Predictor Importance for Selection: Best Practices and Latest Findings

A discussion of conventional wisdom surrounding predictor importance and selection systems

Written by

Dan Hawthorne

Cogniphany

A key issue when designing a selection system is how to determine which selection tool(s) contribute(s) the most to the prediction of the outcome or criterion of interest (e.g. performance on the job). Johnson (2000) provides a good definition of relative importance: "the contribution each variable makes to the prediction of a dependent variable considering both its unique contribution and its contribution, when combined with other variables" (p. 1). An additional clarification is provided by LeBreton (LeBreton, 2007, Slide 6), when he states "Relative importance refers to the proportionate contribution each predictor makes to R^2 considering both its individual effect and its effect when combined with the other variables in a regression equation."

Determining relative importance is essential in most circumstances for maximizing selection system utility. Relative importance is not necessary when a single predictor is used because there are not two variables to relate to one another; however, many organizations generally choose to use a multiple predictor system for two of reasons. First, employers might balk at hiring an applicant based on a single predictor (e.g. only using a résumé and no interview). Secondly, multiple predictors can increase the criterion-related validity of the overall selection system. For example, Schmidt and Hunter (1998) noted, combining a structured interview (r = .51) and a cognitive ability test (r = .51) yields a combined validity of .63. Thus, accurate determination of the best combination of predictors is of utmost importance. Additionally, each selection device may have different challenges associated with it (e.g. adverse impact, significant cost, negative applicant reactions, etc.). So, knowledge of which devices have the most predictive ability relative to other predictors is critical when assessing whether or not to add a device to a selection system.

The first debates about the determination of relative importance began to appear in the 1960s (Darlington, 1968). More recently, VanIddekinge and Ployhart (2008) provided a comprehensive review of strategies for assessing relative importance. They suggest there are 4 major strategies to determine relative predictor importance: (1) regression coefficients, (2) incremental validity, (3) dominance analysis, and (4) relative weight analysis. This paper explains these strategies for the determination of relative predictor importance, and offers practical suggestions for determining which strategy is appropriate.

Strategies to Determine Relative Predictor Importance

As noted above, there are four main strategies to determine relative predictor importance. Until approximately a decade ago, only two options were available: examination of regression coefficients and analysis of correlation coefficients. More recently two additional options have been created: dominance analysis and relative weight analysis. A discussion of both the historically used methods and the more recently created methods follows.

Historically Utilized Strategies for Assessing Relative Importance. Hoffman (1960) presented evidence that the products of each predictor variable's standardized regression coefficient (β_x) and its zero-order correlation (r_{yx}) with the criterion variable summed to R^2. Hoffman further stated that this represented the variable's "**independent** contribution of each predictor" (p.120; emphasis added). Following his seminal work, researchers and practitioners began utilizing regression coefficients and zero-order correlations to evaluate selection systems.

Regression Coefficients**.** The first strategy—and historically the most commonly used—is the magnitude and statistical significance of regression coefficients. A regression coefficient is a constant in a regression equation that represents the rate of change in one variable (criterion) as a function of another variable (predictor). Thus, any predictor variable that expresses a significant regression coefficient can be considered to be an important predictor and its relative importance can be determined by the magnitude and sign of the coefficient value. These coefficients can be interpreted as either unstandardized or standardized values. Unstandardized regression coefficients indicate the predicted change in the criterion variable given a one unit change in the predictor variable, but unstandardized coefficients of multiple predictors cannot be compared directly because of their different units of measurement. Standardization of the regression coefficient—changing the unit of measure to a mean of 0 and variance of 1—allows for easier comparison of predictor variables that were measured on different scales.

Thus, when evaluating a selection system for predictor utility, it is more desirable to use standardized regression coefficients so it is possible to know which predictors provide stronger relative importance. Organizations may weigh these predictors with more relative importance in order to improve a selection system's ability to identify applicants with high performance potential. Conversely, attempting to compare unstandardized regression coefficients would be difficult and confusing because the variables would be compared on a different metric. Thus, any weighting with unstandardized coefficients would also be misleading, because scores would exhibit different values (e.g. 1-100 score range on one predictor versus 1-5 score range on another predictor). In sum, if a regression coefficients strategy is used to determine relative importance of multiple predictors, standardized regression weights should be used.

While the logical and methodological simplicity of this strategy may be considered useful, the strategy does possess an inherent weakness. In his seminal work, Hoffman (1960) claimed that his conception of relative importance measured "**independent** contribution of each predictor" (p. 120; emphasis added). However, other researchers called into question Hoffman's assertion of "independent" contribution of each predictor because it inferred that every other predictor's influence is held constant in the model (Ward, 1962). Hoffman (1962) was forced to reply that his conception of relative weights did not measure independent contribution in that sense.

Essentially, if predictors are uncorrelated or orthogonal, standardized regression coefficients equal zero-order correlations, and if the squared regression coefficients are summed, they equal R^2. However, when predictors are correlated (as they almost always are in selection systems), a change in one predictor variable will almost assuredly result in a change in all other correlated predictors, and summing squared regression coefficients will no longer equal R^2, making decomposition of the effects difficult or near impossible (LeBreton, 2007). That is, estimates of importance from regression coefficients use total effects, but ignore partial and direct effects, masking the effects of correlated predictors (LeBreton, Ployhart, & Ladd, 2004). Consequently, indices of predictor importance that do not consider the existing relationship between all of these different variables can be misleading (Johnson & LeBreton, 2004).

Some suggestions have been made to address this problem. For example, standardization of coefficients using a partial standard deviation that controls for correlations in other predictors has been suggested, but this approach still does not consider the partial effects (Bring, 1994). Thus, it is common to report these as indices of relative predictor importance, but it may be unwise to rely solely on them because they yield highly divergent results compared to more recently developed strategies have demonstrated more accuracy in Monte Carlo simulations (Johnson & LeBreton, 2004; LeBreton, et al., 2004). Specifically, because regression coefficients control for, rather than expose partial effects, they can create confusion around which one in a set of correlated predictors is best, making relative ranking ineffective.

Correlation Coefficients The second major strategy, which is frequently used in conjunction with the first strategy, is the examination of incremental validity in the form of correlations (Azen & Budescu, 2003). This strategy, like the first strategy, is also simple and straightforward in its logic because if one variable is statistically related to another (i.e. criterion variable), then logically it would be important. Essentially, this index gives an indication of how much more of the variance in a criterion variable can be attributed to the inclusion of the selected predictor variable (e.g. increase in R^2). However, there are also inherent weaknesses involved in this strategy. A major weakness is that similar to regression coefficients, correlation coefficients do not consider partial and total effects of relationships between predictors and criterion variables (LeBreton, et al., 2004). Essentially, correlation coefficients cannot partition the variance shared between multiple predictors that should be attributed to each predictor. Consequently, squared correlations will only sum to the model's R^2 if the predictors are all uncorrelated (LeBreton, 2007). Additionally, model order entry can affect predictor importance, which can mask a predictor's relative importance. For example, if X_2 is entered after X_1 and they are both highly correlated, X_1 will look more important, but if the order is reversed then X_2 will look more important.

Weaknesses Associated with Both Strategies. Even though the first two approaches are the most commonly used in research and in practice through selection system design, both share weaknesses that warrant consideration (Azen & Budescu, 2003). Overall, these two indices of predictor importance are difficult to interpret due to model order entry and information that cannot be accounted for in the procedures. Regression coefficients (both unstandardized and standardized) include the effects of all of the other variables in the equation. That is, ordinary least squares (OLS) is used to describe the line of best fit and how much change will occur in a predictor with change in the criterion while holding all other variables constant. Likewise, strategies examining correlations ignore the effects of all other predictors. That is, they describe the relationship between the predictor and criterion in isolation of all other variables. LeBreton and colleagues (2004) performed a Monte Carlo study and found that as the mean validity of the predictors, amount of predictor collinearity, or the number of predictors increased (beyond 3), the interpretability of the beta and correlation coefficients suffered seriously due to coefficient instability. Thus, they warned against using either of these first two strategies in isolation, and instead advocated the use of Dominance Analysis (DA, Azen & Budescu, 2003; D. V. Budescu, 1993) or a form of Relative Weight Analysis (RWA), the epsilon (ε) statistic (Johnson, 2000).

Recently Developed Strategies for Assessing Relative Importance. While many researchers and practitioners continue to use R^2, β, and correlations as indices of predictor importance, research over approximately the last decade provides strategies to more accurately assess relative importance (e.g. D. V. Budescu, 1993). These strategies include DA and RWA.

Dominance Analysis. Dominance Analysis (DA) attempts to account for every possible pairwise combination of predictor variables in the prediction of a criterion variable in an effort to identify a pattern of prediction dominance exhibited by the most important variables (Azen & Budescu, 2003; D. V. Budescu, 1993). Unlike the first two strategies that start with a particular statistic then try to apply meaning to the number, DA attempts to provide clarity to the issue of importance in a straightforward manner using a strategy designed especially for assessing importance. Specifically, the correlation coefficient strategy gives a single statistic to denote the importance of a predictor, while DA produces a table giving a detailed summary of how each predictor contributes differently to all subset models (D. Budescu & Azen, 2004). For example, if a potential full model examines three predictors (e.g. extraversion, emotional stability, and agreeableness), a DA summary table will show all possible combinations of these variables and the average relative predictive indexes (based on squared semipartial correlations) for each model, as well as overall averages for each predictor (see table below). Budescu (1993) originally covered only complete dominance (e.g. the variable is the most important predictor in the entire set of predictors) and undetermined (e.g. complete dominance cannot be determined). However, Azen and Budescu (2003) created a hierarchy of dominance categories: (1) complete dominance, (2) conditional dominance (the variable is the most dominant predictor depending on the subset included in the model), and (3) general dominance (the overall average dominance index of a particular variable in a given subset of variables). Thus, the pattern of dominance can show which variables have the most importance in each subset of the model, as well as within the entire model (LeBreton, 2007).

(DV = JOBSAT)	RSq	EXTRA	STABLE	AGREE
None		0.0478	0.1296	0.0308
EXTRA	0.0478		0.1058	0.0273
STABLE	0.1296	0.0241		0.0021
AGREE	0.0308	0.0443	0.1009	
EXTRA STABLE	0.1536			0.0012
EXTRA AGREE	0.0751		0.0797	
STABLE AGREE	0.1317	0.0231		
General Dominance Weights:		0.0350	0.1042	0.0156
Rescaled Weights:		0.2263	0.6731	0.1006

Note: Adapted from LeBreton (2007)

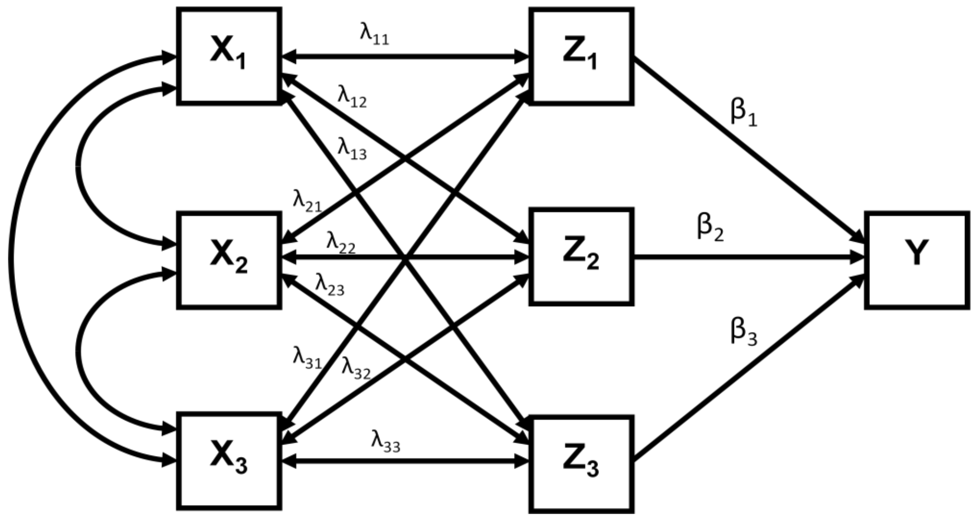
One of this approach's strengths is that DA can detect and identify suppressor variables because the DA table will show a negative dominance index, instead of it being masked as in other strategies. An additional strength of this approach is the ability to perform a constrained DA. That is, one predictor variable can be constrained as being necessary to a model, to find which predictors are the most important and best complement the variable required in the model (Azen & Budescu, 2003). This can be a benefit to a practitioner redesigning a selection system who has been given direction that a specific tool(s) must remain in the system. A final strength of this approach is that importance estimates can be divided by the R^2 to calculate the percentage of explainable variance in the overall R^2 by a given predictor variable (LeBreton, 2007).

A weakness of this strategy is that DA is not designed to address the hierarchical order of predictors. This can be done through multiple runs of a constrained DA, but is more difficult to interpret due to the amount of rendered data (Azen & Budescu, 2003). Additionally, as the number of analyzed predictors increases, DA becomes more computationally difficult because of the exponentially increasing number of submodels involved (Johnson & LeBreton, 2004). This is because for *p* predictors, there are 2^p

– 1 submodels. For example, with three predictors a DA summary table will have seven possible models (i.e. (1) X_1, (2) X_2, (3) X3, (4) X_1 & X_2, (5) X_2 & X_3, (6) X_1 & X_2, (7) X_1, X_2, & X_3). Likewise, 10 predictors will yield 1023 possible submodels.

Relative Weights Analysis. The final strategy for identifying relative importance of predictors is RWA (Johnson, 2000). This strategy involves variable transformation of the original predictors into orthogonal (uncorrelated) variables that are related to criterion variables, but not to each other, which are then related back to the original predictors (Van Iddekinge & Ployhart, 2008). This strategy attempts to deal with the problem inherent with many of the strategies: if predictor variables are correlated, they influence all other derived relative importance indices. Historically, one of the major weaknesses of this strategy was in how the orthogonal variables were calculated. The regression weights used to create them are still coefficients from regressions on correlated variables. Thus, because the new supposedly orthogonal variables were regressed upon the original correlated predictors, this reintroduced the problem of correlated variables back into the equation.

Johnson's (2000) epsilon statistic (ε_j) addresses this weakness and has become one of the more suggested approaches to relative weights analysis (Johnson & LeBreton, 2004; Van Iddekinge & Ployhart, 2008). An epsilon statistic is calculated for each predictor variable that is its relative importance. Epsilon can also be easily transformed into a statistic that can be interpreted as the percentage of the model R^2 associated with each predictor. The figure below provides a detailed illustration of how the epsilon statistic is calculated. The X_j variables represent three predictor variables, which are transformed into variables (represented by the Z_k variables) that are as related to the criterion variable as possible, but uncorrelated with the original predictor variables. The Z variables are then used to predict the criterion variable Y, and the regression coefficients for this are represented by β_k. Finally, the regression coefficients for X_j on Z_k are represented by λ_{jk}. Since the new Z variables are uncorrelated, the regression coefficients λ_{jk} are equal to the correlation coefficients between X_j and Z_k. So, any λ_{jk}^2 is equal to the proportion of variance in Z_k that is accounted for by X_j. To calculate the epsilon statistic for any predictor, the proportion of variance accounted for each Z_k by each X_j is multiplied by the proportion of variance accounted for by each Z_k in Y, and then all of the products are summed. In other words, the original predictors are regressed on the orthogonal variables, instead of regressing the orthogonal variables on the original predictors. For example, to calculate epsilon for predictor variable 1 (X_1) in the figure, the following equation would be used: $\varepsilon_1 = \lambda^2_{11}\beta^2_1 + \lambda^2_{12}\beta^2_2 + \lambda^2_{13}\beta^2_3$.



A major strength of Johnson's (2000) epsilon statistic over DA is that its ease of computation with an unlimited number of predictors. Furthermore, this calculation eliminates the problem that earlier attempts at RWA had with correlated variables. RWA also presents surprising convergent validity with DA. This is because each is based on different mathematical processes, but arrives at almost identical results (Johnson & LeBreton, 2004). Additionally, RWA and DA are generally no more computationally complex than conducting a regression analysis, and a number of researchers have made SAS and SPSS syntax available, as well as Excel spreadsheets to make the analyses more accessible (c.f. LeBreton, 2007; LeBreton, Hargis, Griepentrog, Oswald, & Ployhart, 2007). The primary weakness is that the relative weights are conducted on the full model containing all predictors, so looking for a pattern of relative predictor importance is difficult (LeBreton, 2007).

As an example of RWA's utility, LeBreton and colleagues (2007) reanalyzed data from Mount, Witt, and Barrick (2000), which had examined the incremental validity of biodata measures. LeBreton and colleagues conducted a RWA and found that the relative importance of a biodata-work habits measure was two times larger than the incremental importance reported by Mount, et al (2000). They also found that the RWA index for the biodata-problem solving measure was three times larger than its incremental importance, completely dominating all of the other predictors in the study. In the original study, the biodata scales accounted for small to moderate increases in R^2; using RWA, they not only added to the prediction, but were identified as the most important predictors of job performance.

Practical Recommendations. LeBreton, et al. (2007), presented a three-step set of guidelines for practitioners regarding relative importance of predictor variables. They suggest that first, all predictor variables should be examined for bivariate correlations. The rationale in this step is that if a predictor fails to have a significant correlation with the criterion, there is little reason to proceed further. Second, if predictor variables exhibit significant correlations with the criterion, a hierarchical regression analysis should be performed to examine the incremental variance of each predictor. *F* tests, examining the change in R^2 associated with each predictor can be conducted, or the statistical significance of the *t*-tests and unstandardized regression coefficients can be examined, as they are statistically identical to the same tests to acquire change in R^2. As a final step, RWA or DA should be conducted to find the relative weights or dominance weights for each predictor. That is, RWA or DA should be conducted as supplemental indices of relative predictor importance (Van Iddekinge & Ployhart, 2008). It is also suggested that if one wishes to see the pattern of dominance exhibited by different predictors in the full model or different subsets of the model that a DA can be conducted, because it gives pattern information that RWA cannot provide (LeBreton, 2007).

Research following LeBreton and colleagues (2007) present other pieces of information that should be presented in a relative importance analysis. First, research has developed methods for using RWA to assess multiple predictors on multiple criterion variables (LeBreton & Tonidandel, 2008). As suggested by LeBreton (2007) in his CARMA presentation on the topic, this could allow organizational practitioners to assess the relative importance of predictors on multiple outcomes, such as the examination of a new biodata measure on the job performance and turnover rates simultaneously, or on multiple aspects of an outcome measure (e.g. task performance, OCBs, and CWBs). Secondly, until recently there was no way to measure the statistical significance of relative weights. Now, researchers using RWA can calculate power and statistical significance instead of making inferences based on confidence intervals (Tonidandel, LeBreton, & Johnson, 2009). In addition to these two new pieces of information, relative weights (or general dominance weights) can be divided by R^2 to create rescaled relative importance weights that make for easier communication to human resource executives and other stakeholders. For example, it should be easier to convince an organizational decision-maker that a new selection device is valuable when, "it is described as accounting for 25% of the predictable variance in the criterion or outcome than […] when it is described as increasing R^2 by .03" (LeBreton, et al., 2007, p. 481).

Summary

Whether considering the use of multiple selection devices in a complete selection system, being able to accurately discern which predictors are the most important relative to other devices being considered is critical to the design. Regression coefficients and correlation correlations have been used in isolation for decades to determine relative predictor importance. However, newly designed procedures like DA and RWA have presented strong evidence of their ability to provide accurate information about relative importance without some of the inherent weaknesses of more traditionally accepted methods. Thus, in addition to reporting these traditional indices such as regression coefficients and incremental validity, we would be wise to follow the recommendations that several researchers have made about using RWA and DA as supplementary indications of which predictors are the most important to making selection decisions.

References

Azen, R., & Budescu, D. V. (2003). The dominance analysis approach for comparing predictors in multiple regression. *Psychological Methods, 8*(2), 129-148.

Bring, J. (1994). How to standardize regression coefficients. *The American Statistician, 48*(3), 209-213.

Budescu, D., & Azen, R. (2004). Beyond global measures of relative importance: Some insights from dominance analysis. *Organizational Research Methods, 7*(3), 341-350. doi: 10.1177/1094428104267049

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.

Darlington, R. B. (1968). Multiple regression in psychological research and practice. *Psychological Bulletin, 69*(3), 161-182. doi: 10.1037/h0025471

Hoffman, P. J. (1960). The paramorphic representation of clinical judgment. *Psychological Bulletin, 57*(2), 116-131. doi: DOI: 10.1037/h0047807

Hoffman, P. J. (1962). Assessment of the independent contributions of predictors. *Psychological Bulletin, 59*(1), 77-80. doi: DOI: 10.1037/h0044373

Johnson, J. (2000). A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate Behavioral Research, 35*(1), 1-19. doi: citeulike-article-id:4498196

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

LeBreton, J. (2007, October 26, 2007). *Relative importance of predictors with regression models.* Paper presented at the Center for Advancement of Research Methods and Analysis, Wayne State University.

LeBreton, J., Hargis, M., Griepentrog, B., Oswald, F., & Ployhart, R. (2007). A mulitdimensional approach for evaluating variables in organizational research and practice. *Personnel Psychology, 60*(2), 475-498. doi: 10.1111/j.1744-6570.2007.00080.x

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

LeBreton, J., & Tonidandel, S. (2008). Multivariate relative importance: Extending relative weight analysis to multivariate criterion spaces. *Journal of Applied Psychology, 93*(2), 329-345. doi: Doi: 10.1037/0021-9010.93.2.329

Mount, M., Witt, L., & Barrick, M. (2000). Incremental validity of empirically keyed biodata scales over GMA and the five factor personality constructs. *Personnel Psychology, 53*(2), 299-323. doi: 10.1111/j.1744-6570.2000.tb00203.x

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*(2), 262-274. doi: 10.1037/0033-2909.124.2.262

Tonidandel, S., LeBreton, J., & Johnson, J. (2009). Determining the statistical significance of relative weights. *Psychological Methods, 14*(4), 387-399. doi: DOI: 10.1037/a0017735

Van Iddekinge, C. H., & Ployhart, R. E. (2008). Developments in the criterion-related validation of selection procedures: A critical review and recommendations for practice. *Personnel Psychology, 61*(4), 871-925. doi: 10.1111/j.1744-6570.2008.00133.x

Ward, J. H., Jr. (1962). Comments on "the paramorphic representation of clinical judgment.". *Psychological Bulletin, 59*(1), 74-76. doi: 10.1037/h0045341

