In regression, a series of additive and multiplicative weights are applied to the independent variable(s) to create a predicted value of the dependent variable (called y-hat, or ŷ).

The weights are such that the difference between ŷ and the actual values of y is as small as possible. This difference is called error.

- The ability to predict is a function of the relationship between two variables.
- If a relationship is perfect, and we know the individual’s value on x, we can determine that individual’s value on y.
- When the relationship is perfect, every individual with a given score on x will have the same score on y.
- When the relationship is not perfect, y varies at every value of x.
- When the relationship is not perfect, y varies at every value of x.
\[ \hat{y} = a + b(x) \]

**The predicted value of y**

**Additive constant**

**Multiplicative constant**

Flashback to high school: What does this look like?

\[ b = r_{xy} \left( \frac{s_y}{s_x} \right) \]

1. The standard deviation of x is set to 1

2. The standard deviation of x is set to the standard deviation of y

3. The fact that the relationship between x and y may not be perfect is taken into account

4. The mean of the transformed x variable is then set to equal the mean of y

\[ a = \overline{Y} - b \left( \overline{X} \right) \]
\[ \hat{y} = a + b \ (x) \]

The y – intercept is the predicted value of y if x = 0.

Predicted UGPA = 1.9 + .003 (SAT_v)

If someone got a zero on the SAT Verbal, we would expect them to get an UGPA of 1.9.

Predicted UGPA = 1.9 + .003 (SAT_v)

The slope tells us by how much the predicted value of y increases for every point increase in x.

Predicted UGPA = 1.9 + .003 (SAT_v)

For every point above zero, someone scores on the SAT_V, we expect their GPA to go up by .003.

Predicted UGPA = 1.9 + .003 (SAT_v)

What would happen if we converted x and y into z-scores before calculating the constants?

\[ b = r_{xy} \left( \frac{s_y}{s_x} \right) \]
\[ a = \bar{Y} - b \left( \bar{X} \right) \]
What would happen if we converted $x$ and $y$ into z-scores before calculating the constants?

$$b = r_{xy} \left( \frac{1}{1} \right) \quad a = 0 - b \left( 0 \right)$$

What would happen if we converted $x$ and $y$ into z-scores before calculating the constants?

$$\beta = r_{xy} \quad a = 0$$

Instead of “$b$” the multiplicative constant is called a Beta-weight.

There is no additive constant – the line crosses the y-axis at the origin.

A correlation is the slope of the line describing the relationship between 2 z-scores.

Put another way, a correlation tells you how many standard deviations you can expect $y$ to increase for every standard deviation $x$ increases from the mean.

$$\hat{y} = a + b \left( x \right)$$

$$Z_{\hat{y}} = \beta \left( Z_x \right)$$
$Z\hat{Y} = \beta (Zx)$

- The predicted value of $Z_y$
- The additive constant for $z$-scores is always 0.
- The correlation between $x$ and $\hat{y}$
- The slope of the standardized regression line
- The expected increase in $x$ for every standard deviation unit that $x$ increases.

What if we want to use more than 1 independent variable to predict a dependent variable?

- **Multiple regression** uses more than 1 independent variable to predict a single dependent variable.

What if the multiple predictors are correlated with one another?
- The regression equation looks the same.

**Multiple Regression**

\[ \hat{y} = a + b_1 (x_1) + b_2 (x_2) \]

\[ Z\hat{y} = \beta_1 (Z_{x1}) + \beta_2 (Z_{x2}) \]

But now, $X_1$ and $X_2$ overlap with one another.

The $R^2$ still partitions the variance of $Y$ into $\hat{y}$ and error.
- The variance of $\hat{y}$ is now the variance of $Y$ that is explained by either $X_1$ or $X_2$.

We can now think of $\hat{y}$ as being comprised of the portion of $Y$ that:
- Is predicted only by $X_1$
- Is predicted only by $X_2$
- Is predicted both by $X_1$ and $X_2$. 
The zero-order correlation between $X_1$ and $Y$ tells you how much $X_1$ directly contributes to the prediction of $Y$.

The zero-order correlations tell you how much a variable directly contributes to the prediction of $Y$.

$R^2$ can no longer equal the sum of $r_{1y}^2$ and $r_{2y}^2$, because that would mean the center portion would be counted twice.

Rather, $R^2$ is simply the percentage of $Y$ that is $\hat{y}$.

$R^2$ is a “model statistic” that tells you how much the combination of predictors explains.

In simple regression:

- $R = \hat{\beta} = r_{xy}$

In multiple regression with uncorrelated predictors:

- $R \neq \hat{\beta} = r_{xy}$
- $R^2 = \Sigma r_{iy}^2$

However, in multiple regression with correlated predictors:

- $R \neq \hat{\beta} \neq r_{xy}$
- $R^2 \neq \Sigma r_{iy}^2$

In multiple regression with correlated predictors:

$\hat{\beta}_1 = r_{1y} - (r_{y2})(r_{12}) \over 1 - r_{12}^2$

This is the formula for a beta-weight when you have two correlated predictors.

This formula grows more complex as you add more predictors.

You will not be asked to calculate anything using this formula, however keep in mind that a beta weight...
\[ \beta_1 = \frac{r_{1y} - (r_{y2})(r_{12})}{1 - r_{12}^2} \]

- In the case of uncorrelated predictors, \( r_{12} = 0 \).

\[ \beta_1 = \frac{r_{1y} - (r_{y2})(r_{12})}{1 - r_{12}^2} \]

\[ \beta_1 = \frac{r_{1y} - (r_{y2})(0)}{1 - 0^2} \]

- In the case of uncorrelated predictors, \( r_{12} = 0 \).

\[ \beta_1 = r_{1y} \]

- In the case of uncorrelated predictors, \( r_{12} = 0 \).

\[ \beta_1 = \frac{r_{1y} - (r_{y2})(r_{12})}{1 - r_{12}^2} \]

- In the case of correlated predictors, \( r_{12} \neq 0 \) and \( \beta_{12} \neq r_{12} \).
- Beta-weights are model-dependent. This means that if we remove a predictor from a model, all of the beta-weights will change.
Zero-order correlations tell us how much a variable directly contributes to the prediction of $Y$.

Beta-weights tell us how much credit a variable gets for the prediction of $Y$, in the context of the other predictors.

Both zero-order correlations and Beta-weights are standardized coefficients; therefore, we can compare the absolute values of these coefficients across predictors to learn something about the relative strength of predictors.

Remember, we compare correlations to correlations and Betas to Betas. We do not compare Betas to correlations.

While in regression we can use zero-order correlations to examine the direct contribution of a predictor, other statistics use structure coefficients.

A structure coefficient is a correlation between a latent/synthetic variable and a measured/observed variable.

In multiple regression, the structure coefficients are the correlations between the predictors and the predicted value of $\hat{y}$:

$$r_s = r_{y\hat{y}}$$

How many variables?

There are 5 variables in this multiple regression equation:

- 3 are measured or observed: $X_1$, $X_2$, $Y$
- 2 are latent or synthetic: $\hat{y}$ or error

Structure coefficients and zero-order correlations both give the same information, though in different metrics; either can be used for the same information.
<table>
<thead>
<tr>
<th>$r_{xy}$ or $r_s$</th>
<th>$\beta$</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>High, relative to the other predictors</td>
<td>High, relative to the other predictors</td>
<td>This is a strong overall predictor – it directly contributes a lot, and gets a lot of credit</td>
</tr>
<tr>
<td>Low, relative to the other predictors</td>
<td>Low, relative to the other predictors</td>
<td>This is a lousy overall predictor – it directly contributes very little, and gets very little credit</td>
</tr>
<tr>
<td>High, relative to the other predictors</td>
<td>Low, relative to the other predictors</td>
<td>This predictor directly contributes a lot, but doesn't get much credit for what it predicts.</td>
</tr>
</tbody>
</table>

- What would this look like as a Venn diagram?
- The variable doesn’t get credit for what it directly contributes because other predictors also explain the same parts of $Y$, and they get credit for it instead.
<table>
<thead>
<tr>
<th>$r_{xy}$ or $r_\beta$</th>
<th>$\beta$</th>
<th>Interpretation²</th>
</tr>
</thead>
<tbody>
<tr>
<td>High, relative to the other predictors</td>
<td>High, relative to the other predictors</td>
<td>This is a strong overall predictor – it directly contributes a lot, and gets a lot of credit.</td>
</tr>
<tr>
<td>Low, relative to the other predictors</td>
<td>Low, relative to the other predictors</td>
<td>This is a lousy overall predictor – it directly contributes very little, and gets very little credit.</td>
</tr>
<tr>
<td>High, relative to the other predictors</td>
<td>Low, relative to the other predictors</td>
<td>This predictor directly contributes a lot, but doesn’t get much credit for what it predicts.</td>
</tr>
<tr>
<td>Low, relative to the other predictors</td>
<td>High, relative to the other predictors</td>
<td>This predictor directly contributes very little, but is getting credit for what it doesn’t directly predict.</td>
</tr>
</tbody>
</table>

What would this look like as a Venn diagram?  

A suppressor variable is a variable that makes other predictors better at predicting the dependent variable.  
A suppressor variable is generally useless as a predictor by itself.  
A suppressor variable is only useful as a predictor in the context of other variables.  

Suppressor effects as basketball.  
http://en.wikipedia.org/wiki/Point_(basketball)

To interpret multiple regression results:  
1. **Check the $R^2$.** This is the percentage of variance in $Y$ explained by the model and is an indication of, overall, how accurate the model is in predicting $Y$.  
2. **Check the Beta weights** – consider each variable’s Beta relative to the Betas of the other variables in the model. Think of this as indicating how much credit a variable gets for predicting $Y$.  
The sign of the Beta tells you the direction of the relationship between the variables.
3. **Check the zero-order correlations** – consider each variable’s correlation relative to the correlations of the other variables in the model. Think of this as indicating how much a variable directly contributes to the prediction of Y.

4. **Compare the relative Betas to the relative correlations**. Strong disagreement between the two can be interpreted as shared variance or a suppressor effect.

<table>
<thead>
<tr>
<th>Descriptive Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variable</strong></td>
</tr>
<tr>
<td>Global Health Questionnaire</td>
</tr>
<tr>
<td>Marital Opinion Questionnaire</td>
</tr>
<tr>
<td>Sternberg Intimacy Scale</td>
</tr>
<tr>
<td>Passionate Love Scale</td>
</tr>
</tbody>
</table>

- Relationship satisfaction has the strongest relationship with health, followed closely by intimacy, with passion a distant third.
- Higher levels of each of the predictors are associated with fewer health problems.
<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>1.733</td>
<td>.166</td>
<td>10.426</td>
<td>.000</td>
</tr>
<tr>
<td>Marriage Opinion Questionnaire</td>
<td>-2.318</td>
<td>.166</td>
<td>-4.582</td>
<td>.000</td>
</tr>
<tr>
<td>Sternberg Intimacy Scale</td>
<td>-.006</td>
<td>.048</td>
<td>-.137</td>
<td>.881</td>
</tr>
<tr>
<td>Passionate Love Scale</td>
<td>-.218</td>
<td>.042</td>
<td>-.015</td>
<td>.911</td>
</tr>
<tr>
<td>Predictor</td>
<td>r</td>
<td>Beta</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfaction (MOQ)</td>
<td>-3.81</td>
<td>-.479</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intimacy</td>
<td>-3.01</td>
<td>-.015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Passion</td>
<td>-0.49</td>
<td>.212</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The b-weights can be used to create the equation used to predict raw scores on the GHQ.

The Beta-weights are an indication of how much credit a variable gets for predicting GHQ scores.

Intimacy, which directly contributes nearly as much as satisfaction to the prediction of health, gets the least credit in the context of the other variables.

Why?
Passion, which directly contributes next-to-nothing (less than 1 percent, 0.2%) to the prediction of health, gets a fair amount of credit in the context of the other variables.

- The model with passion explains 17.6% of the variance in health.
- The same model with passion removed explains 14.8% of the variance in health.
- How can passion, which only directly predicts less than 1% of health cause the model to explain an additional 2.8% of the variance in health?
- Where is the other 2.6% coming from?

<table>
<thead>
<tr>
<th>Predictor</th>
<th>r</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisfaction (MOQ)</td>
<td>-.381</td>
<td>-.479</td>
</tr>
<tr>
<td>Intimacy</td>
<td>-.301</td>
<td>-.015</td>
</tr>
<tr>
<td>Passion</td>
<td>-.049</td>
<td>.212</td>
</tr>
</tbody>
</table>

Table 1
Summary of Multiple Regression Analysis for Relationship Variable Predicting Global Health

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Zero-order</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relationship Satisfaction (MOQ)</td>
<td>-.381</td>
<td>-.479</td>
</tr>
<tr>
<td>Sternberg Intimacy Scale</td>
<td>-.301</td>
<td>-.015</td>
</tr>
<tr>
<td>Passionate Love Scale</td>
<td>-.049</td>
<td>.212</td>
</tr>
</tbody>
</table>

Note. $R^2 = .176$. Dependent Variable = Global Health Questionnaire. MOQ = Marital Opinion Questionnaire.

Part and Partial Correlations

- Zero-order correlations tell you how much a variable directly contributes to the prediction of $Y$.
- Zero-order correlations do not take into account overlap between the predictors.
- If we wanted, we could remove the influence of the other predictors from $X_1$.
- This is called a part correlation or semi-partial correlation.

The part correlation can be squared to let you know the percentage of variance that is uniquely contributed by a predictor.
A variable with a large zero-order correlation and a relatively small part correlation may directly contribute a lot, but what it predicts is largely redundant with the other predictors.

If we want to know what the effect of \( X_1 \) is on \( Y \) after removing the effects of the other predictors, we can use a **partial correlation**.

Essentially, \( X_2 \) is removed from the equation, including any portions of \( X_1 \) and \( Y \) explained by \( X_2 \).

A partial correlation is the relationship between two variables after controlling for the effect of an additional variable(s).

When a variable is “controlled”, it is statistically removed from the other variables.

Partial correlations are often used when the relationship between two variables is suspected to be an artifact of a third variable.

For example, if we consider “days” as cases, the amount of ice cream consumed in a given city has a strong zero-order correlation with the number of drowning deaths in that city.

- Does ice cream cause people to drown?
- Does having a loved one drown cause people to eat ice cream?
- Or is the relationship due to a third variable?

What would happen if we partialed out the daily temperature?

- Suppose the zero-order correlation between gallons of ice cream consumed and number of drowning deaths was .5.
- The partial correlation between ice cream consumption and number of drowning deaths (controlling for temperature) was .001.
- What does this mean?

To get these correlations from SPSS:

- Analyze => Regression => Linear
- Select variables
- Statistics … Part and Partial Correlations
Relationship satisfaction contributes a fair amount to the prediction of health, and much of what it predicts is not predicted by the other variables.

Intimacy directly contributes a fair amount to the prediction of health, but makes only a miniscule unique contribution in the context of the other variables.

This is an indication of how much of what a predictor explains is unique.

How can we explore the relationship between relationship satisfaction and intimacy and passion?

There is a strong relationship between how passionate a relationship is and how satisfying a relationship is. But nearly all of that relationship can be explained by the level of intimacy in a relationship.
<table>
<thead>
<tr>
<th>Model</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
<th>Beta</th>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.097</td>
<td>.201</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.481</td>
<td>.631</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.714</td>
<td>.030</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.862</td>
<td>.23.661</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.000</td>
<td>.865</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.807</td>
<td>.686</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.004</td>
<td>.005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.142</td>
<td>.887</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.526</td>
<td>.008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.004</td>
<td>.004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There is a strong relationship between how passionate a relationship is and how satisfying a relationship is...

But nearly all of that relationship can be explained by the level of intimacy in a relationship.

Conversely, the relationship between satisfaction and intimacy cannot be explained by passion.

---

Categorical Variables in Regression

- So far, we have dealt with regression using continuous variables.
- However, categorical variables can also be used in regression.
- In order to be used in regression, the categorical variables must be modeled as continuous variables.
- We will consider two types of coding:
  - Dummy Coding
  - Contrast Coding

---

Dummy Coding

- Dummy coding involves representing a categorical variable as a series of continuous variables, each with only two possible values: 0 and 1.
- The number of dummy variables necessary to represent a categorical variable is k - 1, where k is the number of categories.
Dummy Coding

- Dichotomous categorical variables (e.g., gender)
- Which status (male or female) is assigned the value of 1 or 0 is arbitrary, but needs to be noted:
  - Male = 0 (female) or Male = 1 (male)
  - Female = 0 (male) or Female = 1 (female)
- This dummy-coded variable can now be used as a predictor in regression

\[ \hat{y} = a + b \text{ (male)} \]

The predicted value of \( y \)

Additive constant
The y-intercept
The point at which the regression line crosses the y-axis.
The predicted value of \( y \) when \( x=0 \)

Multiplicative constant
The slope (rise/run)
The expected increase in \( y \) for every point that \( x \) increases.

\[ \hat{y} = a + b \text{ (male)} \]

The predicted value of \( y \)

The average value of \( y \) for women

The difference between the average score of men and the average score of women:
Sign indicates the direction of the difference

Description Statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>4.035</td>
<td>1.49071</td>
<td>10</td>
</tr>
<tr>
<td>Male</td>
<td>4.025</td>
<td>1.42273</td>
<td>10</td>
</tr>
</tbody>
</table>

The mean of male is the percentage of the sample who are male

Coefficient

<table>
<thead>
<tr>
<th></th>
<th>Constant</th>
<th>Dependent Variable</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.104</td>
<td>Dependent Variable</td>
<td>4.300</td>
<td>.005</td>
</tr>
</tbody>
</table>

The mean \( y \) value of women (male=0) is 3.2

The mean \( y \) of men is 2.4 points lower than the mean \( y \) of women
The same results would be obtained using:
- A one-way ANOVA
- An independent samples t-test
- A point-biserial correlation
- It’s all the general linear model!

Categorical variables with more than 2 categories require more than 1 dummy variable (k-1, to be precise).
Choosing how to model the dummy variables is important, as you will be dealing with planned contrasts, rather than ANOVA-style omnibus tests of effect.
This means you have to think, and consider theory!

Consider a categorical variable “couple type” with four possible values:
- Heterosexual married
- Heterosexual cohabiting
- Gay cohabiting
- Lesbian cohabiting
With a standard approach to dummy coding, one category is selected as the basis for comparison.

\[ \hat{y} = a + b_1 \text{ (cohabit)} + b_2 \text{ (gay)} + b_3 \text{ (lesbian)} \]

\[ \hat{y} = a + b_1 \text{ (cohabit)} + b_2 \text{ (gay)} + b_3 \text{ (lesbian)} \]

<table>
<thead>
<tr>
<th>Couple Type</th>
<th>Cohabit</th>
<th>Gay</th>
<th>Lesbian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterosexual married</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Heterosexual cohabiting</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gay cohabiting</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Lesbian cohabiting</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

The comparison group should be selected based on theory. Here, all couple types are being compared to married heterosexual couples.

<table>
<thead>
<tr>
<th>Mean of married couples</th>
<th>Difference between cohabiting and married couples</th>
<th>Difference between gay and married couples</th>
<th>Difference between lesbian and married couples</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ \hat{y} ]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
If you were theoretically interested in testing, you could also dummy-code these variables to test:

- Same-sex vs. different-sex
- Married vs. Cohabiting
- Heterosexual vs. Gay vs. Lesbian
- If dealing with individuals, gender and same-sex

### Simple Contrast Coding

<table>
<thead>
<tr>
<th>Couple Type</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterosexual married</td>
<td>x1 -1</td>
</tr>
<tr>
<td>Heterosexual cohabiting</td>
<td>1</td>
</tr>
<tr>
<td>Gay cohabiting</td>
<td>0</td>
</tr>
<tr>
<td>Lesbian cohabiting</td>
<td>0</td>
</tr>
</tbody>
</table>

- Compares all groups to a contrast group (as dummy coding)
- The intercept of the equation is the grand mean
- The slopes tell you by how much the group differs from the grand mean
- This is the same as the dummy coding example, but the slopes need to be doubled to indicate the differences between groups.

### Forward Difference Contrast Coding

<table>
<thead>
<tr>
<th>Couple Type</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterosexual married</td>
<td>x1 1</td>
</tr>
<tr>
<td>Heterosexual cohabiting</td>
<td>-1</td>
</tr>
<tr>
<td>Gay cohabiting</td>
<td>0</td>
</tr>
<tr>
<td>Lesbian cohabiting</td>
<td>0</td>
</tr>
</tbody>
</table>

- Each group is compared to the next group down the line — useful when you have an a priori order in which you expect groups to fall.
- X1 compares married and cohabiting heterosexual couples
- X2 compares cohabiting heterosexual and gay couples
- X3 compares gay and lesbian couples

### Backward Difference Contrast Coding

<table>
<thead>
<tr>
<th>Couple Type</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterosexual married</td>
<td>x1 -1</td>
</tr>
<tr>
<td>Heterosexual cohabiting</td>
<td>-1</td>
</tr>
<tr>
<td>Gay cohabiting</td>
<td>0</td>
</tr>
<tr>
<td>Lesbian cohabiting</td>
<td>0</td>
</tr>
</tbody>
</table>

- Each group is compared to the previous group — useful when you have an a priori order in which you expect groups to fall.
- This is the same as the dummy coding example, but the slopes need to be doubled to indicate the differences between groups.
- X1 compares married and cohabiting heterosexual couples
- X2 compares cohabiting heterosexual and gay couples
- X3 compares gay and lesbian couples

### Helmert Contrast Coding

<table>
<thead>
<tr>
<th>Couple Type</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterosexual married</td>
<td>x1 3</td>
</tr>
<tr>
<td>Heterosexual cohabiting</td>
<td>-1</td>
</tr>
<tr>
<td>Gay cohabiting</td>
<td>-1</td>
</tr>
<tr>
<td>Lesbian cohabiting</td>
<td>-1</td>
</tr>
</tbody>
</table>

- Each group is compared to the mean of the following groups.
- This is the same as the dummy coding example, but the slopes need to be doubled to indicate the differences between groups.
- X1 compares married to all other couples
- X2 compares cohabiting heterosexual to gay and lesbian couples
- X3 compares gay and lesbian couples
■ Reverse Helmert
■ User Defined –
  ■ Basically, you can set up any comparisons you want. If you want them to be orthogonal (uncorrelated), you must make sure the contrast codes are uncorrelated with one another
  ■ If you have unequal cell sizes, the contrast numbers must be adjusted to account for this
  ■ http://www.bolderstats.com/orthogCodes/

Moderators

■ Moderators exist when the relationship between two variables differs as a function of a third variable
■ Moderator effects are also sometimes called interactions
■ For example, suppose that attributions moderate the relation between relationship quality and stress.

Relationship between Relationship Quality and Stress at Different Levels of Attributions

■ On average, there was a trend evident with high amounts of stress being associated with low levels of relationship quality.

Regression between Wives’ Relationship Quality and Stress at Different Levels of Relationship Attributions

■ For couples making distress maintaining attributions (RAM averages 1 S.D. higher than the mean), the relationship between stress and relationship quality was more marked and highly statistically significant (p < .001).
Regression between Wives’ Relationship Quality and Stress at Different Levels of Relationship Attributions

- For couples making relationship-enhancing attributions (RAM averages 1 S.D. below the mean), the relationship between stress and relationship quality was negated.

Regression between Wives’ Relationship Quality and Stress at Different Levels of Relationship Attributions

- In actuality, because attributions are continuous, a different regression line exists for every attribution value.

Creating Moderators

- Moderators are created by multiplying the two predictor variables together

\[ +1 \times -1 = -1 \]

\[ +1 \times +1 = +1 \]

\[ -1 \times -1 = +1 \]

\[ -1 \times +1 = -1 \]
However, simply multiplying two variables together creates problems with \textit{multicollinearity} – the parent terms are highly correlated with the interaction term.

To solve this, the parent terms must be \textit{centered} about their means before they are combined to create the interaction term.

\[ X_{\text{centered}} = X - \bar{X} \]

To create a moderator variable, you can use either syntax or point-and-click to get to the compute command.

\[
\text{Compute } AxB = (\text{VarA} - \text{MeanA}) \times (\text{VarB} - \text{MeanB}).
\]

\[ \hat{y} = b_0 + b_1 X + b_2 Z + b_3 XZ \]

**What does the interaction mean?**
- The interaction is the change in \( b_1 \) for every point by which \( Z \) changes.
- The interaction is the change in \( b_2 \) for every point by which \( X \) changes.

**Visualizing interactions:**
- \( \text{http://www.ats.ucla.edu/stat/sas/faq/spplot/reg_int_cont.html} \)
- \( \text{Free download: } \text{http://www.provalisresearch.com/ITALASSI/ITALdowload.html} \)

**Help!**
- **My interaction is statistically significant! What do I do?**
- When we move into HLM, the meaning of the interactions will be self-evident.
- Most often, researchers will plot statistically significant interactions in order to interpret them.

**Plotting Interactions**
- Generally, three regression lines are plotted:
  - The regression line between \( x \) and \( y \) when \( z = 0 \).
  - The regression line between \( x \) and \( y \) when \( z \) is one standard deviation above the mean.
  - The regression line between \( x \) and \( y \) when \( z \) is one standard deviation below the mean.
- \( \pm 1 \sigma \) is used as an arbitrary value, when no other meaningful value presents itself. If there are meaningful anchors, they can be used:
  - Min/Max scores
  - Cut-off scores

**Plotting Interactions**
- If one of the variables is categorical, two regression lines are plotted, one for each of the two categories.
- Depending on the story told by the data, more regression lines can be plotted.
Plotting Interactions

- In order to plot the relation between x and y at different values of z,
  \[ \hat{y} = b_0 + b_1x + b_2z + b_3xz \]
  is restructured to:
  \[ \hat{y} = (b_1 + b_2z)x + (b_2z + b_3) \]
- Plugging in the values of the unstandardized coefficients and Z will give you the equation at different levels of Z.
- You can then plot the 3 different lines in a figure.

Post Hoc Probing

The slope is:
\[ \text{Slope} = b_1 + b_2z \]

The standard error of the slope is:
\[ \text{Standard error} = \sqrt{s_{11} + 2zs_{13} + Z^2s_{33}} \]

- \( s_{11} \) is the squared standard error for slope \( b_1 \)
- \( s_{33} \) is the squared standard error for slope \( b_3 \)
- \( s_{13} \) is the covariance between \( b_1 \) and \( b_3 \)

=> Obtained from statistics … regression coefficients (covariance matrix)

- A simple slope test (t = b / s.e.) with n-k-1 df is then used (k is the number of predictors, not including the intercept)

Create interaction term
- Center the parent terms and multiply them
- Run regression including parent and interaction terms
- If an interaction is stat. sig., plot it to interpret the interaction
- Simple slope tests may also be useful
**Mediators**

- Mediators describe the process through which one variable impacts another.

**Diagram 1:**

```
Depression  -->  Relationship Quality
```

**Mediators**

- The logic of a mediation test is fairly straightforward.
- However, there are several different ways to calculate a mediation test. The following is taken from Baron and Kenny.

**Diagram 2:**

```
X  -->  M  -->  Y
```

- Consider the above model – we would like to know whether M mediates the effect of X on Y.

**Diagram 3:**

```
X  -->  Y
```

- First, the relation between x and y must be established.
- Conduct a regression analysis, predicting Y with X.
  - If this is not statistically significant, there is no effect to mediate.
  - If this is statistically significant, there is an effect to mediate.
  - Make note of the unstandardized coefficient and standard error for path c.

**Diagram 4:**

```
X  -->  M
```

- Second, you must demonstrate a relationship between X and M.
- Conduct a regression analysis, predicting M with X.
  - If this is not statistically significant, there is no mediation.
  - If this is statistically significant, there might be mediation.
  - Make note of the unstandardized coefficient and standard error for path a.
Finally, run a multiple regression analysis predicting Y with both M and X

- In order to differentiate the path from X to Y from the previous analysis, the path is called c’ instead of c
- We expect b to be statistically significant, but c’ to be near zero
- Make note of the path coefficients and standard errors

If a, b, and c are non-zero (or statistically significant) and c’ is zero (or statistically non-significant), then M mediates the relation between X and Y

- This type of mediation (when c’ is 0) is complete mediation
- What about partial mediation?

There are some limits to the Sobel test –
- You should be able to support the causal direction of the model.
- The results are highly suspect to measurement error
- Alternative tests of mediation exist, and are possibly better at testing mediation:
  - For example, bootstrapping:
  - http://www.psychwiki.com/wiki/Mediation

A Sobel test can be used to determine whether the mediated path (a*b) is statistically significant.

- Multiple on-line calculators are available:
  - http://people.ku.edu/~preacher/sobel/sobel.htm
- In essence, the path coefficient (a*b) is divided by its standard error (SQRT(b^2s_e^2 + a^2s_b^2)) and compared to the probabilities of the normal distribution

HLM

PSY 515
Jim Graham
Fixed Effects

- So far, most of we have dealt with uses what are called fixed effects.
- A fixed variable is assumed to be measured without error.
- Regression considers $y$ as having error, but not the predictors:
  \[
  \hat{y} = a + b \ (x) \\
  y = a + b \ (x) + \text{error} \\
  y = \hat{y} + \text{error}
  \]

Random Effects

- The regression coefficients vary within a population.
- The regression equation obtained from one study can differ from the equation obtained from another study using the same population.
- The regression coefficients vary within a population.

Fixed Effects

- In a fixed effects model, the regression coefficients are assumed to be the same for everyone in the population.
- The regression equation obtained from one study should be the same as the equation obtained from another study using the same population.
- The regression coefficients are invariant within a population.

Random Effects

- Fixed effects cannot be generalized beyond the population from which the sample is drawn, or beyond the specific values used in the study.
  - e.g., drug outcome study examining 0, 5, and 10 mg of a drug cannot tell us anything about what taking 7 mg of the drug might do.
- Random effects can be generalized more broadly, and results can be generalized beyond the specific values.
- Ordinary Least Squares: Coefficients minimize the squared difference between the actual and predicted values of $y$; These values are maximized for a given sample; Cross-validation must be used to examine generalizability.
- Maximum Likelihood: Maximizes the fit of the data to the model. This uses an iterative process, where a series of values are used until a good fit is found. These coefficients maximize the likelihood function
- A likelihood function is the probability (or probability density) for the occurrence of a sample configuration, given a known probability density.

**HLM**
- HLM is multi-level mixed-effects regression.
- A great deal of data exist in a nested structure:
  - Assignments within students within classes within teachers within schools within districts within states within regions,
  - Days within individuals within therapists
  - Trials within people within experimental conditions
  - People within families within countries
- Once you start thinking about nested data structures, you start seeing them EVERYWHERE

**Variables in HLM**
- In HLM, variables exist at different levels.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Experience Level</th>
<th>Individual Level</th>
<th>City Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mood State</td>
<td>Vary moment by moment within individuals</td>
<td>Vary from individual to individual</td>
<td>Vary from city to city</td>
</tr>
<tr>
<td>Gender</td>
<td>Is constant within individuals</td>
<td>Varies from individual to individuals</td>
<td>Varies from city to city</td>
</tr>
<tr>
<td>Crime Rate</td>
<td>Are constant within individuals</td>
<td>Are constant for individuals in the same city</td>
<td>Varies from city to city</td>
</tr>
</tbody>
</table>
We are interested in studying the relationships between mood states, stress level, and gender.

- A fixed-effects approach might manipulate the mood of participants and look at the effect on the level of stress,
- or manipulate the level of stress experienced and look at changes in mood.
- Neither of these approaches would allow one to generalize beyond those specific levels.

Using HLM, we might consider experiences nested within individuals.

- Mood states and stress level vary within individuals moment-by-moment and between individuals. Level 1
- Gender (dummy-coded 0=male, 1=female) is constant within individuals moment-by-moment and varies between individuals. Level 2

6 individuals (3 men and 3 women) are followed over the course of a day and their mood and stress level are sampled 10 times.

For each individual, a regression equation using stress level to predict mood

A series of regression equations are then obtained:

For each individual, a regression equation using stress level to predict mood

A series of regression equations are then obtained:

Male 1: Mood = 4 - .5 (Stress) + e
Male 2: Mood = 3 - .6 (Stress) + e
Male 3: Mood = 6 - .4 (Stress) + e
Female 1: Mood = 6 - .9 (Stress) + e
Female 2: Mood = 8 - .8 (Stress) + e
Female 2: Mood = 7 - .7 (Stress) + e

The regression coefficients are NOT treated as constants. Rather, they vary across individuals.

Male 1: Mood = 4 - .5 (Stress) + e
Male 2: Mood = 3 - .6 (Stress) + e
Male 3: Mood = 6 - .4 (Stress) + e
Female 1: Mood = 6 - .9 (Stress) + e
Female 2: Mood = 8 - .8 (Stress) + e
Female 2: Mood = 7 - .7 (Stress) + e

In a general form, each individual has their own regression equation, in the form of:

Mood = B0 + B1 (Stress) + error

Because B0 and B1 are variables (and not constants), they can be used as variables at level 2.
Level 1: Mood = B0 + B1 (Stress) + error

Level 2: B0 = G00 + G01 (Gender) + error

Remember that B0 is the mood of an individual when stress is equal to 0.

This equation predicts mood under no stress with gender.

Gender is dummy coded as 0 = male, 1 = female.

The regression weight B0 from level 1 becomes the outcome variable at level 2.

Remember that B0 is ... ?

The value of B0 when Gender=0

The mood under no stress when Gender=0.
Level 1: \[ \text{Mood} = B_0 + B_1 \text{(Stress)} + \text{error} \]
Level 2: \[ B_0 = G_{00} + G_{01} \text{(Gender)} + \text{error} \]

The mood of men under no stress.

Level 1: \[ \text{Mood} = B_0 + B_1 \text{(Stress)} + \text{error} \]
Level 2: \[ B_0 = G_{00} + G_{01} \text{(Gender)} + \text{error} \]

The increase in mood under no stress for every point increase in gender.

Remember that \( B_1 \) is the increase in mood for every point that stress increases.

This equation predicts the relationship between stress and mood.
Level 1: \[ \text{Mood} = B_0 + B_1 \text{(Stress)} + \text{error} \]
Level 2: \[ B_0 = G_00 + G_01 \text{(Gender)} + \text{error} \]
\[ B_1 = G_{10} + G_{11} \text{(Gender)} + \text{error} \]

The value of $B_1$ for men.

The increase in mood for every point that stress increases for men.

The increase in the relationship between mood and stress for men.

The increase in B1 for every point that gender increases.

The increase in the relationship between mood and stress for men.

The difference between men and women in the relationship between stress and mood.

The relationship between mood and stress for men.

Cross-level Interaction!
### Level 1: Mood $= B_0 + B_1 \text{(Stress)} + \text{error}$

### Level 2: $B_0 = G_{00} + G_{01} \text{(Gender)} + \text{error}$
$B_1 = G_{10} + G_{11} \text{(Gender)} + \text{error}$

- $B_0, G_{00}$ = The mood of men under no stress.
- $B_1, G_{10}$ = The increase in mood for every point that stress increases for men.
- $G_{01}$ = The difference in mood under no stress between men and women
- $G_{11}$ = The difference between men and women in the relationship between stress and mood.

---

### Level 1: Mood $= 5 - .5 \text{(Stress)} + \text{error}$

### Level 2: $B_0 = 5 - 2 \text{(Gender)} + \text{error}$ $B_1 = .5 + .25 \text{(Gender)} + \text{error}$

- What is the mood of men under no stress?
- What is the mood of women under no stress?
- How does men’s mood change for every point increase in stress?
- How does women’s mood change for every point increase in stress?
- What is the difference in mood between men and women?
- What is the difference in the relationship between stress and mood between men and women?

### Centering

- In the preceding examples, variables are entered in their raw, uncentered forms.
- As such, 0 means 0.
- It is often advantageous to change the meaning of 0.
- HLM uses two forms of centering:
  - Group-mean centered: centered about one’s own level 1 “group” average.
    - It is that individual’s average
  - Grand-mean centered: centered about the overall average of everyone.
    - It is the average of everyone

---

### If stress is uncentered

### Level 1: Mood $= B_0 + B_1 \text{(Stress)} + \text{error}$

### Level 2: $B_0 = G_{00} + G_{01} \text{(Gender)} + \text{error}$
$B_1 = G_{10} + G_{11} \text{(Gender)} + \text{error}$

- $B_0, G_{00}$ = The mood of men under no stress.
- $B_1, G_{10}$ = The increase in mood for every point that stress increases for men.
- $G_{01}$ = The difference in mood under no stress between men and women
- $G_{11}$ = The difference between men and women in the relationship between stress and mood.

---

### Relationship between stress and mood

| Level 1: Mood $= B_0 + B_1 \text{(Stress)} + \text{error}$ | B1 | Stress and mood
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 2: $B_0 = G_{00} + G_{01} \text{(Gender)} + \text{error}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$B_1 = G_{10} + G_{11} \text{(Gender)} + \text{error}$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- $B_0, G_{00}$ = The mood of men under no stress.
- $B_1, G_{10}$ = The increase in mood for every point that stress increases for men.
- $G_{01}$ = The difference in mood under no stress between men and women
- $G_{11}$ = The difference between men and women in the relationship between stress and mood.

---

### Relationship between gender and mood

| Level 1: Mood $= B_0 + B_1 \text{(Stress)} + \text{error}$ | B1 | Gender and mood
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 2: $B_0 = G_{00} + G_{01} \text{(Gender)} + \text{error}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$B_1 = G_{10} + G_{11} \text{(Gender)} + \text{error}$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- $B_0, G_{00}$ = The mood of men under no stress.
- $B_1, G_{10}$ = The increase in mood for every point that stress increases for men.
- $G_{01}$ = The difference in mood under no stress between men and women
- $G_{11}$ = The difference between men and women in the relationship between stress and mood.

---

### Relationship between gender by stress and mood

| Level 1: Mood $= B_0 + B_1 \text{(Stress)} + \text{error}$ | B1 | Gender by stress and mood
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 2: $B_0 = G_{00} + G_{01} \text{(Gender)} + \text{error}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$B_1 = G_{10} + G_{11} \text{(Gender)} + \text{error}$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- $B_0, G_{00}$ = The mood of men under no stress.
- $B_1, G_{10}$ = The increase in mood for every point that stress increases for men.
- $G_{01}$ = The difference in mood under no stress between men and women
- $G_{11}$ = The difference between men and women in the relationship between stress and mood.

---

### Notes

- Centering: In the preceding examples, variables are entered in their raw, uncentered forms.
- As such, 0 means 0.
- It is often advantageous to change the meaning of 0.
- HLM uses two forms of centering:
  - Group-mean centered: centered about one’s own level 1 “group” average.
    - It is that individual’s average
  - Grand-mean centered: centered about the overall average of everyone.
    - It is the average of everyone

---

### Questions

1. What is the mood of men under no stress?
2. What is the mood of women under no stress?
3. How does men’s mood change for every point increase in stress?
4. How does women’s mood change for every point increase in stress?
5. What is the difference in mood between men and women?
6. What is the difference in the relationship between stress and mood between men and women?
If stress is group-mean centered

Level 1:  \[ \text{Mood} = B_0 + B_1 (\text{Stress}) + \text{error} \]

Level 2:  
- \[ B_0 = G_{00} + G_{01} (\text{Gender}) + \text{error} \]
- \[ B_1 = G_{10} + G_{11} (\text{Gender}) + \text{error} \]

- \( B_0, G_{00} = \) The average mood of men.
- \( B_1, G_{10} = \) The increase in mood for every point that stress increases for men.
- \( G_{01} = \) The average difference in mood between men and women
- \( G_{11} = \) The difference between men and women in the relationship between stress and mood.

Centering

- Generally, we will:
  - enter dummy variables uncentered
  - Group-mean center continuous variables at level 1
  - Grand-mean center continuous variables at level 2

3-level HLM

- Considering the precious example, suppose the participants we were looking at are members of romantic couples, and we’re interested in examining how the length of the relationship.

  Level 1: Mood, Stress
  - Varies within individuals, within romantic couples, and between couples

  Level 2: Gender
  - Constant within an individual and varies between members of the same romantic couple and between couples

  Level 3: Relationship Length
  - Constant within individuals and members of the same couple, varies between different couples

If stress is group-mean centered and gender is grand-mean centered

Level 1:  \[ \text{Mood} = B_0 + B_1 (\text{Stress}) + \text{error} \]

Level 2:  
- \[ B_0 = G_{00} + G_{01} (\text{Gender}) + \text{error} \]
- \[ B_1 = G_{10} + G_{11} (\text{Gender}) + \text{error} \]

- \( B_0, G_{00} = \) The average mood of the participants.
- \( B_1, G_{10} = \) The increase in mood for every point that stress increases.
- \( G_{01} = \) The average difference in mood between men and women
- \( G_{11} = \) The difference between men and women in the relationship between stress and mood.
**4-level HLM**

- Just kidding

---

**HLM**

- Data must be structured by level and sorted by ID:
  - Level 1 variables
  - Timestamp
  - Individual ID
  - Couple ID
  - Stress
  - Mood
  - Level 2 variables
    - Individual ID
    - Couple ID
    - Gender
  - Level 3 variables
    - Couple ID
    - Relationship Length

---

**Other HLM Considerations**

- Time series designs
- Persons within groups / measures within persons
- Within level interactions
- Effect Sizes, unexplained variance
- Robust vs. non-robust standard errors
Overview of Meta-Analytic Data Analysis

- Transformations, Adjustments and Outliers
  - The Inverse Variance Weight
  - The Mean Effect Size and Associated Statistics
  - Homogeneity Analysis
- Fixed Effects Analysis of Heterogeneous Distributions
  - Fixed Effects Analog to the one-way ANOVA
- Random Effects Analysis of Heterogeneous Distributions
  - Mean Random Effects ES and Associated Statistics
  - Random Effects Analog to the one-way ANOVA
  - Random Effects Regression Analysis

Transformations

- Some effect size types are not analyzed in their “raw” form.
- Standardized Mean Difference Effect Size
  - Upward bias when sample sizes are small
  - Removed with the small sample size bias correction
    \[ ES'_{z} = ES \left( 1 - \frac{3}{4N - 9} \right) \]

Transformations (continued)

- Correlation has a problematic standard error formula.
- Recall that the standard error is needed for the inverse variance weight.
- Solution: Fisher’s Zr transformation.
- Finally results can be converted back into “r” with the inverse Zr transformation

Transformations (continued)

- Analyses performed on the Fisher’s Zr transformed correlations.
  \[ ES_{z} = .5 \ln \left( \frac{1 + r}{1 - r} \right) \]
- Finally results can be converted back into “r” with the inverse Zr transformation.
  \[ r = \frac{e^{2ES_{z}} - 1}{e^{2ES_{z}} + 1} \]

Adjustments

- Hunter and Schmidt Artifact Adjustments
  - measurement unreliability (need reliability coefficient)
  - range restriction (need unrestricted standard deviation)
  - artificial dichotomization (correlation effect sizes only)
  - assumes an underlying distribution that is normal
- Outliers
  - extreme effect sizes may have disproportionate influence on analysis
  - either remove them from the analysis or adjust them to a less extreme value
  - indicate what you have done in any written report

Overview of Transformations, Adjustments, and Outliers

- Standard transformations
  - sample size bias correction for the standardized mean difference effect size
  - Fisher’s Z to t transformation for correlation coefficients
  - Natural log transformation for odds-ratios
- Hunter and Schmidt Adjustments
  - perform if interested in what would have occurred under “ideal” research conditions
- Outliers
  - any extreme effect sizes have been appropriately handled
Independent Set of Effect Sizes

- Must be dealing with an independent set of effect sizes before proceeding with the analysis.
- One ES per study OR
- One ES per subsample within a study (modeled as a nested data set)

The Inverse Variance Weight

- Studies generally vary in size.
- An ES based on 100 subjects is assumed to be a more "precise" estimate of the population ES than an ES based on 10 subjects.
- Therefore, larger studies should carry more "weight" in our analyses than smaller studies.
- Simple approach: weight each ES by its sample size.
- Better approach: weight by the inverse variance.

What is the Inverse Variance Weight?

- The standard error (SE) is a direct index of ES precision.
- SE is used to create confidence intervals.
- The smaller the SE, the more precise the ES.
- Hedges' showed that the optimal weights for meta-analysis are:

\[ w = \frac{1}{SE^2} \]

Inverse Variance Weight for the Major Effect Sizes

- Standardized Mean Difference:

\[ w = \frac{1}{\bar{SE}^2} \]

\[ \bar{SE} = \frac{\frac{1}{n_1} + \frac{1}{n_2}}{\frac{2\ES}{2(n_1 + n_2)}} \]

- Zr transformed Correlation Coefficient:

\[ w = \frac{1}{n - 3} \]

Ready to Analyze

- We have an independent set of effect sizes (ES) that have been transformed and/or adjusted, if needed.
- For each effect size we have an inverse variance weight (w).

The Weighted Mean Effect Size

- Start with the effect size (ES) and inverse variance weight (w) for 10 studies.

\[ \ES = \frac{\sum (w \times \ES)}{\sum w} \]
### The Weighted Mean Effect Size

Start with the effect size (ES) and inverse variance weight (w) for 10 studies. Next, multiply w by ES. Repeat for all effect sizes. Sum the columns, w and ES. Divide the sum of (w*ES) by the sum of (w).

<table>
<thead>
<tr>
<th>Study</th>
<th>ES</th>
<th>w</th>
<th>w*ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.33</td>
<td>11.91</td>
<td>-3.93</td>
</tr>
<tr>
<td>2</td>
<td>0.32</td>
<td>28.57</td>
<td>9.14</td>
</tr>
<tr>
<td>3</td>
<td>0.39</td>
<td>58.82</td>
<td>22.94</td>
</tr>
<tr>
<td>4</td>
<td>0.31</td>
<td>29.41</td>
<td>9.12</td>
</tr>
<tr>
<td>5</td>
<td>0.17</td>
<td>13.89</td>
<td>2.36</td>
</tr>
<tr>
<td>6</td>
<td>0.33</td>
<td>9.80</td>
<td>-3.24</td>
</tr>
<tr>
<td>7</td>
<td>0.15</td>
<td>10.75</td>
<td>1.61</td>
</tr>
<tr>
<td>8</td>
<td>0.00</td>
<td>14.93</td>
<td>0.00</td>
</tr>
</tbody>
</table>

\[
\bar{w} = \frac{\sum_{i=1}^{n}(w_i \times ES_i)}{\sum_{i=1}^{n}w_i} = \frac{41.82}{209.96} = 0.195
\]

### The Standard Error of the Mean ES

The standard error of the mean is the square root of \(1 \div \sum w\). \(SE_{\bar{w}} = \sqrt{\frac{1}{\sum w}} = 0.061\)

### Mean, Standard Error, Z-test and Confidence Intervals

- **Mean ES**: \(\bar{ES} = \frac{\sum_{i=1}^{n}(w_i \times ES_i)}{\sum_{i=1}^{n}w_i} = \frac{41.82}{209.96} = 0.195\)
- **SE of the Mean ES**: \(SE_{\bar{w}} = \sqrt{\frac{1}{\sum w}} = 0.061\)
- **Z-test for the Mean ES**: \(Z = \frac{\bar{ES} \times \sqrt{n}}{SE_{\bar{w}}} = \frac{0.195 \times \sqrt{10}}{0.061} = 4.64\)
- **95% Confidence Interval**: \(\text{Lower} = \bar{ES} - 1.96(0.061) = 0.195 - 1.20 = -0.1\)
\(\text{Upper} = \bar{ES} + 1.96(0.061) = 0.195 + 1.20 = 1.39\)

- **Mean ES**: \(\bar{ES} = 0.195\)
- **SE of the Mean ES**: \(SE_{\bar{w}} = 0.061\)
- **Z-test for the Mean ES**: \(Z = 4.64\)
- **95% Confidence Interval**: \([-0.1, 1.39]\)

- **What is the average effect size?**
- **In essence, we are estimating the population parameter.**
- **Is the average effect size statistically significantly different from 0?**
- **What is the 95% confidence interval about the mean effect size?**
- **Is it reasonable to assume that a single population is being represented?**
**Homogeneity Analysis**

- Homogeneity analysis tests whether the assumption that all of the effect sizes are estimating the same population mean is a reasonable assumption.
- If homogeneity is rejected, the distribution of effect sizes is assumed to be heterogeneous.
  - Single mean ES not a good descriptor of the distribution
  - There are real between study differences, that is, studies estimate different population mean effect sizes.
- Two options:
  - model between study differences
  - fit a random effects model

**Q - The Homogeneity Statistic**

<table>
<thead>
<tr>
<th>Study</th>
<th>ES</th>
<th>w</th>
<th>w*ES</th>
<th>w*ES^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.33</td>
<td>11.91</td>
<td>-3.93</td>
<td>1.30</td>
</tr>
<tr>
<td>2</td>
<td>0.32</td>
<td>28.57</td>
<td>8.40</td>
<td>2.69</td>
</tr>
<tr>
<td>3</td>
<td>0.39</td>
<td>58.82</td>
<td>22.94</td>
<td>6.95</td>
</tr>
<tr>
<td>4</td>
<td>0.11</td>
<td>20.41</td>
<td>2.25</td>
<td>0.26</td>
</tr>
<tr>
<td>5</td>
<td>0.17</td>
<td>13.89</td>
<td>2.39</td>
<td>0.40</td>
</tr>
<tr>
<td>6</td>
<td>0.84</td>
<td>8.35</td>
<td>5.47</td>
<td>3.16</td>
</tr>
<tr>
<td>7</td>
<td>-0.33</td>
<td>9.80</td>
<td>-3.24</td>
<td>1.07</td>
</tr>
<tr>
<td>8</td>
<td>0.35</td>
<td>10.75</td>
<td>1.61</td>
<td>0.24</td>
</tr>
<tr>
<td>9</td>
<td>-0.02</td>
<td>83.33</td>
<td>-1.67</td>
<td>0.03</td>
</tr>
<tr>
<td>10</td>
<td>0.00</td>
<td>14.93</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Calculating Q

We now have 3 sums:

\[ \sum w = 269.96 \]
\[ \sum (w \times ES) = 41.82 \]
\[ \sum (w \times ES^2) = 21.24 \]

Q is calculated using these 3 sums:

\[ Q = \frac{\sum (w \times ES^2) - \left( \frac{\sum (w \times ES)^2}{\sum w} \right) \frac{\sum w}{\sum (w \times ES)}}{2 \sum w} = \frac{21.24 - 41.82^2}{2 \times 269.96} = 21.24 - 6.48 = 14.76 \]

Interpreting Q

- Q is distributed as a Chi-Square
- df = number of ESs - 1
- Running example has 10 ESs, therefore, df = 9
- Critical Value for a Chi-Square with df = 9 and p = .05 is 16.92
- Since our Calculated Q (14.76) is less than 16.92, we fail to reject the null hypothesis of homogeneity.
- Thus, the variability across effect sizes does not exceed what would be expected based on sampling error.

**Heterogeneous Distributions: What Now?**

- Analyze excess between study (ES) variability
  - categorical variables with the analog to one-way ANOVA
  - continuous variables and/or multiple variables with weighted multiple regression
- Assume variability is random and fit a random effects model.

**Random-Effects Meta-Analysis**

- Random-effects models should be familiar from HLM.
- With SPSS, the easiest solution is to use David Wilson’s SPSS macros:
  - [http://mason.gmu.edu/~dwilsonb/ma.html](http://mason.gmu.edu/~dwilsonb/ma.html)
- xexc
### Meta-Analytic Results

#### Distribution Description

<table>
<thead>
<tr>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>26,000</td>
<td>-0.0010</td>
<td>1.3046</td>
<td>1.3046</td>
</tr>
</tbody>
</table>

#### Fixed & Random Effects Model

<table>
<thead>
<tr>
<th>Mean ES</th>
<th>95% CI</th>
<th>SE</th>
<th>Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>1.2046</td>
<td>1.3046</td>
<td>1.3046</td>
<td>1.3046</td>
</tr>
</tbody>
</table>

#### Random Effects Variance Component

<table>
<thead>
<tr>
<th>v</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>1.2046</td>
<td>1.3046</td>
</tr>
</tbody>
</table>

### Homogeneity Analysis

<table>
<thead>
<tr>
<th>Q</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>238.6390</td>
<td>25.0000</td>
<td>.0000</td>
</tr>
</tbody>
</table>

### Random effects v estimated via noniterative method of moments.

#### Regression Coefficients

- **Total**: 19.8337
- **Residual**: 10.1947
- **Model**: 9.6391

### Fixed & Random Effects Model

#### Distribution Description

<table>
<thead>
<tr>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>26,000</td>
<td>-0.0010</td>
<td>1.3046</td>
<td>1.3046</td>
</tr>
</tbody>
</table>

#### Fixed & Random Effects Model

<table>
<thead>
<tr>
<th>Mean ES</th>
<th>95% CI</th>
<th>SE</th>
<th>Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>1.2046</td>
<td>1.3046</td>
<td>1.3046</td>
<td>1.3046</td>
</tr>
</tbody>
</table>

#### Random Effects Variance Component

<table>
<thead>
<tr>
<th>v</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>1.2046</td>
<td>1.3046</td>
</tr>
</tbody>
</table>

### Homogeneity Analysis

<table>
<thead>
<tr>
<th>Q</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>238.6390</td>
<td>25.0000</td>
<td>.0000</td>
</tr>
</tbody>
</table>

### Random effects v estimated via noniterative method of moments.

### Version 2005.05.23

The predictors explained 48.6% of the variance in the ES.

### Version 2005.05.23

This is a random effects regression model, predicting the ES with the % of the sample that was white and the % of the sample that was male.
Run MATRIX procedure:
Version 2005.05.23

***** Inverse Variance Weighted Regression *****
***** Random Intercept, Fixed Slopes Model *****

------- Descriptives -------
Mean ES  R-Square  k
1.5491  .4860  11.0000

------- Homogeneity Analysis -------
<table>
<thead>
<tr>
<th>Q</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residual</td>
<td>10.187</td>
<td>8.000</td>
</tr>
</tbody>
</table>

------- Regression Coefficients -------
<table>
<thead>
<tr>
<th>B</th>
<th>SE</th>
<th>-95% CI</th>
<th>+95% CI</th>
<th>Z</th>
<th>P</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>1.1669</td>
<td>.1770</td>
<td>.8199</td>
<td>1.5139</td>
<td>6.5910</td>
<td>.0000</td>
</tr>
<tr>
<td>white</td>
<td>.3242</td>
<td>.1056</td>
<td>.1172</td>
<td>.5312</td>
<td>3.0691</td>
<td>.0021</td>
</tr>
<tr>
<td>male</td>
<td>.4990</td>
<td>.3816</td>
<td>-.2490</td>
<td>1.2470</td>
<td>1.3076</td>
<td>.1910</td>
</tr>
</tbody>
</table>

------- Maximum Likelihood Random Effects Variance Component -------
v = .00746
se(v) = .00437

------ END MATRIX ------

The remaining variance in ES after the predictors are taken into account is not statistically significantly different from 0. This suggests that the remaining variance is homogeneous.

% white is a statistically significant predictor of ES, with larger ESs coming from samples with higher proportions of white participants.

% male is not a statistically significant predictor of ES. The effect does not appear to differ by gender.

Path Analysis

- More and more, statistical analyses are moving beyond simple hypothesis testing (which doesn’t really tell us anything we don’t already know) to model-building.
- A theoretically-derived model is of much greater interest than a simple test of relationship.
- Path analysis is an extension of multiple regression, and is used to test theoretical models.
Path analysis (and factor analysis) form the basis for structural equation modeling.

- Path analysis uses a series of regression equations to identify a path model (a diagram showing the relationships between variables)
- Think of path models as including a series of independent, mediating, and dependent variables

An exogenous variable has no direct causes (no arrows drawn to them – independent variables)

- An endogenous variable does have a direct cause (mediators or dependent variables)
- Causal paths are shown with single-ended arrows (remember, this represents a causal hypothesis, and does not truly demonstrate causation)
- Correlations between exogenous variables are shown with double-ended arrows.

Path Diagram Rules
- All endogenous variables are assumed to have error, or disturbance terms (representing measurement error and variables not included in the model)
- Paths cannot be recursive

- Path coefficients are just standardized regression weights (beta weights)

Originally, path analyses used Ordinary Least Squares as the method of estimation (using standard multiple regression).

- Now, path analyses typically use Maximum Likelihood.
- Initially, we’ll discuss path analysis from a regression framework, so you can see what’s happening in the procedure.
- The minimum recommended sample size is 10 or 20 times the number of measured variables

Path coefficients are estimated with a series of regression equations.

- Each endogenous variable is used as a dependent variable in its own regression equation
  - any variables (endogenous or exogenous) contributing to an endogenous variable are included simultaneously as predictors
  - The resulting beta-weights (and their associated tests of statistical significance) are used to identify the paths in the path diagram
A direct effect is the effect of one variable on another.

An indirect effect is the mediated effect of one variable on another.

Compound paths are equal to the product of the related direct paths.

Sobel tests (and other techniques for testing mediation) can be used to test the statistical significance of indirect effects.

Example of a path analysis diagram from IARR presentation.

Study 2 and 3 Procedure

Participants completed online surveys

Using free response, described conjoint activities that made them feel good, bad, and neither good nor bad about their relationship

Participants then rated each respective activity on a variety of measures:

- Perceived challenge
- Required skill
- Affect
- Activation
- Relationship satisfaction

Studies 2 and 3

Past research suggests that engaging in challenging and self-expanding activities results in positive affect, which in turn becomes associated with the relationship (Graham, 2008; Lewandowski & Ams, 2003, 2004)

What mediating role do affect and activation play in how challenge and skill impacts relationship quality?

Sample

<table>
<thead>
<tr>
<th></th>
<th>Study 1: Internet Sample</th>
<th>Study 2: Undergraduate Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Participants</td>
<td>289 individuals (80.6% female); 774 activities</td>
<td>232 individuals (66.5% female); 694 activities</td>
</tr>
<tr>
<td>Relationship status</td>
<td>21.9% married, 6.9% engaged, 1.5% legal domestic partners, 41.9% exhibiting, 21.2% cohabiting</td>
<td>21.8% married, 2.6% engaged, 0% legal domestic partners, 6% cohabiting</td>
</tr>
<tr>
<td>Average Relationship Length</td>
<td>4.3 years (range = 0 months to 39.08 years)</td>
<td>4.7 years (range = 0 months to 8.16 years)</td>
</tr>
<tr>
<td>Average Age</td>
<td>27 years</td>
<td>20 years</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>73.3% White</td>
<td>73.4% White</td>
</tr>
<tr>
<td>Level of Education</td>
<td>49.6% some college</td>
<td>100% some college</td>
</tr>
</tbody>
</table>
Study 2: Challenge and Relationship Quality

- Considered alone, challenge was negatively associated with relationship quality:
  - $p < .05$, **$p < .01$, ***$p < .001$. 

Study 2: Challenge, Challenge-Skill and Relationship Quality

- Within the bounds of skill, challenge is positively associated with relationship quality:
  - $p < .05$, **$p < .01$, ***$p < .001$. 

Full Model

- * $p < .05$, ** $p < .01$, *** $p < .001$. 

Study 2: Solved Path Model

- * $p < .05$, ** $p < .01$, *** $p < .001$. 

Study 3:

- * $p < .05$, ** $p < .01$, *** $p < .001$. 

Relationship Quality

- $\beta = .326^{***}$

- $\beta = .386^{***}$

- $\beta = .386^{***}$

- $\beta = .386^{***}$
Path Analysis with AMOS

- If you’d rather use maximum likelihood as the method of estimation, many SEM programs can be used.

- We’ll be using AMOS,

- AMOS also provides fit indices, measures of how closely the correlation matrix described by the path model matches the actual correlation matrix.

- There are a wide variety of fit indices, each with their own strengths and weaknesses.

- Chi-Square – A statistically non-significant chi-square indicates good overall fit, but chi-square is most useful as a measure of relative fit (with a chi-square difference test)

- RMSEA < .08 indicates good fit

- CFI, GFI, or AGFI > .9 indicates good fit
Factor Analysis

- Factor Analysis
  - Analyzes the correlations between variables
  - Creates latent variables (factors) based on the variance shared between variables
- FA can be used to
  - evaluate score validity
  - Develop theory re: the nature of constructs
  - Summarize information about relationships into a parsimonious form that can be used for subsequent analyses

Exploratory Factor Analysis
- The researcher lets the data (partially) dictate the factors that are created

Confirmatory Factor Analysis
- The researcher dictates the factors and tests the data's fit to the theoretical model
- Both EFA and CFA exist along a continuum
- CFA will be covered in SEM

Which item responses group together into which subscales (factors)?
- Which underlying constructs are a group of measures tapping into?
- How many groups of items are there?
- What are the groups of items?

The math of EFA relies on matrix algebra.
- The Thompson book has an accessible introduction.
- For our purposes, we will rely on a conceptual, rather than mathematical understanding of Factor Analysis

Process of Factor Analysis

- While factors are extracted simultaneously, many often conceptualize the process of FA as sequential.
  1. The correlations between all items are considered
  2. The largest area of overlap is “extracted” – this comprises the first factor.
  3. The remaining (residual) variance is examined
4. The next largest area of overlap is extracted – this is the second factor
5. The remaining (residual variance) is examined
6. Etc … this process repeats until the number of factors = the number of variables.

- **Factors**
  - Factors are the latent variables of interest, created from a combination of measured variables
  - Every individual has a **factor score** on each factor
  - Regression equivalent \( \hat{y} \)

- Because factors are extracted from the residuals (what is left after the other factors are removed), they are **orthogonal** (uncorrelated) to one another.

- **Pattern Coefficients**
  - Pattern coefficients are the standardized weights applied to the measured variables in order to create the factors
  - Regression equivalent = Beta Weights

- **Structure Coefficients**
  - Structure coefficients are the correlations between the items and the factors
  - Regression equivalent = structure coefficients, \( r_{ij} \)
  - When the factors are initially extracted, the pattern and structure coefficients are identical!

- **Communality Coefficient** (adding the squared structure coefficients across the rows)
  - The amount of variance in a measured variable that a factor set can reproduce
  - How much of the variance in the measured variable was useful in creating the factor set.
    - Communality of 0 means that variable is not at all represented by the factors
    - Communality of 1 means that all of the variable is reproduced by the factors
    - Communality > 1 is called a “Heywood” case, and is statistically inadmissible

- **Eigenvalues** (adding the squared structure coefficients down the columns; aka, characteristic roots)
  - The number of eigenvalues is equal to the number of measured variables
  - The sum of the eigenvalues is equal to the number of measured variables
  - Eigenvalue/\# of measured variables = the % of variance in a matrix that a given factor reproduces
  - Sum of eigenvalues/\# of measured variables = the % of variance that is reproduced by all of the factors combines

- The book uses an example where 7 students rate how handsome, beautiful, ugly, brilliant, smart, and dumb a professor is.
  - The resulting correlation matrix is:
    
    |      | 1 | 2 | 3 | 4 | 5 | 6 |
    |------|---|---|---|---|---|---|
    | Handsome | 1 |   |   |   |   |   |
    | Beautiful | 1 | 1 |   |   |   |   |
    | Ugly | -1 | -1 | 1 |
    | Brilliant | 0 | 0 | 0 | 1 |
    | Smart | 0 | 0 | 0 | 1 | 1 |
    | Dumb | 0 | 0 | 0 | -1 | -1 | 1 |
The items pertaining to attractiveness are perfectly correlated with one another.

The items pertaining to intelligence are perfectly correlated with one another.

100% of each measured item is reproduced in the factors.

Factor 1 reproduces $\frac{2}{6} = 33.33\%$ of the variance in the correlation matrix.
Factor Analysis

- Are the data appropriate for factor analysis?
  - Sample Size
  - Bartlett’s test of sphericity
  - KMO Measure of Sampling Adequacy

- Factor Analysis is a family of associated statistical procedures, that produce slightly different results depending on the analytic choices made
- “Factor the data by several different analytic procedures and hold sacred only those factors that appear across all the procedures used” (Gorsuch, 1983, p. 330)

Are the data appropriate for factor analysis?

- Factor structures are generally replicable if:
  - $N = 300$
  - $N > 150$ and factors are defined with 10 or more structure coefficients around $|.4|$
  - $N > 0$ and factors are defined by 4 or more structure coefficients greater than $|.6|$
Is the strength of the correlations between measured variables strong enough to warrant factor analysis?

For example, the following matrix is an identity matrix – if factor analyzed, 5 factors (1 for each item) would be produced:

```
   X1  X2  X3  X4  X5
X1  1.00 0.00 0.00 0.00 0.00
X2  1.00 0.00 0.00 0.00
X3  1.00 0.00 0.00
X4  1.00 0.00
X5  1.00
```

**Bartlett's test of sphericity** tests the null hypothesis that the correlation matrix comes from a population in which the variables are noncollinear (i.e., an identity matrix)

- If statistically significant, the data has enough variance for factor analysis
- If statistically non-significant, there is not enough shared variance to factor analyze

The Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) uses partial correlations to determine whether variables are measuring common factors

- KMO ranges from 0 to 1, with higher values indicating more common variance
- A high KMO means that the extracted factors will account for a large amount of the variance in the measured items

The KMOs for each individual item are found on the diagonal of the anti-image correlation matrix.

- If the overall KMO is poor, consider dropping items with low KMOs
- If you find items with overall low individual KMOs, consider dropping them even if the overall KMO is satisfactory
- KMOs are context-dependent and need to be re-calculated each time you drop an item.

**K, M, & O's descriptors**

<table>
<thead>
<tr>
<th>KMO Value</th>
<th>Degree of Common Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.90 to 1.00</td>
<td>Marvelous</td>
</tr>
<tr>
<td>0.80 to 0.89</td>
<td>Meritorious</td>
</tr>
<tr>
<td>0.70 to 0.79</td>
<td>Middling</td>
</tr>
<tr>
<td>0.60 to 0.69</td>
<td>Mediocre</td>
</tr>
<tr>
<td>0.50 to 0.59</td>
<td>Miserable</td>
</tr>
<tr>
<td>0.00 to 0.49</td>
<td>Don't Factor</td>
</tr>
</tbody>
</table>

**Factor Analysis**

- Are the data appropriate for factor analysis?
- **Which matrix of association coefficients should be analyzed?**
- Which methods should be used to extract the factors?
- How many factors should be extracted?
- How should the factors be rotated?
- How should factor scores be computed (if you care?)
Factor Analysis

- Are the data appropriate for factor analysis?
- Which matrix of association coefficients should be analyzed?
- Which methods should be used to extract the factors?
- How many factors should be extracted?
- How should the factors be rotated?
- How should factor scores be computed (if you care)?

Principal Components Analysis
- Default in most statistical packages
- Treats measured variables as perfectly reliable
- Uses 1s on the diagonal of the correlation matrix
- Attempts to reproduce the information in the sample data

Principal Axis Factoring
- Considers the unreliability of measured variables
- Uses communality coefficients instead of 1s on the diagonals of the correlation matrix
- Uses an iterative process
- Uses the PCA communality coefficients, calculates the new communalities, subsums them into the matrix, re-iterates, etc.
- MAKE SURE THE SOLUTION CONVERGED!!!!!!

Alpha Factor Analysis: Creates factors with maximum reliability.
Maximum Likelihood Analysis: Creates factors that reproduce the population matrix
Image Factor Analysis: Creates factors that minimize factors comprised of a single item
Canonical Factor Analysis: Maximizes the relation between the factors and measured variables

- Pearson’s correlation matrix
  - This is the default for SPSS
  - Variables must be intervally scaled
  - Do the two variables order people in the same way, and place them the same distance away from one another
- Spearman’s rho matrix
  - Variables can be intervally or ordinally scaled
  - Do the two variables order people in the same way?
- Covariance matrix
  - Often used in CFA, less often in EFA
  - Considers the correlation, and the variance in both variables
Determining the Number of Factors

- **Bartlett’s test:** Determines whether the matrix of associations is an identity matrix.
  - Used to determine whether the residual matrix after each factor is extracted is an identity matrix.
  - FA generally uses large samples, so tests of statistical significance aren’t that informative.
- **Kaiser Rule:** Factors should have eigenvalues > 1
  - Factors should contain at least as much variance as 1 measured variable.
  - However, keep in mind that eigenvalues have sampling error.

**Scree Plot:** AKA, The Pencil Test - Eigenvalues are plotted by factor.
- The “mountain” is the good factors.
- The “scree” or rubble is the bad factors.
- This is subjective, though some have attempted to make the decision more objective.

---

**Factor Analysis**

- Are the data appropriate for factor analysis?
- Which matrix of association coefficients should be analyzed?
- Which methods should be used to extract the factors?
- How many factors should be extracted?
- **How should the factors be rotated?**
  - How should factor scores be computed (if you care)?

- Each item exists in “factor space”, with their position defined by the pattern coefficients on each factor.
- Consider a 2 factor test – the pattern coefficients define the x:y coordinates of every item in factor space.
This item loads highly on both factors

This item loads highly on the vertical factor, but near 0 on the horizontal factor

It can be difficult to interpret factors when items load equally on each factor.

The interpretation can be simplified by rotating the axes that define factor space.

Now, the factor structures are more clearly defined.

When the factors are initially extracted, they are orthogonal (uncorrelated) to one another.

When we keep the factor axes at 90 degrees from one another, the factors stay uncorrelated with one another.

This is called an orthogonal rotation

We could make the factors even more distinct by breaking the 90-degree angles
In this case, each item loads more distinctly on a single factor – this makes the factors easier to interpret.

An **Oblique rotation** allows the factors to correlate with one another.

Rotations make factors easier to interpret by letting items “load” on a single factor.

However, if the factors are correlated, pattern and structure coefficients are no longer the same thing.

Several rotation options exist.

**Simple Structure:** When each item loads on a single factor.
Orthogonal Rotations

- **Varimax**: Maximizes the differences between squared pattern/structure coefficients on each factor
  - About 85% will yield simple structure
- **Quartimax**: Best when you have a single large factor that is saturated with variables
- **Equamax**: A compromise between the two

Oblique Rotations

- **Promax**: A good 1st choice
  - Starts with varimax
  - Takes varimax coefficients to the nth power to maximise simple structure
  - Uses a procrustean rotation fits the matrix to a target matrix
  - The degree of correlation between factors can be influenced by changing the pivot power

Oblique Rotations

- **Oblimin**: Uses “delta” to control the factor correlations
  - Delta of 0 makes highly correlated factors
  - Large negative values make uncorrelated factors
  - Promax is usually a good choice

Oblique rotations produce orthogonal factors if those are the best fit.

Running a Factor Analysis

- Check that the data is appropriate for factor analysis – Bartlett’s, KMO, etc.
- Drop bad items
- Choose an extraction method
- Determine the number of factors – K1 & scree
- Choose a rotation method
- Run the final factor analysis
  - Oftentimes, folks run FA with multiple techniques and compare the results – a good factor structure should generalize

<table>
<thead>
<tr>
<th>Factor</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>PQY1</td>
<td>0.38</td>
<td>0.37</td>
</tr>
<tr>
<td>PQY2</td>
<td>0.67</td>
<td>0.66</td>
</tr>
<tr>
<td>PQY3</td>
<td>-0.64</td>
<td>0.61</td>
</tr>
<tr>
<td>PQY4</td>
<td>-0.73</td>
<td>0.65</td>
</tr>
<tr>
<td>PQY6</td>
<td>0.79</td>
<td>0.64</td>
</tr>
<tr>
<td>PQY7</td>
<td>0.70</td>
<td>0.54</td>
</tr>
<tr>
<td>PQY8</td>
<td>0.36</td>
<td>0.32</td>
</tr>
<tr>
<td>PQY9</td>
<td>0.30</td>
<td>0.30</td>
</tr>
<tr>
<td>PQY10</td>
<td>0.27</td>
<td>0.37</td>
</tr>
</tbody>
</table>

**Factor 1**: Worry, strain, overcome difficulties, depressed, low confidence

<table>
<thead>
<tr>
<th>Factor</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>PQQ1</td>
<td>0.38</td>
<td>0.37</td>
</tr>
<tr>
<td>PQQ2</td>
<td>0.67</td>
<td>0.66</td>
</tr>
<tr>
<td>PQQ3</td>
<td>-0.64</td>
<td>0.61</td>
</tr>
<tr>
<td>PQQ4</td>
<td>-0.73</td>
<td>0.65</td>
</tr>
<tr>
<td>PQQ6</td>
<td>0.79</td>
<td>0.64</td>
</tr>
<tr>
<td>PQQ7</td>
<td>0.70</td>
<td>0.54</td>
</tr>
<tr>
<td>PQQ8</td>
<td>0.36</td>
<td>0.32</td>
</tr>
<tr>
<td>PQQ9</td>
<td>0.30</td>
<td>0.30</td>
</tr>
<tr>
<td>PQQ10</td>
<td>0.27</td>
<td>0.37</td>
</tr>
</tbody>
</table>
Factor Analysis

- Are the data appropriate for factor analysis?
- Which matrix of association coefficients should be analyzed?
- Which methods should be used to extract the factors?
- How many factors should be extracted?
- How should the factors be rotated?
- How should factor scores be computed (if you care)?

- Factor scores can be saved if you want to use them in variables for further analyses
- Regression with factors
- Higher-order factor analysis (Factor analysis of factors)
- **Regression**: Most commonly used
  - Standardizes the items, and combines them with their pattern coefficients to calculate the factor scores.
  - Resulting factor scores are also z-scores
Meta-Analysis

Slides adapted and/or copied from David Wilson’s website: http://mason.gmu.edu/~dwilsonb/m a.html

History of MA

- 1952: Hans J. Eysenck concluded that there were no favorable effects of psychotherapy, starting a raging debate
- 20 years of evaluation research and hundreds of studies failed to resolve the debate
- 1978: To prove Eysenck wrong, Gene V. Glass statistically aggregated the findings of 375 psychotherapy outcome studies
- Glass (and colleague Smith) concluded that psychotherapy did indeed work
- Glass called his method “meta-analysis”

Logic of Meta-Analysis

- Traditional methods of review focus on statistical significance testing
- Significance testing is not well suited to this task
  - Highly dependent on sample size
  - Null finding does not carry the same “weight” as a significant finding
- Meta-analysis changes the focus to the direction and magnitude of the effects across studies
- Isn’t this what we are interested in anyway?
- Direction and magnitude represented by the effect size

Two-Mode techniques

- R – people are rows, variables are columns
- Q – variables are rows, people are columns
- O – variables are rows, occasions are columns
- P – occasions are rows, variables are columns
- T – people are rows, occasions are columns
- S – occasions are rows, participants are columns

Standardized noncentered factor scores:
- Developed when the primary goal is to be able to compare factor scores across groups
- See the Thompson book

Ideas behind meta-analysis predate Glass’ work by several decades
- R. A. Fisher (1944)
  - “When a number of quite independent tests of significance have been made, it sometimes happens that although few or none can be claimed individually as significant, yet the aggregate gives an impression that the probabilities are on the whole lower than would often have been obtained by chance” (p. 99).
- W. G. Cochran (1953)
  - Source of the idea of cumulating probability values
  - Discusses a method of averaging means across independent studies
  - Laid-out much of the statistical foundation that modern meta-analysis is built upon (e.g., inverse variance weighting and homogeneity testing)
Questions for Meta-Analysis

- Meta-analysis is applicable to collections of research that
  - Are empirical, rather than theoretical
  - Produce quantitative results, rather than qualitative findings
  - Examine the same constructs and relationships
  - Have findings that can be configured in a comparable statistical form (e.g., As effect sizes, correlation coefficients, odds-ratios, proportions, etc.)
  - Are “comparable” given the question at hand

“Simple” Meta-Analysis can be used to:
- Central Tendency: Summarize the size of an effect across studies
- Variability: Describe the variability of effect sizes across studies
- Moderators: Explain the variability of effect sizes across studies using characteristics of the study, sample, etc.
- Ultimately, meta-analytic data can be used to create meta-data that can be used in:
  - Factor Analysis
  - Structural Equation Modeling
  - Anything in the General Linear Model

Effect Sizes in MA

- The effect size makes meta-analysis possible
  - It is the “dependent variable”
  - It standardizes findings across studies such that they can be directly compared
- Any standardized index can be an “effect size” (e.g., Standardized mean difference, correlation coefficient, odds-ratio, reliability coefficient) as long as it meets the following
  - Is comparable across studies (generally requires standardization)
  - Represents the magnitude and direction of the relationship of interest
  - Is independent of sample size
  - Different meta-analyses may use different effect size indices

Questions for Meta-Analysis: Pure vs. Conceptual Replication

- You must be able to argue that the collection of studies you are meta-analyzing examine the same relationship.
- This may be at a broad level of abstraction, such as the relationship between criminal justice interventions and recidivism or between school-based prevention programs and problem behavior.
- Alternatively it may be at a narrow level of abstraction and represent pure replications (e.g., the effect of a single manualized intervention program on smoking rates).
- The closer to pure replications your collection of studies, the easier it is to argue comparability

Which Studies to Include?

- It is critical to have an explicit inclusion and exclusion criteria
  - The broader the research domain, the more detailed they tend to become
  - Refine criteria as you interact with the literature
  - Components of a detailed criteria
    - distinguishing features
    - research respondents
    - key variables
    - research methods
    - cultural and linguistic range
    - time frame
    - publication types

Which Studies to Include?: Methodological Quality

- Include or exclude low quality studies?
  - The findings of all studies are potentially in error (methodological quality is a continuum, not a dichotomy)
  - Being too restrictive may restrict ability to generalize
  - Being too inclusive may weaken the confidence that can be placed in the findings
  - Methodological quality is often in the “eye-of-the-beholder”
  - You must strike a balance that is appropriate to your research question
Which Studies to Include?
- The “we only included published studies because they have been peer-reviewed” argument
- Significant findings are more likely to be published than non-significant findings
- Critical to try to identify and retrieve all studies that meet your eligibility criteria
  - Conference proceedings
  - Contacting principal researchers
  - Dissertation abstracts
  - Announcements on list-serve, etc.

Identifying Studies
- Potential sources for identification of documents
  - Computerized bibliographic databases
  - Authors working in the research domain
  - Conference programs
  - Dissertations
  - Review articles
  - Hand searching relevant journal
  - Government reports, bibliographies, clearinghouses

Identifying Studies
- Rapidly changing area
- Get to know your local librarian!
- Searching one or two databases is generally inadequate
- Use “wild cards” (e.g., random? will find random, randomization, and randomize)
- Throw a wide net; filter down with a manual reading of the abstracts

Strengths
- Imposes a discipline on the process of summing up research findings
- Represents findings in a more differentiated and sophisticated manner than conventional reviews
- Capable of finding relationships across studies that are obscured in other approaches
- Protects against over-interpreting differences across studies
- Can handle a large numbers of studies (this would overwhelm traditional approaches to review)

Weaknesses
- Requires a good deal of effort
- Mechanical aspects don’t lend themselves to capturing more qualitative distinctions between studies
- “Apples and oranges” criticism
- Most meta-analyses include “blemished” studies to one degree or another (e.g., a randomized design with attrition)
- Selection bias poses a continual threat
- Negative and null finding studies that you were unable to find
- Outcomes for which there were negative or null findings that were not reported
- Analysis of between study differences is fundamentally correlational

Effect Sizes
- The effect size (ES) makes meta-analysis possible
- The ES encodes the selected research findings on a numeric scale
- There are many different types of ES measures, each suited to different research situations
- Each ES type may also have multiple methods of computation
Effect Sizes

- Standardized mean difference
- Group contrast research
- Treatment groups
- Naturally occurring groups
- Inherently continuous construct
- Odds-ratio
- Group contrast research
- Treatment groups
- Naturally occurring groups
- Inherently dichotomous construct
- Correlation coefficient
- Association between variables research

What makes it an Effect Size?

- The type of ES must be comparable across the collection of studies of interest
- This is generally accomplished through standardization
- Must be able to calculate a standard error for that type of ES
- The standard error is needed to calculate the ES weights, called inverse variance weights (more on this latter)
- All meta-analytic analyses are weighted

The Standardized Mean Difference

$$ES = \frac{\bar{X}_{G2} - \bar{X}_{G1}}{s_{pooled}}$$

$$s_{pooled} = \sqrt{\frac{s_{G1}^2 (n_1 - 1) + s_{G2}^2 (n_2 - 1)}{n_1 + n_2 - 2}}$$

- Represents a standardized group contrast on an inherently continuous measure
- Uses the pooled standard deviation (some situations use control group standard deviation)
- Commonly called “d” or occasionally “g”

The Correlation Coefficient

$$ES = r$$

- Represents the strength of association between two inherently continuous measures
- Generally reported directly as “r” (the Pearson product moment coefficient)

The Odds-Ratio

- The odds-ratio is based on a 2 by 2 contingency table, such as the one below

<table>
<thead>
<tr>
<th></th>
<th>Success</th>
<th>Failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Group</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>Control Group</td>
<td>c</td>
<td>d</td>
</tr>
</tbody>
</table>

$$ES = \frac{ad}{bc}$$

- The Odds-Ratio is the odds of success in the treatment group relative to the odds of success in the control group.
- 1 means success is just as likely in one group as the other.

Unstandardized Effect Size Metric

- If you are synthesizing are research domain that using a common measure across studies, you may wish to use an effect size that is unstandardized, such as a simple mean difference (e.g., dollars expended)
- Multi-site evaluations or evaluation contracted by a single granting agency
Methods of Calculating the Standardized Mean Difference

- The standardized mean difference probably has more methods of calculation than any other effect size type.

\[ ES = \frac{\bar{X}_1 - \bar{X}_2}{s_{pooled}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} \]

Methods of Calculating the Standardized Mean Difference

Algebraically Equivalent Formulas:

- \( ES = t \frac{n_1 + n_2}{n_1 n_2} \) independent t-test
- \( ES = \sqrt{\frac{F(n_1 + n_2)}{n_1 n_2}} \) two-group one-way ANOVA

Exact p-values from a t-test or F-ratio can be converted into t-value and the above formula applied.

Methods of Calculating the Standardized Mean Difference

Close Approximation Based on Continuous Data -- Point-Biserial Correlation. For example, the correlation between treatment/no treatment and outcome measured on a continuous scale.

\[ ES = \frac{2r}{\sqrt{1 - r^2}} \]

Methods of Calculating the Standardized Mean Difference

Estimates of the Denominator of ES --

Pooled Standard Deviation

\[ s_{pooled} = \sqrt{\frac{n_1 - 1}{n_1} + \frac{n_2 - 1}{n_2}} \]

Methods of Calculating the Standardized Mean Difference

Estimates of the Numerator of ES --

- The Mean Difference
- Difference between gain scores
- Difference between covariance adjusted means
- Unstandardized regression coefficient for group membership
Methods of Calculating the Standardized Mean Difference

Estimates of the Denominator of ES --

Pooled Standard Deviation

\[ s_{pooled} = \sqrt{\frac{MS_{between}}{F}} \] one-way ANOVA > 2 groups

\[ MS_{between} = \frac{\sum X_j^2/n_j - (\sum X_j/n)^2}{k-1} \]

Methods of Calculating the Standardized Mean Difference

Estimates of the Denominator of ES --

Pooled Standard Deviation

\[ s_{pooled} = \frac{s_{gain}}{\sqrt{2(1-r)}} \] standard deviation of gain scores, where \( r \) is the correlation between pretest and posttest scores

Methods of Calculating the Standardized Mean Difference

Estimates of the Denominator of ES --

Pooled Standard Deviation

\[ s_{pooled} = \frac{MS_{error} \cdot df_{error} - 1}{1 - r^2} \] ANCOVA, where \( r \) is the correlation between the covariate and the DV

\[ MS_{error} = \frac{\sum X_j^2/n_j - (\sum X_j/n)^2}{k-1} \]

Methods of Calculating the Standardized Mean Difference

Estimates of the Denominator of ES --

Pooled Standard Deviation

\[ s_{pooled} = \sqrt{\frac{SS_B + SS_{AB} + SS_Y}{df_B + df_{AB} + df_Y}} \] A two-way factorial ANOVA where \( B \) is the irrelevant factor and \( AB \) is the interaction between the irrelevant factor and group membership (factor \( A \)).

Methods of Calculating the Standardized Mean Difference

- Get the picture?
- Standardized Mean Difference effect sizes can be calculated from almost anything.
- http://mason.gmu.edu/~dwilsonb/ma.html
Formulas for the Correlation Coefficient

- Results typically reported directly as a correlation
- Any data for which you can calculate a standardized mean difference effect size, you can also calculate a correlation type effect size.

Data to Code Along With the ES

- The effect size
- May want to code the data from which the ES is calculated
- Confidence in ES calculation
- Method of calculation
- Any additional data needed for calculation of the inverse variance weight
- Sample size
- ES specific attrition
- Construct measured
- Point in time when variable measured
- Reliability of measure
- Type of statistical test used
- Anything you may want to use to explain variance in effect size!

The Hierarchical Nature of Meta-Analytic Data

- Meta-analytic data is inherently hierarchical
- Multiple outcomes per study
- Multiple measurement points per study
- Multiple sub-samples per study
- Results in multiple effect sizes per study
- Any specific analysis can only include one effect size per study (or one effect size per sub-sample within a study)
- Analyses almost always are of a subset of coded effect sizes. Data structure needs to allow for the selection and creation of those subsets
- The solution? Nested HLM files!

Coding studies

- Paper Coding
  - include data file variable names on coding form
  - all data along left or right margin cases data entry
  - Coding directly into a computer database
Coding Directly into a Computer Database

- Advantages
  - Avoids additional step of transferring data from paper to computer
  - Easy access to data for data cleanup
  - Database can perform calculations during coding process (e.g., calculation of effect sizes)
  - Faster coding

- Disadvantages
  - Can be time consuming to set up
    - the bigger the meta-analysis the bigger the payoff
  - Requires a higher level of computer skill

GLM

- All general linear modeling procedures:
  - Start with measured variables
  - Create a latent variable with a series of additive and multiplicative weights
  - This latent variable is the focus of the analysis

- All GLM procedures produce:
  - Standardized weights
  - Structure Coefficients
  - R-squared effect sizes

Regression
  - Latent variable: Y-hat
  - Weights: Beta weights
  - Structure Coefficients: structure coefficients
  - Effect Size: $R^2$

ANOVA
  - Latent variable: Main and interaction effects
  - Weights: eta
  - Structure Coefficients: structure coefficients
  - Effect Size: $\eta^2$

Factor Analysis
  - Latent variable: Factors
  - Weights: Pattern coefficients
  - Structure Coefficients: structure coefficients
  - Effect Size: Eigenvalues

Multiple Regression
**Discriminant Analysis**

- Describes the multivariate differences between groups
  - Only one grouping variable, any number of groups
  - Multivariate one-way ANOVA
- Descriptive
  - Describes differences between groups
  - Example: How do sex-offenders who re-offend differ from those that do not on a variety of personality measures?
- Predictive
  - Produces discriminant functions that correctly categorize people into their respective groups
  - Example: How can personality measures best be used to classify individuals based on their risk for recidivism?

**Descriptive Discriminant Analysis**

- Latent variable: Discriminant Functions
- Weights: canonical discriminant function coefficients
- Structure Coefficients: structure coefficients
- Effect Size: Canonical Correlation (Eigenvalue)

**Predictive Discriminant Analysis**

- NOT PART OF THE GLM!!!!!!

---

**Cluster Analysis**

- Kind of like q-technique Factor Analysis, but not quite: http://www.qmethod.org/blog.php
- Examines a set of measured variables, and creates a number of groups of people with common response patterns across measured variables
- Example Question: How many types of prisoners are there in respect to MMPI scores? What personalities define these types?

---

**MANOVA**

- Multivariate ANOVA
  - As ANOVA, but with multiple DVs
  - Unlike DDA, can have multiple IVs and test interactions
  - Tests Omnibus Multivariate Hypotheses
  - Often followed up with multiple univariate ANOVAs to interpret (which takes the original results out of context)
  - Better to follow up with DDAs
- Example question: How do a large group of variables differ between and within genders and relationship types?

**Canonical Correlation Analyses**

- Multivariate Regression
- Predicting multiple DVs with multiple IVs
  - Latent variable: Canonical Functions
  - Weights: canonical function coefficients
  - Structure Coefficients: structure coefficients
  - Effect Size: Canonical Correlation (Eigenvalue)
- Not used much anymore – SEM is more easily customized and interpreted
SPSS performs canonical correlation using the `manova` command.

The `manova` is not in the point-and-click analysis menu, rather it is one of SPSS's hidden abilities.

Used with the `discrim` option, `manova` will compute the canonical correlation analysis.

```
manova dv1 dv2 dv3 with iv1 iv2 iv3 iv4
  / discrim all alpha(1)
  / print=sig(eigen dim) .
```