Psychological Bulletin
1993, Vol. 114, No. 3, 542–551

0033-2909/93/$3.00

Dominance Analysis: A New Approach to the Problem of Relative Importance of Predictors in Multiple Regression

David V. Budescu

Whenever multiple regression is used to test and compare theoretically motivated models, it is of interest to determine the relative importance of the predictors. Specifically, researchers seek to rank order and scale variables in terms of their importance and to express global statistics of the model as a function of these measures. This article reviews the many meanings of importance of predictors in multiple regression, highlights their weaknesses, and proposes a new method for comparing variables: dominance analysis. Dominance is a qualitative relation defined in a pairwise fashion: One variable is said to dominate another if it is more useful than its competitor in all subset regressions. Properties of the newly proposed method are described and illustrated.

An important aspect of any multiple regression analysis is the determination of the relative importance of the various predictors. The quest for a single-valued, meaningful index of importance for each variable is almost as old as the model itself (Yule, 1899). However, recent articles by Kruskal (1987a, 1987b), Kruskal and Majors (1989), Pratt (1987), and Theil (1987) have vividly illustrated that no single "best" solution to this problem has emerged. The goal of this article is to critically review some of the standard measures of importance, point out their major problems, and propose a new method for determining the importance of variables.

There are numerous ways of using regression analysis in behavioral research (e.g., Hocking, 1976). I approached the problem from the point of view of researchers interested in the theoretical implications of specific models. For example, one may be trying to determine the relative importance of education and economic status in determining opinions (F. Williams & Mosteller, 1947); monetary and fiscal policy in determining gross national product; teacher quality, school effects, and peer quality as determinants of school achievement (Mood, 1969); personality traits and family cohesiveness as predictors of ability to cope effectively with stress; or genetic and environmental factors as predictors of intelligence.

The Problem

Consider a univariate multiple regression model in which, in a certain population, a single criterion, y, is described in terms of a linear combination of p predictors, $x_1, \ldots, x_p$, and a constant term ($x_0 = 1$):

$$y = \sum_{j=0}^{p} \beta_j x_j + e. \tag{1}$$

Under alternative versions of the model (e.g., Graybill, 1961), the predictors are assumed to be constant numerical values (the fixed model), random variables with a p-variate normal distribution (the random model), or a mixture of the two (the mixed model). In each case, the residual component is normally distributed with zero mean and fixed variance, σ^2. β_j ($j = 1 \cdots p$) are population parameters known as regression coefficients. It is often convenient to reexpress all of the variables included in the model in standardized form with zero mean and unit variance. Let Zy and Zx_j be standardized values of y and x_j ($j = 1, \ldots, p$). Then

$$Zy = \sum_{j=1}^{p} \beta_j^* Zx_j + e, \tag{2}$$

where β_j^* are the standardized regression coefficients.

Various notions of importance have been suggested in the literature. Regardless of the particular interpretation adopted, users typically attempt to rank order the p predictors in terms of their importance, scale the variables' importance (i.e., locate them along a continuum with interval or ratio properties), and relate the importance measures to certain global statistics such as the model's squared multiple correlation, $\rho_{y \cdot x_1 \cdots x_p}{}^2$, or the criterion expected value, μ_y. Sometimes, however, researchers have more specific and focused goals, such as identifying the "most important" predictor or determining whether x_i is more, or less, important than x_j ($i \neq j$).

Part of the confusion regarding the variables' importance may be attributed to the fact that in a special and simple case, most measures coincide to yield a unique solution. If all p predictors are uncorrelated, then

$$\rho_{y \cdot x_1 \cdots x_p}{}^2 = \sum_{j=1}^{p} \rho_{yx_j}{}^2 = \sum_{j=1}^{p} \beta_j^{*2}. \tag{3}$$

Thus, the squared zero-order correlations of the variables with

This article benefited from comments provided by numerous colleagues. I wish to thank Ayala Cohen, Yoav Cohen, Robyn Dawes, Steve Klepper, Amnon Rapoport, Benjamin Reiser, Yossi Shavit, three anonymous reviewers, and participants at quantitative psychology seminars at the University of North Carolina and the University of Illinois for their important insights and contributions.

Correspondence concerning this article should be addressed to David V. Budescu, Department of Psychology, University of Illinois, 603 East Daniel Street, Champaign, Illinois 61820. Electronic mail may be sent to DBUDESCU@S.PSYCH.UIUC.EDU.

the criterion (which, in this case, happen to coincide with the standardized coefficients) provide a meaningful scale along which one can rank and locate the variables' importance in the population. Furthermore, the sum of these measures of importance yields the population's squared multiple correlation.

Unfortunately, this simple and elegant result does not hold for any arbitrary pattern of correlations among the predictors. Although many of the regularly used measures of importance reduce to the decomposition of Equation 3 under the special circumstances just mentioned, they generalize this decomposition in different ways that imply different notions of importance. However, there is no compelling reason for preferring any one of these generalizations over the others.

Various Implied Meanings of Importance

This article does not include a detailed review of all techniques proposed to determine the relative importance of predictors in regression. Partial reviews are available in such standard references as Achen (1982); Darlington (1968, 1990); Cohen and Cohen (1983); Kruskal and Majors (1989); Lindeman, Merenda, and Gold (1980); and Pedhazur (1975, 1982). My goal is simply to show that these measures rely on different and, in my opinion, unsatisfactory notions of importance. To simplify matters, I present and discuss the various methods using population terms. Later I address the issue of sampling error.

It is sometimes suggested that the predictors' zero-order correlations with the criterion, ρ_{yx_j}, or squared values, $\rho^2_{yx_j}$, be used as measures of their importance (e.g., Darlington, 1968). This implies that a predictor's importance is defined by its unique and direct predictive ability, ignoring all other variables in the model.

It is well known that if a model is misspecified, either by omitting important variables or by including unnecessary ones, some of its attractive properties do not hold (e.g., Hocking, 1976). For example, the inclusion of unnecessary variables inflates the standard errors of the estimates, and the omission of predictors that belong in the model induces bias in the parameter estimates. Thus, another simple approach to the problem would be to define all variables in the correct model as equally important.

Undoubtedly, the most commonly invoked interpretation of importance is a conditional one. The predictor's importance depends on its contribution to the full model (Equation 1) conditional on the contribution of the other $p - 1$ variables. There are two classes of statistics in this category. The first includes all slope-based measures (Kruskal, 1984): regression coefficients, their standardized counterparts, the normalized standardized coefficients (sometimes referred to as direction cosines), and semistandardized coefficients (Stavig, 1977). These measures can be interpreted as the rate of change in the criterion as a function of a standard (unit) change in each predictor while the other predictors are being held constant. Achen (1982) advocated the product $\beta_j\mu_{x_j}$ as a measure of level importance. Its absolute value, $|\beta_j\mu_{x_j}|$, is proportional to the predictor's elasticity, a measure often used in econometrics. This measure offers a simple decomposition of the criterion's mean:

$$\mu_y = \sum_{j=1}^{p} \beta_j\mu_{x_j}.$$

The second class consists of the different variance reduction measures such as the squared partial correlation and the squared semipartial correlation, which is also known as the predictor's usefulness (Darlington, 1968). These correlational measures can be interpreted as proportions of the criterion's variance that can be reproduced by each predictor conditional on the other variables' contribution (e.g., Cohen & Cohen, 1983). They allow for sequential decompositions of $\rho_{y\cdot x_1\cdots x_p}{}^2$, the model's overall fit. In terms of partial correlations, there is a multiplicative decomposition:

$$(1 - \rho_{y\cdot x_1\cdots x_p}{}^2) = (1 - \rho_{y\cdot x_1}{}^2)(1 - \rho_{yx_2\cdot x_1}{}^2)\cdots (1 - \rho_{yx_p\cdot x_1\cdots x_{p-1}}{}^2) = \prod_{j=1}^{p} (1 - \rho_{yx_j\cdot x_1\cdots x_{j-1}}{}^2). \quad (4)$$

In terms of semipartial correlations, there is an additive decomposition:

$$\rho_{y\cdot x_1\cdots x_p}{}^2 = \rho_{y\cdot x_1}{}^2 + \rho_{y(x_2\cdot x_1)}{}^2 + \cdots + \rho_{y(x_p\cdot x_1\cdots x_{p-1})}{}^2 = \sum_{j=1}^{p} \rho^2_{y(x_j\cdot x_1\cdots x_{j-1})}. \quad (5)$$

For natural, or theoretically meaningful, orders (e.g., successive terms in a polynomial regression), Kruskal (1987a, 1987b) and Lindeman, Merenda, and Gold (1980) recommended these sequential partial and semipartial correlations, respectively, as measures of importance. However, they also pointed out that, in most cases, no such ordering exists. For these cases, they suggested averaging the sequential components (squared partial or semipartial correlations) across all $p!$ possible orderings of the predictors. Thus, a predictor's importance is defined as its average contribution to the fit of the model when all $p!$ orderings are considered.

Building on Kruskal's approach, Theil (1987) and Theil and Chung (1988) recommended measuring each variable's importance by the number of bits of information of the specific variable. Let $I(x) = -\log_2(1 - x)$ for $0 \le x \le 1$. Then Equation 4 can be reexpressed as the sum of p measures of information:

$$I(\rho_{y\cdot x_1\cdots x_p}{}^2) = I(\rho_{y\cdot x_1}{}^2) + I(\rho_{yx_2\cdot x_1}{}^2) + \cdots + I(\rho_{yx_p\cdot x_1\cdots x_{p-1}}{}^2) = \sum_{j=1}^{p} I(\rho_{yx_j\cdot x_1\cdots x_{j-1}}{}^2). \quad (6)$$

By averaging across all $p!$ orders, one obtains another scale of importance yielding a natural decomposition of the model's overall goodness (in this case, the total number of bits of information).

One of the best-known measures of importance is the product of the variable's direct (its correlation with the criterion) and total (its standardized regression coefficient) effects. This measure provides a decomposition of the model's fit.

$$\rho_{y\cdot x_1\cdots x_p}{}^2 = \sum_{j=1}^{p} \rho_{yx_j}\beta^*_j. \quad (7)$$

Although often criticized (e.g., Darlington, 1968) and ridiculed (Kruskal, 1984; Ward, 1969), the measure has been recommended and justified recently by Pratt (1987).

A factor-analytical approach to multiple regression suggests quantifying a predictor's importance as its loading on the predicted composite: its squared correlation with the predicted values. It is possible to show that

$$\rho_{x_i\hat{y}}{}^2 = \sum_{j=1}^{p} \beta_j^{*2}\rho_{x_jx_i}/\rho_{y\cdot x_1\cdots x_p}{}^2. \qquad (8)$$

Thus, each variable's importance is a weighted average of all p squared standardized regression coefficients. The weights are proportional to the squared correlations of the predictor of interest with all other variables.

Critique of the Various Measures of Importance

An examination of these measures indicates that most share certain problems that render them less than attractive to users and researchers. The most glaring weakness is that they all rely on too narrow, restrictive, and quite arbitrary definitions of importance that do not necessarily match the regular connotations attached to the term in substantive scientific applications. Concepts are often vague and ill defined, and it may be impossible to offer a precise universal definition of importance. I suspect, however, that most people have a broader understanding of importance than implied by any of these definitions, and I do not expect either of the definitions described earlier to provide fully satisfactory answers to questions such as the relative importance of home and school effects in academic achievement in the sense intended by educational or developmental researchers.

The various definitions do not necessarily agree with each other. This is not surprising if one keeps in mind that certain measures (e.g., the variable's direct effect) depend only on each predictor's relationship with the criterion, others (e.g., the conditional measures) depend on the whole system of $p + 1$ variables, and others (e.g., the average measures) are defined across all submodels. This gives rise to serious problems in interpretation of results and communication among users. Numerous illustrative examples in which various measures of importance rank and scale the predictors differently are available. Some are related to problems of collinearity (e.g., Cramer, 1974) and suppressor variables (e.g., Tzelgov & Henik, 1991), but the problem is not restricted to such pathological instances (e.g., Darlington, 1968, 1990; Mosteller & Tukey, 1977; E. J. Williams, 1978).

Most of the approaches reviewed earlier lead to conclusions that are model dependent and are not invariant to subset selection. Thus, it is possible that $\beta_1^* > \beta_2^*$ in the presence of x_3 and x_4 but that $\beta_1^* < \beta_2^*$ when x_3 or x_4 are excluded. Two researchers who are interested in the relative importance of x_1 and x_2 may reach different conclusions depending on the other predictors they include in their models. This situation reflects, in part, a true state of affairs induced by the nature of the relationships between the different predictors involved. However, it would be to everyone's advantage to have a method of determining importance that is invariant under all subset selections.

An assumption implicit in all of the approaches is that the p variables can always be ordered (in a weak sense) according to their importance. In fact, in all of the methods reviewed earlier indices are calculated separately for each predictor and importance is inferred from the order of these indices. This precludes the possibility that, for some variables, it might be impossible to determine an ordering. In such cases, the best one can hope for is a partial ordering of the predictors' importance. Pratt (1987) argued that looking for relative importance may be analogous to "dividing the indivisible" because the predictors are often interrelated and intertwined in extremely complex and complicated patterns. It is important to distinguish between those cases in which a ranking, or scaling, by importance is feasible and meaningful and those instances in which such a ranking is futile and meaningless.

To determine the importance of x_i and x_j, one should ideally compare a certain function of x_i (unaffected by x_j) with a similar function of x_j (unaffected by x_i). Yet, many of the methods reviewed earlier yield measures that are correlated and interdependent. Thus, it is questionable whether their comparison can lead to meaningful conclusions about relative importance. For example, when comparing two regression coefficients (or partial correlations), one is contrasting the effect of x_i given x_j (and possibly other variables) with the effect of x_j given x_i (and possibly others).

A New Approach to Determining Relative Importance

An alternative approach to the problem of importance is to formulate a set of requirements and conditions to be satisfied by a new method. Specifically, a method designed to identify the variables' relative importance must meet the following three conditions: (a) Importance should be defined in terms of a variable's "reduction of error" in predicting the criterion, y; (b) The method should allow for direct comparison of relative importance instead of relying on inferred measures; (c) Importance should reflect a variable's direct effect (i.e., when considered by itself), total effect (i.e., conditional on all other predictors), and partial effect (i.e., conditional on subsets of predictors). I make no claims of exclusivity or uniqueness with regard to these desiderata; some may consider this selection of conditions quite arbitrary and may suggest others. This particular set of requirements is consistent with a fairly general and broad interpretation of importance. Furthermore, any technique satisfying these requirements will also fulfill the requirements of most of the measures currently available.

The first requirement explicitly equates importance with prediction or error reduction, which appears to be the most prevalent and popular interpretation in the social sciences. The second condition induces a mechanism for distinguishing between cases in which variables can be ranked according to their importance and those in which they cannot be ordered by this criterion. Finally, the third condition outlines the three components of reduction in error associated with importance: direct, partial, and total.

Next I describe a new methodology (not necessarily the only one) that satisfies the three requirements: dominance analysis. Dominance is defined as a pairwise relationship that can be tested for all $p(p-1)/2$ pairs of variables included in the model. Consider any pair of predictors, x_i and x_j, and let x_h stand for any subset of the remaining $p-2$ variables (i.e., excluding x_i and x_j). Define variable x_i to "weakly dominate" variable x_j if, and only if,

$$\rho_{y\cdot x_ix_h}{}^2 \geq \rho_{y\cdot x_jx_h}{}^2 \qquad (9)$$

Table 1
Dominance Analysis for p = 3

Variable(s)	ρ^2	Contribution of x_1	Contribution of x_2	Contribution of x_3
—	0	$\rho_{y\cdot x_1}{}^2$	$\rho_{y\cdot x_2}{}^2$	$\rho_{y\cdot x_3}{}^2$
x_1	$\rho_{y\cdot x_1}{}^2$	—	$\rho_{y\cdot x_1x_2}{}^2 - \rho_{y\cdot x_1}{}^2$	$\rho_{y\cdot x_1x_3}{}^2 - \rho_{y\cdot x_1}{}^2$
x_2	$\rho_{y\cdot x_2}{}^2$	$\rho_{y\cdot x_1x_2}{}^2 - \rho_{y\cdot x_2}{}^2$	—	$\rho_{y\cdot x_2x_3}{}^2 - \rho_{y\cdot x_2}{}^2$
x_3	$\rho_{y\cdot x_3}{}^2$	$\rho_{y\cdot x_1x_3}{}^2 - \rho_{y\cdot x_3}{}^2$	$\rho_{y\cdot x_2x_3}{}^2 - \rho_{y\cdot x_3}{}^2$	—
x_1, x_2	$\rho_{y\cdot x_1x_2}{}^2$	—	—	$\rho_{y\cdot x_1x_2x_3}{}^2 - \rho_{y\cdot x_1x_2}{}^2$
x_1, x_3	$\rho_{y\cdot x_1x_3}{}^2$	—	$\rho_{y\cdot x_1x_2x_3}{}^2 - \rho_{y\cdot x_1x_3}{}^2$	—
x_2, x_3	$\rho_{y\cdot x_2x_3}{}^2$	$\rho_{y\cdot x_1x_2x_3}{}^2 - \rho_{y\cdot x_2x_3}{}^2$	—	—

for all $2^{(p-2)}$ possible selections of x_h, including the null set. The notation x_i D x_j indicates dominance. An alternative form of the definition involves the variables' usefulness: x_i D x_j if, and only if,

$$(\rho_{y\cdot x_ix_h}{}^2 - \rho_{y\cdot x_h}{}^2) \geq (\rho_{y\cdot x_jx_h}{}^2 - \rho_{y\cdot x_h}{}^2) \tag{10}$$

for all possible choices of x_h; x_i is said to weakly dominate (i.e., be at least as important as) x_j if, in all subset models that do not include either of the two predictors, x_i is at least as useful as x_j (i.e., x_i adds to the overall fit of the model at least as much as x_j). One variable is more important than its competitor if its predictive ability exceeds the other's in all subset regressions.

According to this definition, two predictors, x_i and x_j, can be involved in one of the following four exclusive relationships: (a) x_i D x_j, (b) x_j D x_i, (c) x_i D x_j and x_j D x_i, or (d) neither x_i D x_j nor x_j D x_i. The first two situations define the simple cases in which one of the variables dominates the other, the third case identifies equally important predictors, and the last one indicates situations in which the relative importance of two variables cannot be determined and they cannot be rank ordered meaningfully.

Examples

Consider first the case of $p = 2$ predictors. In this trivial situation, x_i D x_j if, and only if, $\rho_{y\cdot x_i}{}^2 \geq \rho_{y\cdot x_j}{}^2$ (note that this implies that the contribution of x_i to a submodel consisting of x_j only is greater than the additional contribution of x_j to x_i alone).

Table 1 presents a prototype of dominance analysis for $p = 3$ predictors. The first column identifies the variables in each submodel, and the second describes the fit of that model.

The next three columns (one for each predictor) describe the increase in the model's fit as a result of the addition of that particular variable. For example, the first row describes the increase in goodness of fit of the null model associated with the addition of each variable, and the second row describes the degree to which a model consisting of x_1 is improved by adding to it one of the additional predictors. To determine pairwise dominance, one compares each pair of columns (predictors) across all rows (submodels) for which both have nonempty entries. For $p = 3$, this amounts to two comparisons for each pair. For example, when comparing x_1 and x_2, one examines their direct contributions ($\rho_{y\cdot x_1}{}^2$ vs. $\rho_{y\cdot x_2}{}^2$) and their additional contributions to the model including x_3. If both differences have the same sign (e.g., x_1 has a greater contribution than x_2 in both instances), a dominance is established (x_1 D x_2).

As an illustration of this analysis, consider the following hypothetical example. A large university is studying the importance of various factors in prediction of performance in graduate school. The three predictors are as follows: x_1 = letters of recommendations, x_2 = Graduate Record Examination (GRE) scores, and x_3 = undergraduate grade point average (GPA). The criterion, y, is the cumulative GPA after 2 years of graduate school. Data are available for all graduate students over a period of several years. Because no sampling is involved, these data are treated as population values. Table 2 presents the $(2^3 - 1) = 7$ squared multiple correlations of all the possible models involving the $p = 3$ predictors. The values are presented in lexicographical order.

Table 2 also presents the calculations for dominance analysis. Dominance is examined for all pairs of predictors: GRE scores contribute more to the prediction of graduate performance than do letters of recommendation alone (.20 > .10) and in the presence of GPA (.17 > .06). Thus, GRE dominates recommendations. GRE scores are also better than GPA alone (.20 > .15) and in addition to recommendations (.15 > .11), so GRE dominates GPA. Finally, it is easy to verify that GPA dominates recommendations. Thus, the three predictors have been ranked ac-

Table 2
A Hypothetical Example With p = 3 Predictors

Variable(s)	ρ^2	Dominance analysis: Additional contribution of x_1	x_2	x_3
—	0	.10	.20	.15
x_1	.10	—	.15	.11
x_2	.20	.05	—	.12
x_3	.15	.06	.17	—
x_1, x_2	.25	—	—	.12
x_1, x_3	.21	—	.16	—
x_2, x_3	.32	.05	—	—

Note. x_1 = recommendations; x_2 = Graduate Record Examination (GRE); x_3 = grade point average.

cording to their relative importance: GRE, GPA, and letters of recommendation.

Properties of Dominance Analysis

By definition, dominance is obtained only if one variable betters the other in all models; thus, this approach guarantees consistency across all subsets and allows more general inferences. Also, all of the comparisons are easy to interpret because each pair of variables is compared only with respect to submodels excluding both variables of interest. The following is a partial list of the method's properties:

1. Importance depends only on the first two moments, and joint moments, of the $p + 1$ variables involved.
2. Dominance is invariant across any linear transformation applied to any subset of variables.
3. If x_i D x_j, it can be shown that the usefulness of x_i is not smaller than that of x_j; that is,

$$(\rho_{y\cdot x_1\cdots x_p}{}^2 - \rho_{y\cdot x_h x_j}{}^2) \geq (\rho_{y\cdot x_1\cdots x_p}{}^2 - \rho_{y\cdot x_h x_i}{}^2).$$

Thus, dominance analysis allows inferences regarding models including both variables as well.

4. Dominance is not affected by elimination of any subset of predictors from the model.
5. Dominance is not affected by (a) addition of any new "noise" variable uncorrelated with the criterion and the original variables or (b) addition of any new variable uncorrelated with $x_1 \cdots x_p$.
6. Dominance implies several of the better known measures of importance. In particular, if x_i D x_j, it must also be true that

$$\rho_{xy_i}{}^2 \geq \rho_{yx_j}{}^2 \ (x_h = \text{empty set}),$$

$$\rho_{y(x_i\cdot x_h)}{}^2 \geq \rho_{y(x_j\cdot x_h)}{}^2 \text{ for all } x_h,$$

and

$$\rho_{yx_i\cdot x_h}{}^2 \geq \rho_{yx_j\cdot x_h}{}^2 \text{ for all } x_h.$$

Also, the mean semipartial correlation of x_i must be equal to or greater than the mean semipartial correlation of x_j (Lindeman et al.'s 1980, measure), the mean partial correlation of x_i must be equal to or greater than the mean partial correlation of x_j (Kruskal's, 1987a, 1987b, measure), and the mean information of x_i must be equal to greater than the mean information of x_j (Theil's, 1987, measure).

7. Dominance is a transitive relationship. The following conditions apply for any triple (x_i, x_j, x_k): if x_i D x_j, then

$$\rho_{y\cdot x_i x_h}{}^2 \geq \rho_{y\cdot x_j x_h}{}^2.$$

If x_j D x_k, then

$$\rho_{y\cdot x_j x_h}{}^2 \geq \rho_{y\cdot x_k x_h}{}^2.$$

It follows that

$$\rho_{y\cdot x_i x_h}{}^2 \geq \rho_{y\cdot x_k x_h}{}^2,$$

implying that x_i D x_k.

8. Lack of dominance, however, is not necessarily transitive. It is conceivable that relative importance (dominance) cannot be determined for the pairs $\{x_i,x_j\}$ and $\{x_j,x_k\}$, yet x_i D x_k, x_k D x_i, or both.

A Quantitative Measure of Importance

Assume now that, in a system with p predictors, all $p(p-1)/2$ pairs can be ordered (i.e., dominance or equality was identified in each pair). In this case, it is possible to determine a meaningful quantitative measure of importance that is fully consistent with the current approach and yields a useful decomposition of the models' squared multiple correlation.

The qualitative notion of dominance relies on the comparison of the usefulness of the predictors across all subsets. The quantitative measure provides a summary of these usefulness measures. In particular, let $C_{x_i}^{(k)}$ be the mean usefulness of x_i across all $\binom{p-1}{k}$ models consisting of $k + 1$ variables (x_i and k additional variables):

$$C_{x_i}^{(k)} = \sum (\rho_{y\cdot x_i x_h}{}^2 - \rho_{y\cdot x_h}{}^2) \Big/ \binom{p-1}{k}, \tag{11}$$

where x_h is any subset of k predictors, x_i excluded. Thus, $C_{x_i}{}^{(k)}$ is the mean usefulness of x_i when it is added to k ($k = 0 \cdots p - 1$) additional predictors. Finally, by averaging these values across all orders, one obtains C_{x_i}, the variable's average usefulness:

$$C_{x_i} = \sum_{k=0}^{p-1} C_{x_i}^{(k)}/p. \tag{12}$$

$C_{x_i} \geq 0$ because it is a combination of positive components ($C_{x_i} = 0$ if, and only if, x_i is a noise variable uncorrelated with the criterion and the other predictors). Computationally, C_{x_i} is equivalent to Lindeman et al.'s (1980) measure (the mean squared semipartial correlation across all $p!$ permutations). Therefore, it provides the same decomposition of the model's total fit:

$$\rho_{y\cdot x_1\cdots x_p}{}^2 = \sum_{j=1}^{p} C_{x_j}. \tag{13}$$

With $p = 3$ variables, one simply calculates the predictor's average contribution for each class of models with $k = 0$, 1, or 2 predictors and then averages this contribution across all models. Thus, one obtains, for x_1,

$$C_{x_1}{}^{(0)} = \rho_{y\cdot x_1}{}^2,$$

$$C_{x_1}{}^{(1)} = [(\rho_{y\cdot x_1x_2}{}^2 - \rho_{y\cdot x_2}{}^2) + (\rho_{y\cdot x_1x_3}{}^2 - \rho_{y\cdot x_3}{}^2)]/2,$$

and

$$C_{x_1}{}^{(2)} = (\rho_{y\cdot x_1x_2x_3}{}^2 - \rho_{y\cdot x_2x_3}{}^2).$$

The variable's importance, C_{x_i}, is the mean of these three components.

It is interesting to note that, after some simple algebra, this equation can be reexpressed as follows:

$$C_{x_1} = \{[(\rho_{y\cdot x_1}{}^2 - (\rho_{y\cdot x_2}{}^2 + \rho_{y\cdot x_3}{}^2)/2] + [(\rho_{y\cdot x_1x_2}{}^2 - \rho_{y\cdot x_1x_3}{}^2)/2 - \rho_{y\cdot x_2x_3}{}^2] + \rho_{y\cdot x_1x_2x_3}{}^2\}/3. \tag{14}$$

Table 3
Quantitative Measures of Importance for the Hypothetical Example With Three Predictors

k	Recommendations	GRE	Grade point average
0	.10	.20	.15
1	.055	.16	.115
2	.05	.15	.12
$M(C_{x_i})$	.069	.173	.128
%	18.6	46.8	34.6

Note. GRE = Graduate Record Examination.

This particular form provides an interesting insight into the meaning of C_{x_1}: It averages, across all orders, the difference between the fit of the models including x_1 and the models excluding it. Thus, the first term compares the fit obtained by x_1 alone and the average of the direct contributions of x_2 and x_3; the second term compares the fit of all models with two predictors, including x_1, and the fit of a model including x_2 and x_3; and so on.

Table 3 illustrates these calculations for the hypothetical example presented in Table 2. The entries in Table 3 are the predictors' mean usefulness for all models of order $k = 0$, 1, and 2. Consistent with the definition of dominance, the value for GRE is the highest in each row, and the one for recommendations is the lowest. Thus, after ranking the importance of the three variables, we were able to scale them at .173, .128, and .069. These three values add up to the (full) model's squared multiple correlation (.37) and attribute 46.8% of the variance to GRE, 34.6% to GPA, and the remaining 18.6% to letters of recommendation.

Computation and Sampling Theory

The definition of dominance (Equation 9) indicates that, for each pair of variables, one needs to compare squared multiple correlations from several partial models. If one is to compare all $p(p-1)/2$ pairs of variables, all subset multiple correlations must be involved. Thus, any calculation of dominance relationships must start by obtaining all $2^p - 1$ squared multiple correlations (e.g., the RSQUARE procedure in SAS [SAS Institute, 1985]). Let ρ be the $(2^p - 1) \times 1$ vector of these subset squared multiple correlations, lexicographically ordered, and let $\mathbf{A}_{ij}$ be any $(2^{p-2}) \times (2^p - 1)$ matrix of contrasts identifying the 2^{p-2} models relevant for the comparison of x_i and x_j. The product, $\Delta_{ij} = \mathbf{A}_{ij}\rho$, is a $(2^{p-2}) \times 1$ vector including all of the relevant differences. If all of its entries are nonnegative, then x_i D x_j. If all of its entries are nonpositive, then x_j D x_i. If all of its entries are zero, x_i and x_j are equally important; all other cases indicate that x_i and x_j cannot be ordered with respect to their importance. Table 4 illustrates the calculations for the case of $p = 3$. The table presents the vector ρ; three matrices of contrasts, $\mathbf{A}_{12}$, $\mathbf{A}_{13}$, and $\mathbf{A}_{23}$, necessary for the comparisons of x_1 and x_2, x_1 and x_3, and x_2 and x_3, respectively; and the difference vectors Δ_{12}, Δ_{13}, and Δ_{23}.

Up to this point, the discussion has been in terms of population values and has ignored any sampling problems. Let $\mathbf{R}$ be the vector of the $2^p - 1$ sample squared multiple correlations, lexicographically ordered, based on n observations. It is well known that $\mathbf{R}$ is a biased estimator of ρ (Olkin & Pratt, 1958). Olkin and Siotani (1976) showed that the asymptotic distribution of $\mathbf{R}$ is multivariate normal, and Hedges and Olkin (1981) showed that, as $n \rightarrow \infty$,

$$\sqrt{n}\mathbf{A}(\mathbf{R} - \rho) \sim N(0, \mathbf{A}\Psi\mathbf{A}'), \qquad (15)$$

where ψ is the asymptotic variance–covariance matrix of the various squared multiple correlations (its form is outlined in the Appendix). Thus, in testing for dominance of one variable over the other, it is possible to use the just-described result to calculate $(1 - \alpha)$ 100% asymptotic confidence interval for any Δ_{ij}. If all lower bounds of Δ_{ij} are nonnegative, the results are supportive of the claim that x_i D x_j; if all upper bounds are nonpositive, they support the notion that x_j D x_i. If any of the intervals includes a zero, the relationship between the two predictors cannot be determined.

No precise small-sample model of inference is available for dominance analysis. One feasible approximate solution for small samples is to "jacknife" the estimates (e.g., Arvesen & Salsburg, 1975; Miller, 1974; Mosteller & Tukey, 1977). By eliminating one observation at a time, one can obtain n pseudo-independent estimates of all the relevant squared multiple correlations, estimate the variance–covariance matrix, and obtain approximate confidence intervals.

As an illustration of large-sample inference regarding dominance, consider an additional example. Pedhazur (1982, p. 202) presented data regarding the prediction of college GPA by three predictors—socioeconomic status (SES), IQ, and need achievement (nAch)—for a sample of 300 individuals. Table 5 presents the correlation matrix and the vector of squared multiple correlations in this sample, ordered lexicographically.

Table 6 summarizes the calculation of asymptotic confidence intervals for all three pairs of predictions. The first column presents the difference between the relevant squared multiple correlations (see Table 4). It appears that this is a clear case of domi-

Table 4
Calculation of All Three Pairwise Differences in a Model With p = 4 Predictors

$$\underbrace{\begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 \end{pmatrix}}_{A_{12}} \times \begin{pmatrix} \rho_{y\cdot 1}^2 \\ \rho_{y\cdot 2}^2 \\ \rho_{y\cdot 3}^2 \\ \rho_{y\cdot 12}^2 \\ \rho_{y\cdot 13}^2 \\ \rho_{y\cdot 23}^2 \\ \rho_{y\cdot 123}^2 \end{pmatrix} = \underbrace{\begin{pmatrix} \rho_{y\cdot 1}^2 - \rho_{y\cdot 2}^2 \\ \rho_{y\cdot 13}^2 - \rho_{y\cdot 23}^2 \end{pmatrix}}_{\Delta_{12}}$$

$$\underbrace{\begin{pmatrix} 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 \end{pmatrix}}_{A_{13}} \times \rho = \underbrace{\begin{pmatrix} \rho_{y\cdot 1}^2 - \rho_{y\cdot 3}^2 \\ \rho_{y\cdot 12}^2 - \rho_{y\cdot 23}^2 \end{pmatrix}}_{\Delta_{13}}$$

$$\underbrace{\begin{pmatrix} 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 \end{pmatrix}}_{A_{23}} \times \rho = \underbrace{\begin{pmatrix} \rho_{y\cdot 2}^2 - \rho_{y\cdot 3}^2 \\ \rho_{y\cdot 12}^2 - \rho_{y\cdot 13}^2 \end{pmatrix}}_{\Delta_{23}}$$

Table 5
A Numerical Example With p = 3

	Variable			
Variables	SES	IQ	nAch	GPA
SES	1			
IQ	.30	1		
nAch	.41	.16	1	
GPA	.33	.57	.50	1

Vector of squared multiple correlations

Variables in model	R^2
SES	.109
IQ	.325
nAch	.250
SES, IQ	.353
SES, nAch	.269
IQ, nAch	.496
SES, IQ, nAch	.496

Note. SES = socioeconomic status; nAch = need achievement; GPA = grade point average. (From *Multiple Regression in Behavioral Research, Explanation and Prediction* [p. 203] by E. J. Pedhazur, 1982, New York: Holt, Rinehart and Winston, Inc. Copyright 1982 by Holt, Rinehart and Winston, Inc. Reprinted by permission.)

nance: SES is dominated by IQ and nAch, and IQ seems to dominate nAch. The second column presents the standard errors of the differences (using Hedges and Olkin's, 1981, approach presented in the Appendix), and the last two columns show the lower and upper bounds of the 95% asymptotic confidence intervals. These bounds support the inferiority of SES as a predictor in the population. However, the comparison of IQ and nAch is inconclusive: The lower bounds are negative and the upper bounds are positive. Thus, although the difference between the squared multiple correlations favors IQ, one cannot reject the hypothesis that the predictors are equally important.

Table 6
Asymptotic 95% Confidence Intervals for All Pairwise Differences (n = 300)

Variables compared (i, j)	Difference	Asymptotic *SE*	95% confidence bounds: Lower	95% confidence bounds: Upper
SES–IQ	−.216	0.050	−.314	−.118
	−.228	0.039	−.300	−.158
SES–nAch	−.141	0.045	−.230	−.053
	−.144	0.031	−.204	−.084
IQ–nAch	.075	0.064	−.045	.195
	.084	0.055	−.023	.191

Note. SES = socioeconomic status; nAch = need achievement. Each pair of variables is compared across two submodels. (From *Multiple Regression in Behavioral Research, Explanation and Prediction* [p. 203] by E. J. Pedhazur, 1982, New York: Holt, Rinehart and Winston, Inc. Copyright 1982 by Holt, Rinehart and Winston, Inc. Reprinted by permission.)

Table 7
A Numerical Example With p = 4

	x_1	x_2	x_3	x_4	y
x_1	1				
x_2	.683	1			
x_3	.154	−.050	1		
x_4	.460	.297	.006	1	
y	.618	.461	.262	.507	1

Vector of squared multiple correlations

Variables	R^2	Variables	R^2
x_1	.382	x_3x_4	.324
x_2	.213		
x_3	.069	$x_1x_2x_3$	.419
x_4	.257	$x_1x_2x_4$	.448
x_1x_2	.385	$x_1x_3x_4$	.480
x_1x_3	.410	$x_2x_3x_4$	.439
x_1x_4	.445		
x_2x_3	.294	$x_1x_2x_3x_4$	.491
x_2x_4	.363		

Note. From *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (p. 99) by J. Cohen & P. Cohen, 1983, Hillsdale, NJ: Erlbaum. (Copyright 1983 by Lawrence Erlbaum Associates, Inc. Reprinted by permission.)

Finally, consider an example with $p = 4$ predictors. Cohen and Cohen (1983, p. 99) presented data for the prediction of salary (y) from years since doctoral degree (x_1), number of publications (x_2), sex (x_3), and number of citations (x_4). It is interesting to note that standard measures of importance disagree with regard to the ordering of the predictors: (a) The simple correlations indicate that $\rho_{yx_1} > \rho_{yx_4} > \rho_{yx_2} > \rho_{yx_3}$; (b) The standardized coefficients rank order the predictors as follows: $b_1 >$

Table 8
Dominance Analysis of a Numerical Example With p = 4

		Contribution of			
Variable(s)	R^2	x_1	x_2	x_3	x_4
—	0	.382	.213	.069	.257
x_1	.382	—	.003	.028	.063
x_2	.213	.172	—	.081	.150
x_3	.069	.341	.225	—	.255
x_4	.257	.188	.106	.067	—
x_1x_2	.385	—	—	.034	.063
x_1x_3	.410	—	.009	—	.070
x_1x_4	.445	—	.003	.035	—
x_2x_3	.294	.125	—	—	.145
x_2x_4	.363	.085	—	.076	—
x_3x_4	.324	.156	.115	—	—
$x_1x_2x_3$	.419	—	—	—	.072
$x_1x_2x_4$	.448	—	—	.043	—
$x_1x_3x_4$	.480	—	.011	—	—
$x_2x_3x_4$	.439	.052	—	—	—

Note. From *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (p. 99) by J. Cohen & P. Cohen, 1983, Hillsdale, NJ: Erlbaum. (Copyright 1983 by Lawrence Erlbaum Associates, Inc. Reprinted by permission.)

$b_4 > b_3 > b_2$; (c) The usefulness of the variables is as follows: $U(x_4) > U(x_1) > U(x_3) > U(x_2)$. The data are summarized in Table 7 and analyzed in Table 8.

As all six pairs of variables are compared, it becomes apparent that x_1 and x_4 dominate x_2 and x_3 but that the relationships between x_1 and x_4 and between x_2 and x_3 are indeterminate. Note, for example, that x_1 is a better predictor than x_4 by itself and in the presence of another variable (x_2 or x_3); however, x_4 betters x_1 in the presence of both x_2 and x_3. This result indicates that the four variables cannot be fully ordered and that their relative contribution cannot be assessed quantitatively. One can achieve a partial order only with (x_1 or x_4 or both) D (x_2 or x_3 or both). One possible way of "solving" this problem is to fit a simpler model. If, for example, one eliminates x_2, the overall fit drops only slightly (to 0.48), the three variables can be ordered by dominance (x_1 D x_4 D x_3), and their importance measures are 0.26, 0.16, and 0.05, respectively.

Concluding Remarks

The main goals of this article have been to illustrate the many meanings that have been explicitly and implicitly attached to the term *importance* in the context of multiple regression, to highlight certain weaknesses shared by most of these meanings, and to propose a new approach to this important problem. What is the role of dominance analysis in the standard use of multiple regression in behavioral research applications? This new technique was designed to characterize (and sometimes quantify) the relative importance of the p predictors in a specific model. Thus, it does not compete with the various methods used to identify the model, nor is it biased in favor of one of these methods. It is best to think of dominance analysis as complementing the stage of model identification and parameter estimation. Its major contribution is in the interpretation of the results.

One can think of regression analysis as a three-stage procedure. The first stage involves selection of the model, the second stage consists of a qualitative dominance analysis, and the final stage involves a quantitative analysis.

Selection of the Model

In many confirmatory applications, researchers consider the exact model to be known and specified on the basis of prior theoretical and empirical considerations. In other exploratory applications, researchers attempt to identify the "best" subset of predictors for the explanation and prediction of y. There are numerous techniques for selecting this group of variables, and many measures of goodness. A review and comparison of these methods is beyond the scope of this article (for partial reviews, however, see Cohen & Cohen, 1983; Darlington, 1990; Draper & Smith, 1981; or Hocking, 1976). Only after the completion of this model specification stage does it make sense to apply the dominance analysis. Needless to say, dominance analysis is conditional on the identification of the correct model. If the model is misspecified, the results of the dominance analysis will also be incorrect. This may cause problems, especially if certain predictors are incorrectly omitted from the model. As argued earlier, adding irrelevant (noise) variables should not affect the results of dominance comparison in the population. However, omission of variables that belong in the model can bias these comparisons (just as it causes bias in the regression coefficient estimates).

Qualitative Dominance Analysis

Dominance is determined through a pairwise comparison of all predictors. The final result of this stage is the establishment of a complete (e.g., the example in Table 2) or partial (e.g., the example in Table 8) ordering of the p predictors. The determination of this qualitative relationship involves a stringent criterion incorporating all relevant models for each pair of predictors. The two obvious advantages of this approach are that (a) the operational definition of importance is more compelling and matches the intuitive meaning of the term and (b) it eliminates most of the confusion due to the inconsistency between different measures (given a specific model) and between various submodels (assuming a certain measure). Recall that x_i D x_j if it predicts y better by itself, in the presence of all other $p - 2$ predictors, or in any other submodel. Thus, if dominance is established, x_i is a more important predictor than x_j by any of the other regular measures of importance. In those cases in which only a partial order can be established (see Table 8), the qualitative stage of the analysis pinpoints the potential sources of ambiguity in inference by identifying the subset of variables that cannot be ordered unequivocally and the source of the ambiguity (x_i may be better than x_j by itself and in the presence of all other variables, but only in certain submodels).

An important distinction, often overlooked when interpreting regression analysis, is the one between the descriptive and the inferential implications of the results. It is not uncommon to encounter research articles that report numerous significant tests to identify the best fitting model and interpret the estimates from the sample as if they were population values. Dominance analysis is, clearly, not immune to similar problems. Establishing that x_i dominates x_j in a certain sample does not guarantee that the same pattern will be replicated in other samples or in the population (see Table 6 for an example). Users of dominance analysis should make every effort to maintain this important distinction.

Quantitative Analysis

The last stage of the analysis invokes a quantitative approach and, in fact, makes use of an existing measure. However, it is important to emphasize that this quantitative stage is meaningful only under a certain pattern of qualitative conclusions. It is inconsistent with the basic approach to calculate Lindeman et al.'s (1980) measure for a set of predictors that cannot be ordered: Assigning unequal numerical values to such variables would imply that one is better or more important than the other, contradicting the result of the qualitative analysis. Assigning equal values to these variables would imply that they are equally important, which would, again, contradict the qualitative results.

Some Generalizations

In many cases, the most interesting question from a theoretical point of view is whether variable x_1 (e.g., SES) is more or less important than x_2 (e.g., IQ) in predicting a criterion, y, such as scholastic achievement, regardless of the other variables involved. The regular measures infer relative importance of the two predictors from their rank in the vector of p measures (e.g., standardized coefficients). In contrast, dominance analysis provides a methodology for a direct test of this question without explicitly referring to the importance of the other $p - 2$ variables.

Although dominance analysis was introduced as a method of comparing specific variables, it can be easily applied to groups (sets) of inseparable predictors. For example, consider a regression in which religion is operationalized by three binary dummy variables, x_1, x_2, and x_3, and marital status is operationalized by two additional dummy variables, x_4 and x_5. The meaningful comparison involves the set of religion variables against the set of marital status predictors rather than all 10 pairwise comparisons because, from a theoretical point of view, the dummy variables in each set will always be considered as an integral group. Similarly, when fitting a polynomial regression, it seems sensible to combine all terms of x_1 (linear, quadratic, cubic, etc.) and of x_2 for the purposes of importance comparisons of the two predictors.

References

Achen, C. H. (1982). *Interpreting and using regression.* Beverly Hills, CA: Sage.

Arvesen, J. N., & Salsburg, D. S. (1975). Approximate tests and confidence intervals using the jacknife. In R. M. Elashoff (Ed.), *Perspectives in biometrics* (pp. 123–149). New York: Academic Press.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cramer, E. M. (1974). Brief report: The use of highly correlated predictors in regression analysis. *Multivariate Behavioral Research, 9,* 241–244.

Darlington, R. B. (1968). Multiple regression in psychological research and practice. *Psychological Bulletin, 69,* 161–182.

Darlington, R. B. (1990). *Regression and linear models.* New York: McGraw-Hill.

Draper, N. R., & Smith, H. (1981). *Applied regression analysis* (2nd ed.). New York: Wiley.

Graybill, F. P. (1961). *An introduction to linear statistical models.* New York: McGraw-Hill.

Hedges, L., & Olkin, I. (1981). The asymptotic distribution of communality components. *Psychometrika, 46,* 331–337.

Hocking, R. R. (1976). The analysis and selection of variables in linear regression. *Biometrics, 32,* 1–49.

Kruskal, W. (1984). Concepts of relative importance. *Qüestiiô, 8,* 35–45.

Kruskal, W. (1987a). Relative importance by averaging over orderings. *American Statistician, 41,* 6–10.

Kruskal, W. (1987b). Correction to Kruskal. *American Statistician, 41,* 341.

Kruskal, W. P., & Majors, R. (1989). Concepts of relative importance in recent scientific literature. *American Statistician, 43,* 2–6.

Lindeman, R. H., Merenda, P. F., & Gold, R. Z. (1980). *Introduction to bivariate and multivariate analysis.* Glenview, IL: Scott, Foresman.

Miller, R. J. (1974). The jacknife: A review. *Biometrika, 61,* 1–15.

Mood, A. M. (1969). Macro-analysis of the American educational system. *Operations Research, 17,* 770–784.

Mosteller, F., & Tukey, J. W. (1977). *Data analysis and regression: A second course in statistics.* Reading, MA: Addison-Wesley.

Olkin, I., & Pratt, J. W. (1958). Unbiased estimation of certain correlation coefficients. *Annals of Mathematical Statistics, 39,* 201–211.

Olkin, I., & Siotani, M. (1976). Asymptotic distribution of functions of a correlation matrix. In S. Ideka (Ed.), *Essays in probability and statistics* (pp. 235–251). Tokyo: Shinho Tsusho.

Pedhazur, E. J. (1975). Analytic methods in studies of educational effects. In F. N. Kerlinger (Ed.), *Review of research in education* (Vol. 3, pp. 243–286). Itasca, IL: Peacock.

Pedhazur, E. J. (1982). *Multiple regression in behavioral research: Explanation and prediction* (2nd ed.). New York: Holt, Rinehart & Winston.

Pratt, J. W. (1987). Dividing the indivisible: Using simple symmetry to partition variance explained. In T. Pukilla & S. Duntaneu (Eds.), *Proceedings of Second Tampere Conference in Statistics* (pp. 245–260). University of Tempere, Finland.

SAS Institute. (1985). *SAS user's guide: Statistics, Version 5 edition.* Cary, NC: Author.

Stavig, G. S. (1977). The semistandardized regression coefficient. *Multivariate Behavioral Research, 12,* 255–258.

Theil, H. (1987). How many bits of information does an independent variable yield in a multiple regression? *Statistics and Probability Letters, 6,* 107–108.

Theil, H., & Chung, C. F. (1988). Information-theoretic measures of fit for univariate and multivariate linear regressions. *American Statistician, 42,* 249–252.

Tzelgov, J., & Henik, A. (1991). Suppression situations in psychological research: Definitions, implications and applications. *Psychological Bulletin, 109,* 524–536.

Ward, J. H. (1969). Partitioning of variance and contribution or importance of a variable: A visit to a graduate seminar. *American Educational Research Journal, 6,* 467–474.

Williams, E. J. (1978). Linear hypotheses: Regression. In W. H. Kruskal & J. M. Tanur (Eds.), *International encyclopedia of statistics* (pp. 523–541). New York: Free Press.

Williams, F., & Mosteller, F. (1947). Education and economic status as determinants of opinion. In H. Cantril (Ed.), *Gauging public opinion* (pp. 195–208). Princeton, NJ: Princeton University Press.

Yule, G. V. (1899). An investigation into the causes of change in pauperism, chiefly during the last two intercensal decades (Part I). *Journal of the Royal Statistical Society, 62,* 249–295.

(Appendix follows on next page)

Appendix

Asymptotic Sampling Distribution of All Squared Multiple Correlations (from Hedges & Olkin, 1981)[1]

Let ρ be the vector of all 2^{p-1} squared multiple correlations ordered lexicographically, and let $\mathbf{R}$ be its sample counterpart. Let a and b denote nonempty sets of the indices of the predictors $(1, 2, \ldots, p)$, and let $\rho_{y \cdot a}{}^2$ and $\rho_{y \cdot b}{}^2$ stand for the squared multiple correlations of y with the variables included in subsets a and b, respectively. Define a^+ and b^+ as $a^+ = a \cup y$ and $b^+ = b \cup y$. For any nonempty subsets i and j of $\{0, 1, \ldots, p\}$, define the matrix of bivariate population and sample correlations whose subscripts lie in i and j, respectively, by P_{ij} and R_{ij}. Thus, P_{ij} is the matrix of correlations ρ_{st} such that $s \in i$, $t \in j$, and R_{ij} is the sample analogue of P_{ij}.

THEOREM:

Let $(y, x_1, \ldots, x_p)$ have a multivariate normal distribution; then, as $n \to \infty$,

$$\sqrt{n}(R - \rho) \sim N(0, \Psi),$$

where $\psi = \|\psi_{ab}\|$,

$$\psi_{aa} = 4\rho_{y \cdot a}{}^2(1 - \rho^2{}_{y \cdot a})^2,$$

$$\psi_{ab} = \phi(a^+, b^+)\left[\frac{1}{|P_{aa}|\,|P_{bb}|}\right] + \phi(a, b)\left[\frac{|P_{a^+a^+}|\,|P_{b^+b^+}|}{|P_{aa}|^2|P_{bb}|^2}\right] - \phi(a^+, b)\left[\frac{|P_{b^+b^+}|}{|P_{aa}|\,|P_{bb}|^2}\right] - \phi(a, b^+)\left[\frac{|P_{a^+a^+}|}{|P_{aa}|^2|P_{bb}|}\right],$$

and

$$\phi(x, y) = 2|P_{xx}|\,|P_{yy}|\operatorname{tr}\{(P_{xx}^{-1} - I)P_{xy}(P_{yy}{}^{-1} - I)P'_{xy}\}, \quad x, y \in \{a, a^+, b, b^+\}.$$

[1] From "The Asymptotic Distribution of Communality Components" by L. Hedges and I. Olkin, 1981, *Psychometrika*, *46*, p. 333. Copyright 1981 by the Psychometric Society. Reprinted by permission.

Received June 14, 1991
Revision received February 1, 1993
Accepted April 19, 1993 ■