SAMPLE 1

(FIRST 3 PAGES ONLY)

Lecture 8: Mediation & Moderation (Field Ch. 10)

Path Analysis

- Path analysis may be undertaken in SPSS when you want to model observed variables through the direct and indirect effects of a mediating variable (**mediation**) or, alternatively, when you want to assess the impact (change) of the interaction of two variables on the relationship between two other variables (IV & DV) (**moderation**)
- To undertake this in SPSS, "a series of multiple regression analyses are performed where variables are regressed on prior variables in the model" (Hills, 2007, p. 277)
- Assuming large sample sizes, Hills (2011) recommends at least **10 participants for each path parameter** to be estimated in path analysis

Form of Path Analysis

- Increasingly, path analysis is being used in psychology
- Baron and Kenny's (1986) work on mediation is often quoted
- The methods outlined in their paper are used when "a third variable plays an important role in *governing the relationship between two other variables*"
- A path model is a diagram that outlines independent, intermediary, and dependent variables
- Moderation and mediation analyses may be undertaken using path analysis in SPSS with simplified path models
- More complex designs require other statistical software (e.g. AMOS, MPlus, etc.)

Extension of Regression – Designs of Moderation & Mediation

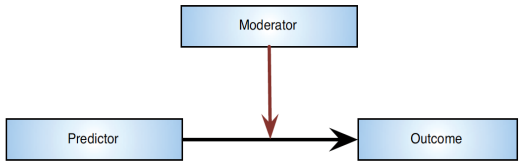
- In regression, we generally ask a question like "Does X predict Y?"
- Research questions can move beyond simplified forms of multiple regression analyses
- Mediators address "**how" or "why**" X causes Y, providing greater explanatory power in interpretation ★★

- Moderators address "**when" or "for whom**" X causes Y (simple slopes evaluation) ★★
  - Through the nature of assessing differing slopes, moderation is looking at for whom or when this may assist in the explanation
- Mediation:
  - "If you have a hypothetical **causal sequence** of three (or more) variables, the middle variable is considered a mediator (indirect effect), suggesting representation of at least part of the chain of events leading to changes in the DV" (Tabachnick & Fidell, 2007, p. 159)
  - "The term mediation means the same thing as **indirect effect**. When we say that motivation affects achievement through homework, this is the same as saying that motivation has an indirect effect on achievement through homework or that homework mediates the effect of motivation on achievement"
  - The major purpose of mediation is to **evaluate the impact of the indirect effect** of a 3rd variable that may in part be driving **the direct effect between the IV & DV**

Defining Mediation & Moderation

- Mediation occurs when a variable impacts the relationship between the IV that influences the DV
  - Attempts to identify a variable/s through the IV, which acts to influence the DV
  - "To "mediate" something is to stand in between two other things and pass on the effect of one to the other"
  - Mediator should be measured as a **continuous variable**
- Moderation occurs when the relationship between the IV and DV *changes* as a function of the **level** of a 3rd variable (the moderator)
  - If the predictor here was anxiety and the outcome was aggression, then this may vary according to a moderator of say, sex, which may provide information on the variation of this effect according to whether the participant is a male or female
- A moderator is an **interaction effect**, whereas a mediator is where an **indirect effect** impacts on the DV

FIGURE 10.4
Diagram of the *conceptual* moderation model



1. Moderation

- The combined effect of two variables on another is *known conceptually as* moderation, and in statistical terms as an interaction effect ★★
- "Moderation means the same thing as interaction. When we say that ability moderates the effect of TV viewing on achievement, this is the same thing as saying that ability and TV viewing interact in their effect on achievement"
- The major purpose of moderation relates to the interaction of 2 variables, regardless of whether they are a combination of categorical or continuous measures
- Creating an interaction term, variables need to be centred; that is, scores need to be converted to deviation scores (value minus mean, or converting to a mean of 0) to assist and avoid multicollinearity being an issue
    - In addition, Field suggests that in moderation you need to place 3 variables in the analysis
    - The interaction term may make the *b*s for the main predictors uninterpretable in many situations
    - For this reason, it is common to transform the predictors using grand mean centering
    - Field suggests "Centring refers to the process of transforming a variable into deviations around a fixed point"

Data Preparation for Moderation

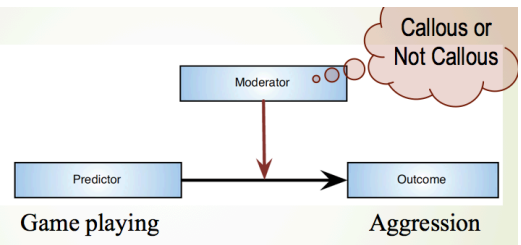
- Creating new moderators (interactions)
- A moderator variable can be a combination of **2 continuous variables or a discrete variable and a continuous variable**
- Using 2 continuous variables, you multiply the 2 variables together after you centre them
- Create centred variables in their own column (value minus mean for each variable – continuous only), then multiply these 2 columns into a single newly created centred variable (moderator)
- Centring accommodates and controls for multicollinearity when the original values are already in the regression
- Therefore, following these instructions (further outlined in Field) you have 3 new variables, one for each column of the two variables and then a new column of the interaction variable
- This complex aspect is not undertaken when using PROCESS
- Think of it as being similar to the interaction in ANOVA
- Although here we have to prepare and identify the interactions we wish to assess, you should be aware problems may occur when multiplying 2 columns
- Some of you may have seen this already
- What happens when 2 negatives are multiplied together creating a positive value?
- Jose (2013) suggests centering only one of the variables will avoid this problem and there will be ease of interpretation when only one has the negative value
- Alternatively, if you are using PROCESS, then add values for one variable so that there are no negative values and instead in the interaction this would equate to a low level of that continuous variable

When to Use Moderation

- Moderator variables may strengthen or weaken the relationships between IV & DV
- **Strong theoretical underpinnings** are required for moderation to be used in research
    - E.g. You're interested in the effect of education and sex on income levels
    - Really asking whether levels of education show a different relationship for males or females on income, or: do males and females present differing slopes in education on the income in the regression equation?
- In moderation one variable depends on the level of another variable
- Same principle of main effects and interactions in ANOVA

Conceptual Moderation Model Example

- If callous-unemotional traits were a moderator then we're saying that the strength or direction of the relationship between game playing and aggression is affected by callous-unemotional traits



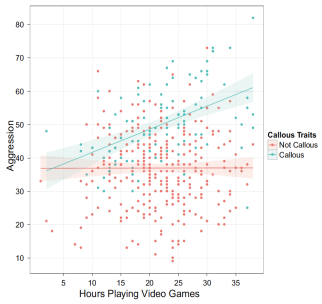
Moderation Example

- Do violent video games make people antisocial?
- Participants
  - 442 youths
- Variables
  - Aggression
  - Callous Unemotional Traits (CaUnTs)
  - # of hours spent playing video games per week

Differing Slopes in Moderation

- You can visually see that there is variation in levels of aggression from the predictor of hours playing video games when the personality variable of having callous traits or not having callous traits is incorporated in the analysis
- From the figure below, you can see that Callous traits as two groups shows a different pattern when # of hours playing video games is regressed onto Aggression

FIGURE 10.5
A categorical moderator (callous traits)



- When the moderator does not have an effect, then there is no difference (LHS)
- However, when the level of the moderator does have an effect, the results are very different (RHS)
- We can see in the diagram on the right that callous traits does indeed have an effect on levels of aggression and the hours of video games being played
- It provides greater explanatory power; in effect, it explains the "**when" or "for whom**" the said relationship occurs between the two variables
- These two graphs outline there is something happening, which suggests that variability exists at levels depending on the third level "for whom" (those who are Callous compared to those who are Not Callous) level of callous unemotional traits when moderation was used
- It can simply be undertaken on two variables where you want to separate the variance between the DV and IV in an interaction

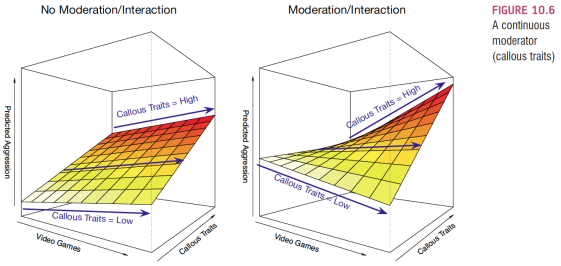
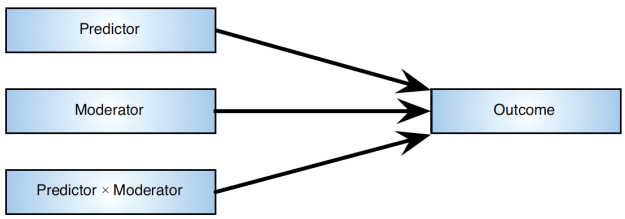


FIGURE 10.6
A continuous moderator (callous traits)

The Statistical Moderation Model

- *Differs from ANOVA in that predictors are continuous (rather than categorical)*

FIGURE 10.7
Diagram of the *statistical* moderation model



$$Y_i = \left(b_0 + b_1 A_i + b_2 B_i + b_3 AB_i\right) + \varepsilon_i$$

$$\text{Aggression}_i = \left(b_0 + b_1 \text{Gaming}_i + b_2 \text{Callous}_i + b_3 \text{Interaction}_i\right) + \varepsilon_i$$

SAMPLE 2

(FIRST 3 PAGES ONLY)

Lecture 10: Factor Analysis & Principal Components Analysis (PCA) (Field Ch. 17)

Purpose of Factor Analysis

- "PCA and FA have considerable utility in reducing numerous variables down to a few factors. Mathematically, PCA and FA produce several linear combinations of observed variables, each linear combination a factor"
- The goal of research using PCA or FA is to concisely describe (and perhaps understand) the relationships among observed variables or to test theory about underlying processes
- As Tabachnick and Fidell (2013) suggest, in EFA "one seeks to describe and summarize data by grouping together variables that are correlated"

Differences Among PCA & EFA

- NB: If an exam question is referring to components, it can't be referring to factors; the difference is in the variance that is explained ★★
- Principal Components Analysis (PCA) produces results that allow for the components to account for **all the variance** in the variables imputed
    - *Whether the variance is unique, shared, or error*
    - *In the communalities in the output, you will find that Initially the extraction is 1 because it is all the variance – this doesn't occur in factor analysis*
- Exploratory Factor Analysis (EFA) produces factors that **only account for the variance that is shared** by the variables
- "FA and PCA differ on the variance that is analyzed. In PCA, all the variance in the observed variables is analyzed. In FA, only shared variance is analyzed; attempts are made to estimate and eliminate variance due to error and variance that is unique to each variable."

Different Forms of Extraction in SPSS for FA

- In SPSS several extractions are available such as;
    - Principal components analysis (PCA): The mathematically determined solution with the common, unique and error variances mixed into the components
    - Principal factor extraction / Principal axis factoring: Estimates communalities in an attempt to eliminate unique and error variances from variables – only shared variance is evaluated

Confirmatory Factor Analysis (CFA)

- Confirmatory factor analysis is a separate technique that is not performed through SPSS
- It tests theoretical and conceptual underpinnings of an a-priori model with items loading on specific factors
- It measures both the amount of variance explained and also through uniqueness the amount that is not accounted for within the model (the residuals at each level of the model)
- *Much more powerful form of analysis*
- *How construct validity is obtained*

Purpose of Factor Analysis

- Research using PCA or FA's goal is to describe the relationships among observed variables or test theory about underlying processes
- To test for clusters of variables or measures
    - As Tabachnick and Fidell (2013) suggest in EFA, "one seeks to describe and summarize data by grouping together variables that are correlated"
- See whether different measures tap a common dimension
- Data summarization
- Data reduction
    - "PCA and FA have considerable utility in reducing numerous variables down to a few factors. Mathematically, PCA and FA produce several linear combinations of observed variables, each linear combination a factor"

Summarization & Reduction

1. Data summarization = Derives underlying dimensions that, when interpreted and understood, describe the data in a much smaller number of concepts than the original individual variables
2. Data reduction = Extends the process of data summarization by deriving an empirical value (factor score or summated scale) for each dimension (factor) and then substitutes this value for the original values
    - *Use the latent construct rather than the original value*

Factors and Components

- Factor analysis attempts to achieve *parsimony* by explaining the maximum amount of **common variance** in a correlation matrix using the smallest number of explanatory constructs
    - These 'explanatory constructs' are called factors
- PCA tries to explain the maximum amount of **total variance** in a correlation matrix
    - Done by transforming original variables into set of linear components

Matrices

- The component matrix represents the loadings between the variables and components
- The rotated component matrix represents the loadings between the variables and rotated components and may be used to assist researchers interpret the components by representing the best simple **structure** after rotation to the best, simplest **solution**

R-Matrix

- In Factor Analysis & PCA we look to reduce the R-matrix (correlation matrix) into a smaller set of uncorrelated dimensions

FIGURE 17.2
An *R*-matrix

	Talk 1	Social Skills	Interest	Talk 2	Selfish	Liar
Talk 1	1.000					
Social Skills	.772	1.000				
Interest	.646	.879	1.000			
Talk 2	.074	−.120	.054	1.000		
Selfish	−.131	.031	−.101	.441	1.000	
Liar	.068	.012	.110	.361	.277	1.000

Factor 1

Factor 2

Graphical Factor Plot

- *Graphical presentation of mean values of variables*
- You can see that Selfish, Talk 2 and Liar are high on Consideration, but Liar is positioned lower and positively on Sociability, whereas Talk 2 and Selfish are positioned low but negatively
- Interest, Talk 1 and Social Skills are positioned high on Sociability but low on Consideration, with Social Skills just on the negative side

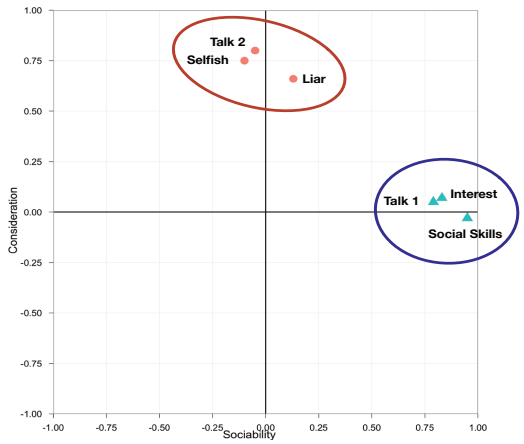
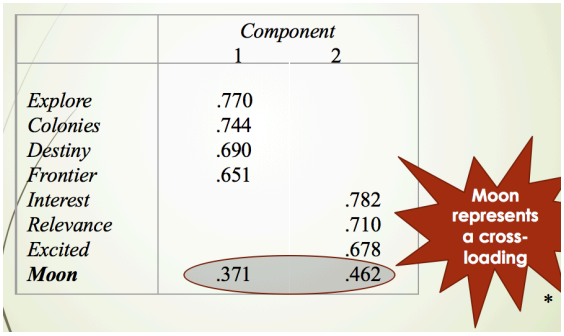


FIGURE 17.3
Example of a factor plot

Interpretation of the Components

- It is useful at this point when investigating the Rotated Component Matrix to suppress loadings less than .3
- Where loadings are seen to be greater on more than one component, this interpretation becomes complex
- The next step is to look carefully at each component and review the wording of each item
- What does the grouping of these items tell you about the factor so that you may label it?
- *If you have complex variables (which have a loading >0.3 for more than 1 variable), you must inform the reader*
- *If you are conducting an exploratory pilot study, you go back and look at the item*
- *Options for dealing with the complex variable(s):*
    - *Adjust the wording if necessary (e.g. If there too much information in the question?)*
    - *Collect more data or data from another sample to see if it can be discriminated*
    - *Eliminate the item and retest*



- Labelling the components can be quite difficult
- As suggested by T & F (2007), to properly create a questionnaire, often it takes many attempts to carefully identify and then measure a given abstract concept, that may be both reliably and validly used
- Often it entails several attempts at reliably obtaining information and many pilot studies undertaken
- If you view the rotated components matrix, you can see that the complex variable Moon clearly falls between the other two components
- What might Component 1 indicate or what would you label this component after reading the information on each item? – Hills suggests that the importance of space exploration, so perhaps Perceived Importance
- Component 2 seems to be more about the level of interest a person has about space and space exploration
- Perhaps this could be labelled Personal Interest
- However, when you undertake this exercise it is important that you sit and reflect on what you would think would be appropriate labelling
- IMPORTANTLY, Moon cross loads onto both components
    - That is, it loads above the value of .3 onto the 2 components
    - Complex variable – carefully look at the wording and the intent of the question being posed in the questionnaire

Rotation of Factors

- Factor rotation = The reference axes of the factors are turned about the origin until some other position has been reached
- Since unrotated factor solutions extract factors based on how much variance they account for, with each subsequent factor accounting for less variance
- The ultimate effect of rotating the factor matrix is to redistribute the variance from earlier factors to later ones to achieve a simpler, theoretically more meaningful factor pattern

What Do We Mean by Rotation?

- Rotation is required because without rotation it would be difficult to interpret the results
- In SPSS 2 main rotation methods are used
- Orthogonal rotation*:* Varimax rotation is orthogonal rotation that simplifies the factors by setting levels on a simplicity criterion
    - Default option in SPSS
    - The goal of Varimax is to maximize the variance of the factor loadings by making high loadings higher and low ones lower for each factor
    - Orthogonal = Axes are maintained at 90 degs
- Oblique rotation*:* Uses the orthogonally rotated solution on rescaled factor loadings, therefore the solution may be oblique with respect to the original factor loadings
    - The factors are often correlated in oblique rotation (Tabachnick & Fidell, 2013)
    - Oblique = Axes are not maintained at 90 degs