General Linear Models with More than One Conceptual Predictor

- Topics:
  - Review: specific and general model results
  - Unique effect sizes: standardized slopes; semi-partial (part) and partial versions of correlation and their squared versions
  - Special cases of GLM (and corresponding effect sizes):
    - "Multiple (Linear) Regression" with 2+ quantitative predictors
    - "Analysis of Covariance" (ANCOVA) with both categorical and quantitative predictors—requires joint significance tests and effect sizes
  - Some examples of unexpected results

Review: Specific Info for Fixed Effects

- The role of **each predictor variable** $\boldsymbol{x_i}$ in creating a custom expected **outcome** $\boldsymbol{y_i}$ is described using <u>one or more fixed slopes</u>:
  - **One slope** is sufficient to capture the mean difference between two categories for a **binary** $\boldsymbol{x_i}$ or to capture a **linear effect of a quantitative** $\boldsymbol{x_i}$ (or an exponential-ish curve if $\boldsymbol{x_i}$ is log-transformed)
  - **More than one slope** is needed to capture other **nonlinear effects** of a **quantitative** $\boldsymbol{x_i}$ (e.g., quadratic curves or piecewise spline slopes)
  - $\boldsymbol{C-1}$ **slopes** are needed to capture the mean differences in the outcome across a **categorical predictor** with C categories
    - # possible pairwise mean differences = $\frac{C!}{2!(C-2)!}$, but only $C-1$ are given directly

- For each fixed slope, we obtain an **unstandardized** solution:
  - **Estimate, SE**, $\boldsymbol{t}$-value, $\boldsymbol{p}$-value (in which [Est-0]/SE = t, in which $DF_{num}=1$ and $DF_{den}=N-k$ are used to find the p-value; this is a "Univariate Wald Test" (or a "modified" test given use of t, not z)
  - Effect size found by converting $\boldsymbol{t}$**-value** into **partial** $\boldsymbol{r}$ or **Cohen's** $\boldsymbol{d}$

GLMs with Single Predictors: Review of Fixed Effects

- Predictor $\boldsymbol{x1_i}$ alone: $\boldsymbol{y_i} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1}(\boldsymbol{x1_i}) + \boldsymbol{e_i}$
  - $\boldsymbol{\beta_0}$ = **intercept** = expected $\boldsymbol{y_i}$ when $\boldsymbol{x1_i} = \mathbf{0}$
  - $\boldsymbol{\beta_1}$ = **slope of $\boldsymbol{x1_i}$** = difference in $\boldsymbol{y_i}$ per one-unit difference in $\boldsymbol{x1_i}$
    - Standardized slope for $\boldsymbol{\beta_1}$ = Pearson's r for y_i with $x1_i$ ($\beta_{1std} = r_{y,x1}$)
  - $\boldsymbol{e_i}$ = discrepancy from $y_i - \hat{y}_i$ where $\hat{\boldsymbol{y}}_{\boldsymbol{i}} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1}(\boldsymbol{x1_i})$

- Predictor $\boldsymbol{x2_i}$ alone : $\boldsymbol{y_i} = \boldsymbol{\beta_0} + \boldsymbol{\beta_2}(\boldsymbol{x2_i}) + \boldsymbol{e_i}$
  - $\boldsymbol{\beta_0}$ = **intercept** = expected $\boldsymbol{y_i}$ when $\boldsymbol{x2_i} = \mathbf{0}$
  - $\boldsymbol{\beta_2}$ = **slope of $\boldsymbol{x2_i}$** = difference in $\boldsymbol{y_i}$ per one-unit difference in $\boldsymbol{x2_i}$
    - Standardized slope for $\boldsymbol{\beta_2}$ = Pearson's r for y_i with $x2_i$ ($\beta_{2std} = r_{y,x2}$)
  - $\boldsymbol{e_i}$ = discrepancy from $y_i - \hat{y}_i$ where $\hat{\boldsymbol{y}}_{\boldsymbol{i}} = \boldsymbol{\beta_0} + \boldsymbol{\beta_2}(\boldsymbol{x2_i})$

Review: Joint F-Test of Fixed Effects

- Whether the **set of fixed slopes describing the relation of $\boldsymbol{x_i}$ with $\boldsymbol{y_i}$** significantly explains y_i variance (i.e., if $R^2 > 0$) is tested via a "**Multivariate Wald Test**" (usually a "modified" version with F using denominator DF, or χ^2 otherwise)
  - $F(DF_{num}, DF_{den}) = \frac{SS_{model}/(k-1)}{SS_{residual}/(N-k)} = \frac{(N-k)R^2}{(k-1)(1-R^2)} = \frac{known}{unknown}$
  - ***F* test-statistic** ("F-test") evaluates model R^2 *per slope spent to get to it AND per slope leftover* (F is a weighted ratio of info *known* to info *unknown*)
  - $\boldsymbol{R^2} = \frac{SS_{total} - SS_{residual}}{SS_{total}}$ = square of $\boldsymbol{r}$ of predicted $\boldsymbol{\hat{y}_i}$ with $\boldsymbol{y_i}$
  - $\boldsymbol{R^2}$ = proportion reduction in residual variance relative to empty model
  - $\boldsymbol{R^2_{adj}} = 1 - \frac{(1-R^2)(N-1)}{N-k-1} = 1 - \frac{MS_{residual}}{MS_{total}}$ = correction used for small N

- For GLMs with **only one fixed slope**, the Univariate Wald (t) test for that slope is the same as the Multivariate Wald (F) Test for the model R^2
  - Slope $\boldsymbol{\beta_{unstandardized}}$: $t = \frac{Est(-H_0)}{SE}$, $\boldsymbol{\beta_{standardized}}$ = Pearson r
  - Model: $F = t^2$, $R^2 = r^2$ because predicted $\boldsymbol{\hat{y}_i}$ only uses $\boldsymbol{\beta_{unstd}}$

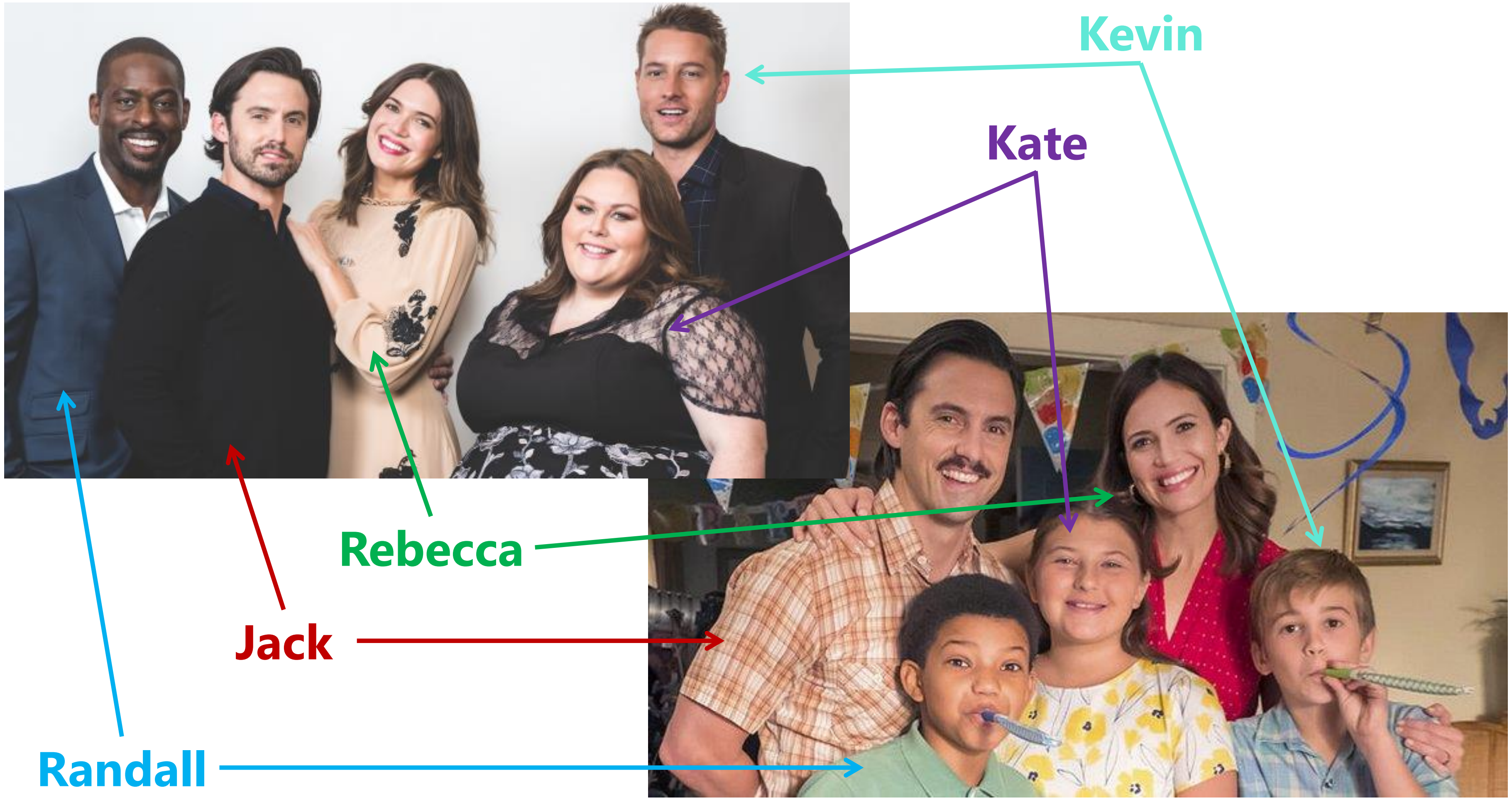
Moving On: GLMs with Multiple Conceptual Predictors

- So far, each **set of fixed slopes** within a separate model have **worked together** to describe the effect of a **single conceptual variable**
  - Thus, the F-test of the model R^2 has reflected the contribution of **one predictor variable conceptually** in forming $\hat{y}_i$, albeit with one or more fixed slopes to capture its relationship to y_i

- Now we will see what happens to the fixed slopes for each variable when combined into a single model that includes **multiple conceptual predictor variables**, each with its own fixed slope(s)
  - Short answer: fixed slopes go from representing "**bivariate**" to "**unique**" relationships (i.e., controlling for the other predictors), and predicted $\hat{y}_i$ is created from all predictors' fixed slopes simultaneously
    - Standardized slopes are no longer equal to bivariate Pearson's r
    - Multiple possible metrics by which to quantify "unique" effect size

A Real-World Analog of "Unique" Effects

- House cleaning with the Pearsons—the cast from "This is Us"



A Real-World Example of "Unique" Effects

- Scenario: Rebecca HAS HAD IT with 3 messy tween-agers and decides to provide an incentive for them to clean the house
  - Let's say the Pearson house has 10 cleanable rooms: 4 bedrooms, 2 bathrooms, 1 living area, 1 kitchen area, 1 dining area, 1 garage
- Generous incentive system for each cleaner (3 children and dad Jack):
  - Individual: one Nintendo game per room cleaned by yourself
  - Family Bonus: if ≥ 8 rooms are clean, the family gets a new TV! (8 = average of 2 rooms per person)
- Rebecca decides to let the family decide what rooms they will each be responsible for while she is shopping for necessities
  - She returns home to a cleaner house, and asks who did what…

Pearson House: Who Cleaned What?

Room	Jack	Kevin	Kate	Randall
Master bedroom	x			
Kevin bedroom		x		
Kate bedroom			x	
Randall bedroom				x
Bathroom 1				x
Bathroom 2				x
Living area		x	x	x
Kitchen area	x			x
Dining area	x			x
Garage (didn't get cleaned)				

- 9/10 rooms are cleaned, so the family gets a new TV—hooray!
- But what should each person receive for their **individual** effort?

Pearson House: Who Cleaned What?

Room	Jack	Kevin	Kate	Randall
Master bedroom	x			
Kevin bedroom		x		
Kate bedroom			x	
Randall bedroom				x
Bathroom 1				x
Bathroom 2				x
Living area		x	x	x
Kitchen area	x			x
Dining area	x			x
Garage (didn't get cleaned)				

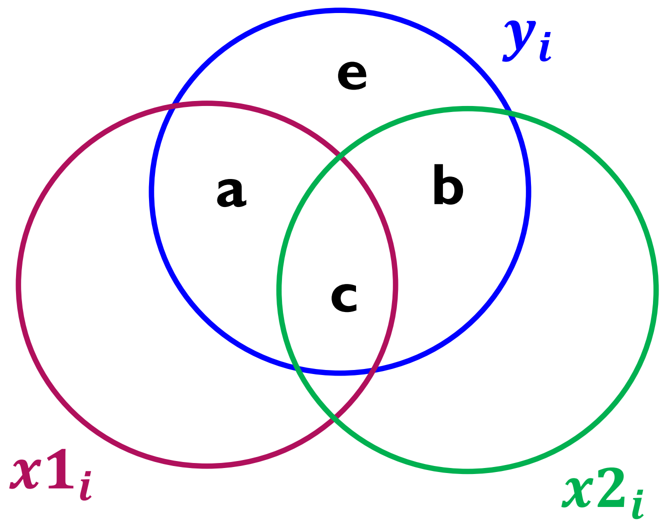
- Jack, Kevin, and Kate: only one Nintendo game each for cleaning one unique room (can't assign rewards for overlapping rooms)
- Randall: three Nintendo games for three unique rooms
- No one gets credit for overlapping rooms (but the family gets the TV)

From Cleaning to Modeling: 2 Goals

1. **General Utility:** Do the model predictors explain a significant amount of variance?
   - Is the model R^2 (the $\boldsymbol{r}^2$ of $\hat{\boldsymbol{y}}_i$ with $\boldsymbol{y}_i$) significantly > 0 (is F-test significant)?
   - **Model R^2 includes both shared AND unique effects of predictor variables: for diagram on right, $R^2 = \frac{a+b+c}{a+b+c+e}$**

2. **Specific Utility:** What is each predictor's **unique** contribution to the model R^2 after discounting (i.e., controlling for) its redundancy with the other predictors?
   - **No predictors get credit for what they have in common** (area $\boldsymbol{c}$ on the right) in predicting $\boldsymbol{y}_i$, even though that shared variance still increases the R^2

Areas below describe partitions of y_i variance:
a = y_i unique to $x1_i$
b = y_i unique to $x2_i$
c = y_i shared by $x1_i$ and $x2_i$
e = y_i leftover (residual)



GLMs with Multiple Predictors: New Interpretation of Fixed Effects

- GLM with 2 predictor variables: $\boldsymbol{y_i} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1}(\boldsymbol{x1_i}) + \boldsymbol{\beta_2}(\boldsymbol{x2_i}) + \boldsymbol{e_i}$
  - $\boldsymbol{\beta_0}$ = **intercept** = expected $\boldsymbol{y_i}$ when $\boldsymbol{x1_i} = \mathbf{0}$ AND when $\boldsymbol{x2_i} = \mathbf{0}$
  - $\boldsymbol{\beta_1}$ = **slope of $\boldsymbol{x1_i}$** = unique difference in $\boldsymbol{y_i}$ per one-unit difference in $\boldsymbol{x1_i}$ "controlling for" or "partialling out" or "holding constant" $\boldsymbol{x2_i}$ (so β_{1std} ≠ Pearson's bivariate $r_{y,x1}$ if $r_{x1,x2} \neq 0$)
    - But $\boldsymbol{\beta_1}$ is still assumed to be constant over all values of $\boldsymbol{x2_i}$ (and $\boldsymbol{x1_i}$)
  - $\boldsymbol{\beta_2}$ = **slope of $\boldsymbol{x2_i}$** = unique difference in $\boldsymbol{y_i}$ per one-unit difference in $\boldsymbol{x2_i}$ "controlling for" or "partialling out" or "holding constant" $\boldsymbol{x1_i}$ (so β_{2std} ≠ Pearson's bivariate $r_{y,x2}$ if $r_{x1,x2} \neq 0$)
    - But $\boldsymbol{\beta_2}$ is still assumed to be constant over all values of $\boldsymbol{x1_i}$ (and $\boldsymbol{x2_i}$)
  - Here $x1_i$ and $x2_i$ have "additive effects" (effect = slope in this context)…
    stay tuned for "multiplicative effects" via interaction terms in unit 5!

Btw: From Pearson Correlations and Covariances to Standardized Slopes

- For a **one-predictor** model: $y_i = \beta_0 + \beta_1(x1_i) + e_i$
  - Unstandardized: $\beta_0 = M_y - (\beta_1 M_{x1}), \beta_1 = r_{y,x1}\frac{SD_y}{SD_{x1}}, \beta_1 = \frac{Cov_{x1,y}}{SD^2_{x1}}$
  - Standardized: $\beta_0 = 0, \beta_{1std} = \beta_1 \frac{SD_{x1}}{SD_y}$ **(so $\beta_{1std} = r_{y,x1}$ here)**
  - Btw, you reported standardized slopes in HW2 with one predictor

- For a **two-predictor** model: $y_i = \beta_0 + \beta_1(x1_i) + \beta_2(x2_i) + e_i$
  - Unstandardized: $\beta_0 = M_y - (\beta_1 M_{x1}) - (\beta_2 M_{x2})$
  - Standardized: $\beta_{1std} = \frac{r_{y,x1} - (r_{y,x2} * r_{x1,x2})}{1 - R^2_{x1,x2}}, \beta_{2std} = \frac{r_{y,x2} - (r_{y,x1} * r_{x1,x2})}{1 - R^2_{x1,x2}}$
  - Standardized to unstandardized: $\beta_1 = \beta_{1std}\frac{SD_y}{SD_{x1}}, \beta_2 = \beta_{2std}\frac{SD_y}{SD_{x2}}$

Where the "Common" Area c Goes

- Model R^2 can be understood in many ways—here, for two slopes:
  - Old: R^2 is the square of the r between predicted $\hat{\boldsymbol{y}}_{\boldsymbol{i}}$ and $\boldsymbol{y}_{\boldsymbol{i}}$
  - Old R^2 said differently: $R^2 = \frac{Var_{\hat{\boldsymbol{y}}_{\boldsymbol{i}}}}{Var_{\boldsymbol{y}_{\boldsymbol{i}}}} = \frac{explained\ variance}{total\ variance}$
  - New: $R^2 = \beta_{1std}^2 + \beta_{2std}^2 + \boxed{\left(2 * \beta_{1std} * \beta_{2std} * r_{x1,x2}\right)}$

- In general: $\boldsymbol{R^2}$ **= unique effects + function of common effects**
  - $\boldsymbol{R^2}$ = general effect size for magnitude of prediction by the model

- The standard errors of each "unique" slope also must be adjusted to reflect the <u>unique</u> variance of its predictor variable relative to the other predictor variables…

Standard Errors of Each Fixed Slope

- Standard Error (SE) for fixed effect estimate βx in a one-predictor model (recall that SE is like the SD of the estimated slope across samples):

$$\mathrm{SE}_{\beta_x} = \sqrt{\frac{\text{residual variance of } y_i}{\mathrm{Var}(x_i) * (N-k)}}$$

N = sample size
k = number of fixed effects

- **When more than one predictor is included, SE turns into:**

$$\mathrm{SE}_{\beta_x} = \sqrt{\frac{\text{residual variance of } y_i}{\mathrm{Var}(x_i) * \left(\mathbf{1} - \boldsymbol{R_x^2}\right) * (N-k)}}$$

R_x^2 = x_i variance accounted for by other predictors, so **$1 - R_x^2$ = unique x_i variance**

- So all things being equal, SE (index of inconsistency) is smaller when:
  - More of the y_i variance has been reduced (with a better predictive model)
    - So fixed slopes can become significant if added later (if R^2 is higher than before)!
  - The predictor has less correlation with other predictors
    - Best case scenario: each x_i is uncorrelated with all other predictors

- If SE is smaller → t-value (or z-value) is bigger→ p-value is smaller

Section Review: Shared Variance

1. In what situation would you expect the amount of variance captured by a predictor to be the same regardless of what other predictors are included?

2. How could it be possible for a model to have a significant R^2 yet have nonsignificant fixed slopes for the individual predictors?

What about "Multicolinearity"? Meh.

- A frequently worried-about problem is "**multicolinearity**" *(see also "multicollinearity" or just "colinearity" or "collinearity")*
- SE for a predictor's slope will be greater to the extent that the predictor has in common (more correlation) with other predictors—makes it **harder to assess its unique effect**
- Diagnostics for this overhyped danger are given in many forms
  - "**tolerance**" = unique predictor variance = $1 - R^2_x$ (< .10 = "bad")
  - "**variance inflation factor**" (VIF) = 1/tolerance (> 10 = "bad")
  - Computers used to have numerical stability problems with high collinearity, but these problems are largely nonexistent nowadays (especially for these GLMs)

- **Only when you have "singularity" is it truly a problem**—when a predictor is a perfect linear combination of the others (redundant)
  - e.g., when including two subscale scores AND their total as predictors
  - e.g., when including intercept + 3 binary predictors for 3 groups
  - You will get a row of dots (or a missing term) instead of results for redundant predictors

Addressing (Multi)Collinearity

- Use the bivariate relationships among your to-be-considered predictors to guide the possibility of "equivalent" models
  - e.g., invasive biological measure vs. highly related but non-invasive alternative measure—can one sufficiently replace the other?

- Such questions require comparing non-nested models
  - **Nested** = one model **is a subset** of other (model A vs. model A+B+C)
    - Btw, it is possible to test nested models using just one model (just not easily in R)
  - **Non-nested** = models are **not subsets** (model A+B vs. model A+C)
    - "Hotelling's t" can be used for significance test of *r* from each model (must save predicted outcome $\hat{y}_i$ for each model and compute their correlation first)
  - See also "dominance analysis" (see Darlington & Hayes, 2016 sec. 8.3)

- Or just try to reduce the slope SEs by adding predictors that are related to y_i but that are (mostly) unrelated to other predictors
  - Less residual variance → smaller SE for each predictor → more power

Metrics of Effect Size per Fixed Slope

- **Unstandardized** fixed slopes cannot inform the relative importance of each predictor because they are **scale-dependent** → differences in what a "unit" is matter

- So we also need to report **some kind of "unique" effect size**
  - Relevant **per fixed slope** (for predictors whose effect on y_i is described by a single slope) and **per conceptual predictor** (for predictors whose effect on y_i require multiple slopes)
  - Why? Beyond putting the slope magnitudes on same scale, specific effect sizes are also used in **meta-analyses** and to **predict power**
  - Choices in $\boldsymbol{r}$ **metric**: standardized slopes (which are not really correlations, see next slide), semi-partial r, or partial r → r gets called $\boldsymbol{\eta}$ ("eta" when using R^2) or $\boldsymbol{\omega}$ ("omega" when adjusted by N, to be used with adjusted R^2)
    - Btw, Cohen's d in a standardized mean difference metric is from the "partial" family
  - Also sometimes used and in $\boldsymbol{R^2}$ metric: Cohen's $\boldsymbol{f^2}$

- *Let's examine more closely how these differ from each other…*

Standardized Slopes: Confusing and Limited

- **Standardized slopes** (solution using z-scored variables, each with $M = 0$ and $SD = 1$) are supposed to describe the **change in $\boldsymbol{y_i}$ per "SD" of $\boldsymbol{x_i}$**
  - Provided in R using lm.beta; can also get by z-scoring all variables, then doing usual GLM (i.e., as implemented in R's lm function by putting scale( ) around each variable)

- Although **standardized slopes** (β_{std}) are often used to index effect size in GLMs and path models, they **are confusing and limited in scope**:
  - They range from ±∞, not ±1 (so they are not correlations), because the SD of original x_i is almost always larger than the SD for "unique" x_i variance
    - Btw, multiplying β_{std} by unique SD of x_i (as $\sqrt{Tolerance}$) = semi-partial r
  - Yield ambiguous results for quadratic or multiplicative terms (z-scored product of 2 variables is not equal to product of two z-scored variables)
  - Differences in sample size across subgroups create different standardized slopes for categorical predictors given the same unstandardized mean difference (see Darlington & Hayes, 2016, sec. 5.1.5 and ch. 8)
  - Do not readily extend to more complex types of prediction models (e.g., generalized linear models, multilevel or "mixed-effects" models)

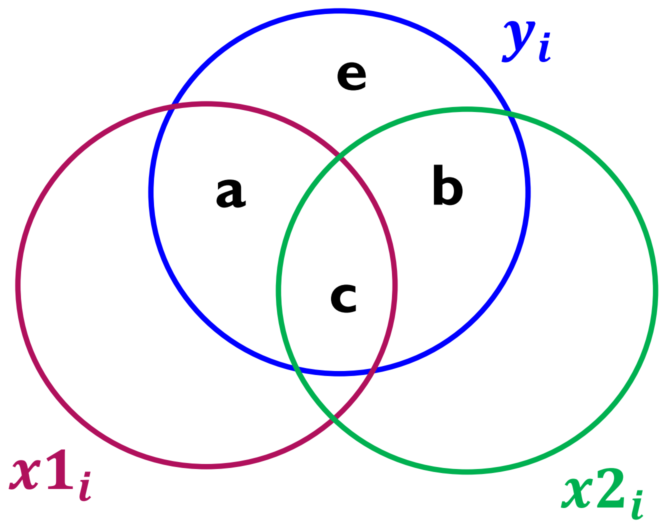
For a helpful blog post on this topic, see: http://www.daviddisabato.com/blog/2016/4/8/on-effect-sizes-in-multiple-regression

Semi-Partial (aka, "Part") Eta-Squared

- Given this GLM: $\boldsymbol{y_i} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1}(\boldsymbol{x1_i}) + \boldsymbol{\beta_2}(\boldsymbol{x2_i}) + \boldsymbol{e_i}$
- For $\boldsymbol{x1_i}$, **semi-partial** $\boldsymbol{\eta^2} = \boldsymbol{sR^2} = \frac{\boldsymbol{SS_{x1}}}{\boldsymbol{SS_{total}}} = \frac{\boldsymbol{a}}{\boldsymbol{a+b+c+e}}$
  - "Unique" sums of squares / total sums of squares: **amount of model $\boldsymbol{R^2}$** that is due to $x1_i$ → directly intuitive ☺
  - Will NOT be influenced by adding extra predictors to the model to explain residual variance → comparability across studies ☺
  - Btw, $\boldsymbol{sr}$ can be found directly from $\boldsymbol{t}$-value:
    - $\boldsymbol{sr = t_{x1}\sqrt{\frac{1-R^2}{DF_{den}}}}$

      SQRT part → prop. *unexplained* variance
    - Btw, there is no semi-partial analog to Cohen's d (b/c group is needed in the model)

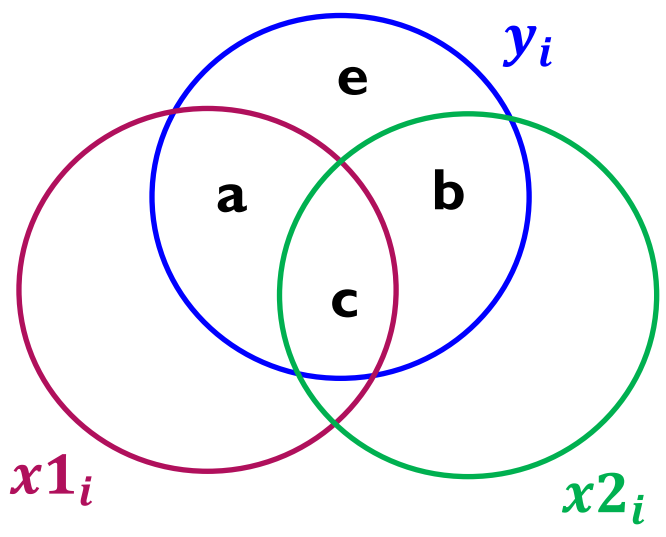
Overall model $\boldsymbol{R^2} = \frac{\boldsymbol{a+b+c}}{\boldsymbol{a+b+c+e}}$



Partial Eta and Eta-Squared

- Given this GLM: $y_i = \beta_0 + \beta_1(x1_i) + \beta_2(x2_i) + e_i$
- For $x1_i$, **partial** $\eta^2 = pR^2 = \frac{SS_{x1}}{SS_{x1}+SS_{residual}} = \frac{a}{a+e}$
  - Unique SS / (unique SS + residual SS) → R^2 **for what's left**
  - WILL BE influenced by adding extra predictors to explain residual variance → lack of comparability across models/studies ☹
  - More useful pr version can also be found from t-value:
    - Partial $\eta = pr = \frac{t}{\sqrt{t^2+DF_{den}}}$
    - Btw, Cohen's d for mean differences in SD units is also in the partial family: $d = \frac{2t}{\sqrt{DF_{den}}}$
  - The word "**partial**" is a synonym for "**unique**"

Overall model $R^2 = \frac{a+b+c}{a+b+c+e}$



Summarizing Effect Sizes (for $x1_i$ here)

- **Semi-partial** $\boldsymbol{\eta^2 = sR^2 = \frac{a}{a+b+c+e}}$
  - Unique / total: amount of model R^2 due to $x1_i$ (directly useful)
- **Partial** $\boldsymbol{\eta^2 = pR^2 = \frac{a}{a+e}}$
  - Unique / (unique+residual): $x1_i$ contribution setting aside $x2_i$
  - Given that it describes a subset of model R^2, pr (or d) version can be less prone to misinterpretation (which is why I use pr)
- Cohen's $\boldsymbol{f^2} = \frac{a}{e} = ?????$
  - But is often used in power analysis!

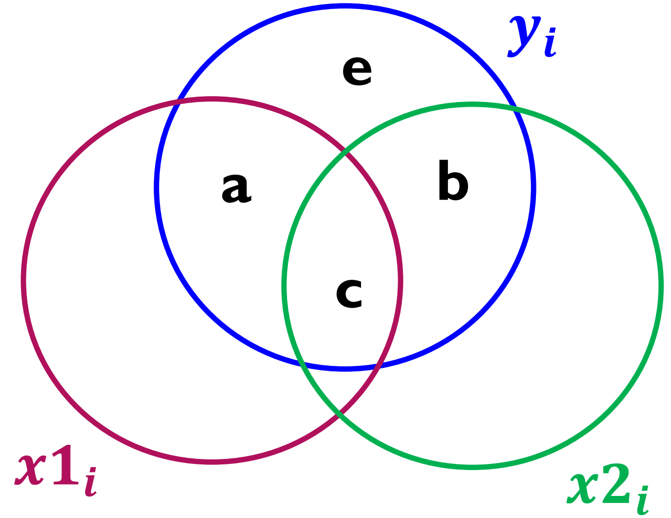
Areas below describe partitions of y_i variance:

a = y_i unique to $x1_i$

b = y_i unique to $x2_i$

c = y_i shared by $x1_i$ and $x2_i$

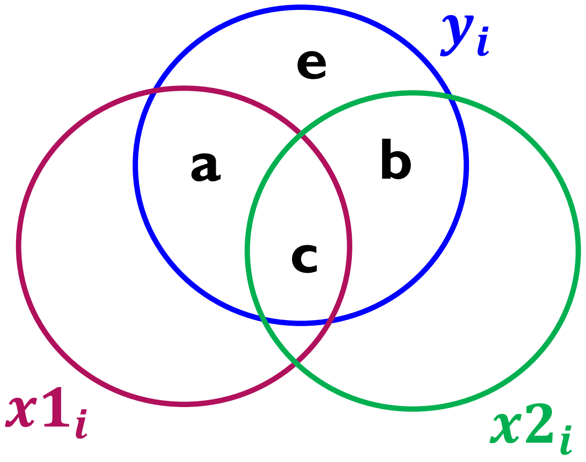
e = y_i leftover (residual)



Model $\boldsymbol{R^2} = \frac{a+b+c}{a+b+c+e}$

Interpreting Effect Sizes with the Pearsons

- Effect sizes for $x1_i$
  - **Semi-partial** $\boldsymbol{\eta^2 = sR^2} = \frac{a}{a+b+c+e} = \frac{unique}{total}$
  - **Partial** $\boldsymbol{\eta^2 = pR^2} = \frac{a}{a+e} = \frac{unique}{unique+residual}$
    - Should not be compared across studies whose models differ in predictor content—here's why:



- Using the Pearsons—of 10 rooms, Randall cleaned 4 rooms, Kevin cleaned 1 room, and Randall and Kevin cleaned 2 common rooms
  - Randall: a = 4, Kevin: b =1, common: c = 2, residual: e = 3 (for this example)
  - Randall: $sR^2 = \frac{4}{4+1+2+3} = .40$, $pR^2 = \frac{4}{4+3} = .57$
    - Randall cleaned 40% of the house, or 57% of the house *that Kevin didn't*
  - Kevin: $sR^2 = \frac{1}{4+1+2+3} = .10$, $pR^2 = \frac{1}{1+3} = .25$
    - Kevin cleaned 10% of the house, or 25% of the house *that Randall didn't*

Section Review: Effect Size

1. What is the primary difference between the partial and semi-partial effect size families?

2. Which family will always yield a larger effect size than the other?

Single-Slope Predictors in "Multiple Linear Regression"

Sum of separate $R^2 = .1986$

- Separate models from Example 2 (in which $R^2 = sR^2 = pR^2$)
  - Empty: $\boldsymbol{income_i} = \boldsymbol{\beta_0}$, $R^2 = 0$
  - Education: $\boldsymbol{income_i} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1}(\boldsymbol{educ_i} - \mathbf{12}) + \boldsymbol{e_i}$, $R^2 = .1480$
  - Marital Status: $\boldsymbol{income_i} = \boldsymbol{\beta_0} + \boldsymbol{\beta_2}(\boldsymbol{marry01_i}) + \boldsymbol{e_i}$, $R^2 = .0506$

- Combined model: $\boldsymbol{income_i} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1}(\boldsymbol{educ_i} - \mathbf{12}) + \boldsymbol{\beta_2}(\boldsymbol{marry01_i}) + \boldsymbol{e_i}$
  - $R^2 = .1903$ for both $<$ sum of separate $R^2 = .1986$ b/c of common
  - Education $\boldsymbol{\beta_1}$: semi-partial $sR^2 = .1396$, partial $pR^2 = .1471$ (t → sig*)
    - Explained 13.96% of income variance (14.71% of variance unexplained by marital)
  - Marital $\boldsymbol{\beta_2}$: semi-partial $sR^2 = .0423$, partial $pR^2 = .0496$ (t → sig*)
    - Explained 4.23% of income variance (4.96% of variance unexplained by educ)

- Significance of effect sizes given directly *per conceptual predictor* (linear education and binary marital status require 1 slope each)

Single-Slope Predictors in Multiple Linear Regression

Combined model: $income_i = \beta_0 + \beta_1(educ_i - 12) + \beta_2(marry01_i) + e_i$

```
print("Model 1: Linear Education and Binary Marital Status")
Model1 = lm(data=Lecture4, formula=income~1+educ12+marry01)
obj=LMsummary(Model1, effectsizes=TRUE)  # Custom output
```

```
 Sums of Squares Table
              SS  DF         MS      F      p    R2
Model  26530.412   2 13265.206 85.894 <0.001 0.190
Error 112892.820 731   154.436
Total 139423.232 733   190.209

 Fixed Effects Table
             Est    SE      t      p    LCI    UCI
Intercept 11.474 0.678 16.935 <0.001 10.144 12.804
educ12     1.774 0.158 11.228 <0.001  1.464  2.084
marry01    5.695 0.922  6.179 <0.001  3.885  7.504

 Effect Sizes for Fixed Effects Table
          Est      p     d    pr   sR2
educ12  1.774 <0.001 0.831 0.384 0.140
marry01 5.695 <0.001 0.457 0.223 0.042
```

sR^2 → education explained 14.0% of the variance while marital status explained 4.2% (with the rest of the model R^2 due to their shared variance with income instead)

Multi-Slope Predictors in Multiple Linear Regression

Combined: $Income_i = \beta_0 + \beta_1 (LvsM_i) + \beta_2(LvsU_i) + \beta_3(Age_i - 18) + \beta_4(Age_i - 18)^2 + \beta_5(LessHS_i) + \beta_6(GradHS_i) + \beta_7(OverHS_i) + e_i$

```
print("Model 2: Three-Category Workclass, Quadratic Age, and Piecewise Education")
Model2 = lm(data=Lecture4, formula=income~1+LvM+LvU+age18+I(age18^2)+lessHS+gradHS+overHS)
obj=LMsummary(Model2, effectsizes=TRUE)  # Custom output
```

```
Sums of Squares Table
              SS  DF        MS      F      p    R2
Model  40246.424   7 5749.489 42.088 <0.001 0.289
Error  99176.808 726  136.607
Total 139423.232 733  190.209
```

The semi-partial R^2 (sR^2) values computed by my function cannot by directly combined across the rows for the same conceptual variable in this table! Here's why…

```
Effect Sizes for Fixed Effects Table
              Est      p      d     pr   sR2
LvM         6.060 <0.001  0.475  0.231 0.040
LvU         7.209  0.008  0.198  0.099 0.007
age18       1.070 <0.001  0.646  0.307 0.074
I(age18^2) -0.018 <0.001 -0.571 -0.275 0.058
lessHS      0.259  0.645  0.034  0.017 0.000
gradHS      3.157  0.073  0.133  0.067 0.003
overHS      1.528 <0.001  0.545  0.263 0.053
```

Multi-Slope Predictors in Multiple Linear Regression

- Combined: $\boldsymbol{Income_i = \beta_0 + \beta_1(LvsM_i) + \beta_2(LvsU_i) + \beta_3(Age_i - 18) + \beta_4(Age_i - 18)^2}$
  $\boldsymbol{+ \beta_5(LessHS_i) + \beta_6(GradHS_i) + \beta_7(OverHS_i) + e_i}$
  - *Btw, this model might also be called "Analysis of Covariance" (or ANCOVA)*

- Semi-partial R^2 (sR^2) effect size per slope is problematic for two conceptual predictors:
  - **Working Class:** slopes $\boldsymbol{\beta_1}$ and $\boldsymbol{\beta_2}$ share a common reference (low group) and imply 3 pairwise group differences (2 in model; 1 given as linear combination; other differences could be requested as needed)
    - So the $\boldsymbol{sR^2}$ values across three possible group differences will sum to more than they should (for a single 3-category predictor)
  - **Age:** Linear age slope $\boldsymbol{\beta_3}$ is specific to centered age = 0, so its $\boldsymbol{sR^2}$ would change if age were centered differently; also, the unique $\boldsymbol{sR^2}$ values for linear and quadratic age cannot be summed directly to create total $\boldsymbol{sR^2}$ for age because of the correlation among the two predictors
  - **Education:** although the unique $\boldsymbol{sR^2}$ values for $\boldsymbol{\beta_5}$, $\boldsymbol{\beta_6}$, and $\boldsymbol{\beta_7}$ are ok to use in this case, they also cannot be summed directly to create total $\boldsymbol{sR^2}$ for education because of the correlation among the three predictors

Effect Sizes for a Set of Slopes

- To get a **valid sR^2 (and F-test) for slopes that work together** to describe the effect of one conceptual predictor, we must obtain their **joint contribution to the model sums of squares** by estimating a nested model without those slopes
- How to compute effect sizes for a set of slopes using contribution to the sums of squares (SS)—done for you using my custom function R2compare (that starts with hierarchical_lm from the lmhelprs R package)
  - Step 1: From the full model, get model SS for the model: SS_{Full}
    From the full model, get residual SS for the model: $SS_{Residual}$
    From the full model, get total SS (as empty model): SS_{Total}
  - Step 2: Get the model SS from a reduced model without the slopes for which you want a joint test: $SS_{Reduced}$
  - Step 3: Compute SS difference b/t models: $SS_{Effect} = SS_{Full} - SS_{Reduced}$
  - Step 4: Compute effect sizes: $sR^2 = \frac{SS_{Effect}}{SS_{Total}}$, $pR^2 = \frac{SS_{Effect}}{SS_{Effect}+SS_{Residual}}$
  - Step 5: Repeat steps 1–4 per set of slopes to be tested
- But sequential models are more common, so HW4 will give you practice using those
  - Then the change in the model R^2 after adding new slopes will directly provide sR^2 for the new slopes (at each step, so these contributions will differ from what they would be in a full simultaneous model)

Effect Size and Significance for a Set of Slopes

```
print("Model 2: Three-Category Workclass, Quadratic Age, and Piecewise Education")
Model2 = lm(data=Lecture4, formula=income~1+LvM+LvU+age18+I(age18^2)+lessHS+gradHS+overHS)
obj=LMsummary(Model2, effectsizes=TRUE)  # Custom output

### How to get F-test and change in R2 for multiple fixed slopes at once ###

# Fit model without fixed slopes of interest (LvM and LvU for workclass here)
Model2NoClass = lm(data=Lecture4, formula=income~1+age18+I(age18^2)+lessHS+gradHS+overHS)
# Get F-test and effect sizes for fixed slopes of interest using custom function
obj=R2compare(ReducedModel=Model2NoClass, FullModel=Model2, PredName="Workclass", explain=TRUE)
```

```
F-Test and R2 Change for Workclass Slopes
 R2reduced R2full R2diff DFnum DFden      F      p   pR2   sR2
     0.246  0.289  0.043     2   726 21.821 <0.001 0.057 0.043

Explanation:
R2reduced and R2full = Reduced and Full Model R-squares, R2diff = Change in Model R-square
DFnum and DFden = Numerator and Denominator Degrees of Freedom for Change in R-square,
F = F test-statistic for change in R-square, p = two-sided p-value
pR2 = Partial R-square, sR2 = Semi-Partial R-square
```

Effect Size and Significance for a Set of Slopes

```
print("Model 2: Three-Category Workclass, Quadratic Age, and Piecewise Education")
Model2 = lm(data=Lecture4, formula=income~1+LvM+LvU+age18+I(age18^2)+lessHS+gradHS+overHS)
obj=LMsummary(Model2, effectsizes=TRUE)  # Custom output

# Repeat for age slopes
Model2NoAge = lm(data=Lecture4, formula=income~1+LvM+LvU+lessHS+gradHS+overHS)
obj=R2compare(ReducedModel=Model2NoAge, FullModel=Model2, PredName="Age")

# Repeat for education slopes
Model2NoEduc = lm(data=Lecture4, formula=income~1+LvM+LvU+age18+I(age18^2))
obj=R2compare(ReducedModel=Model2NoEduc, FullModel=Model2, PredName="Education")
```

```
 F-Test and R2 Change for Age Slopes
  R2reduced R2full R2diff DFnum DFden      F      p   pR2   sR2
      0.208  0.289  0.081     2   726 41.080 <0.001 0.102 0.081


F-Test and R2 Change for Education Slopes
 R2reduced R2full R2diff DFnum DFden      F      p   pR2   sR2
     0.208  0.289  0.081     3   726 27.455 <0.001 0.102 0.081
```

Multi-Slope Predictors in Multiple Linear Regression

- Separate models from Example 3 (here, $R^2 = sR^2 = pR^2$)
  - 3-Category Workclass (2 slopes): $R^2 = .1034$
  - Linear + Quadratic Age (2 slopes): $R^2 = .1139$
  - Piecewise Education (3 slopes): $R^2 = .1643$

Sum of separate $R^2 = .3816$

- Combined: $\boldsymbol{Income_i} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1}(\boldsymbol{LvsM_i}) + \boldsymbol{\beta_2}(\boldsymbol{LvsU_i}) + \boldsymbol{\beta_3}(\boldsymbol{Age_i} - \mathbf{18}) + \boldsymbol{\beta_4}(\boldsymbol{Age_i} - \mathbf{18})^2 + \boldsymbol{\beta_5}(\boldsymbol{LessHS_i}) + \boldsymbol{\beta_6}(\boldsymbol{GradHS_i}) + \boldsymbol{\beta_7}(\boldsymbol{OverHS_i}) + \boldsymbol{e_i}$
  - $R^2 = .2887$ for all < sum of separate $R^2 = .3816$ (b/c of what's in common across predictors)
  - **Workclass** $\boldsymbol{\beta_1}$, $\boldsymbol{\beta_2}$: semi-partial $sR^2 = .0428$, partial $pR^2 = .0567$
    - Explained 4.28% of income variance (or 5.67% of variance unexplained by others)
  - **Age** $\boldsymbol{\beta_3}$, $\boldsymbol{\beta_4}$: semi-partial $sR^2 = .0805$, partial $pR^2{:} = .1017$
    - Explained 8.05% of income variance (or 10.17% of variance unexplained by others)
  - **Education** $\boldsymbol{\beta_5}$, $\boldsymbol{\beta_6}$, $\boldsymbol{\beta_7}$: semi-partial $sR^2 = .0807$, partial $pR^2{:} = .1019$
    - Explained 8.07% of income variance (or 10.19% of variance unexplained by others)

Sequential Example: Testing R^2 vs. Change in R^2

Example Model Fixed Effects	MSE residual variance (leftover)	Model R2 (relative to empty model)	Change in R2 from new slopes = Semipartial r2
1. intercept	200	0.00	
2. intercept + A	180	0.10	0.10
3. intercept + A + B	140	0.30	**0.20**
4. intercept + A + B + C + D	80	0.60	**0.30**

- F-tests assess the significance of a set of multiple slopes
  - F-test for model $\boldsymbol{R^2}$ is given by default (for all slopes in model)
- To assess the **change in the $\boldsymbol{R^2}$** after adding new slopes:
  - 1 slope? Its p-value tests R^2 change directly (e.g., model 2 to 3)
  - 2+ slopes? Must request a **separate $\boldsymbol{F}$-test** for new slopes added
    - e.g., for R^2 change from model 3 to 4, use my custom function R2compare

Recommendation: What to Report for Effect Size

- **I recommend reporting partial effect sizes per slope (r or d), but reporting semi-partial R^2 per conceptual predictor instead**
- Default output for my functions that compute effect sizes for fixed effects (LMsummary) and for linear combinations of fixed effects (glhtSummary):
  - (Partial) Cohen's d: mean difference in SD units (from partial family for "unique")
  - Partial r: correlation of that slope's x_i and y_i controlling *both* for all other predictors to indicate the size of the predictor's "unique" effect in an r metric
    - Why not pR^2? To avoid misinterpretation as contribution to model R^2! But I did compute it (so you could edit the function to print it if you want)
  - Semi-partial R^2: amount of model R^2 that slope is responsible for
    - May not be added directly together for slopes for the same conceptual predictor!
    - I also compute (but not print) semi-partial r that has two other definitions
      - Correlation of that slope's x_i and y_i controlling only that x_i for all other predictors
      - Standardized slope β_{std} multiplied by the unique SD of that x_i
- My custom function R2compute computes and prints both pR^2 and sR^2, in which sR^2 is the same thing as the change in overall R^2 across models

Recommendation: Model-Building Strategies

- **Step 0:** Create new variables out of each conceptual predictor
  - Quantitative: center (subtract a constant) so that 0 is meaningful
  - Categorical: represent group differences using binary (0/1) predictors

- **Step 1**: Examine **bivariate** relations of each conceptual predictor with y_i
  - "Bivariate" = "zero-order" relation for two variables (x_i and y_i)
  - For a quantitative or binary predictor that has a linear relation with y_i, its bivariate relation is given by Pearson correlation r (and a correlation matrix is useful for many variables at once)
    - Square of Pearson r = "shared variance" for x_i and y_i
  - Otherwise, you need a GLM for each conceptual predictor in order to include multiple fixed slopes (e.g., 3+ categories; linear+quadratic slopes)
    - Model R^2 = "shared variance" for x_i and y_i

- **Step 2**: Examine bivariate relations of each conceptual predictor with the other predictors—useful to get a sense of how they will compete with each other when combined in the same model for y_i
  - Via correlation matrices when possible, using models otherwise
  - Quantify shared variance using same process as in step 1

Recommendation: Model-Building Strategies

- **Step 3**: Combine conceptual predictors into the same model in whatever way corresponds to your **research questions**… here are two examples:

- **Simultaneous**: How does y_i relate to $x1_i$, $x2_i$, and $x3_i$?
  - Put all slopes into same model in one step—report model test (F for R^2), as well as direction, significance, and effect size per predictor (with partial or semi-partial effect size options)
  - Then results for significance and effect size are not biased by order of inclusion in the model!

- **"Stepwise" using R^2 change:** (a) After controlling for $x1_i$, how does $x2_i$ predict y_i?
  (b) After controlling for $x1_i$ AND $x2_i$, how does $x3_i$ predict y_i?
  - (a) **Put $x1_i$** into model and report its direction, significance, and effect size. **Add $x2_i$** into model—report model test (F for R^2), change in model test (F for R^2 change if 2+ slopes), as well as $x2_i$ direction (also significance and effect size per slope). Comment on how the slope(s) for $x1_i$ changed after $x2_i$.
  - (b) **Add $x3_i$** into model—report model test (F for R^2), change in model test (F for R^2 change if 2+ slopes), as well as $x3_i$ direction (also significance and effect size). Report how $x1_i$ and $x2_i$ slopes changed after $x3_i$.
- A stepwise strategy is useful if there is a clear hierarchy for the inclusion of predictors (i.e., based on previous findings or your research questions). But if not, a **simultaneous strategy is likely more defensible** because a variable's contribution will differ before vs after adding other predictors!

Unexpected Results: Suppression

- In general, the semi-partial r for each predictor (and its unique standardized slope) will be smaller than the bivariate r (and its standardized slope when by itself) with y_i
- However, this will not always be the case given **suppression**: when the relationship between predictors is hiding (suppressing) their "real" relationship with the outcome
  - Occurs given $r_{y,x1} > 0$ and $r_{y,x2} > 0$ in three conditions: (a) $r_{y,x1} < r_{y,x2} * r_{x1,x2}$, (b) $r_{y,x2} < r_{y,x1} * r_{x1,x2}$, or (c) $\boldsymbol{r_{x1,x2} < 0}$ (**negative correlation among predictors**)
  - For example: Consider y_i = sales success as predicted by $x1_i$= assertiveness and $x2_i$= record-keeping diligence
    - $r_{y,x1} = .403, r_{y,x2} = .127$, and $r_{x1,x2} = -.305$ (so is condition c)
    - Standardized: $\hat{y}_i = 0 + 0.487(x1_i) + 0.275(x2_i)$
    - So these standardized slopes (for the predictors' unique effects) are greater than their bivariate correlations with the outcome!
- This is one of the reasons why you cannot anticipate just from bivariate correlations what will happen in a model with multiple predictors…

Example taken from Cohen, Cohen, Aiken, & West (2003)

Unexpected Results: Multivariate Power

Correlations

		Y	X1	X2	X3	X4	X5
Y	Pearson Correlation	1	.191	.192	.237	.174	.110
	Sig. (2-tailed)	.	.119	.117	.081	.155	.371
	N	68	68	68	68	68	68
X1	Pearson Correlation	.191	1	-.250*	-.077	-.079	-.110
	Sig. (2-tailed)	.119	.	.039	.535	.521	.371
	N	68	68	68	68	68	68
X2	Pearson Correlation	.192	-.250*	1	-.077	.361**	.013
	Sig. (2-tailed)	.117	.039	.	.532	.003	.917
		68	68	68	68	68	68
X3	Pearson Correlation	.237	-.077	-.077	1	.203	.219
	Sig. (2-tailed)	.081	.535	.532	.	.098	.073
	N	68	68	68	68	68	68
X4	Pearson Correlation	.174	-.079	.361**	.203	1	.162
	Sig. (2-tailed)	.155	.521	.003	.098	.	.187
	N	68	68	68	68	68	68
X5	Pearson Correlation	.110	-.110	.013	.219	.162	1
	Sig. (2-tailed)	.371	.371	.917	.073	.187	.
	N	68	68	68	68	68	68

*. Correlation is significant at the 0.05 level (2-tailed).

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-350.742	195.472		-1.794	.078
	X1	3.327	1.376	.290	2.418	.019
	X2	2.485	1.185	.271	2.098	.040
	X3	3.125	1.479	.257	2.112	.039
	X4	.366	1.342	.035	.273	.786
	X5	.844	1.309	.077	.644	.522

Even though these five **predictors have no significant bivariate correlations with $\boldsymbol{y_i}$**, they still combined to create a significant model R^2

$$F(5,62) = 2.77,$$
$$MSE = 27{,}2631.57,$$
$$p = .025, R^2 = .183$$

This is most likely when the predictors have little correlation amongst themselves (and thus can contribute uniquely)

Example originally borrowed from: https://psych.unl.edu/psycrs/statpage/mr_rem.pdf

Unexpected Results: Null Washout

Correlations

		P1	P2	P3	P4	P5	P6	P7	P8	P9
Y	Pearson Correlation	.230	.059	.004	.079	-.100	-.028	-.040	-.007	.013
	Sig. (2-tailed)	.002	.432	.953	.294	.186	.709	.595	.927	.863
	N	177	177	177	177	177	177	177	177	177

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	100.454	17.866		5.623	.000
	P1	.115	.038	.233	3.047	.003
	P2	4.511E-02	.077	.044	.583	.561
	P3	-1.93E-02	.076	-.019	-.254	.800
	P4	7.511E-02	.076	.075	.988	.325
	P5	-9.22E-02	.070	-.099	-1.320	.189
	P6	6.555E-04	.077	.001	.009	.993
	P7	-4.86E-02	.076	-.048	-.640	.523
	P8	-4.13E-02	.073	-.044	-.568	.571
	P9	6.592E-03	.076	.007	.087	.931

Even though P1 has a significant bivariate correlation with y_i and a significant unique effect, the **model R^2 is not significant**—because it measures the average predictor contribution

$$F(9,167) = 1.49,$$
$$MSE = 93.76,$$
$$p = .155, R^2 = .074$$

Example originally borrowed from: https://psych.unl.edu/psycrs/statpage/mr_rem.pdf

Unexpected Results: A Significant Model R^2 with No Significant Predictors???

		Y	P1	P2	P3	P4	P5
Y	Pearson Correlation	1	.298**	.198**	.221**	.221**	.251**
	Sig. (2-tailed)	.	.000	.008	.003	.003	.001
	N	177	177	177	177	177	177
P1	Pearson Correlation	.298**	1	.689**	.712**	.742**	.728**
	Sig. (2-tailed)	.000	.	.000	.000	.000	.000
	N	177	177	177	177	177	177
P2	Pearson Correlation	.198**	.689**	1	.499**	.500**	.520**
	Sig. (2-tailed)	.008	.000	.	.000	.000	.000
	N	177	177	177	177	177	177
P3	Pearson Correlation	.221**	.712**	.499**	1	.471**	.494**
	Sig. (2-tailed)	.003	.000	.000	.	.000	.000
	N	177	177	177	177	177	177
P4	Pearson Correlation	.221**	.742**	.500**	.471**	1	.593**
	Sig. (2-tailed)	.003	.000	.000	.000	.	.000
	N	177	177	177	177	177	177
P5	Pearson Correlation	.251**	.728**	.520**	.494**	.593**	1
	Sig. (2-tailed)	.001	.000	.000	.000	.000	.
	N	177	177	177	177	177	177

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	93.378	1.899		49.184	.000
	P1	.115	.080	.244	1.441	.151
	P2	-1.23E-02	.073	-.017	-.169	.866
	P3	1.555E-02	.076	.022	.206	.837
	P4	-4.41E-03	.077	-.006	-.057	.954
	P5	5.211E-02	.074	.076	.707	.481

This model R^2 is definitely significant:

$F(5{,}171) = 3.455$,
$MSE = 89.85$,
$p = .005, R^2 = .190$

Yet no predictor has a significant unique effect—this is because of their sizeable correlations with each other (and "common" still contributes to R^2)

Example originally borrowed from: https://psych.unl.edu/psycrs/statpage/mr_rem.pdf

GLM with Multiple Predictors: Summary

- **For any GLM with multiple fixed slopes**, we want to know:
  - Do the slopes join to create a model $R^2 > 0$? Check p-value for model F
  - What is the model's effect size? Check $R^2 = (\boldsymbol{r}$ of $\hat{\boldsymbol{y}}_{\boldsymbol{i}}$ with $\boldsymbol{y}_{\boldsymbol{i}})$ squared
  - Is each slope significantly ≠ 0? Check p-value for $t = (Est - H_0)/SE$
  - How large is each slope's "unique" effect? Report partial r (pr) and/or Cohen's d

- To assess the contribution of **multiple fixed slopes for the same conceptual predictor variable:**
  - Conduct a nested model comparison by removing those slopes
  - Report the change in model R^2 (= semi-partial R^2) and its F-test
  - I would not report separate semi-partial R^2 for each slope to avoid confusion

Example Results Table

Table 1

Final Model Results

Fixed Effect	Est	SE	$p <$	Cohen's d	Partial r
Intercept	−3.687	2.005	.001		
Lower vs Middle Class	6.060	0.947	.001	0.475	.231
Lower vs Upper Class	7.209	2.698	.008	0.198	.099
(Middle vs Upper Class)	1.148	2.708	.672	0.031	.016
Linear Age Slope	1.070	0.123	.001	0.646	.307
Quadratic Age Slope	−0.018	0.002	.001	−0.572	−.275
Education: 2 to 11 years	0.259	0.561	.645	0.034	.017
Education: 11 to 12 years	3.157	1.757	.073	0.134	.067
Education: 12 to 20 years	1.528	0.208	.001	0.546	.263

Note: Cohen's d and partial r effect sizes were computed as: $d = \frac{2t}{\sqrt{DF_{den}}}$; $r = \frac{t}{\sqrt{t^2+DF_{den}}}$. Model-implied effects are given in parentheses, computed as linear combinations of the fixed effects.

Example Results (continuing from Lecture 3)

After examining the bivariate contributions of three-category self-reported working class membership, linear and quadratic years of age, and piecewise linear slopes for years of education in separate models, we then estimated a combined model to examine their unique contributions after controlling for the contribution of each other construct. The model including all seven fixed slopes captured a significant amount of variance in annual income, $F(7, 726) = 42.09$, MSE = 136.61, $p < .001$, $R^2 = .289$. Parameter estimates and effect sizes are given in Table 1. Semi-partial eta-squared (η^2) effect sizes and corresponding multivariate Wald F-tests were obtained to evaluate the amount of total variance captured by distinct sets of predictor slopes.

The omnibus unique effect of three-category self-reported working class membership remained significant, $F(2, 726) = 21.83$, MSE = 136.61, $p < .0001$, semi-partial $\eta^2 = .043$. As shown in Table 1, relative to lower-class respondents (the reference group), after controlling for years of age and years of education, annual income was still significantly higher for both middle-class and upper-class respondents (by 6.060 and 7.209 thousand dollars, respectively). Middle-class and upper-class respondents still did not differ significantly in predicted annual income.

Example Results (continued)

The omnibus unique effect of quadratic years of age (centered at 18) also remained significant, $F(2, 726) = 41.08$, MSE = 136.61, $p < .0001$, semi-partial $\eta^2 = .081$. As shown in Table 1, after controlling for self-reported working class and years of education, annual income was expected to be significantly higher by 1.070 thousand dollars per year of age at age 18; this instantaneous linear age slope was predicted to become significantly less positive per year of age by twice the quadratic coefficient of −0.018. As given by the quantity (−1*linear slope) / (2*quadratic slope) + 18, the age of maximum predicted personal income was 48.56 (i.e., the age at which the linear age slope = 0).

The omnibus unique effect of piecewise years of education (centered at 11) also remained significant, $F(3, 726) = 27.46$, MSE = 136.61, $p < .0001$, semi-partial $\eta^2 = .081$. As shown in Table 1, after controlling for self-reported working class and years of age, annual income was expected to be nonsignificantly higher by 0.259 thousand dollars per year of education from 2 to 11 years, to be nonsignificantly higher by 3.157 thousand dollars for those achieving a high school degree, and to be significantly higher by 1.528 thousand dollars per year of additional education past 12 years. Notably, the effect of a high school degree (the difference between 11 and 12 years of education) was no longer significant after controlling for age and self-reported working class membership.

[The rest of the text would need to emphasize why it matters based on your research questions that the predictors had significant unique effects. This is the part that must be customized per research study!]