Testing Directional Research Hypotheses

Keith McNeil
New Mexico State University

Theory, literature review, and past research results will guide the development and testing of most research questions. This paper argues that most research questions will be directional, instead of nondirectional, particularly since most researchers want to make a directional conclusion. Although many researchers incorrectly make directional conclusions after finding significance with a nondirectional test, tests of directional hypotheses are the only ones that allow directional conclusions.

Most computer packages only report the nondirectional probability. Therefore, an adjustment is necessary when a directional research hypothesis has been tested. Exhibits are provided for testing both directional and nondirectional hypotheses regarding a) the difference between two means, b) single population correlation, c) traditional covariance, d) interaction between two dichotomous predictors, e) interaction between one continuous variable and one dichotomous variable, f) contribution of a variable, and g) selected non-linear hypotheses.

Researchers have a choice of various statistical tools; readers of this journal realize that most research hypotheses can be tested with the GLM. Each statistical tool can be used to test both nondirectional research hypotheses and directional research hypotheses. The researcher has to decide whether the research hypothesis is directional or nondirectional. The choice should not be difficult, as the decision is affected by theory, literature review, and past research. If these areas do not provide a clue, then the researcher should consider the desired conclusion. If the researcher is content with stating, "There is a difference between Treatment and Comparison," then the nondirectional research hypothesis is appropriate. But if all the forces point to desiring to make the directional conclusion, "Treatment is better than Comparison," then a directional research hypothesis is appropriate. The choice of a directional or nondirectional research hypothesis is not a statistical one. The choice is driven by the research base and tied to one's desired conclusion.

A sample of three recent statistics texts illustrates the confusion related to this issue. Grimm (1993) waffles on the use of directional research hypotheses.

> Research hypotheses (scientific hypotheses) are usually stated as predictions about the expected direction of an experimental effect. For Exhibit, persuasion technique A will induce <u>greater</u> attitude change than technique B; subjects' perceptions of control over a stressor will <u>decrease</u> stress reactions; or higher levels of physiological arousal will create <u>stronger</u> emotions. Researchers typically frame their statistical hypotheses in a nondirectional form. In other words, even though the research hypothesis makes a prediction about which of two means will be larger, the null and alternative hypotheses allow the investigator to discover if a treatment effect is opposite to the predicted effect. (p. 184)

His major concerns are that choice of the direction should be made <u>before</u> data are collected, a valid concern. But the other concern is that results in the opposite direction are ignored with a directional test. If one is theory building, then one may want to investigate those anomalous results to see if, in fact, they are replicable. Grimm (1993) does not treat directional hypotheses with statistical tests other than the difference between two means, although directional interpretations are often made with nondirectional tests of significance.

Sprinthall (1990) introduces directionality when discussing differences between two means, but treats the concept as a mechanical issue, "Remember, in terms of technique, the only difference between a one-tail and a two-tail t is how we look up the significance level" (p.185). He also doesn't discuss directional tests of significance for other tests of significance, but makes directional conclusions from several nondirectional research hypotheses. Several of his examples are stated as directional, but tested as nondirectional. Sprinthall (1990) points out that "the alternative hypothesis for F can never be directional. That is, if t is computed by taking the square root of F, then its significance must be evaluated against the critical values in the two-tailed t table "(p.275). That this is not so will be demonstrated later.

Shavelson (1988) is more in line with the essence of this paper. He introduces directional research hypotheses with the very first statistical test, even discussing the directional hypothesis before the nondirectional. In discussing most subsequent tests, he uses the same approach. He continuously emphasizes that "if both theory and empirical evidence suggest the outcome of a study, a directional research hypothesis should be used" (p. 251). He discusses directional hypothesis testing for a single mean, difference between two means, correlation, planned comparisons, and difference between two correlations. He does not discuss directional hypotheses in terms of ANOVA, ANCOVA, or multiple regression. Because he doesn't discuss the use of one degree of freedom $\underline{F}$ tests, he doesn't attend to the issue of computer-generated probabilities discussed in this paper.

Rationale for Directional Research Hypotheses

In the case of a new treatment, a researcher should show that it is more effective, costs less, is quicker to administrate, has longer lasting impact, etc. Who would care if the new treatment is worse than the existing comparison treatment? Any idiot can design a new treatment that is worse, costs more, is slower to administer, has a shorter lasting impact, etc. What would the research community learn from such findings? Over many years of experience with this issue, it has become apparent that nondirectional research hypotheses are only useful in dredging data in search of hypotheses for another researcher with some other data to verify. If a researcher has a good grasp of the content area, a directional research hypothesis will be desired.

Model Structure

An area of confusion is that both directional and nondirectional research hypotheses are tested by the same null hypothesis. For instance, if the research hypothesis is directional, "Treatment is more effective than Comparison," the statistical hypothesis is "Treatment is as effective as Comparison." If the research hypothesis is nondirectional, "Treatment and Comparison are not equally effective," the statistical hypothesis is "Treatment is as effective as Comparison." Most statistics texts illustrate this fact, but give primary coverage to the nondirectional hypothesis. Unfortunately, statistics texts do not emphasize the permissible conclusions of the two. Indeed, some statistics texts confuse the issue by making directional conclusions from nondirectional research hypotheses. Journal reviewers and editors reinforce the confusion by allowing only the statistical hypothesis to be reported. Why not force the author to state what is desired?

From a GLM perspective, the Full Model and Restricted Model are identical. The difference is the desired algebraic status of the weighting coefficient which will be identified as "want" in the following exhibits. Statistical packages (e.g., SAS, SPSS, BMDP) report only one probability value--that for the nondirectional research hypothesis. Consequently, many users mistakenly report that nondirectional probability when they have tested a directional research hypothesis.

Adjustment of Computed Probability

Statistics texts make the case that the required critical value depends upon whether one has a directional or nondirectional research hypothesis. We have all seen pictures of alpha in one tail of the t-distribution for a directional hypothesis, and alpha split between the two tails for a nondirectional research hypothesis. We also all remember that the relationship between $\underline{t}$ and $\underline{F}$ is $\underline{t}^2 = \underline{F}$. Thus the tails of the negative and positive sides of the $\underline{t}$ distribution both constitute the right-hand tail of the $\underline{F}$ distribution, as in Figure 1. What this means is that we would get a large $\underline{F}$ value half the time when sample $\text{mean}_T >$ sample mean_C and half the time when sample $\text{mean}_T <$ sample mean_C. If our research hypothesis was directional, then we would be interested only in one of the two halves of the $\underline{F}$ distribution in Figure 1. If the calculated $\underline{F}$ was 4.24, then the reported (nondirectional) probability would be .05. But if we had a directional research hypothesis, (say population $\text{mean}_T >$ population mean_C) and the results were in line with our research hypothesis (sample $\text{mean}_T = 15$, sample $\text{mean}_C = 10$) instead of being exactly opposite, say (sample $\text{mean}_C = 15$, sample $\text{mean}_T = 10$), then we would obtain a $\underline{t}$ value of 2.06 and we would need to divide the reported probability by 2, as discussed in Figure 2.

On the other hand, if our results did turn out opposite to expectations, we would not want to say we had "significant results." Suppose our results produced a $\underline{t}$ value of -2.06 at @ in Figure 1. Although that $\underline{t}$ value translates to an $\underline{F}$ value of 4.24, one cannot rely on the $\underline{F}$ value (and the probability associated with it). One must check the data to see if the results are in the direction hypothesized. If the results are in the hypothesized direction (the shaded area in the bottom of the $\underline{F}$ distribution), then the computed probability must be divided by 2. If the results are not in the desired direction, then the computed probability must be divided by 2 and subtracted from 1.00. These procedures are outlined in Figure 2 and apply to each of the following exhibits.

Figure 1 Relationship Between *t* and *F* with Respect to Directional and Non-Directional Hypotheses

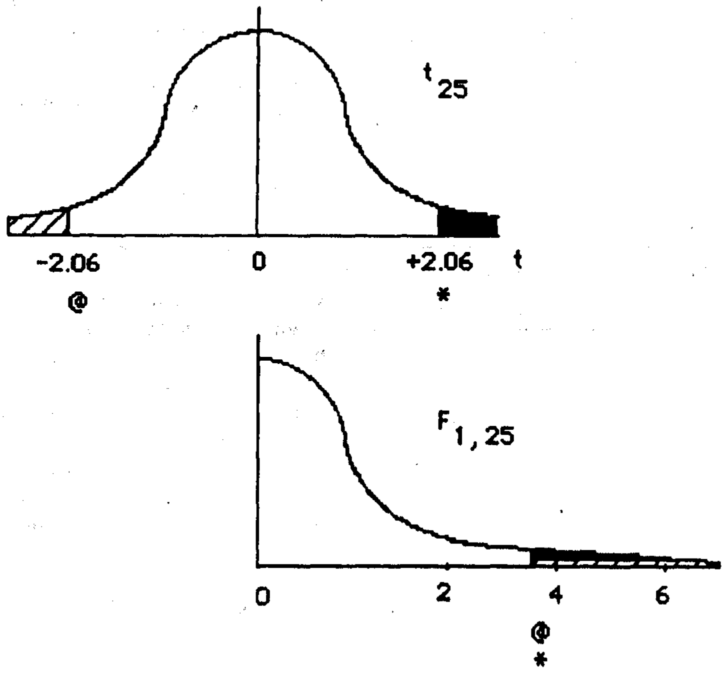


Figure 2 Procedures for Changing Computer-Generated Nondirectional Probability of F-tests to Directional Probabilities

Check to see whether **Condition I** or **Condition II** holds.

Condition I: If results (means, correlations, difference between means, etc.) are in the hypothesized direction: Divide nondirectional computer probability by 2.

Example: Nondirectional probability on printout is .08. Therefore the directional probability is (.08 / 2) .04, which is the probability that should be reported, and is indicated by the * in Figure 1.

Condition II: If results (means, correlations, differences between means, etc.) are opposite to the hypothesized direction, divide nondirectional computer probability by 2 and subtract the resulting value from 1.00.

Example: Nondirectional probability on printout is .08. Therefore the directional probability is 1 - (.08 / 2), or .96, which is the probability that should be reported, and is indicated by the @ in Figure 1.

Note. The directional research hypothesis could only have been tested when the numerator degrees of freedom are equal to 1.

Examples

Difference Between Two Means

Exhibit 1 contains both the directional research hypothesis and the nondirectional research hypothesis for testing the difference between two means. Notice that both research hypotheses use the same statistical hypothesis. The two Full Models are exactly the same, and the two Restricted Models are exactly the same. The difference is in the "want." The different wants require that different actions be taken on the computed probability, as discussed in the previous section. The different wants also impact the permissible conclusions.

Exhibit 1 Difference Between Two Population Means

Directional Research Hypothesis: For the population of interest, Group A has a higher mean than Group B on the criterion Y.

Nondirectional Research Hypothesis: For the population of interest, Group A and Group B are not equally effective on the criterion Y.

Statistical Hypothesis: For the population of interest, Group A and Group B are equally effective on the criterion Y.

Full Model: $Y = a0U + aGA + E3$

Want (for directional RH) $a > 0$; restriction: $a = 0$.
Want (for nondirectional RH) $a \neq 0$; restriction: $a = 0$.

Restricted Model: $Y = a0U + E4$

Where: Y = criterion; U = 1 for all subjects; GA = 1 if subject in Group A, 0 if subject in Group B; and a0 and a are least squares weighting coefficients calculated so as to minimize the sum of the squared values in the error vectors.

```
PROC REG; MODEL Y = GA;
TEST GA = 0;
```

Correlation

The above discussion is also appropriate to testing correlations. If a new testing instrument is developed, one would hope that it is reliable and valid. These conclusions require positive correlations, not correlations different from 0. If a theory posits that X and Y are related, the theory should specify if that relationship is positive or negative. If one is going to consider studying for a test, one needs to know if the relationship between studying and exam grade is positive or negative! Exhibit 2 provides the complete GLM solution of a research hypothesis regarding directional correlation.

Exhibit 2 Correlation

Directional Research Hypothesis: For some population, X is positively related with Y.

Nondirectional Research Hypothesis: For some population, X is related with Y.

Statistical Hypothesis: For some population, X is not related with Y.

Full Model: $Y = a0U + bX + E1$

Want (for directional RH) $b > 0$; restriction: $b = 0$.
Want (for nondirectional RH) $b \neq 0$; restriction: $b = 0$.

Restricted Model: $Y = a0U + E2$

Where: Y = criterion; U = 1 for all subjects; X = predictor score for subject; a0 and b are least squares weighting coefficients calculated so as to minimize the sum of the squared values in the error vectors.

```
PROC REG; MODEL Y = X;
TEST X = 0;
```

Analysis of Covariance

Assume that you have a Treatment and Comparison situation as previously described, and you want to adjust the posttest scores for initial differences in pretest scores. You would want the Treatment group to be higher than the Comparison group on the adjusted posttest scores. Again, who would be interested in a treatment that produced lower adjusted posttest scores? Exhibit 3 provides the GLM solution for both the nondirectional and directional analysis of covariance research hypothesis. The directional research hypothesis in ANCOVA is applicable only when there are two groups being compared, resulting in one degree of freedom in the numerator of the *F*. When there is more than one degree of freedom in the numerator, only a nondirectional research hypothesis can be tested.

Exhibit 3 Analysis of Covariance

Research Hypothesis: For a given population, Method A is better than Method B on the criterion Y, over and above the covariable C.

Nondirectional Research Hypothesis: For a given population, Method A and Method B are differentially effective on the criterion Y, over and above the covariable C.

Statistical Hypothesis: For a given population, Methods A and B are not differentially effective on the criterion Y, over and above the covariable C.

Full Model: Y = a0U + a2G2 + c1C + E1

Want (for directional RH) a2 < 0; restriction: a2 = 0
Want (for nondirectional RH) a2 not equal 0; restriction: a2 = 0

Restricted Model: Y = a0U + c1C + E2

Where: Y = criterion; U = 1 for each subject; G2 = 1 if subject received Method B, 0 if Method B; C = covariable score; and a0, a2, and c1 are least squares weighting coefficients calculated so as to minimize the sum of the squared values in the error vectors.

```
PROC REG;  MODEL Y = G2  C;
TEST  G2 = 0;
```

Interaction Between Two Dichotomous Variables

Suppose you have two treatments and two levels of motivation, and are interested in Posttest scores. Traditional analysis of variance tests for the interaction effect first, and then proceeds to the main effects if the interaction is not significant, and to simple effects if the interaction effect is significant. The interaction effect usually is treated as an assumption, or as an effect that is preferably not in existence. But the interaction effect may be the researcher's primary hypothesis, and it may be either directional or nondirectional. (In traditional analysis of variance it is always nondirectional, unless tested as an a priori contrast.)

Suppose that the treatment was designed to be particularly responsive to highly motivated students. Based on the assumption that there might be ways to increase student's motivation, you expect the directional interaction pictured in Figure 3. Your expectation is that "Students with high motivation will do better on the Posttest than students with low motivation, and the difference will be greater for the Treatment than for the Comparison." The focus of the directional interaction could just as well have been on treatments, with the expectation being "Treatment students will do better on the Posttest than Comparison students, and the difference will be greater for high motivated students than for low motivated students." The two statements are equivalent and both identify directional interaction. The complete GLM solution is provided in Exhibit 4. Notice again that the only difference between directional and nondirectional is in the "want," in the adjustment of the probability, and the permissible conclusion. Again, the directional interaction can be tested only if there is one degree of freedom in the numerator of the F.

Figure 3 Directional Interaction Between Two Dichotomous Predictors

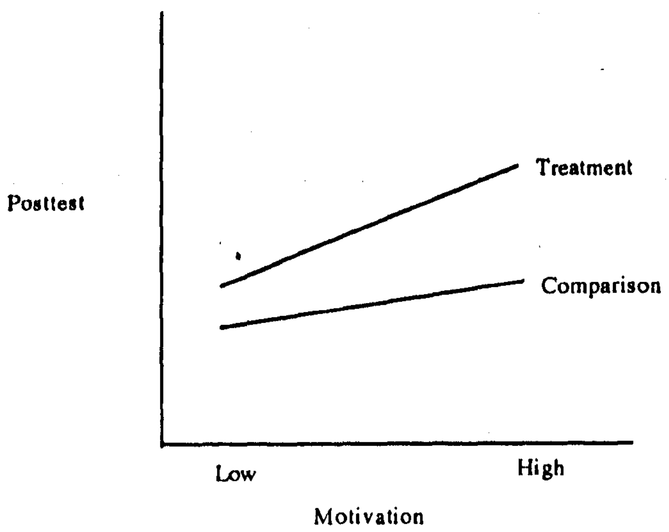


Exhibit 4 Directional Interaction Between Two Dichotomous Predictors

Directional Research Hypothesis: For a given population, the relative effectiveness of Method A (X10) as compared to Method B (X11) on the criterion of interest (X9) will be greater for Group A (X12) than for Group B (X13).

Nondirectional Research Hypothesis: For a given population, the relative effectiveness of Method A (X10) as compared to Method B (X11) on the criterion of interest (X9) will be different for Group A (X12) than for Group B (X13).

Statistical Hypothesis: For a given population, the relative effectiveness of Method A (X10) as compared to Method B (X11) on the criterion of interest (X9) will be the same for Group A (X12) as for Group B (X13).

Full Model: X9 = a0U + b(X10*X13) + c(X11*X12) + d(X11*X13) + E1

Want (for directional RH) (c) > (b - d);
restriction: (c) = (b - d)
Want (for nondirectional RH) (c) not equal (b - d);
restriction: (c) = (b - d)

Restricted Model: X9 = a0U + eX10 + fX12 + E2

```
PROC REG; MODEL X9 = X10*X13 X11*X12
X11*X13;

TEST (X11*X12) = (X10*X13 - X11*X13);
```

Interpretation: If the weighting coefficient c is numerically larger than (b - d), the directional probability is appropriate and the following conclusion can be made: For a given population, the relative effectiveness of Method A (X10) as compared to method B (X11) on the criterion of interest (X9) will be greater for Group A (X12) than for Group B (X13).

Interaction Between One Continuous Variable and One Dichotomous Variable

An extension of the previous section would be to consider motivation as a continuous variable instead of as a dichotomous variable. The same rationale applies, although now since motivation is being considered as a continuous variable two lines will be fit to the date, not four means. Figure 4 depicts the expected directional interaction. Note that Figure 4 appears very similar to Figure 3, the only difference is that motivation is considered as a continuous variable in Figure 4. The directional interaction research hypothesis would be, "As motivation increases, the relative superiority of Treatment over Comparison increases." Shavelson (1988) presents a directional example of this type, framed as the "test for difference between regression slopes from two independent samples." His presentation is in terms of a complicated *t* test. The GLM approach illustrates the similarity of all directional research hypotheses and relies on the same model comparisons as all the previous examples. Exhibit 5 contains the complete GLM solution for interaction between one continuous variable and one dichotomous variable.

Exhibit 5 Interaction Between One Continuous Variable And One Dichotomous Variable

Directional Research Hypothesis: For a given population, as X increases, the relative superiority of Method A over Method B on Y will linearly increase.

Nondirectional Research Hypothesis: For a given population, as X increases, the relative superiority of Method A over Method B on Y will linearly change.

Statistical Hypothesis: For a given population, as X increases, the difference between Method A and Method B on Y will remain the same.

Full Model: Y = aU + a1U1 + b1X1 + b2X2 + E1

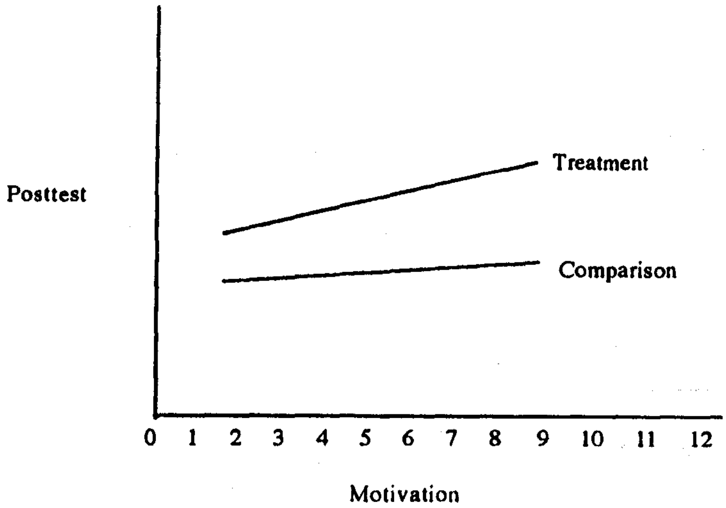
Want (for directional RH) b1 > b2;
restriction: b1= b2
Want (for nondirectional RH) b1 not equal b2;
restriction: b1 = b2

Restricted Model: Y = aU + b3X + E8

Where: Y = the criterion; U1 = 1 if the score on the criterion is from a subject in Method A, 0 otherwise; X = the continuous predictor variable; X1 = (U1*X) = the continuous predictor variable if the criterion is from a subject in Method A, 0 otherwise; U2 = 1 if the score on the criterion is from a subject in Method B, 0 otherwise; X2 = (U2*X) = the continuous predictor variable if the criterion is from a subject in Method B, 0 otherwise; and a, a1, b1, b2, and b3 are least squares weighting coefficients calculated so as to minimize the sum of the squared values in the error vectors.

```
PROC REG; MODEL Y = U1 X1 X2;
TEST X1 = X2;
```

Figure 4 Directional Interaction Between One Continuous Predictor (Motivation) and One Dichotomous Predictor (Type of Treatment)



Non-Linear Relationships

If all the predictor variables of interest are polynomial terms, the directional research hypothesis is still appropriate. Consider the case in which the linear and second-degree terms are under consideration. The second-degree curve can be either an inverted U or U-shaped. The U-shaped curve identifies a "trough" of minimum performance on the criterion, whereas the inverted U identifies a "peak" of maximum performance on the criterion. These are two very different conclusions and are a function of the sign of the second-degree term. The curves are identified in Figure 5 and the GLM solution is in Exhibit 6.

Exhibit 6 Non-linear Hypotheses

Directional Research Hypothesis: For a given population, there is a positive second degree effect of X on Y, over and above the linear effect of X.

Nondirectional Research Hypothesis: For a given population, there is a second degree effect of X on Y, over and above the linear effect of X.

Statistical Hypothesis: For a given population, there is not a positive second degree effect of X on Y, over and above the linear effect of X.

Full Model: $Y = a0U + aX + bX16 + E1$
Where: $X16 = X*X$

Want (for directional RH) $b > 0$; restriction: $b = 0$
Want (for nondirectional RH) b not equal 0; restriction: $b = 0$

Restricted Model: $Y = a0U + aX + E2$

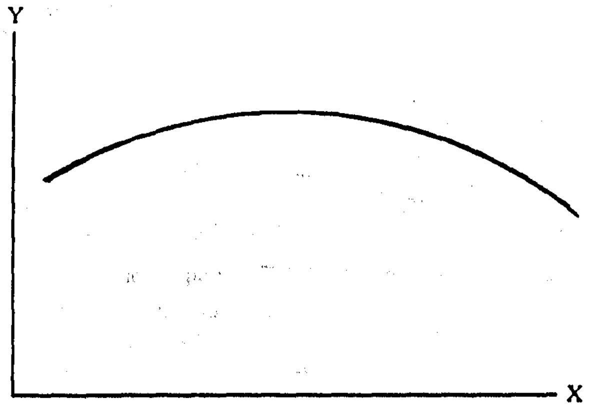
PROC REG; MODEL Y = X X16;
TEST X16 = 0;

Contribution of One Variable, Over and Above Other Variables

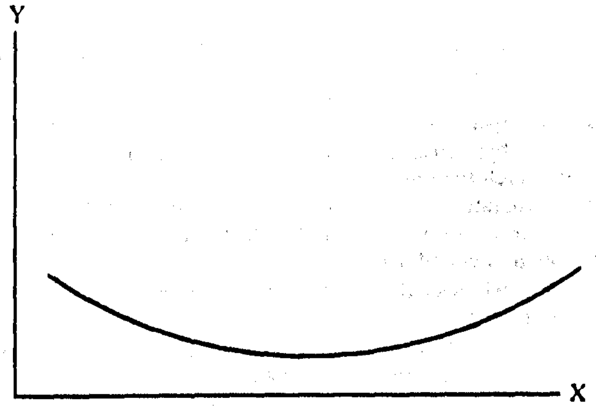
A researcher may be interested in how a variable is related to a criterion, after the effects of several other variables have been "statistically adjusted." If the variable is dichotomous (say study or not study), then this question is simply an extension of the analysis of covariance discussion into more than one covariable. The GLM solution would simply have the multiple covariables in the Full Model as well as in the Restricted Model as in Exhibit 3.

If the variable under concern is a continuous variable (say hours of studying), then whether the variable relates positively or negatively to the criterion after adjustment for the covariables would be of interest in the directional situation. Again, knowing that studying is predictive of the criterion (over and above the other variables) is not that informative; what is informative is knowing whether studying is positively related or negatively related to the criterion. If one wanted to use these results to recommend trying to increase the criterion, one would have to know the directional relationship between studying and the criterion. The GLM solution is provided in Exhibit 7.

Figure 5 U-Shaped Curves Resulting From Negative and Positive Weights of Second Degree Terms



$Y = aU + bX = cX^2 + E$, where c is negative.



$Y = aU + bX + cX^2 + E$, where c is positive.

Exhibit 7 General Over and Above

Directional Research Hypothesis: For a given population, X6 is positively predictive of the criterion Y, over and above X1, X2, X3, and X4.

Nondirectional Research Hypothesis: For a given population, X6 is predictive of the criterion Y, over and above X1, X2, X3, and X4.

Statistical Hypothesis: For a given population, X6 is not predictive of the criterion Y, over and above X1, X2, X3, and X4.

Full Model: Y = a0U + a1X1 + a2X2 + a3X3 + a4X4 + a6X6 + E1

Want (for directional RH) a6 > 0; restriction: a6 = 0
Want (for nondirectional RH) a6 not equal 0; restriction: a6 = 0

Restricted Model: Y = a0U + a1X1 + a2X2 + a3X3 + a4X4 + E2

Where: Y = the criterion; X1, X2, X3, X4, X6 = continuous or categorical information; and a0, a1, a2, a3, a4, and a6 are least squares weighting coefficients calculated so as to minimize the sum of the squared values in the error vectors.

```
PROC REG; MODEL Y = X1 X2 X3 X4 X6;
TEST X6 = 0;
```

Summary

Researchers often do not follow the knowledge base by stating a directional research hypothesis. Often, though, directional conclusions are made from testing non-directional research hypotheses. Since the statistical (or null) hypothesis is the same for directional and non-directional research hypotheses, researchers often overlook the distinction. In addition, all canned computer packages report only the non-directional probability. This paper has illustrated how the GLM can be used for directional hypothesis testing and for obtaining the correct directional probability.

All the previous exhibits are subsets of the same general situation described in Exhibit 7. The differences depend on the number of predictors, number of covariates (many, one, none), and whether the variable tested is continuous or dichotomous. In all the statistical tests discussed, a directional research hypothesis can be tested if there is a directional expectation. If there is a directional research hypothesis, there is only one want, one restriction, and one degree of freedom in the numerator of the F-test. In all cases the reported nondirectional probability must be adjusted based on how the sample results match the directional research hypothesis. These are all essential elements of a directional hypothesis.

References

Grimm, L. G. (1993). *Statistical applications for the behavioral sciences*. New York: Wiley.

Shavelson, R. J. (1988). *Statistical reasoning for the behavioral sciences*. (2nd ed.). Boston: Allyn and Bacon.

Sprinthall, R. C. (1990). *Basic statistical analyses*. Englewood Cliffs, NJ: Prentice Hall.