Multivariate Analyses & Programmatic Research

- Re-introduction to Multivariate research
- Re-introduction to Programmatic research
- Factorial designs → “It Depends”
  - Importance of Interactions
  - Application of Factorial Designs in Programmatic research
- Multiple regression → “Unique Contributions”
  - Importance of collinearity & suppressor effects
  - Application of Multiple regression in programmatic research
- Path Analysis → “Direct & Indirect Effects

During the first lecture of this section we talked about the importance of going beyond bivariate research questions and statistical analyses to multivariate questions and analyses.

Why multivariate research designs?  ➔ Multicausality

Multicausality is the idea that behavior is complicated – that any behavior has multiple causes, and so, can be better studied using multivariate research designs with multiple IVs than with bivariate research designs with only a single IV!!!

So, multivariate research can be used to …
- involve multiple IVs in a single study → to get a more complete picture of the interrelationships among the behaviors we are studying
- “check up” on previous results from bivariate research → to see if the results we got “hold up” within a multivariate context
- is “the effect” we found with the bivariate analysis what we “thought it was”?

As we’ve discussed, there are two fundamental questions about multicausality that are asked in multivariate research…

1. Interactions
   - does the effect of an IV upon the DV depend upon the value of a 2nd IV?
   - Studied using Factorial Designs
2. Unique contributions
   - Does an IV tell us something about a DV that other IVs don’t?
   - Studied using Multiple Regression
3. Causal Structures
   - Is a DV an IV for another DV?
   - Behaviors are effects of some things and causes of others
   - Structural Modeling & Path Analysis

Again, all of these are used to because bivariate “snapshots” of complex behaviors can be incomplete & inaccurate!
Using Factorial Designs in Programmatic Research

Adding a 2nd Treatment

Perhaps the most common application of factorial designs is to look at the separate (main) and combined (interaction) effects of two IVs.

Often our research starts with a simple RH: that requires only a simple 2-group BG research design.

Computer     Lecture

Keep in mind that to run this study, we made sure that none of the participants had any other treatments!

Factorial Designs – Separate (Main) and combined (interaction) effects of two treatments

At some point we are likely to use Factorial designs to ask ourselves about how a 2nd IV also relates to the DV.
Using Factorial Designs in Programmatic Research II

"Correcting" Bivariate Studies

Our well sampled, carefully measured, properly analyzed study showed …

\[ \begin{array}{cc}
\text{High} & \text{Low} \\
40 & 40 \\
\end{array} \]

… nothing!

Our well sampled, carefully measured, properly analyzed study showed …

\[ \begin{array}{cc}
\text{Male} & \text{Female} \\
40 & 40 \\
\end{array} \]

… nothing!

Looks like neither IV is related to the DV !!!

Using Factorial Designs in Programmatic Research I

However, when we analyzed the same data including both variables as IVs …

\[ \begin{array}{cc}
\text{Tx1} & \text{Tx2} \\
\text{High} & \text{Low} \\
40 & 40 \\
40 & 40 \\
60 & 20 \\
20 & 60 \\
\end{array} \]

There are treatment effects both for those who are "Low" & those who are "High" – the marginal means are an "aggregation error".

There are High-Low effects both for those in Tx1 & those in TX2 – the marginal means are an "aggregation error."

So, instead of the “neither variable matters” bivariate results, the multivariate result shows that both variables related to the DV and they interact too !!!!!

Using Factorial Designs in Programmatic Research III

Between Groups Factorial Designs – Generalization across Populations, Settings & Tasks

Often our research starts with a simple RH: that requires only a simple 2-group BG research design.

Keep in mind that to run this study, we had to make some choices/selections:

For example:

- population \rightarrow \text{College Students}
- setting \rightarrow \text{Lecture setting}
- stim/task \rightarrow \text{teach Psychology}
Using Factorial Designs in Programmatic Research III

When we've found and replicated an effect, making certain selections, it is important to check whether changing those selections changes the results.

If there is no interaction – if the results "don't depend upon" the population, task/stimulus, setting, etc – we need to know that, so we can apply the results of the study to our theory or practice, confident in their generalizability.

If there is an interaction – if the results "depend upon" the population, task/stimulus, setting, etc – we need to know that, so we can apply the "correct version" of the study to our theory or practice.

At some point we are likely use BG Factorial designs to ask ourselves how well the results will generalize to:

- other populations – college vs. high school
- other settings – lecture vs. laboratory
- other tasks/stimuli – psyc vs. philosophy

Notice that each factorial design includes a replication of the earlier design, which used the TX instructional methods to:
- teach Psychology
- to College Students
- in a Lecture setting

Each factorial design also provides a test of the generalizability of the original findings:
- w/ Philosophy vs. Psychology
- to High School vs. College Students
- in an Online vs. Lecture setting
Using Factorial Designs in Programmatic Research  IV

Mixed Groups Factorial Designs – Do effects “depend upon” length of treatment?

As before, often our research starts with a simple RH: that requires only a simple 2-group BG research design.

Time Course Investigations

In order to run this study we had to select ONE treatment duration (say 16 weeks):
• we assign participants to each condition
• begin treatment of the Tx groups
• treat for 16 weeks and then measured the DV

Mixed Groups Factorial Designs for Time Course Investigations

Using this simple BG design we can “not notice” some important things. A MG Factorial can help explore the time course of the Tx effects.

By using a MG design, with different lengths of Tx as the 2nd IV, we might find different patterns of data that we would give very different interpretations

Using Factorial Designs in Programmatic Research  V

Mixed Groups Factorial Designs – Evaluating Initial Equivalence when Random assignment is not possible

As before, often our research starts with a simple RH: that requires only a simple 2-group BG research design.

Initial Equivalence Investigations

In order to causally interpret the results of this study, we’d have to have initial equivalence
• but we can’t always RA & manipulate the IV
• So what can we do to help interpret the post-treatment differences of the two treatments?
• Answer – compare the groups before treatment too!
Mixed Groups Factorial Designs to evaluate Initial Equivalence

By using a MG design, we can compare the groups pre-treatment and use that information to better evaluate post-treatment group differences (but can’t really infer cause). For which of these would you be more comfortable concluding that \(Tx_1 > Tx_2\)?

<table>
<thead>
<tr>
<th></th>
<th>Pre</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Tx_1)</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td>(Tx_2)</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

Nah – \(Tx_1\), lowered score

Maybe – more increase by \(Tx_1\)

As good as it gets!

Moving on to Multiple Regression …..

We often perform both bivariate (correlation) and multivariate (multiple regression) analyses – because they tell us different things about the relationship between the predictors and the criterion…”

Correlations (and bivariate regression weights) tell us about the “separate” relationships of each predictor with the criterion (ignoring the other predictors)

Multiple regression weights tell us about the relationship between each predictor and the criterion that is unique or independent from the other predictors in the model.

- What does this predictor tell us about the criterion that no other predictors tell us?
- Is “the predictor” we studied with the bivariate analysis “what we thought it was”???

There are 5 patterns of bivariate/multivariate relationship

Simple correlation with the criterion

<table>
<thead>
<tr>
<th>Simple correlation with the criterion</th>
<th>0</th>
<th>+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bivariate relationship and multivariate contribution (to this model) have same sign</td>
<td>“Suppressor effect” – no bivariate relationship but contributes (to this model)</td>
<td>“Suppressor effect” – bivariate relationship &amp; multivariate contribution (to this model) have different signs</td>
</tr>
<tr>
<td>Non-contributing – probably because colinearity with one or more other predictors</td>
<td>Non-contributing – probably because of weak relationship with the criterion</td>
<td>Non-contributing – probably because colinearity with one or more other predictors</td>
</tr>
<tr>
<td>“Suppressor effect” – bivariate relationship &amp; multivariate contribution (to this model) have different signs</td>
<td>“Suppressor effect” – no bivariate relationship but contributes (to this model)</td>
<td>Bivariate relationship and multivariate contribution (to this model) have same sign</td>
</tr>
</tbody>
</table>
Among these 9 outcomes are 5 “kinds” of results

Bivariate relationship and multivariate contribution (to this model) have same sign -- bivariate results “still look ok”

Non-contributing – probably because of weak relationship with the criterion -- bivariate results “still look ok”

Non-contributing – probably because collinearity with one or more other predictors -- variables has no unique information that is unique to the others in the model

“Suppressor variable” – no bivariate relationship but contributes (to this model) -- bivariate results “missed” an important variable

“Suppressor variable” – bivariate relationship & multivariate contribution (to this model) have different signs -- bivariate results “got the relationship wrong”

An Example ….

We discussed the “curious” bivariate finding of a substantial relationship between ice cream sales & violent crimes, say \( r(363) = .32, p < .01 \)

Our hypothesis was that this finding is a result of a third variable “temperature” that is causing both violent crimes and ice cream sales. To test this we might put both temperature and ice cream sales into a multiple regression model to predict violent crimes.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>ice cream sales</th>
<th>temp</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r(p) )</td>
<td>.32 (&lt;.01)</td>
<td>.55 (&lt;.001)</td>
</tr>
<tr>
<td>( b(p) )</td>
<td>.07(.82)</td>
<td>.61(.01)</td>
</tr>
</tbody>
</table>

Ice cream sales has a nice correlation; temperature has a stronger correlation

However, when we put them both into the model, only temperature has a unique contribution. Why??

Here’s the Venn diagram for this multivariate model

Both ice cream sales and temperature are correlated with violent crimes. But notice that they are highly correlated with each other (collinear)!

This collinearity means that part of what each predictor shares with the criterion they also with the other predictor.

The result of this collinearity is…

-- most of what ice cream sales shares with violent crimes it shares with temperature, resulting in a small \( b \) value for ice cream sales

-- however, much of what temperature shares with violent crimes is independent of ice cream sales, resulting in a large \( b \) value for temperature
Another Example ....

Nearly everybody who looks for it finds a relationship between “practice” and “performance”.

For example, in a recent semester the correlation between % EDU homeworks completed and Exam 1% grade was r(127) = .33, p<.01. This would be interpreted as, “Those who completed more of the EDU homeworks tended to have higher grades on Exam 1.

However, this is not an experimental study, so the “EDU effect” may be confounded by lots of other variables.

While we can’t measure or even think all of the possible confounding variables, we can consider what are things that might be related to both % EDUs completed and Exam scores ????

We chose 3 for study: motivation, exam study time, GPA

Bivariate & Multivariate contributions – DV = Exam 1% grade

<table>
<thead>
<tr>
<th>predictor</th>
<th>Motiv</th>
<th>St. Time</th>
<th>GPA</th>
<th>% Pink</th>
</tr>
</thead>
<tbody>
<tr>
<td>r(p)</td>
<td>.28(&lt;.01)</td>
<td>.45 (&lt;.01)</td>
<td>.46 (&lt;.01)</td>
<td>.33(&lt;.01)</td>
</tr>
</tbody>
</table>

All of these predictors have substantial correlations with Exam grades!!

| β(p) | .32 (.02) | -.25 (.04) | .09 (51) | .58 (.01) |

GPA does not have a significant regression weights – after taking the other variables into account, it has no unique contribution!

Exam study time has a significant regression weight, however, notice that it is part of a suppressor effect! After taking the other variables into account, those who study more for the test actually tend to do poorer on the exam.

%Pink does have a significant regression weight. Even after taking the other variables into account, those who do more EDU exercises tend to do better on the exam.

Motivation does have a significant regression weight. Even after taking the other variables into account, those who are more motivated tend to do better on the exam.

Notice that only two of the 4 predictors had the same “story” from the bivariate and multivariate analysis!!!!

Path Analysis – allows us to look at how multiple predictors relate to the criterion – considering both “direct” and “indirect” relationships!!

<table>
<thead>
<tr>
<th>Motiv</th>
<th>- .31</th>
<th>St Time</th>
<th>-.25</th>
</tr>
</thead>
<tbody>
<tr>
<td>%Pink</td>
<td>.32</td>
<td>Exam 1%</td>
<td>.58</td>
</tr>
<tr>
<td>GPA</td>
<td>.21</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

GPA → no direct effect – but indirect effects thru %pink & St Time
Motiv → direct effect – also indirect effects thru %pink & St Time
%Pink → direct effect – also indirect effect thru St Time

-β for St Time? Less %Pink predicts more St Time, suggesting that those who study more were those who did less work before they started to study for the exam, and they also did poorer on the exam!