DOCUMENT RESUME

ED 466 699 TM 034 278

AUTHOR Herring, Jennifer C.

TITLE Multiple Regression Methodology in the "Journal of Education for Students Placed at Risk": Effect Sizes and Structure Coefficients.

PUB DATE 2002-02-00

NOTE 15p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (Austin, TX, February 14-16, 2002).

PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)

EDRS PRICE EDRS Price MF01/PC01 Plus Postage.

DESCRIPTORS *Effect Size; *Regression (Statistics); *Research Reports; Scholarly Journals; *Statistical Analysis

IDENTIFIERS *Structure Coefficients

ABSTRACT

This study reviewed the statistical practices in published research articles in the "Journal of Education for Students Placed at Risk" to determine the reporting of effect sizes and structure coefficients. Of the 12 quantitative studies found in the last 3 volumes of the journal, only 3 were identified as using multiple regression analysis. Two of the three indicated use of effect size magnitudes, but the third reported only "F" and "p-" values. None of these articles gave structure coefficients. Reporting effect sizes highlights the difference between statistical and practical significance. The paper also recommends the use of "what-if" analysis when using regression. (SLD)

Running head: EFFECT SIZES AND STRUCTURE COEFFICIENTS

Multiple Regression Methodology in the *Journal of Education for Students Placed at Risk*: Effect Sizes and Structure Coefficients

Jennifer C. Herring

University of North Texas

Paper presented at the annual meeting of the Southwest Educational Research Association, February 14-16, 2002, Austin, TX.

Multiple Regression Methodology in the *Journal of Education for Students Placed at Risk*: Effect Sizes and Structure Coefficients

In any research study, the integrity of the research is only as strong as the weakest link. Study validity can be threatened from myriad fronts, including attacks to what is commonly called statistical conclusion validity. Statistical conclusion validity refers to the appropriate use and interpretation of the statistical methodology used in a given study (Babbie, 1992). When methods are not appropriately employed, the conclusions that can be drawn from a line of inquiry are weakened. Accordingly, the purpose of the present paper is to review the statistical practices in published research articles in the *Journal of Education for Students Placed at Risk* concerning two important issues in regression research.

First, as in all general linear model analyses (cf. Cohen, 1968; Knapp, 1978) and consistent with current trends in best statistical practice (Thompson, 1999; Wilkinson & American Psychological Association (APA) Task Force on Statistical Inference, 1999), the interpretation of effect sizes in regression research is critical to understanding one's study results. Effect sizes help researchers understand the magnitude of relationships among their variables, and therefore shed light on result importance (Thompson, 1999). Effect size reporting is becoming increasingly important as researchers become aware of the limits of statistical significance testing. As the APA Task Force recommended:

> It is hard to imagine a situation in which a dichotomous accept-reject decision is better than reporting an actual *p*-value or, better still, a confidence

interval... Always provide some effect-size estimate when reporting a *p*-value (p. 599, emphasis added).

Second, and also in all general linear model analyses, structure coefficients are critical to interpreting variable importance. Indeed, the common practice of examining only standardized weights (i.e., beta weights) in regression research can lead to erroneous conclusions (Thompson & Borrello, 1985;Courville & Thompson, 2001). Because they represent the simple bivariate correlation between a predictor variable and the latent predicted variable (i.e. Yhat), structure coefficients tell us the contribution each predictor could have made in accounting for overall effect. Structure coefficients are not adversely affected by multicollinearity (as beta weights are), and therefore overcome the so-called "bouncing beta" problem in which the beta weights can dramatically change as predictor intercorrelations fluctuate across studies. Accordingly, both beta weights and structure coefficients should almost always be interpreted in regression research.

Several studies have examined the reporting of effects and structure coefficients in research (cf. Courville & Thompson, 2001; Henson & Smith, 2000; Vacha-Haase, Nilsson, Reetz, Lance, & Thompson, 2000). In large part, and unfortunately, both effect sizes and structure coefficients are not frequently reported and interpreted in the research literature. Kirk (1996) suggested increase frequency of reporting for multiple regression articles due to the fact that statistical software often readily provide R^2 effects in regression output.

Purpose

To further examine the current status of best reporting practice, the present paper will examine the use and interpretation of structure coefficients and effect sizes in

multiple regression articles in the *Journal of Education for Students Placed at Risk.* Therefore, the present paper will extend the methodological examination to the multicultural and at-risk education literature.

Articles using multiple regression were first identified in the last three volumes of the journal. These articles were then examined and coded concerning whether the article (a) reported structure coefficients, (b) interpreted structure coefficients, (c) reported effect sizes, and (d) interpreted effect sizes. Along with this data, the paper also documents the types of effect sizes used. Examples of best and poor practice are noted and recommendations for improved reporting practice are offered.

Method

Article Selection

A total of twelve quantitative type studies were found in the last three volumes of *The Journal of Education for Students Placed At Risk*. Only three of those twelve articles were identified as using multiple regression analysis, of which this paper addresses.

Effect Size Reporting

Two of the three articles examined using multiple regression in *The Journal of Education for Students Placed At Risk* indicated use of effect size magnitudes. The third article only reported F and p-values.

Effect Size Interpretation

Kirk (1996) identified 16 varieties of magnitude effect reporting in social science journals reviewed in one study, however only r, adjusted r^2 and eta^2 were interpreted in the articles reviewed that used multiple regression in *The Journal of Education for Students Placed At Risk.* Of the two studies using effect sizes, adjusted r^2 was only

referenced once; eta^2 was also reported in one article. The article reporting adjusted r^2, provided an interpretation of the data results in both the Results and Discussion sections; the other article reporting eta^2 provided an interpretation in the Results section only.

Structure Coefficient Reporting and Interpretation

In examining studies that used multiple regression analysis in *The Journal of Education for Students Placed At Risk,* all of those studies were quantitative in nature, reporting statistics such as the mean, standard deviation, F, and p-values. However, none reported or interpreted structure coefficients or beta weights.

Reporting and Interpreting Structure Coefficients

Reporting structure coefficients as an important measure in research studies has yet to be found as a common method in examining and analyzing data, in the general linear model (Courville & Thompson, 2001). However, as noted by Thompson (1992), the reporting of beta weights and structure coefficients are both important. Thompson (1992) further argued that "the thoughtful researcher should always interpret either (a) both the beta weights and the structure coefficients or (b) both the beta weights and the bivariate correlations of the predictors with Y" (p. 14).

Unfortunately, in regression research, heavy emphasis has been placed on solely interpreting beta weights to the exclusion of structure coefficients. For example, one study (Courville & Thompson, 2001) noted that only three of twenty articles reviewed in *The Journal of Counseling Psychology* between January 1990 and April 1993 reported structure coefficients.

Since structure coefficients can help explain the predictive importance of a predictor variable's contribution to the overall effect, reporting these measures may help

in explaining the theoretical and practical ramifications of the results. Courville and Thompson (2001) suggested that if a general linear model analysis is used in a study where predictors are correlated, as they often are, then "when interpreting regression results, once noteworthy effects have been detected it may be best to consult the full system of results, just as we routinely would in applications of other members of the general linear model analytic family" (p. 45).

A structure coefficient is a correlation between an observed variable (e.g. predictor) and a synthetic variable (e.g. Yhat); an example structure coefficient formula for predictor X1 would be: r structure = r Y with X1/R (cf. Courville & Thompson, 2001; Daniel & Onwuegbuzie, 2001). Structure coefficients are not readily obtainable for multiple regression using a statistical software package (i.e. SPSS), although it is as much a correlation analysis as Pearson's r, and can be calculated using the statistical package (cf. Daniel & Onwuegbuzie, 2001). All researchers need to do is to save the unstandardized predicted values (i.e. Yhat) as part of the regression analysis and then correlate these Yhat scores with the observed predictors. When one squares the structure coefficient (r_s^2), the result is the percentage of variance the predictor could contribute to the explained effect.

Reporting Effect Sizes

Reporting effect sizes alongside *p*-values, given the limitations of statistical significance testing, is long overdue for educational journals (Kirk, 1996). As noted by Kirk (1996), "researchers have been encouraged to supplement reports of null hypothesis tests with measures of effect magnitude" for over 70 years now.

Effect sizes can generally be grouped into two categories. Category one includes mean difference effects such as Cohen's d, which represents the difference between experimental and control groups divided by the pooled standard deviation of both groups. Category two includes variance-accounted-for effects such as r^2, eta^2 and adjusted r^2 (Snyder & Lawson, 1993).

Reporting an effect size is necessary in that statistical significance testing alone only answers the question "*If* the sample represents the population, how likely is the obtained [sample] result?" (Thompson, 1998, p. 34, emphasis in original). Too, statistical conclusion on the hypothesis is a function of obtained *p*-values, which, as Henson and Smith suggested (2000, p.286, emphasis in original) only "indicate the probability (0 to 1.0) of attaining the *presently* observed results from the *present* sample assuming that the null hypothesis is *exactly* true in the population." This makes it necessary to supplement such judgments with alternative reporting methods, such as effect size measures to indicate the magnitude of the finding, not just the likelihood of the finding.

According to Kirk (1996), reporting a statistic that indicates a specific magnitude of effect, rather than reporting a general reference to the use of effect sizes, is more favorable. He further noted that r^2 and eta^2 are readily recorded as such, when using statistical software packages, such as SPSS (Kirk, 1996).

Practices and Recommendations

Of the articles reviewed in *The Journal of Education for Students Placed At Risk* using multiple regression, three of them reported p-values or F statistics. Of these however, only two offered effect sizes magnitude; none gave structure coefficients.

It is poor practice to report statistical significance testing values without also reporting and interpreting effect size values (Thompson, 1998). Thompson also posited "most authors will simply not change their practices until editorial policies require them to do so (Thompson, 1998, p. 35). In this same article, he also confronted the wording of the American Psychological Association's 4th edition publication manual in reference to reporting practices for effect sizes of "encouraging" authors to report effect sizes; he cited this wording as "too vague to enforce " (Thompson, 1998, p. 35). It is interesting to note here that the APA's 5th edition (2001) now reads:

> For the reader to fully understand the importance of your findings, it is <u>almost always necessary</u> to include some index of effect size or strength of relationship in your Results section. You can estimate the magnitude of effect or the strength of the relationship with a number of common effect size estimates, including (but not limited to) r^2, η^2, ω^2, R^2, φ^2, Cramer's V, Kendall's W, Cohen's d and κ, Goodman-Kruskal's λ and γ, Jacobson and Truax's (1991) and Kendall's (1999) proposed measure of clinical significance, and the multivariate Roy's Θ and the Pillai-Barlett V (pp. 25-26, emphasis added).

The Journal of Education for Students Placed At Risk shows evidence in the quantitative articles reviewed that effect size reporting and interpretation is practiced. It must be noted her that *The Journal of Education for Students Placed At Risk* does include both quantitative and qualitative articles. Of these articles (forty-five), however, only twelve of them were quantitative studies. In examining the twelve quantitative studies, only three used multiple regression, but of these three, two of them reported effect sizes.

The practice of reporting structure coefficients in *The Journal of Education for Students Placed At Risk,* is not popular to say the least (none were noted). Courville and Thompson (2001) cited "in most cases, regression researchers ought to interpret β weights and structure coefficients (or else bivariate correlations of predictors with the criterion) once a noteworthy omnibus effect is detected" (p. 244). Also, Courville and Thompson (2001) noted the importance of interpreting structure coefficients in that the reason for calling "structure coefficients" such is "that they provide insight regarding what is the nature or structure of the underlying synthetic variables of the actual research focus. (p 238, emphasis added).

Recommendations

It is recommended that authors report *p*-values and test statistics from the general linear model when necessary, but also include the reporting and interpreting of effect sizes to explain magnitude of effect. When notable effects are observed, then and only then, report and interpret both standardized weights and structure coefficients to explain how much each predictor is contributing to the model (Daniel & Onwuegbuzie, 2001). Reporting effect sizes also "highlight the distinction between *statistical* and *practical* significance. Results that are statistically significant are not always noteworthy in a practical sense" (Snyder & Lawson, 1993, pp. 334-335, emphasis in original) but effect sizes "tells us how much of the dependent variable can be controlled, predicted, or explained by the independent variable(s)" (Snyder & Lawson, 1993, p. 335).

It is also recommended that researchers conduct a newer, precise "what-if analysis" when using regression, as noted by Henson and Smith (2000) and proposed by Kieffer and Thompson (1999); this newer "what-if analysis," provides for corrected

effect estimates (Henson & Smith, 2000). A "what if" analysis resembles a post hoc power analysis and helps researchers place statistical significance tests in the context of sample size (Henson & Smith, 2000).

Because most researchers learn their methodology skills in educational statistics courses in university settings, it is recommended that courses and textbooks reflect such practices. Thompson (1998, emphasis in original) noted "*empirical* studies consistently show that many researchers do not fully understand the logic of statistical tests" (p. 34). He further stated, "Similarly, many textbooks teach misconceptions regarding these tests (Thompson, 1998, p. 34).

Critical to statistical conclusion validity when reporting *p*-values is to always report and interpret effect sizes. To analyze how much a variable(s) is contributing to that effect, researchers should almost always report structure coefficients. Conducting a "what-if" analysis when interpreting data results and reporting statistical significance findings in terms of "sample size," also helps facilitate understanding of results. As suggested by Thompson (1997):

> All analyses are part of one general linear model. When interpreting results in the context of this model, researchers should generally approach the analysis hierarchically, by asking two questions:
>
> - Do I have anything? (Researchers decide this question by looking at some combination of statistical significance tests, effect sizes...and replicability evidence.)

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

References

All articles used in the review are marked with an asterisk (*).

American Psychological Association. (1994). *Publication manual of the American Psychological Association (4th ed.)*, Washington, DC: Author.

American Psychological Association. (2001). *Publication manual of the American Psychological Association (5th ed.)*, Washington, DC: Author.

Babbie, E. (1992). *The Practice of social research.* Belmont, CA: Wadsworth.

Cohen, J. (1968). Multiple regression as a general data analytic system. *Psychological Bulletin, 70,* 426-443.

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

Daniel, L. & Onwuegbuzie, A. (2001, February). *Multiple regression: A leisurely primer*. Paper presented at the annual meeting of the Southwest Educational Research Association. New Orleans, LA.

Freiberg, H. & Connell, M. (2001). *Effects of consistency management® on student mathematics achievement in seven Chapter I elementary schools, 6(3),* 249-270.

Henson, R. K., & Smith, A. D. (2000). State of the art in statistical significance and effect size reporting: A review of the APA Task Force report and current trends. *Journal of Research and Development in Education, 33,* 285-296.

Kieffer, K. M. & Thompson, B. (1999, November). *Interpreting statistical significance test results: A proposed new "what if" method.* Paper presented at the annual meeting of the Mid-South Educational Research Association, Point Clear, AL.

Kirk, R. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement, 56,* 746-759.

*Kushman, J. & Yap, K. (1999). What makes the difference in school improvement? An impact study of Onward to Excellence in Mississippi schools. *Journal of Education for Students Placed At Risk, 4(3),* 277-99.

Knapp, T. R. (1978). Canonical correlation analysis: A general parametric significance testing system. *Psychological Bulletin, 85,* 410-416.

*Padron, Y. & Waxman, H. (1999). Classroom behavior and learning environment differences between resilient and non-resilient elementary school students. *Journal of Education for Students Placed At Risk, 4(1),* 65-83.

*Sachs, J. (2000). Inequities in early care and education: What is America buying? *Journal of Education for Students Placed At Risk, 5(4),* 384-397.

Snyder, P. & Lawson, S, (1993). Evaluating results using corrected and uncorrected effect size estimates. *Journal of Experimental Education, 61,* 334-349.

Thompson, B. (1992, April). *Interpreting regression results: beta weights and structure coefficients are both important.* Paper presented at the annual meeting (session #20.19) of the American Educational Research Association. San Francisco, CA.

Thompson, B. (1997). Editorial policies regarding statistical significance tests: Further comments. *Educational Researcher, 26(5),* 29-32.

Thompson, B. (1998). Statistical significance and effect size reporting: Portrait of a possible future. *Research in the Schools, 5,* 33-38.

Thompson, B. (1999). Statistical significance tests, effect size reporting, and the vain pursuit of pseudo-objectivity. *Theory & Psychology, 9,* 191-196.

Thompson, B., & Borrello, G. M. (1985). The importance of structure coefficients in regression research. *Educational and Psychological Measurement, 45,* 203-209.

Vacha-Haase, T., Nilsson, J.E., Reetz, D. R., Lance, T.S., & Thompson, B. (2000). Reporting practices and APA editorial policies regarding statistical significance and effect size, Theory *& Psychology, 10,* 413-425.

Wilkinson, L., & American Psychological Association (APA) Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54,* 594-604. (Reprint available through the APA Home Page: http://www.apa.org/journals/amp/amp548594.html).

TM034278

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title:	Multiple Regression Methodology in The Journal of Education for Students Placed At Risk: Effect Sizes and Structure Coefficients
Author(s):	Jennifer C. Herring
Corporate Source: University of North Texas	Publication Date: February 2002

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1

☒

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A

☐

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B

☐

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature: Jennifer C. Herring	Printed Name/Position/Title: Jennifer C. Herring / Res. Associate
Organization/Address: University of North Texas home address: 1109 Aurora Dr, Richardson TX 75081	Telephone: 972-699-1437 FAX: 469-232-1844 E-Mail Address: jherring@ursuline.pvt.k12.tx.us Date: 2/13/02

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

University of Maryland
ERIC Clearinghouse on Assessment and Evaluation
1129 Shriver Laboratory
College Park, MD 20742
Attn: Acquisitions

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
1100 West Street, 2nd Floor
Laurel, Maryland 20707-3598

Telephone: 301-497-4080
Toll Free: 800-799-3742
FAX: 301-953-0263
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com