

Multiple Linear Regression with Qualitative and Quantitative Independent Variables

Multiple regression with both quantitative and qualitative independent variables proceeds in a manner identical to that described previously for regression. The primary difference, now, is how one interprets the estimated regression coefficients.

The Regression Equation and Coefficient Interpretation

For expository purposes, only one qualitative and one quantitative variable will be described in the following discussion. Realize that the use of more variables is a simple extension of the procedures discussed. For the hypothetical example, assume that a researcher wishes to know whether one of three conditions provides higher levels of mathematics achievement: peer tutoring, jigsaw, or think-pair-share. The researcher also has standardized measures of students' academic ability in the form of IQ scores. Thus, the researcher has one categorical variable, type of instruction, and one quantitative variable, IQ scores. The question of interest is whether achievement levels differ among types of instruction once intelligence is taken into account. In this example we are most interested in differences among types of instruction, and IQ is entered only as a control variable, i.e., we wish to statistically control or equate groups. In analysis of covariance (ANCOVA) parlance, IQ is called a *covariate*.

The sample regression equation is thus formulated as follows:

$$Y_i = b_0 + b_1JIG_{1i} + b_2PAIR_{2i} + b_3IQ_{3i} + e_i, \quad (12)$$

where b_0 is the sample intercept; b_1 is the partial regression coefficient for the *adjusted* difference between mean levels of achievement between peer-tutoring (the reference category) and jigsaw; b_2 is the partial difference between think-pair-share and peer-tutoring; b_3 is partial relationship between IQ and Y ; and e_i is the sample error term.

What is meant by the term adjusted? To understand, first consider the following research scenario. Assume that the three groups receiving the three treatments do not have identical levels of intelligence. For example, suppose the students in the jigsaw group have an average IQ level of 110, the students in the think-pair-share group have an average IQ level of 105, and the peer tutoring group an IQ level of 100. Based upon this information, which group do you think would perform best on the achievement test? Probably the jigsaw group because they have the highest IQ level.

Entering IQ into the multiple regression equation allows one to statistically control for the effects of IQ, and thus gain a better estimate of the true difference in achievement due to the effects of instruction. That is, multiple regression allows us to partial-out the effects of IQ so that these effects will no longer be confounded with instruction. So when one compares mean differences in instruction, such as the difference between jigsaw and peer-tutoring, the mean difference comparison should be based upon the partialled effects (IQ partialled-out), and these partialled estimates of the mean differences are based upon the adjusted means for the groups. Adjusted means are estimated means in achievement, controlling for IQ.

Adjusted means may be obtained by substituting into the sample regression equation the sample mean of the covariate or covariates. For example, the sample data for the hypothetical research question is given in Table 1.

Table 1: Fictional Achievement Scores by Type of Instruction and IQ

Achievement	Instruction	IQ
73	peer	101
72	peer	100
70	peer	105
75	peer	95
80	jig	100
79	jig	107
83	jig	101
84	jig	99
87	pair	104
89	pair	108
88	pair	103
87	pair	105

The sample regression equation for this data is

$$Y' = 110.275 + 9.565(\text{JIG}) + 17.04(\text{PAIR}) + -0.377(\text{IQ}). \quad (13)$$

Recall previous discussions of calculating predicted means for each of the three groups with dummy variables in regression. For instance, the predicted mean for the peer-tutoring group would be the constant, or 110.275, and the predicted mean for the jigsaw group would be $b_0 + b_1$ or $110.275 + 9.565 = 119.84$.

Since these predicted means are well above the actual scores obtained—which range from 70 to 89—it should be obvious the procedure of obtaining predicted means is flawed because the statistical effects of IQ were not taken into account. To correct this problem, now one must take into account the effect of IQ on the predicted means. To do so requires only that the regression coefficient corresponding to the covariate be multiplied by the covariate sample mean, e.g., b_3 , the coefficient for IQ, should be multiplied by the sample mean of IQ, which is 102.333. For example, the adjusted mean for the peer tutoring group is:

$$\begin{aligned} Y' &= 110.275 + 9.565(0) + 17.04(0) + -0.377(102.333), \\ Y' &= 110.275 + \quad \quad \quad -38.58, \\ Y' &= 71.695. \end{aligned}$$

The adjusted mean for the jigsaw group is:

$$\begin{aligned} Y' &= 110.275 + 9.565(1) + 17.04(0) + -0.377(102.333), \\ Y' &= 110.275 + 9.565(1) \quad \quad \quad + -38.58, \\ Y' &= 119.84 + -38.58, \\ Y' &= 81.26, \end{aligned}$$

and the adjusted mean for the think-pair-share group is:

$$\begin{aligned} Y' &= 110.275 + 9.565(0) + 17.04(1) + -0.377(102.333), \\ Y' &= 110.275 \quad \quad \quad + 17.04(1) + -38.58, \\ Y' &= 127.315 + -38.58, \\ Y' &= 88.735. \end{aligned}$$

The adjusted means just calculated reflect the statistical effects of IQ upon achievement controlling for type of instruction. Table 2 provides both the adjusted and observed means for each group.

Note that only slight differences exist; this suggests that the average level of IQ was about the same for the three groups since little difference exists between observed and adjusted means.

Table 2

Group	Observed Mean	Adjusted Mean
Peer	72.50	71.70
Jigsaw	81.50	81.26
Pair	87.75	88.74

Regression coefficients are interpreted in a manner described previously under multiple regression. The coefficients for dummy variables represent the mean difference between the reference group and the group represented by the dummy, controlling for the effect of the covariates or other independent variables in the regression model. For example, the difference between the jigsaw group and the peer tutoring group is estimated to be 9.565, which is the same difference between the adjusted means for the two group, within rounding error, i.e., $81.26 - 71.70 = 9.56$. And the coefficient for IQ, b_3 , is simply the estimated relationship between IQ and achievement, controlling for instruction (note that in this case, the relationship is expected to be negative!).

Overall Model Fit and Statistical Inference

The usual measures of fit and inference for the overall models continues to apply here.

Global Effects, ΔR^2 , and the Partial F Test of ΔR^2

Statistical inference regarding the global effect, as measured by $\Delta R^2(X_k)$, continues to hold here. To illustrate, the overall statistical effect of instruction upon achievement will be tested. The reduced model contains only IQ:

$$Y_i = b_0 + b_3IQ_{3i} + e_i, \quad (14)$$

and the full model contains all variables, which, in this case, is instruction and IQ:

$$Y_i = b_0 + b_1JIG_{1i} + b_2PAIR_{2i} + b_3IQ_{3i} + e_i. \quad (12)$$

The null hypothesis for the instruction effect is

$$H_0: \Delta R^2(\text{instruction}) = \Delta R^2(\text{JIG, PAIR}) = 0.00.$$

This null states that the combination of dummies representing instruction, JIG and PAIR, do not make an increase in the overall model fit as measured by the model R^2 .

The resulting statistics from the reduced model are: $R_r^2 = .373$ and $df_{2r} = 10$. The full model statistics needed are: $R_f^2 = .964$ and $df_{2f} = 8$. The increment in the model R^2 due to adding instruction to the model, denoted $\Delta R^2(\text{JIG,PAIR})$ or $\Delta R^2(\text{instruction})$, is

$$\Delta R^2(\text{instruction}) = R_f^2 - R_r^2 = .964 - .373 = 0.591.$$

Degrees of freedom for this test are

$$df_1 = df_{2r} - df_{2f} = 10 - 8 = 2, \text{ and}$$

$$df_2 = df_{2f} = 8.$$

Calculation of the partial F statistic follows:

$$F = \frac{\Delta R^2(X_k)/(df_{2r} - df_{2f})}{(1 - R^2)/df_{2f}}$$
$$= \frac{.591/(10-8)}{(1 - .964)/8} = \frac{.2955}{(.036)/8} = \frac{.2955}{0.0045} = 65.66.$$

The critical value at the 5% significance level is $.05F_{2,8} = 4.46$. Since 65.66 is clearly larger than 4.46, the null is rejected, and one may conclude that inclusion of instruction statistically reduces variance error in Y.

Inferential Procedures for Regression Coefficients

Following the reject of overall model null, $H_0: R^2 = 0.00$, one should next test the partial effect of each distinct independent variable. This has just been illustrated for instruction using $\Delta R^2(X_2, X_3)$ and the partial F test. Since the remaining variable, IQ, is quantitative, the simple t-ratio will suffice for testing $H_0: \beta_3 = 0.00$.

Pairwise Comparisons Among Categories and Control of the Familywise Type I Error Rate

Since instruction is a categorical variable with more than two categories, some type of control must be introduced to check the possible inflation of the Type I error rate. Assuming the experiment was developed with the express purpose of comparing these three techniques, one may use the Bonferroni procedure for control of α_{fw} .

There are a total of three pairwise comparisons, so the α_{fw} will be $.05/3 = 0.0167$, or, similarly, the corrected Bonferroni critical t is ± 3.005 . So when constructing a table of comparisons, the calculated t-ratio for each comparison, where $t = \text{estimated mean difference}/\text{standard error of difference}$, should be compared against the Bonferroni critical t. Also, one may construct .95CIs using the Bonferroni adjustment for each comparison.

Reporting Multiple Regression

The SPSS results for this data analysis are reported below.

* * * * M U L T I P L E R E G R E S S I O N * * * *

	Mean	Std Dev	Label
ACH	80.583	6.762	
JIG	.333	.492	
PAIR	.333	.492	
IQ	102.333	3.701	

N of Cases = 12

Correlation, 1-tailed Sig:

	ACH	JIG	PAIR	IQ
ACH	1.000	.100	.783	.373
	.	.378	.001	.116
JIG	.100	1.000	-.500	-.116
	.378	.	.049	.359
PAIR	.783	-.500	1.000	.532
	.001	.049	.	.037
IQ	.373	-.116	.532	1.000
	.116	.359	.037	.

Multiple R	.98189
R Square	.96410
Adjusted R Square	.95064
Standard Error	1.50226

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	3	484.86232	161.62077
Residual	8	18.05435	2.25679

F = 71.61523 Signif F = .0000

----- Variables in the Equation -----

Variable	B	SE B	95% Confdnce	Intrvl B	Beta
JIG	9.565217	1.085107	7.062958	12.067476	.696516
PAIR	17.039855	1.272936	14.104462	19.975248	1.240801
IQ	-.376812	.147664	-.717326	-.036297	-.206246
(Constant)	110.275362	14.822398	76.094887	144.455838	

----- in -----

Variable	T	Sig T
JIG	8.815	.0000
PAIR	13.386	.0000
IQ	-2.552	.0341
(Constant)	7.440	.0001

Table 3
Descriptive Statistics and Correlations Among Achievement, IQ and Type of Instruction

Variable	Correlations			
	Achievement	Jigsaw	Pair	IQ
Achievement	---			
Jigsaw	.10	---		
Pair	.78*	-.50	---	
IQ	.37	-.12	.53	---
Mean	80.58	0.33	0.33	102.33
SD	6.76	0.49	0.49	3.70

Note. Jigsaw and Pair are dummy variables representing students in the jigsaw group (1 = jigsaw, 0 = others) and think-pair-share group (1 = pair, 0 = others); n = 12.

*p<.05.

Table 4
Regression of Mathematics Achievement Scores on Instruction and IQ

Variable	b	se	ΔR^2	95%CI	F	t
IQ	-0.38	0.15	.03	-0.72, -0.04		-2.55*
Instruction			.83		91.92*	
Jigsaw	9.57	1.09		7.06, 12.07		8.82*
Pair	17.04	1.27		14.10, 19.98		13.39*
Intercept	110.29	14.82		76.10, 144.46		7.44*

Note. $R^2 = .96$, adj. $R^2 = .95$, $F_{3,8} = 71.62^*$, $MSE = 2.26$, $n = 12$. ΔR^2 represents the semi-partial multiple correlation or the increment in R^2 due to adding the respective variable. Jigsaw and Pair are dummy variables representing students in the jigsaw group (1 = jigsaw, 0 = others) and think-pair-share group (1 = pair, 0 = others)

*p<.05.

Table 5
Comparisons of Mathematics Achievement Scores Among Instruction Types

Contrast	Estimated Adjusted Mean Difference	Standard Error of Difference	Bonferroni Adjusted .95CI
Jigsaw vs. Peer	9.565*	1.085	12.83, 6.30
Pair vs. Peer	17.040*	1.273	20.87, 13.21
Jigsaw vs. Pair	-7.475*	1.166	-3.97, -10.98

*p<.05, where p-values are adjusted using the Bonferroni method.

Both intelligence and instruction were statistically related to mathematics achievement. All pairwise comparisons were performed and all were statistically significant at the 5% level using the Bonferroni adjustment. Students in the think-pair-share strategy showed the highest levels of achievement, and students using peer tutoring demonstrated the lowest levels of achievement.

Exercises

Assume all comparisons were planned in advanced.

(1) According to the leadership literature, there are a number of different leadership styles. Listed below are scores obtained from an instrument designed to measure a particular leadership style, which will be referred to as style X. Of interest is whether X differs by school district type in terms of urbanity, and by years of experience. A stratified random sample of school principals were selected from three district types (mostly urban, mostly suburban, and mostly rural).

The scores on style X range from 100 to 0. The closer the score to 100, the more the respondent conforms to style X, while the closer the score to 0, the less the respondent conforms to style X.

Is there any evidence that X differs among the three district types, once years of experience is taken into account?

Years of Experience	District Type	Style X
21	urban	85
20	urban	98
18	urban	75
10	urban	63
17	urban	91
5	urban	49
4	urban	62
7	suburban	49
1	suburban	48
9	suburban	56
10	suburban	78
2	suburban	35
4	suburban	50
1	rural	33
16	rural	95
3	rural	26
2	rural	11
2	rural	33
8	rural	25
14	rural	65

(2) A researcher is interested in learning whether frequency of reading at home to elementary-aged children produces differential effects on reading achievement. After obtaining information from a randomly selected sample of parents about this behavior, the following classifications and standardized achievement scores were recorded. (Note: frequency classifications as follows: a = less than once per month, b = once to three times per month, c = more than three times per month.) In addition to reading frequency, information regarding the family's status concerning whether or not the family's child receives either free or reduced lunch is recorded as a proxy for SES. Also, students' IQ scores were recorded.

IQ	SES	Freq. of Reading	Achievement
100	4	a	48
101	3	a	37
98	3	a	47
105	5	a	65
86	4	b	57
92	3	b	39
117	5	b	49
99	2	b	45
105	5	c	61
93	3	c	55
101	5	c	51
103	1	c	30

Note. SES scores range from a low of 1 to a high of 5.

Is frequency of reading at home related to student reading achievement once SES and IQ are taken into account?

(3) An administrator wishes to know whether student behavioral problems can be linked to student performance. If students were suspended or reprimanded more than once, they are classified as having behavioral problems. In addition, each student's SES is known, and should be taken into account. The administrator randomly selects 13 students and collects the appropriate data.

Student	GPA	Student SES	Behavioral Problems
Bill	3.33	5	n
Bob	1.79	1	y
Stewart	2.21	4	n
Linda	3.54	5	y
Lisa	2.89	4	n
Ann	2.54	3	n
Fred	2.66	5	y
Carter	1.10	1	y
Bill	3.10	4	n
Sue	2.10	2	y
Kara	2.07	2	y
Loser	2.31	3	n
Kathy	3.67	4	n

Note. SES scores range from a low of 1 to a high of 5.

Answers to be added.

Analysis of Covariance (ANCOVA) Results for Sample Data

ANOVA VARIABLES=ach BY instruct(1 3) WITH iq /MAXORDERS ALL
/STATISTICS REG /METHOD UNIQUE /FORMAT LABELS .

* * * A N A L Y S I S O F V A R I A N C E * * *

ACH
by INSTRUCT
with IQ

UNIQUE sums of squares
All effects entered simultaneously

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Covariates	14.696	1	14.696	6.512	.034
IQ	14.696	1	14.696	6.512	.034
Main Effects	414.904	2	207.452	91.923	.000
INSTRUCT	414.904	2	207.452	91.923	.000
Explained	484.862	3	161.621	71.615	.000
Residual	18.054	8	2.257		
Total	502.917	11	45.720		

Covariate Raw Regression Coefficient

IQ -.377

12 cases were processed.
0 cases (.0 pct) were missing.

OLS Assumptions and Assessment of Violations

Linear regression requires several assumptions in order for proper coefficient estimates be produced. The assumptions are:

(1) *Fixed X*: This means that the values of X are known and can be held constant or controlled by the researcher, such as treatments in an experiment. Usually in correlational or non-experimental research, such as data derived from surveys, this assumption is not met. Fortunately, results obtained from linear regression do not necessarily need for this assumption to be held. In short, violations of this assumption typically are not serious.

(2) *No Measurement Error for X*: This means that the instruments used to derive scores for X are perfectly reliable (and valid). Of course this will never happen in practice. Unreliable measures result in increases in coefficient standard errors, and therefore a decrease in precision. As a result, the power of statistical tests is decreased—thus it is more difficult to obtain statistically significant results as measurement error (or unreliability) increases. Also, regression coefficients for X will tend to be underestimated, in absolute value, thus further decreasing power and precision. When multiple X's are present in the regression model, the effects of measurement error are less predictable, but very problematic in terms of power and precision.

(3) *Regression of Y on X is Linear*: This means that the relationship between X and Y can be depicted using a straight line. This assumption is not stringent, and non-linear models involving X and Y will be discussed under the topic of polynomial regression. In short, should a non-linear or curvilinear regression situation arise, it is very easy to alter the regression equation to better model this non-linearity.

(4) *Assumptions for Errors*: Recall that residuals, or errors, in the regression equation are defined as $Y' - Y$, predicted Y minus observed Y. Residuals, then, simply represent variation in Y left over from that which is explained using X (or X's). For proper OLS estimation, the following assumptions concerning the residuals are needed:

(a) *mean of errors is zero*: For each level of X, the mean of the residuals should be equal to zero.

(b) *errors are uncorrelated*: Each residual is uncorrelated with other residuals. This assumption simply means that the observations of Y are independently distributed. Should the data be derived in a manner in which correlated observations are collected, then OLS is not the most efficient estimation procedure and alternative regression techniques will be needed.

(c) *homoscedasticity*: For each level of X, the variance of the residuals should be constant (or approximately equal).

(d) *errors not correlated with X's*: The errors should be uncorrelated with observations on X. If not, then additional IVs are needed in the model to remove this lack of independence.

(5) *Normality*: For statistical tests in regression, the residuals are assumed to be normally distributed for each level of X. Normality is only needed for inferential tests, and does not affect OLS estimation. Normality typically is not a problem, even for severely non-normal distributions, due to the central limit theorem.

Another very important error in regression is the *specification* error. Specification errors occur when the model specified by the regression equation is not correct, thus resulting in *model*

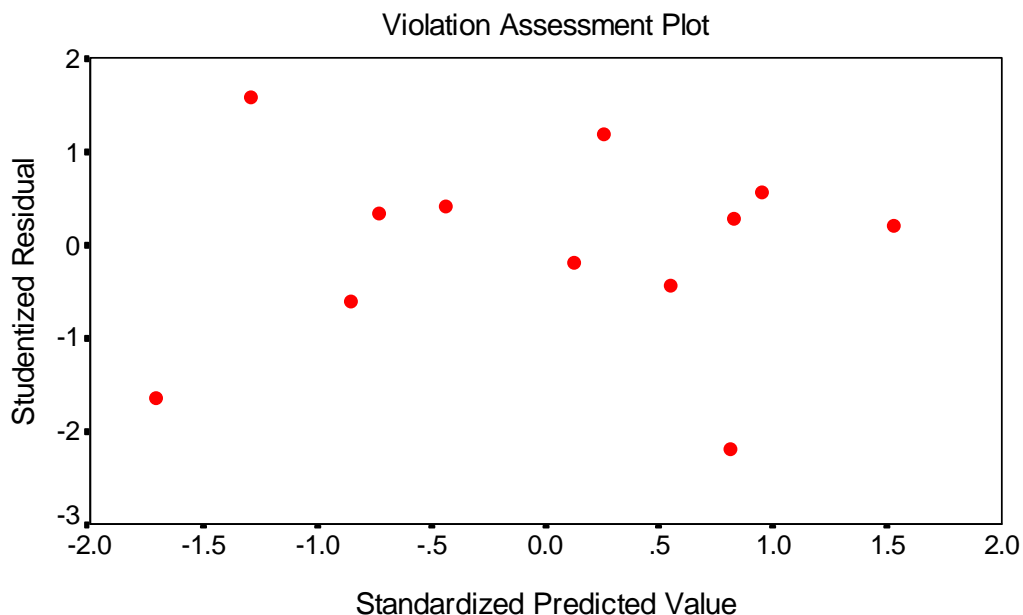
misspecification. Model misspecification refers to the "(1) omission of relevant variables from the equation; (2) inclusion of irrelevant variables in the equation; [or] (3) specifying that the regression is linear when it is curvilinear" (Pedhazur, 1982, p. 35). Misspecification also results when one fails to properly model interactions between IVs (to be discussed later). Model misspecification will lead to errors in hypothesis testing and interpretation of regression coefficients. This is a serious problem as model misspecification could potentially lead a researcher to make unwarranted conclusions.

Misspecification is the most common problem in all regression analyses. Perhaps the best way to avoid misspecification is to (1) think, (2) understand the theory behind the model, and (3) examine your data. Carefully consider possible relationships among potential X's and Y, and always consider interactions among the X's and possible curvilinear relationships.

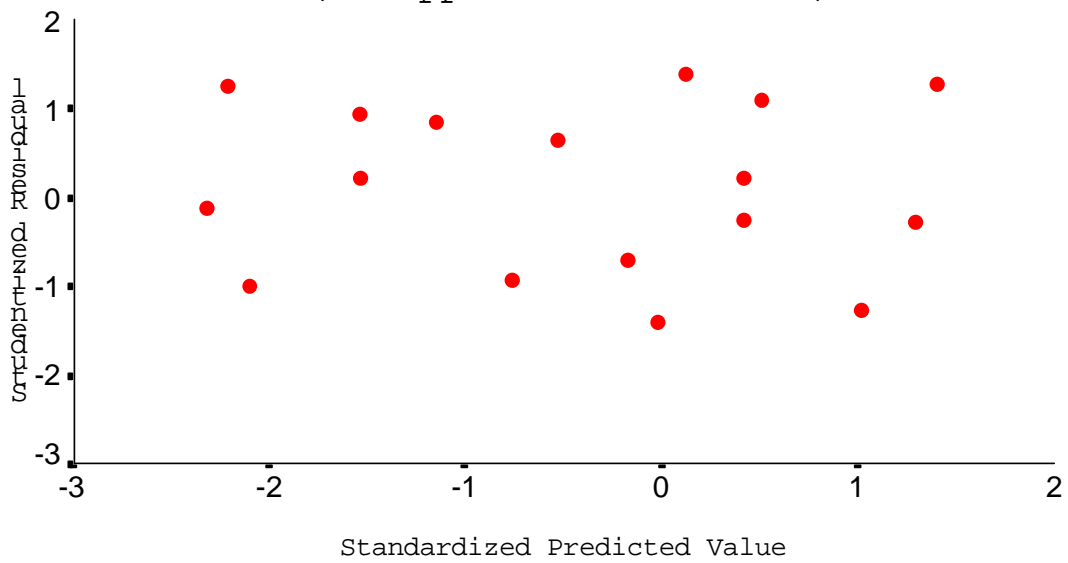
The standard procedure for assessing whether any of the above assumptions may be violated for a given regression requires that residuals and predicted values be plotted in a scatterplot. Both should be in standardized form. Specifically, the residuals should be Studentized Residuals (SRESID), and the predicted values should be in Z score form (e.g., ZPRED in SPSS).

Once the regression is completed, SPSS saves the predicted and residuals values into new columns that may be plotted together using the GRAPH command. On the Y axis should be SRESID and on the X axis should be ZPRED.

The pattern of dots in the scatterplot should form, roughly, a rectangle with no clear pattern or grouping of the scatter. Below is an example of a residual plot in which no violations are apparent.

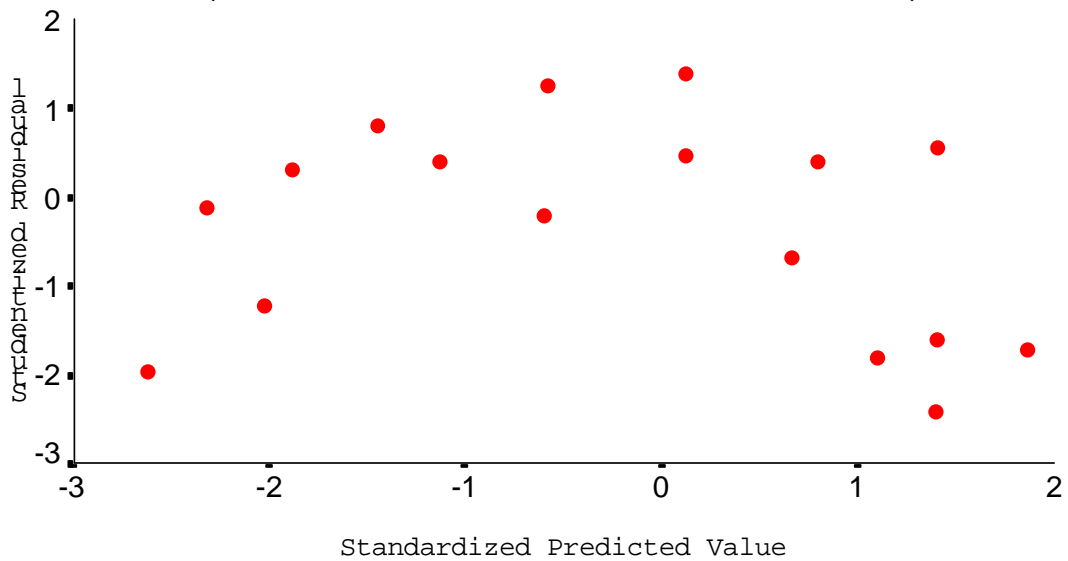


Assessing Assumption Violations
(No Apparent Violations)

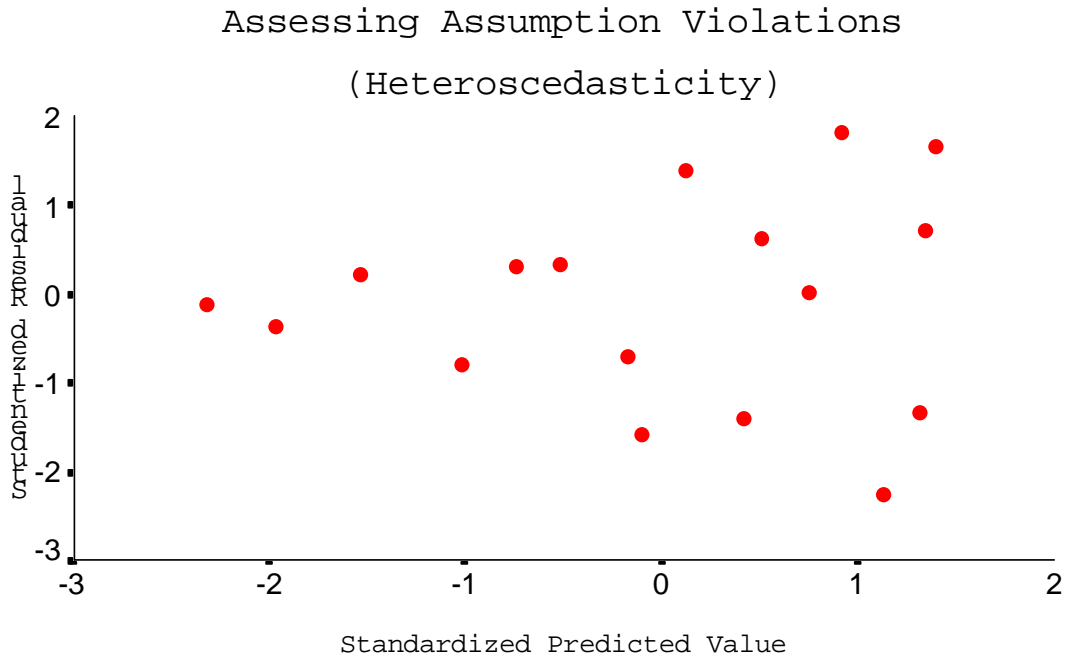


The scatter below shows a clear violation of the linearity assumption.

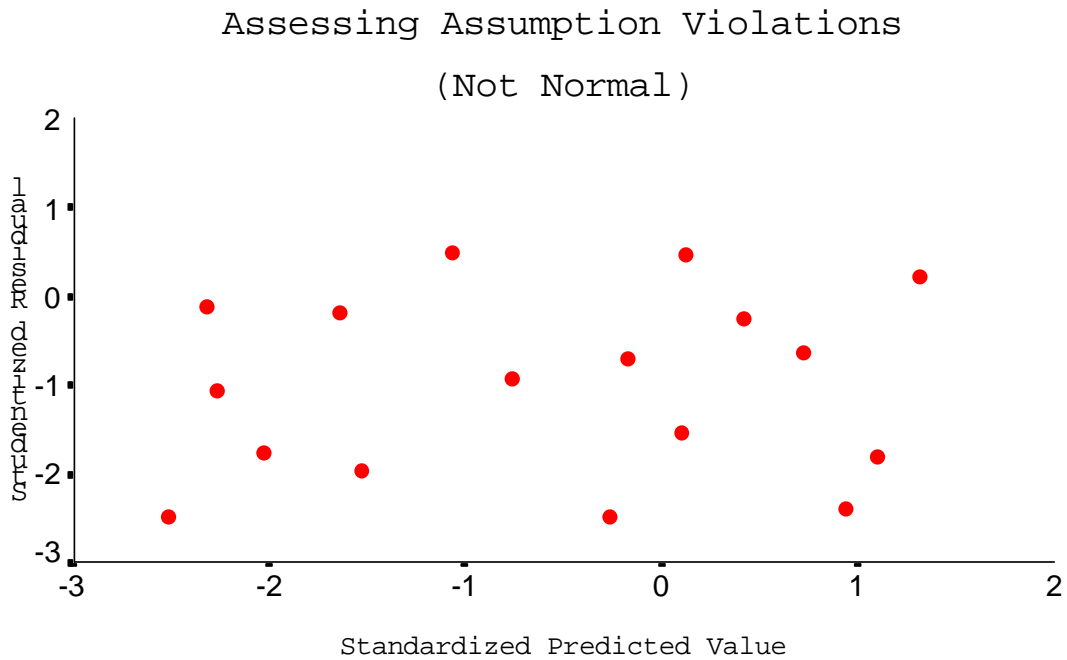
Assessing Assumption Violations
(Correct Fit Problem--Not Linear)



This scatterplot reveals a possible violation with constant variance (homoscedasticity).

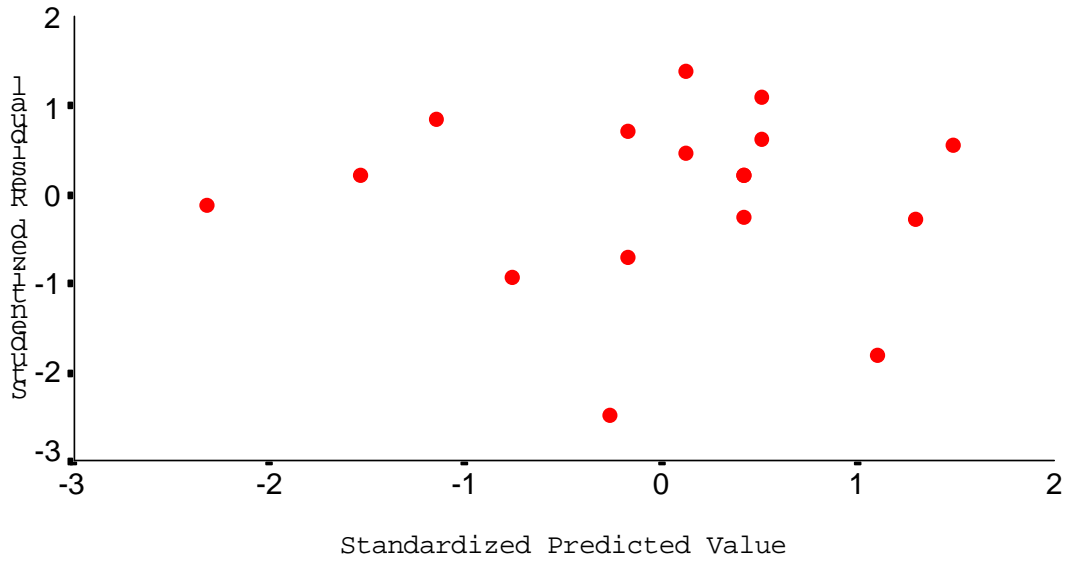


The scatterplot below reveals what may be a violation of the normality assumption. Note that most points are below 0.00 on the vertical axis.



The scatterplot below was taken from the example data on ratings and grades. It is difficult to determine whether a violation exists, therefore the lack of clear evidence suggests that regression will be appropriate for this data.

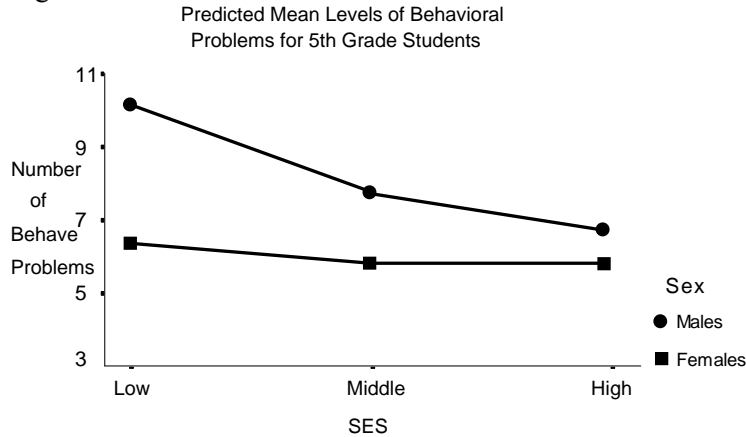
Assessing Assumption Violations (Student ratings and percent A's)



Interactions

An interaction occurs when an independent variable's statistical effects (or differences) upon the dependent variable vary or differ across levels of a second independent variable. For example, if one were interested in examining the relationship between SES, sex, and the number of behavioral problems displayed among 5th grade students, one may find a pattern such as the one depicted in Figure 1.

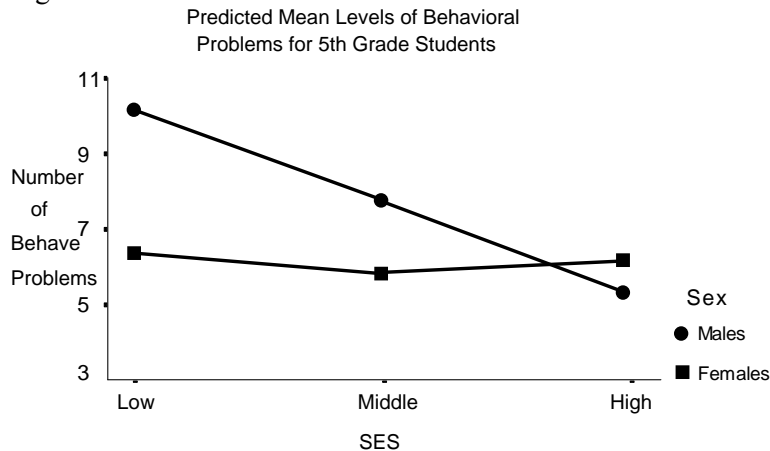
Figure 1



As illustrated in Figure 1, the difference in behavioral problems between males and females changes across levels of SES; thus, there is an interaction between SES and sex.

Note that an interaction only occurs between independent variables; the dependent variable does not interact with an independent variable. There are two types of interactions, ordinal and disordinal. Figure 1 illustrates an ordinal interaction. An ordinal interaction occurs when one group's predicted means is always greater than another group's predicted means. For example, the predicted male means are always greater than predicted female means, yet the differences between males and females varies by SES, therefore an ordinal interaction results.

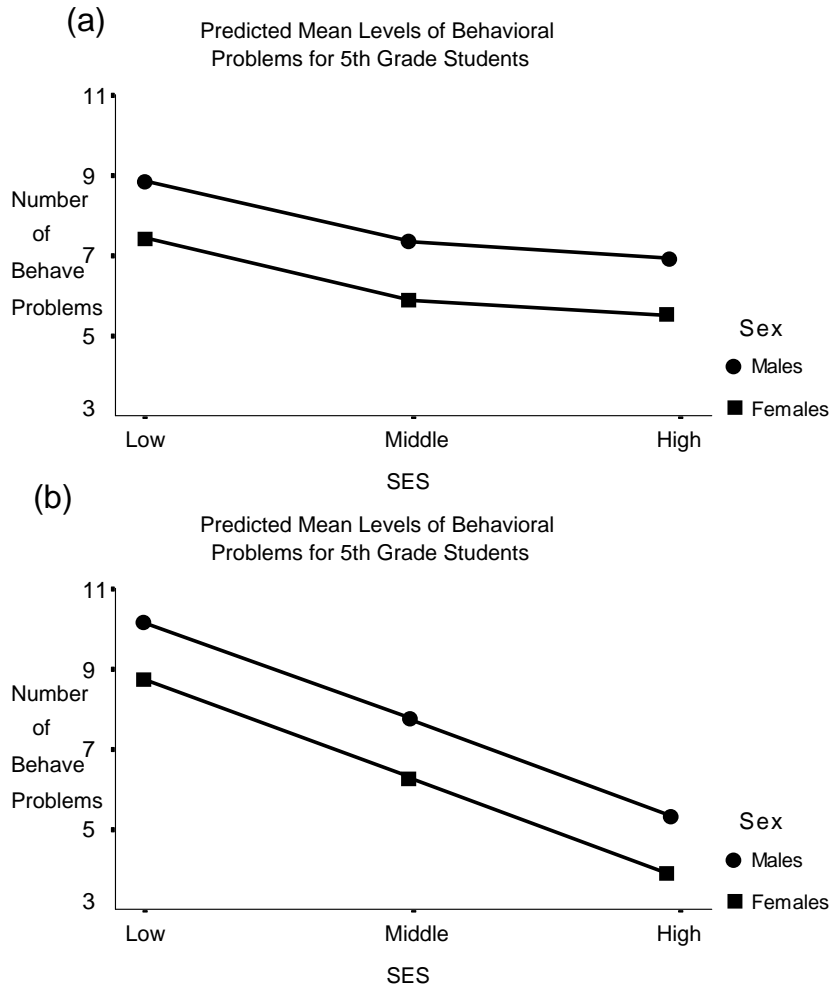
Figure 2



When two or more group means switch or cross, a disordinal interaction occurs. Figure 2 illustrates a disordinal interaction. Note that predicted male means are higher for low and middle levels of SES, but the predicted female mean for high levels of SES is greater than the predicted male mean.

Figures 3a and 3b, which contain no interactions, are included as a reference. Note that the differences between males and females is constant in both a and b, i.e., the sex difference does not vary by SES level.

Figure 3



The Regression Equation for Interaction Between Two Categorical IVs

Assume that we are interested in modeling the number of behavioral problems displayed by 5th grade students, and the two variables used to model this behavior are sex and SES. We may expect that the differences between males and females will not be constant for each level of SES; therefore an interaction is expected. The sample regression model *without* the interaction is

$$Y' = b_0 + b_1\text{MALE} + b_2\text{HIGH} + b_3\text{LOW} \tag{1}$$

where MALE is a dummy coded 1 if the student is male, and 0 if female; HIGH is a dummy indicating whether the student has a high level of SES (1) or not (0); and LOW is a dummy coded 1 for students with lower levels of SES and 0 if otherwise.

To model an interaction, one must simply create a multiplicative term, i.e., for two independent variables, X_1 and X_2 , the interaction is found by including in the regression model the following product:

interaction = $X_1 * X_2$.

When one or more of the independent variables is categorical, each dummy representing the variable must be multiplied. For the current example, there will be two multiplicative terms to include in the regression model:

$$INT_1 = MALE * HIGH,$$

and

$$INT_2 = MALE * LOW.$$

Entering these terms into the sample regression model results in the following equations:

$$Y' = b_0 + b_1MALE + b_2HIGH + b_3LOW + b_4INT_1 + b_5INT_2 \quad (2)$$

or

$$Y' = b_0 + b_1MALE + b_2HIGH + b_3LOW + b_4(MALE*HIGH) + b_5(MALE*LOW). \quad (3)$$

Sample data for this model is given in Table 1.

Table 6
Fictional Data Demonstrating Interaction

behave	sex	ses	MALE	HIGH	LOW	MALE*HIGH	MALE*LOW
9	m	h	1	1	0	1	0
11	m	h	1	1	0	1	0
12	m	h	1	1	0	1	0
10	m	h	1	1	0	1	0
4	f	h	0	1	0	0	0
5	f	h	0	1	0	0	0
6	f	h	0	1	0	0	0
4	f	h	0	1	0	0	0
5	m	m	1	0	0	0	0
6	m	m	1	0	0	0	0
4	m	m	1	0	0	0	0
6	m	m	1	0	0	0	0
2	f	m	0	0	0	0	0
4	f	m	0	0	0	0	0
3	f	m	0	0	0	0	0
3	f	m	0	0	0	0	0
6	m	l	1	0	1	0	1
7	m	l	1	0	1	0	1
5	m	l	1	0	1	0	1
5	m	l	1	0	1	0	1
3	f	l	0	0	1	0	0
2	f	l	0	0	1	0	0
1	f	l	0	0	1	0	0
3	f	l	0	0	1	0	0

Hypothesis Testing

As usual, the first hypothesis to test is that of the overall model: $H_0: R^2 = 0.00$. If this is rejected, the next hypothesis to test is that of the interaction effect, i.e., $H_0: \Delta R^2(\text{interaction}) = 0.00$ or $H_0: \Delta R^2(\text{MALE*HIGH, MALE*LOW}) = 0.00$. This will be illustrated for the sample data.

For the sample data, the reduced model R^2 is .836 with $df_2 = 20$. The full model R^2 is .9035 with $df_2 = 18$, so

$$\Delta R^2 = .9035 - .836 = 0.0675.$$

The partial F ratio is

$$F = \frac{\Delta R^2(X_k)/(df_{2r} - df_{2f})}{(1 - R_f^2)/df_{2f}}$$
$$= \frac{.0675/(20-18)}{(1 - .9035)/18} = \frac{0.0337}{0.0054} = 6.24.$$

which is statistically significant at the .05 level.

When an interaction is statistically significant, one need not perform significance testing for the independent variables associated with the interaction, i.e., one need not compute ΔR^2 and partial F tests for independent variables associated with the interaction.

Coefficient Interpretation

With interactions, one cannot simply interpret main effects; that is, one cannot simply rely upon the regression coefficients produced for the two independent variables. When interpreting the results, one must take into account the interaction. For example, consider interpretations for the current sample data. The regression results are

$$Y' = 3.00 + 2.25(\text{MALE}) + 1.75(\text{HIGH}) + -0.75(\text{LOW}) \\ + 3.50(\text{MALE*HIGH}) + 1.25(\text{MALE*LOW}).$$

One might be tempted to claim that the estimated difference between males and females is 2.25 since this is the main effect of sex. However, since there are interactions present between sex and SES, one must examine the *simple main effects* for each variable.

What is meant by simple main effects? This means that one must interpret the effects of sex for each level or category of the variable in which sex interacts. That is, indicate whether males or females differ, and by how much, for each unique level of SES. Similarly, for each sex, indicate how the SES levels differ.

The easiest way to show these effects is to calculate the predicted means for each cell or combination of variables and to plot these on a graph like Figures 1 through 3. For example, the predicted means for males and females for each level of SES can be found using the regression equation, as demonstrated below:

$$Y' = 3.00 + 2.25(\text{MALE}) + 1.75(\text{HIGH}) + -0.75(\text{LOW}) \\ + 3.50(\text{MALE*HIGH}) + 1.25(\text{MALE*LOW}).$$

SES = LOW
SEX = MALE

$$Y' = 3.00 + 2.25(1) + 1.75(0) + -0.75(1) + 3.50(0) + 1.25(1). \\ Y' = 3.00 + 2.25 - 0.75 + 1.25 = 5.75.$$

SES = LOW
SEX = FEMALE

$$Y' = 3.00 + 2.25(0) + 1.75(0) + -0.75(1) + 3.50(0) + 1.25(0). \\ Y' = 3.00 - 0.75 = 2.25.$$

SES = MIDDLE
SEX = MALE

$$Y' = 3.00 + 2.25(1) + 1.75(0) + -0.75(0) + 3.50(0) + 1.25(0). \\ Y' = 3.00 + 2.25 = 5.25.$$

SES = MIDDLE
SEX = FEMALE

$$Y' = 3.00 + 2.25(0) + 1.75(0) + -0.75(0) + 3.50(0) + 1.25(0). \\ Y' = 3.00 = 3.00.$$

SES = HIGH
SEX = MALE

$$Y' = 3.00 + 2.25(1) + 1.75(1) + -0.75(0) + 3.50(1) + 1.25(0). \\ Y' = 3.00 + 2.25 + 1.75 + 3.50 = 10.5.$$

SES = HIGH
SEX = FEMALE

$$Y' = 3.00 + 2.25(0) + 1.75(1) + -0.75(0) + 3.50(0) + 1.25(0). \\ Y' = 3.00 + 1.75 = 4.75.$$

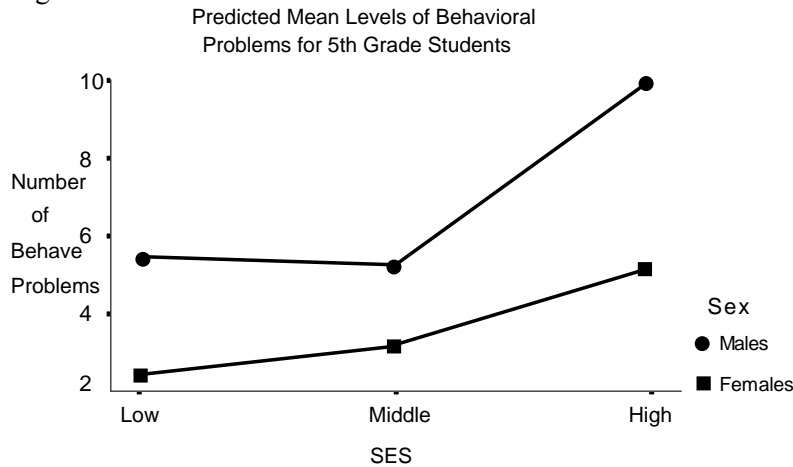
So the predicted means are:

Table 7

SEX	SES		
	Low	Middle	High
Male	5.75	5.25	10.5
Female	2.25	3.00	4.75

which can be better illustrated via a graph, such as Figure 4.

Figure 4



Now the only thing left to do test the statistical significance of the various comparisons. This is easily accomplished using the computer procedure to calculate standard errors. For example, the difference between males and females, when SES is held constant at the middle level is $b_1 = 2.25$, which has a standard error of 0.707. In Table 2 the difference between males and females for each level of SES is calculated and reported.

Table 8

SEX	SES		
	Low	Middle	High
Male	5.75	5.25	10.5
Female	2.25	3.00	4.75
Sex Difference (standard error)	3.50 (0.707)	2.25 (0.707)	5.75 (0.707)

To find the estimated difference between males and females for low levels of SES, simply re-specify the regression model so that low SES represents the omitted reference group. When doing this, be careful that the proper interaction terms are also included (in this case, no interactions based upon the LOW dummy should be included). For example, to find the estimated difference between the sexes for low levels of SES, the revised regression equation is

$$Y' = b_0 + b_1\text{MALE} + b_2\text{HIGH} + b_3\text{MIDDLE} + b_4(\text{MALE}*\text{HIGH}) + b_5(\text{MALE}*\text{MIDDLE}). \quad (4)$$

The resulting statistics are:

$$Y' = 2.25 + 3.50(\text{MALE}) + 2.50(\text{HIGH}) + 0.75(\text{MIDDLE}) + 2.25(\text{MALE}*\text{HIGH}) + -1.25(\text{MALE}*\text{MIDDLE}),$$

and the standard error for the difference between men and women is 0.707. To find the estimated difference and standard error for sex at high levels of SES, the following equation should be used:

$$Y' = b_0 + b_1\text{MALE} + b_2\text{MIDDLE} + b_3\text{LOW} + b_4(\text{MALE}*\text{MIDDLE}) + b_5(\text{MALE}*\text{LOW}). \quad (5)$$

The resulting statistics are:

$$Y' = 4.75 + 5.75(\text{MALE}) - 1.75(\text{MIDDLE}) - 2.50(\text{LOW}) \\ - 3.50(\text{MALE} * \text{MIDDLE}) - 2.25(\text{MALE} * \text{LOW}).$$

One may also compare the three levels of SES for each distinct sex category using a method similar to the one just described. For this set of comparisons, either Scheffé or Bonferroni should be used for each comparison set (i.e., each set of comparisons per sex category).

Example Output

Results for the first regression, equation (3), are below.

Listwise Deletion of Missing Data

	Mean	Std Dev	Label
BEHAVE	5.250	2.848	
HIGH	.333	.482	
LOW	.333	.482	
MALE	.500	.511	
MA_LOW	.167	.381	
MAL_HI	.167	.381	

N of Cases = 24

Multiple R	.95052
R Square	.90349
Adjusted R Square	.87668
Standard Error	1.00000

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	5	168.50000	33.70000
Residual	18	18.00000	1.00000

F = 33.70000 Signif F = .0000

----- Variables in the Equation -----

Variable	B	SE B	95% Confdnce	Intrvl B	Beta
HIGH	1.750000	.707107	.264424	3.235576	.295936
LOW	-.750000	.707107	-2.235576	.735576	-.126830
MALE	2.250000	.707107	.764424	3.735576	.403570
MA_LOW	1.250000	1.000000	-.850922	3.350922	.167113
MAL_HI	3.500000	1.000000	1.399078	5.600922	.467916
(Constant)	3.000000	.500000	1.949539	4.050461	

----- in -----

Variable	T	Sig T
HIGH	2.475	.0235
LOW	-1.061	.3029
MALE	3.182	.0052
MA_LOW	1.250	.2273
MAL_HI	3.500	.0026
(Constant)	6.000	.0000

Table 9

Descriptive Statistics and Correlations Among Behavioral Problems, SES, Sex, and the SES×Sex Interaction

Variable	Behave	Correlations				
		High	Low	Male	Male * Low	Male * High
Behave	---					
High	.602*	---				
Low	-.317	-.500*	---			
Male	.688*	.000	.000	---		
Male * Low	.080	-.316	.632*	.447*	---	
Male * High	.842*	.632*	-.316	.447*	-.200	---
Mean	5.25	.33	.33	.50	.17	.17
SD	2.85	.48	.48	.51	.38	.38

Note. Male is a dummy variable (1 = male, 0 = female), as are High (1 = high SES, 0 = otherwise) and Low (1 = low SES, 0 = otherwise); n = 24.

*p<.05.

Table 10

Regression of Behavioral Problems on Sex and SES

Variable	b	se	ΔR^2	95%CI	F	t
Male	2.25	0.71	.05	0.76, 3.74	10.13*	3.18*
SES			.07		6.58*	
Low	-0.75	0.71		-2.24, 0.74		-1.06
High	1.75	0.71		0.26, 3.24		2.48*
Sex × SES			.07		6.29*	
Male × High	3.50	1.00		1.40, 5.60		3.50*
Male × Low	1.25	1.00		-0.85, 3.35		1.25
Intercept						

Note. $R^2 = .90$, adj. $R^2 = .88$, $F_{5,18} = 33.70^*$, MSE = 1.00, n = 24. ΔR^2 represents the semi-partial multiple correlation or the increment in R^2 due to adding the respective variable; Male is a dummy variable (1 = male, 0 = female), as are High (1 = high SES, 0 = otherwise) and Low (1 = low SES, 0 = otherwise).

*p<.05.

Table 5

Simple Main Effect Comparisons of Behavioral Problems by Sex Across Levels of SES

Contrast	Male vs. Female Mean Difference	Standard Error of Difference	Bonferroni Adjusted .95CI
SES Level			
Low	3.50*	0.71	1.64, 5.36
Middle	2.25*	0.71	0.39, 4.11
High	5.75*	0.71	3.89, 7.61

*p<.05, where p-values are adjusted using the Bonferroni method.

[Note:

1. Instead of sex pairwise comparisons for each level of SES in Table 5, one could also calculate and display SES comparisons separately for males and for females

2. The Bonferroni critical t used was 2.631 (df = 18, comparisons = 3).]

An interactive graphical of these results can be seen here:

<http://tinyurl.com/2vk7l4s>

Also here:

<https://spreadsheets.google.com/ccc?key=0ArHM99WFArmdGpYRzIRQ3FHQ09NWVVqcEdWOWhp cGc&hl=en&authkey=CMeshLcC>

Means Per Cell Combination

	Male	Female
High	10.5	4.75
Middle	5.25	3
Low	5.75	2.25

Figure Showing Interaction: Sex Differences Across Levels of SES

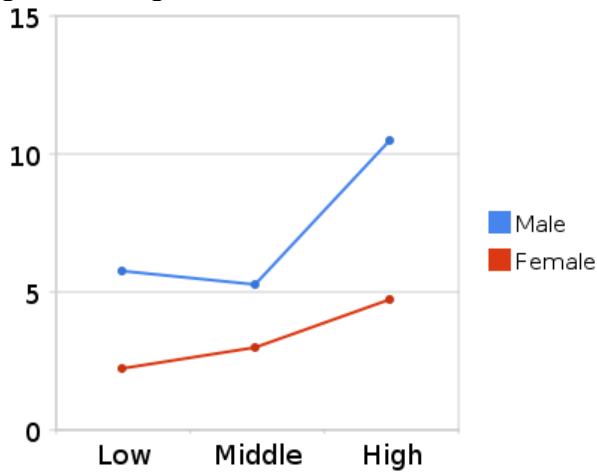


Figure Showing Same Data and Interaction: SES Differences Across Levels of Sex

