects: Advantages and limitations. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 5, pp. 139-148). Stamford, CT: JAI.
- Lehman, W., & Simpson, D. (1992). Employee substance use and on-the-job behaviors. *Journal of Applied Psychology*, 77, 309-321.

- Levine, M. S. (1977). *Canonical analysis and factor comparison*. Beverly Hills, CA: Sage.
- Maslyn, J., & Fedor, D. (1998). Perceptions of politics: Does measuring different foci matter? *Journal of Applied Psychology*, 84, 645-653.
- Mellor, S., Mathieu, J., & Swim, J. (1994). Cross-level analysis of the influence of local union structure on women's and men's union commitment. *Journal of Applied Psychology*, 79, 203-210.
- Meredith, W. (1964). Canonical correlations with fallible data. *Psychometrika*, 29, 55-65.
- Parker, S. (1998). Enhancing role breadth self-efficacy: The roles of job enrichment and other organizational interventions. *Journal of Applied Psychology*, 83, 835-852.
- Pedhazur, E. J. (1982). *Multiple regression in behavioral research: Explanation and prediction* (2nd ed.). New York: Holt, Rinehart & Winston.
- Pedhazur, E. J. (1997). *Multiple regression in behavioral research* (3rd ed.). Ft. Worth, TX: Harcourt Brace.
- Podsakoff, P., Ahearne, M., & MacKenzie, S. (1997). Organizational citizenship behavior and the quantity and quality of work group performance. *Journal of Applied Psychology*, 82, 262-270.
- Staw, B., Barsade, S., & Koput, K. (1997). Escalation at the credit window: A longitudinal study of bank executives' recognition and write-off of problem loans. *Journal of Applied Psychology*, 82, 130-142.
- Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Mahwah, NJ: Erlbaum.
- Tannenbaum, S., Mathieu, J., Salas, E., & Cannon-Bowers, J. (1991). Meeting trainees' expectations: The influence of training fulfillment on the development of commitment, self-efficacy, and motivation. *Journal of Applied Psychology*, 76, 759-769.
- Thompson, B. (1984). *Canonical correlation analysis: Uses and interpretation*. Newbury Park, CA: Sage.
- Thompson, B. (1991). A primer on the logic and use of canonical correlation analysis. *Measurement and Evaluation in Counseling and Development*, 24, 80-95.
- Thompson, B. (1995). Stepwise regression and stepwise discriminant analysis need not apply here: A guidelines editorial. *Educational and Psychological Measurement*, 55, 525-534.
- Thompson, B. (1997a). Editorial policies regarding statistical significance tests: Further comments. *Educational Researcher*, 26(5), 29-32.
- Thompson, B. (1997b). The importance of structure coefficients in structural equation modeling confirmatory factor analysis. *Educational and Psychological Measurement*, 57, 5-19.
- Thompson, B. (1999a, April). *Common methodology mistakes in educational research, revisited, along with a primer on both effect sizes and the bootstrap*. Invited address presented at the annual meeting of the American Educational Research Association, Montreal. (ERIC Document Reproduction Service No. ED 429 110)
- Thompson, B. (1999b). Five methodology errors in educational research: A pantheon of statistical significance and other faux pas. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 5, pp. 23-86). Stamford, CT: JAI.
- Thompson, B. (2000). Canonical correlation analysis. In L. Grimm & P. Yarnold (Eds.), *Reading and understanding more multivariate statistics* (pp. 285-316). Washington, DC: American Psychological Association.
- Thompson, B., & Borrello, G. M. (1985). The importance of structure coefficients in regression research. *Educational and Psychological Measurement*, 45, 203-209.
- Thompson, B., & Daniel, L. G. (1996). Seminal readings on reliability and validity: A "hit parade" bibliography. *Educational and Psychological Measurement*, 56, 741-745.
- Thorndike, R. M. (1978). *Correlational procedures for research*. New York: Gardner.
- Vliert, E., Euwema, M., & Huismans, S. (1995). Managing conflict with a subordinate or a superior: Effectiveness of conglomerated behavior. *Journal of Applied Psychology*, 80, 271-281.

- Wanberg, C., Watt, J., & Rumsey, D. (1996). Individuals without jobs: An empirical study of job-seeking behavior and reemployment. *Journal of Applied Psychology, 81*, 76-87.
- Wells, R. D. (1999). Factor scores and factor structure and communality coefficients. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 5, pp. 123-138). Stamford, CT: JAI.
- Wilkinson, L., & APA Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*, 594-604.
- Woolley, K. K. (1997, January). *How variables uncorrelated with the dependent variable can actually make excellent predictors: The important suppressor variable case*. Paper presented at the annual meeting of the Southwest Educational Research Association, Austin. (ERIC Document Reproduction Service No. ED 407 420)