Practical Assessment, Research, and Evaluation

Volume 17 *Volume 17, 2012* Article 9

2012

Interpreting Multiple Linear Regression: A Guidebook of Variable Importance

Laura L. Nathans

Frederick L. Oswald

Kim Nimon

Recommended Citation

Nathans, Laura L.; Oswald, Frederick L.; and Nimon, Kim (2012) "Interpreting Multiple Linear Regression: A Guidebook of Variable Importance," *Practical Assessment, Research, and Evaluation*: Vol. 17 , Article 9.
DOI: https://doi.org/10.7275/5fex-b874
Available at: https://scholarworks.umass.edu/pare/vol17/iss1/9

Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Volume 17, Number 9, April 2012 ISSN 1531-7714

Interpreting Multiple Linear Regression: A Guidebook of Variable Importance

Laura L. Nathans, *University of North Texas*
Frederick L. Oswald, *Rice University*
Kim Nimon, *University of North Texas*

Multiple regression (MR) analyses are commonly employed in social science fields. It is also common for interpretation of results to typically reflect overreliance on beta weights (cf. Courville & Thompson, 2001; Nimon, Roberts, & Gavrilova, 2010; Zientek, Capraro, & Capraro, 2008), often resulting in very limited interpretations of variable importance. It appears that few researchers employ other methods to obtain a fuller understanding of what and how independent variables contribute to a regression equation. Thus, this paper presents a guidebook of variable importance measures that inform MR results, linking measures to a theoretical framework that demonstrates the complementary roles they play when interpreting regression findings. We also provide a data-driven example of how to publish MR results that demonstrates how to present a more complete picture of the contributions variables make to a regression equation. We end with several recommendations for practice regarding how to integrate multiple variable importance measures into MR analyses.

Across behavioral science disciplines, multiple linear regression (MR) is a standard statistical technique in a researcher's toolbox. An extension of simple linear regression, MR allows researchers to answer questions that consider the role(s) that multiple independent variables play in accounting for variance in a single dependent variable. Researchers tend to rely heavily on beta weights when interpreting MR results (e.g., Nimon, Gavrilova, & Roberts, 2010; Zientek, Carpraro, & Capraro, 2008). Presumably, this is because most statistical packages automatically output these weights by default, which can then be easily rank ordered based on their magnitudes. But beta weights serve as only one method for answering the proverbial fairy-tale question: *Mirror, mirror on the wall, what is the best predictor of them all?* As others have shown (e.g., Courville & Thompson, 2001) and readers will see, it is often not best to rely only on beta weights when interpreting MR results. In MR applications, independent variables are often intercorrelated, resulting in a statistical phenomenon that is referred to as *multicollinearity* when correlations between predictors are high, and, more generally, associations when there are correlations between independent variables. The more predictors there are in the model, the greater the potential there is for multicollinearity or association between variables (Pedhazur, 1997; Zientek & Thompson, 2006). The current paper demonstrates that there are several approaches researchers can use to gain insight into the role(s) that predictors play in MR that are particularly important to use in the presence of associations or correlations between variables. Each approach, we show, yields different perspectives and insights regarding the importance of independent variables in a regression equation, as well as often different rank orderings of those independent variables in terms of their contributions to the regression equation. We support Azen and Budescu's (2003) assertion that assessment of *variable importance* is contingent on how importance is defined and quantified and that it is therefore impossible to generate a, "precise universal definition of importance" (Budescu, 1993, p. 544). The term *variable importance* in and of itself is not a meaningful term; rather, it must be discussed (if it is discussed at all) in the context of a specific metric for it to have any meaning to the researcher or the reader. It is probably better for researchers to emphasize the specific ways in which *variable importance* is operationalized (e.g., dominance analysis, commonality analysis) rather than to focus on the blanket term and assume that the reader knows what this term means. Through gaining an understanding of multiple methods of assessing *variable importance* and how they complement each other, a researcher should be able to avoid dichotomous thinking (e.g., yes it is an 'important' predictor, or no it is not), and instead understand the importance of independent variables in more nuanced terms. This means that there is no single 'right' way to interpret regression results, and although reliance on beta weights may feel right because it is normative practice, it provides very limited information. We want to open researchers' eyes to the multiple lenses through which MR can be profitably viewed; we also want researchers to promote the value of the multiple-lens approach in their publications and with their colleagues and graduate students.

Several statistical techniques have been developed to determine independent variables' contributions to MR models. Our focus is on two general families of techniques. One family provides different methods of rank ordering individual predictors' contributions to an overall regression effect or R^2 (e.g., Pratt's measure, dominance analysis, relative weights), and the other family involves partitioning R^2 into the unique and shared variance contributions of the independent variables (e.g., commonality analysis, squared semi-partial correlations). These two families are aligned with the framework of LeBreton, Ployhart, and Ladd (2004), who categorized methods of *variable importance* into those that assess (a) *direct effects*, which quantify the contribution of each independent variable to the regression equation when measured in isolation from other independent variables; (b) *total effects*, which quantify the each independent variable's contribution to the regression equation when the variance contributions of all other predictors in the regression model have been accounted for; or (c) *partial effects*, which quantify the each independent variable's contribution to the regression equation while accounting for the contributions to regression models of a specific subset or subsets of remaining independent variables. It is important to clarify that a *direct effect* in this paper refers to a relationship between an independent and a dependent variable that does not incorporate variance contributions of other independent variables, as opposed to a direct effect in a path model, which refers to an independent variable that directly impacts another variable in the model (either an independent or dependent variable). Different techniques for assessing *variable importance* from these different categories may potentially (but not in all cases) yield different rank orderings of independent variables. Use of multiple techniques that reflect these three types of effects when assessing variable importance will yield the richest and most complete picture regarding the relationships between independent variables and the dependent variable.

Keeping this context in mind, the goal of our paper is to present a set of measures or *lenses* that provide different yet complementary ways of viewing the role(s) that each independent variable plays in a MR equation. Our paper is structured by (a) defining each measure, (b) highlighting each measure's advantages and limitations, (c) describing how each measure can help identify suppression effects, (d) outlining specific research questions that each measure can address, and (e) detailing when the researcher should select each particular index to assess *variable importance*. This guidebook should serve as a practical resource for researchers to define and identify the measure or set of measures with which to analyze their data using MR. In conclusion, we present a data-driven example and results section that researchers can use as a template for interpreting and reporting MR results. Lastly, we present recommendations for practice for selecting and reporting of the variable importance measures included in our guidebook.

Beta Weight

Research shows that beta weights are heavily relied on to assess variable importance (e.g., Courville & Thompson, 2001; Nimon, Gavrilova, & Roberts, 2010; Zientek, Carpraro, & Capraro, 2008). The regression weight for each given independent variable is interpreted as the expected difference in the dependent variable score between people who differ by one unit on that independent variable, with all other independent variable scores held constant (Hoyt, Leierer, & Millington, 2006; Johnson, 2004). When variables are standardized (i.e., converted to z-scores), regression weights are called beta weights. A beta weight for an independent variable indicates the expected increase or decrease in the dependent variable, in standard deviation units, given a one standard-deviation increase in independent variable with all other independent variables held constant. When variables are not standardized (i.e., scaled in their original metric), regression weights are called B weights. A B weight also indicates how much a one unit increase in the independent variable results in an increase in the dependent variable with all other variables held constant. However, this increase is scaled in terms of the variable's original scaling metric, rather than in a standardized metric. Our focus in this paper is on beta weights rather than B weights, because beta weights are more comparable across independent variables due to being scaled in the same standardized metric.

According to Pedhazur (1997), beta weights are computed to weight the independent variables so that when the weights are multiplied by variable scores, their sum is maximally correlated with the dependent variable. This computation process minimizes the sum of squared errors between the dependent variable scores and the dependent variable scores predicted by the regression equation (called Y-hat or $\hat{y}$) (Pedhazur, 1997). Because the beta weight calculation process accounts for the contributions of all variables in the model to the regression equation, each beta weight is a measure of the total effect of an independent variable (cf., LeBreton, Ployhart, & Ladd, 2004). If beta weights are rank-ordered by their absolute values, it is important to understand that the rank ordering represents solely this type of contribution (and not others that we will describe shortly).

According to Pedhazur (1997), sole reliance on using beta weights to interpret MR is only justified in the case where predictors are perfectly uncorrelated. In the absence of shared variances between independent variables, each standardized beta weight is equal to the zero-order correlation between the independent and dependent variable. In such a case, variable importance can easily by determined by squaring the standardized beta weights; thus, it is not necessary to calculate more complicated MR indices.

The major advantage of beta weights is that they provide a measure of variable importance that is easily computed and provides an initial rank ordering of independent variables' contributions to a MR equation that accounts for contributions of other independent variables. However, this lens has a limited focus due to associations between independent variables. A given beta weight may receive the credit for explained variance that it shares with one or more independent variables (Pedhazur, 1997). As such, the other weights are not given credit for this shared variance, and their contribution to the regression effect is thus not fully captured in the beta weight value.

Beta weights are also limited in their ability to determine suppression in a regression equation. An independent variable that contributes little or no variance to the dependent variable may have a large non-zero beta weight because it "purifies" one or more independent variables of their irrelevant variance, thereby allowing it or their predictive power to increase (Capraro & Capraro, 2001). In such a case, the beta weight's value may lead the researcher to erroneously conclude that it directly predicts the dependent variable (Pedhazur, 1997). Thus, after obtaining beta weights, researchers should not cease to analyze their MR results, but rather should consider additional measures to gain a broader and fuller perspective on the contributions that independent variables make to the regression equation.

Beta weights are best used as an initial "starting point" from which to begin exploring the issue of independent variables' contributions to a regression equation. It is recommended that all researchers begin MR analyses with beta weights, as they are easily computed with most statistical software packages and can provide an initial rank ordering of variable contributions in one computation. If there are no associations between independent variables or the model is perfectly specified (Courville & Thompson, 2001), no other techniques need to be employed. However, in the more realistic case of correlated predictors, researchers should select from at least one of the other techniques discussed in this guidebook to determine the impact of shared variance between variables on the regression equation.

Beta Weight Research Question: *What is the contribution of each independent variable to the regression equation, holding all other independent variables constant?*

Zero-Order Correlation

As applied to MR, zero-order correlations reflect the bivariate relationships between independent and dependent variables. According to Hinkle, Wiersma, and Jurs (2003), the correlation coefficient is, "an index that describes the extent to which two variables are related" (p. 98). The correlation coefficient reflects both the magnitude and direction of the relationship between two independent variables. Its values range from -1.0 to 1.0. If a correlation coefficient is negative, the values of the two variables that are correlated are inversely related; as one variable's scores increase, the other variable's scores decrease. If a correlation coefficient is positive, an increase (or decrease) in one variable is related to an increase (or decrease) in the other variable in the coefficient. The closer the value of the correlation coefficient is an absolute value of 1.0, the larger the magnitude of the relationship is between the two variables. If the value of the correlation coefficient is zero, there is no relationship between the two variables. It is important to note that the strength/magnitude of a correlational relationship is not related to the sign of the correlation coefficient; thus, equivalent predictive power can be attributed to correlations of equivalent magnitude but different signs.

Zero-order correlations are measures of direct effect (cf., LeBreton, Ployhart, & Ladd, 2004), as they determine the magnitude of the bivariate relationship between the independent and dependent variable without accounting for the contributions of other variables in the regression equation. These coefficients are also commonly referred to as "validities" (Nunnally & Bernstein, 1994). In the case where independent variables are uncorrelated, zero-order correlations are equivalent to beta weights and are sufficient to rank order independent variables. When squared, they add up to the model R2, and thus partition the regression effect (Pedhazur, 1997). Conversely, in the case where at least some of the independent variables are correlated, squared zero-order correlations will generally add up to a total that is greater than the model R^2, as shared variance in the dependent variable is added multiple times into the overall sum for R^2. Thus, in the latter case, researchers should employ other statistics to determine how the regression equation is affected by shared variance among the independent variables.

According to Nunnally and Bernstein (1994), there are several advantages to use of zero-order correlations. First, if a researcher has to select one independent variable to target for research or intervention after obtaining several beta weights of similar magnitude, s/he should choose the variable with the highest zero-order correlation. Second, zero-order correlations are less sensitive to effects of sampling error than are beta weights. Third, this measure is the only measure presented in this guidebook that is able to quantify how much variance is directly shared between the independent and dependent variable without being affected by shared variance between independent variables.

According to Nunnally and Bernstein (1994), there are limitations to the use of zero-order correlations. First, an independent variable may have the largest zero-order correlation with the dependent variable yet make the smallest (or potentially no) contribution to the regression equation when measures of total effect are calculated due to variance it shares with other variables. As such, several independent variables may show high zero-order correlations, yet may not be, "particularly important to independent prediction" (p. 192) when other independent variables' contributions to the regression equation are accounted for. Thus, it is best to complement use of zero-order correlations with measures of total and partial effects that consider contributions of other variables to the regression equation in their assessments of variable importance.

Zero-order correlations are often useful to identify the occurrence of suppression when viewed in concert with beta weights, although they cannot identify suppression effects in and of themselves. If an independent variable has a near-zero or negligible zero-order correlation with the dependent variable and a large and statistically significant beta weight, these two statistics suggest that the variable is a suppressor. This variable shares no variance directly with the dependent variable and thus contributes to the regression equation through removing irrelevant variance from other independent variables.

Researchers should choose this measure as a complement to beta weights when they are interested in determining the magnitude and direction of the bivariate relationship between an independent variable and the dependent variable without accounting for other predictors in the regression equation.

Zero-Order Correlation Research Question: *What is the magnitude and direction of the relationship between the independent and dependent variable, without considering any other independent variables in the MR model?*

DOI: https://doi.org/10.7275/5fex-b874

Product Measure

Pratt (1987) proposed the product measure, which is calculated by multiplying the variable's zero order correlation (its relationship to the dependent variable in isolation from of other independent variables) by its beta weight (which accounts for contributions of all other predictors to the regression equation). Thus, the product measure uniquely reflects in one statistic both direct and total effects (cf. LeBreton, Ployart, & Ladd, 2004).

A benefit of product measures is that they partition the regression effect; thus, their sum across all independent variables equals the multiple R^2 for the regression model, even when independent variables are correlated with one another (Azen & Budescu, 2003). This measure thus enables rank orderings of variable importance based on the partitioning of the regression effect. The major difficulty with this measure is that it may also yield negative values for an independent variable even when the independent variable accounts for a large amount of variance in the dependent variable, particularly if either the zero-order correlation or beta weight for an independent variable is negative (Darlington, 1968). If large negative values are generated for independent variables, it renders their product measure values substantively meaningless in terms of quantifying their contribution to R^2 because their variance contributions to the regression effect are subtracted from rather than added to the overall R^2. Thus, they should not be used in such cases.

Concerning suppression, Thomas, Hughes, and Zumbo (1998) explained that the product measure can identify suppressor variables. Multiplying an independent variable's beta weight by the small or negligible zero-order correlation of the suppressor variable with the dependent variable will yield a small or negligible product measure value, thereby demonstrating that the variable did not directly contribute to the regression effect. Small negative values may also indicate the presence of suppression if a near-zero correlation coefficient is multiplied by a negative beta weight (a classic profile for suppression).

This measure should be selected as an initial, easily computed (as opposed to general dominance and relative weights, which will be discussed later) method of partitioning R^2 in the presence of correlated predictors. It should be used to complement beta weights to provide a different "perspective" on variable importance because its values, unlike those for beta weights, do sum to R^2. It can be compared and contrasted with other methods for partitioning the regression effect as well.

Product Measure Research Question: *How can the regression effect be partitioned into non-overlapping partitions based on the interaction between the beta weight of each independent variable and its zero-order correlation with the dependent variable?*

Relative Weights

Another method for partitioning the R^2 in MR between independent variables in the model is through *relative weight analysis* (RWA; Fabbris, 1980; Genizi, 1993; Johnson, 2000). Johnson (2000) explained that when independent variables are uncorrelated, the relative weights are computed by calculating the squared zero-order correlation between the independent variable and the dependent variable (also the standardized beta weight) and dividing this number by R^2. In contrast, when independent variables are correlated, relative weights address this problem by using principal components analysis to transform the original independent variables into a set of uncorrelated principal components that are most highly correlated with the original independent variables (Tonidandel & LeBreton, 2010). These components are then submitted to two regression analyses. The first analysis is a regression that predicts the dependent variable from these uncorrelated principal components. Next, the original independent variables are regressed onto the uncorrelated principal components. Finally, relative weights are computed by multiplying squared regression weights from the first analysis (regression of dependent variables on components) with squared regression weights the second analysis (regression of independent variables on components). Each weight can be divided by R^2 and multiplied by 100 so that the new weights add to 100%, with each weight reflecting the percentage of predictable variance. Relative weights are unique as a measure of *total effect* in that they provide rank orderings of individual independent variables' contributions to a MR effect in the presence of all other predictors based on a computational method that addresses associations between independent variables between variables by creating their uncorrelated "counterparts" (Johnson & LeBreton, 2004).

There are several strengths of relative weights. First, relative weights add up to R^2 (Tonidandel & LeBreton, 2011). They partition the regression effect based on a procedure that addresses the problem of associations between independent variables through the use of uncorrelated principal components. Relative weights will thus, "perform appropriately and much better than regression weights" when partitioning variance in the presence of correlated independent variables (Tonidandel & LeBreton, 2011, p. 5). Additionally, relative weights are easy to explain to researchers, because, unlike beta weights, they account for all variance explained by the regression model and at the same time address the minimize the impact of multicollinearity on weights attributed to each independent variable (Johnson & LeBreton, 2004). These points are probably more important than another advertised benefit of RWA, which is computational ease (Johnson, 2000).

According to Tonidandel and LeBreton (2011), there are also several limitations of relative weight analysis. First, these weights are highly dependent upon the independent variables in the regression model, and are thus "susceptible to model misspecification" (p. 5). Second, although the computational procedure for relative weights presents a unique way of dealing with associations between independent variables, it is a common "myth" (p. 5) among users of this method that this method "fixes" all related problems. Relative weights are still affected by associations between independent variables; as each weight generally contains variance that is shared with other independent variables as well as unique variance, associations between independent variables are not completely eliminated. We would also add as a potential limitation that relative weights only identify suppression effects indirectly, when the weights sum to a total that is larger than R^2 or account for greater than 100% of the variance in R^2 when relative weights are converted to percentages.

The researcher should generally select relative weights when there is multicollinearity between independent variables. It serves as a strong complement to beta weights as, it is uniquely able to partition R^2 while minimizing the impact of associations between variables, and can thus present a more accurate picture of variables' contributions to a regression effect than other R^2 partitioning methods (such as the product measure).

Relative Weights Research Question: *How do independent variables contribute to the dependent variable when the regression effect is partitioned as a joint function of (a) how highly related independent variables are to their uncorrelated counterpart and (b) how highly related the uncorrelated counterparts are to the dependent variable?*

Structure Coefficients

A structure coefficient is the bivariate correlation between an independent variable variable and the predicted y value resulting from the MR model, where y represents the predicted dependent variable scores (Courville & Thompson, 2001). A structure coefficient in MR analyses is a useful measure of a variable's *direct* effect (LeBreton, Ployhart, & Ladd, 2004), as it quantifies the magnitude of the bivariate relationship between each independent variable and y in isolation from other independent variable-y correlations. However, it is important to note that other independent variables do contribute indirectly to structure coefficient values, as they are used in computation of the y scores. The major difference between a zero-order correlation and a structure coefficient is that the structure coefficient is scaled to remove the difference of the multiple R^2.

According to Courville and Thompson (2001), there are two ways a structure coefficient can be computed. First, for a given independent variable, X, the structure coefficient is:

$$r_s = r_{X.Y}/R$$

where $r_{X.Y}$ is the bivariate correlation between the independent variable (X) and the dependent variable (Y), and R is the multiple correlation for the regression containing all independent variables. Second, the structure coefficient may be calculated by computing the correlation between a given independent variable and the predicted y scores, or:

$$r_s = r_{X.y}$$

When squared, structure coefficients represent the amount of variance that an independent variable shares with the variance from the predicted y scores.

According to Courville and Thompson (2001), a beneficial property of structure coefficients is that they

DOI: https://doi.org/10.7275/5fex-b874

are not affected by associations between independent variables, as a structure coefficient is simply a Pearson *r* between an independent variable and $\hat{y}$. Thus, structure coefficients enable an understanding of how much variance each independent variable shares with the predicted $\hat{y}$ scores that are the actual analytic focus of the study. If an independent variable has a small beta weight but explains substantial variance in $\hat{y}$ as reflected a large squared structure coefficient, the researcher then knows that (a) there is shared variance between this variable and another variable and (b) that the beta weight calculation process assigned that shared variance to another independent variable. However, structure coefficients are limited in and of themselves due to being solely a measure of *direct effect* that does not identify which independent variables jointly share variance in predicting the dependent variable or quantify the amount of this shared variance.

A special case that highlights the usefulness of structure coefficients in identifying how the variance assignment process for a particular regression equation occurs is the suppression case. If a structure coefficient is near zero or zero in magnitude, that independent variable contributes little or no direct variance to $\hat{y}$. If that independent variable has a substantial beta weight, it can be determined that the particular independent variable is a suppressor. Notably, a similar process is used when comparing zero-order correlations to beta weights to determine the presence of a suppression effect. The major difference between the two processes is that the independent variable in the zero-order correlation in a suppression effect will share little or no variance with the dependent variable, while the independent variable in the structure coefficient will share little or no variance with the $\hat{y}$ scores that are a portion of the dependent variable. However, the researcher does not know what variables are being suppressed and must rely on other techniques, such as commonality analysis (to be discussed), to determine the magnitude and loci of suppression.

As asserted by Courville and Thompson (2001) and we recommend, the researcher should select this measure in addition to beta weights in the presence of correlated predictors, as it is able to determine both (a) the variance each independent variable shares with $\hat{y}$ and (b) if a variable's contribution to the regression equation was "distorted" or minimized in the beta weight calculation process due to assignment of variance it shares with another independent variable to another beta weight.

Squared Structure Coefficients Research Question: *How much variance in the predicted scores for the dependent variable ($\hat{y}$) can be attributed to each independent variable when variance is allowed to be shared between independent variables*?

Commonality Coefficients

Developed in the 1960s as a method of partitioning variance (Mayeske et. al, 1969; Mood, 1969, 1971; Newton & Spurrell, 1967), commonality analysis partitions the R^2 that is explained by all independent variables in a MR into variance that is unique to each variable and variance that each possible subset of independent variables share (Onwuegbuzie & Daniel, 2003; Rowell, 1996). For example, if a MR models the effects of X_1, X_2 and X_3 in predicting Y (the dependent variable), then a commonality analysis calculates the amount of variance in Y that is predicted by 7 subsets that together add up to the model R^2: the unique variances of each variable X_1, X_2, and X_3; the shared variances of variables taken two at a time, $\{X_1, X_2\}$, $\{X1, X_3\}$, and $\{X_2, X_3\}$; and the shared variance of all three variables $\{X_1, X_2, X_3\}$. It is important to note that this process partitions the regression effect into nonoverlapping components of variance that can thus be easily compared.

There are two types of commonality coefficients: unique effects and common effects. Unique effects reflect how much variance an independent variable contributes to a regression equation that is not shared with other independent variables (Zientek & Thompson, 2006). This statistic is also termed the independent variable's usefulness or squared semipartial correlation. If independent variables are all uncorrelated, all independent variable contributions are unique effects, as no variance is shared between independent variables in predicting the dependent variable. In this case, unique effects are identical in value to squared zero-order correlations and squared beta weights, and variable importance can be determined by rank orderings of unique effects. A unique effect is a measure of *total effect*, as it is only calculated when all independent variables have been entered into the regression equation.

In contrast, common effects provide detailed information that identifies and quantifies the extent and pattern of the independent variables' "overlap" in predicting dependent variable variance (Mood, 1971). Common effects are measures of *total effect* (LeBreton et al., 2004), as they quantify the contribution to the regression effect that each variable shares with every other variable set. By knowing which variables share variance in R^2, researchers have knowledge of how particular sets of variables operate in combination in predicting an outcome, and can thus generate recommendations regarding how to jointly target these variables to produce desired effects (see Seibold & McPhee, 1979 for an example). Commonality coefficients sum to the multiple R^2 for the regression model. Thus, they can be used to determine how the regression effect is partitioned into percentages of unique and shared variance.

A unique property of commonality analysis is that the researcher can add the common effects for each independent variable and compare them with its unique effect to determine whether a variable contributes more to a regression effect when operating in combination with other variables or independently of them. The resulting data provide rich interpretation of the regression effect. Seibold and McPhee (1979) stressed the importance of decomposing R^2 into constituent parts:

> Advancement of theory and the useful application of research findings depend not only on establishing that a relationship exists among independent variables and the dependent variable, but also upon determining the extent to which those independent variables, singly and in all combinations, share variance with the dependent variable. Only then can we fully know the relative importance of independent variables with regard to the dependent variable in question. (p. 355)

There are two major difficulties with the commonality procedure. First, as the number of independent variables increases, the number of commonality coefficients increases exponentially (Mood, 1971). The number of commonality coefficients is 2^p-1, where p equals the number of independent variables in the MR model (Mood, 1971). For example, with three, four, or five independent variables, the number of commonality coefficients is 7, 15, and 31 respectively. Therefore, with large numbers of predictors, there are large numbers of commonalities to report and interpret. However, software in PASW (Nimon, 2010) and R (Nimon, Lewis, Kane, & Haynes, 2008) exists to compute these coefficients. Researchers will also need to summarize multiple combinations of variables represented by commonality coefficients. For example, with six variables, there are second-order commonalities between two variables, third-order commonalities between three variables, fourth-order commonalities between four variables, fifth-order commonalities between five variables, and a sixth-order commonality between all six of the variables (Amado, 1999). It may be hard for the researcher to assign clear meanings and interpretations to higher-order commonalities that reflect combinations of varied constructs (DeVito, 1976), but we argue that the attempt to do so can be worthwhile.

The results of a commonality analysis can aid in identifying where suppressor effects occur and also how much of the regression effect is due to suppression. Negative values of commonalities generally indicate the presence of suppressor effects (Amado, 1999). In the suppression case, a variable in a particular common effect coefficient that does not directly share variance with the dependent variable suppresses variance in at least one of the other independent variables in that coefficient. The suppressor removes the irrelevant variance in the other variable or variables in the common effect to increase the other variable(s)' variance contributions to the regression effect (DeVito, 1976; Zientek & Thompson, 2006). Commonality analysis is uniquely able to both identify which variables are in a suppressor relationship and the specific nature of that relationship. The researcher can look across a table of commonality coefficients and see if a particular variable is a part of multiple negative common effects, which suggests that it is a suppressor for at least one of the other variables in the common effects. Additionally, researchers can compare common effects with structure coefficients and zero-order correlations to help identify suppressors. If a variable has a small or negligible structure coefficient and zero-order correlation (and thus does not directly share variance with y or the dependent variable) and that same variable is part of a negative common effect, it can be determined that the variable is a suppressor for the other variables that are part of this common effect. Summing all negative common effects for a regression equation can quantify the amount of suppression present in the regression equation.

Researchers should select commonality analysis when they have identified that shared variance between independent variables is impacting the regression equation (usually through comparisons of beta weights with structure coefficients) so that they can determine the patterns and extent of shared and unique variance. Additionally, commonality analysis should always be selected for use in the presence of suppression effects, as it is uniquely able to both identify the variables that are in a suppressor relationship and quantify the amount of suppression present in a particular regression equation.

Commonality Analysis Research Question: *How much variance in the dependent variable is uniquely vs. jointly explained by each predictor or predictor set?*

Dominance Weights

Dominance analysis is a technique developed by Budescu (1993) and refined by Azen and Budescu (2003) to determine variable importance based on comparisons of unique variance contributions of all pairs of variables to regression equations involving all possible subsets of predictors. Budescu (1993) initially proposed dominance analysis as a means to improve upon previous relative importance measures by quantifying an independent variable's *direct effect* in isolation from other independent variables (as the subset containing no other independent variables includes squared zero-order correlations), *total effect* (as it compares independent variables' unique variance contributions when all predictors are included in the model), and *partial effect* (as it compares independent variables' unique variance contributions for all possible subsets of independent variables).

According to Azen and Budescu (2003), there are three types of dominance. First, an independent variable shows *complete dominance* over another independent variable across all submodels if the former independent variable always shows a higher unique variance contribution than the latter independent variable when it is entered last into regression equations containing all possible subsets of independent variables. Budescu (1993) lists four "exclusive" complete dominance relationships that an independent variable can have in relation to another independent variable: (a) X_1 completely dominates X_2; (b) X_2 completely dominates X_1; (c) X_1 and X_2 contribute equally to prediction of variance in the dependent variable; or (d) neither independent variable dominates the other across all possible model subsets (i.e., they each dominate different model subsets or no subsets).

Azen and Budescu (2003) developed a weaker type of dominance, termed *conditional dominance*. *Conditional dominance* is determined by first calculating the averages of independent variables' contributions to all models of the same subset size (e.g., how much unique variance, on average, independent variables add to models with no independent variables, one other independent variable, and two other independent variables if the model has three independent variables). If, on average, an independent variable contributes more unique variance than another independent variable across model of all subset sizes, an independent variable is said to conditionally dominate another independent variable.

Lastly, Azen and Budescu (2003) define a third and weakest type of dominance, termed *general dominance*. General dominance reflects the average additional unique variance contribution of each independent variable to *all* subset models; this can be computed by averaging across all conditional dominance statistics. Interpretively, general dominance represents the, "average difference in fit between all subset models (of equal size) that include X_i and those that do not include it" (Azen & Budescu, 2003, p. 137). An important property of general dominance variance averages is that they partition the total R^2, enabling the researcher to rank order independent variables' contributions to the regression effect based on their average contributions across all possible subsets of independent variables. Notably, this method of partitioning R^2 is different from that used with both the product measure, which is based upon zero-order correlations and beta weights rather than averages of uniqueness values, and relative weights, which are calculated on the basis of the entire regression model and not the average contribution across all sub-models.

There are many advantages of the dominance analysis procedure. First, dominance analysis enables comparisons between the uniqueness values of all independent variables for all possible subsets of independent variables in one technique (Budescu, 1993). Second, dominance analysis rank orders independent variables equivalently across multiple measures of fit, including Mallow's C, Akaike's information criterion, Bayesian information criterion, adjusted R^2, and many other measures (Azen & Budescu, 2003). Most importantly, dominance analysis is a comprehensive technique in that it is the only measure that involves calculation of all three types of effects in LeBreton et al. (2004)'s theoretical framework.

There are also several limitations of dominance analysis in addition to its many advantages. First, dominance analysis involves computation of numerous statistics, and the number of models increases exponentially with the number of independent variables. For example, Johnson (2000) noted that there are 1,023 models for 10 independent variables and 32,767 models for 15 independent variables. The researcher must construct tables of all R^2s and perform extensive visual comparisons to establish dominance relationships for multiple pairs of variables. A SAS macro is listed in Azen and Budescu (2003) to automate the process, however. Second, Budescu (1993) stressed that dominance analysis is contingent upon identification of the correct regression model; rank orderings will less informative if important variables are excluded and unimportant variables are included. Third, Baltes, Parker, Young, Huff, and Altman (2004) did not find that in their data, dominance analysis added new information to that found with more traditional variable importance measures except in the presence of independent variables that do not contribute substantially to the regression effect; thus, the extensive computations and comparisons that dominance analysis requires may not be justified in such cases.

Azen and Budescu (2003) explained how suppression effects can be determined through examination of conditional dominance statistics. Unique variance contributions of a "regular" independent variable averaged over all models of a given size will typically decrease with increasing numbers of independent variables. This trend occurs because variance is typically shared between the independent variables, which reduces the amount of variance an independent variable uniquely explains in the presence of other independent variables. However, in the suppression case, a suppressor will contribute more variance on average across subsets with greater numbers of predictors, as it only contributes variance to the regression equation through suppressing variance in other predictors. Viewing this trend for an independent variable that contributes no or little variance to the subset containing no predictors (and thus is not a direct contributor to variance in the dependent variable) will demonstrate that the independent variable is a classic suppressor.

Researchers should select this technique when they are interested in understanding the dynamics of all possible subsets of independent variables, which is not captured in any other measure in this guidebook. If the researcher wants to make pairwise comparisons between variables, this technique is solely suited for those comparisons, as well. This technique is also well-suited for selection of the subset of the most significant independent variables for future analyses.

Dominance Analysis Research Question: *Does one independent variable contribute more variance that another independent variable to the regression effect for models containing all or some subsets of independent variables, or on average across all possible subsets of independent variables?*

The preceding discussions are summarized in Table 1.

A Caveat Regarding Theory-Driven Regression Methods and Stepwise Regression

Because of the importance of regression analysis for answering theory-driven—as opposed to data-driven—research questions, we provide a discussion of how variable importance measures can be used in theory-based research. Theoretical considerations can be factored into both assessments of variable importance and inclusion of variables in regression models. According to Schafer (1991), the purpose of theory-based hierarchical regression analysis is to enter variables into regression equations in a predetermined order that is relevant to the theory underlying development of the regression model. This process enables determinations of both the (a) incremental predictability at each regression step and (b) the variance explained by the variable(s) entered at each step. The researcher is able to control the variance contributions of several "control" variables before entering the variables of "primary importance" to the theory with such methods. Researchers can examine variable importance in the same hierarchical manner, and they can frame variable importance in the context of the theory underlying the regression model.

Table 1. The Multiple Lenses of Multiple Regression: Ways to Operationalize Independent Variable Contributions to R^2

	Always Total R^2	Identifies Suppressor	Direct Effect	Total Effect	Partial Effect	Values are Identical When Predictors Uncorrelated	Identifies Multi-collinearity
β Weights				X		X	
Zero-order *r*			X			X	
Product Measure	X		X	X			
Structure Coefficient			X				
Commonality Coefficients	X	X					
Unique				X		X	
Common				X			X
Dominance Analysis							
Complete			X	X	X		
Conditional		X					
General	X						
Relative Weights	X			X			

Generally, in the presence of strong theory, regression equations can be used to answer three types of theoretical research questions: (a) can specific combinations of independent variables predict or explain variance in the dependent variable?; (b) is a specific variable in a set of independent variables necessary to predict or explain variance in the dependent variable; and (c) can specific combinations of independent variables predict or explain variance in the dependent variable, given a strong theoretical rationale for including control variables as predictors? (Thayer, 2002). In the context of intervention research, Groemping (2007) also argued that researchers can use regression results to make comparative judgments of which independent predictors can produce specific effects on a particular dependent variable that are used to inform intervention-related theory in such a case.

Stepwise Regression Methods

Stepwise regression methods are sometimes relied upon to determine a set of independent variables that purportedly represent the "best" set of predictors of a particular dependent variable. Hinkle, Wiersma, and Jurs (2003) outlined the steps involved in conducting forward stepwise regression. The first independent variable is selected for entry into the regression equation that demonstrates the highest bivariate correlation with the dependent variable. The second independent variable selected produces the highest increase in R^2 after accounting for the prediction of the first variable. After this second independent variable is added, a second significance test is conducted to determine if the first independent variable remains a statistically significant predictor; if it is not, it is dropped from the equation. This process repeats until either (a) all independent variables have been entered into the equation or (b) entry of the remaining independent variables into the stepwise solution does not produce a statistically significant increase in R^2.

There are several significant difficulties inherent in stepwise regression analyses (Thompson, 1989, 1995) that caution against its widespread use as a measure of variable importance. First, as Thompson (1995) pointed out, the analysis does not incorporate the correct number of degrees of freedom for each statistical significance test. The method assigns one degree of freedom to each variable that is tested at each step; however, because each independent variable is selected from all remaining independent variables in the regression equation at each step, a degree of freedom should be assigned to *all* considered variables. This oversight increases the F value at each step and its likelihood of being statistically significant. The miscalculation of degrees of freedom will thus result in more Type I errors overall (rejecting the null hypothesis when it is true; Thompson, 1989).

Even if this issue were addressed, a second significant difficulty with the stepwise regression procedure is that the independent variables that are selected are conditional on the variance contributions of the variables that have already been entered into the regression (Thompson, 1995). All results hinge on the variable that is selected as the first predictor; accordingly, the researcher will get different entry orders at different steps depending on the variable that "starts" the stepwise solution. Because of this conditionality, the stepwise regression process does not answer the question that it claims to answer, that is, "what is the best set of independent variables to predict variance in a particular dependent variable?" In fact, it is likely if stepwise regression methods are used that (a) other models with the same number of independent variables may have a larger R^2; (b) models with fewer independent variables may predict an equivalent R^2 as the models selected by the stepwise solution; (c) independent variables not included in the stepwise solution may be just as significant independent variables as those that are included; and (d) the independent variables will not enter the model in the order of importance that they would in a final model of independent variables selected simultaneously (Thayer, 1990).

Lastly, and perhaps most importantly, selection of variables in stepwise solutions capitalizes hugely on sampling error (Thompson, 1995). An independent variable that contributes an amount of variance that explains an extremely miniscule amount of greater variance than another variable due to sampling error may be chosen as the "best" predictor for a particular step, and all following variable selections will thereby result from a "chance" selection at this previous step. Due to these three major problems with the stepwise regression method, we generally proscribe use of this technique in assessing variable importance. We dedicated considerable discussion to these issues because the method remains in popular use by researchers.

Illustration

Below we present an example of a results section that employs all variable importance measures discussed in our guidebook to interpreting MR results, in hopes that this can serve as a template for other researchers to use in their own work. Data for this example was obtained from the Holzinger and Swineford (1939)'s dataset. This dataset contains assessments of 301 subjects from two high schools (Paster School and White Grant School) on a battery of 26 tests that measured verbal, spatial, and mathematical abilities. This example used three tests relevant to subjects' mathematical aptitude: (a) *numeric*, (b) *arithmetic*, and (c) *addition* to predict *deductive mathematical reasoning* (i.e., *reasoning*).

The regression equation for this analysis was: *reasoning* = -2.92 + 1.43*numeric* + 0.89*arithmetic* – 0.12*addition*. A comparison across all statistics presented in Table 2 highlighted that *numeric* was the strongest direct predictor of *reasoning* across multiple

Table 2. Summary of Statistics Determining Independent Variable Contributions to Regression Effects

Variable	β	r_s	r_s^2	r	Pratt	Unique	Common	GDW	RWI	RWI %
Numeric	.35	.87	.76	.40	.140	.092	.065	.124	.124	60.0
Arithmetic	.23	.72	.52	.33	.075	.038	.070	.072	.073	35.5
Addition	-.16	.09	.01	.04	-.006	.022	-.020	.011	.009	4.5

indices. *Numeric* obtained the largest beta weight (β = .35, $p < .001$), demonstrating that it made the largest contribution to the regression equation, while holding all other predictor variables constant. The zero-order correlation of *numeric* with the *reasoning* (r = .40), when squared, showed that *numeric* shared the largest amount (16%) of its variance with *reasoning*. The squared structure coefficient ($r_s^2 = .76$) demonstrated that *numeric* explained the largest amount (76%) of the variance in $\hat{y}$, the predicted values of *reasoning*. Product measure results demonstrated that *numeric* accounted for the largest partition of variance in *reasoning* (.140, 67.6% of the regression effect) when multiplying the beta weight (.35) by the zero-order correlation (.40). Notably, relative weight results supported that *numeric* explained a large portion of the overall regression effect (.124, 60%) when partitioning that effect based on creation of variables' uncorrelated or independent counterparts. Dominance analysis results (see Table 4) demonstrated complete dominance of *numeric* over *arithmetic* and *addition*, as it contributed more unique variance in the regression effect than the other two independent variables across all 3 MR sub-models that include that variable. This complete dominance can be determined by looking across each row in Table 4 and seeing how each value for *numeric* was larger than the values for *arithmetic* and *addition*.

Table 3. Commonality Coefficients

Effect	Coefficient	Percent
Unique to Numeric	.092	44.5
Unique to Arithmetic	.038	18.5
Unique to Addition	.022	10.7
Common to Numeric and Arithmetic	.075	36.2
Common to Numeric and Addition	-.015*	-7.4*
Common to Arithmetic and Addition	-.011*	-5.3*
Common to Numeric, Arithmetic, and Addition	.006	2.9
Total	.207	100.0

*Negative values represent suppression effects

Table 4. Complete Dominance Weights

Variable(s)	Model R^2	Additional Contribution of: Numeric	Arithmetic	Addition
Subset Containing No Predictors		.158	.108	.002
Numeric	.152		.027	.011
Arithmetic	.108	.073		.008
Addition	.002	.167	.113	
Numeric and Arithmetic	.185			.022
Numeric and Addition	.169		.038	
Arithmetic and Addition	.115	.092		
Numeric, Arithmetic, and Addition	.207			

Because *numeric* was a completely dominant independent variable over *arithmetic* and *addition*, it necessarily showed conditional dominance (see Table 5) over *arithmetic* and *addition*. This conditional dominance can be determined by looking across all rows in Table 5 and seeing how *numeric* contributed more unique variance on average to regression effects for models of all subset sizes than *arithmetic* and *addition*. As *numeric* was completely and conditionally dominant over the other two predictors in the model, it also showed general dominance over both of these variables (see Table 2). These statistics are equal to the averages of the conditional dominance weights shown in Table 5. Notably, for this example, the relative weight (.124) was equivalent to the general dominance weight (.124); thus, its contribution to the regression effect assessed in terms of averages of unique variance contributions to all possible subsets or through the creation of uncorrelated counterpart variables was the same.

Table 5. Conditional Dominance Weights

Subset Size	Numeric	Arithmetic	Addition
0	.158	.108	.002
1	.122	.070	.009
2	.092	.038	.022

Arithmetic clearly emerged as the second strongest direct predictor of *reasoning*. In terms of the beta weight ($\beta = .23$, $p < .001$), it made the second largest contribution to the regression equation when holding all other predictors constant. Its zero-order correlation ($r = .33$) was also the second largest in the model, which, when squared, demonstrated that *arithmetic* shared the second largest amount (10.9%) of its variance with *reasoning*. The squared structure coefficient ($r_s^2 = .52$) illustrated that *arithmetic* shared the second largest amount (52%) of variance with $\hat{y}$. Product measure (.075) results showed that *arithmetic* accounted for 36.2% of R^2 when it was partitioned based on multiplying the beta weight (.23) by the zero-order correlation (.33). *Arithmetic*'s relative weight (.073) was nearly identical to its product measure, demonstrating that *arithmetic* accounted for 35.5% of the regression effect when partitioning it based on creation of the independent variables' uncorrelated counterparts. Thus, *arithmetic* accounted for the second largest amount of variance in the regression equation across multiple measures.

Complete dominance analysis results (see Table 4) supported the fact that *arithmetic* was completely dominant over *addition* (see Table 4), as a perusal of all rows in Table 4 show that it contributed more unique variance to all rows of all subsets sizes than *addition*. As *arithmetic* was completely dominant over *addition*, it also showed conditional dominance (see Table 5) in and general dominance (see Table 2) over this variable. Once again, the general dominance weight for *arithmetic* was nearly identical to the relative weight, reflecting that *arithmetic* explained the second largest and a substantial 35.5% of the variance in R^2 when partitioning the regression effect based on the average difference in fit between subsets that do and do not include *arithmetic*—a different partitioning method than that used for relative weights.

Although other statistics clearly showed how *numeric* was the strongest direct contributor to the regression equation, followed by *arithmetic*, they did not show exactly how those variables contributed unique and shared variance to the regression equation. Thus, commonality coefficients were consulted to obtain this information (see Table 3). When viewing commonality analysis results, unique effect results demonstrated that *numeric* contributed more unique variance (44.5%) to the regression effect than did *arithmetic* (18.5%). The common effect for *numeric* and *arithmetic* reflected that both variables also contributed substantial shared variance (36.2%) to the regression effect. These results highlighted that *numeric* and *arithmetic* partially operated in combination in predicting *reasoning*. Viewing unique and common effect columns in Table 2 illustrated how *numeric* contributed more unique than shared variance to variance in *reasoning* (.09 vs. .07, respectively), while, conversely, *arithmetic* contributed more shared than unique variance to variance in the dependent variable (.04 vs. .07, respectively).

The results for different measures presented different *lense*s on the predictive power of *addition* to predicting variance in the dependent variable. The beta weight for *addition* (β = -.16, *p* = .004) suggested that addition played a more minor but still statistically significant role in the regression effect when holding all other variables constant. However, the squared structure coefficient (r_s^2 = .001) showed that *addition* contributed negligible variance to $\hat{y}$, and the zero-order correlation (*r* = .04) showed that *addition* explained 0.2% of the variance in the overall dependent variable. These statistics clearly demonstrated that *addition* did not directly share variance with either the obtained effect or the dependent variable as a whole. When partitioning R^2 based on multiplying the beta weight (-.16) by the zero-order correlation (.02), product measure (-.006) results reflected that *addition* once again contributed very little variance to the regression effect (-2.9%). The relative weight for *addition* (.009) demonstrated that *addition* contributed very little variance the regression effect (4.5%) when partitioning it based on variables' uncorrelated counterparts. The fact that the beta weight for *addition* was statistically significant but that *addition* shared little variance with *reasoning* when examining (a) its shared variance with $\hat{y}$, (b) its shared variance with the dependent variable, and (c) partitioning the regression effect based on two measures suggested that *addition* was a suppressor variable, which removed irrelevant variance in at least one independent variable, thereby allowing its (or their) contributions to the regression effect to become larger.

Conditional dominance weights confirmed that *addition* was a suppressor, as its contribution to regression models increased across subset sizes. Its contribution to the subset containing no other predictors was .002, which illustrated that it contributed nearly negligible variance to the regression effect in isolation from other variables in the regression equation. Its average unique variance contribution for the models containing only one of the other two predictors was .009, and for the subset containing both other predictors it was .022. Notably, *addition*'s contribution to the regression model containing both *numeric* and *arithmetic* was greater than its contribution to models containing either *numeric* or *arithmetic*, suggesting that *addition* might be a suppressor for both *numeric* and *arithmetic*.

Commonality coefficient findings supported that *addition* was a suppressor for both *numeric* and *arithmetic*, as the common effects between *numeric* and *addition* (-.015) and *arithmetic* and *addition* (-.011) were negative, and suppression is generally demonstrated in negative commonalities. Summing the percentages of variance that each of these common effects contributes to the overall regression effect (7.4% for *numeric* and 5.3%, for *addition*, respectively) showed that 12.7% of the regression effect is due to suppression, and thus that addition removes 12.7% of irrelevant variance from *numeric* and *arithmetic* to increase the amount of variance in the dependent variable explained by these two independent variables by 12.7%.

Overall, these findings supported how both *numeric* was the most significant direct contributor and *arithmetic* was the second most important direct contributor to predicting variance in *reasoning*, as reflected across different measures of *direct, total*, and *partial effects*. It is important to note that this is not always the case: One independent variable may be deemed the most important through one *lens*, and another independent variable may achieve that status through another *lens*. Results also supported from multiple lenses how *addition* functioned as a suppressor in this regression equation. Reliance on beta weights alone would not have pointed out the nature of the suppressor effect.

Table 6 is offered as a summary of the purposes of each of the discussed statistics.

List of Recommendations for Practice

We present a list of recommendations for reporting of regression results based on the methods discussed in this guidebook to assist researchers in selection and synthesis of multiple variable importance measures. Prior to outlining our recommendations, we would like to state a general recommendation to run scatterplots of all independent variables and the dependent variable to find outliers and/or unexpected patterns in the data prior to conducting MR analyses. Clean up the data

Table 6. Purposes of Each Statistical Measure

Statistic	Purposes of Measure
Beta weight	Determination of independent variables' contributions to prediction within a linear regression equation while holding all other independent variables constant
Zero-order *r*	Determination of magnitude and direction of bivariate linear relationship between each independent variable and the dependent variable
Product measure	Determination of variable importance based on partitioning model R^2 as a function of a predictor's beta weight multiplied by its zero-order *r*
Squared structure coefficients	Determination of how much variance each independent variable contributes to y
Commonality coefficients:	
Unique effects	Determination of variance each independent variable contributes to a regression equation that is not shared with other independent variables (squared semipartial correlation)
Common effects	Determination of which independent variables share variance in predicting the dependent variable as well as quantification of how much variance is shared between independent variables
Dominance weights	Determination of whether one independent variable contributes more variance than another independent variable to models: (a) across all subsets of independent variables (complete dominance) (b) on average across models for all subset sizes (conditional dominance) (c) on average across all models (general dominance)
Relative weights	Determination of variable importance based on method that addresses multicollinearity by creating variables' uncorrelated "counterparts"

as appropriate before proceeding with summarizing them with regression and other statistical analyses. Also, previous articles have recommended including tables of means, standard deviations, and intercorrelations between independent variables to accompany regression summary tables (cf. Schafer, 1991). Reporting reliability coefficients for all variables can also be helpful (e.g., independent variables with low alpha reliability might explain low intercorrelations and weak prediction). Previous authors have also recommended (a) including the specific regression equation(s) in a footnote or in text (i.e., do not merely list the total R^2 or the statistical significance of the equation; cf. Schafer, 1992) as well as (b) including a general regression summary table that reports the values of the zero-order correlation between each independent variable and the dependent variable, F and $MS_{residual}$ values, and p values for each individual predictor to illustrate statistical significance (cf. Schafer, 1992).

Specific to reporting multiple indices of variable importance, we offer following recommendations for practice:

1) Do not rely solely upon beta weights when interpreting regression findings, except in the case of uncorrelated predictors or when the model is perfectly specified (see, e.g., Courville & Thompson, 2001). Beta weights are only informative with regard to prediction; they do not tell the researcher other important information provided by the other metrics we have reviewed.
2) Use different tables to help interpret different indices of variable importance.
   a. Include one table that enables visual comparisons across indices for each independent variable. Joint comparisons across indices can aid in identification of associations between variables and the presence of suppression in a regression equation. Table 2 of this paper is recommended if all indices are compared. If specific subsets of indices are used for a specific research purpose, then one can still

present those indices in the format we display.

   b. If commonality analysis is used, include a table (see Table 3) that lists the unique variance contributions of each independent variable and the common variance contributions for all possible subsets of independent variables to the regression equation. Also report these portions of R^2 as percentages of R^2, as well.
   c. If dominance analysis is used, include a table (see Table 4) of unique variance contributions of each independent variable to all possible subset sizes to enable pairwise comparisons of unique variance contributions of each independent variable to all subsets of predictors. Additionally, include a table (see Table 5) of conditional dominance weights that averages unique variance contributions of each independent variable to models of all subset sizes. This table is useful in identifying suppression effects if values increase across subset sizes with more independent variables.

3) When reporting commonality analysis results in text, describe both the unique and shared variance contributions of all independent variables <u>and</u> whether each variable contributes more shared or unique variance in its contribution to R^2.
4) When reporting dominance analysis results in text, include all three types of dominance (complete, conditional, and general), as dominance relationships can be established at lesser levels of dominance (i.e., general dominance) if not at higher levels (i.e., complete dominance).
5) Calculate squared structure coefficients when independent variables are correlated in order to determine the role of shared variance in the regression equation (cf. Courville & Thompson, 2001). If there is divergence between beta weight and structure coefficient results (supportive of multicollinearity/associations between variables and/or suppression), it is suggested that researchers employ commonality analysis to determine the location and extent of this shared variance and/or suppression.
6) Include a variance partitioning statistic in addition to beta weights in the presence of correlated predictors (e.g., general dominance weights, relative weights, and Pratt's measure). Draw comparisons in text between techniques that partition R^2 if multiple techniques are used.
7) In the presence of suppression in a regression equation, always calculate commonality coefficients, as this technique is uniquely able to identify which variables are being suppressed and quantify the overall amount of suppression present in an equation.
8) Always consider the statistics used in calculating each index when evaluating variable importance. For example, if a Pratt measure value is near-zero, the researcher should verify that the zero-order correlation used in computing the Pratt measure and not the beta weight shows a near-zero or zero value when determining if a variable is a suppressor, as suppression is only demonstrated in the former case. The researcher should also consider the value of R^2 that is expressed as percentages in commonality analysis and relative weights analysis. It is reasonable to think that percentages of a small R^2 might be interpreted differently than the same percentages based on a large R^2.

Conclusion

This paper has illustrated how researchers' conceptualizations and assessments of variable importance can be enhanced by viewing MR results through multiple lenses. As it has shown that each "lens" has distinct advantages and limitations, and that multiple statistical measures complement each other in the perspectives they provide regarding regression findings, we hope that this paper will encourage researchers to "look beyond" beta weights and employ the other measures discussed in our guidebook. Ideally, this practice would become a matter of routine among researchers, in our peer-reviewed journals, and in teaching MR within graduate-level statistics curricula. The hope is that our data-driven example will allow researchers to write up their own findings in a similar manner and thus will be able to better represent the richness of their regression findings, and how these findings are impacted by such issues as suppression and patterns of shared variance that go undiscovered through heavy reliance on beta weights alone.

References

Amado, A. J. (1999, January). *Partitioning predicted variance into constituent parts: A primer on regression commonality analysis*. Paper presented at the annual meeting of the Southwest Educational Research Association, San Antonio, TX.

Azen, R., & Budescu, D. V. (2003). The dominance analysis approach for comparing predictors in multiple regression. *Psychological Methods, 8(*2), 129-148. doi: 10.1037/1082-989X.8.2.129

Baltes, B. B., Parker, C. P., Young, L. M., Huff, J. W., & Altmann, R. (2004). The practical utility of importance measures in assessing the relative importance of work-related perceptions and organizational characteristics on work-related outcomes. *Organizational Research Methods, 7*(3), 326-340. doi: 10.1177/1094428104266014

Budescu, D. V. (1993). Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. *Psychological Bulletin, 114*(3), 542-551.

Capraro, R. M. & Capraro, M. M. (2001). Commonality analysis: Understanding variance contributions to overall canonical correlation effects of attitude toward mathematics on geometry achievement. *Multiple Linear Regression Viewpoints, 27*(2) 16-23.

Courville, T., & Thompson, B. (2001). Use of structure coefficients in published multiple regression articles: ß is not enough. *Educational and Psychological Measurement, 61*, 229-248.

Darlington, R. B. (1968). Multiple regression in psychological research and practice. *Psychological Bulletin, 69*(3), 161-182.

DeVito, P. J. (1976, May). *The use of commonality analysis in educational research*. Paper presented at the annual meeting of the New England Educational Research Association, Provincetown, MA.

Fabbris, L. (1980). Measures of predictor variable importance in multiple regression: An additional suggestion. *Quality & Quantity, 4*, 787–792.

Genizi, A. (1993). Decomposition of R^2 in multiple regression with correlated regressors. *Statistica Sinica*, *3,* 407-420.

Grömping, U. (2007). Estimators of relative importance in linear regression based on variance decomposition. *The American Statistician, 61*, 139-147.

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Applied statistics for the behavioral sciences* (5th ed.). Boston, MA: Houghton Mifflin.

Holzinger, K. J., & Swineford, F. (1939). *A study in factor analysis: The stability of a bi-factor solution* (No. 48). Chicago, IL: University of Chicago.

Hoyt, W. T., Leierer, S., & Millington, M. (2006). Analysis and interpretation of findings using multiple regression techniques. *Rehabilitation Counseling Bulletin, 49*, 223-233.

Johnson, J. W. (2000). A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate Behavioral Research, 35*(1), 1-19.

Johnson, J. W., & LeBreton, J. M. (2004). History and use of relative importance indices in organizational research. *Organizational Research Methods, 7*(3), 238-257. doi: 10.1177/1094428104266510

LeBreton, J. M., Ployhart, R. E., & Ladd, R. T. (2004). A Monte Carlo comparison of relative importance methodologies. *Organizational Research Methods, 7*(3), 258-282. doi: 10.1177/1094428104266017

Mayeske, G. W., Cohen, W. M., Wisler, C. E., Okada, T., Beaton, A. E., Proshek, J. M., Weinfeld, F. D., & Tabler, K. A. (1969). *A study of our nation's schools.* Washington, DC: U.S. Department of Health, Education, and Welfare, Office of Education.

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

Mood, A. M. (1971). Partitioning variance in multiple regression analyses as a tool for developing learning models. *American Educational Research Journal*, *8*, 191-202.

Newton, R. G., & Spurrell, D. J. (1967). A development of multiple regression for the analysis of routine data. *Applied Statistics, 16*, 51-64.

Nimon, K. (2010). Regression commonality analysis: Demonstration of an SPSS solution. *Multiple Linear Regression Viewpoints, 36*(1), 10-17.

Nimon, K., Gavrilova, M., & Roberts, J. K. (2010). Regression results in human resource development research: Are we reporting enough? In C. Graham & K. Dirani (Eds.), *Proceedings of the Human Resource Development 2010 International Conference* (pp. 803-812), Knoxville, TN: AHRD.

Nimon, K., Lewis, M., Kane, R., & Haynes, R. M. (2008). An R package to compute commonality coefficients in the multiple regression case: An introduction to the package and a practical example. *Behavior Research Methods*, *40*, 457-466. doi: 10.3758/BRM.40.2.457

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill, Inc.

Onwuegbuzie, A. J., & Daniel, L. G. (2003, February 19). Typology of analytical and interpretational errors in quantitative and qualitative educational research. *Current Issues in Education* [On-line], *6*(2). Available: http://cie.ed.asu.edu/volume6/number2/

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Stamford, CT: Thompson Learning.

Pratt, J. W. (1987). Dividing the indivisible: Using simple symmetry to partition variance explained. In T. Pukkila & S. Puntanen (Eds.), *Proceedings of the second international Tampere conference in statistics* (pp. 245-260). Tampere, Finland: University of Tampere.

Rowell, R. K. (1996). Partitioning predicted variance into constituent parts: How to conduct regression commonality analysis. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 4, pp. 33-43). Greenwich, CT: JAI Press.

Schafer, W. D. (1991). Reporting hierarchical regression results. *Measurement and Evaluation in Counseling and Development*, *24*, 98-100.

Schafer, W. D. (1992). Reporting nonhierarchical regression results. *Measurement and Evaluation in Counseling and Development*, *24*, 146-149.

Thayer, J. D. (1990, April). *Implementing variable selection techniques in regression.* Paper presented at the annual meetings of the American Educational Research Association, Boston, MA.

Thayer, J. D. (2002, April). *Stepwise regression as an exploratory data analysis procedure.* Paper presented at the annual meetings of the American Educational Research Association, New Orleans, LA.

Thomas, D. R., Hughes, E., & Zumbo, B. D. (1998). On variable importance in linear regression. *Social Indicators Research: An International and Interdisciplinary Journal for Quality-of-Life Measurement, 45*, 253-275.

Thompson, B. (1989). Why won't stepwise methods die?. *Measurement and Evaluation in Counseling and Development, 21*, 146-148.

Thompson, B. (1995). Stepwise regression and stepwise discriminant analysis need not apply here: A guidelines editorial. *Educational and Psychological Measurement, 55*(4), 525-534.

Tonidandel, S., & LeBreton, J. M. (2010). Determining the relative importance of predictors in logistic regression: An extension of relative weight analysis. *Organizational Research Methods, 13*(4), 767-781. doi: 10.1177/1094428109341993

Tonidandel, S., & LeBreton, J. M. (2011). Relative importance analysis – A useful supplement to regression analyses. *Journal of Business and Psychology, 26,* 1-9. doi: 10.1007/s10869-010-9204-3

Zientek, L. R., Capraro, M. M., & Capraro, R. M. (2008). Reporting practices in quantitative teacher education research: One look at the evidence cited in the AERA panel report. *Educational Researcher, 34*, 208-216. doi: 10.3102/0013189X08319762

Zientek, L. R., & Thompson, B. (2006). Commonality analysis: Partitioning variance to facilitate better understanding of data. *Journal of Early Intervention, 28*, 299-307. doi: 10.1177/105381510602800405
.

Citation:

Nathans, Laura L., Oswald, Frederick L., & Nimon, Kim (2012). Interpreting Multiple Linear Regression: A Guidebook of Variable Importance. *Practical Assessment, Research & Evaluation*, 17(9). Available online: http://pareonline.net/getvn.asp?v=17&n=9

Corresponding Author:

Laura L. Nathans
University of North Texas
laura_nathans [AT] yahoo.com