TEACHING STATISTICAL INTERACTION IN ANOVA AND OLS REGRESSION IN THE CONTEXT OF IDEOLOGY IN CRIMINAL JUSTICE PROCESSING

Huei-Hsia Wu, Boise State University, and
Anthony Walsh, Boise State University

ABSTRACT

In the social world we often come across circumstances where the effect of one variable on another variable varies depends on different levels of a third variable. In statistical terms this is known as Interaction. The concept of interaction can often be confusing to students, who tend to think in terms of straight linear relationships. The present study offers a simple and effective way to teach this complex concept in the context of sociopolitical ideology. We examine the role of ideology in determining the sentencing recommendations of probation officers using ANOVA and OLS regression and contend that interaction terms can be presented to students in an easily understandable manner. Our approach helps integrate substantive theory with statistics to the benefit of both.

INTRODUCTION

A recent paper addressing the quantitative literacy gap among social science students calls for integrating data analysis (IDA) into the general curriculum (Howery & Rodriguez 2006). The basic idea of IDA is that teaching quantitative and qualitative material simultaneously (as is typically the case in the physical and natural sciences) increases the understanding of both. Howery and Rodriguez are positive about IDA and conclude that:

> The IDA project has shown that introducing data analysis early, frequently, and sequentially throughout the curriculum can have a major impact on students' professional development and in developing hands-on experiences that will enhance their technical and methodological skills and allows for a better understanding of the substantive and theoretical applications of the discipline. (2006 37)

It is well known that quantitative skills among sociology majors are not on par with students of most other majors (Bushway & Flowers 2002; Proctor 2002, 2005). This may be a function of the discipline being attractive to students with poor quantitative skills, or a function of the paucity of classes in which quantitative materials are used. Proctor's (2006) recent paper is concerned with integrating computer software into statistics classes to enhance student interest and understanding. Howery and Rodriguez (2006) are primarily concerned with using quantitative methods to enhance students' grasp of substantive and theoretical concepts. The present paper combines the concerns of both, but focuses on using substantive and theoretical concepts to enhance understanding of the basic concept of statistical interaction.

STATISTICAL INTERACTION

Teaching the concept statistical interaction in introductory statistics courses is perhaps the most challenging component. There are other more familiar meanings to the term *interaction* that may act as roadblocks to understanding its use in statistics (Aiken & West 1991). For instance, in sociology we learn that interaction refers to mutual or reciprocal actions of individuals in groups and in chemistry we learn that it means the transfer of energy between elementary particles. These usages connote situations in which two or more objects have reciprocal effects on one another and their relationship is easily grasped intuitively.

Likewise, students have little problem understanding so-called "main effects" because such effects are linear; i.e., the independent variable X has averaged constant effects on dependent variable Y. Positive linear relationships such as "the more one studies the better the grades one gets," or negative linear relationships such as "the lower the social class the greater the number of medical problems," put little strain on students' intuitive skills. If we add a third variable to a model such as "the more one studies and the higher one's IQ the better grades one will get," we are still in the realm of linear relationships in the sense that both *time*

spent studying (X_1) and *IQ* (X_2) are assumed to be interpretable alone and to have simple additive effects.

If the magnitude of the effect of one independent variable (X_1) on the dependent variable (Y) varies as a function of the values of a second independent variable (X_2), we say that the effect of X_1 on Y is conditional upon the value of X_2. This is known as an interaction effect. Whenever we examine the effects of two or more independent variables on a dependent variable we must be alert to the possibility of interaction effects. However, we tend to ignore interaction effects, usually because we are unaware of them because the only frequently used technique that uncovers and reports them is the analysis of variance (ANOVA) (Curran, Bauer, & Willoughby 2004). If interaction effects exist and we do not account for them in our model estimations we have committed any number of specification errors. This paper attempts to provide a simplified account of how this important concept can be presented to statistics students in a relatively painless way for students and professors alike.

TWO EXAMPLES OF STATISTICAL INTERACTION FROM SENTENCING DATA

The two examples we use to illustrate interaction pertain to the relationship between political ideology and the dispensing of justice in the form of sentencing recommendations that probation officers make to the courts for offenders convicted of felony sex crimes. Students understand technical information more readily in the context of examples that are both inherently interesting and pertinent to their field of study. The concepts of justice and ideology are dear to most social scientists' hearts, so our examples should strike a responsive chord with all social science students.

Decisions about the sentencing of criminal defendants are made within the boundaries of the law, but within those boundaries there is ample opportunity for extralegal considerations (Roberts & Stalens 2000). Although judges make the final decision regarding the fate of convicted felons, according to Kaplan and Skolnick:

> The modern day sentencing judge is like the 17th century monarch, who possessed absolute power in theory but in practice was frequently manipulated by the ministers who stood quietly behind the throne and controlled the flow of information to him. (1982 567)

Probation officers (POs) are the modern judges' "ministers" controlling the flow of information to them via presentence reports (PSI) and sentencing recommendations. Officers are somewhat constrained in their sentencing recommendations in some jurisdictions by sentencing guidelines that instruct them to assign scores to relevant elements of the offense before the bench and the offender's criminal record. These scores assign offenders to a particular presumptive sentencing category.

The operationalization of legally relevant variables is intended to render justice objectively in that similarly situated offenders are treated similarly (Lauen 1997). However, the guidelines used by the courts from which our data come contain sections in which an officer can impose his or her subjective interpretation. This is most strongly so in the offense section, which contains subsections for aggravating circumstances, for which points are added to an offender's score, and mitigating circumstances, for which points are deducted. For example, one of the three aggravating circumstances is: "Offender was engaging in continuing criminal activity as a source of income or livelihood." An example of one of the four mitigating circumstance is: "There was substantial provocation, justification, or excuse for the offense." Even in the putatively more objective areas of the guideline there are grey areas where some officers may add points to an offender's score while others will not.

It is reasonable to assume that officers' causal attributions will bias (consciously or not) their scoring of the guidelines. According to Sims (2003), people who attribute criminal behavior to the free choice of individuals lacking in moral character and short on self-control demand swift and severe punishment of miscreants, while those who view it as a result of forces external to the individual such as poverty and discrimination will advocate for rehabilitation. These two positions reflect the conservative and liberal view of crime and punishment, respectively (Walsh & Ellis 2004). Conservative officers are thus more likely to recommend harsher sentences than liberal officers and to justify them with higher guideline scores for offenders pro-

Table 1: Mean Recommendations of Liberal and Conservative Probation Officers for Categories of Victim-Offenders Relationship: Main and Interaction Effects

Grand Mean	555.5	n=413

Victim Offender Relationship			Probation Officer Ideology		
Family	286.1	n=104	Liberal	399.5	n=242
Acquaintance	409.7	n=196	Conservative	711.4	n=171
Stranger	970.7	n=113			

Probation Officer Ideology / Victim Offender Relationship Interaction

	Liberal		Conservative	
Family	243.6	n= 61	328.5	n=43
Acquaintance	290.9	n=124	528.6	n=72
Stranger	664.2	n= 57	1277.2	n=56

	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta
Vic/Off Relationship	30557384.5	2	15278692.2	38.0	.000	.396
P.O. Ideology	9022838.7	1	9022838.7	22.5	.000	.228
Two-Way Interaction	4108756.1	2	2054378.1	5.1	.006	.158

cessed by them.

METHODS

Because this article is concerned with pedagogy and not a research piece, we briefly describe the sample and operationalization of the variables. The dependent variable is sentencing recommendations (Porec) made by POs. Porec was operationalized by asking all 31 probation officers in this study to assess the relative severity of a number of sentencing options given a base value of 10 severity points for each year of probation. The results, averaged and rounded, were as follows:

Each year of probation	10.0 points
Each $25 of a fine	1.0 point
Each day in county jail	1.0 point
Each two days in work release	1.0 point
Each day in state prison	1.1 points

For instance, a sentence of three years probation, 60 days in jail, and a $250 fine would translate in a Porec score of 30 + 60 + 10 = 100, or two years imprisonment and a $1,000 fine would equal 365 x 2 x 1.1 + 40 = 843 points.

The first independent variable is victim/offender relationship (V/Orel) divided into three categories (1=family, 2=acquaintance, and 3=stranger) according to information contained in case records. Probation officer ideology (Pold) is divided into two categories (1=liberal; 2=conservative) determined by officers' scores on our criminal justice ideology scale, which is presented in the appendix. We obtained an alpha reliability coefficient of 0.71 for this instrument.

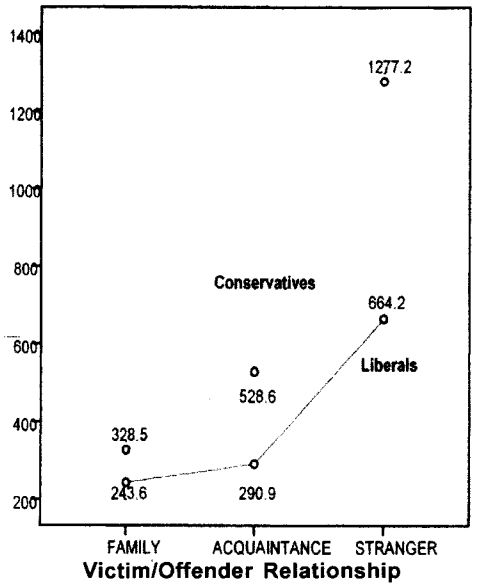
The ideology scale has a possible range from zero to 40, with the actual range based on the scores of 31 probation officers being 14 through 28. Officers scoring 14 through 20 were considered conservatives (n=14), and those scoring 21 through 28 were considered liberals (n=17). Eleven of the 14 conservative officers (78.6%) considered their role to be primarily law enforcement while 11 of the 17 liberal officers (64.7%) considered their role to be primarily social work. These differing role descriptors further underline the punitive versus rehabilitative differences between liberal and conservative officers.

An additional independent variable used in our OLS example is crime seriousness (Crser). This variable and its operationalization was discussed earlier, and had scores ranging from zero (in rare cases in which aggravating factors were canceled by mitigating factors) to 10, with a mean of 3.4.

ANOVA AND INTERACTION

Our first task is to illustrate interaction in a 2-way ANOVA. In our statistics classes we first present an overhead of Table 1 showing the results of a 2-way ANOVA printout condensed to provide the most relevant information. We have the grand mean PO sentencing recommendation and those recommendations broken down by 1) victim/offender relationship, 2) PO ideology, and 3) broken down by PO ideology and victim/of-

Figure 1: Sentence Recommendation Means by Victim/Offender Relationship and Probation Officers Ideology



fender relationship simultaneously. We note that there are large jumps in recommendation severity as we move from the family to the acquaintance to the stranger condition. This is a good point to ask students why this might be. Point out that while it is especially true for conservative officers it is true for liberal officers also. The primary reason is that mean guideline crime seriousness scores significantly increase from family to acquaintance to stranger for both liberals ($F=29.7$, $p < .0001$) and conservatives ($F=31.3$, $p< .0001$). Why this is so should also lead to a lively discussion (it has mostly to do with strangers inflicting greater physical harm on their victims).

We next explain that the concept of interaction can be seen as analogous with parallelism explaining that if no interaction exists, lines drawn between the means of the dependent variable (Porec) for each category of the first independent variable (V/Orel) will be parallel for each of the categories of the second independent variable (Pold). Figure 1 shows that this is plainly not the case. Although we see that there is a linear relationship between sentencing recommendations and categories of victim-offender relationship in the same direction for both sets of officers, the lines connecting the means are far from parallel. We have found that presenting a figure such as this is an excellent aid for student understanding. Graphic displays such as this are excellent pedagogical tools to help students to grasp somewhat troubling statistical concepts such as interaction.

INTERACTION AND OLS REGRESSION

Our second example involves the inclusion of an interaction term in an OLS regression to show how including interactions helps to get a better understanding of the data. As previously noted, many interactions in a data set go unnoticed and unanalyzed in our haste to enter the relevant variables into multivariate models (Curran, Bauer, & Willoughby 2004). If preliminary analysis or theoretical knowledge of the subject matter under investigation leads us to suspect significant interaction we must deal with it in our multivariate models or risk misspecification. We now illustrate the usefulness in terms of garnering additional understanding of the data gained by including interaction terms in multivariate OLS regression. Again, we wish to illustrate interaction using the impact of ideology on probation officers' sentencing recommendations for convicted felons. In this case we use crime seriousness as our interaction term.

We have seen that theory informs us that conservative officers will be more punitive than liberal officers and that they may attempt to justify their greater punitive recommendations by assigning higher guideline scores. As expected, the mean crime seriousness score assigned by conservative officers is significantly greater than the mean score assigned by liberal officers ($t=3.5$, $p<.001$). This can just as easily be interpreted as liberals underscoring the guideline as conservatives overscoring them (probably a bit of both), but the fact remains that the mean scores are significantly different. This difference cannot be interpreted as officers retrospectively identified as conservatives getting the most serious cases, or vise versa, because cases are assigned to officers on a strictly rotational basis and is thus random assignment. Mean prior record scores did not differ between conservatives and liberals, primarily because there is far less room in this portion of the guideline to exercise subjective judgments.

Table 2: Illustrating Interaction Effects in OLS Regression

(A) Standard Model					
Variable	**b**	**s.e.**	**beta**	**t**	**Sig.**
Crime Seriousness	239.51	37.83	.687	23.95	.0000
PO Ideology	161.97	40.20	.116	4.03	.0001
(Y-intercept)	320.80	37.83		-8.48	.0000
Adjusted r-squared = .505					
(B) Liberal Officers Only					
Variable	**b**	**s.e.**	**beta**	**t**	**Sig.**
Crime Seriousness	178.50	13.16	.582	13.52	.0000
(Y-intercept)	-151.20	42.98		-3.52	.0000
Adjusted r-squared = .338					
(C) Conservative Officers Only					
Variable	**b**	**s.e.**	**beta**	**t**	**Sig.**
Crime Seriousness	288.57	14.76	.776	19.55	.0000
(Y-intercept)	-322.16	59.45		-5.42	.0000
Adjusted r-squared = .600					
(D) Liberal and Conservative Officers and Crime Seriousness X Ideology Interaction					
Variable	**b**	**s.e.**	**beta**	**t**	**Sig.**
Interaction Term	110.06	19.34	.350	5.60	.0000
Crime Seriousness	178.50	14.62	.512	12.21	.0000
PO Ideology	-170.96	71.84	-.122	-2.40	.0170
(Y-intercept)	-151.19	47.74		-3.18	.0020
Adjusted r-squared = .531					

The issue we wish to examine in the context of OLS regression is the effect of crime seriousness and PO ideology on sentencing recommendations. If the two variables have only additive effects the regression equation predicting those effects is:

$$Y' = b_0 + b_1 X_1 + b_2 X_2 + e$$

If significant interaction exists, this additive model is inadequate to describe the relationship. The OLS equation required to analyze a 2-way interaction is:

$$Y' = b_0 + b_1 X_1 + b_2 X_2 + b_3 (X_1 X_2) + e$$

where $b_3 (X_1 X_2)$ is the interaction between crime seriousness and probation officer ideology. It is a product or constituent term obtained by multiplying the two interacting variables together.

Table 2 presents four separate regression models. Model A is a standard model examining the effects of Pold and crime seriousness on Porec. As expected, with crime seriousness in the model conservative officers still recommend significantly more serious sentences (b=161.97, ß =.116, p<001). If this model were our only concern we would feel satisfied, but we need to go further.

Models B and C are simple bivariate models for the effects of crime seriousness for liberal and conservative officers, respectively. The adjusted r squared values for liberals (0.338) and conservatives (0.600) support the contention that the seriousness of the crime has a greater impact on the recommendations made by conservatives, and that liberals take factors other than crime seriousness more into consideration than conservatives when making their sentencing recommendations. For conservatives, the impact of each additional crime seriousness point is 288.57 sentencing recommendation points whereas for liberals it is only associated with an additional 178.50 points. The substantive difference between the slopes is 110.06 recommendation points. Crime seriousness accounts for 26.2 percent more of the variance in recommendation for conservatives than for liberals, which emphasizes the more single-minded approach to justice held by conservatives (the punishment must fit the crime, with little else being of major importance).

Model D is a multiple regression model that includes crime seriousness, PO ideology, and the interaction term. The important point to note in the context of teaching statistical interaction is that the value of the slope for the interaction term is 110.06. This is the value we got from subtracting the liberal slope in model B from the conservative slope in model C. The test of statistical significance

for the difference between slopes B and C is simply the t value (5.60) for the interaction effect.

Note that probation officer ideology remains a significant predictor of the dependent variable. However, also note that the coefficients are negative, indicating that liberal officers recommend *harsher* sentences after the effects of the interaction term and crime seriousness have been accounted for. Although the standardized beta is very weak it is a statistically significant finding which would not have been uncovered without including the interaction term. Given that the direction of the relationship is reversed in model C from what it was in model A, the researcher is now confronted with the task of delving deeper into his or her data to discover why, although this is beyond the purpose of this paper. This opportunity to more fully understand the nature of the data would not have been available without due attention being paid to the possibility of interaction.[1]

CONCLUSION

We hope that this simplified example of how the concept of interaction can be meaningfully explained to beginning statistics students proves as useful to others as it has over the years to us. As previously mentioned, students always find technical concepts easier to grasp if such concepts are presented in tandem with examples taken from their fields of interest (Gregory & O'Toole 1987). Both examples presented here are relatively straightforward with mean recommendation values presented in both tabular and graphic form and should be easily understood.

We believe that it is particularly important to present interaction in the context of OLS regression, primarily because interaction effects tend to go unanalyzed when multivariate regression models are used (Curran, Bauer, & Willoughby 2004). Our demonstration showed how the interaction term is simply the difference between the unstandardized betas for separate models for a dichotomized independent variable. It also demonstrated analyzing interactions can uncover unexpected and challenging findings in the data. Further discussion of the unexpected finding presented in this paper is beyond the purpose of this pedagogical piece, however.

The major limitation in terms of pedagogy is the lack of a method of assessing the efficacy of our learning strategy. That is, we have no comparison between students exposed to it and those not exposed to it in terms of their understanding of the concept of interaction. Such a strategy would require random assignment to different statistics classes in which one class was exposed to this material and one class not, which is plainly impossible. However, because the strategy of integrating data analysis with the teaching of substantive material has a long history of positive outcomes in the hard sciences and because the IDA project has proven successful in sociology (Howery & Rodriguez 2006), we see no reason why our method should not enhance students' understanding of a difficult statistical concept and an important substantive concept in criminal justice.

ENDNOTE

[1]This anomalous finding does not appear to be a function of collinearity. The collinearity diagnostics reveal a tolerance of .298 for PO ideology and .425 for crime seriousness in the final model (C) and variance inflation factors of 3.35 and 2.35, for PO ideology and crime seriousness, respectively.

REFERENCES

Aiken L. & S. West. 1991. *Multiple Regression: Testing and Interpreting Interactions*. Thousand Oaks, CA: Sage.

Bushway S. & S. Flower. 2002. Helping criminal justice students learn statistics: a quasi experimental evaluation of learning assistance. *J Criminal Just Edu* 13 35-56.

Curran P., D. Bauer & M. Willoughby. 2004. Testing main effects and interactions in hierarchical linear growth models. *Psychological Methods* 9 220-237.

Gregory S. & R. O'Toole. 1987. Teaching sociological research methods to medical students. *Teaching Sociology* 15 128-135.

Howery C. & H. Rodriguez. 2006. Integrating data analysis (IDA): working with sociology departments to address the quantitative literacy gap. *Teaching Sociology* 34 23-38.

Kaplan J. & J. Skolnick. 1982. *Criminal Justice*. NY: Foundation Press.

Lauen R. 1997. *Positive Approaches to Corrections: Research, Policy, and Practice*. Lanham, MD: American Correctional Association.

Proctor J. 2002. SPSS vs. Excel: computing software, criminal justice students, and statistics. *J Criminal Just Edu* 13 433-442.

_____. 2005. Academic achievement and statistical knowledge: a comparison of criminal

justice and noncriminal justice majors. *J Criminal Just Edu* 17 143-161.
Roberts J. & L. Stalans. 2003. *Public Opinion, Crime, and Criminal Justice*. Boulder, CO: Westview.
Sims B. 2003. The impact of causal attribution on correctional ideology: a national study. *Criminal Just Rev* 28 1-25.
Walsh A. & L. Ellis. 2004. Ideology: criminology's Achilles' heel? *Qrtly J Ideology* 27 1-25.

Appendix
The Liberal/Conservative Scale*

1. Those who commit the common forms of street crimes are essentially forced into such behavior by destructive social conditions.

2. The root cause of crime is the erosion of the moral values that have traditionally deterred criminality.

3. Most people whom we label criminals are victims of racism, discrimination, class bias, and poverty.

4. The bulk of serious crime is committed by individuals with limited self-control who are oriented to the present, and who have underdeveloped moral consciences.

5. Crime is best attacked by dealing with offenders in a firm, decisive, and forceful way.

6. Crime is best attacked by concentrating not on the individual offender but on efforts to meliorate the source of crime: malignant environments.

7. Prison regimens should involve hard work and strict discipline in order to instill in offenders good work habits and moral fiber.

8. Only the most dangerous offenders should be incarcerated. Maximum use should be made of community-based programs.

9. Punishment should fit the offender and be appropriate to rehabilitation, with the nature of the crime being a secondary consideration.

10. Individuals are responsible for their own behavior. Since they have the capacity to choose between right and wrong, it makes no sense to explain criminal behavior in terms of "bad environments."

*Items are scored in Likert fashion ranging from "strongly agree" to "strongly disagree."

Alpha reliability = .71