DOCUMENT RESUME

ED 416 215 TM 028 067

AUTHOR	Henard, David H.
TITLE	Suppressor Variable Effects: Toward Understanding an Elusive Data Dynamic.
PUB DATE	1998-01-00
NOTE	16p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (Houston, TX, January 1998).
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Heuristics; *Multiple Regression Analysis; *Predictive Measurement; *Predictor Variables; *Suppressor Variables

ABSTRACT

The important and sometimes difficult-to-grasp concept of regression suppressor variable effects is explored. An inquiry into the phenomenon of suppressor effects is accomplished via a synthesis of the existing literature and the use of a small heuristic data set to improve the accessibility of the concept. Implications for researchers are also forwarded and it is suggested that the search for suppressor variables in an effort to remove unwanted predictor variable variance may prove less efficient than the search for additional predictor variables which directly explain variance in the criterion variable of interest. However, when they are present, suppressor effects can be critical to note and interpret. (Contains 1 figure, 3 tables, and 14 references.) (Author)

Suppressor Variable Effects: Toward Understanding An Elusive Data Dynamic

David H. Henard
Texas A&M University
College Station, TX 77843-4112

Paper presented at the annual meeting of the Southwest Educational Research Association, Houston, January, 1998.

Abstract

The important and sometimes difficult-to-grasp concept of regression suppressor variable effects is explored. An inquiry into the phenomenon of suppressor effects is accomplished via a synthesis of the existing literature and the use of a small heuristic data set to improve the accessibility of the concept. Implications for researchers are also forwarded and it is argued that the search for suppressor variables in an effort to remove unwanted predictor variable variance may prove less efficient than the search for additional predictor variables which directly explain variance in the criterion variable of interest. However, when they are present, suppressor effects can be critical to note and interpret.

The author wishes to thank Bruce Thompson for his comments on an earlier version of this manuscript.

It is generally understood that the 'usefulness' of a given predictor variable can be measured by the impact that it has on explaining the variance in a respective criterion, or dependent, variable. What may be overlooked, however, is a second way in which a predictor variable can impact predictability--namely, through increasing the predictive power of other predictor, or independent, variables. This phenomenon, termed *suppression*, was first forwarded by Horst (1941) who claimed that suppression resulted from the inclusion (in a regression equation) of a predictor variable having a zero or near-zero correlation with the criterion variable while being correlated with at least one other predictor variable.

The format and intent of this manuscript is threefold. First, a synthesis of existing literature concerning suppressor effects is pursued with the intent of providing the reader with a basic primer on the topic. Second, a heuristic data set that illustrates the various dimensions of the phenomenon is presented. Finally, a discussion as to the relevance of investigating the topic of suppressor effects is forwarded noting some advantages and pitfalls which are likely to be encountered in applied research.

CONCEPTUAL DEVELOPMENT

Classical Definition

In what has become known as the *classical* definition of a suppressor variable, Horst (1966, p. 355) relayed an experience in the prediction of success in pilot training programs during World War II. Comprising the examination battery given to prospective pilots were tests of mechanical, numerical, spatial and verbal ability. Each of the first three predictor variables had positive correlations with the criterion variable, success as a pilot. The fourth, verbal ability, was virtually uncorrelated with pilot success yet had a fairly high correlation with each of the other three predictors. An interesting finding was that when verbal ability was included as a variable in the regression equation, the $\underline{R}^2$ of the model as a whole increased despite the near-zero $\underline{r}$ between verbal ability and pilot training success.

Horst concluded that while verbal ability was not a noteworthy predictor of pilot success, verbal ability <u>indirectly</u> improved the prediction by making the other predictor variables purer measures and thus, improved their predictive power. Simply put, measurement artifacts contaminated the results (Thompson, 1992). By removing (suppressing) the variance accounted for by verbal ability from the equation, the predictive efficiency of the remaining variables was improved.

INSERT FIGURE 1 ABOUT HERE

Figure 1 provides a graphical illustration of how a *classical* suppressor variable operates. X_p represents the first predictor variable. Notice that the correlation between the suppressor variable (X_s) and Y is zero while a correlation does exist between X_p and X_s. This relationship necessarily implies that the structure coefficient (Thompson, 1997; Thompson & Borrello, 1985) equals zero while the $\underline{r}$ between predictor variables is non-zero. The suppressor variable impacts the criterion variable primarily through its impact on the beta weight of the other predictor variable.

Expanded Definitions

Central to the Horst (1941, 1966) definition is the existence of a near-zero structure coefficient and a non-zero beta coefficient. In fact, perfect prediction would result if X_s explains all of the 'remaining' variance in X_p (Smith et al., 1992). In essence,

$$r^2_{Y.Xp.Xs} = 1 - r^2_{Y.Xp}$$

It is a purely theoretical exercise to contemplate a situation in which the suppressor variable will have exactly a zero correlation with the criterion variable. Darlington (1968) forwarded a more general definition of a suppressor by defining it as a variable which, if included in a set of predictors which are positively correlated with the dependent variable, receives a negative regression coefficient when regressed on the population of variables. By introducing the concept of "population", Darlington averted the issue of sampling error (Conger, 1974). This classification is termed *negative* suppression (due to the negative value of the coefficient), yet the fundamental behavior of the variable acts just as it does under the Horst (1941) definition, notably, by increasing predictive power of regression variables through the removal of irrelevant variance in other predictor variables.

The preceding definitions of the suppressor effect are based primarily upon examination of the suppressor beta weights and zero-order correlations. What had been absent in the study of suppressors was the impact that each suppressor variable's beta weight had on the beta of the other predictor variable(s).

Taking this logic further, a suppressor variable is not uniquely defined by its own beta weight but rather through its impact on the weights given to the other predictor variables (Conger, 1974). It should be reiterated here that beta weights are impacted by correlations among the various predictors and may therefore change in value if the correlations or the predictors change (Thompson, 1992). Beta weights should not be interpreted as constants when they are clearly context specific.

Conger (1974) discovered that for both *classical* and *negative* suppressor variables, the regression weights of the suppressed (X_p) variables are increased with the addition of the suppressor variable (X_s). According to Conger (1974, p. 36), a suppressor variable is defined as "a variable which increases the predictive validity of another variable (or set of variables) by its inclusion in the regression equation. This variable is a suppressor only for those variables whose regression weights are increased."

This definition, termed *reciprocal* suppression, subsumes all previous typologies while expanding the application beyond mere two-variable equations to the k-variable case. Thus, a suppressor variable is identifiable by its impact on the beta weights of other independent variables in the regression system rather than merely by its own weight.

HEURISTIC EXAMPLE

To this point, the purpose of this manuscript has been to familiarize the reader with the conceptual development of the suppressor effect in the hope of providing a foundation upon which to build a knowledge base. An examination into the dynamics of each of the three recognized types of suppressors should now prove more beneficial. Let us first consider a bivariate regression equation where Y is regressed on X_1:

$$(1) \quad \mathbf{Yhat = a + \beta_1 X_1}$$

Invoking the *classical* suppressor definition, the addition of a second variable, X_2, into the equation would traditionally have a structure coefficient $r_{Y.X1} = 0$ and a non-zero beta value. The introduction of X_2 suppresses the irrelevant variance in X_1 thereby increasing the $\underline{R}^2$ of the multiple regression equation. Due to its impact on the predictive power of the equation, X_2 will receive a beta coefficient despite its lack of correlation with the criterion variable, Y. The resulting equation will be in the form

$$(2) \qquad \mathbf{Yhat_{CLASSICAL} = a + \beta_1 X_1 + \beta_2 X_2}$$

where β_2 is not equal to zero and β_1 and β_2 have opposite signs.

Assuming the *negative* suppression definition, X_2 is positively correlated with the dependent variable Y yet receives an unanticipated negative beta coefficient value. The equation for negative suppression,

$$(3) \qquad \mathbf{Yhat_{NEGATIVE} = a + \beta_1 X_1 + \beta_2 X_2},$$

resembles equation (2) with the caveats that under negative suppression, $\beta_2 < 0$ and the structure coefficient, $r_{Y.X2} > 0$.

In the *reciprocal* suppressor circumstance, the regression equation will be similar to equations (2) and (3) but with some key differences. First, X_2 will be positively correlated with Y and will receive a positive beta coefficient. Secondly, X_2 will be correlated with X_1 and the addition of X_2 into the regression equation leads to an increase in the value of β_1 over what it would have been had X_2 not entered the equation in addition to an increase in β_2 over what it would have been if X_2 were entered as the sole predictor variable. Thus,

$$(4) \qquad \mathbf{Yhat_{RECIPROCAL} = a + \beta_1 X_1 + \beta_2 X_2}$$

where $\beta_2 > 0$, $\beta_{1\ RECIPROCAL} > \beta_{1\ YHAT}$ and $\beta_{2\ RECIPROCAL} > \beta_{2\ YHAT}$. This interesting dynamic arises when the two independent variables, X_1 and X_2, mutually suppress irrelevant variance in each other (Conger, 1974; Lutz, 1983).

Table 1 is a simple heuristic data set developed by Lutz (1983) which is useful in facilitating an understanding of the phenomenon by highlighting all three suppressor effect types. The variable definitions utilized in the table are:

Y	Criterion Variable Score
X_1	1st Predictor Variable Score
X_{2C}	*Classical* Suppressor Variable Score
X_{2N}	*Negative* Suppressor Variable Score
X_{2R}	*Reciprocal* Suppressor Variable Score

INSERT TABLE 1 ABOUT HERE

From Table 2, we see the inter-variable simple correlations. Only those relationships of interest to our study have their values listed. Notice that the structure coefficients for each of the X_2 variables satisfies the definitional characteristics set forth in the preceding pages. X_{2C} has a zero correlation with Y ($\underline{r}$ = 0.00) and a positive correlation with X_1. Likewise, X_{2N} has a positive correlation with both Y and X_1 while $\underline{r}_{Y.X2R} = 0.46$ yet is negatively correlated with X_1 in the reciprocal case.

INSERT TABLE 2 ABOUT HERE

Finally, by regressing the criterion values against the respective predictor values one obtains the standardized regression coefficients detailed in Table 3. Notice that for each of the three suppressor types, the beta coefficient for X_1 is increased over what it would have been had no additional independent variable been introduced. The beta for X_1, assuming no additional predictors, would have simply been the zero-order correlation between Y and X_1 given in Table 2.

INSERT TABLE 3 ABOUT HERE

In the *classical* example, X_{2C} has a structure coefficient of zero and a non-zero beta. Also true to form, the *negative* suppressor type has a positive structure coefficient and a negative beta while the negative relationship between X_{2R} and X_1 coupled with the positive correlation between X_{2R} and Y produce the definitive *reciprocal* suppressor effect. This small data set was designed to facilitate an understanding of the various effect types. Empirical investigations of suppressor effects broach topics ranging from the relationship between shyness and alcohol consumption (Bruch et al., 1992) to the association between goal orientation and task performance (Hofmann, 1993) and can provide the reader with a more complete understanding of the effect.

IMPLICATIONS FOR RESEARCHERS

Given the previous discussion, at least two implications of suppressor effects emerge as potentially important issues for researchers using regression methods. On an interpretation level, given that a suppressor variable does not require modification of the basic regression model, a practical problem is that suppressor variables may simply be overlooked because of their low zero-order correlations (Velicer, 1978) or seemingly counter-productive beta weights.

From an experimental validity level, the researcher might even question whether or not the pursuit of suppressors is worthwhile. Conger and Jackson (1972, p. 597) claimed that "a suppressor for any given degree of correlation does not yield as much incremental validity as an additional predictor." The question then becomes 'what strategy regarding predictor variables will yield the greatest reward in the explanation of variance?'

Thompson (1992) stated that interpreting beta weights alone is insufficient in analyzing regression results and that further analysis, such as examination of structure coefficients or simple correlations, is necessary if conducting a thorough examination. Evaluation of the structure coefficients or correlation of the criterion-predictor variables will lead to identical conclusions (Pedhazur, 1982; Thompson, 1992) given that they are merely expressed in a different metric.

Given the *classical* suppressor example values from Tables 1 and 2, a less thoughtful researcher may incorrectly assume that X_{2C} provides nothing in the form of predictive power if they simply looked at either the $\beta = -1.00$ or the $r_{Y.X2C} = 0.00$. Likewise, a cursory examination of results from a *reciprocal* effect situation may lead to inaccurate interpretation when both independent variables mutually benefit from each others presence and removal of either one would diminish the predictive power of the regression equation.

As to the incremental validity question, Conger and Jackson (1972)[1] warn that researchers should not expect to find suppressor-predictor r's to be much larger than the

[1] For a more detailed mathematical explanation see pp. 592-596.

criterion-predictor r. A problem frequently arises when large suppressor-predictor correlations are necessary to increase the explanatory power of a set of variables. For example, to increase the R of a regression equation from 0.40 to 0.50, a suppressor variable with $r_{X1.X2} = 0.60$ is required. This is no small requirement.

This leads to the belief that it may be more beneficial for the researcher to focus on uncovering additional predictor variables which explain variance in the criterion variable rather than seeking suppressor variables. Assuming the researcher is not 'data mining' and has developed a thoughtful theoretical rationale for their experimental design, it makes intuitive sense that meaningful predictor variables would be identified in the literature review phase of the research.

That is not to say that suppressor effects should not be considered when they occur. A thorough scrutiny of all relevant variables is expected of a thorough researcher. It is through this scrutiny of both beta coefficients and structure coefficients that the suppressor effect will likely be uncovered and correct interpretations will be formulated. The point, however, is that efforts to increase the predictive power of an experiment may be better served by a sound conceptual development which uncovers noteworthy predictor variables. More may be gained from the discovery of that portion of the criterion variable which is not being predicted than that part of the predictor variable not being used (Conger & Jackson, 1972).

The suppressor variable effect, while empirically rare, provides an opportunity to reiterate the necessity for the researcher to thoroughly analyze and interpret experimental results. Cursory data examination can lead to the dismissal of important variables as well as an incorrect interpretation of system effects. An understanding of suppressor effects in regression analyses can lead to a more complete interpretation of the phenomena under investigation by alerting the researcher to potential areas of insight unseen by less thorough peers.

REFERENCES

Bruch, M.A., Heimberg, R.G., Harvey, C., McCann, M., Mahone, M., & Slavkin, S.L. (1992). Shyness, alcohol expectancies, and alcohol use: discovery of a suppressor effect. Journal of Research In Personality, 26, 137-149.

Conger, A.J. (1974). A revised definition for suppressor variables: a guide to their identification and interpretation. Educational and Psychological Measurement, 34, 35-46.

Conger, A.J., & Jackson, D.N. (1972). Suppressor variables, prediction, and the interpretation of psychological relationships. Educational and Psychological Measurement, 32, 579-599.

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

Hofmann, D.A. (1993). The influence of goal orientation on task performance: a substantively meaningful suppressor variable. Journal of Applied Social Psychology, 23 1827-1846.

Horst, P. (1941). The prediction of personal adjustment. Social Science Research Council Bulletin, 48.

Horst, P. (1966). Psychological measurement and prediction. Belmont, CA: Wadsworth.

Lutz, J.G. (1983). A method for constructing data which illustrate three types of suppressor variables. Educational and Psychological Measurement, 43, 373-377.

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

Smith, R.L., Ager, J.W., & Williams, D.L. (1992). Suppressor variables in multiple regression/correlation. Educational and Psychological Measurement, 52, 17-29.

Thompson, B. (1997). The importance of structure coefficients in structural equation modeling confirmatory factor analysis. Educational and Psychological Measurement, 57, 5-19.

Thompson, B. (1992, April). Interpreting regression results: beta weights and structure coefficients are both important. Paper presented at the annual meeting of the American Educational Research Association, San Francisco. (ERIC Document Reproduction Service No. ED 344 897)

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

Velicer, W.F. (1978). Suppressor variables and the semipartial correlation coefficient. Educational and Psychological Measurement, 38, 953-958.

Table 1
Data Set

id	Y	X_1	X_{2C}	X_{2N}	X_{2R}
1	-1.5	-0.3	0.9	0.2	-0.3
2	-1.0	-0.8	-0.6	-0.3	0.0
3	-0.5	-0.9	-1.4	-0.6	0.2
4	0.0	-0.3	-0.6	-0.2	0.1
5	0.0	0.3	0.6	0.2	-0.1
6	0.5	0.9	1.4	0.6	-0.2
7	1.0	0.8	0.6	0.3	0.0
8	1.5	0.3	-0.9	-0.2	0.3

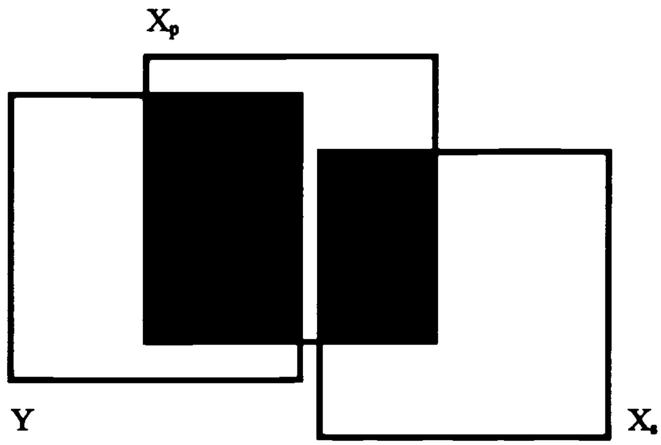
Table 2
Bivariate Correlations

	Y	X_1	X_{2C}	X_{2N}	X_{2R}
Y	--	0.70	0.00	0.23	0.46
X_1		--	0.70	0.85	-0.31
X_{2C}			--	--	--
X_{2N}				--	--
X_{2R}					--

Table 3
Regression Beta Weights

Suppressor Type	β_1	β_2
Classical (X_{2C})	1.40	-1.00
Negative (X_{2N})	1.87	-1.37
Reciprocal (X_{2R})	0.94	0.75

Figure 1
Venn Diagram of Elements of a Suppressor Relationship



TM028067

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: SUPPRESSOR VARIABLE EFFECTS: TOWARD UNDERSTANDING AN ELUSIVE DATA DYNAMIC

Author(s): DAVID H. HENARD

Corporate Source:

Publication Date: 1/22/98

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/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. 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 the identified document, please CHECK ONE of the following options and sign the release below.

[X] Sample sticker to be affixed to document

Check here
Permitting microfiche (4"x 6" film), paper copy, electronic, and optical media reproduction

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
DAVID H. HENARD
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 1

Sample sticker to be affixed to document [ ]

or here
Permitting reproduction in other than paper copy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY
Sample
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 2

Sign Here, Please

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical 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."

Signature: [signature]	Position: RES ASSOCIATE
Printed Name: DAVID H. HENARD	Organization: TEXAS A&M UNIVERSITY
Address: TEXAS A&M UNIVERSITY COLLEGE STATION, TX 77843-4112	Telephone Number: (409) 845-1335
	Date: 1/19/98

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 this 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 which cannot be made available through EDRS).

Publisher/Distributor:	
Address:	
Price Per Copy:	Quantity Price:

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

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

Name and address of current copyright/reproduction rights holder:

Name:

Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

If you are making an unsolicited contribution to ERIC, you may return this form (and the document being contributed) to:

ERIC Facility
1301 Piccard Drive, Suite 300
Rockville, Maryland 20850-4305
Telephone: (301) 258-5500