**Psy 633 Multiple Regression
Questions for Discussion**

1. In regression, an independent variable is called the predictor (x) variable and the dependent variable is called the criterion (y) variable.
2. What is it called when more than one independent variable is used to predict the dependent variable? Multiple regression
3. Describe the "least squares estimation" procedure. This is the mathematical process involved in establishing the best fitting line through the data points. It minimized squared deviations of the data points from the line of best fit. The total amount of prediction error (both positive and negative) is as small as possible, giving us the best mathematically achievable line through the set of points.
4. In specifying the regression equation, the constant represents the y-intercept . The regression coefficients represent the slope for their particular variable.
5. Explain what a partial correlation is.
The regression coefficients represent the independent contribution of each independent variable to the prediction of the dependent variable. Independent contribution refers to amount variable *x1* is correlated with *y* variable after controlling for all other independent variables; this independent contribution is a partial correlation.
6. 1-r2 = residual variance.  Explain. Since r2 is the total amount of variability explained by the regression model, 1-r2 is the total amount of **unexplained** variability.
7. What is the easiest/most common way to reduce unexplained residual variability? Add another predictor *(x)* variable.
8. How many variables should one include in a regression equation? We should have 15-20 respondents for each variable included in the regression equation. Otherwise, the estimates will be very unstable.
9. What is multicollinearity?  Why is this a problem in multiple regression?
Multicollinearity exists when there are moderate to high intercorrelations among IVs to be used in a regression equation. The variables are redundant (explain the same variability in y). Multicollinearity is a problem because 1) distorts the size of R because the IVs are going after much of the same variability on the DV; 2) results in difficulty interpreting individual effects of variables because they are confounded; and 3) increases variances, which ultimately results in a more unstable predictions.
10. What is the difference between hierarchical and stepwise regression? Both hierarchical and stepwise regression evaluate the significance of the contributions of individual variables at each stage of model building. With hierarchical regression, the researcher (usually influenced by theory) determines the order in which the variables are added. With stepwise regression, a computer determines the order in which the variables are added.