psyc3010 lecture 10

review of regression topics

last lecture: mediation and moderation analyses
next lecture: within-participants ANOVA

Quiz 2 next Sunday & Monday

preparation
• review Lectures 6, 7, 8, and 9
• organize your notes so you know where to find information
• look at quiz tips + complete the practice questions online
• complete the Practice Quiz on Blackboard

taking the quiz
• opens @ 9am Sunday, closes @ 9pm Monday
• no time restrictions + can return to active quiz
• can submit quiz only once, and must do so by closing date
Assignment 2 due 23 May @ 12 noon

- submission deadline = noon on Monday of Week 12
- submission via Turnitin on Blackboard only. (If late, e-mail to your tutor.)
- tutorials this week (Week 10) provide feedback on Assignment 1 and tips for Assignment 2
- tutorials next week (Week 11) provide consultation

topics for this lecture

- least-squares criterion
- standard error of the estimate
- partial correlations and semi-partial correlations
- SMR: questions, tests, results
- HMR: questions, tests, results
- ANCOVA: questions, tests, results
- meaning of an interaction in MR
- steps in moderated multiple regression
- Indirect effects and mediation

Optional:
- Intro to SEM
error variance in regression

- research often investigates whether scores on one variable can be predicted by scores on another variable(s)
- the **systematic variance** in this case is the **association** or **relationship** between the criterion and the predictor(s)

**IF** there were no error variance in the dataset, every observed score would reflect the systematic relationship between variables

**BUT** due to random or unknown systematic influences, there will be additional unexplained or error variance in the observed scores

least squares criterion

- regression line represents the relationship between a criterion and 1+ predictors
  - estimated based on observed scores in the dataset
- since there is always some error variance in a dataset, observed scores will not fall in a perfect line

**⇒ so how do we decide where to draw the line?**

- aim: accurately capture the systematic association in the data, while limiting the influence of error variance

**least-squares criterion:**
minimise the (squared) distance between each observed score and the regression line
**semi-partial correlation squared**

- scale-free measure of association between two variables, independent of other IVs
- proportion of total variance in DV uniquely accounted for by one IV

![Diagram showing semi-partial correlation squared](image)

\[
\text{spr}^2 \text{ for IV1:} \\
\text{unique IV-DV variance} = A \\
\text{total DV variance} = A + B + C + D
\]

**partial correlation squared**

- scale-free measure of association between two variables, independent of other IVs
- proportion of residual variance in DV (after taking out the effect of the other IVs) uniquely accounted for by IV

![Diagram showing partial correlation squared](image)

\[
\text{pr}^2 \text{ for IV1:} \\
\text{unique IV-DV variance} = A \\
\text{residual DV variance} = A + B
\]
• Standard Multiple Regression (SMR)
• Hierarchical Multiple Regression (HMR)

questions in SMR

- how much of the variance in a criterion variable can be explained by a set of predictor variables?
  - \( R^2 \) = proportion of variance in one variable (i.e. the criterion) that is explained by others (i.e. all the predictors together)
  - assess magnitude of \( R^2 \) with an \( F \) test:
    - is \( R^2 \) significantly different from 0?

- how important is each individual predictor in explaining variance in the criterion?
  - independent predictive effect of each predictor is represented by its slope: \( b \) or \( \beta \)
  - assess magnitude of \( b / \beta \) with a \( t \) test:
    - is \( b / \beta \) significantly different from 0?
SMR: predictors entered simultaneously

- model $R^2$ assessed in 1 step
- $b$ for each IV based on unique contribution

---

a new example

- did you give your mother a card or gift on mother’s day?
- possible predictors:
  - how many hints mom gave you
  - how much you support capitalism
  - how much you love your mom

- **how much variance ($R^2$) can the predictors explain as a set?**
- **what is the relative importance ($b, \beta, pr^2, sr^2$) of each predictor?**
testing the magnitude of $R^2$

<table>
<thead>
<tr>
<th>Model</th>
<th>Sums of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>4346.03</td>
<td>3</td>
<td>1448.68</td>
<td>16.23</td>
<td>.000</td>
</tr>
<tr>
<td>Residual</td>
<td>2321.43</td>
<td>26</td>
<td>89.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>6667.46</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

"The model including hints, support for capitalism, and love accounted for significant variation in gift-giving behaviour, $F(3, 26) = 16.23, p < .001, R^2 = .65.""

standard vs hierarchical multiple regression

- **standard regression**
  - all predictors are entered *simultaneously*
  - each predictor is evaluated in terms of what it adds to prediction beyond that afforded by all others

- **hierarchical regression**
  - predictors are entered *sequentially* in a pre-specified order based on logic and/or theory
  - each predictor is evaluated in terms of what it adds to prediction *at its point of entry* (i.e., independent contribution relative to predictors already in the model; may be assigned variance shared with variables entered in later steps)
  - order of prediction based upon theory
hierarchical multiple regression

- each step adds more IVs
- model $R^2$ assessed in more than 1 step

$b$ at each step based on unique contribution, independent of other IVs in *current and earlier steps*

---

research questions in HMR

- still testing magnitude of $R^2$ and individual $b / \beta$
- order of entry can help address specific questions:
  1. **demonstrate that hypothesised predictor(s) explains significantly more variance than control variable(s)**
     - similar idea as ANCOVA:
       - predictor at step 1 is like the covariate
  2. **build a sequential model according to theory**

- order is *crucial* to outcome and interpretation
- predictors entered *singly* or in *blocks* of > 1
- can test *increment* in prediction at each block: $R^2$ change and $F$ change
steps and models in HMR

R ch.  \( \rightarrow \) criterion

\( R^2 \) ch.  \( \rightarrow \) criteria

F ch.  \( \rightarrow \) criteria

model 1

steps and models in HMR

R ch.  \( \rightarrow \) criterion

\( R^2 \) ch.  \( \rightarrow \) criteria

F ch.  \( \rightarrow \) criteria

model 2
steps and models in HMR
back to our example…

• suppose we want to repeat our Mother’s Day gift-giving study using hierarchical regression

• further suppose our real interest is in the variables of hints from mom and support for capitalism: we want to show that these explain significantly more variance than love for mom:
  – enter love at step 1
  – enter hints and capitalism support at step 2

• model would be assessed sequentially
  – step 1: prediction by love for mom
  – step 2: prediction by hints and capitalism support, above and beyond variance explained by love

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R²</th>
<th>R² adj</th>
<th>change statistics</th>
<th>sig F ch</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R² ch</td>
<td>F ch</td>
</tr>
<tr>
<td>1</td>
<td>.505</td>
<td>.255</td>
<td>.228</td>
<td>.255</td>
<td>9.584</td>
</tr>
<tr>
<td>2</td>
<td>.813</td>
<td>.652</td>
<td>.612</td>
<td>.397</td>
<td>14.836</td>
</tr>
</tbody>
</table>

for model 1:  
$R^2_{ch} = R^2$ because it simply reflects the change from zero
### model summary

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R²</th>
<th>R² adj</th>
<th>change statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R² ch</td>
</tr>
<tr>
<td>1</td>
<td>.505</td>
<td>.255</td>
<td>.228</td>
<td>.255</td>
</tr>
<tr>
<td>2</td>
<td>.813</td>
<td>.652</td>
<td>.612</td>
<td>.397</td>
</tr>
</tbody>
</table>

#### for model 2:
- **R² ch** tells us that including Hints and Capitalism Support increases the amount of variance accounted for by 40%.
- alternatively, **R² ch** tells us that after controlling for Love, Hints and Capitalism Support explain 40% of the variance. **standard MR can’t do that...**
testing the magnitude of $R^2$ ch

<table>
<thead>
<tr>
<th>Model</th>
<th>$R$</th>
<th>$R^2$</th>
<th>$R^2$ adj</th>
<th>change statistics</th>
<th>$R^2$ ch</th>
<th>$F$ ch</th>
<th>df1</th>
<th>df2</th>
<th>sig $F$ ch</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.505</td>
<td>.255</td>
<td>.228</td>
<td></td>
<td>9.584</td>
<td>1</td>
<td>28</td>
<td>.004</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.813</td>
<td>.652</td>
<td>.612</td>
<td>14.836</td>
<td>16.23</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

for model 2:

$F$ ch tells us that this increment in the variance accounted for is significantly different from 0

→ again, standard MR can’t do that...

---

testing the magnitude of $R^2$

Summary Table for Analysis of Regression:

<table>
<thead>
<tr>
<th>Model</th>
<th>Sums of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>$F$</th>
<th>sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Regression</td>
<td>1702.901</td>
<td>1</td>
<td>1702.901</td>
<td>9.584</td>
<td>.004</td>
</tr>
<tr>
<td>Residual</td>
<td>4964.567</td>
<td>28</td>
<td>177.306</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>6667.46</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Regression</td>
<td>4346.03</td>
<td>3</td>
<td>1448.68</td>
<td>16.23</td>
<td>.000</td>
</tr>
<tr>
<td>Residual</td>
<td>2321.43</td>
<td>26</td>
<td>89.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>6667.46</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

for model 1: details are the same as reported in the change statistics section (as the change was relative to zero)
### testing the magnitude of $R^2$

Summary Table for Analysis of Regression:

<table>
<thead>
<tr>
<th>Model</th>
<th>Sums of Squares</th>
<th>Mean Square</th>
<th>F</th>
<th>sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Regression</td>
<td>1702.901</td>
<td>1702.901</td>
<td>9.584</td>
<td>.004</td>
</tr>
<tr>
<td>Residual</td>
<td>4964.567</td>
<td>177.306</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>6667.466</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Regression</td>
<td>4346.03</td>
<td>1448.68</td>
<td>16.23</td>
<td>.000</td>
</tr>
<tr>
<td>Residual</td>
<td>2321.43</td>
<td>89.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>6667.466</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Model 2**: $F$ tests the overall significance of the model (thus, exactly the same as if we had done an SMR with these three predictors)

### testing magnitude of coefficients

<table>
<thead>
<tr>
<th>Model</th>
<th>B</th>
<th>SE</th>
<th>$\beta$</th>
<th>t</th>
<th>sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 constant</td>
<td>-80.233</td>
<td>7.595</td>
<td>7.009</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>LOVE</td>
<td>5.268</td>
<td>1.700</td>
<td>.505</td>
<td>3.009</td>
<td>.004</td>
</tr>
<tr>
<td>2 constant</td>
<td>-95.02</td>
<td>3</td>
<td>1448.68</td>
<td>16.23</td>
<td>.000</td>
</tr>
<tr>
<td>LOVE</td>
<td>1.678</td>
<td>1.437</td>
<td>.16</td>
<td>1.168</td>
<td>.253</td>
</tr>
<tr>
<td>HINTS</td>
<td>.789</td>
<td>.208</td>
<td>.42</td>
<td>3.785</td>
<td>.000</td>
</tr>
<tr>
<td>CAPITALISM</td>
<td>1.453</td>
<td>.484</td>
<td>.46</td>
<td>3.000</td>
<td>.005</td>
</tr>
</tbody>
</table>

**for model 1**:

Coefficient for Love as sole predictor of gift-giving (i.e., the variable included at step 1)
testing magnitude of coefficients

<table>
<thead>
<tr>
<th>Model</th>
<th>B</th>
<th>SE</th>
<th>β</th>
<th>t</th>
<th>sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-80.233</td>
<td>7.595</td>
<td>7.009</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>constant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOVE</td>
<td>5.268</td>
<td>1.700</td>
<td>.505</td>
<td>3.009</td>
<td>.004</td>
</tr>
<tr>
<td>2</td>
<td>-95.02</td>
<td>3</td>
<td>1448.68</td>
<td>16.23</td>
<td>.000</td>
</tr>
<tr>
<td>constant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOVE</td>
<td>1.678</td>
<td>1.437</td>
<td>.16</td>
<td>1.168</td>
<td>.253</td>
</tr>
<tr>
<td>HINTS</td>
<td>.789</td>
<td>.208</td>
<td>.42</td>
<td>3.785</td>
<td>.000</td>
</tr>
<tr>
<td>CAPITALISM</td>
<td>1.453</td>
<td>.484</td>
<td>.46</td>
<td>3.000</td>
<td>.005</td>
</tr>
</tbody>
</table>

for model 2:
identical to the coefficients table we would get in SMR with the three predictors entered simultaneously

ANOVA

→ adjusted treatment means assume that covariate means are the same at each level of the focal IV
→ thus, any differences in the adjusted treatment means can be attributed to the focal IV only

“would groups differ on the DV if they were equivalent on the covariate?”

- **refines error term** by subtracting variation that is predictable from covariate
  - larger adjustment when covariate-DV relationship is strong

- **refines treatment effect** to adjust for any systematic group differences on covariate that existed before experimental treatment
### Between-Subjects Factors

<table>
<thead>
<tr>
<th>Value Label</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>nationality</td>
<td>75</td>
</tr>
<tr>
<td>-1.00 Other</td>
<td></td>
</tr>
<tr>
<td>1.00 Australian</td>
<td>74</td>
</tr>
</tbody>
</table>

### Tests of Between-Subjects Effects

**Dependent Variable:** Intentions to use sun protection behaviour

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5.845</td>
<td>1</td>
<td>5.845</td>
<td>1.997</td>
<td>.161</td>
</tr>
<tr>
<td>att1</td>
<td>16.053</td>
<td>1</td>
<td>16.053</td>
<td>5.484</td>
<td>.022</td>
</tr>
<tr>
<td>nationality</td>
<td>10.160</td>
<td>1</td>
<td>10.160</td>
<td>3.471</td>
<td>.045</td>
</tr>
<tr>
<td>Error</td>
<td>427.342</td>
<td>146</td>
<td>2.927</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>459.400</td>
<td>148</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. R Squared = .381 (Adjusted R Squared = .359)

"In an ANCOVA controlling for attitudes, which covaried significantly with intentions, $F(1, 146)= 5.48, p=.022$, nationality was associated with significant differences in intentions to use sun protection, $F(1, 146) = 3.47, p=.045$" (Report some sort of effect size measure plus means, SD).

- interactions in multiple regression
- steps in Moderated Multiple Regression
interactions in MR

- relationship between a criterion and a predictor varies as a function of a second predictor
- the second predictor is usually called a moderator
- moderator enhances or attenuates the relationship between criterion and predictor
- example:
  * hints → gift-giving, moderated by capitalism support
- moderated regression achieves the same purpose as examination of interactions in factorial ANOVA: effect of X at different levels of Z

questions in moderated regression

1. does the XZ interaction contribute significantly to the prediction of Y?
   in hierarchical regression:
   - Additive direct effects accounted for in 1st block
   - contribution of interaction term assessed in later block → significant R² ch indicates a significant interaction

2. how do we interpret the effect Z has on the X → Y relationship?
   - in ANOVA, we examine the simple effects of IV1 at different levels of IV2
   - similarly, in moderated regression, we examine the simple slopes of X→Y lines at different values of Z
making sense of the interaction

→ *simple slopes help us interpret a significant interaction*

- simple effects in ANOVA: examine effect of Factor A at different levels of Factor B
- simple slopes in MMR: examine IV-DV relationship at different levels of the moderator
- predictors in MMR are continuous – have no levels
- we *select* critical values of the moderator where it is interesting to examine the simple slopes of the association between the predictor and Y
- we use logical grounds, usually +1 and -1 SD of moderator (“high” and “low” levels of Z)

animated graphical representation

![Regression plane with interactive effects, varying b3](image)
another look at simple slopes

we examine the relationship between X and Y at *high* and *low* values of Z
(and typically the values chosen are ± 1SD of Z)
### (3) testing simple slopes – strategy

- our original MMR gave us the slope for OP (X→Y) when motivation (Z) = 0
- 0 is the mean of Z (centered)
- we want SPSS to test the slope of OP (X→Y) at high and low values of motivation (Z)
- we will create two new variables for Z (high and low)
- we will then re-run the MMR regression at each value of Z to get the simple slopes of OP

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Std. Error</th>
<th>Beta</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(Constant)</td>
<td>5.475</td>
<td>.045</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C_OP</td>
<td>.059</td>
<td>.014</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C_MOT</td>
<td>.212</td>
<td>.047</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>(Constant)</td>
<td>5.478</td>
<td>.043</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C_OP</td>
<td>.059</td>
<td>.013</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C_MOT</td>
<td>.229</td>
<td>.046</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C_INT</td>
<td>-.047</td>
<td>.014</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. Dependent Variable: GPA

### (3) testing simple slopes - steps

**first:**
- create two new variables for the moderator: high and low
- formulae: add or subtract 1 S.D. (standard deviation)
  - high level of moderator = \( \text{Mod}_{\text{ABOVE}} = cmod - \text{SD} \)
  - low level of moderator = \( \text{Mod}_{\text{BELOW}} = cmod + \text{SD} \)

**second:**
- yes, it’s counter-intuitive
- create 2 new interaction terms: 1 for each level of moderator (centered IV x centered high moderator; centered IV x centered low moderator)

**third:**
- re-run MMR for each of the new values of moderator (additive effects in Step 1, interaction in Step 2)
- examine slope of predictor X in Step 2 – this is the relationship between X and Y at high / low levels of Z
(3a) test simple slope of OP at high motivation

- center motivation at +1SD and re-run the MMR:
  1. subtract 0.9528 from C_MOT to give C_MOT+
  2. recalculate interaction term (C_OP x C_MOT+) to give INT+
  3. re-run MMR: Step 1 = additive effects, Step 2 = interaction

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td>t</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Constant)</td>
<td>5.592</td>
<td>0.067</td>
<td>0.925</td>
<td>89.253</td>
</tr>
<tr>
<td>C_OP</td>
<td>0.05</td>
<td>0.014</td>
<td>0.019</td>
<td>0.764</td>
</tr>
<tr>
<td>MOTHIGH</td>
<td>0.229</td>
<td>0.046</td>
<td>0.041</td>
<td>4.982</td>
</tr>
<tr>
<td>C_INT_HI</td>
<td>-0.047</td>
<td>0.014</td>
<td>-0.019</td>
<td>-3.421</td>
</tr>
</tbody>
</table>

a. Dependent Variable: GPA

b₁ is now the slope for X → Y at high Z

(3b) test simple slope of OP at low motivation

- center motivation at -1SD and re-run the MMR, i.e.,
  1. subtract -0.9528 from C_MOT to give C_MOT-
  2. recalculate interaction term (C_OP x C_MOT-) to give INT-
  3. re-run MMR: Step 1 = additive effects, Step 2 = interaction

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td>t</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Constant)</td>
<td>5.260</td>
<td>0.061</td>
<td>0.855</td>
<td>85.551</td>
</tr>
<tr>
<td>C_OP</td>
<td>-0.1</td>
<td>0.104</td>
<td>0.119</td>
<td>5.517</td>
</tr>
<tr>
<td>MOTHLOW</td>
<td>0.229</td>
<td>0.046</td>
<td>0.041</td>
<td>4.982</td>
</tr>
<tr>
<td>C_INT_LO</td>
<td>-0.047</td>
<td>0.014</td>
<td>-0.019</td>
<td>-3.421</td>
</tr>
</tbody>
</table>

a. Dependent Variable: GPA

b₁ is now the slope for X → Y at low Z
interpreting the results

- we have gone through 2 steps of a moderated multiple regression:
  - we identified a significant interaction
  - we decomposed the interaction by examining the simple slopes

- so now we have an answer to our question...

- our analysis showed that the relationship between OP and GPA is significant at lower levels of motivation but not higher levels
  - high motivation attenuates or buffers against the effects of poor prior academic performance on current academic performance

---

HMR: Indirect effects & mediation

<table>
<thead>
<tr>
<th>Model</th>
<th>B</th>
<th>SE</th>
<th>β</th>
<th>t</th>
<th>sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 constant</td>
<td>-80.233</td>
<td>7.595</td>
<td>-</td>
<td>7.009</td>
<td>.000</td>
</tr>
<tr>
<td>LOVE</td>
<td>5.268</td>
<td>1.700</td>
<td>-.51</td>
<td>3.009</td>
<td>.004</td>
</tr>
<tr>
<td>2 constant</td>
<td>-95.02</td>
<td>3</td>
<td>1448.68</td>
<td>16.23</td>
<td>.000</td>
</tr>
<tr>
<td>LOVE</td>
<td>1.678</td>
<td>1.437</td>
<td>.16</td>
<td>1.168</td>
<td>.253</td>
</tr>
<tr>
<td>HINTS</td>
<td>.789</td>
<td>.208</td>
<td>.42</td>
<td>3.785</td>
<td>.000</td>
</tr>
<tr>
<td>CAPITALISM</td>
<td>1.453</td>
<td>.484</td>
<td>.46</td>
<td>3.000</td>
<td>.005</td>
</tr>
</tbody>
</table>

In original HMR: reported love as sig in B1 and Hints & Capitalism as sig in B2

In mediation analysis: Interested in change in effect for love when hints & capitalism are controlled
Conditions for mediation in MR

- Strong theory!
  1. IV should predict mediator
  2. IV should predict DV in Block 1
  3. Mediator should predict DV in Block 2 (i.e., when IV = controlled)
  4. Coefficient for IV should decrease to ns (full mediation) or in size but still sig (partial mediation)
  5. Sobel test should be significant or bootstrapping analyses should be significant

Hypothetical write-up focusing only on mediation

Two standard multiple regressions were conducted predicting each of the mediators from the distal independent variable (love) while controlling for the alternative mediator. Love was associated independently with self-reported hints ($\beta = .66, p < .001$) but not with support for capitalism ($\beta = .04, p = .898$). Accordingly, only self-reported hints met the conditions for mediation analyses (Baron & Kenny, 1986).

A hierarchical multiple regression was conducted to predict gift-giving from the distal independent variable (love) and control variables (support for capitalism) in Block 1 and the potential mediator (self-reported hints) in Block 2 ...
Write up cont’d....

The model is depicted in Figure 1. Consistent with expectations, love was a significant predictor of gift-giving in Block 1 ($\beta = .55$, $p = .002$), but declined to non-significance in Block 2 ($\beta = .16$, $p = .364$) when hints was included in the model ($\beta = .44$, $p = .034$). A Sobel test confirmed the indirect effect of love via hints was significant, $z=2.14$, $p=.006$. 

\[ \text{Love} \rightarrow \text{Hints} \rightarrow \text{Gifts} \]
But what if both mediators were associated with DV?

- With multiple IVs and multiple mediators and multiple DVs, mediation analyses get increasingly cumbersome in MR world
- Move to world of structural equation modelling (SEM)
- Cool thing #1 SEM does – allows alternative to “unit-weighted” scales. Instead of averaging (each item = weighted equally), each item counts towards the scale in proportion to its correlation with the other items
Cofig thing #2: Complex mediational models including multiple mediators and causal sequences.

Interpreting SEM results
(the short version)

- Like all regressions, a SEM has two key components – **overall model fit + coefficients for individual paths**

- Model fit is reported in the text. Look for:
  - Chi square test for overall model meant to be ns (but it never is)
  - GFI, CFI, TLI over .9 (ok) or over .95 (good)
  - RMSEA < .08 (ok) or < .05 (good)
  - Larger R² accounted for in DV(s)
  - There are a million fit statistics and no strong convention as to which one is best – journals differ

- With poor model fit, researchers frequently add new paths suggested by “Modification Indices” – data driven not theory-driven – Type 1 error problem

- If allow error terms to correlate fit goes up (often not shown in model)
Direct and indirect effects

- Coefficients for direct effects may be reported in text or simply in figure
  - You want significant coefficients
  - But NB: whether it's a unidirectional arrow in one direction, a bi-directional arrow, or a unidirectional arrow in the other direction is based on theory (causality is usually inferred)

- Indirect effects generally reported in the text or a table. In text “The indirect effect was significant (IE=.xx, UL=.xx, LL = .xx)“.
  - IE is a coefficient for the indirect effect; UL and LL are upper and lower limit confidence intervals. Sig if UL to LL does not include zero. E.g., UL= .15 LL = .03 is sig. UL = .15 LL = -.01 is ns.

- In a full “effects decomposition” researchers report estimates for the total effect (TE), direct effect (DE), and indirect effect (IE) – generally in a Table.

A SEM writeup example excerpts from Barlow, Louis & Hewstone (2009)

- The model provided a good fit to the data, $x^2(70, N = 272) = 136.30, p < .001, x^2/df = 1.95; CFI = 0.98; RMSEA = .059; SRMR = .061$. The results of the present SEM are displayed in Figure 1. A summary of the effects decomposition analysis is shown in Table 1.
Fig 1. Model stresses direct effect coefficients (always) plus R² (not always). Note how the structural model is hidden here. Error terms also hidden.

Table 1. Generally need to look at effects decomposition table along with figure. Note how multiple effects tested at once (exact ps hidden). Type 1 error avoided ostensibly by good model fit based on theory. No real effect sizes.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Cognition of rejection</th>
<th>Intergroup anxiety</th>
<th>Old-fashioned racism</th>
<th>Issue avoidance</th>
<th>Active avoidance</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross-group friendship</td>
<td>−.23**</td>
<td>−.19***</td>
<td>−.24**</td>
<td>−.18**</td>
<td>−.22**</td>
<td>−.23**</td>
</tr>
<tr>
<td>Cognitions of rejection</td>
<td>.58***</td>
<td>.43***</td>
<td>.43***</td>
<td>.31**</td>
<td>.29**</td>
<td>.58***</td>
</tr>
<tr>
<td>Intergroup anxiety</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old-fashioned racism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Issue avoidance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active avoidance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. * p < .05; ** p < .01.
Take home SEM messages

- SEM is a powerful technique that most readers, reviewers, and researchers do not fully understand.
- SEM makes it possible to represent and test complex causal models in a way that hides a lot of the math.
  - Simultaneously awesome and disturbing.
- At its simplest, can be read like MR – tells you (1) overall model fit; (2) relationships of predictors to DVs.
- Conventions still evolving; prone to confusion & misuse.
  - Common sins: (1) good model fit created by allowing error terms to correlate (hidden off figure); (2) dodgy 'item-parcelling' to avoid bad fit in structural model; (3) paths added based on modification indices (data-driven – may not replicate).

In class next week:
- Within Ps ANOVA

In the tutes:
- This week: Feedback on A1 + Tips on A2
- Next week: Consult for A2

Readings:
- Howell
  - chapter 14
- Field
  - Chapter 11