

Hierarchical Multiple Regression in Counseling Research: Common Problems and Possible Remedies

John V. Petrocelli

A brief content analysis was conducted on the use of hierarchical regression in counseling research published in the Journal of Counseling Psychology and the Journal of Counseling & Development during the years 1997–2001. Common problems are cited and possible remedies are described.



Multiple regression is a powerful set of methods for examining specific scientific hypotheses and relationships among experimental, quasiexperimental, and nonexperimental data. Typically, multiple regression is used as a data-analytic strategy to explain or predict a criterion (dependent) variable with a set of predictor (independent) variables. Wampold and Freund (1987) provided an important and useful overview of the practical uses of multiple regression procedures for counseling research. They also described the distinction between simultaneous, stepwise, and hierarchical regression. In short, simultaneous regression involves cases in which the investigator enters all of the predictors into the analysis at once. Stepwise regression involves choosing which predictors to analyze on the basis of statistics. Hierarchical regression involves theoretically based decisions for how predictors are entered into the analysis. Simultaneous and stepwise regression are typically used to explore and maximize prediction, whereas hierarchical regression is typically used to examine specific theoretically based hypotheses (Aron & Aron, 1999; B. H. Cohen, 2001). For an extensive description of how these methods of multiple regression are computed, please see Pedhazur (1982).

Although Wampold and Freund (1987) noted that use of multiple regression procedures in counseling research was uncommon, it appears that their overall use has become more frequent in recent years. Wampold and Freund reported that only 14% of the research described in articles published in the *Journal of Counseling Psychology* used multiple regression procedures. During the years 1997–2001, of the quantitative research articles published in the *Journal of Counseling Psychology* and the *Journal of Counseling & Development*, 26.82% (70) have used some form of multiple regression (not including structural equation modeling, hierarchical linear modeling, canonical analysis, or any of the various analysis of variance procedures). Thus, the use of multiple regression in explaining relationships among counseling variables of interest has become quite common.

Until the 1990s, stepwise regression was one of the most frequently used statistical methods in psychological research (Thompson, 1989). Like other researchers who have focused efforts on developing appropriate methods of multiple regression (J. Cohen & Cohen, 1983; Pedhazur, 1982), Wampold and Freund (1987) warned against the routine use of stepwise regression. Stepwise, forward, and backward methods of regression, have received more criticism than any of the other forms of multiple regression (Aron & Aron,



John V. Petrocelli, Indiana University, Bloomington. The author would like to thank Bruce Thompson and Brian A. Glaser for their helpful comments on an earlier draft of this article. Correspondence concerning this article should be addressed to John V. Petrocelli, Department of Psychology, Indiana University, Bloomington, 1101 East 10th Street, Bloomington, IN 47405 (e-mail: johnpetrocelli@lycos.com).

1999; Chatterjee & Price, 1991; B. H. Cohen, 2001). Often, these methods are criticized because they produce unstable results that are sample specific and do not accurately or consistently reflect the existing relationships within the population. Furthermore, stepwise methods have often lead to incorrect computations due to the disregard of appropriate degrees of freedom, as well as inappropriate conclusions regarding the relative importance of predictor variables that are statistically dependent on variables already entered into the analysis (Huberty, 1989; Thompson, 1989). Among articles published in the *Journal of Counseling & Development* and the *Journal of Counseling Psychology* from 1997–2001, only one has reported using stepwise regression. Perhaps, articles such as “Why Won’t Stepwise Methods Die?” (Thompson, 1989), “Stepwise Regression and Stepwise Discriminant Analysis Need Not Apply Here: A Guidelines Editorial” (Thompson, 1995), and “Problems with Step-Wise Regression in Research on Aging and Recommended Alternatives” (Scialfa & Games, 1987) have been effective in getting the message about stepwise regression across to counseling researchers. It is also possible that the nature of counseling research has shifted from an exploratory focus to a clearer focus on theory testing. The latter explanation is, perhaps, a more appropriate reason for not using stepwise regression, because some quantitative methodologists still believe that stepwise regression is appropriate for exploratory purposes (Aron & Aron, 1999; B. H. Cohen, 2001).

Although the routine and inappropriate use of stepwise regression has seemingly been eradicated, new problems surrounding the use of multiple regression procedures in counseling research have emerged from the routine use of another form of regression: hierarchical regression. Almost half of the 70 articles (34, 48.57%) published in the *Journal of Counseling Psychology* and the *Journal of Counseling & Development* that have reported use of multiple regression have specifically used hierarchical multiple regression. As Scarr (1985) implied, the prolonged, routine use of any data-analytic strategy often leads to the violation of assumptions inherent in the strategy and neglect of essential guidelines grounded in its appropriate use. Editors and readers alike should beware not only because hierarchical regression has become a routine analytic procedure, but also because several methodological errors may be found surrounding its use in counseling research.

The intent of the current article is to improve subsequent counseling research investigations using hierarchical regression procedures by refreshing basic and necessary guidelines of hierarchical regression procedures that should no longer be ignored. I conducted a content analysis of quantitative research studies published in the *Journal of Counseling Psychology* and the *Journal of Counseling & Development* (during the years of 1997–2001) that used hierarchical multiple regression as a primary data analytic procedure by taking a closer look at the logic that researchers have used when using this method. A brief example illustrating the importance of the specific hierarchical order of predictor variable entry in hierarchical regression is provided. Finally, in cases for which appropriate guidelines of hierarchical regression are not feasible, common alternatives are briefly discussed.

HIERARCHICAL REGRESSION

Researchers are often interested in testing theoretical assumptions and examining the influence of several predictor variables in a sequential way, such that the relative importance of a predictor may be judged on the basis of how much it adds to the prediction of a criterion, over and above that which can be accounted for by other important predictors. As B. H. Cohen (2001, pp. 523–524) and Wampold and Freund (1987, p. 377) noted, hierarchical regression has been designed to test such specific, theory-based hypotheses.

In stepwise and simultaneous regression, a common focus is on determining the “optimal” set of predictors, by limiting the number of predictors without significantly reducing the R^2 coefficient. These methods may also be used to examine the degree of standardized unit change in the criterion for every standardized unit change in the predictor variable

when holding all other predictor variables in the model constant (at their mean) as indicated by the β coefficient (standardized partial regression coefficient). However, in hierarchical regression, the focus is on the change in predictability associated with predictor variables entered later in the analysis over and above that contributed by predictor variables entered earlier in the analysis. For instance, a researcher may want to know the extent to which measures of positive expectations about counseling and client attendance rate predict therapy outcome over and above preexisting psychopathology variables. In such a case, hierarchical regression analysis would be appropriate, provided that preexisting psychopathology variables are entered into the analysis first, followed by positive expectations about counseling and then attendance rate (because preexisting psychopathology and expectancies precede attendance—an important consideration in hierarchical regression discussed later). Substantive theory would also be strongly considered in specifying the order of entry.

Change in R^2 (ΔR^2) statistics are computed by entering predictor variables into the analysis at different steps. A predetermined, theoretically based plan for the order of predictor variable entry, held at the discretion of the researcher, is imposed on the data. Statistics associated with predictor variables entered in later steps are computed with respect to predictor variables entered in earlier steps. Thus, ΔR^2 and its corresponding change in F (ΔF) and p values are the statistics of greatest interest when using hierarchical regression (Wampold & Freund, 1987). The corresponding ΔF value for ΔR^2 would allow a researcher, interested in the example described above, to determine if the ΔR^2 statistics due to positive expectations about counseling and attendance rate significantly improve the model's ability to predict therapy outcome over and above that which can be predicted by preexisting psychopathology variables. With a focus on ΔR^2 , rather than on β or structure coefficients (Courville & Thompson, 2001; Thompson & Borrello, 1985), less attention is given to how predictor variables are reevaluated on the basis of their corresponding β s and structure coefficients when other predictors are added to the analysis, as was often done in stepwise regression.

Usually, if a β coefficient associated with a predictor variable is reported in a hierarchical regression study, it is that which was computed for the step in which it was first entered. Thus, the reported β of the predictor variable entered in Step 2 is computed while statistically controlling for the variable entered in Step 1; the reported β of the predictor variable entered in Step 1 is not that which is reevaluated in Step 2. Sometimes, experimenters report all of the coefficients for each variable at each step, including a variable's second, third, or fourth reevaluated β coefficient (for comments on reporting hierarchical regression results, see Schafer, 1991). These experimenters tend to subsequently focus their discussion on the overall model. Perhaps this pattern of analysis is evidence of the experimenter's misunderstanding of hierarchical regression or the experimenter's temptation to answer a different question than the one he or she conducted the analysis for in the first place. In such cases, a simultaneous regression may be more appropriate. However, the choice among methods of multiple regression depends on the research question being asked, the hypothesis being tested, and the logic behind the research design.

COMMON PROBLEMS OF HIERARCHICAL REGRESSION IN COUNSELING RESEARCH

Each article published in the *Journal of Counseling Psychology* and the *Journal of Counseling & Development* during the years 1997–2001 was identified as a qualitative-based, quantitative-based, or research review article. During this time, 261 quantitative articles were published in these two journals. These articles were then sorted in terms of the particular data-analytic strategies used with particular attention to the use of multiple regression procedures. A total of 70 (26.82%) of the quantitative articles reported the use of

some form of multiple regression (excluding structural equation modeling, hierarchical linear modeling, canonical analysis, or any of the various analysis of variance procedures). Of these remaining articles, 34 (48.57%) used hierarchical regression procedures. Each of these articles was then closely examined in light of the theory being tested, the appropriateness of research questions designed for examination with hierarchical regression, and the logic behind the structure of the design and statistical procedures used.

Six problems with the use of hierarchical regression are evident in these articles that involve (a) lack of clarity as to whether the study was designed to explain or predict specific outcomes; (b) hypotheses that are not consistent with those that are testable with hierarchical regression; (c) lack of a clear, explicit rationale, consistent with commonly adopted principles, for selecting hierarchical regression as a primary data-analysis strategy; (d) a focus on maximizing prediction rather than on theory-testing and the relative importance of particular predictor variables; (e) failure to examine and address probable problems of multicollinearity that may negatively affect the interpretability of the results; and (f) a discussion of results that focuses primarily on the overall model and not on the differences found through comparing progressive steps. Furthermore, several articles reporting the use of interaction terms in hierarchical regression fail to note the centering of continuous predictor variables, as strongly recommended by Aiken and West (1991), nor do they appropriately report the probing of significant interaction terms. Although neglect of these procedures certainly affects hierarchical regression results and interpretation, discussion on these issues is beyond the scope of the current analysis.

These problems, and perhaps others not detected, may be the results of four basic errors that were frequently found in the current analysis of articles. They are (a) neglect of theoretical basis for use of hierarchical multiple regression, (b) violation of causal priority, (c) use of hierarchical regression in an exploratory manner, and (d) interpretation of hierarchical regression results.

With J. Cohen and Cohen's (1983) conceptualization of appropriate uses of hierarchical regression, it follows that if one misuses this analysis in one step, subsequent steps conducting and interpreting it will lead to additional errors. However, in the current analysis only the primary misuse of hierarchical regression was recorded for each report with an error.

Neglect of Theoretical Basis

Some researchers have neglected to provide the necessary, detailed, theoretically based hypothesis or some other theoretical basis for their use of hierarchical regression. A total of 3 (8.8%) articles reporting the use of hierarchical regression failed to do this in particular.

A representative example of how hierarchical regression may have been used inappropriately in a report can be found in a study of correlates of self-reported multicultural competencies (MCCs) conducted by Sadowsky, Kuo-Jackson, Richardson, and Corey (1998). These were the stated hypotheses of the study:

After the significant relationships of multicultural social desirability and race, each, with self-reported MCCs were taken into account, respondents' locus of control racial ideology and feelings of social inadequacy would overall and individually significantly predict self-reported MCCs, and respondents' multicultural training activities would overall and individually significantly predict self-reported MCCs. (p. 257)

The researchers attempted to examine an overall score of MCCs as a criterion by entering multicultural social desirability (MCSD) in Step 1, race in Step 2, locus of control racial ideology (LCRI) and feelings of social inadequacy (FSI) in Step 3, and four types of multicultural training activities (MTA) in Step 4. From the stated hypotheses, it can be

argued that simultaneous regression would be as appropriate, if not more appropriate, because the variables listed do not imply a specific order and simultaneous regression can also take into account the significant relationships of multicultural social desirability and race with self-reported MCCs. Had the hypotheses read “We predicted that respondents’ LCRI and FSI would account for a significant amount of variance in self-reported MCCs over and above that accounted for by MCSD and race; we also predicted that respondents’ MTA would account for a significant amount of variance in self-reported MCCs over and above that accounted for by MCSD, race, LCRI, and FSI,” then hierarchical regression would have been clearly an appropriate data-analytic strategy.

It is also important to remember that in a hierarchical regression the statistical results associated with later steps depend largely on what is entered into the analysis during earlier steps. As J. Cohen and Cohen (1983) argued, demographic variables are typically good candidates for initial step entry; however, Sodowsky and colleagues (1998) entered race in Step 2 after MCSD was entered in Step 1. Furthermore, it can be argued strongly that MTAs greatly influence both MCSD and LCRI, because a primary purpose of requiring such activities in training programs is to increase knowledge and sensitivity to the experience of multicultural people. Thus, it may be more appropriate to enter MTAs in an earlier step than MCSD and LCRI. In support of this later argument, Constantine (2001) apparently believed that the number of multicultural counseling courses were predictor variables that should be considered before MCCs in a hierarchical regression of observer ratings of multicultural counseling competence in Black, Latino, and White American trainees. The Constantine study also serves as a good example for appropriately worded hypotheses that are testable with hierarchical regression (see p. 457).

Perhaps, the arguments presented in regard to the Sodowsky and colleagues (1998) study are serious threats to the validity of the report. Possibly, the researchers ignored causal priority and ordered predictor variables on the basis of J. Cohen and Cohen’s (1983) second guideline (research relevance). If this was the case, the researchers should have provided some rationale for why the relative importance of the variables studied was given priority over presumed causality. In either case, a reader should not be forced to assume the logic of the researchers’ design. If using hierarchical regression as the primary data-analytic strategy, researchers need to generate a clear and logical rationale for its use, the selection of predictor variables, and their specific order of entry.

Another example of neglecting to provide a theoretical basis for using hierarchical regression can be found in a study conducted by Lucas (1997). In an attempt to examine the effects of career development and psychological separation on moratorium and achieved identity statuses, Lucas reported entering a block of five career development measures in Step 1 and then a block of four psychological separation variables in Step 2. It can be argued that this order violated the principle of causal priority: For example, self-exploration would seem to precede comfort with career decision, not the reverse, and psychological separation variables are likely to develop along with career development variables, not after. For this reason, a rationale was needed to justify the hierarchical relevance of the predictors to the criterion. Lucas did not provide a rationale or hypothesis as to why separation variables would affect identity above and beyond the effects of career development variables, or why these variables were being examined in this specific way.

A good example of how hierarchical regression may still be appropriate for examining the effects of constructs that develop simultaneously may be found in the arguments made by McCracken and Weitzman (1997), in which a hierarchical model was computed on the basis of construct stability. Another good example of the appropriate use of hierarchical regression can be found in a study of the validity and construct contamination of the Racial Identity Attitude Scale–Long Form conducted by Fischer, Tokar, and Serna (1998). In this particular study, the hypotheses stated are clearly consistent with those that can be tested with hierarchical regression, a rationale for the predictor variables selected is clear,

as well as the specific order of their entry on the basis of causal priority and structural properties of the research design.

Violation of Causal Priority

If using hierarchical regression as the data-analytic strategy it is important to understand that results may depend largely on the order in which variables are entered into the analysis (J. Cohen & Cohen, 1983, pp. 120–123). Thus, it is necessary for questions, hypotheses, and the rationale behind the order of predictor variable entry to be very specific and theory-based. An “atheoretical” use of hierarchical regression may be just as inappropriate as using exploratory-based analyses such as stepwise regression.

J. Cohen and Cohen (1983) described three basic principles that should underlie the hierarchical order of predictor variable entry. First, predictor variable entry should respect presumed causal priority (the direction of causal flow). In other words, if there are causal relationships among the predictor variables the causes should be entered into the analysis before their effects. J. Cohen and Cohen stated that “ideally, no independent variable entering later should be a presumptive cause of an independent variable that has been entered earlier” (p. 120). Adherence to this guideline, they argue, leads to two major advantages of hierarchical regression: (a) extraction of as much causal inference as the data will allow and (b) a unique partitioning of the total variance of the criterion that can be accounted for by individual predictors as indicated by the increase in R^2 . When using hierarchical regression, the variance of the criterion attributed to a predictor variable depends on its relationship with the criterion and on what has already been entered into the model. Thus, if causal priority is ignored in the ordering of predictor variables, a researcher risks attributing changes in the explained variance of the criterion to an effect that would otherwise be attributed to a cause. J. Cohen and Cohen also argued that “[t]his stolen (spurious) variance will then mislead the investigator about the relative importance to Y of the cause and its effect” (p. 121). Thus, the most careful of researchers tend to enter static variables of interest (e.g., gender, age, or race) before entering dynamic variables in subsequent steps.

Second, in some investigations one predictor may be more relevant than another on the basis of some theoretical position. The order of predictor variable entry may be dictated by the hierarchical relevance of each predictor to the criterion. Selection of the order of predictor entry may still reflect causal priority or a theoretical or psychometric rationale. As J. Cohen and Cohen (1983) stated, adherence to this guideline will lead to “clarity in interpretation of the influence of X_1 and X_2 that is likely to result from this approach . . . the statistical power of the test of the major hypothesis is likely to be maximal when the appropriate error model is used” (p. 123).

Sometimes researchers collect data on variables that have curvilinear relationships or interactions among them. Thus, as J. Cohen and Cohen (1983) stated, some predictors “have characteristics that make assessment of their contribution to R^2 meaningful only after related variables have been partialled, thus mandating a specific order” (p. 123).

Some researchers have neglected the principle of causal priority when using hierarchical regression. A total of 8 (23%) articles reporting the use of hierarchical regression examined in the current analysis failed to do this. On rare occasions, adherence to this principle is not mandated. However, in this case the researcher should provide some rationale for why the analysis of potentially causal variables was meaningful only after their effects have been partialled. Violations of causal priority appear to be the most common problem associated with the use of hierarchical regression in counseling research.

A representative example of violating causal priority can be found in a study conducted by McCullough, Worthington, Maxey, and Rachal (1997). These researchers used hierarchical regression to reveal whether or not the effect of counselor gender on participants’

attitudes toward their counselor was mediated by perceptions of counselor religiousness. They appear to have appropriately tested the first three conditions of a mediating variable recommended by Baron and Kenny (1986). Baron and Kenny, as well as others (Kenny, Kashner, & Bolger, 1998), have argued that when the proposed mediator is statistically controlled the effect of the initial predictor (in this case gender of counselor) on the criterion (attitude toward counselor) should be reduced to nonsignificance. They have provided a modified version of Sobel's (1982) test to examine whether or not the reduction in the effect is statistically significant. However, McCullough and colleagues (1997) attempted to reveal this reduction by conducting a second hierarchical regression model in which gender was entered in the second step, following perceptions of counselor religiousness. Not only does this method violate the principle of causal priority but it also fails to adequately test the reduction of the effect.

Ruelas, Atkinson, and Ramos-Sanchez (1998) provided one of the most obvious cases of violating causal priority. In this study, the researchers entered their experimental manipulation as a variable before participant acculturation level. Acculturation level may be measured subsequent to the experimental manipulation in the actual experiment; however, when analyzing the data, the effect of the experimental manipulation would have been examined appropriately had acculturation level been entered in a step prior to the manipulation variable.

An example of how the order of predictor variable entry makes a difference in the results and subsequent conclusions is illustrated with three different ordering methods used by three different researchers. Suppose all three researchers were interested in determining the degree to which time spent in psychotherapy improves preexisting global functioning aside from that of pretherapy data. Specifically, they wanted to know the degree of variance in Global Assessment of Function (GAF) score increase (from the time of intake) that time spent in psychotherapy (measured by number of psychotherapy sessions) accounts for over and above the variance accounted for by pretherapy depression scores and age. An artificial data set with 10 participants (clients reporting to psychotherapy, complaining of severe depression for the past 3 months), three predictor variables (age, depression scale scores, and number of therapy sessions), and one criterion variable (GAF score increase from the time of intake) is displayed in Table 1. As is typically provided by reports of studies using multiple regression procedures, a correlation matrix is provided in Table 2.

Suppose Researcher 1 ordered the predictors in such a way as to reflect the principle of causal priority and to test the hypothesis that number of therapy sessions account for vari-

TABLE 1
Artificial Data Set

Participant	Age	Pretherapy Depression Score	Number of Therapy Sessions	GAF Increase
1	21	52	6	25
2	19	61	17	25
3	18	50	5	0
4	30	57	31	43
5	27	54	10	18
6	39	55	27	35
7	36	61	11	51
8	20	52	7	20
9	23	51	8	16
10	35	55	10	47

Note. GAF = Global Assessment of Functioning.

TABLE 2
Intercorrelations Between All Artificial Study Variables

Variable	1	2	3	4
1. Age	—			
2. Pretherapy depression score	.42	—		
3. Number of therapy sessions	.51	.47	—	
4. Increase in overall GAF scale score	.79**	.70*	.49	—

Note. GAF = Global Assessment of Functioning.

* $p < .05$. ** $p < .01$.

ance in psychotherapy success (as indicated by the total GAF score increase from the time of intake), over and above that which is accounted for by pretherapy depression scale scores and age. Age is a demographic variable that is not the effect of either of the other two predictors. Thus, Researcher 1 analyzed the data in Table 1 with hierarchical regression, in which GAF score increase was regressed on age, depression scale scores, and number of therapy sessions entered separately in Step 1 through Step 3 respectively. Researcher 1 would report the results displayed in the top of Table 3 (Model 1). Using this specific order of entry would lead Researcher 1 to conclude that depression scale scores accounted for 16% of the variance in GAF score increase that is over and above the variance accounted for by age. Researcher 1 would also conclude that this finding was a statistically significant increase. Finally, Researcher 1 would conclude that the number of therapy sessions does not significantly increase the amount of variance accounted for in GAF score increase over and above that which is accounted for by age and depression scale scores. Thus, in this case, age and depression scale scores may be judged as more relevant than time spent in therapy in understanding how GAF scores are likely to increase after 15 weeks of therapy.

Suppose Researcher 2 analyzed the same data displayed in Table 1 with hierarchical regression but haphazardly ordered the predictors such that number of therapy sessions, age, and depression scale scores were entered separately in Step 1 through Step 3 respectively. The results computed by Researcher 2 (displayed as Model 2 in Table 3) are notice-

TABLE 3
Hierarchical Regression Analysis Summary for GAF Increase on Age, Pretherapy Depression Score, and Number of Therapy Sessions

Model, Step, and Predictor Variable	R^2	ΔR^2	ΔF	df
Model 1				
1. Age	.63**	.63	13.63**	(1, 8)
2. Pretherapy depression score	.79**	.16	5.56*	(1, 7)
3. Number of therapy sessions	.79**	.00	0.03	(1, 6)
Model 2				
1. Number of therapy sessions	.25	.25	2.63	(1, 8)
2. Age	.64*	.39	7.66**	(1, 7)
3. Pretherapy depression score	.79*	.15	4.50	(1, 6)
Model 3				
1. Age	.63**	.63	13.62**	(1, 8)
2. Pretherapy depression score/number of therapy sessions	.79*	.16	2.41	(1, 6)

Note. GAF = Global Assessment of Functioning; ΔR^2 = change in R^2 ; ΔF = change in F .

* $p < .05$. ** $p < .01$.

ably different from the results computed by Researcher 1 (Model 1 of Table 3). Researcher 2 would be led to conclude that age accounted for 39% of the variance in GAF score increase that is over and above the variance accounted for by number of therapy sessions. Researcher 2 would also conclude that this finding was a statistically significant increase. Researcher 2 would report that the pretherapy depression score increased R^2 by 15%, but that this finding was not a statistically significant increase over and above that accounted for by age and time spent in therapy. However, the order of predictor variable entry used by Researcher 2 is inappropriate not only because of the original research question, but also because it violates causal priority. Researcher 2 may then conclude that the depression scale scores do not significantly help explain GAF score increase when statistically controlling for age and number of therapy sessions, despite the evidence provided by Researcher 1 that supports the notion that depression scale scores do significantly help explain GAF score increases over and above age.

The overall variance in GAF score increase accounted for by the predictor variables is the same for both methods used by Researchers 1 and 2. However, R^2 is associated with a model's degree of predictability of a criterion variable. As described earlier, hierarchical regression is designed for examining theory-based hypotheses and not for maximizing prediction (Aron & Aron, 1999; B. H. Cohen, 2001; Wampold & Freund, 1987). The ΔR^2 is an indicator of the relevance of particular predictor variables entered in later steps, relative to those entered in earlier steps. A focus on ΔR^2 should not be confused with a concern for maximizing predictability. Rather, the purpose of using hierarchical regression is to test theoretical assumptions and to determine the degree to which variables entered later in the analysis account for variance in the criterion over and above that which is accounted for by variables entered earlier in the analysis.

Sometimes researchers choose to enter sets of variables, or blocks, in a hierarchical regression analysis. Again, such procedures should not be used haphazardly because they may also affect the results. Suppose Researcher 3, analyzed the data displayed in Table 1 using hierarchical regression but decided to use a block method without a theoretical reason for doing so. Suppose Researcher 3 entered age in Step 1 and depression scale scores and number of therapy sessions as a block in Step 2. Researcher 3 would report the results displayed as Model 3 in Table 3. Thus, Researcher 3 may conclude that depression scale scores and number of therapy sessions increased R^2 by 16%, but again this finding was not a statistically significant increase over and above that accounted for by age. Thus, Researcher 3 may conclude that both pretherapy depression level and time spent in therapy do not significantly contribute to explaining improvements in GAF scores over and above that which is explained by age.

This example demonstrates that the order of predictor entry, whether individual or in blocks, makes a difference in the results and conclusions. If there is high multicollinearity among the predictor variables, as is found in many of the investigations conducted by counseling researchers, results may depend even more on the order of predictor variable entry. One reason for this is that if there is high multicollinearity among predictors, it is almost impossible to determine which variables are contributing to the explained variance of the criterion when using regression procedures. The example also emphasizes the importance of J. Cohen and Cohen's (1983) recommendations that the researcher have some theoretical, psychometric, or causal-based rationale for ordering the predictor variables.

Use of Hierarchical Regression in an Exploratory Manner

Some researchers have appeared to deliberately, yet inappropriately, use hierarchical regression in an exploratory manner. A total of three (8.8%) articles, examined in the current analysis, have used hierarchical regression in this way.

A study conducted by Lee and Robbins (1998) provides a representative example of such use of hierarchical regression. The analysis in this study reflects not only the neglect

of causal priority but also the uncertainty about appropriate use of hierarchical regression as a primary data analytic strategy. Lee and Robbins explored the unique role of social connectedness, collective self-esteem, and social support in level of trait anxiety by testing three different hierarchical regression models in which each of the predictors was entered in the final step of separate regression models, after controlling for other predictors of trait anxiety, respectively. Aside from apparently ignoring both (a) the possibility that social connectedness and social support may be causes of collective self-esteem and (b) the multicollinearity among these predictors, there was obvious neglect of the fact that hierarchical regression is to be used as a theory-driven analysis. Although interesting results were found that highlighted the importance of social connectedness over and above collective self-esteem and social support, the speculation that the results were clouded by multicollinearity may have been reduced with a more theoretically driven use of hierarchical regression in this study.

Interpretation of Hierarchical Regression Results

Inappropriate use of hierarchical regression is almost always going to lead to inaccurate interpretation of results. Thus, in a sense, each study reporting an error in the use of hierarchical regression leads to the misinterpretation of results. However, a total of 3 (8.8%) articles, in particular, appeared to have followed standard guidelines of use but either misinterpreted or underinterpreted the results of hierarchical regression.

A representative example of this is found in a study of the effects of parental divorce and parent-child bonds on adult attachment avoidance and anxiety conducted by Lopez, Melendez, and Rice (2000). These researchers appeared to have followed appropriate guidelines of hierarchical regression; however, when interpreting the results of each of the ten models they tested, the researchers gave attention to the significance or nonsignificance of ΔR^2 statistics in only two of the models; the ΔF values and the question of whether or not ΔR^2 associated with progressive steps were neglected from the statistics summary tables. More important to the argument that this particular article misinterpreted the results of hierarchical regression, the researchers focused squarely on the β coefficient as well as the overall R^2 , suggesting that results were interpreted as if they had been obtained from the computation of simultaneous regression models. By failing to examine whether or not parent-child bonds significantly explained adult attachment avoidance and anxiety above and beyond age and parental marital status, the hierarchical regression results were undermined by poor interpretive statements. A similar instance of underinterpretation of hierarchical regression can be found in a hierarchical model of predictors, designed to explain client attitudes toward spirituality in therapy, tested by Rose, Westefeld, and Ansley (2001). These researchers reported the overall R^2 and the significance of the overall model, but neglected to report any of the statistics unique to hierarchical regression.

POSSIBLE REMEDIES

As in other research paradigms, those that use a multiple regression procedure as a data-analytic strategy must be preceded by the formation of specific questions (Huberty & Hussein, 2001). Hierarchical regression should only be used (a) when the question is theoretically based and concerned with the degree to which predictor variables entered later in the analysis account for variance in the criterion over and above those entered earlier in the analysis and (b) when the guidelines derived by J. Cohen and Cohen (1983, pp. 120–123) are feasible. It should be clear to researchers that if recommended guidelines for the use of hierarchical regression and predictor variable entry cannot be met, methods such as simultaneous regression may be more appropriate.

If hierarchical regression is not an appropriate strategy, given its assumptions and guidelines of appropriate use, other multiple regression procedures may sometimes be used.

One method that some statisticians adopt is to first compute a simultaneous regression model. Then, compute several subsequent models (equaling the number of predictor variables) each with a different predictor absent from the analysis. Finally, the relative importance of a predictor is judged on the basis of how much its absence decreases R^2 . The advantage of this method is that the focus is still on the change in ΔR^2 . The ΔR^2 for each of these subsequent models is also equal to the semipartial (or part) r^2 from the initial simultaneous regression.

Montgomery and Peck (1992) have also suggested that the investigator order predictor variables on the basis of the computed t statistic corresponding to β coefficients. The search on t method begins with a simultaneous regression by entering all of the predictors at once. Then, predictor variables are ordered by their absolute value of t . Finally, a hierarchical regression analysis may be computed by sequentially adding one variable at a time in descending order of the original corresponding t values. This method is similar to forward selection regression; is likely to produce unstable, sample-specific results; and meets the same criticisms as other stepwise methods.

Hierarchical regression does not necessarily solve all of the problems associated with high multicollinearity among some or all of the predictor variables. Thus, methods such as principal components regression, as described by Massy (1965) and McCallum (1970), and ridge regression, as described by Hoerl and Kennard (1970a, 1970b), Marquardt and Snee (1975), and Price (1977), have been developed.

However, the problem with all of the aforementioned methods described is that they are still somewhat data-driven and influenced by fluctuations in samples (a frequent criticism of stepwise regression). Thus, they do not appear to be appropriate replacements of hierarchical regression and the theory-based testing inherent in its use.

If the unique contribution of a single predictor is of interest, the ΔR^2 is identical to a semipartial (or part) r^2 . Furthermore, when F has 1 degree of freedom in the numerator and k degrees of freedom in the denominator, the ΔF is equal to the square of the t statistic for that variable's regression coefficient from a simultaneous regression. In hierarchical regression, the degrees of freedom in the numerator will equal 1 in each step, and the degrees of freedom in the denominator will be $N - 1$ —the number of predictors currently in the model. A major advantage of hierarchical regression is of course the ability to examine the significance in the incremental increases in R^2 when more than one predictor is of interest or a set of predictors that share some relevant commonality are of interest (blocks). The entry of a block of predictors is unique to hierarchical regression and often particularly useful to counseling researchers. For instance, a researcher may want to know whether counseling approach can explain variation in outcome, over and above the client's level of anxiety. Because anxiety may be measured by more than one instrument, it is reasonable to enter multiple anxiety measures as a block of predictors after controlling for counseling approach.

In either case, B. H. Cohen (2001) argued, a researcher will eventually need some way of ordering predictor variables to test theory-based hypotheses. Persistent efforts to answer specific research questions should not continue at the expense of an appropriate analysis of experimental or nonexperimental data. The ability of a researcher to answer quantitative research questions is always dependent on the statistical-analytic strategies available. Data-driven methods of multiple regression are not appropriate replacements of hierarchical regression. Sometimes considered as extensions of multiple regression, relatively new developments in structural equation modeling (Jöreskog, 1993; Maruyama, 1998), hierarchical linear modeling (Bryk & Raudenbush, 1992), and discriminant analysis (Huberty, 1994) may be appropriate alternatives to hierarchical regression, depending on the particular research hypotheses being tested.

Commonality Analysis

In hierarchical regression, the importance of predictors are judged on the basis of ΔR^2 . However, it is also useful to examine common effects (in general, the difference between

the sum of all of the unique effects of all predictors and the total explained variance) because the total explained variance equals the sum of all of the unique and common effects. The common effects can be examined by using a follow-up procedure known as commonality analysis (Rowell, 1996; Thompson, 1985). Commonality analysis is useful if the researcher wants to find the unique variance of the criterion explained by a predictor as well as the degree to which the predictability of the predictor is common to and available from other predictor variables. Although hierarchical regression is certainly used much more frequently than commonality analysis in counseling research, commonality analysis has compelling advantages that should be seriously considered when selecting multiple regression strategies to test experimental hypotheses. For instance, commonality analysis allows a researcher to “decompose” R^2 by examining the proportion of explained variance of the criterion associated with the common effects of the predictors. Sometimes, a predictor’s unique effect is relatively small but its associated common effects are relatively large, thus, rendering it an important variable after all. If researchers are attempting to determine a predictor’s (or set of predictors’) predictability over and above predictors considered in earlier steps, they may add to the interpretation and understanding of their results by examining common effects through commonality analysis as a follow-up procedure to hierarchical regression. For descriptive examples of how to conduct a commonality analysis, researchers are encouraged to consult Rowell (1996) and Seibold and McPhee (1979).

CONCLUSION

Misuse of hierarchical regression can produce results that are just as spurious as those produced through stepwise regression. One of the most common errors in the use of hierarchical regression in counseling research appears to be the violation of causal priority in the models of predictors tested. Researchers need to provide not only an appropriate rationale for using hierarchical regression, but also logical reasoning for why predictor variables were ordered as they were. It is hoped that attention to the common errors and adherence to the recommendations described here will lead to more sophisticated and valid counseling research that uses hierarchical regression as a data-analytic strategy.

REFERENCES

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. New York: Sage.
- Aron, A., & Aron, E. N. (1999). *Statistics for psychology* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis*. Thousand Oaks, CA: Sage.
- Chatterjee, S., & Price, B. (1991). *Regression analysis by example* (2nd ed.). New York: Wiley.
- Cohen, B. H. (2001). *Explaining psychological statistics* (2nd ed.). New York: Wiley.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Constantine, M. G. (2001). Predictors of observer ratings of multicultural counseling competency in Black, Latino, and White American trainees. *Journal of Counseling Psychology, 48*, 456–462.
- Courville, T., & Thompson, B. (2001). Use of structure coefficients in published multiple regression articles: β is not enough. *Educational and Psychological Measurement, 61*, 229–248.
- Fischer, A. R., Tokar, D. M., & Serna, G. S. (1998). Validity and construct contamination of the Racial Identity Attitude Scale—Long Form. *Journal of Counseling Psychology, 45*, 212–224.
- Hoerl, A. E., & Kennard, R. W. (1970a). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics, 1*, 55–67.
- Hoerl, A. E., & Kennard, R. W. (1970b). Ridge regression: Applications to nonorthogonal problems. *Technometrics, 1*, 68–82.

- Huberty, C. J. (1989). Problems with stepwise methods—Better alternatives. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 1, pp. 43–70). Greenwich, CT: JAI Press.
- Huberty, C. J. (1994). *Applied discriminant analysis*. New York: Wiley.
- Huberty, C. J., & Hussein, M. H. (2001). Reporting information in multiple correlation and multiple regression studies. In E. I. Farmer & J. W. Rojewski (Eds.), *Research pathways: Writing professional papers, theses, and dissertations in workforce education* (pp. 325–348). New York: University Press of America.
- Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 294–316). Thousand Oaks, CA: Sage.
- Kenny, D. A., Kashner, D. A., & Bolger, N. (1998). Data analysis in social psychology. In D. Gilbert, S. T. Fiske, & G. Lindsey (Eds.), *Handbook of social psychology* (4th ed., Vol. 1, pp. 233–265). New York: McGraw-Hill.
- Lee, R. M., & Robbins, S. B. (1998). The relationship between social connectedness and anxiety, self-esteem, and social identity. *Journal of Counseling Psychology, 45*, 338–346.
- Lopez, F. G., Melendez, M. C., & Rice, K. G. (2000). Parental divorce, parent–child bonds, and adult attachment orientation among college students: A comparison of three racial/ethnic groups. *Journal of Counseling Psychology, 47*, 177–186.
- Lucas, M. (1997). Identity development, career development, and psychological separation from parents: Similarities and differences between men and women. *Journal of Counseling Psychology, 44*, 123–132.
- Marquardt, D. W., & Snee, R. D. (1975). Ridge regression in practice. *The American Statistician, 29*, 3–20.
- Maruyama, G. (1998). *Basics of structural equation modeling*. Thousand Oaks, CA: Sage.
- Massy, W. F. (1965). Principal components regression in exploratory statistical research. *Journal of the American Statistical Association, 60*, 234–256.
- McCallum, B. T. (1970). Artificial orthogonalization in regression analysis. *The Review of Economics and Statistics, 52*, 110–113.
- McCraken, R. S., & Weitzman, L. M. (1997). Relationship of personal agency, problem-solving appraisal, and traditionality of career choice to women's attitudes toward multiple role planning. *Journal of Counseling Psychology, 44*, 149–159.
- McCullough, M. E., Worthington, E. L., Maxey, J., & Rachal, K. C. (1997). Gender in the context of supportive and challenging religious counseling interventions. *Journal of Counseling Psychology, 44*, 80–88.
- Montgomery, D. G., & Peck, E. A. (1992). *Introduction to linear regression analysis* (2nd ed.). New York: Wiley.
- Pedhazur, E. J. (1982). *Multiple regression in behavioral research: Explanation and prediction* (2nd ed.). New York: Holt, Rinehart, and Winston.
- Price, B. (1977). Ridge regression: Application to non-experimental data. *Psychological Bulletin, 84*, 759–766.
- Rose, E. M., Westefeld, J. S., & Ansley, T. N. (2001). Spiritual issues in counseling: Clients' beliefs and preferences. *Journal of Counseling Psychology, 48*, 61–71.
- 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. 33–44). Greenwich, CT: JAI Press.
- Ruelas, S. R., Atkinson, D. R., & Ramos-Sanchez, L. (1998). Counselor helping model, participant ethnicity and acculturation level, and perceived counselor credibility. *Journal of Counseling Psychology, 45*, 98–103.
- Scarr, S. (1985). Constructing psychology: Making facts and fables of our times. *American Psychologist, 40*, 499–512.
- Schafer, W. D. (1991). Reporting hierarchical regression results. *Measurement and Evaluation in Counseling and Development, 24*, 98–100.
- Scialfa, C. T., & Games, P. A. (1987). Problems with step-wise regression in research on aging and recommended alternatives. *Journal of Gerontology, 42*, 579–583.
- Seibold, D. R., & McPhee, R. D. (1979). Commonality analysis: A method for decomposing explained variance in multiple regression analysis. *Human Communication Research, 5*, 355–365.
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological methodology 1982* (pp. 290–312). San Francisco: Jossey-Bass.
- Sodowsky, G. R., Kuo-Jackson, P. Y., Richardson, M. F., & Corey, A. T. (1998). Correlates of self-reported multicultural competencies: Counselor multicultural social desirability, race, social inadequacy, locus of control racial ideology, and multicultural training. *Journal of Counseling Psychology, 45*, 256–264.

- Thompson, B. (1985). Alternate methods for analyzing data from education experiments. *Journal of Experimental Education, 54*, 50–55.
- 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*, 525–534.
- Thompson, B., & Borrello, G. M. (1985). The importance of structure coefficients in regression research. *Educational and Psychological Measurement, 45*, 203–209.
- Wampold, B. E., & Freund, R. D. (1987). Use of multiple regression in counseling psychology research: A flexible data-analytic strategy. *Journal of Counseling Psychology, 34*, 372–382.