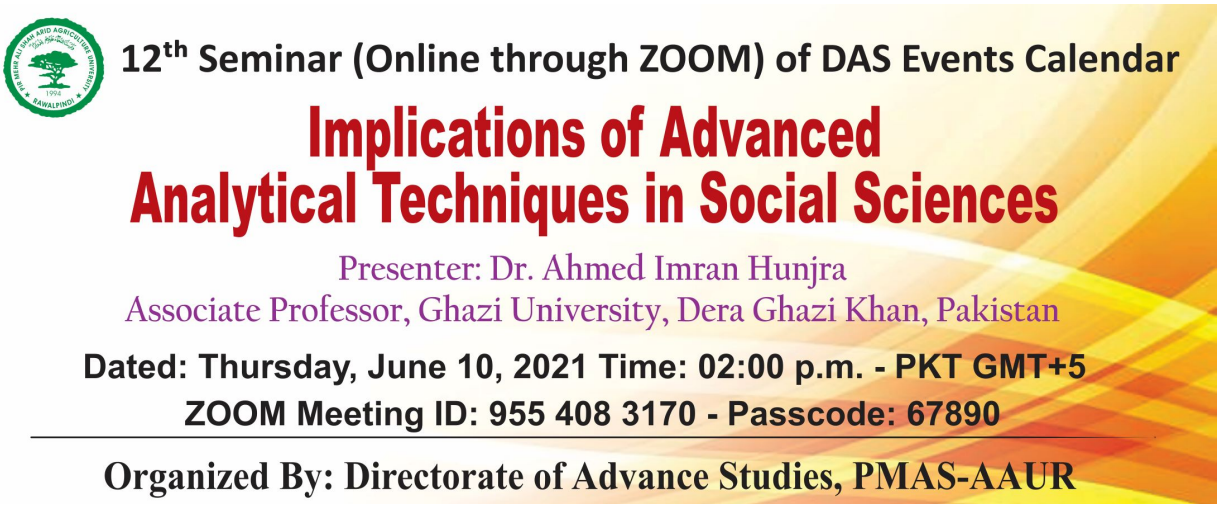
DIRECTORATE OF ADVANCED STUDIES

EVENT CATALOGUE

2021

12TH SEMINAR OF DAS EVENTS CALENDAR – 2021

IMPLICATIONS OF ADVANCED ANALYTICAL TECHNIQUES IN SOCIAL SCIENCES



ACTIVITIES

Nature of Data

- **Qualitative**

- **Quantitative**

There are three different types of Quantitative data:

- *Cross-sectional*
- Time series
- Panel & Pool

Importance of Analytical Techniques

- "The unexamined life is not worth living" *Socrates*
- "Everything must be taken into account. If the fact will not fit the theory---let the theory go." *Agatha Christi*
- "It is a capital mistake to theorize before one has data." *Arthur Conan Doyle*
- "Knowing is not enough, we must apply. " *Bruce Lee*
- "Data! Data!" he cried impatiently. "I can't make bricks without clay." *Bharat Rajpurohit*
- "There are fifty who can reason synthetically for one who can reason analytically." *Christopher Sims*

Analytical Techniques

Outline:

- Nature of Data
- Scale of Measurement/Instrument
- Pilot Testing (Subjective measure & numerical Measure)
- Validity of the newly developed Instrument by (EFA)
- Validity of the Instrument by (CFA)
- Treatment of missing Values
- Reliability of the Data
- Normality of Data
- Descriptive Statistics
- Inferential Statistics: (Measure of Differences and Measure of Association)
- Measure of Differences (Mean differences between the groups, one sample T-test, Independent sample T-test, Paired sample T-test, One way ANOVA)
- Measure of Association (Regression, Correlation)
- Structural Equation Model (SEM) & Mediation Analysis

Instrument

- Instrument is a tool which is used to collect the primary data and it is collected through questionnaires, interviews, observations etc..

Pilot Testing

To check the measurement problems and the reliability and validity of the questions first pilot testing should be done.

There are two methods to refine (validate) the instrument

- ***Subjective Measure (Face or content Validity)***

- ***Numerical Measure (Exploratory Factor Analysis & Confirmatory Factor Analysis)***

Scales of Measurement

- Ratio scales: used when exact numbers are given for e.g. how many orders do you operate?
- Interval scale: used for responses to various items on 5/7 points.
- Ordinal scale: for preference in use
- Nominal scale: used for personal data

Scales of Measurement

- Scales of measurement include:
  - Nominal
  - Ordinal
  - Interval
  - Ratio
- The scale determines the amount of information contained in the data.
- The scale indicates the data summarization and statistical analyses that are most appropriate.

Face Validity

The extent to which the content of the items is consistent with the construct definition, it totally based on the researcher's judgment.

An expert of the specific area check the face validity.

It is a subjective measure

Confirmatory Factor Analysis

In order to check the validity of the instrument and model fit CFA is used.

Construct validity of measured variables actually represent the theoretical latent construct which is designed to measure. It is made up of three components:

- ***Convergent Validity (****Factor Loadings, Average Variance Extracted (AVE) & Construct Reliabilities)*
- ***Discriminant Validity***
- ***Nomological Validity***

Confirmatory Factor Analysis

- Confirmatory factor analysis (CFA) is another process to validate the instrument.

- It also checks the goodness fit of the model discloses the measurement errors in the model.

- According to Steenkamp and Baumgartner (2000) CFA is the best method to understand either the questions/items/statements have strong interest or not to explain the variable.

Exploratory Factor Analysis (EFA)

When the new instrument is introduced to measure any concept, it should be checked through EFA. Followings are the criterias to validate the instrument:

- ***Extraction Value/Load Factor***
- ***Kaiser-Meyer-Olkin (KMO) Measure and Bartelett test***
- ***Total Variance Explained***
- ***Scree plot***

Structural Model Analysis

The structural model is the second part of SEM that demonstrates the direct and indirect relationship between the variables.

- ***Relative chi-square (CMIN/DF)***
- ***Root Mean Square Error of Approximation (RMSEA)***
- ***Goodness-of-Fit Index (GFI)***
- ***Adjusted Goodness-of-Fit Index (AGFI)***
- ***Comparative Fit Index (CFI)***

These are the multiple indices to assess structural model.

Treatment of Missing Values

Missing values may exists in the collected data:

- If 5% of the respondents did not give the response of the total sample size then replace the missing values by the mean of series or replace with mean of the five above & below of the missing values.
- If the response is missing between 5% to 25% then revisit the respondents.
- If its more than 25% then revisit yourself and your questionnaire.

Diagnosing Measurement Model Problems

In addition to evaluating goodness-of-fit (GOF) statistics, the following diagnostic measures for CFA should be checked:

Path Estimates:

Connector Among Error Terms:

Standardized Residuals Covariance

Standardized Values of Adopted Goodness of Fit Statistics

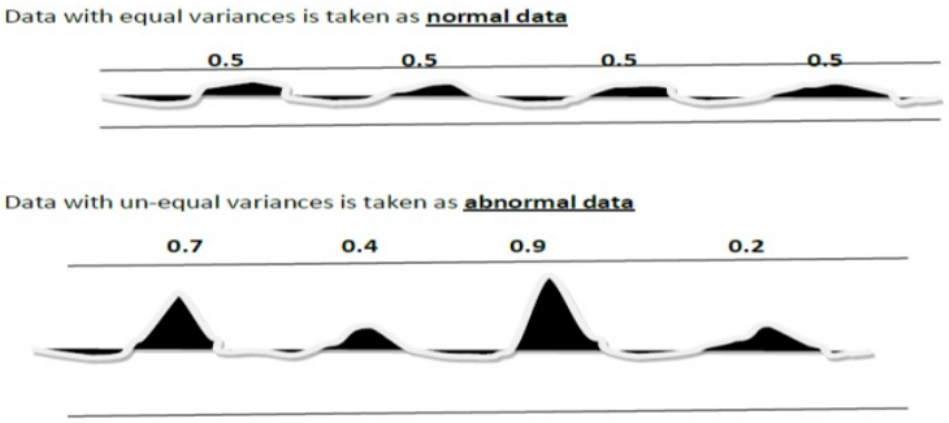
Sr.#	Fit indices	Ranges and acceptance criteria
1	Relative / Normal chi-square (CMIN/ DF)	(1.00< CMIN/DF <5.00) Best/excellent fit: 1-3 Reasonably acceptable: 3-5 Poor fit: above 5
2	Goodness Fit Index (GFI)	(0.90< GFI <1.00) Best/excellent fit: ≥ 0.95 Reasonably acceptable: ≥ 0.90
3	Adjusted Goodness Fit Index (AGFI)	(0.80< AGFI <1.00) Best/excellent fit: ≥ 0.90 Reasonably acceptable: ≥ 0.80
4	Comparative Fit Index (CFI)	(0.90< CFI <1.00) Best/excellent fit: ≥ 0.95 Reasonably acceptable: ≥ 0.90
5	Root Mean Square Error of Approximation (RMSEA)	(0.01< RMSEA <1.00) Best/excellent fit: ≤ 0.05 Reasonably acceptable: 0.06 – 0.08 Poor fit: above 0.10

Reliability Analysis

- The Confirmatory Factor Analysis (CFA) finalized/refined the instrument and that was used for final survey.
- The extent of consistency between the numerous measurements of a variable is called as reliability (Hair et al., 1998).
- Internal consistency of the scale is measured by reliability analysis and is most widely used. Cronbach's alpha is obtained by the Coefficient of Alpha.
- Its limit is 0 to 1 and data will be reliable if its value is more than or equal to 0.70

Normality of Data

- Normality of the study variables (scaled variables) should be checked. We can name it test of homogeneity
- There might be normal or abnormal data but most of the time data will be abnormal.
- H_0: There is no significant difference in the variance of data. (Normal data)
- H1: There is significant difference in the variance of data. (Abnormal data)



Reliability Analysis

- The Confirmatory Factor Analysis (CFA) finalized/refined the instrument and that was used for final survey.
- The extent of consistency between the numerous measurements of a variable is called as reliability (Hair et al., 1998).
- Internal consistency of the scale is measured by reliability analysis and is most widely used. Cronbach's alpha is obtained by the Coefficient of Alpha.
- Its limit is 0 to 1 and data will be reliable if its value is more than or equal to 0.70

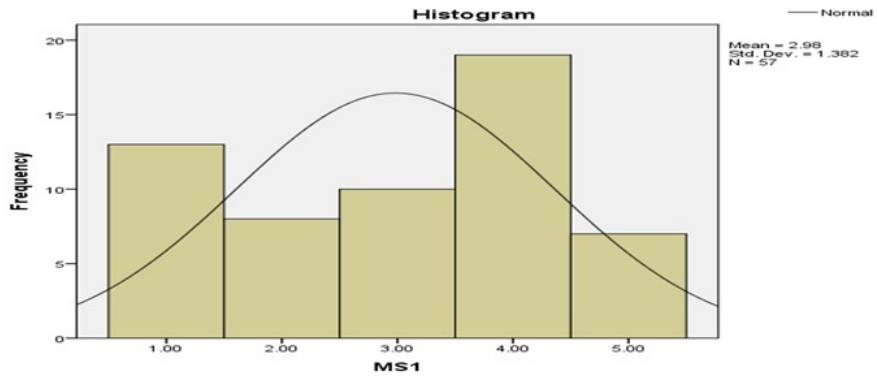
Treatment of Missing Values

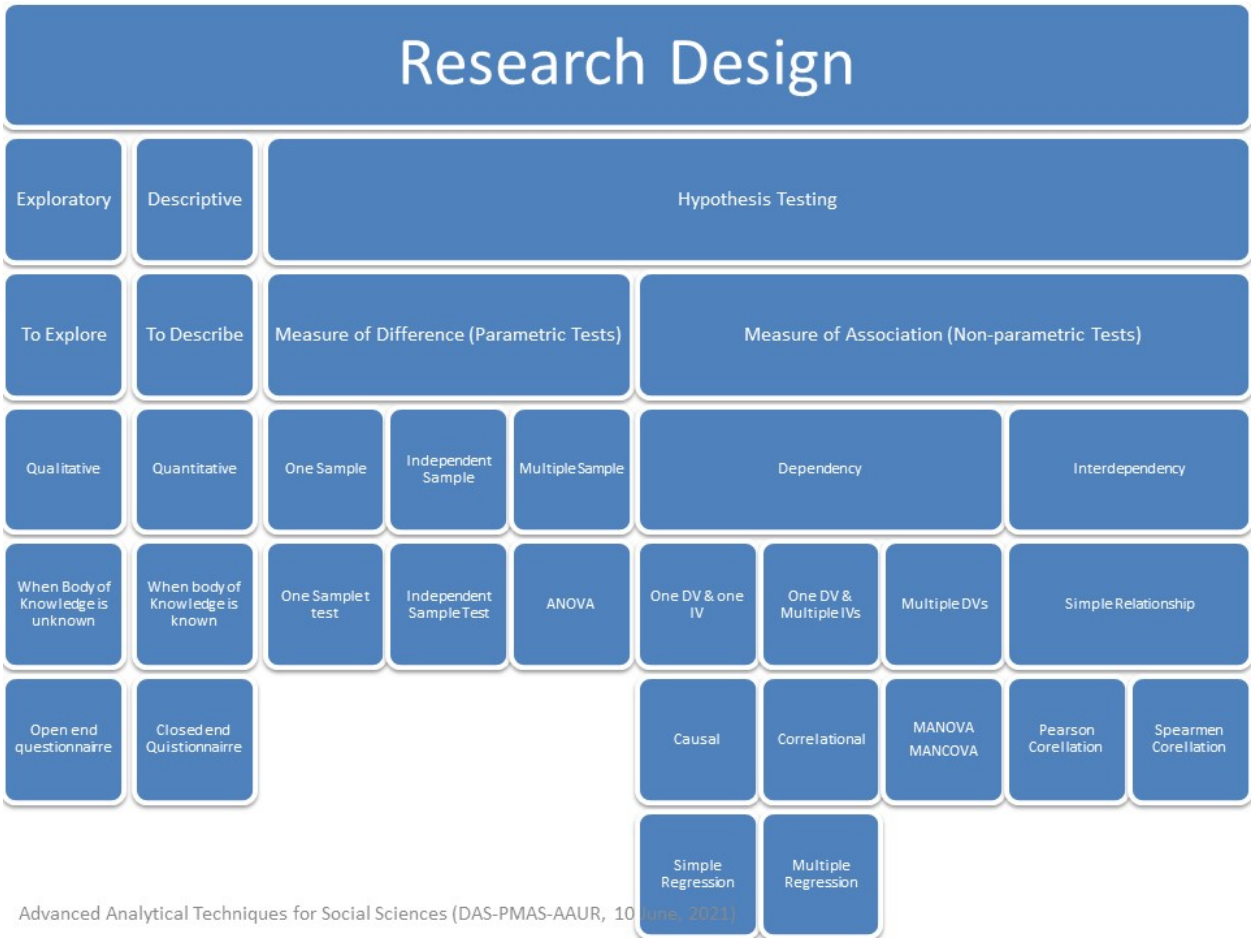
Missing values may exists in the collected data:

- If 5% of the respondents did not give the response of the total sample size then replace the missing values by the mean of series or replace with mean of the five above & below of the missing values.
- If the response is missing between 5% to 25% then revisit the respondents.
- If its more than 25% then revisit yourself and your questionnaire.

Cont.....

- Descriptive are (Mean, std. deviation, skewnesses and kurtosis) the lenient criterias to detect the normality of data.
-





Descriptive Statistics

- Descriptive statistics is used to check the strength of response of the respondents about the questions.
- Descriptive statistics employ three kinds of indicators: dispersion measures (std. deviation, range, skewnes), central tendency measures (mean, median, mode, kurtosis) and frequency distribution.

Cont.……

- Some are very strict (conservative) criterias to detect the normality. Like Kolmogorov-Smirnov and Shapiro-Wilk, decision is made on the bases of significant values.
- If the sample size is < 50 then consider the Shapiro-Wilk and if sample size is > 50 then Kolmogorov-Smirnov significant value is considered for decision of the normality

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
MS1	.225	57	.000	.872	57	.000

Measure of Differences (Parametric tests)

These tests are applied with demographic variables and also called comparative analysis.

- One Sample t-test
- Independent Sample t-test
- Paired Sample t-test
- One way ANOVA

Structural Equation Model (SEM)

- SEM was applied to identify the direct, indirect and mediation relationship between the dependent (Endogenous variable), independents (Exogenous variable) and mediators. The confidence interval may be 95 % with 5 % level of significance. The level of significance of the variable is checked at 5 %.

There are Two steps:

- First, the direct effect of the independents to the dependent variable is to be examined.
- Second, the indirect effect of the independent to the dependent variable is to be examined through mediating variable.
- If the relationship of the independent variable to the dependent variable is insignificant through the mediating variable, it is complete (full) mediation; but if it is still significant and the path declines, it is partial mediation (Prabhu, 2007; Hoyle and Smith, 1994; Baron and Kenny, 1986)

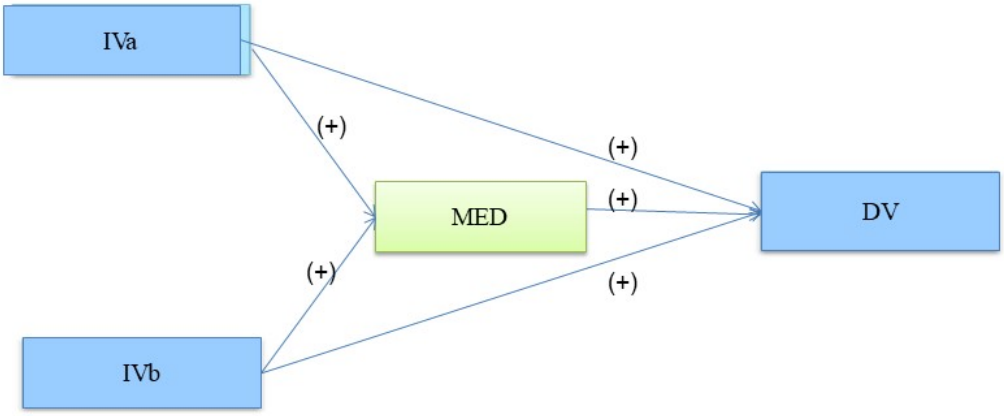
Regression

- In statistics, **regression analysis** is a statistical process for estimating the relationships among variables. When the focus is on the relationship between a dependent variable and one or more independent variables.
- The two basic types of regression are simple linear regression and multiple linear regression. Linear regression uses one independent variable to explain and/or predict the outcome of Y, while multiple regression uses two or more independent variables to predict the outcome.

Correlation Coefficient

- Degree and type of relationship between any two or more variables in which they vary together over a period.
- Correlation can vary from +1 to -1. Values close to +1 indicate a high-degree of positive correlation, and values close to -1 indicate a high degree of negative correlation.
- Values close to zero indicate poor correlation of either kind, and 0 indicates no correlation at all.

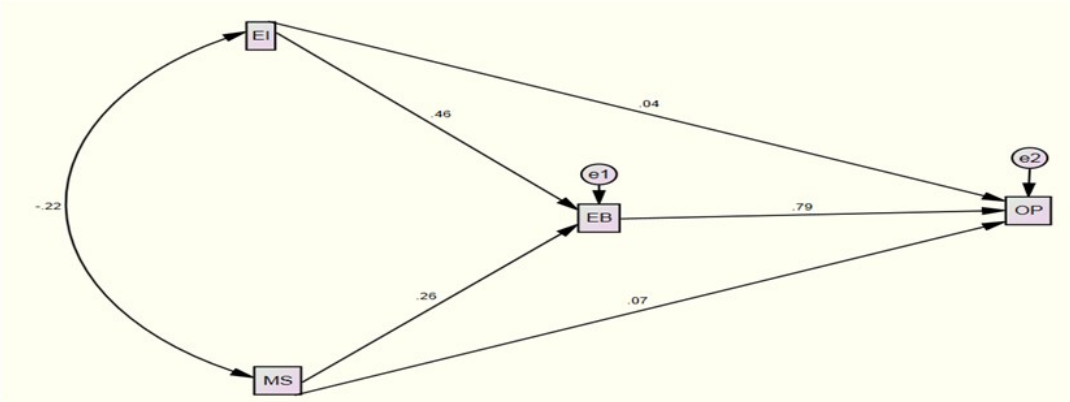
Theoretical Framework & Hypotheses



Mediation Effects (Indirect Effects)

Variables			Direct Effects		Indirect Effects		Hypothesis Support
			Estimate	P-Value	Estimate	P-Value	
DV	←	IVa	0.46	***	0.04	0.690	H6 is accepted.
DV	←	IVb	0.26	0.033	0.07	0.358	H7 is accepted.

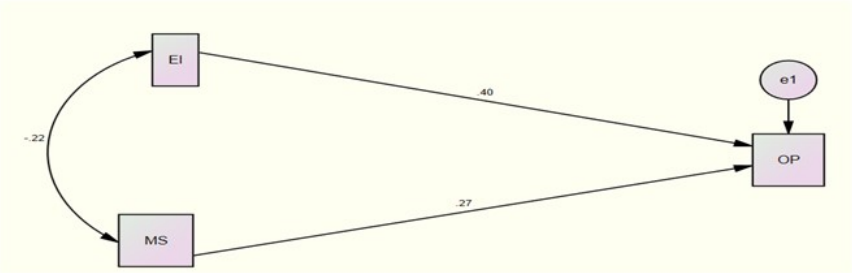
IVs on MV & MV on DV



Variables			Estimate	P-Value	Hypothesis Support
MV	←	IVa	0.46	***	H3 is accepted
MV	←	IVb	0.26	0.033	H4 is accepted
DV	←	MV	0.79	***	H5 is accepted

Direct Effects

(The direct effect of the independent to the dependent variable)



Variables		Estimate	P-Value	Hypothesis Support
DV	IVa	0.40	0.011	H1 is Supported
DV	IVb	0.27	0.026	H2 is Supported