

MANUAL ON GOOD PRACTICES IN EXTENSION RESEARCH & EVALUATION

Regression Data Interpretation
Qualitative **Extension**
Methods Thematic Analysis Designs
Standard Deviation Descriptive Validity
Learning Management Reliability Tools
Ethics **Research** Margin of Error
Journals Diffusion
Modeling Monitoring Constructs
Policy Implications Impact Assessment
Experimental Variance **Evaluation**
Inference Case Study Variables
Replication Psychometrics



PUBLISHED BY

Agricultural Extension in South Asia (AESAs)

www.aesa-gfras.net

C/o Centre for Research on Innovation and Science Policy (CRISP)

Road No. 10, Banjara Hills, Hyderabad-500 034, INDIA

www.crispindia.org

In collaboration with



Manual on Good Practices in Extension Research & Evaluation

Editors: P Sethuraman Sivakumar, B S Sontakki, Rasheed Sulaiman V, Saravanan R and Nimisha Mittal

P Sethuraman Sivakumar, Senior Scientist, ICAR-Central Tuber Crops Research Institute (ICAR-CTCRI), Thiruvananthapuram, India <http://www.ctcri.org>

B S Sontakki, Head-EXM Division, ICAR-National Academy of Agricultural Research Management (ICAR-NAARM), Hyderabad, India <https://naarm.org.in>

Rasheed Sulaiman V, Director, Centre for Research on Innovation and Science Policy (CRISP), Hyderabad, India <http://www.crispindia.org>

Saravanan R, Director (Agricultural Extension), National Institute of Agricultural Extension Management (MANAGE), Hyderabad, India <http://www.manage.gov.in/>

Nimisha Mittal, Program Manager, Centre for Research on Innovation and Science Policy (CRISP), Hyderabad, India <http://www.crispindia.org>

Extracts from this material may be reproduced for personal and not-for-profit use with due acknowledgement to AESA.

February 2017

MANUAL ON
GOOD PRACTICES IN EXTENSION
RESEARCH & EVALUATION

Editors

P Sethuraman Sivakumar

B S Sontakki

Rasheed Sulaiman V

Saravanan R

Nimisha Mittal



About AESA

Agricultural Extension in South Asia (AESAs) is a network of all those who are interested and involved in Extension and Advisory Services (EAS) in South Asia (www.aesa-gfras.net).

Our vision is to improve food and nutrition security through efficient and effective EAS in South Asia. Our mission is to promote sharing, learning and networking for building effective and efficient EAS.

AESA is part of the Global Forum for Rural Advisory Services (GFRAS).

The Centre for Research on Innovation and Science Policy (CRISP) hosts the Secretariat of AESAs. CRISP conducts policy relevant research on agricultural extension and rural innovation.

ACKNOWLEDGEMENT

This Manual has been developed by AESA in collaboration with ICAR-NAARM, CRISP, ICAR-CTCRI and MANAGE through a research grant from the Global Forum for Rural Advisory Services (GFRAS). The editors also benefitted greatly from the inputs provided at the Workshop on 'Good practices in extension research and evaluation' from 29 November to 2 December 2016, held in India.

The editors would like to thank **Dr R. Kalpana Sastry, Acting Director** ICAR-National Academy of Agricultural Research Management (ICAR-NAARM); **Mrs V. Usha Rani, IAS, Director General**, National Institute of Agricultural Extension Management (MANAGE), (Hyderabad, India); and **Dr James George, Director**, ICAR-Central Tuber Crops Research Institute (ICAR-CTCRI), Thiruvananthapuram, India, for their assistance in organizing this workshop.

Our sincere gratitude to **Ms Kanika Singh**, Senior Research Fellow, Centre for Research on Innovation and Science Policy (CRISP) for her support in organizing the workshop and to **Ms Rajeswari Aluri**, Administrative Assistant, CRISP for her support in the compilation of this manual.

We also thank **Ms Niharika Lenka** for the design and layout of this manual and **Ms Alphy J Plakkoottam** for language editing.

The contribution and support provided by everyone is greatly appreciated. We apologize if we have unintentionally overlooked anyone.

Any opinions expressed here are those of the editors.

P Sethuraman Sivakumar
B S Sontakki
Rasheed Sulaiman V
Saravanan R
Nimisha Mittal



Photo Credit: Subash S P

CONTENTS

Acknowledgement	i
Background	xi
Introduction	xiii
Module 1: Basics of Extension Research	1
Module 2: Extension Research Problems	5
Module 3: Research Designs and Methods in Extension	9
Module 4: Methods for Identifying Concepts, Constructs, Variables and Hypotheses	17
Module 5: Using Thematic Analysis for Construct Conceptualisation	25
Module 6: Sampling Methods for Quantitative Extension Research	33
Module 7: Sampling Methods for Qualitative Extension Research	41
Module 8: Sample Size Estimation	51
Module 9: Best Practices in Data Collection – Errors and Biases	59
Module 10: Online Data Collection	67
Module 11: Data Exploration and Preparation – Missing Data Treatment	75
Module 12: Data Exploration and Preparation – Outlier Treatment	87
Module 13: Data Exploration and Preparation – Testing Statistical Assumptions	97
Module 14: Statistical Designs in Data Analysis – Descriptive to Inferential Statistics	107
Module 15: A Few Guidelines for Data Interpretation and Inference	115
Module 16: Focus Group Discussion	121
Module 17: Case Study Research	129
Module 18: Human Subjects Research- Ethical Practices in Data Collection	139
Module 19: Modeling Adoption of Agricultural Technologies	145
Module 20: Scale Development in Extension Research	153
Module 21: Developing Summated Scales - Latent Factor Identification	163
Module 22: Developing Summated Scales – Structural Equation Modeling	175
Module 23: Methods for Evaluation of Extension Programmes	181
Module 24: Methods of Analysing Impact of Agricultural Technologies	191
Module 25: Publishing Extension Research	203
Module 26: Multi-Disciplinary Extension Research	209
Final Remarks: Where do we go from here?	215
Annexure A-1: Guidelines for Designing Web Surveys	217
Annexure A-2: Advanced methods of data collection	220
Annexure A-3: Using SPSS for basic statistical analysis	223
Annexure A-4: Guidelines for avoiding biases in selecting items	243
Annexure A-5: IBM SPSS AMOS	237
About the Editors and Authors	253

List of Tables

Table No	Title	Page No
1	Overview of the Manual	xv
1.1	Goals of scientific research	2
1.2	Major paradigms in research and evaluation (adapted from Mertens & Wilson 2012)	3
3.1	Research strategies used in extension studies (Adapted from Creswell, 2009; Diem, 2002; McMillan and Schumacher, 2006)	10
3.2	Research methods and techniques used in quantitative, qualitative and mixed methods designs (adapted from Creswell, 2009)	15
4.1	Popular theories used in extension research	19
5.1	Tabulation of codes against each verbatim	29
5.2	Preliminary themes, sub-themes and codes	29
5.3	Results of thematic analysis of computer use behavior	31
7.1	Sample size requirement for qualitative studies	49
8.1	Necessary sample size to detect a given effect size for multiple regression, ANOVA ("t" test)(= 0.05 and = 0.20)	56
8.2	Necessary sample sizes based on the number of independent variables for multiple regression (r =0.30; Power = 0.80)	57
10.1	Cost and data quality in traditional and online data collection methods	68
10.2	Advantages and disadvantages of online data collection	68
10.3	Web survey response types and their use	71
14.1	Guidelines for choosing statistical tests	109
14.2	Selecting statistical tests based on research questions and hypotheses	113
14.3	Selecting statistical tests based on purpose of research and extension	113
15.1	Levels of the ladder along with associated actions	118
17.1	Approaches to scientific knowledge	134
17.2	Positivist framework	135
18.1	A few examples of human subject research in extension	140
19.1	Varieties of rice crops suitable for climate change; An option survey of farmers' selected particular option	147
20.1	Analysis of judge's rating of items	156
21.1	Difference between exploratory factor analysis and principal component analysis	167
21.2	Statistical significance of factor loadings based on sample size	170
22.1	Minimum sample size requirements (Hair Jr. et al., 2006)	176
22.2	Elements and relationships of factors and variables in CFA (Hair Jr. et al., 2006)	178
24.1	Impact Assessment Methods and Techniques	201
26.1	Differences between various types of research collaborations (Choi and Pak, 2006)	210

Table No	Title	Page No
26.2	Multidisciplinary research work and the role of extension scientists	212
A-4.1	Minimum Values of CVR critical for One Tailed Test ($p = .05$)	235
A-5.1	Criteria for model fit assessment, item validity and reliability	244

List of Boxes

Box No	Title	Page No
2.1	How to write a good research problem statement?	6
2.2	Using narrative hooks in research topic statement	7
4.1	Ways of Conceptualizing a Phenomenon (Kaplan, 1964)	20
4.2	Tips for writing good research objectives	23
5.1	An Example of defining the research problem and research question	26
5.2	Sample transcripts from an audio-recorded in depth personal interview	27
5.3	Guidelines for selecting data extracts for coding	28
5.4	Theme-piling method	29
6.1	Some terms related to sampling	33
7.1	Steps in respondent-assisted sampling	47
8.1	Sample size estimation procedure for descriptive studies	54
8.2	Online sample size calculators	57
10.1	Online survey using Google Forms®	70
10.2	A few Do's and Don'ts when using Google Forms as a web survey tool	70
10.3	Steps in conducting Online Bulletin Board Research	72
10.4	Software required for Qualitative Online Research	73
14.1	A few terms used in statistical analysis	108
14.2	Choosing type of graphs	112
16.1	Criteria for assessing success of focus groups	123
17.1	Points to be considered while doing an individual case study	134
17.2	Skills required for a researcher for conducting an interview for case study research	134
17.3	Example of building theories from case study research	136
18.1	Elements of the informed consent sheet	143
19.1	Suitability of econometric models for adoption	146
23.1	Steps in programme evaluation	182
23.2	Evaluation tools and techniques by programme stage	184
23.3	Basic participatory and collective evaluation methods	185
23.4	Participatory evaluation of disease control through vaccination campaigns	186
23.5	Logical Framework Analysis evaluation of Kisan Call Centers in Tamil Nadu	187

Box No	Title	Page No
23.6	Bennett's hierarchy applied in evaluation of contract broiler farming (CBF) and non- contract broiler farming (NCBF) systems	188
24.1	Definitions of elements in the Impact Assessment Framework	192
24.2	Procedure for selecting impact indicators	193
24.3	Method for Impact Assessment of Poverty Alleviation Projects, 1998	200
25.1	Common reasons for rejection of research articles	205
A-3.1	How to write the output of independent samples "t" test	225
A-3.2	How to write the output of paired samples "t" test	227
A-3.3	How to write the output of Chi-square test	230
A-3.4	How to write the output of One-way ANOVA	233

List of Figures

Fig No	Title	Page No
1	Steps in the Research Process	xiv
1.1	Extension Research process (Adapted from Harvey, 2015)	3
2.1	Distinguishing among Research Topic, Research Problem, Research Purpose & Research Questions	5
3.1	The research design framework (Adapted from Creswell, 2009)	9
3.2	Forms of data collection in survey research (Adapted from Creswell, 2012)	13
4.1	Construct conceptualisation process	18
5.1	A sample mind map showing relationship between themes, sub-themes and codes	30
6.1	Sampling strategy (Adapted from Robinson, 2014)	34
6.2	Sampling strategies used in extension research	35
7.1	Types of generalisations in empirical studies (Adapted from Onwuegbuzie and Leech, 2007)	42
7.2	Types of purposive sampling (Adapted from Daniel, 2012)	44
10.1	Popular online web survey tools	69
14.1	Decisions based on Type I and Type II errors	111
14.2	Critical region	112
15.1	Wisdom pyramid	116
17.1	Decision tree for choosing a case	130
17.2	Decision process for choosing case study research (Thomas and Myers, 2015)	133
17.3	Decision process for choosing a case study research on farmer's suicides	133
18.1	A sample informed consent form	143

Fig No	Title	Page No
20.1	An example of Q- Grid	157
21.1	Modern scale development process (Adapted from Churchill, 1979; Hinkin et al., 1997 and DeVellis, 2003)	164
22.1	Visual diagram of the Computer Use Behaviour constructs or latent factors and their items	177
24.1	Impact assessment framework (Adapted from Devos Vaughan and Associates - http://devosvaughan.com/services_programme-review.html)	192

ABBREVIATIONS

AESA	Agricultural Extension in South Asia
AMOS	Analysis of Moment Structures
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
ATARI	Agricultural Technology Application Research Institute
ATMA	Agricultural Technology Management Agency
BCR	Benefit Cost Ratio
C-OAR-SE	Construct definition–Object representation–Attribute classification–Ratter entity identification–Selection of item type and answer scale–Enumeration and scoring rule
CALIS	Covariance Analysis and Linear Structural Equations
CBF	Contract Broiler Farming
CD	Capacity Development
CFA	Confirmatory Factor Analysis
CRISP	Centre for Research on Innovation and Science Policy
CTCRI	ICAR-Central Tuber Crops Research Institute
DAC	Department of Agriculture and Cooperation
DID	Difference in Differences
DIKW	Data Information Knowledge Wisdom Hierarchy
EAS	Extension and Advisory Services
EFA	Exploratory Factor Analysis
EQS	Equations
ERB	Ethics Review Board
FAQs	Frequently Asked Questions
FB	Facebook
FCR	Feed Conversion Ratio
FGD	Focus Group Discussion
FLD	Front Line Demonstration
GFRAS	Global Forum for Rural Advisory Services
GMO	Genetically Modified Organism
GoI	Government of India
HTML	Hyper Text Markup Language
IA	Impact Assessment
ICAR	Indian Council of Agricultural Research
ICMR	Indian Council of Medical Research
ICT	Information and Communications Technology
IEC	Institutional Ethics Committee
IFAD	International Fund for Agricultural Development
IPM	Integrated Pest Management
IRB	Institutional Review Board
IRR	Internal Rate of Return
IRT	Item Response Theory
IV	Instrumental Variable
IVRS	Interactive Voice Response System
JAEE	Journal of Agricultural Education and Extension
KASA	Knowledge, Attitude, Skill and Aspirations
KCC	Kisan Call Center
KNN	K Nearest Neighbor Analysis
KVK	Krishi Vigyan Kendra

LISREL	Linear Structural Relationships
LLM	Local Linear Matching
LOF	Local Outlier Factor
MANAGE	National Institute of Agricultural Extension Management
MAPP	Method for Impact Assessment of Poverty Alleviation Projects
MCAR	Missing Completely At Random
MGNREGA	Mahatma Gandhi National Rural Employment Guarantee Scheme
MLE	Measurement, Learning and Evaluation Projects
MoA	Ministry of Agriculture
NAARM	ICAR-National Academy of Agricultural Research Management
NARES	National Agricultural Research and Education System
NARS	National Agricultural Research System
NCBF	Non Contract Broiler Farming
NN	Nearest Neighbor
NPV	Net Present Value
OFT	On Farm Testing
OLS	Ordinary Least Squares
OR	Odds Ratio
OS	Organisational Support
PAME	Participatory Assessment Monitoring and Evaluation
PBT	Partial Budgeting Technique
PCA	Principal Component Analysis
PI	Principal Investigator
PIA	Participatory Impact Assessment
PLM	Participatory Learning Methods
PRA	Participatory Rural Appraisal
PRR	Private Rate of Return
REB	Research Ethics Board
ROC	Receiver Operating Characteristic curve
SaaS	Software as a Service
SEM	Structural Equation Modeling
SHG	Self Help Group
SMART	Specific, Measurable, Attainable, Realistic, Time bound
SMCR	Sender Message Channel Receiver
SMS	Subject Matter Specialist
SPSS	Statistical Package for the Social Sciences
SRI	System of Rice Intensification
SRR	Social Rate of Return
SWOT	Strengths, Weakness, Opportunities and Threats
ToT	Training of Trainers
WOR	Without Replacement
WR	With Replacement
YPARD	Young Professionals for Agricultural Development



Photo Credit: M. J. Chandre Gowda

Background

This document on *Good Practices in Extension Research and Evaluation* is developed as a hands on reference manual to help young researchers, research students, and field extension functionaries in choosing the right research methods for conducting quality research and evaluation in extension.

This manual has been compiled by the resource persons who participated in the Workshop on 'Good Practices in Extension Research and Evaluation' jointly organised by the ICAR-National Academy of Agricultural Research Management (NAARM), Centre for Research on Innovation and Science Policy (CRISP), Agricultural Extension in South Asia (AESAs), ICAR-Central Tuber Crops Research Institute (ICAR-CTCRI), and the National Institute of Agricultural Extension Management (MANAGE) at NAARM, Hyderabad (India), during 29 November-2 December 2016.

This manual builds on the experiences derived from organising this workshop and the feedback received during the workshop.



Photo Credit: Kanika Singh

Introduction

*Rasheed Sulaiman V, P Sethuraman Sivakumar, B S Sontakki, Saravanan R and
Nimisha Mittal*

Need for good practices in extension research and evaluation

Growth of any discipline is directly proportional to the creation of knowledge in that discipline. Extension research is the backbone of extension discipline. Though extension research was envisaged to develop sound methods and models to help the field functionaries in effective delivery of extension services, very little progress has been made in the past few decades. Absence of a 'rigorous approach' in advancing extension science is the primary reason for these lacunae and there is an urgent need to look into the 'mechanics' of conducting extension research (Sivakumar, 2013).

While research tools and techniques in core disciplines from which extension borrowed its research methods have evolved significantly, extension research still depends heavily on many of the outdated tools. "Though a number of analytical frameworks can be used to study extension, the applications of these frameworks have been limited so far. Most studies which evaluate extension methods have focussed on the impact of extension on farm productivity. Assessment of cost effectiveness and scalability have often been missing" (Ragasa and Babu, 2013).

Over the past 2-3 years, many have expressed their concerns on the declining quality and contribution of extension research (through the AESA Blogs) and the limited influence extension research has on improving extension practice, designing new programmes or developing and influencing policies. These have also resulted in low respect and recognition of extension researchers.

Therefore, re-orienting extension research is urgent, and calls for a coordinated approach by integrating state-of-the-art methods from other sciences in order to improve the utility and visibility of the extension research outcomes. Adopting several good practices, such as the following, can enhance the quality of extension research:

- Creative generation of relevant research ideas using an intuitive/common sense approach;
- Selection of a rigorous and robust research design;
- Choice of right variables following alternate criterion-referenced validity assessment procedures;
- Selection of appropriate sample sizes to maximise generalisability;
- Estimation of reliability and validity through robust modelling procedures, such as Structural Equation Modelling;
- Deployment of resource and time saving but accurate tools, such as shortened paper surveys and e-surveys;
- Compensation of respondents so as to maximise the accuracy of responses;
- Data cleaning by employing missing value estimation and assumption testing tools, and multivariate data modelling (Sivakumar and Sulaiman, 2015).

Extension research is a unique social science inquiry where research ideas are gathered from the field problems and put through a systematic cycle of objective investigations that result in significant solutions. Apart from developing theories and models that advance scientific knowledge, extension research should also provide new insights for improving extension policy and practice.

Workshop

A Workshop was organised on ‘Good Practices in Extension Research and Evaluation’ at the ICAR-National Academy of Agricultural Research Management (NAARM), during 29 November-2 December 2016, for young extension researchers and PhD students in extension and this Manual is the outcome of this workshop.

How to use the manual

This document is a manual which researchers could use to refine and update their knowledge on how to approach research in extension in a more systematic and scientific manner. This manual delves into the six steps in the research process (Fig. 1) which have to be undertaken in order to reach the final output (publication).



Fig. 1: Steps in the Research Process

The process begins from selection of the research problem, research designs and methods, sampling, data collection and analysis, interpretation and inference – and together all these culminate in a publication. These steps are critical for a robust output irrespective of the theme of one’s research, which could be evaluation, adoption modelling, impact assessment, etc.

Our aim in producing this manual is to provide researchers with materials that cover the field of extension research along with appropriate tools/methods. This manual is divided into two sections:

- I. Basics to Essentials; and
- II. Path Finder.

Section I (Basics to Essentials) elaborates on the basic aspects of the extension research process, such as research problems, designs and methods in conducting quality research. It integrates various qualitative

and quantitative methods with appropriate statistical tools to provide practical guidelines for young researchers in choosing and using the various methods (Modules 1-18).

Section II (Path Finder) describes the advanced techniques in modelling adoption, using structural equation modelling for scale development, monitoring, evaluation and impact assessment (Modules 19-26). This section also discusses publishing research outputs in different formats.

The manual is organised in such a way that it is user-friendly. Each module has a clearly defined objective. It starts with a brief introduction followed by a discussion interspersed with illustrations, examples, and tips for the user. Some of the modules have an annexure that provides more detailed information on some of the methods. The key points from the discussion are summarised at the end of each module.

Overview and structure of the manual

The 26 modules of the Manual cover the following areas (Table 1):

Table 1: Overview of the Manual

Module	Title and objectives
Module 1	Basics of Extension Research Introduce the philosophical foundations of scientific social research. Explain functions, goals and paradigms of scientific social research. Explain the process of extension research.
Module 2	Extension Research Problems Explain basics aspects of extension research problems and questions. Provide good practice guidelines for selecting and writing research problems and questions.
Module 3	Research Designs and Methods in Extension Provide an overview of research strategies and designs used in extension. Provide guidelines for selecting right research designs and methods.
Module 4	Methods for Identifying Concepts, Constructs, Variables and Hypotheses Introduce methods for identifying concepts and constructs for extension research. Explain thematic analysis as a method of constructs identification.
Module 5	Using Thematic Analysis for Construct Conceptualisation Introduce thematic analysis as a tool for construct conceptualisation. Explain the method for identification of themes using thematic analysis.
Module 6	Sampling Methods for Quantitative Extension Research Introduce sampling methods used in extension research. Explain the probabilistic and non-probabilistic sampling methods.
Module 7	Sampling Methods for Qualitative Extension Research Provide an overview of various sampling strategies followed in qualitative research. Provide guidelines for sample size for qualitative research.

Module	Title and objectives
Module 8	<p>Sample Size Estimation</p> <p>Provide an overview on methods used in estimating sample size in extension studies. Explain sample size estimation methods. Elaborate on guidelines for choosing adequate samples.</p>
Module 9	<p>Best Practices in Data Collection-Errors and Biases</p> <p>Provide an overview of errors and biases in data collection. Explain strategies to avoid or minimize them.</p>
Module 10	<p>Online Data Collection</p> <p>Introduce various online data collection tools. Explain the tools in details and how to use them.</p>
Module 11	<p>Data Exploration and Preparation- Missing Data Treatment</p> <p>Make readers understand the importance of checking and cleaning data prior to statistical analysis. Explain various methods and tools for examining the data before statistical analysis.</p>
Module 12	<p>Data Exploration and Preparation- Outlier Treatment</p> <p>Make readers understand the importance of checking and cleaning data prior to statistical analysis. Explain various methods and tools for identifying and treatment of outliers in the data.</p>
Module 13	<p>Data Exploration and Preparation- Testing Statistical Assumptions</p> <p>Make readers understand the importance of checking and cleaning data prior to statistical analysis. Explain various methods and tools for examining the data for statistical assumption prior to analysis.</p>
Module 14	<p>Statistical Designs in Data Analysis- Descriptive to Inferential Statistics</p> <p>Provide an overview of statistical tests available for interpreting the data. Provide guidelines for choosing specific tests based on the nature of research.</p>
Module 15	<p>A Few Guidelines for Data Interpretation and Inference</p> <p>Provide an overview of interpretation and inference making from the data. Explain various approaches for effectively interpreting the data and draw inferences and conclusions.</p>
Module 16	<p>Focus Group Discussion</p> <p>Introduce focus group discussion (FGD) as a tool for data collection. Explain steps in conducting, coding and analysis of FGD data.</p>
Module 17	<p>Case Study Research</p> <p>Introduce the case study research approach. Explain the method and techniques for doing case study research. Suggest how to carryout case study research with a rigorous methodology.</p>

Module	Title and objectives
Module 18	Human Subject Research-Ethical Practices in Data Collection Provide an overview of human subjects research. Describe the human subjects research guidelines followed in India.
Module 19	Modeling Adoption of Agricultural Technologies Introduce modeling adoption of agricultural technologies. Explain the multinomial and ordinal logit models. Suggest future work on adoption modeling.
Module 20	Scale Development in Extension Research Introduce various scale development approaches followed in extension research. Explain the scale construction approaches in the classical measurement paradigm.
Module 21	Developing Summated Scales- Latent Factor identification Provide an overview of summated scales. Describe steps in developing summated rating scales.
Module 22	Developing Summated Scales- Structural Equation Modeling Provide an overview of structural equation modeling for scale development. Describe various steps involved in confirmatory factor analysis for constructing summated scales.
Module 23	Methods for Evaluation of Extension Programmes Understand the meaning of evaluation and differentiate between appraisal, monitoring, evaluation and impact assessment. Discuss different types, tools, techniques, designs, approaches and challenges of evaluation.
Module 24	Methods of Analysing impact of Agricultural Technologies Introduce advanced methods of assessing impact of agricultural technologies. Explain the quantitative methods applied in assessing impact. Suggest future work for improving impact assessment.
Module 25	Publishing Extension Research Introduce the importance of publishing extension research. Explain aspects related to preparation of good manuscripts for publishing in journals. Introduce different non-journal formats for publishing research outputs.
Module 26	Multi-Disciplinary Extension Research Provide an overview of multidisciplinary extension research. Recommend multi-disciplinary research areas to extension scientists.

This manual is intended to be of relevance to the following learner groups:

1. Research scholars;
2. Young researchers;
3. Faculty teaching social science research methods;
4. Research guides mentoring research scholars;

5. Evaluators; and
6. Extension Managers.

With this Manual we hope to make a significant contribution towards the process of strengthening extension research and evaluation in India.

References

- Ragasa, C and Babu S. (2013.) A Review of Frameworks, Analytical Tools and Remaining Knowledge Gaps in Agricultural Extension, Paper presented at the Workshop on “Research on Agricultural Extension Systems: What Have We Learned, and Where Do We Go From Here?” CGIAR Research Program on Policies, Institutions and Markets (October 15-16, 2013), Washington DC, USA. (Available at http://pim.cgiar.org/files/2013/07/Session-2_Best-fit-practices_Catherine-Ragasa.pdf).
- Sivakumar, P S. (2013.) Research in Extension: New tools to reinvent its future, AESA Blog 4, Agricultural Extension in South Asia. (March 2013). (Available at <http://www.aesa-gfras.net/admin/kcfinder/upload/files/Blog%204%20Enhancing%20the%20potential%20of%20quality%20videos%20for%20farmers.pdf>).
- Sivakumar P,S and Sulaiman, R. V. (2015.) Extension Research in India: Current Status And Future Strategies, AESA Working Paper 2, Agricultural Extension in South Asia (December 2015). (Available at <http://www.aesa-gfras.net/admin/kcfinder/upload/files/Final%20one%20WP2.pdf>).

01

Basics of Extension Research

P Sethuraman Sivakumar

Objectives

1. Introduce the philosophical foundations of scientific social research
2. Explain functions, goals and paradigms of scientific social research
3. Explain the process of extension research

1.1 Introduction

Extension research has long played a significant role in maximising the utility of socio-economic interventions. Even though extension research has evolved by integrating scientific research methods and tools from other social sciences, there is still a need to maximise its utility to address the complex changes happening in agriculture.

Understanding the foundations of social science research is a prerequisite for right selection and use of research approaches and methods. This module describes the different philosophical foundations, goals and steps in extension research.

1.2 Discussion

1.2.1 Four methods of knowing

In general, there are four ways or methods by which we can ascertain the truth of something (Kerlinger and Lee, 2000).

- Method of tenacity or method of established belief – This considers something as true because it has been considered true since time immemorial. It says that knowledge is transferred from tradition or culture. This method follows the logic that something is true because it has always been considered true. Its validity and reliability is unquestionable and there is no chance of verification.
- Method of authority – It is a method of established belief. When an eminent person or expert tells something, it is accepted. The emphasis is on not questioning authority because what authority says is true. The emphasis is on the source, not on the methods the source may have used to gain the information.
- Method of intuition (a priori method) – In this method, truth is seen as 'true' because it is logical. It derives from reasoning but does not carry empirical support.
- Method of science – External experience through observation is the foundation of science. It has a unique character, i.e., self-correction.

1.2.2 Functions of science

The functions of science are viewed in two different forms – static and dynamic (Kerlinger and Lee, 2000).

- Static view – In the static view, science is considered as an activity that develops and provides systematized information about a phenomenon. The scientists are expected to discover new facts and to add them to the already existing body of information, such as laws, theories and principles.
- Dynamic view – Science is viewed as an activity where discoveries are used to advance knowledge and provide goods and services which improve human life.

As a field-oriented social science, extension research follows the dynamic approach to develop materials and services which improve extension work.

1.2.3 Scientific social research

Scientific social research is a process of systematic, controlled, empirical and critical investigation that is designed to collect, analyse, interpret and use data about a natural phenomenon (adapted from Kerlinger and Lee, 2000).

Goals of scientific social research

Broadly, the goals of scientific social research are to answer questions and acquire new knowledge. Precisely, the four goals are: description, prediction, explanation and application (Shaughnessy et al., 2012) (Table 1.1).

Table 1.1: Goals of scientific research

Goal	What is accomplished	Example
Description	Defining, classifying, or categorizing events and relationships to describe the phenomena of interest	Documenting and describing the farming systems followed in a region to understand inter- and intra-regional differences under each system
Prediction	Identifying relationships among variables of a particular phenomenon (through correlations) for estimating its outcome in the future (to prediction)	Predicting adoption of a new crop variety or animal/fish breed through correlation between farmers' intention to use and their decision making styles (behavioural attributes)
Explanation	Establishing causal relationships among variables using empirical data helps to understand the dynamics of the phenomenon	Understanding and explaining the factors i.e. personal, socio-economic, etc., responsible for success of extension intervention in a village
Application	Applying the knowledge generated through research to change behaviour of farmers and other stakeholders	Conducting Front Line Demonstrations (FLDs) in the village to show worthiness of the technology to the villagers. E.g., Action research

1.2.4 Major paradigms in scientific research

A paradigm is composed of certain philosophical assumptions that guide and direct thinking and action.

Four different research paradigms (Mertens, 2009), which are shaping social science research are: (i) post-positivist; (ii) pragmatist; (iii) constructivist; and (iv) transformative. These paradigms reflect the approaches of inquiry based on philosophical thoughts and belief systems (Table 1.2).

Table 1.2: Major paradigms in research and evaluation (adapted from Mertens & Wilson 2012)

Paradigm	Branch	Description
Post-positivist	Methods	<ul style="list-style-type: none"> • Theory-driven, scientific approach relies on empirical data to prove or disprove theories • Focuses primarily on quantitative designs of data
Pragmatic	Use	<ul style="list-style-type: none"> • Focuses on the utility of the research • Advocates the use of mixed methods
Constructivist	Values	<ul style="list-style-type: none"> • Seeks to understand stakeholder values and perspectives on a phenomenon • Relies heavily on qualitative methods
Transformative	Social justice	Focuses primarily on viewpoints of marginalized groups and interrogates the systemic power structure through mixed methods to further social justice and human rights

1.2.5 Extension research

Extension research is a unique social science inquiry where research ideas are gathered from the field problems and put through a systematic cycle of objective investigation that results in meaningful solutions to solve these problems while advancing knowledge (Fig. 1.1).

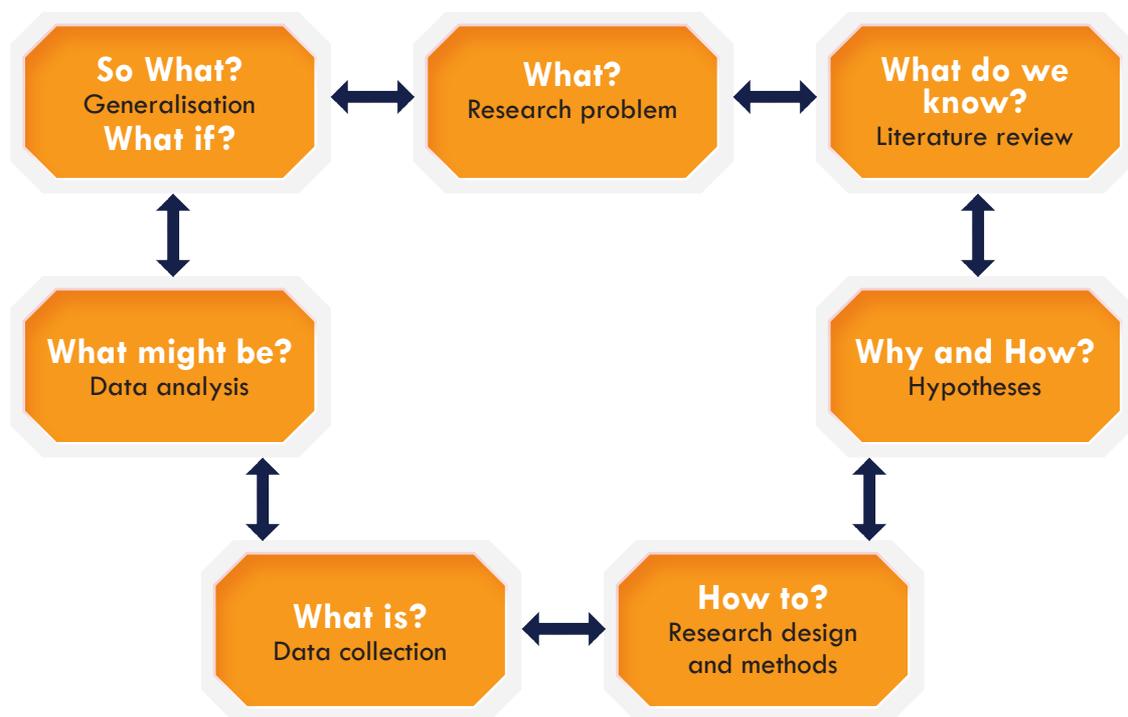


Fig. 1.1: Extension Research process (Adapted from Harvey, 2015)

The extension research process begins with a 'What' question explaining the research problem. Once the problem is identified, an extensive literature review is performed to support the problem and identify research gaps. An elaborate review leads to hypothesis development and its testing through various research designs and methods. The data collection methods are decided based on objectives and hypotheses, and collected data is analysed to make a generalisation on the phenomenon.

1.3 Key Points

- Human knowledge develops through four methods of knowing: Method of tenacity or method of established belief, authority, intuition and science.
- Scientific research provides an objective way of advancing knowledge through description, prediction, explanation and application.
- Four research paradigms – (i) post-positivist; (ii) pragmatist; (iii) constructivist; and (iv) transformative – are guiding various forms of social science research.

References

- Harvey, D. (2015, October). ACE8001: What do we mean by Research? and Can we hope to do genuine Social Science Research. Retrieved from: <https://www.staff.ncl.ac.uk/david.harvey/AEF801/What.html>
- Kerlinger, F. N. and Lee, H. B. (2000.). Foundations of behavioral research. Fourth Edition. Fort Worth: TX: Harcourt College Publishers.
- Mertens, D.M. and Wilson, A.T. (2012.) Program evaluation theory and practice. New York, NY: Guilford.
- Shaughnessy, J. J., Zechmeister, E. B. and Zechmeister, J.S. (2012.) Research methods in psychology. Ninth Edition. New York, NY: McGraw Hill.

02

Extension Research Problems

P Sethuraman Sivakumar

Objectives

1. Explain basic aspects of extension research problems and questions
2. Provide good practice guidelines for selecting and writing research problems and questions

2.1 Introduction

The extension research process starts by identifying the research problem which logically guides the researcher towards identifying the appropriate research designs and methods. This module explains the inter-relationships among various components of the research problem, topic and other elements (displayed in Figure 2.1).

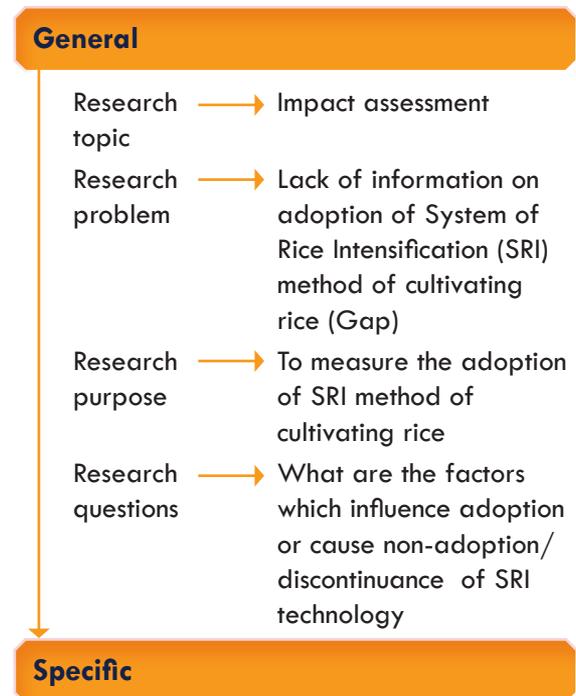


Figure 2.1: Distinguishing among Research Topic, Research Problem, Research Purpose & Research Questions

This module provides a detailed description of various elements of research problems, designs and methods.

2.2 Discussion

2.2.1 Research Problem

A research problem is a precise statement about a concern – a gap in knowledge, skills and/or practice, a condition to improve upon, a difficulty to be eliminated, or an unanswered question that exists from field or theory or academic/field practice, which needs a systematic and empirical investigation (USC, 2016).

Types of research problems

- Practical research problems – Problems derived from field experiences, management issues, future needs, etc.
- Research-based problems – To fill research gaps, support or contradict existing theories

Template for research problem statement

According to Creswell (2009), the problem statement has five basic elements:

1. The research topic;
2. The research problem;
3. A justification of the importance of the problem;
4. The deficiencies in the evidence;
5. The audience that will benefit from a study of the problem.

A good problem statement begins by introducing the broad area of the selected research work, gradually leading the reader to more specific issues of the work (Box 2.1).

Box 2.1: How to write a good research problem statement?

A good research problem should have:

1. A compelling topic: The research problem should be important for the researcher, his/her organisation, the research or academic community, and for the general society.
2. Supports multiple perspectives: A social research problem should provide multiple perspectives on the phenomenon under study. A good problem will elicit diversified views from a composite audience.
3. Feasible: The research problem can be researched in a specific context with the available resources and expertise.
4. Supported in multiple ways: The research problem should be clearly expressed and justified using personal experiences, past research, field experiences and organisational skills.

(i) Research topic

A research topic is the starting point in any research, which should be stated in a clear and concise form to capture the attention of readers. The research topic should specify the problem and research questions to be answered. A research topic is a specific issue in a broad subject area that a researcher wishes to address through a study. It must be narrowed down to a precise research problem and specific research questions. Using a few narrative hooks can help to capture the attention of readers while stating the research problem (Box 2.2).

¹A narrative hook is a journalistic way of writing an opening sentence of a story that “hooks” or captures the reader’s attention and compels them to read the rest of the story.

Good narrative hooks in a problem statement can cause the reader to pay attention, elicit emotional or attitudinal responses, spark interest, and encourage the reader to continue reading (Creswell, 2012).

Box 2.2: Using narrative hooks in research topic statement

A good narrative hook contains:

1. Significant statistical data related to the problem (e.g., Only 40% of all agricultural technologies generated are adopted by farmers.);
2. A provocative question (e.g., Why do farmers commit suicide when the government implements many farmer-friendly policies?);
3. A clear need for research (e.g., climate change and its effects on agriculture is increasingly attracting the attention of extension researchers);
4. The intent or purpose of the study (e.g., the intent of this study is to measure the adoption of Integrated pest management (IPM) practices for mealy bugs in papaya).

(ii) Research problem

As indicated in sections 2.2.1 through 2.2.3, the research problem should be expressed explicitly in a clear and concise form/words.

(iii) Justification of the importance of the problem

Justification of the research problem means presenting reasons supporting the importance of studying a particular issue. Personal field or management experiences, experiences of others, experiences in the workplace and suggestions or findings of past research can be used to establish the importance of a selected research problem.

(iv) Deficiencies in the evidence

A deficiency in the evidence indicates that the past literature or field experiences have not adequately addressed the specific research problem. In general, deficiencies indicate the need to extend the research, replicate a past work, explore a topic further, or if there is a need to add something to improve existing practice.

(v) The audience

The audience include the academicians, students, extension managers, policy makers, farmers and other people who will read the research report for the purpose of gaining knowledge, improving their existing practice, and making management decisions. The problem statement should end with stating how this research work will benefit the audience.

2.2.2 The Research Purpose Statement

A research purpose statement is a declarative sentence which summarizes the specific topic and goals of a document. The purpose is written in the “Statement of problem” section and appears after the last sentence of the introduction. A good purpose statement is specific, precise, concise, clear and goal-oriented. For example, examine the following purpose statement – ‘The purpose of this study is to discuss computer use in Krishi Vigyan Kendras (KVKs)’. Do you deem it to be a good purpose statement? Clearly it’s not a good purpose statement since it is vague and not very precise. So, if we want to create a good purpose statement we should have a statement such as this one – ‘The purpose of this study is to examine the relationship between organisational support for encouraging Subject Matter Specialist’s (SMS) computer use in field extension at Krishi Vigyan Kendras in Punjab and their actual use of computer systems’.

2.2.3 Research Questions

Research questions consist of interrogative sentences which operationalize the purpose of the research. They indicate how the goals of research – such as describing the phenomenon, predicting its outcomes, explaining the processes and outcomes and applying it in improving an existing practice – can be achieved.

A few examples of research questions:

1. What are factors influencing adoption of Genetically Modified cotton varieties by Maharashtra farmers? (Description)
2. What will be the consumer demand for gluten-free pasta in 2020? (Prediction)
3. How does computer anxiety of SMS in KVKs affect his/her computer use for field extension? (Explanation)
4. How do we use this new bio-organic pesticide to control mealy bugs in papaya? (Application)

2.3 Key Points

- Extension research starts with identifying and specifying a relevant research problem.
- Elements in research problem specification include: research topic, research problem, justification of the importance of the problem, deficiencies in the evidence and audience.
- Narrative hooks are journalistic styles used at the beginning of the research problem statement to capture reader's attention.

References

- Creswell, J. W. (2009.) Research design: Qualitative, quantitative, and mixed methods approaches. Third edition. Thousand Oaks: Sage Publications.
- Creswell, J. W. (2012.) Educational research: Planning, conducting, and evaluating quantitative and qualitative research. Fourth edition. Pearson.
- USC. (2016.) Organizing your social sciences research paper: the research problem/question. University of South Carolina Research Guides. (Available at: <http://libguides.usc.edu/writingguide/introduction/researchproblem>)

03

Research Designs and Methods in Extension

P Sethuraman Sivakumar

Objectives

1. Provide an overview of research strategies and designs used in extension
2. Provide guidelines for selecting right research designs and methods

3.1 Introduction

- A research design is the plan or proposal to conduct research which entails the intersection of philosophy, strategies of inquiry and specific methods.
- Choosing a research design involves various intricate decisions based on the purpose and objectives of the study.
- Creswell (2009) has suggested a framework for selecting an appropriate research design based on philosophical paradigms as described in Section 2.1 (Module 2), along with strategies of inquiry and research methods (Figure 3.1).

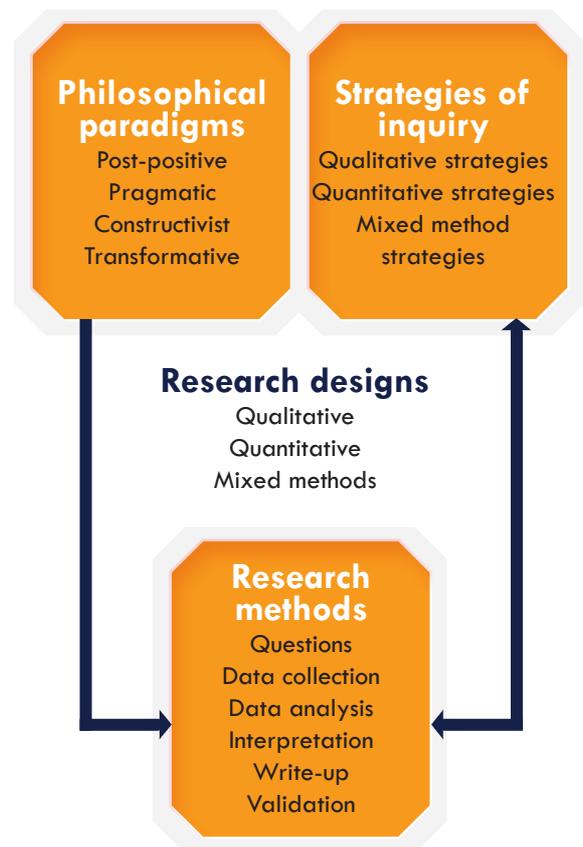


Fig. 3.1: The research design framework (Adapted from Creswell, 2009)

3.2 Discussion

- Strategies of inquiry are types of qualitative, quantitative, and mixed methods, designs, or models that provide specific direction for procedures in a research design (Creswell, 2009).

- Various research strategies are displayed in Table 3.1.

Table 3.1: Research strategies used in extension studies (Adapted from Creswell, 2009; Diem, 2002; McMillan and Schumacher, 2006)

Quantitative		Qualitative		Mixed methods
Experimental	Non-experimental	Interactive	Non-interactive or analytical	
<ul style="list-style-type: none"> • Between-group designs • True experimental • Quasi-experimental • Factorial 	Descriptive	Ethnographic <ul style="list-style-type: none"> • Realist ethnography • Case study • Critical ethnography 	Non-interactive or analytical	Explanatory
With-in group designs <ul style="list-style-type: none"> • Time series experiments • Repeated measures experiments • Single subject experiments 	Comparative	Phenomenological	Concept analysis	Exploratory
	Correlational <ul style="list-style-type: none"> • Explanatory • Prediction 	Grounded theory	Historical analysis	Triangulation
	Survey <ul style="list-style-type: none"> • Cross-sectional • Longitudinal – trend, cohort and panel 	Narrative research	Policy analysis	
	<i>Ex post facto</i>			
	Secondary data analysis			

3.2.1 Quantitative research strategies

A. Experimental strategies

- The researcher manipulates the conditions of what the subjects will experience by structuring actual situations, introducing or controlling certain variables in order to measure their effect on each other, or by systematically imposing or withholding specified conditions.
- Types - Between group and within group designs (Creswell, 2012)

a. Between-group designs

The between-group design is used when the researcher wishes to compare two or more groups.

(i) True experimental design

- Cause-effect analytical experiments which involve random assignment of subjects and conditions to the subjects for objectively measuring the research phenomenon by controlling external interference and errors.
- For example, the influence of technology subsidies on adoption may be studied by assigning subsidy (experimental group) and without subsidy (control group) and assessing adoption by comparing both groups.

(ii) Quasi-experimental design

Quasi-experiments are non-randomised experiments conducted in non-laboratory situations where independent variables are manipulated to assess their influence on dependent variables.

Quasi-experiments differ from true experiments in that subjects are not randomly assigned to conditions. It has both pre-, post-, as well as post-only test designs

Quasi-experimental designs are of three types (Creswell, 2009):

- Non-equivalent (Pre-Test and Post-Test) Control-Group Design – The experimental and control groups are selected without random assignment. Only the experimental group receives the treatment.
- Single-Group Interrupted Time-Series Design – The researcher performs pre- and post-tests in a single group prior to, and after, treatments.
- Control-Group Interrupted Time-Series Design – A modified version of single group time-series design where two non-random groups of participants (control and treatment) are studied over a period of time.

(iii) Factorial designs

- In the factorial design, the researcher studies two or more categorical, independent variables that are analysed at two or more levels (Vogt, 2005).
- The purpose of this design is to study the independent and simultaneous effects of two or more independent treatment variables on an outcome.

b. With-in group designs

Within-group design is employed when the total number of participants is very limited. This design analyses the variations present within the group.

(i) Time series experiments – In this design, studies are conducted on a single group over a period of time, through a series of pre-test and post-test observations. For example, diffusion of new agricultural technologies guided by a series of technology interventions can be studied within a specific geographical or social system over years.

(ii) Repeated measures experiment – In a repeated measures design, multiple treatments are administered on a single group and the researcher compares its performance across these experiments.

(iii) Single-Subject Designs – Single-case designs are usually single-subject or group research (without control), used to evaluate the extent to which a causal relation exists between introduction of an 'intervention' and change in a specific dependent variable. For example, a study focusing on the influence of quality seed material provision to a single farmer or a group of farmers focusing on the

livelihood security aspect may be termed as a single-case research design.

B. Non-experimental strategies

In non-experimental designs, the researcher may examine the relationship between things without manipulating the conditions. Examples include:

Descriptive research – Descriptive research provides a detailed summary of an existing phenomenon by assigning numbers to characteristics of objects or subjects involved.

Comparative research – In comparative designs, the researcher investigates the presence of difference between two or more groups on the phenomenon studied. Comparative research is often multi-disciplinary and utilises quantitative techniques to study the phenomenon. Comparative research methods are used extensively in cross-cultural studies.

Correlational research – Assessing the direction and strength of association between two or more phenomena, without manipulating them. Types of correlational designs: concurrent or explanatory and predictive (Babbie, 1990; Creswell, 2009).

Survey research design

The survey research design provides a quantitative description of trends, attitudes, or opinions of a population by studying a sample of that population. It includes cross-sectional and longitudinal studies using questionnaires or structured interviews for data collection, with the intent of generalizing from a sample to a population (Creswell, 2009).

a. Cross-sectional

In a cross-sectional survey design, the researcher collects data at one point in time. Cross-sectional designs are of several types:

- Single group study measuring current attitudes, beliefs, opinions, or practices;
- Comparing attitudes, beliefs, opinions, or practices between groups.

b. Longitudinal

Collecting data over a period of time from single or multiple groups. There are three types of longitudinal designs:

- Trend study – Conducted every year on a specific aspect, but with a different sample. However, the sample size remains the same. For example, assessing trends in farmers' input utilisation in potato over the years.
- Cohort – A group of subjects are identified and a specific phenomenon is studied over a period of time to assess changes. Although the same population is studied each year, the sample from that population is different for each year. For example, studying food grain consumption patterns through cohorts over a period of time.
- Panels – An identical sample selected at the beginning is used for collecting data every year to assess the changes over time. For example, studying career progression of agricultural students who have graduated from a particular university.

Forms of data collection

The survey is conducted primarily through questionnaires and interviews. The types of questionnaires and

interviews used in quantitative survey research are displayed in Fig. 3.2 below.

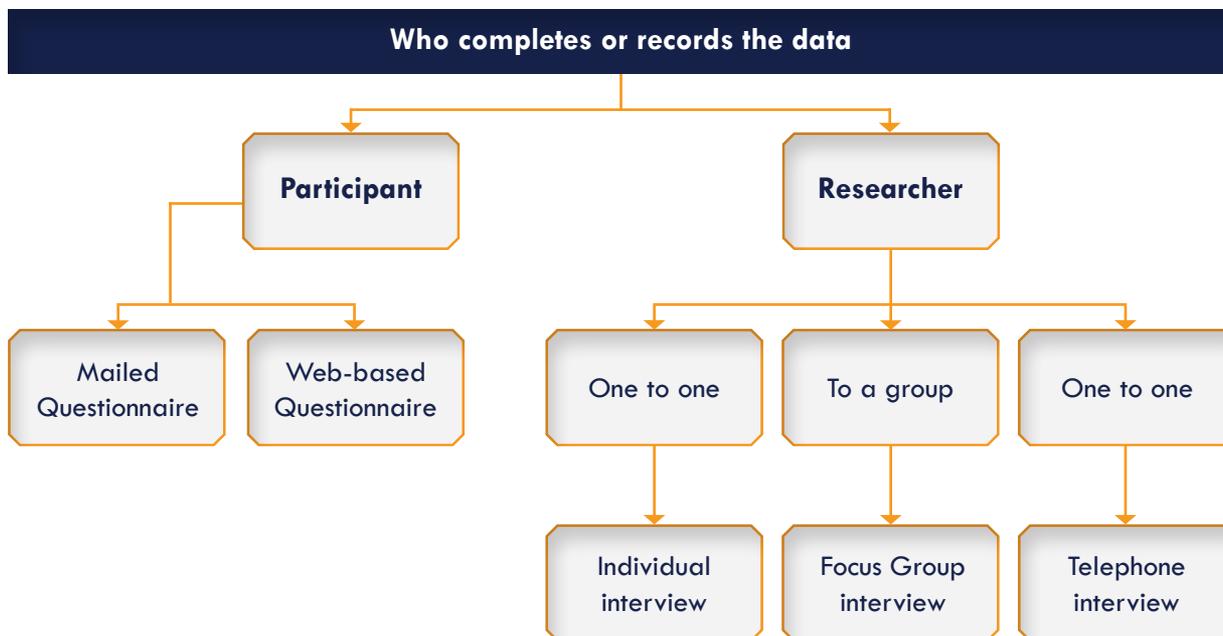


Fig. 3.2: Forms of data collection in survey research (Adapted from Creswell, 2012)

Ex post facto

Ex post facto study or after-the-fact research is a category of research design in which the investigation starts after the phenomenon occurred, without any intervention from the researcher. It is primarily a quasi-experimental study examining how an independent variable, present prior to the study in the participants, affects a dependent variable. This lack of direct control over the independent variable and the non-random selection of participants are the most important differences between ex post facto research and the true experimental research design. Most extension research follows the ex post facto approach.

Secondary data analysis

Secondary data analysis involves the analysis of an existing dataset, which had previously been collected by another researcher, usually for a different research question. This strategy is cost-effective and will produce good results if used intelligently.

3.2.2 Qualitative research strategies

Qualitative designs are concerned with describing or interpreting a phenomenon without manipulating its conditions. It is broadly classified as interactive or non-interactive, based on researchers’ involvement in the inquiry.

a. Interactive

Ethnography – Strategy of inquiry in which the researcher studies an intact cultural group in a natural setting over a prolonged period of time by collecting, primarily, observational and interview data (Creswell, 2009).

Phenomenological research – The researcher attempts to understand and explain how an individual or a group of individuals experience a particular phenomenon from the individual's or individuals' own perspective (Kalaian, 2008) through in-depth interviews.

Grounded theory – Grounded theory is a qualitative strategy for developing theories where the researcher derives an abstract theory of a process, action, or interaction from the views of participants (Channaz, 2006).

Narrative research – In narrative research, the researchers explain the lives of individuals, collect and tell stories about people's lives, and write narratives of individual experiences (Connelly and Clandinin, 1990). Narrative research can be effectively used in documenting indigenous technical knowledge, develop success stories of agricultural technologies, etc., using autobiographies, narrative interviews and oral histories.

b. Non-interactive

Concept analysis

Concept analysis is a research strategy where concepts, their characteristics and relations to other concepts are examined for the purpose of identifying different meanings of the same concept. For example, concepts such as food security, poverty, livelihood security, etc., can be assessed by comparing and contrasting the meanings attributed by various stakeholders like farmers, farm women, farm youth, local traders, panchayat officials, etc.

Historical analysis

Historical analysis is a method of interpreting and understanding the past through a disciplined and systematic analysis. It involves a detailed examination of the 'traces' of past through artefacts, texts, images and old buildings, etc. For example, indigenous technical knowledge can be collected and elaborated upon by key informant interviews, village records, artefacts, old photographs and drawings made by elders.

Policy analysis

Policy analysis is a research strategy for generating information which helps in formulating and implementing policies and then assessing their impact. It uses both quantitative and qualitative methods for collecting and analyzing the information related to policy.

3.2.3 Mixed methods strategy

The mixed methods research design is a procedure for collecting, analyzing, and combining both quantitative and qualitative methods in a single study or a series of studies to understand a research problem (Creswell and Plano Clark, 2011). Combining quantitative and qualitative strategies will provide a comprehensive view of the phenomenon under study. Various types of mixed-methods researches are described below:

(i) Convergent parallel design – Quantitative and qualitative data are collected simultaneously, analysed and interpreted in order to gain a comprehensive view of the phenomenon. It is possible to offset the biases, errors and gaps in the data collected through one form in this design.

(ii) Explanatory sequential design – In this design, quantitative data is collected first and qualitative

data is gathered at a later stage for explaining the results of quantitative analysis. Quantitative data results provide a general idea about the phenomenon and qualitative data refines and explains this view.

(iii) Exploratory sequential design – In this method, qualitative data is collected first to explore the phenomenon under study, and then the quantitative data are gathered to explain the relationships among elements of the phenomenon explained through qualitative data. This method is widely used in developing new scales.

(iv) Embedded design – In this design, both quantitative and qualitative data are collected simultaneously or sequentially, and one form of data is used only as supportive material to justify the results from another set of data. Both first and second set of data are either qualitative or quantitative.

(v) Transformative design – This approach uses any one of the above stated mixed methods designs, but fits the data within a transformative framework (Creswell and Plano Clark, 2011). Examples of transformative frameworks are feminism, gender, ethnicity, disability and racism.

(vi) Multiphase design – The researchers or a team of researchers study a problem through a series of separate studies.

3.2.4 Action research designs

Action research is a problem-oriented design where the researcher systematically gathers information about field practice for improving the effectiveness of field work. Action research can be of two types:

- Practical action research – A small-scale research work narrowly focuses on a specific field problem or issue undertaken in a specific area;
- Participatory action research – A social process in which the researcher deliberately explores the relationship between the individual and other people for the purpose of improving the quality of life.

3.2.5 Research methods

Research methods refer to all forms of data collection, analysis and interpretation used in a research study (Creswell, 2009). It comprises methods and techniques related to research instrument development and testing, methods of formulating items and questions in a research instrument, data collection techniques (interview or questionnaire), data coding, analysis and interpretation. Various research methods or techniques used for quantitative, qualitative and mixed methods research strategies are presented in Table 3.2.

Table 3.2: Research methods and techniques used in quantitative, qualitative and mixed methods designs (adapted from Creswell, 2009)

Aspect	Quantitative Methods	Mixed Methods	Qualitative Methods
Methods	Pre-determined	Pre-determined and emerging	Emerging
Instruments	Structured instruments - Interview schedule, questionnaires, etc.	Both structured and partially structured instruments	Partially structured instruments - Checklists, schedules, interview guides

Aspect	Quantitative Methods	Mixed Methods	Qualitative Methods
Type of questions	Close-ended	Open and close ended	Open-ended
Types of data collected	Numerical	Numerical and descriptive	Numerical and descriptive
Nature of data collected	Technology performance data, psychological attribute data, economic data, social data and census data	Multiple forms	Technology performance data, observational data, document data and audio-visual data
Types of analysis	Numerical and statistical analysis	Numerical, statistical, text and image analysis	Text and image analysis
Interpretation	Statistical interpretation	Multiple forms	Themes and patterns interpretation

3.3 Key Points

- Choosing a research design involves various intricate decisions based on the purpose and objectives of the study.
- A research design is based on the philosophical paradigm of the research, strategies of inquiry and methods.
- The quantitative, qualitative, and mixed method designs provide wide choice and greater flexibility for conducting quality extension research.
- Though extension research is conducted largely as ex post facto, experimental strategies will help to develop theories.

References

- Babbie, E. (2010.) The practice of social research. Twelfth edition. Wadsworth Cengage Learning.
- Charmaz, K. (2006.) Constructing grounded theory. London: Sage Publications.
- Connelly, F. M. and Clandinin, D. J. (1990.) Stories of experience and narrative inquiry. *Educational Researcher*, 19(5), 2–14.
- Creswell, J. W. (2009.). *Research design: Qualitative, quantitative, and mixed methods approaches*. Third edition. Thousand Oaks: Sage Publications.
- Creswell, J. W. (2012.) *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Fourth edition. Pearson.
- Creswell, J. W. and Plano Clark, V. L. (2011.) *Designing and conducting mixed methods research*. Second edition. Thousand Oaks, CA: Sage.
- Diem, K. (2002.) Using research methods to evaluate your Extension program. *Journal of Extension*, 40(6). (Available at: <http://www.joe.org/joe/2002december/a1.php>)
- Kalaian, S. (2008.) Research design. In: P. Lavrakas (ed.) *Encyclopaedia of survey research methods*, (pp. 725-732). Thousand Oaks, CA: Sage.
- McMillan, J. H. and Schumacher, S. (2001.) *Research in education. A conceptual introduction*. Fifth edition. New York: Longman.
- Vogt, W. P. (2005.) *Dictionary of statistics and methodology: A nontechnical guide for the social sciences*. Third edition. Thousand Oaks, CA: Sage.

04

Methods for Identifying Concepts, Constructs, Variables and Hypotheses

P Sethuraman Sivakumar

Objectives

1. Introduce methods for identifying concepts and constructs for extension research
2. Explain thematic analysis as a method of constructs identification

4.1 Introduction

Conceptualisation is a mental process whereby fuzzy and imprecise notions (concepts) are made specific and precise (Babbie, 2010). It is an important part of extension research, which connects theory with practice. During the conceptualisation process, the researcher deliberately analyses the phenomenon to identify its elements and their relationships for the purpose of developing propositions and hypotheses. In extension research, the conceptualisation process is largely performed based on literature review and judges' rating. However, emerging utilitarian perspectives place high emphasis on deriving concepts from stakeholders' perspectives for maximising the impact of the programme. In this module, the various elements and processes involved in the conceptualisation of constructs are described.

4.2 Discussion

The conceptualisation process begins after identification of the research problem and writing research questions and ends with specifying hypotheses for the study. The conceptualisation process is depicted in Fig. 4.1.

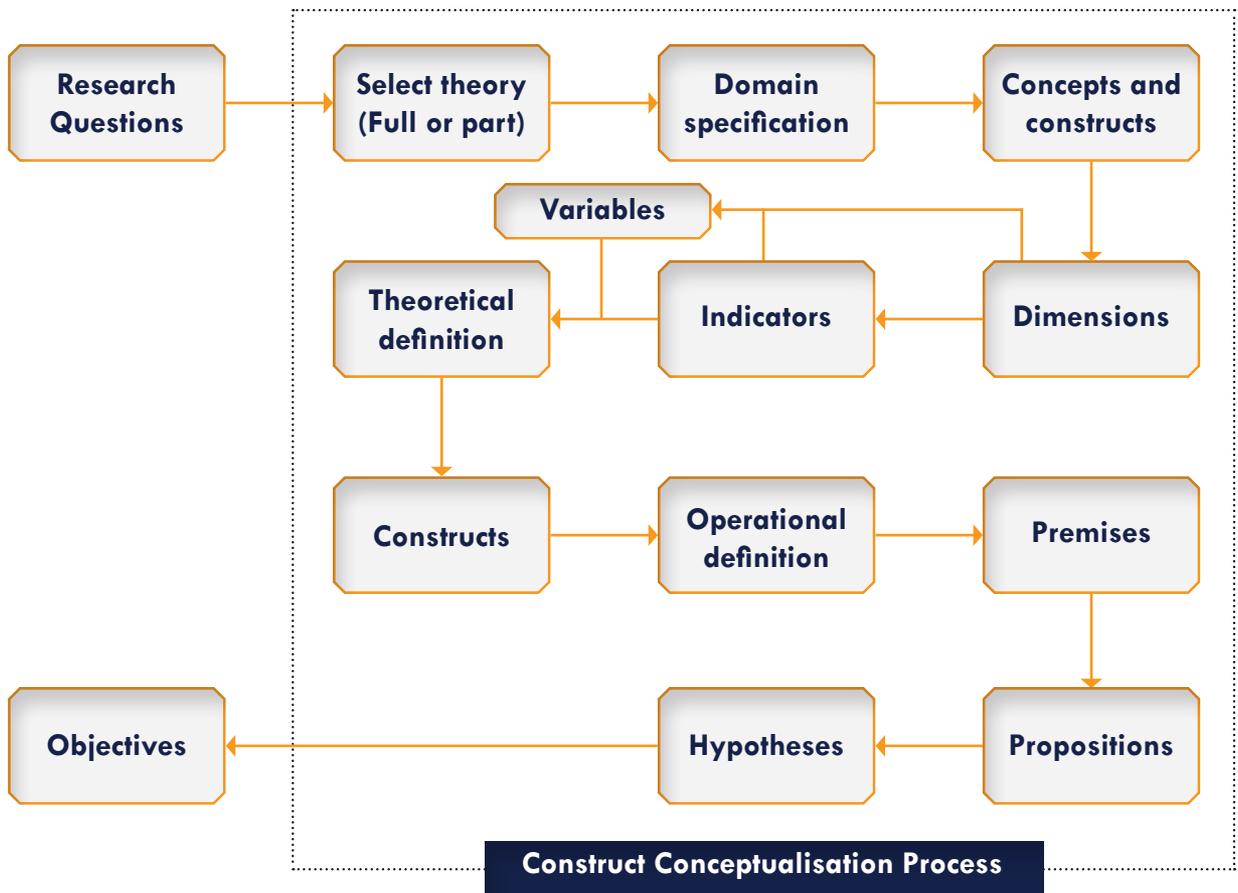


Fig. 4.1: Construct conceptualisation process

When research questions are specified, the researcher begins to look for theory to guide the problem-solving process.

4.2.1 Theory

A theory is a simplified representation of a limited part of the reality (Pawar, 2009). It explains various aspects of the limited part of reality. A theory can be divided into two parts: one that specifies relationships between theoretical constructs; and another that describes relationships between constructs and measures (Bagozzi & Phillips, 1982; Costher, 1969).

4.2.2 Role of theory in extension research

All extension researches are based fully or partly on theory. Both theory and empirical research are interdependent on each other.

- Theory provides the theoretical framework for conducting empirical research
- Research generates the information, that verifies and suggests changes in theory
- Helps to develop useful practice, discovery, explanations and predictions
- A few popular theories which guide extension research are displayed in Table 4.1.

Table 4.1: Popular theories used in extension research

Type of extension research	Theory
Adoption studies	Longitudinal approach - Diffusion theory (Rogers, 1995) Cross-sectional approach - Theory of planned behaviour (Davis, 1989)
Impact assessment	Theory of change (Gertler et al., 2010)
Technology intervention (FLD)	Theory of imitation (Tarde, 1890)
Training effectiveness	Three stage model of planned change (Lewin, 1947)

4.2.3 How research contributes to theory

- Establishing a causal relationship between aspects of reality (e.g., factors influencing Information and communication technology (ICT) use)
- Examination of a specific aspect in a theory (e.g., audience in Sender Message Channel Receiver (SMCR) theory – Knowledge test)
- Method of creating a product (e.g., Instructional Design Theory for Multimedia)
- Constructs development (e.g., communication effectiveness, empowerment)
- Predictive model development (e.g., forecasting of adoption or impact)
- Examination of the impact of an aspect on the nature of the documented problem (e.g. Yield Analysis)

4.2.4 Units of theory

A theory is composed from certain aspects, such as assumptions about the reality, concepts and relationships among concepts (Bacharach, 1989; Pawar, 2009). The units of theory are: (i) domain; (ii) concepts; (iii) variables; (iv) definitional statements; (v) premises; (vi) propositions; and (vii) hypotheses.

(a) Domain

- Domain is a boundary developed by conceptual narrowing or broadening of a phenomenon;
- Represents a limited aspect of reality, which is denoted by theory
- Explains the phenomenon (what, when, and how) in terms of space (geographical locations), time (before and after), and assumptions or values about the nature of the phenomenon;
- An example: Domain - Impact of Front Line Demonstrations (Time)
- A broader domain could be divided into multiple narrower domains, and each of the narrower domains then could have a separate theory focusing on it. E.g., Job satisfaction of extension supervisors and grassroots workers.

Scope of theory: The part of reality delimited by the domain constitutes the scope of a theory.

(b) Concepts

- Concepts are labels or terms formed in the minds of researchers to represent various aspects of the reality (Pawar, 2009).
- Concepts have three components:
 1. The aspect of the reality to be examined;
 2. Its conception in the mind of a researcher; and
 3. The term or label used for it in the vocabulary of science.

Role of Concepts in Theory Building

Concepts specify various aspects of the subject matter or phenomenon that the theory is dealing with. They form a part of “what” of a theory (Whetten, 1989). So concepts:

1. facilitate description and identification of various aspects of the reality;
2. facilitate classification of various occurrences in the reality;
3. provide simplified representation of a complex reality;
4. facilitate a movement from the plane of the empirical world to a theoretical or conceptual plane, and they also facilitate the formation of a relationship between a concept and other concepts.

Kaplan (1964) has categorised concepts into four categories (depicted in Box 4.1) based on (i) degree of directness of the relationship between a concept or a term/label; and (ii) the actual aspect of the reality associated with the concept or term/label.

Box 4.1: Ways of Conceptualizing a Phenomenon (Kaplan, 1964)

CATEGORY 1: Directly observable objects (through direct observation of concrete items); close to reality; **CONCRETE and DIRECTLY OBSERVABLE:** – E.g., rice grain yield.

CATEGORY 2: **CONCRETE and INDIRECTLY OBSERVABLE.** Observable indirectly through complex inferences or operations. E.g., Heartbeats.

CATEGORY 3: **ABSTRACT and INDIRECT OBSERVATION through CONFIGURATION of MULTIPLE OBSERVABLE ENTITIES.** Defines a phenomenon that does not exist as a concrete entity in the empirical world, but the phenomenon can be defined as a configuration of multiple observable entities – Empowerment (income generation capacity, ownership), Food security (availability of food, affordability).

CATEGORY 4: **ABSTRACT and INDIRECTLY OBSERVABLE BY THEORETICAL INTERPRETATION.** Indirect, abstract items which are defined within a theoretical system rather than in reference to any observable entities – E.g., All psychological attributes.

The first three categories of concepts are empirically observable concepts while the fourth is the theoretical or universal concept. Concepts are highly abstract; so they need to be simplified by deriving dimensions (less abstract) and indicators (observable, concrete). Complete conceptualization involves both, specifying dimensions and identifying the various indicators for each.

Construct

Constructs are special types of concepts, which are deliberately and consciously invented or adopted for a scientific purpose (Kerlinger and Lee, 1999). They are abstract and indirectly observable through configuration of multiple observable entities (CATEGORY 3 Concepts), for example: Food security, or empowerment.

(c) Dimensions of the concept

Dimension is a specific aspect of the concept. A concept may have more than one dimension. For example, if food security is identified as a concept, purchase power and food availability are its dimensions.

(d) Indicators

Indicators are concrete and observable forms of a concept which we use to explain the concept. Common in multi-dimensional concepts.

Selected as a variable when dimensions are not available (e.g. uni-dimensional concept) or not observable.

Interchangeability of indicators – When a concept has several indicators, then all of them will behave the same way when the concept is manipulated.

One example: Concept - Food security; Dimensions - purchasing power and food availability; Indicators of purchasing power are monthly income and credit score.

(e) Variables

Variables are empirically observable manifestations of a concept.

Symbol or label that can take different values for different objects depending on the extent to which the property specified by a concept is present in different objects.

Dimensions and indicators are often selected as variables.

(f) Definitional statements

Definitional statements explain the variables in both theoretical and operational terms and assign a label to them.

I. Theoretical definition

- Specifies the verbal meaning attached to the concept label.
- Helps in conveying meaning of the concept without confusion or distortion.
- Is required because the scientific method demands that others understand our theory and be able to criticize and reproduce our observations.

II. Operational definition

- Translates the verbal meaning provided by the theoretical definition into a prescription for measurement.
- Describes the unit of measurement (e.g. minutes, count, weight, height).
- Specifies the level of measurement (e.g. binary, nominal, ordinal).
- Establishes the validity of the measurement.

III. Labels

- Formed based on meanings of both theoretical and operational definitions;
- Facilitate communication – communicate meaning of a concept in an unambiguous fashion.

(g) Premises or assumptions

- Premises are assumptions about the nature of the phenomenon under study.
- Derived from laws of interaction – how the different elements of a phenomenon interact in relation to each other.
- Source – empirical findings coming from past studies, established principles, theoretical views suggested by people, or just assumptions of the researcher, findings from several studies and conjectures in the literature. An example, food availability in the village depends on local production.

(h) Propositions

- Propositions are statements of prediction or statements specifying conjectural relationships between concepts in a theory.
- Logically true statements derived from premises.
- Types of relationships
 - Specifying association – Organisational support is positively related to employee job satisfaction;
 - Specifying causation – Good organisational support results in high job satisfaction.
- One example, food availability in the village is positively related to local production.

(i) Hypotheses

A hypothesis is a conjectural statement about the relationship between variables. Propositions are statements specifying relationships between concepts (part of theory). When the concepts in propositions are replaced with corresponding variables, the resulting statement is a hypothesis (part of the empirical world). An example in point is food availability scores in the village are positively related to local production scores. There are three kinds of hypotheses stated in a research study – Null hypothesis (no difference), alternate hypothesis (research hypothesis), and hypothesis given in question form.

A. Null hypothesis (H₀): Non-directional hypothesis is one which states that no significant difference or relationship exists. It is used as a 'Control' hypothesis which is tested for rejection. It is an initial claim that researchers make based on previous research.

B. Alternate or research hypothesis: Declarative sentence about a phenomenon, which the researcher believes to be true and employs empirical research to prove it. There are three types of research hypotheses – attributive, associative, and causal (Birren, 1981).

(i) Attributive or descriptive research hypothesis: States that something or a phenomenon exists that can be measured and distinguished from another similar thing or phenomenon. It is a Univariate hypothesis which measures only one thing, behaviour or variable. One example, the scale to measure farmers' attitude towards genetically modified cotton.

(ii) Associative research hypothesis: States that a relationship exists between two behaviours – a specific magnitude of change in one kind of behaviour helps to extend change in another, different kind of behaviour. It is a Bivariate hypothesis which deals with two behaviours or variables. For example: the students' aptitude scores in a given test will predict his/her choice of career.

(iii) Causal or explanatory research hypotheses: States that differences in amount or kind of one variable causes/produces/creates/changes differences in amount or kind in another variable. It is a Bivariate hypothesis –changing value of one variable can 'cause', 'produce', 'lead to', 'change', 'increase', 'decrease' the value of the other variable. For example: extent of farmers' participation in the training programme will increase the adoption of Artificial Insemination technology.

The associative and causal research hypotheses can be classified into two types based on the 'direction of relationships; specified by them

(a) Directional hypothesis: Specifies both presence and direction of relationships between variables. The direction of relationship is expressed through comparative adjectives, such as positive/ negative, more/less, increased/decreased, higher/lower, etc. For example: the age of the farmer is negatively correlated with knowledge gain from the training programme (Associative directional hypothesis). Higher group cohesion among members of an SHG will increase its sustainability (Causal directional hypothesis).

(b) Non-directional hypotheses

Indicate presence of relationship between variables, but doesn't specify its direction.

C. Question form hypothesis: In the question form hypothesis, a question is asked as to what the outcome will be, instead of stating what outcome is expected. For example: will the front line demonstration of improved variety in Niger lead to its complete adoption?

(j) Research Objectives

A research objective is a concrete statement describing what the research is trying to achieve (Sharma, 2014). The objectives systematically link the theoretical relationships specified in hypotheses to analytical and methodological procedures required for empirical research (Andrew and Hildebrand, 1976). Research objectives may be of two types – general and specific. While general objectives specify the broad goals to be achieved, specific objectives break them into smaller and achievable activities. The research questions and objectives are written at the end of the Statement of the problem. Tips for writing good research objectives are given in Box 4.2.

Box 4.2: Tips for writing good research objectives

- A well-written objective follows the SMART principle –Simple, Measurable, Attainable, Realistic and Time-bound
- Brief and precise
- Provide a comprehensive view of the research problem and contributing factors of the phenomenon under study
- Stated in operational terms
- Use only action verbs. A few of the action verbs which are commonly used in various research strategies are given below (adapted from Sharma, 2014):
 - Quantitative – Describe, Identify, Assess, Find out, Describe, Compare, Contrast
 - Correlational – Determine, Examine, Identify, Discriminate
 - Experimental – Determine, Examine, Investigate, Measure, Detect
 - Ethnography – Assess, Describe, Examine, Discover, Drive
 - Grounded theory – Develop, Extend, Identify, Validate
 - Phenomenological – Describe, Develop, Generate, Record, Express

4.3 Key Points

- Construct conceptualisation is an important phase of extension research which identifies constructs and variables of the study.
- A theory forms the foundation of conceptualisation and has elements such as (i) domain; (ii) concepts; (iii) variables; (iv) definitional statements; (v) premises; (vi) propositions; and (vii) hypotheses.

References

- Andrew, C.O. and P.E. Hildebrand. (1976.) Planning and conducting applied research. New York, NY: MSS Information Corporation.
- Babbie, E. (2010.) The practice of social research. Twelfth edition. Wardsworth Cengage Learning.
- Bacharach S.B. (1989.) Organizational theories: Some criteria for evaluation. Acad Manage Rev. 14:490–495.

- Bagozzi, R. P. and Phillips, L. W. (1982.) Representing and testing organizational theories: A holistic construal. *Administrative Science Quarterly*, 27:459–489.
- Birren, J. E. (1981.) *Developmental psychology - a life-span approach*. Boston: Houghton Mifflin.
- Blalock, H.M. (1971.) Causal models involving unobserved variables in stimulus-response situations. In H. M. Blalock (ed.), *Causal models in the social sciences* (pp. 335-347). Chicago: Aldine.
- Costner, H. L. (1969.) Theory, deduction, and the rules of correspondence. *American Journal of Sociology*, 75:245–263.
- Creswell, J. W. (2012.) *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* Fourth edition. Pearson.
- Davis, F.D. (1989.) Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 13(3):319-340.
- DeVellis, R.F. (1991.) *Scale development: Theories and applications*. Newbury Park, CA: Sage.
- Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B. and Vermeersch, C. M. J. (2010.) *Impact evaluation in practice*. Washington DC: The International Bank for Reconstruction and Development / The World Bank.
- Kerlinger F.N. and Lee H. (1999.) *Foundations of Behavioral Research*. Fourth edition. Wadsworth Publishing.
- Lewin, K. (1947.) *Frontiers in group dynamics: Concept, method and reality in social science; social equilibria and social change*. *Human Relations*. 1:5–41.
- Messick, S. (1995.) Validity of psychological assessment. *American Psychologist* 50:741–749.
- Pawar, B. (2009.) *Theory building for hypothesis specification in organizational studies*. Thousand Oaks, CA: Sage.
- Rogers, E. (1995.) *Diffusion of innovations*. Fourth edition. New York, NY: The Free Press.
- Sharma, S. (2014.) *Nursing research and statistics*. Fourth edition. Elsevier.
- Tarde, G. (1890.) *Les lois de l'imitation*. Paris: Félix Alcan.
- Whetten, D.A. (1989.) What constitutes a theoretical contribution? *Academy of Management Review* 14:490–495.

05

Using Thematic Analysis for Construct Conceptualisation

P Sethuraman Sivakumar

Objectives

1. Introduce thematic analysis as a tool for construct conceptualisation
2. Explain the method for identification of themes using thematic analysis

5.1 Introduction

In general, the conceptualisation process in extension research is largely based on past studies, theses in the department, judges' rating and from field experiences. This approach suffers due to the following aspects:

- Stereotypic;
- Irrelevant to context;
- Mis-specification and over identification;
- Lack of field application.

Thematic analysis is a qualitative approach used for construct conceptualisation. It is hugely popular in management research. This module describes thematic analysis as a method/tool to identify themes.

Thematic analysis

Thematic analysis is a qualitative method for identifying, analysing, and reporting patterns within data (Barun and Clarke, 2006). These patterns are termed as 'themes' which represent the dimensions of the phenomenon under study. For example, if a researcher wishes to understand the dimensions of 'learning effectiveness of a multimedia module', he/she conducts thematic analysis to identify the dimensions/aspects of learning effectiveness. From the analysis, the researcher can identify the various dimensions, such as knowledge gain and symbolic adoption or intention to use new learning. These dimensions are essentially themes derived from the data. In other words, these themes are variables identified for representing the latent construct 'learning effectiveness'.

Advantages

- Flexible – A wide range of analytical techniques can be employed to identify hidden patterns in the data.
- Participatory approach – Can be integrated with participatory methods like in-depth interviewing, key informant interviews, etc.
- Simplifying complex data - Can be used to summarise key features of a large body of data.

- Low cost – Themes can be generated from secondary sources, such as published interviews, audio-taped conversations, etc.

5.2 Discussion

Steps in thematic analysis

Step 1: Define the research problem and question

Select the research problem and develop a clear, concise and sound research question (See the example in Box 5.1).

Box 5.1: An Example of defining the research problem and research question

Research problem: An ATARI Director wishes to computerize the work environment of all the KVKs in their Zone by providing computers to all SMS. The purpose of this effort is to improve the work efficiency of the SMS in performing extension and office tasks. Since computerization is an expensive task, the Director wishes to know how the computerization will be effective when it is done on a full scale. Precisely, he wishes to know the factors that will influence computer utilization by SMS. This task is assigned to an SMS (Extension) who is working in Agricultural Technology Application Research Institute (ATARI).

Research question: What are the factors which influence computer utilization behavior of SMS in the KVK?

Step 2: Decide on the sources of data

The data can be collected from the following primary and secondary sources:

- Audio/video recorded personal interviews;
- Audio-recorded telephonic interviews;
- Blogs on a specific topic;
- Social media posts on a specific topic;
- Case studies/success stories;
- Newspaper reports;
- Other audio or video recordings of events, views, etc.;
- Reports;
- Policy documents;
- User feedback forms.

Step 3: Collect data in a systematic way

The data will be collected in a systematic way from:

- Personal and telephonic interviews – Follow sampling, develop interview guide, pre-test and collect data;
- Audio and/or video recordings – Select the sources, develop screening criteria for choosing recordings and decide the final sample;
- Blog and social media posts – Sources (e.g., Facebook or Twitter or other), select time period, select search key words, screening criteria for selecting specific posts;
- News reports – Select media monitoring programme (e.g. Clipit), key words, screening criteria for selecting clippings;
- Case studies/success stories – Sources, key words, screening criteria for selecting cases;

- Reports and policy documents – Select reports and policy documents on a specific topic or on a similar topic and time period.

After systematically collecting the data, a data corpus containing all the information gathered for the thematic analysis is prepared.

Step 4: Transcription and translation of verbal data

Transcription

Transcription is the process of converting the verbal data i.e., interviews, audio/video clips and speeches into written form of the same language (Barun and Clarke, 2006). Thematic analysis requires a rigorous and thorough 'orthographic' transcript – a written form of all verbal and non-verbal (e.g., coughs, time taken to answer, etc.) information in verbatim. A few sample transcripts of an indepth-interview conducted to study the computer use behaviour example cited in Step 1 are given in Box 5.2.

Box 5.2: Sample transcripts from an audio-recorded indepth personal interview

Verbatim 1: "It is a good idea to have computer in my desk and I would love to use it. Though I have a laptop at my home, my son always uses it and I hardly get time to work with it. If it is in my office desk, I can try whenever I get free time and explore new applications. When I was in (National Institute of Agricultural Extension Management) MANAGE, Hyderabad, last month, I learned how Facebook and Twitter can be used to create farmers' groups and exchange information on new agricultural technologies. Though I am not good in managing Facebook (FB), I can get help from colleague SMS (Horticulture), who has an active FB account. I know computers can help to improve the extension work, I am still ignorant on how it will be used for field-oriented programmes." (Dr Ganesh Sharma, SMS Plant Protection, Age 43)

Verbatim 2: "I don't know how computerisation will help in improving the extension work. As we know, the computers are very expensive and providing computers to everyone need huge investment. Some of my colleagues are obsessed with computers and I don't know what they do when sitting with their laptops for hours. The Directors always want to impress their bosses and do something which improves their visibility. At my personal level, I am no good with computers and don't think computerisation will do any good for the KVKs."(Mr. Sushil Kant, SMS, Horticulture, Age 53)

Translation

In this phase, the researcher translates the transcribed verbatim into English or any other language in written form. It is usually conducted after themes and indicators are identified (after Step 10). The steps in translating qualitative data from the original language into English (Chen and Boore, 2009) are described below:

1. Verbatim transcription of the content in original language and then conducting thematic analysis;
2. After the concepts and indicators are identified, two bilingual translators will translate them into English anonymously and the final English version is developed from two independently translated scripts through mutual agreement between both translators;
3. Back translation - A bilingual person takes the English version and back translates the concepts and indicators from English to the original language;
4. An expert panel committee examines all the drafts and decides on the final draft of translated material.

Step 5: Familiarising yourself with your data

- Take a quick glance at the verbatim/transcripts as a whole and take notes from your first impressions;
- Read the verbatim/transcripts thoroughly line-by-line by looking for meanings and patterns from the data;
- If the data is already transcribed, the researcher has to spend more time familiarising with the data, and must also check the transcripts back against the original audio or video recordings for accuracy.

Step 6: Generating initial codes or labels

After initial reading, identify the data extracts or specific word(s) which represent a dimension of the research problem. For example, the Verbatim 1 displayed in Box 5.2, the words “... I would love to use it ...” represents a data extract as it indicates the respondents’ ‘intention to use the computer’. A few guidelines for selecting data extracts for coding are given in Box 5.3 below.

Box 5.3: Guidelines for selecting data extracts for coding

What to code or label

- Repeated in several places (Repetition)
- Surprises you (Newness)
- The interviewee explicitly states that it is important (Importance)
- You have read something similar in previously published papers, reports (Familiar)
- It reminds you of a theory or concept (Familiar)
- Something you think which is relevant to the research question (Relevant)

Other points

- Code for as many potential themes/patterns as possible
- Stay unbiased
- Code extracts of data inclusively – code along with the context
- Remember that you can code individual extracts of data in as many different ‘themes’ as they fit into – so an extract may be un-coded, coded once, or coded many times, as relevant.

- Coding or labelling or indexing – Involves the production of initial codes or concise labels for important features of the data, based on the research question. Codes are labels given to a group into which the related data extracts are grouped. Codes are essentially indicators of the concept.
- Please highlight or underline the specific data extracts while reading the transcript. The researcher can also write them down in a separate notebook for grouping in the later stages.
- After identifying the data extract, assign a code to it based on your perception of what it signifies. For example, “...its a good idea...” in the first verbatim may be coded as having ‘Positive attitude’ towards computer use. The data extracts may be coded and tabulated as shown in Table 5.1.
- Group all the codes along with relevant data extracts and prepare a long list of codes and data extracts.

Table 5.1: Tabulation of codes against each verbatim

Verb. No.	Verbatim/ specific transcript	Initial codes against data extract(s)
1	<p>“ it is a good idea to have computer in my desk and I would love to use it. Though I have a laptop at my home, my son always uses it and I hardly get time to work with it. If it is in my office desk, I can try whenever I get free time and explore new applications.</p> <p>When I was in MANAGE, Hyderabad last month, I learned how Facebook and Twitter can be used to create farmers’ groups and exchange information on new agricultural technologies. Though I am not good in managing Facebook, I can get help from colleague SMS (Horticulture), who has an active FB account... I know computers can help to improve the extension work, I am still ignorant on how it will be used for field oriented programmes.....” (Dr. Ganesh Sharma, SMS – Plant Protection, Age 43)</p>	<p>Positive attitude Intention; Availability Problem Availability</p> <p>Action Knowledge</p> <p>Support Positive attitude</p>

Step 7: Searching for themes

- This phase involves sorting the identified codes into preliminary levels of themes based on their perceived closeness, and pooling all the relevant coded data extracts into the identified themes.
- A theme is a coherent and meaningful pattern in the verbatim/transcript relevant to the research question. It is technically a construct or a dimension of the construct.
- Visual techniques, such as mind maps, tables and cards may be used to pool the relevant codes into preliminary themes. A theme piling method is described in Box 5.4.

Box 5.4: Theme-piling method

- Write all the initial codes identified on separate cards.
- Have a look at all the codes and try to find
 - Codes that can be combined
 - Irrelevant codes – outliers.
- Based on the relationship between codes, compile them into themes and sub-themes and include the codes and data extracts into each main and sub-theme.
- From the given example, the codes ‘intention’ (symbolic adoption) and ‘actual use’ (adoption) represent two dimensions of adoption, so these can be kept under the category ‘adoption’.

- At the end of this phase, the preliminary themes and sub-themes, and all extracts of data that have been coded in relation to them, are identified and plotted as in Table 5.2.

Table 5.2: Preliminary themes, sub-themes and codes

Theme/ Construct	Sub-theme/ dimensions	Codes/ indicators	Related data extracts
1	1 2		

Theme/ Construct	Sub-theme/ dimensions	Codes/ indicators	Related data extracts
2	1		
	2		
3	1		
	2		

- Develop a mind map showing the relationship between themes, sub-themes and codes as displayed in Fig. 5.1.

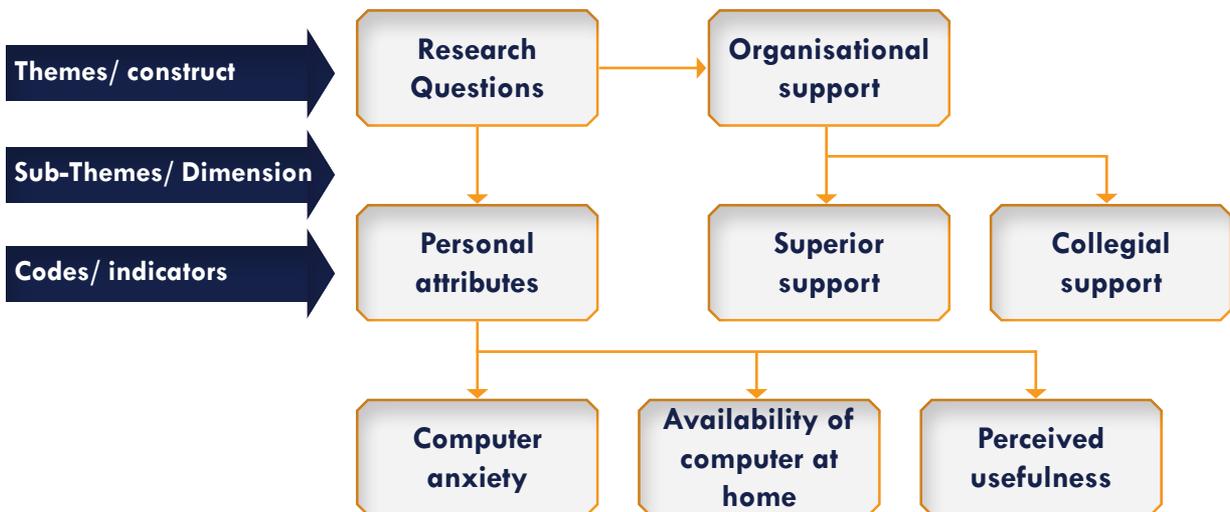


Fig 5.1: A sample mind map showing relationship between themes, sub-themes and codes

- Do not abandon anything at this stage, as without looking at all the extracts in detail (next step) it is uncertain whether the themes can stay as they are, or whether some need to be combined, refined and separated, or discarded.

Step 8: Reviewing themes

Involves reviewing and refining themes. Based on Patton's (1990) dual criteria for judging categories – internal homogeneity and external heterogeneity. This step involves two levels of reviewing and refining your themes.

Level 1: Reviewing at the level of the coded data (Reviewing the codes and data extracts)

- Read all the collated codes and respective data extracts for each theme and sub-theme to check if the data forms a coherent pattern.
- If the main and sub-themes do not fit, you should rework your theme, create a new theme, and find a home for those extracts that do not fit.

Level 2: Overall reviewing of themes with the data set

- Consider each theme in relation to your data corpus.
- Thematic map – Do the relationships between the themes reflect the meaning of your data as a whole? If not, return to Step 7.
- If they do, move on to Step 9 (defining and naming themes).

Step 9: Defining and naming themes

- Capture the essence of what each theme is about and what aspect of the data each theme captures.
- Examine the themes, and sub-themes carefully to see that they are coherent and internally consistent.
- Give names for each theme – The names must be concise, punchy and immediately give the reader a sense of what the theme is about.
- Draw the final thematic map and describe each theme/sub-theme in one or two sentences.
- At the end of this phase, the final thematic map with clearly defined themes is ready.

Step 10: Using the themes

- The themes are essentially the latent factors which represent a specific dimension of the problem. They are constructs of the study.
- The sub-themes are dimensions of the constructs.
- The codes are ‘indicators’ or ‘scale items’ which essentially represent the variable along with several others.
- Both dimensions and indicators are designated as variables.
- Once the themes, sub-themes and codes are available, subject them to a content validity assessment by 10-20 trained experts who will examine them along with all verbatims to check if the themes are clearly etched out.
- A sample of themes/constructs, sub-themes/dimensions and codes (indicators or scale items) are given in Table 5.3.

Table 5.3: Results of thematic analysis of computer use behavior

Construct/ Theme	Dimensions/ Sub-themes	Indicators/ Scale items/ codes	Actual statements
Perceived use- fulness	-	Increase in productivity	Using computers in my job would increase my productivity
		Easy to do job	Using computers would make it easier to do my job
		Better decision	Computers could provide me with information that would lead to better decisions
		Enhance professional image	Using computer gives me the opportunity to enhance my professional image
Organisational support	Superior sup- port	Organising work	Computer technologies can be used to assist me in organizing my work
		Encouragement to use computers in job	My superiors encourage me to use computers in the job
		Deputing for computer training	My superiors support me to undergo computer-related training
		Expectation to use computer	My superiors expect me to use computers as an extension tool
		Providing adequate financial support for computers	My superiors do not provide adequate financial support for purchase/repair of computer facilities

Construct/ Theme	Dimensions/ Sub-themes	Indicators/ Scale items/ codes	Actual statements
	Collegial support	Sharing computer facilities Recognising computer expertise Encouragement to use computer in extension	My colleagues are not willing to share their computer facilities with me My colleagues recognize my computer expertise My colleagues encourage use of computers in extension

5.3 Key Points

- Thematic analysis is a popular tool for generating themes from users' perspectives.
- The user/respondent interviews generate verbatim which contain statements related to the overall theme.
- The relevant statements/data extracts are selected, grouped and coded using representative labels.
- Related codes are grouped into sub-themes and themes using the mind-mapping technique.
- After repeated review of classification of themes, sub-themes, codes and data extracts, they are compiled into a questionnaire for validity assessment and data collection.

References

- Babbie, E. R. (2008.) The basics of social research. Belmont, CA: Thomson/Wadsworth.
- Braun, V. and Clarke, V. (2006.) Using thematic analysis in psychology. *Qualitative Research in Psychology* 3(2):77–101.
- Chen, H-Y. and Boore, J.R.P. (2009.) Translation and back-translation in qualitative nursing research: Methodological review. *Journal of Clinical Nursing* 19:234–239.
- Patton, M. Q. (1990.) *Qualitative evaluation and research methods*. Second edition. Newbury Park, CA: Sage.

06

Sampling Methods for Quantitative Extension Research

A Dhandapani

Objectives

1. Introduce sampling methods used in extension research
2. Explain the probabilistic and non-probabilistic sampling methods

6.1 Introduction

Sampling is a crucial step in scientific extension research. It involves identifying the units of analysis – individuals, groups or communities in specific geographical or professional domains. Selecting a representative sample in a systematic way is an essential prerequisite for conducting quality research. This module provides an overview of sampling techniques commonly used in extension research.

6.2 Discussion

6.2.1 Representative sampling

- Sampling is a procedure by which a part of a population is used for representing the whole population.
- Selecting representative sample is the key to conducting quality research.
- Representativeness is the quality of a sample having the same distribution of characteristics as the population from which it was selected. However, obtaining a true representative sample is difficult.
- A few terms related to sampling are given in Box 6.1.

Box 6.1: Some terms related to sampling

- Sampling Frame – That list or quasi list of units composing a population from which a sample is selected
- Sampling Units – Unit on which observations are collected
- Sample Size – Number of sampling units selected for observation
- Population Size – Total Number of units in the population
- Target Population – Population of interest
- Sampled Population – Population from which samples were collected
- Parameter – Summary description of a given variable in a population
- Statistic – Summary description of a variable in a sample.

6.2.2 Sampling strategy

- Strategy is the plan devised by the researcher to ensure that the sample chosen for the research work represents the selected population.
- Choosing an appropriate sampling strategy is a key aspect of the research design.
- Robinson (2014) proposed a four-point sampling process for systematically selecting adequate samples for obtaining quality results (Fig. 6.1).

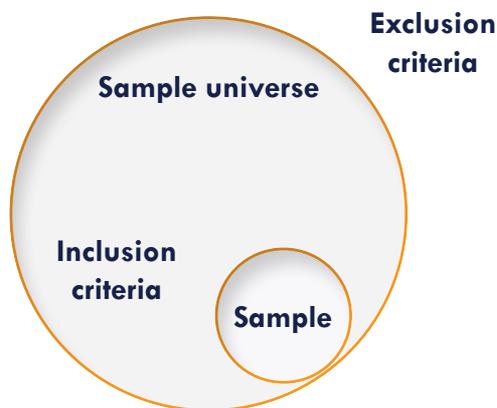


Fig. 6.1: Sampling strategy (Adapted from Robinson, 2014)

1. Define a sample universe

Establish a sample universe, by specifying a set of inclusion and/or exclusion criteria.

- Inclusion criteria – attribute(s) that respondents must possess to qualify for the study;
- Exclusion criteria – attributes that disqualify a case from the study.

For example, in a research investigation focusing on the 'Information source utilisation of Ber growers', the inclusion criteria is 'Ber grower' (Current/ past specified in years), while the exclusion criteria is 'growers of other crops'. During the selection, the homogeneity of the samples i.e., demographic (e.g., youth), geographical (e.g., Maharashtra or Tamil Nadu), physical (e.g., female workers), psychological (e.g., progressive farmers), and life history (e.g., migrant workers) should be considered.

2. Deciding on sample size

- Based on theoretical and practical considerations.
- Theoretical considerations for quantitative studies – nature of problem, population size and the type of analytical strategies used; while qualitative investigations focus on the saturation and redundancy of the data collection methods (Robinson, 2014).
- Practical aspects – time and resource availability, researcher capability and purpose of research work (e.g., for dissertations or sponsored research).

3. Selecting a sample strategy

The popular sampling methods in quantitative research are probabilistic and non-probabilistic sampling, while qualitative research uses random/convenience sampling and purposive sampling strategies. After deciding on the sampling strategy, the respondents required for each sample category (e.g., strata) is decided from the overall sample size.

4. Sourcing sample

When the sample universe, size and strategy are decided, the researcher needs to recruit the participants from the real world. Voluntary participation, recruiting students from the subject pools, advertising in social and print media for recruiting community members, online surveys with jackpot provisions are a few ways for recruiting participants for research work. In this phase, the researcher should follow ethical guidelines (if suggested by the ethics committee) in advertising, selection and handling participants, confidentiality of research data, compensating participants for their time and effort, etc. Extension research in India is conducted without following any ethical practices as suggested by various Human Subject Research regulatory agencies. A detailed description of Human Subject Research Guidelines can be found in the Module 18.

6.2.3 Types of quantitative sampling strategies

There are two types of sampling procedures commonly used – probability and non-probability samples (Fig. 6.2).

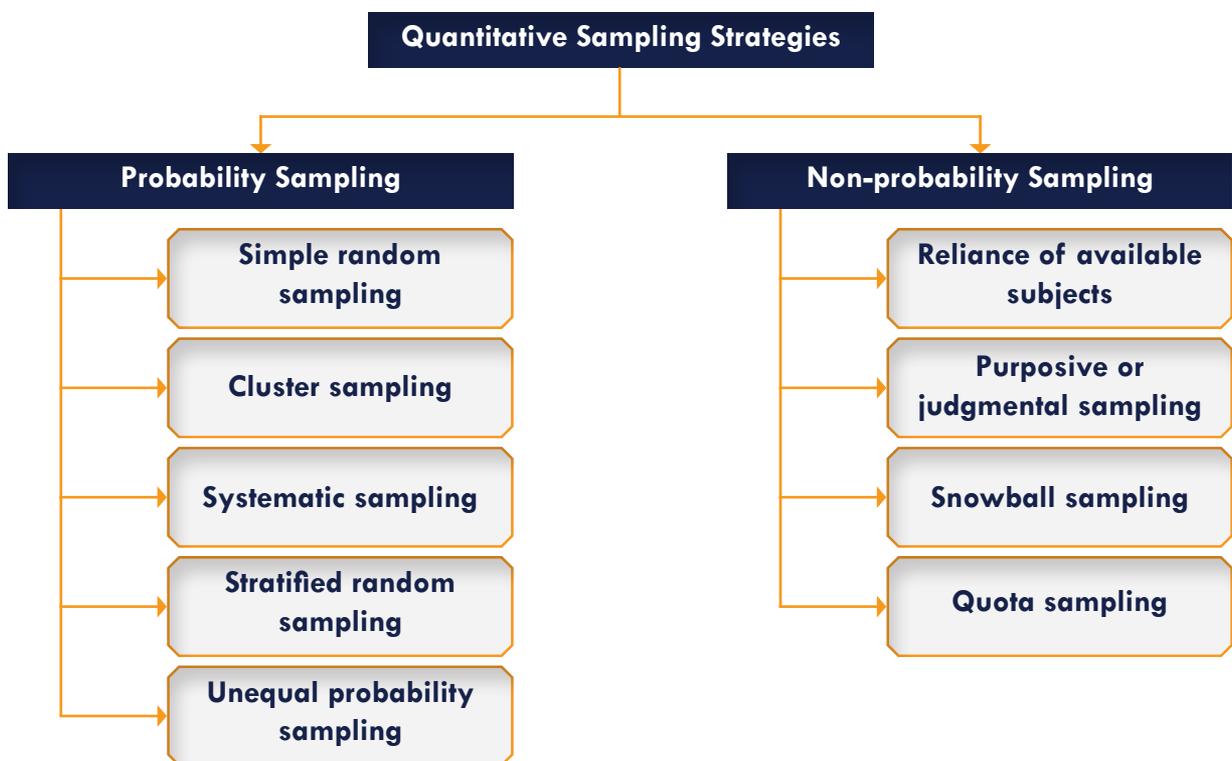


Fig 6.2: Sampling strategies used in extension research

Probability samples

- In probability samples, each unit in the population has known the probability of being selected in the sample.
- Most of the statistical techniques require probability sampling.
- Most rigorous form of sampling and researcher can make generalizations to the population.
- Application of statistical theory possible.
- Provides estimates of 'Error due to sampling'.

Equal probability sampling techniques

- A sample design in which each member of a population has the same chance of being selected into the sample (Babbie, 2010).
- The easiest sampling techniques one can use is to provide equal probability to every unit in the population.
- This is intuitive as there is no reason to prefer one unit over another unit, and all units are equally preferable.
- Some of the commonly used equal probability sampling techniques are: Simple Random Sampling and Systematic Sampling.

A few principles to keep in mind

(a) Sampling frame or target population

- A group of individuals or groups with some common defining characteristic that the researcher can identify and study.
- For example, to select villages from a district, list all villages in that district in the sampling frame.
- Required for probability sampling methods.
- Sampling frames – must be complete and up to date.
- Population census provides list of villages.

(b) With or without replacement of samples

- Without Replacement – Once a unit is selected, it is not considered for selection again.
- With Replacement – the same unit may be selected more than once.
- In general, Without Replacement (WOR) is efficient as compared to With Replacement (WR) Sampling.

(c) Sampling error

- The degree of error to be expected in probability sampling.
- Arises due to observing only part of the population.
- Lower the sampling error, better the estimates.
- Depends on
 - Variability in the measuring parameter;
 - Sample size (most of the time).
- The formula for determining sampling error contains three factors – the parameter, the sample size and the standard error.

(i) Simple Random Sampling

When to use:

- Large population ($N > 30$); population size is known;
- Not suitable for small samples ($N < 30$).

What: Select samples from a population of known size randomly

Procedure: Develop a sampling frame or target population. Assign a number to each individual in the population and then use a random numbers table to select the individuals for the sample.

Example: A student researching job satisfaction of extension professionals in a district collects the full employee list and applies a random number algorithm to this in order to select people to interview.

Pros:

- Unbiased: Each unit has the same chance of being selected. Simple; and provides an unbiased estimate of population parameter.
- Independent: The selection of each unit is not affected by the selection of other units.
- Easy to calculate sampling error.

Cons:

- Unmanageable for a very large population.
- When compared to other sampling methods, such as stratified sampling or unequal probability sampling, simple random sampling may produce less efficient estimators.

(ii) Cluster Sampling or Area Sampling

When to use: When the studied population is spread across a wide area such that a simple random sampling would be difficult to implement in accessing the selected sample.

What:

- Random sample of groups of units (clusters) are selected.
- Grouping of nearby units into clusters.
- Select 'n' clusters randomly from the total number of clusters.
- Once a cluster is selected, all the units within the cluster are selected automatically and information is collected from all the units.

Example: Adjoining two or three villages may be considered as a cluster and once a cluster is selected all villages within that cluster are selected.

Pros: More sampling units at a lower cost.

Cons: Less precise.

(iii) Systematic Sampling

When to use: When a given population is logically homogenous and size is unknown.

What: Selecting every nth individual or event in the population until the desired sample size is reached.

How:

- Decide on the required sample size (n).
- Calculate the sample interval by dividing total number of persons or items in the population (N) by required sample size (n). The sampling interval (k) indicates the distance in which the other sample is selected.
- Arrange the population into a sequential order, ensuring the attribute being studied is randomly distributed.
- Select a random number, x, between 1 and k.
- The first sampled item is the xth. Then select every kth item.

Example: Consider a situation wherein it is expected that 20 houses will be selected from 100 houses. To select the sample using systematic sampling, first divide the population size by sample size, i.e.

100/20 which is 5. Select a random number between 1 to 5 randomly. Suppose, the selected number is 4. Then select every 5th house starting from house number 4.

Pros: No need of sampling frame, only total number of households is required.

Cons: When the population has a trend phenomenon with respect to the study variables, systematic sampling will produce less efficient estimates. It is not possible to provide a measure of sampling error.

(iv) Stratified Random Sampling

When:

- When there are smaller sub-groups that are to be investigated.
- To achieve greater statistical significance in a smaller sample.
- To reduce standard error.

How:

- The population is divided into different groups called strata wherein units within each stratum are homogenous.
- Sampling is carried out in each stratum independently using simple random or any other sampling procedure.
- In case simple random sampling is followed in all strata, then the method is called stratified random sampling.

Types:

- Proportionate stratified sampling takes the same proportion from each stratum.
- Disproportionate stratified sampling takes a different proportion from different strata.

Example: Some of the commonly used strata in Indian conditions include districts, agro-ecological zones, etc.

Pros: Stratified Random Sampling can produce most efficient estimates as compared to simple random sampling or cluster or systematic sampling procedure. Each stratum must be homogeneous.

Cons: Complex – The criteria may be difficult to fulfil and it places a heavy strain on resources.

(v) Unequal Probability Sampling

When to use: Used when the populations of sampling units vary in size

What:

- When all sampling units have an auxiliary variable which is closely related to main variable, it is used as a measure of the size of the unit.
- These units are sampled with probability proportional to size measure.
- The probability of being selected for different units are unequal.

Example:

- In agricultural surveys, area under a crop may be taken as a size measure of farms for estimating the yield of crops.
- Villages growing the crop on a large area should be selected with higher probability than the others.

Pros: High precision. Used in selection of initial stage units.

(vi) Multi-stage Sampling

When to use: Used in large-scale surveys under limited resource situation

What: The sampling is carried out in stages using smaller and smaller sampling units at each stage. In a two-stage sampling design, a sample of primary units is selected and then a sample of secondary units is selected within each primary unit.

How:

- The whole population is divided into first stage sampling units from which a random sample is selected.
- The selected first stage is then subdivided into second stage units from which another sample is selected.
- Third and fourth stage sampling is done in the same manner, if necessary.
- The stages are decided based on the nature of research and the size of the population under study.

Example: First stage – Village; Second stage – Fields within selected village; Third stage –Plants within the selected fields – Fourth Stage – Leaves within the selected plants, etc.

Pros: Cost effective and time saving.

B. Non-probability Sampling

Samples are selected based on subjective criteria – easily available, convenient, and represent some characteristic the investigator seeks to study. Selecting samples but not randomly.

(i) Reliance of Available Subjects/Availability/Accidental Sampling

When to use: When the resources and time are limited

What: Selecting and collecting data from people at a street corner or malls or villages based on researchers' subjective criteria.

Pros: Cost effective

Cons: Inaccurate. Biased selection of samples will result in inefficient estimates.

(ii) Purposive or Judgemental Sampling

When to use: When the researcher is confident of his/her ability to judge the target population to select a representative sample. When time and resources are constrained.

What and how: The sample is selected on the basis of the researcher's judgment about which ones will be the most useful or representative.

Pros: Simple, cost effective

Cons: Subject to bias and error

(iii) Snowball Sampling

When to use: Use when the researcher does not have access to sufficient people with the desirable characteristics.

What: The researcher asks participants to identify others to become members of the sample till he/she achieves the desirable sample size.

Example: When a researcher wishes to study the economic viability of duck farming, he/she could get details from only five farmers. The researcher asks these farmers if they know any more. From the information gathered from respondents, researcher identifies prospective samples.

Used in adoption studies

Pros: Easy, cost effective

Cons: Biased estimates

(iv) Quota Sampling

When to use: Used when researcher knows the proportions of particular sub-groups within a population and want to ensure that each group is proportionately represented.

How: Sampling using any method until the required sample size is reached

Pros: Cost effective

Cons: Biased estimator

6.3 Key Points

- Sampling is a procedure by which part of a population is used for representing the whole population.
- Selecting representative sample is the key to conducting quality research.
- There are two types of sampling procedures commonly used – probability and non-probability samples.
- In probability samples, each unit in the population has known probability of being selected in the sample. Most of the statistical techniques require probability sampling.
- The non-probability sampling uses a deliberate approach to select representative samples based on subjective criteria.
- A good sample, selected through a systematic way using probability sampling will produce sound estimates, which aid in advancement of the discipline.

References

- Babbie, E. (2010.) The practice of social research. Twelfth edition. Wardsworth Cengage Learning.
- Robinson, O.C. (2014.) Sampling in interview-based qualitative research: A theoretical and practical guide. *Qualitative Research in Psychology*, 11:25–41.

07

Sampling Methods for Qualitative Extension Research

P Sethuraman Sivakumar

Objectives

1. Provide an overview of various sampling strategies followed in qualitative research
2. Provide guidelines for sample size for qualitative research

7.1 Introduction

Sampling is an important component of qualitative research design which determines validity and generalizability of the research findings. While quantitative research seeks to analyse a representative sample to make a generalisation for a population, qualitative research aims for in-depth understanding of subjects. Qualitative extension research studies rely heavily on participatory methodologies and focus group discussions involving various stakeholders.

7.2 Discussion

7.2.1 Sampling Strategies

Sampling strategies for qualitative research differs from quantitative methods due to differing goals.

1. Goals: While quantitative research focuses on 'empirical generalisation of results to many in the population' qualitative studies are concerned with achieving 'in-depth understanding' of the phenomenon.
2. Nature of samples: Since qualitative studies focus on the 'quality' of information gathered through a variety of approaches, the samples are called 'participants or informants' rather than 'subjects' as done in quantitative studies. Sampled cases are incidents, slices of life, time periods, or people.
3. Nature of generalisation: In any empirical study, there are three ways to generalise the results to the population (Curtis et al., 2000; Firestone, 1993). Types of generalisations in empirical studies is depicted in Fig 7.1.

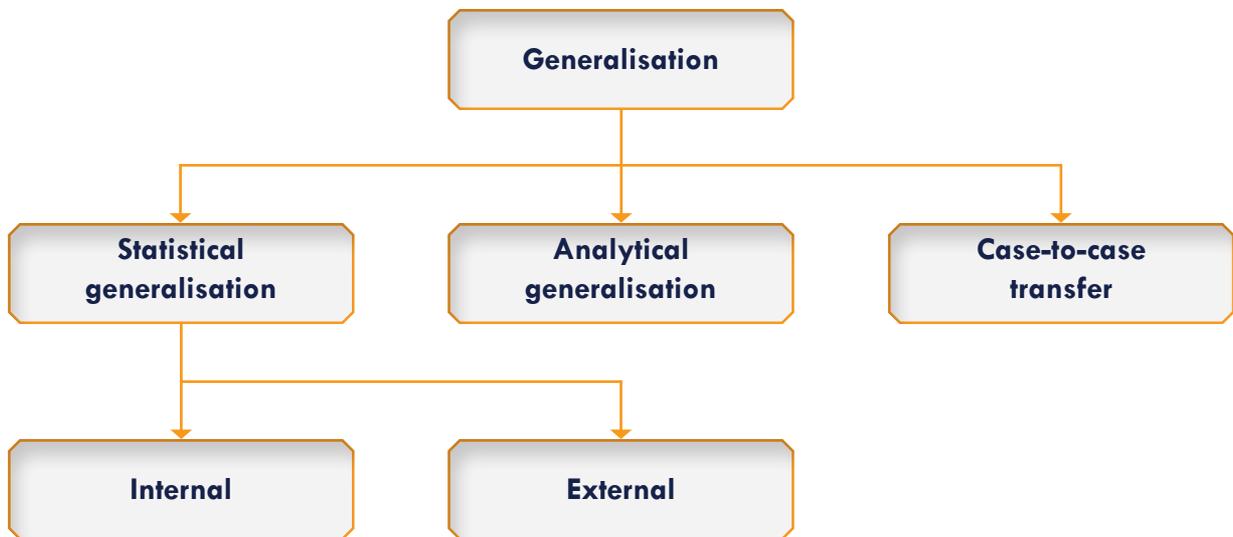


Fig 7.1:Types of generalisations in empirical studies (Adapted from Onwuegbuzie and Leech, 2007)

(a) Statistical generalisation – Making generalizations or inferences on data extracted from a representative statistical sample to the population or samples itself.

- External generalisation – Making generalisations about the population from which the sample is drawn;
- Internal generalisation – Making generalizations from the data extracted from one or more representative participant to the sample from which the participant(s) was drawn.

(b) Analytical generalisation – Making projections about the likely transferability of findings from the empirical study, based on a theoretical analysis of the factors producing outcomes and the effect of context (Macfarlan, 2015).

(c) Case-to-case transfer – Making generalizations from one case to another but a similar case (Firestone, 1993).

While quantitative research focuses on external statistical generalisation, qualitative research is concerned with internal statistical generalizations, analytical generalizations and case-to-case transfers.

Sampling strategies

- Since qualitative research is an in-depth understanding of a phenomenon or units for the purpose of analytical/internal statistical or case-to-case transfer generalisations, only non-probability sampling strategies are used.
- Details of non-probability sampling strategies, such as reliance of available subjects, snowball sampling, and quota sampling, are discussed in Module 6.
- Daniel (2012) has identified four traditions of non-probability sampling:
 1. Availability sampling;
 2. Purposive sampling;
 3. Quota sampling;
 4. Respondent assisted sampling.

A. Availability sampling: The elements are selected from the target population on the basis of their availability, convenience of the researcher, and/or their self-selection (Daniel, 2012). One example is when a researcher requests farmers to fill up a questionnaire for measuring their attitude towards eco-friendly technologies.

Advantage – Most frequently used sampling procedure in research - least time-consuming, least expensive, and the least complicated sampling procedure.

Disadvantage – Works well only with a homogenous sample.

Types

1. Convenience sampling – Based on the convenience of the researcher and the informant;
2. Haphazard sampling – Selected in a haphazard manner;
3. Accidental sampling – Cases selected accidentally;
4. Chunk sampling – Documents or records are studied and then the researcher uses a chunk of data;
5. Grab sampling – Grabbing persons for inclusion in the study;
6. Straw polling – This is done with persons who are available at a meeting, or contacted via advertisements placed on bulletin boards, web sites, or in newspapers requesting persons to volunteer to participate in a study.

B. Purposive sampling: Purposive sampling and its variations are widely used in qualitative research. It is a non-probability sampling procedure in which elements are selected from the target population on the basis of their fit with the purposes of the study and specific inclusion and exclusion criteria (Daniel, 2012). In purposive sampling, the informants are not selected simply on the basis of their availability, convenience, or willingness, but the researcher purposely selects the elements based on the inclusion and exclusion criteria for participation in the study.

Types

- Based on the sample inclusion and exclusion criteria, purposive sampling can be classified into several sub-types.
- The criteria for inclusion and exclusion tend to fall into four major categories (Daniel, 2012):
 1. Criteria based on central tendency;
 2. Criteria based on variability;
 3. Criteria based on theory/model development;
 4. Criteria based on judgment and reputation.

Various sub-types of purposive sampling based on the above criteria are displayed in Fig 7.2.

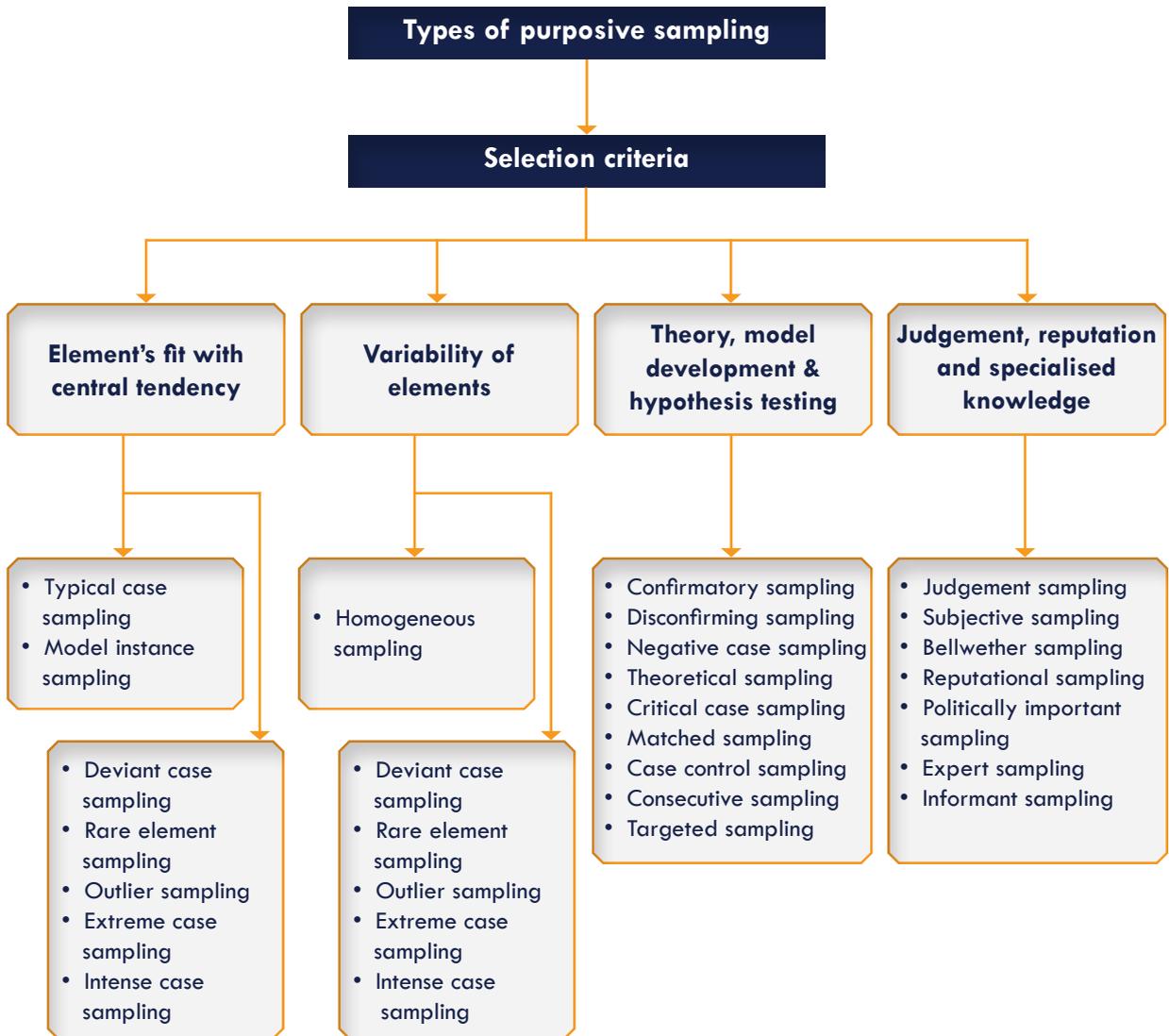


Fig 7.2: Types of purposive sampling (Adapted from Daniel, 2012)

1. Based on central tendency

Inclusion/exclusion criteria – Informants who represent either average or extreme cases of the population.

(i) Typical case and model instance sampling

- Selecting cases which display or are assumed to have average or typical level of a particular characteristic of the population (Daniel, 2012);
- For example, when frontline demonstrations of improved tomato variety are conducted in a few selected places in a village in Manipur, it would be difficult to reach out to the beneficiaries living in remote areas. In such cases, the researcher can choose cases which are representative and accessible for collecting data.

(ii) Extreme/deviant/rare element/outlier case sampling

- The researcher selects atypical, special or unusual cases, which produce significant outcomes,

- success/failures with respect to the phenomenon under study (Daniel, 2012);
- A few rare elements which occur only in a certain time period are also selected;
- Provides significant insight into a particular phenomenon which can act as a lesson that guides future research and practice;
- For example, in diffusion theory laggards are considered as rare or extreme cases who refuse to adopt the technology. These laggards can be studied to understand their unique qualities and reasons for non-adoption.

2. Based on variability

Inclusion/exclusion criteria – Selecting informants for the purpose of either minimising or maximising variability in the sample.

(i) Homogenous sampling

- Selecting cases to achieve a sample which is uniform and possess similar levels of the aspect under study;
- Employed to control extraneous variables;
- Used when the goal of the research is to understand and describe a particular group in depth;
- For example, conducting the multimedia effectiveness study with three homogenous groups of farmers – small, medium and large.

(ii) Heterogeneity/maximum variation/diversity sampling

- Selecting cases to achieve a heterogeneous sample;
- Used to capture a wide range of perspectives relating to the phenomenon;
- Used when the sample size is very small and the population information with respect to the study character is not available;
- For example beneficiary characteristics of MGNREGA (Mahatma Gandhi National Rural Employment Guarantee Scheme) in high performing districts.

3. Based on theory, model development, and hypothesis testing

Inclusion/exclusion criteria – representation of theoretical constructs for confirming or disconfirming theory on the basis of similarities or differences with a comparison group.

(i) Theoretical sampling

- The process of selecting samples on the basis of their potential manifestation or representation of important theoretical constructs (Patton, 200);
- Cases are selected based on their value for confirming (confirmatory sampling) or disconfirming (disconfirming or negative case sampling);
- Suitable when the research focuses on theory and concept development.

Types

a) Critical case sampling

- Selecting a small number of significant cases, which will provide most information and will contribute substantially to advancing the knowledge (Daniel, 2012);
- To identify critical cases, the research team needs to be able to identify the dimensions that make a case critical;

- Used when funds are limited;
- Example, documentation of indigenous technical knowledge of a tribal community using key informant method.

b) Systematic matching and case control sampling

- Persons who have a specific aspect of interest are matched with case(s) that are similar in other characters as well (Daniel, 2012);
- For example, propensity score matching for impact assessment.

c) Consecutive sampling

- Every person who meets the inclusion and exclusion criteria of a sample design within a defined time frame is included in the sample. The sample selection continues till the required sample size is achieved (Daniel, 2012);
- For example, selecting students to participate in a study based on their attitude towards virtual reality-based skill building on using farm implements.

d) Targeted sampling

- Mixed-methods sample design that involves initial rapport building, preliminary qualitative and quantitative study in the geographical area of the target population (Daniel, 2012);
- When rapport is established with the target population, the sample is selected utilizing availability sampling, quota sampling, respondent-assisted sampling and other sampling procedures.

4. Based on judgment, reputation, and specialized knowledge

(i) Judgement or subjective belief sampling

- Sample elements are selected based on the researcher's own knowledge and judgement (judgemental sampling) or subjective belief (subjective sampling) that they possess aspects required for the study;
- For example, selecting key informants for Participatory Rural Appraisal work in a village.

(ii) Bellwether sampling

- The researcher develops a profile of the sample characteristics of the study aspect and selects real world samples which match the profile;
- Used in legal research.

(iii) Reputational case sampling

- The researcher is looking for individuals who have particular expertise that is most likely to advance the researcher's interests and potentially open new doors.

(iv) Politically important sampling

- Selecting cases which represent the political phenomenon under study;
- For example, families of farmers who committed suicide.

(v) Expert and informant sampling

- The sample elements are selected since they have expert knowledge (expert sampling) or special knowledge about the population under study;
- For example, key informant interview.

C. Quota sampling: The population is divided into mutually exclusive sub-categories, and interviewers or other data collectors solicit participation in the study from members of the sub-categories until the target number of elements to be sampled is met. This is very common in qualitative research, with analytical generalizations being the most popular (Daniel, 2012).

Types

(i) Proportional quota sampling: For each quota category, the elements are allotted based on their proportion in the target population. Used when the researcher desires to ensure that the proportional distribution of certain characteristics of the sample is equal to their proportional distribution in the target population.

(ii) Non-proportional sampling: The allocation of the number of elements to be selected for each quota category is not based on their proportions in the target population, but on some other criterion.

(iii) Dimensional sampling: A special case of non-proportional quota sampling, where the researcher selects elements in such a way that at least one element in the sample represents each possible combination of dimensions of the variables targeted in the study.

D. Respondent-assisted sampling: The sample elements are selected from a target population with the assistance of previously selected population elements. Since only a few population elements with characteristics that fit the purposes of a study are sought, it is also viewed as an extension of purposive sampling. Used in studying social networks, rare populations and hidden populations.

Box 7.1: Steps in respondent-assisted sampling

- Define the target population.
- Select inclusion and exclusion criteria for the sample.
- Recruit and select initial 'seeds', - Diverse seeds (e.g., seeds that are heterogeneous in terms of age, gender, and geographical location) may be recruited via social service agencies, stores, social organizations using extension workers and ethnographic methods.
- Interview of the selected seeds.
- After establishing rapport and trust has been established, ask the interviewees for referrals. This initiates a referral chain, with each referral representing a link in the chain.
- Contact those referred.
- Repeat Step 5 and Step 6 with new referrals until targeted sample size or saturation (no new referrals are forthcoming) is achieved.

Questions that arise during the sampling process

1. How many chains should be initiated? – The ideal number of links in a referral chain depends upon the purpose of the study, the nature of the network, and the expected size of the network.
2. How many links in each chain should be targeted? – More links indicate a homogenous population.
3. Should the study have a lot of chains with a fewer number of links, or a smaller number of chains with a relatively large number of links?

Types

- Snowball sampling
- Chain-referral sampling
- Referral sampling
- Nominated sampling
- Multiplicity sampling
- Network sampling

All types of respondent-referral samples involve utilizing well-informed people to identify critical cases or informants who have a great deal of information about a phenomenon. The researcher follows this chain of contacts in order to identify and accumulate critical cases. Often a few key informants or cases will be mentioned multiple times and take on additional importance.

Advantages

- More effective in sampling rare and hard-to-reach populations;
- Appropriate for analysing social networks.

Disadvantages

- Biases;
- Expensive.

7.2.2 Sample size

The sample size requirement for qualitative research depends heavily on two factors: (i) redundancy and saturation level of each method; and (ii) variation in the population (Robinson, 2014).

1. What sample size will reach redundancy or saturation?

- Redundancy is the process of sequentially conducting interviews until all concepts are repeated multiple times without new concepts or themes emerging (Trotter II, 2012).
- In this process, the researcher conducts the first interview and analyses the points that emerge from the interview, and then repeats the same procedure after every session till a few themes emerge out of the analysis.
- The saturation level is reached when all questions have been thoroughly explored in detail and no new concepts or themes emerge in subsequent interviews (Trotter II, 2012).
- The sample size represents the maximum number of interviews or discussions required to achieve the saturation level.

2. How large a sample is needed to represent the variation within the target population?

- Variation in the population is another aspect which determines the sample size.
- As qualitative studies were intended to provide an in-depth and holistic view of the population through the eyes of fewer cases, it is necessary to identify the population variability prior to the sample selection.

Since there are no sample size guidelines available for qualitative extension research, an attempt has been made to identify the minimum sample requirement for various qualitative research methods from the disciplines of sociology, education and psychology. Table 7.1 provides the values of minimum samples required for various qualitative research studies.

Table 7.1: Sample size requirement for qualitative studies

Particulars	Required sample size (No.)
A. Research approach	
1. Action research/ethnographic/ grounded theory studies	20 to 30 interviews (Creswell, 1998)
2. Case study	Individual case study (about a successful farmer): One case or one person (Nastasi, 2016) + 25-30 data sources (Secondary data, such as manuscripts, photographs, audio recordings, etc.) (Yin, 2011) Single case study about an aspect: 15 to 30 interviews (Marshall et al., 2013) Multiple case study: 4 to 10 cases (Stake, 2006)
3. Phenomenology	5 to 25 (Creswell, 1998)
4. Other qualitative studies	Minimum 15 (Guest et al., 2006)
B. Data collection method	
1. Key informant interview	At least 5 (Nastasi, 2016)
2. Focus groups	Group size: 5 to 10 persons; No. of focus groups : 2 to 3 (Guest et al., 2016)
3. Ethnographic surveys	Large sample similar to quantitative study; usually between 30 to 60 depending upon the research question (Bernard, 2000)
4. In-depth interviews	25 to 30 interviews (Dworkin, 2012)
5. Semi-structured interviews	30 to 60 interviews (Morse, 2015)

The length of qualitative interviews often varies with the minimum number of interviews required for the study. Nastasi (2016) recommended a minimum length of interviews as one to two hours for 10 interviews, 30 minutes to one hour for 20 interviews, and 20 to 40 minutes for 30 interviews to attain the required saturation level.

7.3 Key Points

- Sampling for qualitative studies rely heavily on available, purposive, quota and respondent assisted samples;
- Purposive sampling is widely used and it is based on central tendency, variability, theory or model development, and judgement or reputation;
- Various sample size guidelines for qualitative extension research were also developed.

References

- Creswell, J. (1998.) Qualitative inquiry and research design: Choosing among five traditions. Thousand Oaks, CA: Sage.
- Curtis, S., Gesler, W., Smith, G., and Washburn, S. (2000.) Approaches to sampling and case selection in qualitative research: Examples in the geography of health. *Social Science and Medicine* 50:1001-1014.
- Daniel, J. (2012.) Sampling essentials: Practical guidelines for making sampling choices. Thousand Oaks, CA: Sage.
- Dworkin, S. L. (2012.) Sample size policy for qualitative studies using in-depth interviews. *Archives*

of Sexual Behavior, 41(6): 1319-1320.

- Firestone, W. A. (1993.) Alternative arguments for generalizing from data, as applied to qualitative research. *Educational Researcher* 22(4): 16-23.
- Guest G., Namesy E., McKenna K. (2016.) How many focus groups are enough? Building an evidence base for nonprobability sample sizes. *Field Methods*. DOI: 10.1177/1525822X16639015
- Macfarlan, A. (2015.) Analytical generalisation. Better evaluation. (Retrieved from: http://betterevaluation.org/evaluation-options/analytical_generalisation)
- Marshall, B., Cardon, P. W., Poddar, A., and Fontenot, R. (2013.) Does sample size matter in qualitative IS studies? *Journal of Computer Information Systems* 54(1): 11-22.
- Morse, J. M. (2015.) Analytic strategies and sample size. *Qualitative Health Research*, 25(10): 1317-1318.
- Nastasi, B. (2016.) Study notes: Qualitative research: Sampling & sample size considerations. (Available at: https://my.laureate.net/Faculty/docs/Faculty%20Documents/qualit_res__smp_l_size_considerations.doc).
- Onwuegbuzie, A. J., and Leech, N. L. (2007.) Sampling designs in qualitative research: Making the sampling process more public. *The Qualitative Report* 12(2): 238-254.
- Patton, M. Q. (2001.) *Qualitative research and evaluation methods*. Second edition. Thousand Oaks, CA: Sage.
- Robinson, O.C. (2014.) Sampling in interview-based qualitative research: A theoretical and practical guide. *Qualitative Research in Psychology* 11: 25–41.
- Stake, R.E. (2006.) *Multiple case study analysis*. New York: Guilford Press.
- Trotter II, R. T. (2012.) Qualitative research sample design and sample size: Resolving and unresolved issues and inferential imperatives. *Preventive Medicine*, 55(5): 398-400.
- Yin, R. (2011.) *Doing qualitative research from start to finish*. New York: Guilford Press.

08

Sample Size Estimation

P Sethuraman Sivakumar

Objectives

1. Provide an overview on methods used in estimating sample size in extension studies
2. Explain sample size estimation methods
3. Elaborate on guidelines for choosing adequate samples

8.1 Introduction

- Sample size is of primary importance for any applied scientific research as it directly influences the validity and generalisability of the research findings.
- In extension science, empirical research is expected to yield sound extension tools and techniques to help field functionaries effectively implement extension programmes.
- Social science studies conducted with inadequate sample sizes are vulnerable to inconsistencies and are likely to produce contradictory findings when conducted on the same research problem on an identical population (Johnson and Lauren, 2013).
- The purpose of this module is to describe the sample size estimation process and provide guidelines for choosing an adequate sample for both quantitative and qualitative studies in extension research.

8.2 Discussion

8.2.1 Sample size estimation for quantitative extension research

- In quantitative extension research, the sampling methods specify a few guidelines on the number of samples to be selected; the sample size is dependent on various other factors, such as type of study, nature and size of the population and choice of statistical analytical methods for the study.
- Other factors which help in deciding the sample size include the following:
 - confidence level at which the results are interpreted;
 - acceptable levels of sampling errors and precision of the results expected;
 - effect sizes required;
 - variance and standard deviations of the primary variables reported by past work.

The following are the factors to be considered while selecting the sample size for a quantitative study:

(1) Type of research investigation and test population

- Descriptive studies employ minimal statistical estimation procedures, such as proportions and Chi-square tests. For experimental studies involving human subjects (e.g., knowledge gain from multimedia instruction), the sample size depends on the design – replication, randomisation and stratification.
- The test population size also plays a crucial role in sample size estimation and the quantitative methods often require samples representing a maximum of 5% of the total population (Henry, 1990).
- The study population size can also be derived from past studies and secondary data sources (e.g., agriculture census). If the population size is unknown, the sample size can be estimated using the modified procedures as described in Box 8.1 (see section 8.2.2).

(2) Primary variable(s) of measurement

- For estimating the sample size, the researcher should decide the primary variables (dependent and a few significant independent variables) to be included in the study.
- After deciding on the primary variables, the sample sizes are estimated separately for each primary variable or combinations using the formulae given in Box 8.1.
- For example, if a researcher wishes to conduct a study on the factors influencing adoption of IPM for tomato crop, he/she should review the past studies in order to know the variables that have influenced adoption (e.g., gender, educational status, scientific orientation, etc.).
- Using the estimates of those variables (example, educational qualification correlation coefficient with adoption), the researcher can decide on the sample size using Table 8.1.
- In case of proportions, the percentage of population that has adopted (example, male farmers adopted IPM – 60%), the researcher can use the formulae given in equation 1a and 1b to estimate the sample size.
- After estimating the sample size for all primary independent variables individually, the researcher must choose the largest estimated sample size for the investigation.

(3) Acceptable margin of error – confidence intervals and confidence levels

- The margin of error is the error the researcher is willing to accept in the study. It depends on the confidence interval, which is a measure of probability that a population parameter will fall between two set values.
- The confidence interval provides a range of values around the mean (both + or - mean) which represent the value of marginal error. These values represent a population parameter (e.g., adoption level of a crop variety or animal breed in the full population of farmers in the real world) and tell us that these values are true with a probability level (e.g., 90%, 95% or 99%).
- It is necessary to decide on the allowable margin of error prior to the survey for calculating the appropriate sample size. This is decided by scanning through past research studies on the same topic and identifying the reported mean values of primary variables.
- For example, if a researcher wishes to conduct a study on 'Effectiveness of the training programme' with 'Knowledge gain' as the primary variable, he/she should find the knowledge gain mean values reported from past studies and decide on the value to be used for sample size estimation.
- In social research, a maximum of 5 percentage points around the mean is used as marginal error (Krejcie and Morgan, 1970).
- The confidence level indicates an alpha error value in hypothesis testing. During sample size estimation, we are trying to reduce the alpha error by selecting a lower significance level of either 0.05 (95%) or 0.01 (99%) of the test. While an alpha level of 0.05 (5% probability for error) is acceptable for most social research, 0.01 (1% probability for error) is preferred when critical decisions are taken using research results.

- Statistical power for any sample estimation is conventionally set at 0.80 i.e. Beta = 0.20.

(4) Effect size

- The effect size represents the size of the association between variables or difference between treatments the researcher expects to be present in the sample.
- If the researcher expects his/her study to detect even a smaller association or difference between variables with precision, then he/she may need a larger sample size.
- For example, the knowledge gain from multimedia extension module can be detected precisely when the researcher tests the module with a large sample.
- In descriptive studies, the association or difference between the variables is reflected by the amplitude of the confidence interval calculated in the estimation. The effect sizes can be estimated from the reported values of association or effect from previous studies using Cohen's D, odds ratio, correlation coefficient and eta square methods. In general, the effect size (Cohen's D) of 0.2 to 0.3 is considered as 'small', around 0.5 a 'medium' effect and 0.8 to infinity, a 'large' effect (Cohen, 1988).
- As a thumb rule, the associations or differences between variables reported in the past studies with 'small' effect, require a large sample size for further studies. Various online effect size calculators are available on the Psychometrica website (http://www.psychometrica.de/effect_size.html).

(5) Variance or standard deviation

- When the variables analysed in the study are of a quantitative nature, their variability (variance or standard deviation) is considered for sample size estimation. Variance is a measurement of the spread between numbers or observations in a data set and is a square of standard deviation.
- Cochran (1977) listed ways of estimating population variances or standard deviations for sample size estimations: (1) Select the sample in two steps, i.e. select the first sample and estimate the variance through pilot study and use the estimated value for the selection of sample size estimation for the main study; (2) use data from previous studies of the same or a similar population; or (3) estimate or guess the structure of the population assisted by some logical mathematical results. If the researcher finds difficulty in obtaining variance values from the previous study, he/she can use an arbitrary value of 50% (Krejcie and Morgan, 1970).
- In case of descriptive studies involving proportions, the researcher must specify the response distribution (labelled as p in the sample size formula) i.e., the expected proportion of the population that have the attribute the researcher is estimating from the survey. This proportion can be obtained from past studies, a pilot study or through other secondary sources.
- For example, if a researcher wishes to assess the gender differences in effectiveness of training on vegetable cultivation, he/she should review past studies to know the gender difference values (example, percentage of females who are satisfied with training). If this proportion is unknown, it should be arbitrarily set to 50% for use in the equation 1a. In case of descriptive studies involving means, the response distribution is replaced by variance or standard deviation (s^2 in Equation 1b).

8.2.2 Calculation of sample size

(1) Pre-testing (research instrument)

- The main purpose of the pre-test is to verify that the target audience understands the questions, and the proposed response options are used as intended by the researcher, and the respondents are able to answer meaningfully (Perneger et al., 2015).
- Past studies have indicated that a sample size of minimum 30 respondents to achieve a reasonable statistical power to detect problems in the instrument is appropriate (Perneger et al., 2015).

(2) Descriptive studies

- Descriptive studies are conducted to explore and describe a test population or their attributes in a systematic way. These studies are designed to estimate population parameters from the sample which do not involve testing hypotheses. The data generated through these studies are described by presenting frequencies, proportions, and means. The sample size estimation procedures for descriptive studies proposed by Rodríguez del Águilaa and González-Ramírezba (2014) are described in Box 8.1.

Box 8.1: Sample size estimation procedure for descriptive studies

A. FOR FINITE POPULATIONS (KNOWN POPULATION SIZE)

Studies involving categorical variables

When descriptive studies involve only categorical variables (e.g. age, educational qualification, employment, etc.), the researcher can estimate only proportions of a particular attribute, for example, studies aimed at describing a system (e.g. crop or animal production systems, Indigenous Technical Knowledge (ITK) documentation). The sample size for studies involving categorical variables can be computed by the following formula (Rodríguez del Águilaa and González-Ramírezba, 2014):

$$n = \frac{t_{\alpha}^2 * p * q * N}{(N - 1) * e^2 + t_{\alpha}^2 * p * q} \quad \text{Equation 1a}$$

Where n = Sample size to be estimated; t_{α} = value of the normal curve associated to the confidence level; p = expected percentage of population having a particular attribute; q = (p-1); e = accepted margin of error (usually between 5 and 10%) and expressed as percentage and N = Population size.

Studies involving interval or continuous variables

For descriptive studies involving interval or ratio variables (e.g. attitude, knowledge gain), descriptive labels such as mean, mode, median and standard deviation can be computed. The sample size for such studies can be estimated using the following formula:

$$n = \frac{t_{\alpha}^2 * s^2 * N}{(N - 1) * e^2 + t_{\alpha}^2 * s^2} \quad \text{Equation 1b}$$

Where n = Sample size to be estimated; t_{α} = value of the normal curve associated to the confidence level; s^2 = variance of the variable for which we want to estimate the mean; e = accepted margin of error (usually between 5 and 10%) and expressed as percentage and N = Population size.

Correction for estimates exceeding 5% of total population

In case of finite populations (where total population size is known e.g., dairy farmers living in a district). If the calculated sample size exceeds 5% of the population size, Cochran's (1977) correction formula should be used to estimate the final sample size.

$$n1 = \frac{n_0}{(1 + \frac{n_0}{N})} \quad \text{Equation 2}$$

Where N = Population size; n_0 = required return sample size according to Cochran's formula given in equation 1a or 1b; n_1 = required sample size because sample > 5% of population.

Correction for response rate - Surveys

The response rate is a crucial aspect in any research study involving surveys requiring voluntary participation of the respondents. Poor response rates often reduce the sample size and hamper the accuracy of the results. Salkind (1997) recommended oversampling i.e., increasing the sample size to the extent that it will account for anticipated poor return rate. Over-sampling can be achieved through four methods: (1) take the sample in two steps, and use the results of the first step to estimate how many additional responses may be expected from the second step; (2) use pilot study results; (3) use responses rates from previous studies of the same or a similar population; or (4) estimate the response rate through a systematic study. When the response rate is calculated by using any one of the above methods, the final sample size may be calculated using the following formula:

$$n_2 = \frac{n_1}{\text{Anticipated return rate}}$$

where n_2 = sample size required for response rate; n_1 = required sample estimated from equations 1a or 1b or 2.

The formulae and estimation procedure for infinite populations are described in an AESA Blog: Sample Size for Extension Research – Part 1. Quantitative Studies. (Available at: <http://www.aesa-gfras.net/Resources/file/Blog%2063.pdf>).

(3) Analytical studies involving hypothesis testing

Correlation and multiple regression

- Regression analysis is used to examine the relationship between two interval- or ratio-scaled (continuous) variables.
- To estimate the minimum sample size for the multiple regression analysis, it is essential to understand the previously reported relationship/association between the dependent and independent variables.
- For example, if a researcher wishes to identify the factors which determine adoption of a biopesticide, he/she has to derive a value of association from the adoption level and independent variables, such as extension orientation, innovativeness, environmental consciousness, etc., from previous studies.
- The association is represented by the 'reported values of correlation coefficient (r)' between the adoption and independent variables.
- When the correlation coefficient is identified from previous studies, Table 8.1 (Gatsonis and Sampson, 1989) may be used to estimate the required sample size (Weller, 2015). The table provides the sample size requirements for a given effect size (Correlation coefficient r) with default values of $\alpha = 0.05$ and $\beta = 0.20$. The first column contains the minimum correlation that can be detected and the second column contains the minimum total sample size necessary to detect it.
- In case of several independent variables used in a single study, the researcher may calculate sample sizes for all independent variables and choose the largest sample for the study. If the researcher is expecting a higher correlation between the dependent and independent variables from his/her study, the sample size can be selected based on the assumed value. The same procedure can be

used for selecting sample sizes for the study involving estimation of Pearson correlation coefficient.

- In case of research themes with no prior work or the correlation coefficients are not reported in the past studies, the method suggested by Maxwell (2000) may be followed. In this method, the correlations between the variables of interest are assumed as 'medium', ($r = 0.30$ between dependent and independent variable) and the sample size is determined based on the number of independent variables for a default effect size of 0.80. Table 8.2 provides the required sample sizes derived using Maxwell's method (Maxwell, 2000).

Table 8.1: Necessary sample size to detect a given effect size for multiple regression, ANOVA ("t" test) ($\alpha = 0.05$ and $\beta = 0.20$).

Multiple regression		ANOVA and t test	
Correlation coefficient (r)	Reqd. Sample size (N)	Eta (η)	Reqd. Sample size (N)
0.10	782	0.10	396
0.15	346	0.15	176
0.20	193	0.20	99
0.25	123	0.25	64
0.30	84	0.30	44
0.35	61	0.35	33
0.40	46	0.40	25
0.45	36	0.45	20
0.50	29	0.50	16
0.55	23	0.55	14
0.60	19	0.60	11
0.65	16	0.65	10
0.70	13	0.70	9
0.75	11	0.75	8

Source: Gatsonis and Sampson (1989)

ANOVA and "t" test

- In t test and ANOVA analyses, the nominal to interval variable associations are analysed and the association or effect size can be calculated through eta (η) coefficient. The range of η is from 0 to 1, with a larger value indicating a stronger association (Weller, 2015).
- Columns 3 and 4 in Table 8.1 indicate the eta values and corresponding sample size requirement for ANOVA and t tests (Hays, 1963). Note that these estimates assume equal group sizes. The

sample size estimation procedures for t test and ANOVA using online calculators is described in Box 8.2.

Table 8.2: Necessary sample sizes based on the number of independent variables for multiple regression ($r = 0.30$; Power = 0.80)

Number of independent variables	Required sample size
2	141
3	218
4	311
5	419
6	543
7	682
8	838
9	1009
10	1196

Source: Maxwell (2000)

Box 8.2: Online sample size calculators

A downloadable Excel file containing the macros to estimate sample size for descriptive studies (both categorical and continuous variables) developed by the author is available at: <https://drive.google.com/file/d/0BzwtVQNW-WqfYTYzSkVmZ3FDekk/view?usp=sharing>

8.3 Key Points

- Selecting adequate and representative sample is a key component of extension research. This module has elaborated the sample size estimation process for quantitative extension research. Though sample size depends on the nature of the research problem and population, the choice of statistical analytical procedures plays a crucial role in selecting the samples.
- The sample size estimation methods described in this paper are compiled from various published sources and extension scientists can use them effectively for conducting quality research.
- The decision on sample size for the extension study depends largely on past work. Most sample size formulae described here demand coefficients reported from past studies. However, the extension studies published in India haven't reported many coefficients that are essential for estimating sample size for future studies.
- A good publishing practice involves describing, analysing, and reporting the science in a proper way which helps in advancing the knowledge, apart from guiding future research.

References

- Cochran, W. G. (1977.) Sampling techniques. Third edition. New York: John Wiley & Sons.
- Cohen, J. (1988.) Statistical Power Analysis for the Behavioral Sciences. Second edition. Lawrence Erlbaum Associates.
- Gatsonis, C., and A. R. Sampson. (1989.) Multiple correlation: Exact power and sample size calculations. *Psychological Bulletin* 106:516–24.
- Hays, W. L. (1963.) Statistics for psychologists. New York: Rinehart and Winston.
- Henry, G. T. (1990.) Practical sampling, applied social research methods. Series, Vol. 21. Newbury Park, CA: Sage.
- Johnson, D. R. and Lauren, B. (2013.) What can we learn from studies based on small sample sizes? Comment on Regan, Lakhanpal, and Anguiano. (2012.) *Psychological Reports* 113(1):221-224.
- Krejcie, R. V., and Morgan, D. W. (1970.) Determining sample size for research activities. *Educational and Psychological Measurement* 30:607-610.
- Maxwell, S. E. (2000.) Sample size and multiple regression analysis. *Psychological Methods*. 5:434-458.
- Perneger, T. V., Courvoisier, D. S., Hudelson, P. M., and Gayet-Ageron, A. (2015.) Sample size for pre-tests of questionnaires. *Quality of Life Research* 24(1):147–151.
- Rodríguez del Águila, M.M., and González-Ramírez, A.R. (2014.) Sample size calculation. *Allergol Immunopathol (Madr)*. ;42(5):485---492.
- Salkind, N. J. (1997.) Exploring research. Third edition. Upper Saddle River, NJ: Prentice Hall.
- Weller, S. C. (2015.) Sample size estimation: The easy way. *Field Methods* 27(4):333-347.

09

Best Practices in Data Collection – Errors and Biases

Thammi Raju D and P Venkatesan

Objectives

1. Provide an overview of errors and biases in data collection
2. Explain strategies to avoid or minimize them

9.1 Introduction

When the data is collected using survey design, there are several errors and biases occurring throughout the process.

Errors

- Any difference between the average values that were obtained through a study and the true average values of the phenomenon of a particular population being targeted (Fluidsurveys team, 2013).
- Describes how much the results of a study are short of, or exceeds, the real values of the attribute of the population, by comprehending all the flaws in a research study.
- For example, if the study reveals that the adoption of good practices in dairy farming among small dairy farmers is 25% when the actual level is only 20%, the difference could be from a whole range of different biases and errors but the total level of error in this study is considered to be 5%.
- Two types – sampling error and non-sampling error (McNabb, 2014).

Biases

- Bias refers only to error that is systematic in nature, i.e., data's value systematically differs from the true value of the population of interest (Fluidsurveys team, 2013).
- Introduced at various stages of survey – designing, executing, data entry and analysis, and created errors.

9.2 Discussion

Types of errors and biases (McNabb, 2014)

9.2.1 Sampling error

- Sampling error comprises the differences between the sample and the population that are due solely to the particular units that happen to have been selected.
- Occurs when the sample is too small to

adequately infer survey results.

- Can be reduced by following appropriate sampling methods and size guidelines described in Modules 4 and 5.
- Sources – by chance and through sampling bias.

Sampling bias

- Tendency to favour the selection of units that have particular characteristics.
- Cause – poor sampling plan.

9.2.2 Non-sampling or measurement error

- Error that results solely from the manner in which the observations are made.
- The non-sampling error can be measurement and non-response errors.

(a) Measurement error

- The result of poor question wording or questions being presented in such a way that inaccurate or uninterpretable answers are obtained.
- The interviewer error may be survey interviewer error and process error.

A. Survey interviewer error (McNabb, 2014)

(i) Falsification error

- Deliberate falsification of survey item responses in any way, including filling out partial answers, for any reason.

Types

1. Fabricating an interview – the recording of data that are not provided by a designated survey respondent and reporting them as answers of that respondent;
2. Deliberately misreporting disposition codes and falsifying process data (e.g., the recording of a refusal case as ineligible for the sample; reporting a fictitious contact attempt);
3. Deliberately miscoding the answer to a question in order to avoid follow up questions;
4. Deliberately interviewing a non-sampled person in order to reduce effort required to complete an interview; or
5. Intentionally misrepresenting the data collection process to the survey management.

How to avoid or control

- Periodic monitoring of data collection process and selective re-interviewing to detect and deter falsification

(ii) Questioning error

- Interviewer using leading questions or different wording from one respondent to the next

How to avoid or control

- Using structured questions

- Periodic observation of interviewers during work

(iii) Surrogate respondent error

- Deliberate falsification of any portion of subject selection process or specified respondent for interview

How to avoid or control

- Periodic monitoring of data collection process and select re-interviewing to detect and deter falsification

(iv) Response option error

- Interviewer failing to follow response option instructions correctly by reading, or not reading, response options

How to avoid or control

- Comprehensive training of interviewers
- Clear instructions should be given on option-reading policy included in the interviewer's question list

(v) Question error

- Question error occurs when interviewers change the wording or sequence of questions, or when respondents do not provide an answer that relates to the topic construct upon which the question was constructed

How to avoid or control

- Developing a coding scheme for assessing the behaviors displayed by interviewers and respondents during a question-and-answer survey process is one way that researchers can check for question error at the source

B. Process error

(i) Instruction error

- Interviewer skipping or incorrectly paraphrasing instructions that result in respondent deviation from instructions

How to avoid or control

- Providing clear, concise, and unambiguous instructions
- Consistent and complete follow-through by interviewers

(ii) Recording error

- Interviewer deliberately abbreviating or omitting portions of verbal responses to unstructured questions

How to avoid or control

- Supervisor must monitor interview discussion and separately record responses for comparison
- Review of all responses for outliers

(iii) Interpretation error

- Difficulties in interpreting oral or written responses given to open-ended question

How to avoid or control

- Avoid using unstructured questions unless optional answers are provided to respondents

(iv) Scale interpretation error

- Interviewer deliberately simplifying scale options, such as using numbers instead of scale options; failure to note reverse scaling

How to avoid or control

- If scale cards are used, the wording and numbers used on the card should also be included as used on the interviewers working survey

(v) Data capture and recording error

- Process error that is often attributed to deliberate or accidental mistakes made by interviewers in reporting respondents' answers to unstructured (open-ended) questions

(vi) Editing error

- After data collection, the researcher employs a procedure that inadequately locates errors.

(vii) Coding error

- Coding error occurs as responses to open-ended questions are classified and assigned a form that can be used in tabulating and processing survey data

(b) Non-response error

- Non-response error occurs when sampling units selected for a sample are not interviewed
- Sampled units typically do not respond because they are unable, unavailable, or unwilling to do so.

Why problematic

- Introduces systematic bias into the data. This results in poorer data quality and can significantly bias any of the estimates derived from the data.
- Missing completely at random (MCAR) or being systematic.
- MCAR - If non-response is MCAR, then there is no underlying reason why certain sampling units failed to complete the survey (No bias).
- Systematic bias, as discussed above, occurs when there is some underlying reason why sampling units

do not participate in the survey. This will bias any results based upon the data to the extent to which respondents differ from non-respondents on variables of importance to the analysis.

Minimizing non-response (Miller, 2016)

- Call backs/reminders – Researchers should contact sampling units multiple times during the data collection period with reminders to complete the survey.
- Refusal conversions – If an individual has explicitly refused to complete the survey but has not asked the researcher to cease additional contact, a common tactic is to employ staff skilled at response conversion techniques to convince that respondent to participate.
- Incentives – Respondents usually feel no obligation to complete a survey. They usually do not know a researcher and do not care much about the survey itself, if at all. Frankly, respondents are doing the researcher a favor by participating. Offering an incentive to participants can be an extra boost that convinces many to participate.
- Oversampling – If there are certain sub-groups that a researcher suspects will show lower response rates, a common technique is to over-sample that group.

Response bias in self-report questionnaire research

- When the researchers employ a self-report questionnaire for measuring psychological attributes, the following errors and biases are noticed (Bertrand and Mullainathan, 2001).

(1) Ordering effects

- The order in which questions are asked in a questionnaire can have a significant effect on the results.
- The preceding questions provide the context in which the respondent answers an item, and changing this context can make a large difference in the survey results. This is called ‘priming’.
- How to avoid or control – Randomize the questions so that respondents are not answering all questions in the same order.

(2) Wording effects

- Given alternative responses, the respondent’s choices tend to be sensitive to the language used to express an alternative.
- For example, when respondents are asked, “Should the Government permit GMO crops?” most respondents said, “Yes”. But, when asked, “Should the Government ban GMO crops?” most respondents still said, “Yes”. They said so mainly because of their obedience to Government, and not based on their opinion.
- How to avoid or control: Use only neutral questions

(3) Scaling effects

- Response types are arranged in a certain order so that it can influence respondents to respond in the same order.
- For example, when studying farmers’ attitudes towards adopting eco-friendly production practices, the researcher provides response types beginning with ‘Strongly Agree’ to ‘Strongly Disagree’ for all questions. It is quite likely that the respondents will choose only ‘Strongly Agree’ for all statements.
- How to avoid or control: Randomise the question order so that the responses will be independent of scaling effects.

(4) Respondent fatigue

- The fatigue generated from a long or complex questionnaire will lead to non-response or random response.
- How to avoid - Styling and colouring – use plenty of colour combinations, graphics, logos, etc.

(5) Social desirability bias

- Respondent bias created by the unwillingness to provide honest answers stems from the participant's natural desire to provide socially acceptable answers in order to avoid embarrassment or to please the organization conducting the study.
- How to avoid – Avoid using questions which will create social desirability bias.

(6) Non-attitudes

- Respondents may not have an opinion on an issue, but will feel obliged to express one in a questionnaire.
- The attitudes supposed to be measured may not in fact exist in any coherent form.
- How to avoid – Rechecking and triangulation

(7) Acquiescence

- Yea-saying bias – Respondents' tendency to agree with items regardless of their content.
- How to avoid – Rechecking and triangulation

(8) Leniency or harshness

- Individuals systematically may respond more negatively or positively, regardless of the question posed.
- This response tendency is specific and consistent with the respondent.

(9) Critical event and recency

- Associated with respondents' recall.
- Critical event response bias occurs when a dramatic event is given a greater weight in the evaluation than routinely occurring events.
- Recency response bias occurs when events or information presented more recently are weighted more heavily by the respondent than events or information presented in the more distant past.

(10) Halo effect

Halo effect occurs when a participant's response to a previous question serves as a trigger for determining responses to subsequent questions. For example, if a question that asks people if they have heard of 'the profitability of Noni cultivation' and then, later in the questionnaire they are asked to type the names of all commercial crops suitable for their region, there is a good chance that 'Noni' will get a higher result in the latter question than otherwise.

(11) Extreme response style

Extreme response style is the 'tendency to endorse the most extreme response categories regardless of the item content.' When given a survey with a 5-point (Likert) scale, a particular respondent will always

respond using one of the two end points, either a '1' or a '5'" even though the respondents' view of some items may not be either extreme.

(12) Midpoint response style

Reflects the 'tendency to use the middle scale category (or most moderate response alternative) regardless of the content'. When provided with the same 5-point (Likert) scale, a particular respondent repeatedly responds with the neutral mid-point alternative, '3', even though the respondent's view of many items may not be neutral.

9.3 Key Points

- Survey errors are caused by sampling and non-sampling sources.
- Sampling errors are caused by inadequacies and errors during the sampling process.
- The non-sampling errors are researcher-, instrument- or respondent-based.
- The self-report questionnaire can create a halo effect, social desirability bias, central tendency bias, respondent fatigue bias, acquiescence, etc.
- Managing errors and biases require a systematic approach in selecting, designing, executing and coding data.

References

- Bertrand, Marianne and Sendhil, Mullainathan. (2001.) Do people mean what they say? Implications for subjective survey data. Massachusetts Institute of Technology, Department of Economics, Working Paper 01-04.
- Fluidsurveys team. (2013.) How to know the difference between error and bias. (Retrieved from <http://fluidsurveys.com/university/how-to-know-the-difference-between-error-and-bias/>)
- McNabb, D.E. (2014.) Non-sampling error in social surveys. Thousand Oaks, CA: Sage.
- Miller, PR. (2016). Tipsheet - Nonresponse error. DISM Survey methodology. Retrieved from: <http://www.dism.ssri.duke.edu/pdfs/Tipsheet%20-%20Nonresponse%20Error.pdf>



Photo Credit: Kanika Singh

10

Online Data Collection

Saravanan Raj and Suchiradipta Bhattacharjee

Objectives

1. Introduce various online data collection tools
2. Explain the tools in detail and how to use them

10.1 Introduction

Collection of data is a critical part of social and behavioural research; and advancement of Information and Communication Technologies (ICTs) has made the process much more efficient. Data collection has become easier, quicker and cheaper with online tools that can be accessed from computers and mobile phones with equal ease (Skarupova, 2014). The process of primary data collection has evolved from pen and paper personal interviews, postal mails, telephone surveys, computer-assisted personal interview, computer-assisted self-interview, computer-assisted telephone interview, touchstone data entry, interactive voice response/telephone-assisted self-interviewing to online email surveys and online web survey administered within browser applications. Online data collection is a research genre where the researchers employ a variety of designs, methods, and tools to conduct research online primarily using the internet. Online data collection involves several methods that are described below:

- Conducting research online: Employing alternative research designs and methods, such as online surveys, online focus groups to conduct research online. It involves planning, execution and analysis of data through online facilities. A variety of specialized skills are required to design and implement online research.
- Collecting data using online methods: Involves traditional research design and data is collected through online methods, for example - email surveys.
- Conducting research on online behavior: Involves collecting and using data from social media, blogs and bulletin boards for assessing online human behavior.

10.2 Discussion

10.2.1 Need for online data collection

Online research and data collection is emerging as an alternate method for conducting extension research for the following reasons. Extension scientists and professionals have packed schedules with a huge number of

activities, ranging from conducting empirical research to organizing field extension programmes. Since research work involves significant time spent on travel and interacting with respondents, they face challenges in balancing activities within the limited official time on hand. The online data collection methods offer effective methods of conducting research online over distances at schedules convenient to both researcher and respondents. Conducting field extension research is a costly affair. As mentioned earlier, it demands huge resources, time, and effort in designing and conducting research studies. Online data collection methods are cost-effective in terms of respondent reach and also generate data of reasonably good quality (Table 10.1).

Table 10.1: Cost and data quality in traditional and online data collection methods

Type of survey	Cost (Rs.) per person/ group	Data quality	Respondent reach
Postal survey	15.00	Medium	High
Telephone survey	15.00	High	High
Face-to-face interview (excluding cost of travel, boarding and lodging)	12.00	High	High
Focus groups	80.00	High	High
Online survey (Using Email or Free survey tools)	0.00	Medium	Medium

Online research is becoming a popular research approach in social sciences, especially for market research. An ESCOMAR (2012) survey indicated that 21% of all market surveys conducted worldwide was executed through online methods – indicating their importance in social science research.

10.2.2 Advantages and disadvantages of online data collection

While increasing popularity is making online data collection a well-practiced method, it is not free of limitations either. The advantages and disadvantages of online data collection are tabulated in Table 10.2 below (Benfield and Szlemko, 2006; Question Bank, 2007; Sincero, 2012; Skarupova, 2014; Bhaskaran, 2016).

Table 10.2: Advantages and disadvantages of online data collection

Advantages	Disadvantages
<ul style="list-style-type: none"> Democratization of survey research as participation is wide Can be made in real time through emails or chat forums Access to geographically dispersed population Sampling error due to interviewer bias can be avoided Very low cost 	<ul style="list-style-type: none"> Available only to those with access to internet (sampling bias) Only non-probability sampling is applicable Absence of interviewer to clarify and probe further Timely response can be a major constraint Representative general population surveys are not possible

Advantages	Disadvantages
<ul style="list-style-type: none"> • Faster way of conducting research • Convenient for the respondents • Flexibility to logically design surveys • Because there is no interviewer, questions of a sensitive nature can be administered • Audio-visual stimuli can be administered easily • Can be easily accessed on smartphones 	<ul style="list-style-type: none"> • High non-response error • Data protection requires special attention • Computer skill of both researcher and respondent can be a key issue • Many service providers require subscription charges for full access

10.2.3 Online research and data collection methods

Online research and data collection methods are classified into two categories, i.e., Quantitative and Qualitative methods.

A. Quantitative methods

I. Conducting research online

(a) Web Surveys

- A web survey or an INTERNET survey is a questionnaire accessed via an internet browser.
- Created and operated through specialized online survey software.
- Each questionnaire page contains a few questions mostly with multiple choice responses.
- Usually developed by the researcher using tools like Hyper Text Markup Language (HTML), or by subscribing to the existing SaaS (Software as a Service). SaaS services are offered either free of cost or on a payment basis. A few online survey SaaS services are listed in Fig. 10.1.

Rank	Name	Website	Focus	Free plan
1	Qualtrics	http://www.qualtrics.com	Comprehensive	✗
2	Survey Gizmo	http://www.surveygizmo.com	Comprehensive	✗
3	Survey Monkey	http://www.surveymonkey.com	Comprehensive	✓
4	Question Pro	http://www.questionpro.com	Comprehensive	✓
5	Jitsutech	https://www.jitsutech.com/	Comprehensive	✓
6	Fluid Surveys	http://fluidsurveys.com/	Comprehensive	✓
7	Poll Daddy	http://polidaddy.com/	Comprehensive	✓
8	Survs	http://www.survs.com/	Comprehensive	✓
9	Web survey creator	https://www.websurveycreator.com/	Comprehensive	✓
10	Instant	https://www.instant.ly/?&sbredirect	Mobile	✗
11	Allegiance	http://www.allegiance.com/solutions/surveys	Specialized, custom	✗
12	Research now	http://www.researchnow.com/en-US.aspx	Mobile, Custom	✗
13	KeySurvey	http://www.keysurvey.com	Comprehensive	✗
14	Sonar HQ	http://www.sonarhq.com/	Comprehensive	✓
15	Ask Your Target Market	http://askyourtargetmarket.com/	Market Research	✓

Fig. 10.1: Popular online web survey tools

Google Forms®

Google Forms is a freeware online survey tool associated with Google Drive (<https://docs.google.com/forms/u/0/>). It is used for creating web surveys, tests, or web input forms. It allows the researcher to create a simple and user-friendly web survey form, which is linked to a spreadsheet for storing the data. Using Google Forms, the researcher can track the progress of the survey and close it at the completion of data collection. A few significant aspects of Google Forms and its applications for Social Media Web survey are elaborated in Box 10.1.

Box 10.1: Online survey using Google Forms®

'Social media for agricultural extension and advisory services: A global survey' was conducted for the ICT4RAS interest group of GFRAS globally among the extension practitioners and professionals in which 229 respondents from 69 countries participated. Google Forms was used for the survey. To increase visibility and participation, it was promoted through social media platforms like Facebook and Twitter, on web portals of Young Professionals for Agricultural Development (YPARD), eAgriculture, GFRAS and AESA; and email links were sent to some respondents to take the survey. Some advantages of the online survey conducted using Google Forms are:

- a. A global audience was reached in very little time and at no cost at all;
- b. Data visualization was easier;
- c. Accessible through smartphones;
- d. Page breaks make the sections sequential and breaks monotony.

Drawbacks:

- a. Limited audience – as it is accessible only to those with internet access;
- b. Non-response;
- c. Dropping out because of compulsory questions of a sensitive nature;
- d. Limitations in designing the questionnaire.

A few do's and don'ts for using Google Forms as a survey tool are displayed in Box 10.2. Some of

Box 10.2: A few Do's and Don'ts when using Google Forms as a web survey tool

- a. Visualization of output before preparation of schedule is important;
- b. Precise, clear and concise questions;
- c. Choice of online data collection tools should be based on the type of questions to be posed;
- d. Page breaks should be incorporated instead of long scrolling surveys (branching logic);
- e. Sequence of questions should be logical;
- f. Mention tentative time required to take the survey;
- g. Leading questions should be avoided;
- h. If options are non-exhaustive, 'other' option should be incorporated;
- i. Long questions need to be avoided;
- j. Privacy protection measures need special attention;
- k. Communication of survey results to respondents is necessary.

the web survey response types have been illustrated in the table below (Table 10.3). For detailed guidelines on designing online web surveys refer to Annexure A-1.

Table 10.3: Web survey response types and their use

Response type	Example	Used when																																				
Radio buttons	<p>1. How often do you use social media:</p> <table border="1"> <thead> <tr> <th></th> <th>Never</th> <th>Rarely</th> <th>Sometimes</th> <th>Often</th> <th>Always</th> </tr> </thead> <tbody> <tr> <td>Facebook</td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> </tr> <tr> <td>Twitter</td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> </tr> <tr> <td>Linkedin</td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> </tr> <tr> <td>Google+</td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> </tr> </tbody> </table>		Never	Rarely	Sometimes	Often	Always	Facebook	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Twitter	<input type="radio"/>	Linkedin	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Google+	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Used when the respondent must select exactly one choice from a list Multiple choice questions, Likert scale										
	Never	Rarely	Sometimes	Often	Always																																	
Facebook	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>																																	
Twitter	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>																																	
Linkedin	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>																																	
Google+	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>																																	
Check Boxes	<p>4. Which of the following animals live with you? Please check all that apply</p> <p><input type="checkbox"/> Bird</p> <p><input type="checkbox"/> Cat</p> <p><input type="checkbox"/> Dog</p> <p><input checked="" type="checkbox"/> Fish</p> <p><input checked="" type="checkbox"/> Mouse, hamster or other rodent</p> <p><input type="checkbox"/> Snake, iguana, or other reptile</p> <p><input type="checkbox"/> Other animal</p> <p><input type="checkbox"/> None</p>	Used when respondents are permitted to select more than one option from a list, including “select all that apply” type of questions																																				
Drop-Down Menus	<p>Your residential history</p> <table border="1"> <tbody> <tr> <td>Born in</td> <td>-- Please select state --</td> </tr> <tr> <td>Grew up in</td> <td>-- Please select state --</td> </tr> <tr> <td>Live in</td> <td>-- Please select state --</td> </tr> <tr> <td>Work in</td> <td>-- Please select state --</td> </tr> </tbody> </table> <p style="text-align: right;"><input type="button" value="Dropdown List"/></p>	Born in	-- Please select state --	Grew up in	-- Please select state --	Live in	-- Please select state --	Work in	-- Please select state --	Effective when the list of response options is lengthy and would result in excessive scrolling on the page to see all the options and to get to the next question																												
Born in	-- Please select state --																																					
Grew up in	-- Please select state --																																					
Live in	-- Please select state --																																					
Work in	-- Please select state --																																					
Rank-Order Matrices	<p>Please rank the following factors in order of importance to you in your work, where 1 is most important and 5 is least important:</p> <table border="1"> <thead> <tr> <th></th> <th>1</th> <th>2</th> <th>3</th> <th>4</th> <th>5</th> </tr> </thead> <tbody> <tr> <td>Pay</td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input checked="" type="radio"/></td> <td><input type="radio"/></td> </tr> <tr> <td>Benefits</td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input checked="" type="radio"/></td> </tr> <tr> <td>Recognition</td> <td><input checked="" type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> </tr> <tr> <td>Career opportunities</td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input checked="" type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> </tr> <tr> <td>Supportive colleagues</td> <td><input type="radio"/></td> <td><input checked="" type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> </tr> </tbody> </table>		1	2	3	4	5	Pay	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Benefits	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Recognition	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Career opportunities	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	Supportive colleagues	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	To rank a list of options in order of preference or importance; Respondents are not permitted to give the same rank to more than one option, and all options must be ranked before the respondent is allowed to move on to the next question.
	1	2	3	4	5																																	
Pay	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>																																	
Benefits	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>																																	
Recognition	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>																																	
Career opportunities	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>																																	
Supportive colleagues	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>																																	
Constant Sum	<p>2. Please indicate the percentage of your workday you spend on the following activities. (The total must add up to 100%)</p> <p><input type="checkbox"/> Reading e-mail</p> <p><input type="checkbox"/> Answering the phone</p> <p><input type="checkbox"/> Taking orders</p> <p><input type="checkbox"/> Resolving customer complaints</p> <p><input type="checkbox"/> Developing new projects</p> <p><input type="checkbox"/> Other</p>	To assign values or percentages across options so that the total aggregates to a predetermined amount																																				
Open-Ended Text Boxes	<p>Describe your experience today:</p> <div style="border: 1px solid #ccc; height: 40px; width: 100%;"></div> <p style="text-align: center;"><input type="button" value="Back"/> <input type="button" value="Submit"/></p>	To type in free text Qualitative data on reasoning out a particular decision																																				

II. Collecting data using online methods

(a) Email surveys: This involves a researcher mailing his/her survey instruments to a large number of people through email with a request to download and fill it up and send back to them. It is an online data collection method where a traditional survey instrument is converted into a Word document and sent to prospective respondents.

B. Qualitative methods

(a) Online focus groups: Online groups comprise of 6 to 8 people participating in a discussion on a specific topic conducted in real-time, using a dedicated online chat space which is moderated by the researcher. Online focus group software provides for people to attend focus groups in different modes, for example as participant, moderator, admin (administrators) and observer. The minimum number of participants in an online focus group is four to six (Malhotra, 2007), and the optimum length of discussion – 90 minutes (Sweet, 2001).

(b) Bulletin board groups: Asynchronous group discussion that takes place over an extended period of time (from days to months). It is led by a moderator who posts discussions, prompts, replies and tasks on a specially configured bulletin board. The respondents log into the bulletin board, read the latest posts, complete any tasks set for them, and add their comments (Poynter, 2010). Most bulletin board groups are short term, i.e., they tend to run from a few days to three weeks (Schillewaert et al., 2009). The participants are sampled from online access panels or client databases. Various steps in conducting bulletin board research are displayed in Box 10.3.

Box 10.3: Steps in conducting Online Bulletin Board Research

1. Design the bulletin board group;
2. Create the discussion guide;
3. Recruit the participants;
4. Moderate the groups;
5. Extract and report insight;
6. If there is more time, revise 2 and repeat from 4 onwards;
7. Close the group.

Source: Poynter (2010)

(c) Parallel in-depth interviews: These are online, asynchronous, in-depth interviews conducted using an online platform. A typical offline depth interview is conducted in one session, with the interview lasting from 15 minutes to two hours, depending on the project requirements. Sometimes the interviews extend over 15 days depending upon the availability of participants and interviewer. Conducted using bulletin board group software.

Steps in conducting Parallel In-depth Interviews:

1. Create the discussion guide;
2. Recruit the participants;
3. Conduct the interviews;
4. Extract the insight;
5. Close the project.

Various software or online platforms for conducting online qualitative research are described in Box 10.4.

Box 10.4: Software required for Qualitative Online Research

Online qualitative research can be conducted through researcher-created software of SaaS services. The following SaaS services offer multiple facilities to conduct a variety of qualitative online research:

Paid SaaS

FocusGroupIt - <https://focusgroupit.com/>

Focus Vision - <https://www.focusvision.com/>

iTracksChat - <http://www.itracks.com/products-services/chat/>

vBulletin - <https://www.vbulletin.com/>

Free, open source SaaS

PhpBB - <https://www.phpbb.com/>

III. Digitisation of data

1. Use of electronic media, such as tablet to collect, store, manipulate and transmit data relating to interviews conducted between interviewer and respondent(s);
2. Tablets are used in collecting data – cost effective, real time data storage and location identification facilities;
3. Examples – Survey Data ++ (<https://www.seedmanagement.net/survey-data/>).

For advanced methods (simulations, virtual laboratories, large data sets, audience response systems) refer to Annexure A-2.

10.3 Key Points

- Online surveys are becoming more important in social science research because of their time and cost effectiveness.
- Different types of online data collection tools are available but the choice depends strictly on the type of research undertaken.
- Data visualization and management becomes easier with online tools. All the same sampling and low response rates are major problems along with low internet connectivity and poor computer skills.
- Special attention is needed while preparing the type of questionnaire so as to reduce non-response.

References

- Benfield, J.A. and Szlemko, W.J. (2006.) Internet based data collection: Promises and realities. *Journal of Research Practice* 2(2): Article D1. (Retrieved from: <http://jrp.icaap.org/index.php/jrp/article/view/30/51>).
- Bhaskaran, V. (2016.) Online research – A handbook for online data collection. (Retrieved from: <http://www.questionpro.com/images/Online-Research-Handbook.pdf>).
- Poynter, R. (2010.) *The handbook of online and social media research*. Wiltshire: John Wiley & Sons Ltd.

- Question Bank. (2007.) Factsheet 2: Methods of data collection in social surveys. UK: Department of Sociology, University of Surrey. (Retrieved from: https://www.ukdataservice.ac.uk/media/262832/discover_methodsofcollectionfactsheet.pdf).
- Malhotra N. K. (2007.) Marketing research: An applied orientation. Fifth edition. Prentice Hall.
- Schillewaert, N, De Ruyck,T, and Verhaeghe, A. (2009.) Connected research: How market research can get the most out of semantic web waves. *International Journal of Market Research* 51(1):11–27.
- Sincero, S.M. (2012.) Online surveys. (Retrieved from: <https://explorable.com/online-surveys>). (Published on 16.10.2012, Retrieved on 18.12.2016).
- Skarupova, K. (2014.) Computer assisted and online data collection in general population surveys. Lisbon: European Monitoring Centre for Drugs and Drug Addiction. (Retrieved from: http://www.emcdda.europa.eu/system/files/publications/808/Technical_report_Computer_assisted_and_online_data_collection_in_GPS_480810.pdf).
- Sweet, C. (2001.) Designing and conducting virtual focus groups. *Qualitative Market Research: An International Journal* 4(3):130–135.

11

Data Exploration and Preparation – Missing Data Treatment

P Sethuraman Sivakumar

Objectives

1. Make readers understand the importance of checking and cleaning data prior to statistical analysis
2. Explain various methods and tools for examining the data before statistical analysis

11.1 Introduction

- Data exploration and preparation is an essential step in extension research which makes the data amenable to statistical analysis.
- In extension research, the data generated from field research are often subjected to biases and errors which need to be rectified before analysing them.
- This module provides an overview of various data exploration and preparation techniques, such as missing data analysis, outlier detection and treatment, normality testing and linearity assessment.

11.2 Discussion

11.2.1 Data exploration and preparation

- Examining the collected data to understand its characteristics and identify the inadequacies and errors before analysing them.
- Purpose of data exploration
 - Identifying patterns in the data – representativeness of respondents, sampling adequacy;
 - Locating and correcting errors – missing data, outliers;
 - Choosing right data analytical tools by examining whether the data satisfy the necessary statistical assumptions.

11.2.2 Steps in data exploration and preparation

The data exploration and preparation starts after collecting the data. The following procedure suggested by Ray (2016) can be adopted:

- A. Variable identification;
- B. Missing data treatment;
- C. Outlier treatment;
- D. Testing assumptions for statistical analysis;
- E. Data transformation.

In this module, only variable identification and missing value treatment are discussed.

A. Variable identification

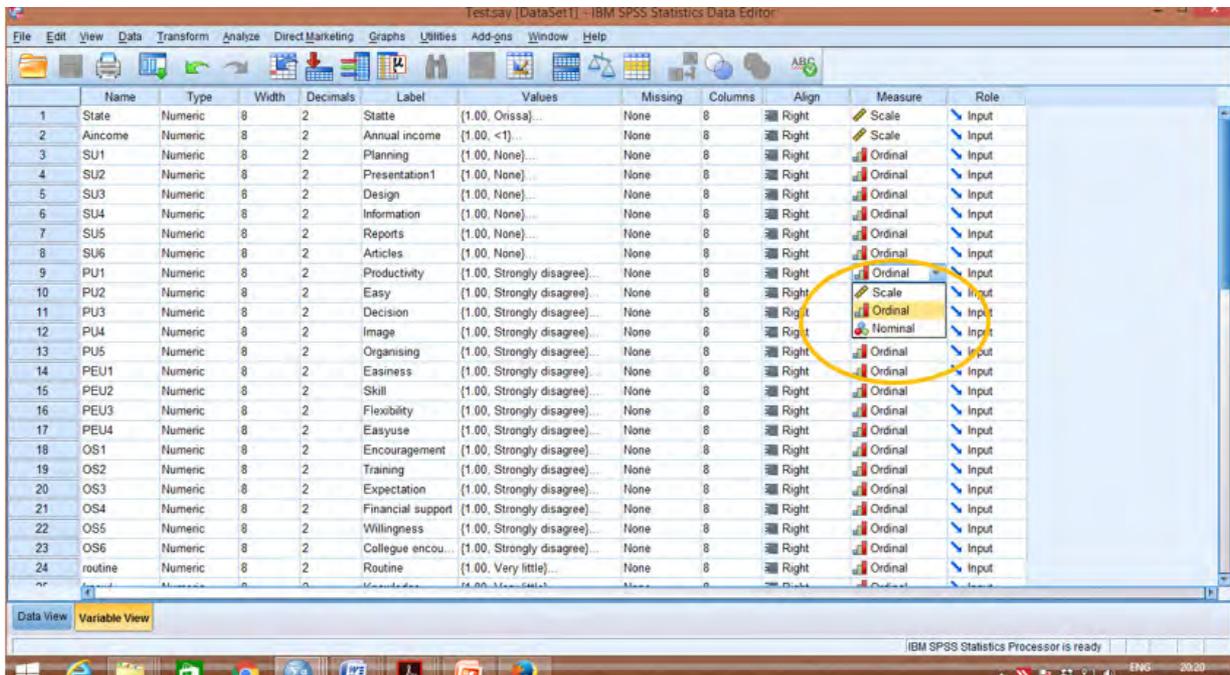
- After collection, the data can be entered in the Spreadsheet (example, SPSS) by identifying its nature (dependent or independent), type (string or numerical) and variable category (categorical and continuous).
- The labels/names of variables and values for each variable (example, Likert responses 1 – ‘Strongly disagree’ to 5 – ‘Strongly agree’) are also entered.
- All this information is entered through ‘Variable View’ mode of SPSS.
- A few sample SPSS data sheets in ‘Variable View’ mode indicating various steps are displayed below.

SPSS Data sheet in ‘Variable View form’

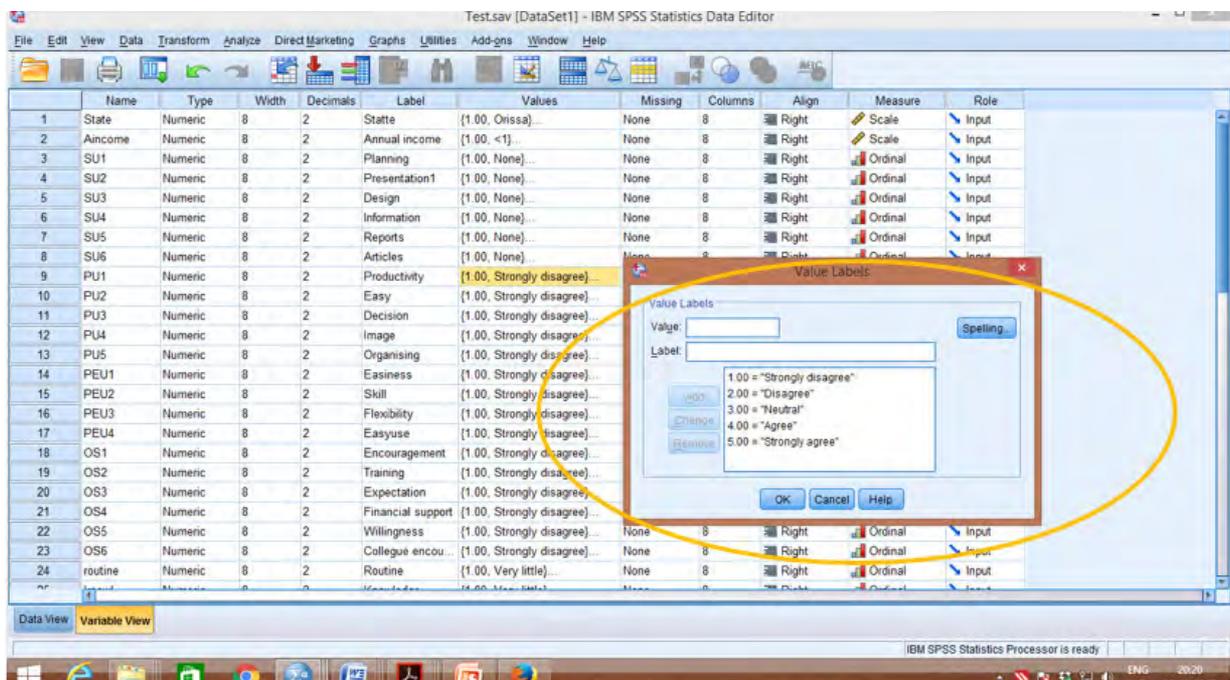
1. Entering variable codes in ‘name’ section and actual names in the ‘labels’ section

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	State	Numeric	8	2	Statte	{1.00, Orissa}...	None	8	Right	Scale	Input
2	Aincome	Numeric	8	2	Annual income	{1.00, <1}...	None	8	Right	Scale	Input
3	SU1	Numeric	8	2	Planning	{1.00, None}...	None	8	Right	Ordinal	Input
4	SU2	Numeric	8	2	Presentation1	{1.00, None}...	None	8	Right	Ordinal	Input
5	SU3	Numeric	8	2	Design	{1.00, None}...	None	8	Right	Ordinal	Input
6	SU4	Numeric	8	2	Information	{1.00, None}...	None	8	Right	Ordinal	Input
7	SU5	Numeric	8	2	Reports	{1.00, None}...	None	8	Right	Ordinal	Input
8	SU6	Numeric	8	2	Articles	{1.00, None}...	None	8	Right	Ordinal	Input
9	PU1	Numeric	8	2	Productivity	{1.00, Strongly disagree}...	None	8	Right	Ordinal	Input
10	PU2	Numeric	8	2	Easy	{1.00, Strongly disagree}...	None	8	Right	Ordinal	Input
11	PU3	Numeric	8	2	Decision	{1.00, Strongly disagree}...	None	8	Right	Ordinal	Input
12	PU4	Numeric	8	2	Image	{1.00, Strongly disagree}...	None	8	Right	Ordinal	Input
13	PU5	Numeric	8	2	Organising	{1.00, Strongly disagree}...	None	8	Right	Ordinal	Input
14	PEU1	Numeric	8	2	Easiness	{1.00, Strongly disagree}...	None	8	Right	Ordinal	Input
15	PEU2	Numeric	8	2	Skill	{1.00, Strongly disagree}...	None	8	Right	Ordinal	Input
16	PEU3	Numeric	8	2	Flexibility	{1.00, Strongly disagree}...	None	8	Right	Ordinal	Input
17	PEU4	Numeric	8	2	Easyuse	{1.00, Strongly disagree}...	None	8	Right	Ordinal	Input
18	OS1	Numeric	8	2	Encouragement	{1.00, Strongly disagree}...	None	8	Right	Ordinal	Input
19	OS2	Numeric	8	2	Training	{1.00, Strongly disagree}...	None	8	Right	Ordinal	Input
20	OS3	Numeric	8	2	Expectation	{1.00, Strongly disagree}...	None	8	Right	Ordinal	Input
21	OS4	Numeric	8	2	Financial support	{1.00, Strongly disagree}...	None	8	Right	Ordinal	Input
22	OS5	Numeric	8	2	Willingness	{1.00, Strongly disagree}...	None	8	Right	Ordinal	Input
23	OS6	Numeric	8	2	Colleague encou...	{1.00, Strongly disagree}...	None	8	Right	Ordinal	Input
24	routine	Numeric	8	2	Routine	{1.00, Very little}...	None	8	Right	Ordinal	Input

2. Specifying variable type



3. Specifying response categories in the 'Values' section



The data entry procedure depends on the type of analysis performed by the researcher. Various types of data entry for different analyses are shown in the following figures.

a. Correlation/ factor analysis

Visible: 4 of 4 Variables

	Adoption_index	Information_seeking_behavior	Extension_participat.	Land_holding	var	var	var	var	var	var	var	var
1	5.30		12.50	19.50	5.50							
2	4.20		18.25	14.25	2.50							
3	9.20		16.50	19.50	9.00							
4	6.50		10.50	11.25	1.50							
5												
6												
7												
8												
9												
10												
11												
12												
13												
14												
15												
16												
17												
18												
19												
20												
21												
22												
23												

b. Independent test

Visible: 4 of 4 Variables

	Adoption_category	Information_seeking_behavior	Extension_participat.	Land_holding	var	var	var	var	var	var	var	var
1	1.00		12.50	19.50	5.50							
2	.00		18.25	14.25	2.50							
3	.00		16.50	19.50	9.00							
4	1.00		10.50	11.25	1.50							
5												
6												
7												
8												
9												
10												
11												
12												
13												
14												
15												
16												
17												
18												
19												
20												
21												
22												
23												

c. ANOVA

	Gender	Media_type	Knowledge_gain	var									
1	Male	Multimedia	12.50										
2	Male	Video	18.25										
3	Female	Video	16.50										
4	Female	Video	10.50										
5	Male	Video	8.50										
6	Female	Multimedia	2.00										
7													
8													
9													
10													
11													
12													
13													
14													
15													
16													
17													
18													
19													
20													
21													
22													
23													

B. Missing data treatment

- Missing data refers to absence of valid values on one or more variables in the data sheet.
- Missing data can reduce the sample size and power of a model leading to estimates with poor model fit.
- Purpose - to identify the patterns and relationships underlying missing data so that the original data distribution pattern can be maintained.

Reasons for missing data

1. Data entry – When the data is transferred from a questionnaire to a data sheet, a few values may be missed by the data entry operator.
2. Data extraction – When the data is extracted from a database or from an Excel spreadsheet to statistical software, like SPSS, a few values may be missed due to extraction errors.
3. Data collection – The errors occurring during data collection are difficult to correct. They can be categorized under four types (Ray, 2016):

(a) Missing completely at random

- The propensity for a data point to be missing is completely random;
- For example, the data on 'land holding' may be missing across the gender categories randomly.

(b) Missing randomly

- The propensity for a data point to be missing is not related to the missing data, but it is related to some of the observed data.
- For example, if a respondent didn't answer a question about 'Annual income of the family', it is because (i) she may be a housewife who didn't know the exact annual income of her husband; or (ii) she may not be aware of income from secondary sources other than the ones she already knows; or (iii) she didn't want to disclose the real income.

(c) Missing values that depend on unobserved predictors

- Where the missing values are not random and are related to the unobserved input variable.
- For example, when conducting effectiveness of multimedia module on 'Ornamental Fish Farming', some participants may not answer the post-test of knowledge as they couldn't understand some part of the module.

(d) Missing that depends on the missing value itself

- Where the probability of missing value is directly correlated with missing value itself.
- For example, people with higher or lower annual incomes are likely to skip the question related to annual income.

Treatment of missing values

Missing values can be ignored, deleted, or remedied based on the nature of missing values and nature of the study. Various methods for treating missing data are explained (Hair Jr et al., 2006; Ray, 2016) below:

1. Ignore the missing values

- If the missing data are about 10% for an individual case or variable, it can be ignored;
- If the cases with no missing data are sufficient for selected data analysis method.

2. Deletion

- Variables with 15% of missing data can be deleted. However, if the missing data exceeds 20%, they can be remedied;
- Cases with missing data in dependent variable are deleted, as they are likely to affect the relationships specified for the study;
- Conduct the statistical analysis with and without the case or variable to check the differences in the coefficients. If there is no marked difference, delete the specific case or variable;
- Deletion methods are used when the nature of missing data is 'Missing completely at random'.

Methods

(i) List-wise deletion or complete-case analysis

- Removes all data for a case or variable that has one or more missing values;
- Used when (a) data are 'Missing completely at random'; (b) deletion will not affect the power of the analysis even with reduced sample size.

(ii) Pair-wise deletion or using all available data approach

- Analysis is performed with all cases in which the variables of interest are present;
- Advantage – Sufficient cases are available for analysis;
- Disadvantage – Uses different sample size for different variables.

3. Mean/Mode/Median Imputation

- Imputation – Method for filling the missing data with estimates from valid values of other cases;
- Uses relationships identified from the valid values of the data for estimating the missing values;
- Popular method where the missing data for a specific attribute are replaced with the mean or median (quantitative attribute) or mode (qualitative attribute) of all known values of that variable.

Types

(i) Generalized Imputation

- The mean or median for all non-missing values of that variable are calculated and substituted for missing values;
- SPSS software can perform the following mean substitution methods (IBM, 2017):

Methods

1. Series mean – Replaces missing values with the mean for the entire series;
2. Mean of nearby points – Replaces missing values with the mean of valid surrounding values;
3. Median of nearby points – Replaces missing values with the median of valid surrounding values;
4. Linear interpolation – Replaces missing values using linear interpolation. The last valid value before the missing value and the first valid value after the missing value are used for the interpolation. If the first or last case in the series has a missing value, the missing value is not replaced;
5. Linear trend at point – Replaces missing values with the linear trend for that point. The existing series is regressed on an index variable scaled 1 to n. Missing values are replaced with their predicted values.

(ii) Similar case Imputation

- The average is calculated for each category of the variable (e.g., gender – male and female; race – Indian and non-Indian) and missing values in the respective categories are substituted with estimated values.

4. Prediction Model

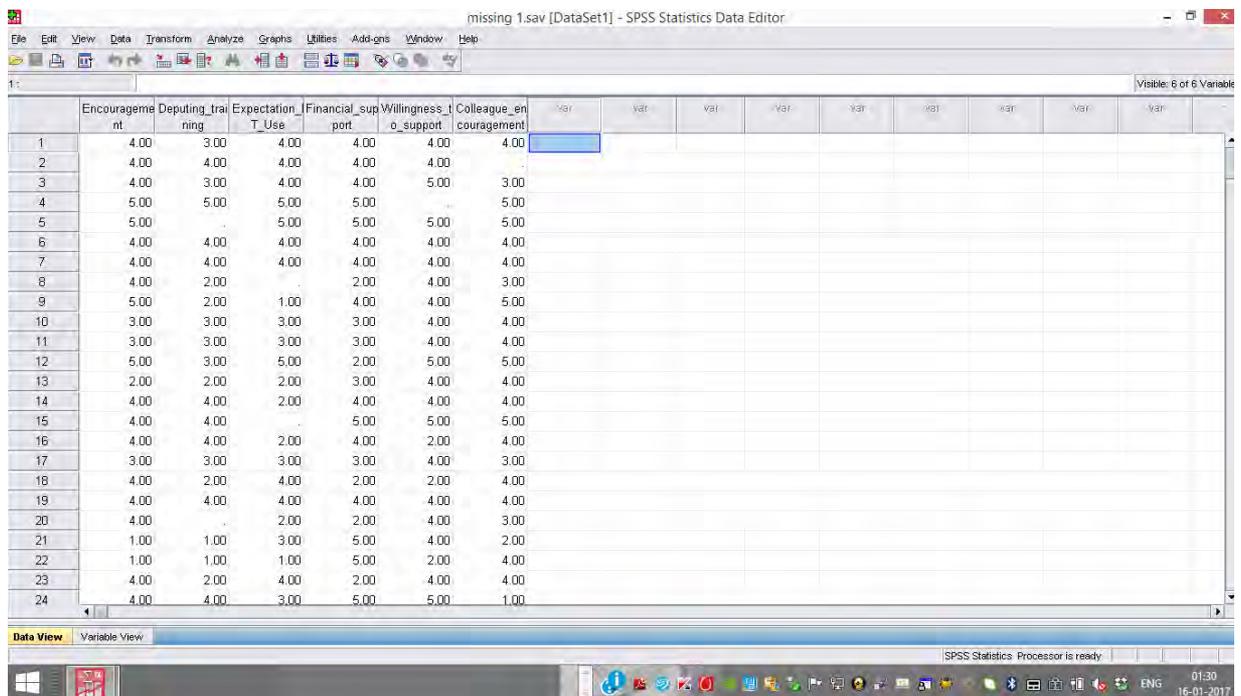
- Creating a prediction model from the valid data to estimate values for substituting missing data;
- Statistical techniques, such as multiple regression, ANOVA, Logistic regression and various modelling techniques are used to estimate missing values.

5. K Nearest-Neighbour (KNN) Imputation

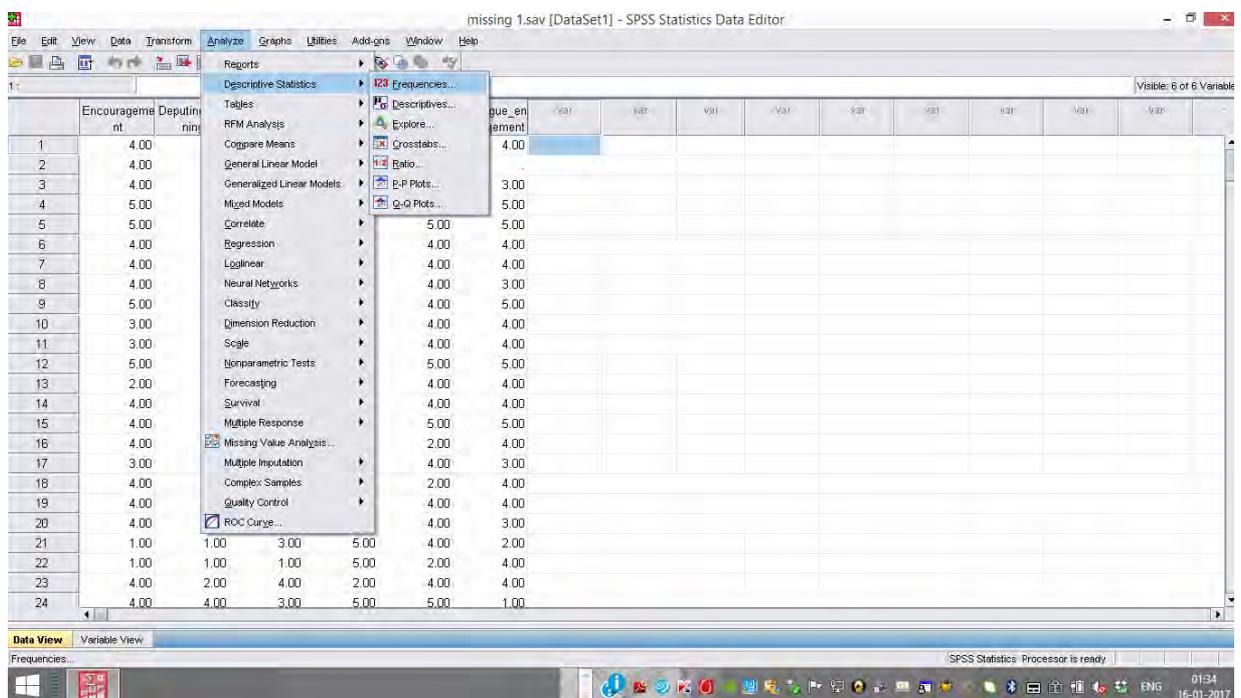
- The missing values of an attribute are imputed using the given number of attributes that are most similar to the attribute whose values are missing, for example, knowledge;
- The similarity of two attributes is determined using a distance function.

Missing value treatment using SPSS

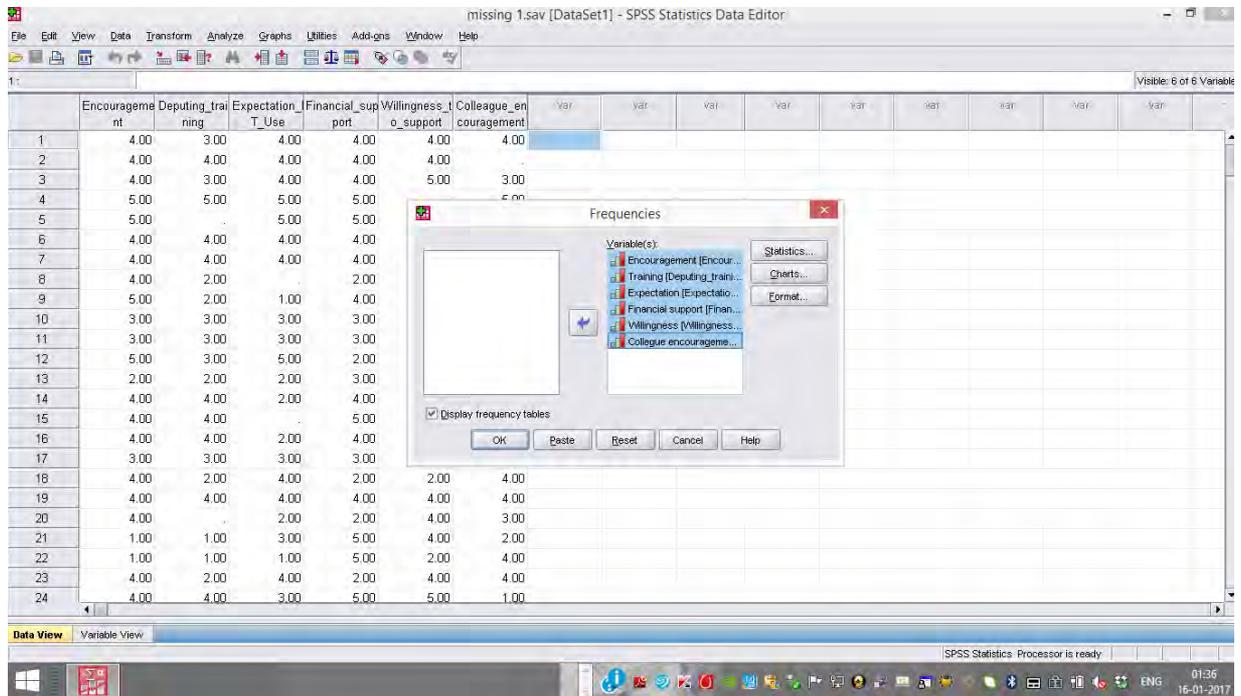
1. SPSS datasheet with missing values



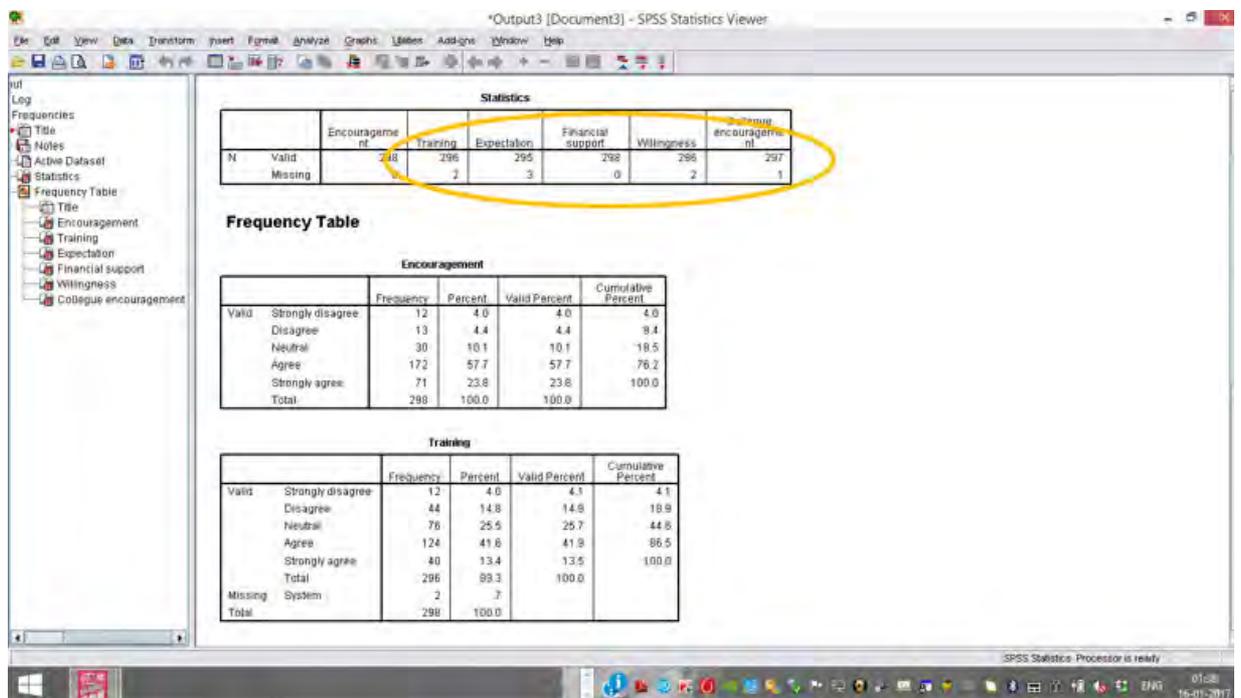
2. To identify the variables with missing data, click Analysis > Descriptive statistics > Frequencies



3. Drag all the variables into the adjacent box and click OK

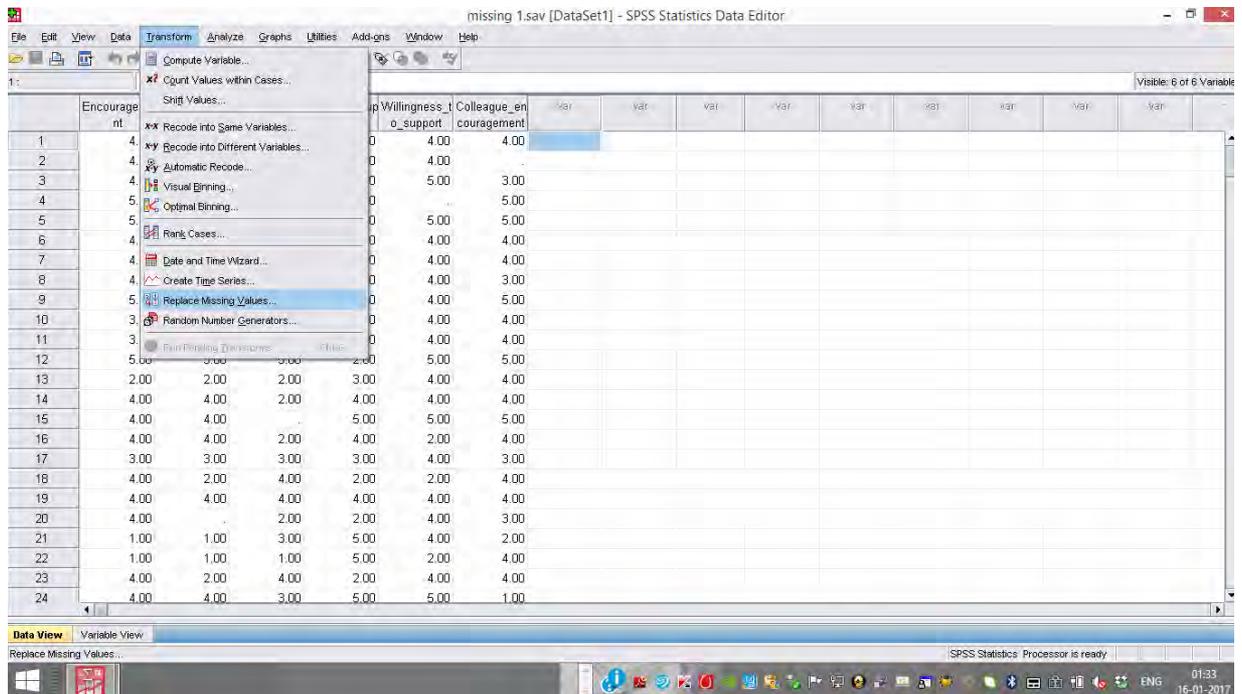


4. Examine the output to identify variables with missing values

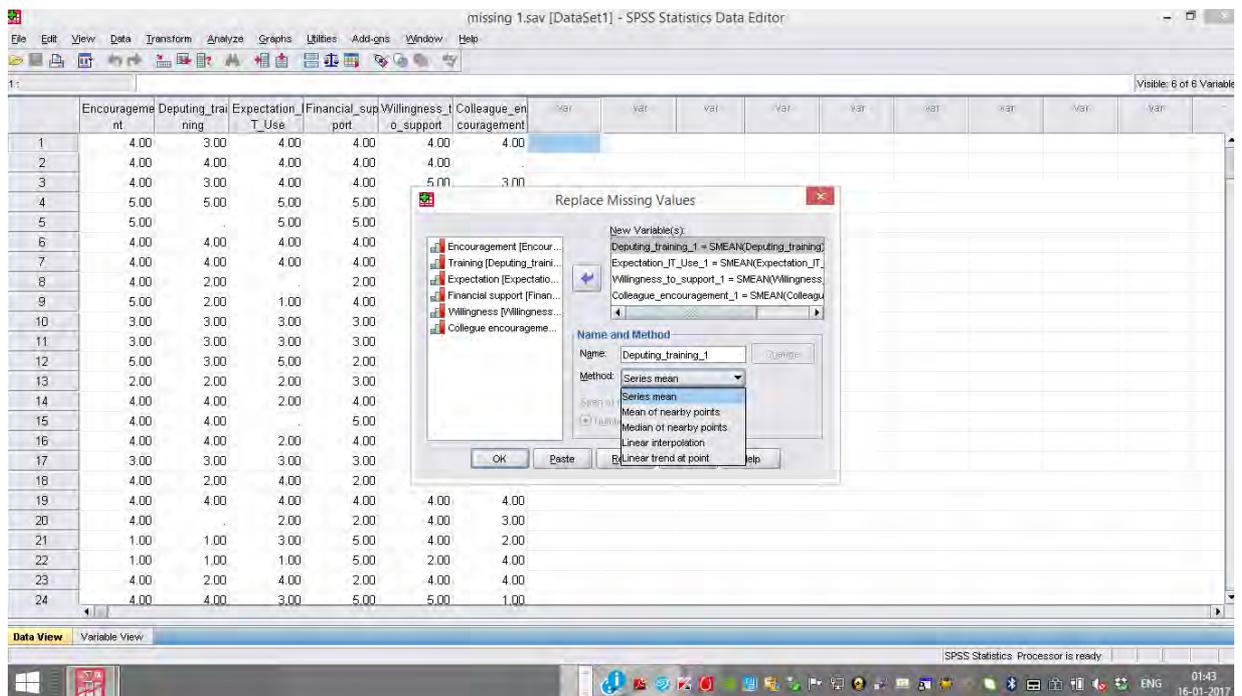


In this example, the variables deputing 'Training', 'Expectation', 'Willingness' and collegial encouragement have missing values.

5. To replace missing values, click Transform > Replace missing values



6. Under the drop menu 'Name and method' select the default 'Series mean'



7. Please check the SPSS data sheet for four newly created variables with missing values substituted.

	financial_support	Willingness_to_support	Colleague_encouragement	Deputing_training_1	Expectation_IT_Use_1	Willingness_to_support_1	Colleague_encouragement_1
1	4.00	4.00	4.00	3.00	4.00	4.00	4.00
2	4.00	4.00		4.00	4.00	4.00	3.85
3	4.00	5.00	3.00	3.00	4.00	5.00	3.00
4	5.00		5.00	5.00	5.00	3.72	5.00
5	5.00	5.00	5.00	3.46	5.00	5.00	5.00
6	4.00	4.00	4.00	4.00	4.00	4.00	4.00
7	4.00	4.00	4.00	4.00	4.00	4.00	4.00
8	2.00	4.00	3.00	2.00	3.78	4.00	3.00
9	4.00	4.00	5.00	2.00	1.00	4.00	5.00
10	3.00	4.00	4.00	3.00	3.00	4.00	4.00
11	3.00	4.00	4.00	3.00	3.00	4.00	4.00
12	2.00	5.00	5.00	3.00	5.00	5.00	5.00
13	3.00	4.00	4.00	2.00	2.00	4.00	4.00
14	4.00	4.00	4.00	4.00	2.00	4.00	4.00
15	5.00	5.00	5.00	4.00	3.78	5.00	5.00
16	4.00	2.00	4.00	4.00	2.00	2.00	4.00
17	3.00	4.00	3.00	3.00	3.00	4.00	3.00
18	2.00	2.00	4.00	2.00	4.00	2.00	4.00
19	4.00	4.00	4.00	4.00	4.00	4.00	4.00
20	2.00	4.00	3.00	3.46	2.00	4.00	3.00
21	5.00	4.00	2.00	1.00	3.00	4.00	2.00
22	5.00	2.00	4.00	1.00	1.00	2.00	4.00
23	2.00	4.00	4.00	2.00	4.00	4.00	4.00
24	5.00	5.00	1.00	4.00	3.00	5.00	1.00

11.3 Key Points

- Data exploration and preparation is an essential step prior to data analysis.
- Missing data can cause serious biases and inaccuracies which will distort the results.
- Missing values may either be deleted (e.g., list-wise deletion) or computed (e.g., series mean, mean of nearby points, etc.).

References

- Hair, Jr., J. F., Black, W. C., Babin, B. J., Anderson, R. E., and Tatham, R. L. (2006.) Multivariate data analysis. Sixth edition. Upper Saddle River, NJ: Pearson Prentice Hall.
- IBM Knowledge Centre. (2017.) Estimation methods for replacing missing values. (Retrieved from: http://www.ibm.com/support/knowledgecenter/en/SSLVMB_24.0.0/spss/base/replace_missing_values_estimation_methods.html)
- Ray, S. (2016.) A comprehensive guide to data exploration. (Retrieved from: <https://www.analyticsvidhya.com/blog/2016/01/guide-data-exploration/#one>)



Photo Credit: Rasheed Gulaiman X

12

Data Exploration and Preparation – Outlier Treatment

P Sethuraman Sivakumar

Objectives

1. Make readers understand the importance of checking and cleaning data prior to statistical analysis
2. Explain various methods and tools for identifying and treatment of outliers in the data

12.1 Introduction

- Outliers in the social data can create serious issues that affect accuracy of the results of data analysis;
- This module provides an overview of outliers, their types, and various methods and strategies for identifying and treating outliers in the data.

12.2 Discussion

Outliers

- Outliers – Observations with a unique combination of characteristics identifiable as distinctly different from the other observations in the data sheet;
- Such an observation appears far away from the other observations and diverges from the overall pattern of the sample;
- Unusually high or low value on a variable.

Causes of outliers (Ray, 2016)

- Data Entry Errors – Human errors caused during data collection, recording, or data entry. Typographical mistakes can add or delete a Zero in the number value and this will create outliers;
- Measurement Error – Occurs when using a faulty instrument;
- Experimental Error – Errors caused when conducting an experiment;
- Intentional Outlier – Reporting false information in the self-reported measures which involves sensitive data (for example, drug abuse data);
- Data Processing Error – Manipulation or extraction errors when capturing data into statistical software from the spreadsheet;
- Sampling Error – Choosing an unrepresentative sample;
- Natural Outlier – Inbuilt into the measurement system.

Impact of outliers (Ray, 2016)

- Increases the error variance and reduces the power of statistical tests;
- If the outliers are non-randomly distributed,

- they can decrease normality;
- They can bias or influence estimates that may be of substantive interest;
- They can also impact the basic assumption of Regression, ANOVA and other statistical model assumptions.

Sources of outliers (Hair Jr et al., 2006)

- Procedural error – Data entry error or mistake in data coding;
- Extraordinary event – Any extraordinary event that occurred in the study area which affects the response (e.g., cyclone wipes out the crop and the yield of improved variety under this disastrous condition is very low);
- Extraordinary observation – A few respondents who are exceptionally different from others (e.g., a small farmer who uses Internet to gather recent developments in agriculture is likely to score high on the INNOVATIVENESS scale);
- Unique in combination – Retain them for analysis.

Types of outliers (Singh and Upadhyaya, 2012)

(a) Global outlier or point anomaly

- Observations which deviate from the rest of the entire data set

(b) Contextual outlier or conditional outlier

- An outlier deviates significantly based on a selected context or behaviour;
- For example, an illiterate farmer showing a high level of comprehension of a printed extension folder;
- Used mostly in time-series data and spatial data.

Two types

- Contextual attributes: defines the context (for example, time and location);
- Behavioral attributes: characteristics of the object, used in outlier evaluation (example, knowledge and intelligence).

(c) Collective outliers

- A group of observations which collectively deviate significantly from the whole data set, even if the individual data objects may not be outliers;
- For example, a small farmer showing attributes of large farmers, where several variables of this specific case deviate from the average of other small farmer cases;
- Collective outliers have been explored for sequence data, graph data and spatial data.

(d) Real outliers

- Real outlying observations which are of interest to the researcher;
- Retained for analysis.

(e) Erroneous outliers

- An observation designated incorrectly as an outlier, due to some inherent problem, or some catastrophic failure.

Methods for detecting outliers (Malik et al., 2014)

Outliers can be detected in two ways.

(i) Based on whether user-labelled examples of outliers can be obtained

1. **Supervised methods:** The training data sets are available and can classify any unusual observation in normal or outlier class by comparing it against a developed model;
2. **Semi-supervised methods:** The training data sets are available only for normal class and any unusual observation is compared against the normal class;
3. **Unsupervised methods:** do not require training data, and thus are most widely applicable. Make the implicit assumption that the data from the normal instances are far more frequent as compared to the outliers.

(ii) Based on assumptions about normal data and outliers

1. Statistical methods

- To detect the outlying observations, and analyze them to study the complete dataset based on them;
- Assume that the normal data follow some statistical model and the data not following the model are outliers;
- Methods – Box plot, maximum likelihood, Grubbs test, Mahalanobis distance and Chi Square test.

2. Proximity-based methods

- An object is an outlier if the nearest neighbors of the object are far away, i.e., the proximity of the object significantly deviates from the proximity of most of the other objects in the same data set;
- Three methods – K Nearest Neighbor analysis (KNN), Clustering method and density based methods.

3. Parametric techniques

- Fitting the data into a model and identifying outliers;
- Popular method – Regression analysis.

4. Non-parametric methods

- No assumptions about distribution of data;
- Methods – histograms, Kernel Density Function or Kernel Feature Space.

5. Distance-based Methods

- Most widely accepted and frequently used techniques in machine learning and data mining;
- Based on Nearest Neighbour principle.

6. Density-based Methods

- Complex method based on Local Outlier Factor (LOF).

7. Clustering-based methods

- Partitioning Clustering Method – Partitioning methods, various centroid based methods, medoids

based methods, PAM, CLARA, k-means, and CLARANS, etc., methods are used;

- Two types – Agglomerative methods and divisive methods;
- Hierarchical Methods – MST clustering, CURE, CHAMALEON and BIRCH.

Statistical methods for detecting outliers

(i) Univariate outliers

- Concerned with a single variable;
- The distribution of observation for every variable is examined as well as select cases falling at the outer ranges (high – low).

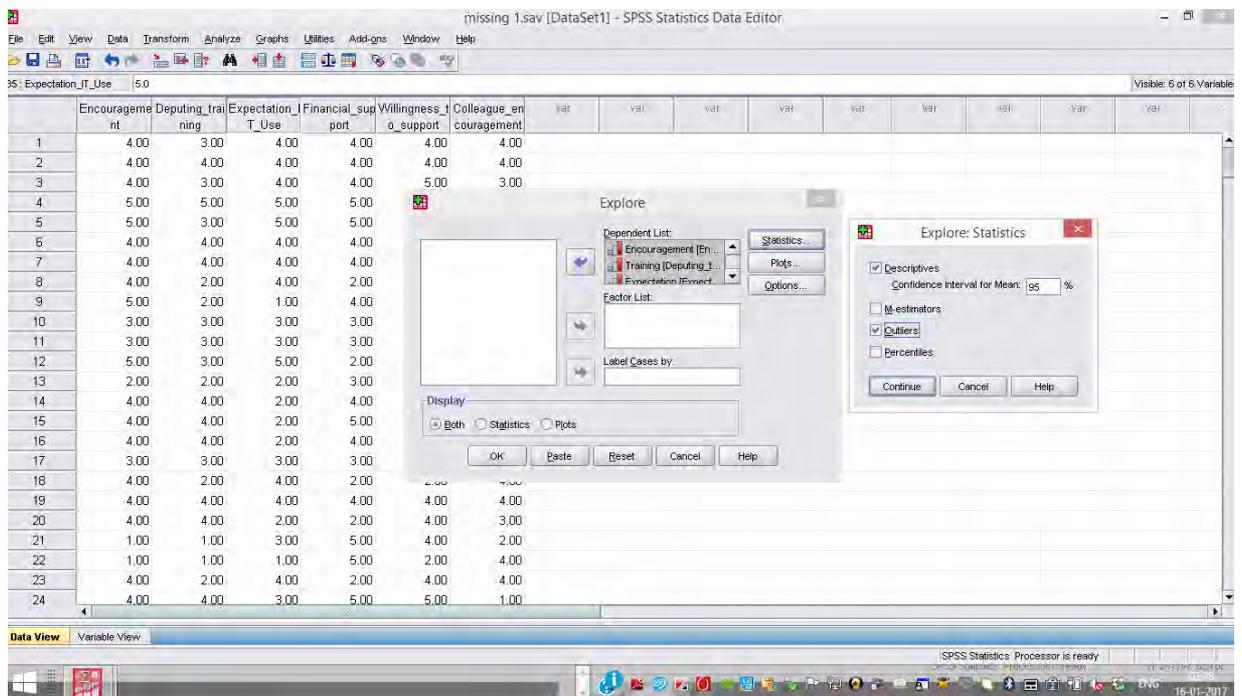
Methods

(i) Box plot

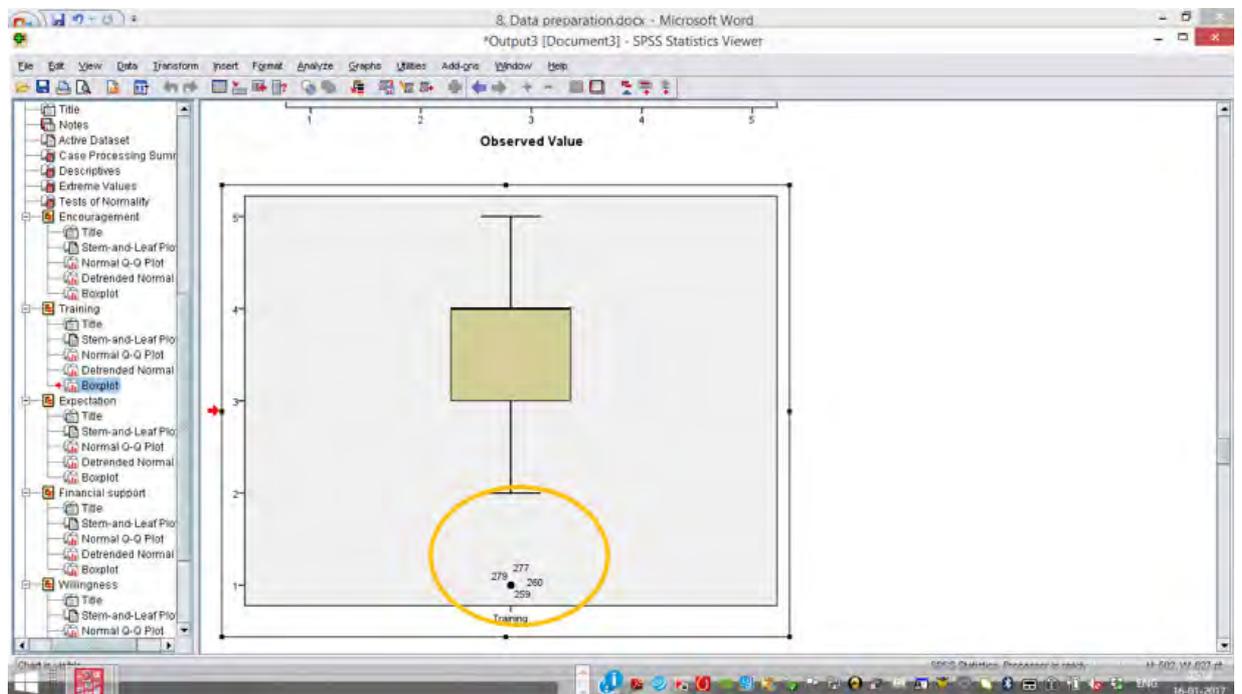
1. Open the SPSS data sheet and select Analyze> Descriptive statistics> Explore

The screenshot shows the SPSS Statistics Data Editor interface. The 'Analyze' menu is open, and the path 'Analyze > Descriptive Statistics > Explore' is highlighted. The data editor shows a table with 24 rows and 6 columns. The first column is labeled '35: Expectation_JT_Use' and has a value of 5.0. The second column is labeled 'Encouragement' and has a value of 4.0. The third column is labeled 'Deputin' and has a value of 5.0. The fourth column is labeled 'gue_ent' and has a value of 4.0. The fifth column is labeled 'ement' and has a value of 4.0. The sixth column is labeled 'var' and has a value of 4.0. The seventh column is labeled 'var' and has a value of 4.0. The eighth column is labeled 'var' and has a value of 4.0. The ninth column is labeled 'var' and has a value of 4.0. The tenth column is labeled 'var' and has a value of 4.0. The eleventh column is labeled 'var' and has a value of 4.0. The twelfth column is labeled 'var' and has a value of 4.0. The thirteenth column is labeled 'var' and has a value of 4.0. The fourteenth column is labeled 'var' and has a value of 4.0. The fifteenth column is labeled 'var' and has a value of 4.0. The sixteenth column is labeled 'var' and has a value of 4.0. The seventeenth column is labeled 'var' and has a value of 4.0. The eighteenth column is labeled 'var' and has a value of 4.0. The nineteenth column is labeled 'var' and has a value of 4.0. The twentieth column is labeled 'var' and has a value of 4.0. The twenty-first column is labeled 'var' and has a value of 4.0. The twenty-second column is labeled 'var' and has a value of 4.0. The twenty-third column is labeled 'var' and has a value of 4.0. The twenty-fourth column is labeled 'var' and has a value of 4.0. The twenty-fifth column is labeled 'var' and has a value of 4.0. The twenty-sixth column is labeled 'var' and has a value of 4.0. The twenty-seventh column is labeled 'var' and has a value of 4.0. The twenty-eighth column is labeled 'var' and has a value of 4.0. The twenty-ninth column is labeled 'var' and has a value of 4.0. The thirtieth column is labeled 'var' and has a value of 4.0. The thirty-first column is labeled 'var' and has a value of 4.0. The thirty-second column is labeled 'var' and has a value of 4.0. The thirty-third column is labeled 'var' and has a value of 4.0. The thirty-fourth column is labeled 'var' and has a value of 4.0. The thirty-fifth column is labeled 'var' and has a value of 4.0. The thirty-sixth column is labeled 'var' and has a value of 4.0. The thirty-seventh column is labeled 'var' and has a value of 4.0. The thirty-eighth column is labeled 'var' and has a value of 4.0. The thirty-ninth column is labeled 'var' and has a value of 4.0. The fortieth column is labeled 'var' and has a value of 4.0. The forty-first column is labeled 'var' and has a value of 4.0. The forty-second column is labeled 'var' and has a value of 4.0. The forty-third column is labeled 'var' and has a value of 4.0. The forty-fourth column is labeled 'var' and has a value of 4.0. The forty-fifth column is labeled 'var' and has a value of 4.0. The forty-sixth column is labeled 'var' and has a value of 4.0. The forty-seventh column is labeled 'var' and has a value of 4.0. The forty-eighth column is labeled 'var' and has a value of 4.0. The forty-ninth column is labeled 'var' and has a value of 4.0. The fiftieth column is labeled 'var' and has a value of 4.0. The fifty-first column is labeled 'var' and has a value of 4.0. The fifty-second column is labeled 'var' and has a value of 4.0. The fifty-third column is labeled 'var' and has a value of 4.0. The fifty-fourth column is labeled 'var' and has a value of 4.0. The fifty-fifth column is labeled 'var' and has a value of 4.0. The fifty-sixth column is labeled 'var' and has a value of 4.0. The fifty-seventh column is labeled 'var' and has a value of 4.0. The fifty-eighth column is labeled 'var' and has a value of 4.0. The fifty-ninth column is labeled 'var' and has a value of 4.0. The sixtieth column is labeled 'var' and has a value of 4.0. The sixty-first column is labeled 'var' and has a value of 4.0. The sixty-second column is labeled 'var' and has a value of 4.0. The sixty-third column is labeled 'var' and has a value of 4.0. The sixty-fourth column is labeled 'var' and has a value of 4.0. The sixty-fifth column is labeled 'var' and has a value of 4.0. The sixty-sixth column is labeled 'var' and has a value of 4.0. The sixty-seventh column is labeled 'var' and has a value of 4.0. The sixty-eighth column is labeled 'var' and has a value of 4.0. The sixty-ninth column is labeled 'var' and has a value of 4.0. The seventieth column is labeled 'var' and has a value of 4.0. The seventy-first column is labeled 'var' and has a value of 4.0. The seventy-second column is labeled 'var' and has a value of 4.0. The seventy-third column is labeled 'var' and has a value of 4.0. The seventy-fourth column is labeled 'var' and has a value of 4.0. The seventy-fifth column is labeled 'var' and has a value of 4.0. The seventy-sixth column is labeled 'var' and has a value of 4.0. The seventy-seventh column is labeled 'var' and has a value of 4.0. The seventy-eighth column is labeled 'var' and has a value of 4.0. The seventy-ninth column is labeled 'var' and has a value of 4.0. The eightieth column is labeled 'var' and has a value of 4.0. The eighty-first column is labeled 'var' and has a value of 4.0. The eighty-second column is labeled 'var' and has a value of 4.0. The eighty-third column is labeled 'var' and has a value of 4.0. The eighty-fourth column is labeled 'var' and has a value of 4.0. The eighty-fifth column is labeled 'var' and has a value of 4.0. The eighty-sixth column is labeled 'var' and has a value of 4.0. The eighty-seventh column is labeled 'var' and has a value of 4.0. The eighty-eighth column is labeled 'var' and has a value of 4.0. The eighty-ninth column is labeled 'var' and has a value of 4.0. The ninetieth column is labeled 'var' and has a value of 4.0. The ninety-first column is labeled 'var' and has a value of 4.0. The ninety-second column is labeled 'var' and has a value of 4.0. The ninety-third column is labeled 'var' and has a value of 4.0. The ninety-fourth column is labeled 'var' and has a value of 4.0. The ninety-fifth column is labeled 'var' and has a value of 4.0. The ninety-sixth column is labeled 'var' and has a value of 4.0. The ninety-seventh column is labeled 'var' and has a value of 4.0. The ninety-eighth column is labeled 'var' and has a value of 4.0. The ninety-ninth column is labeled 'var' and has a value of 4.0. The hundredth column is labeled 'var' and has a value of 4.0.

2. Drag the variables into 'Dependent list' box and click on 'Statistics' and check the box 'Outliers' and click 'Continue'



3. Examine the output 'Boxplot' for each variable. The unusual observations will be found both 'above' and 'below' the plot.



Identification of outliers

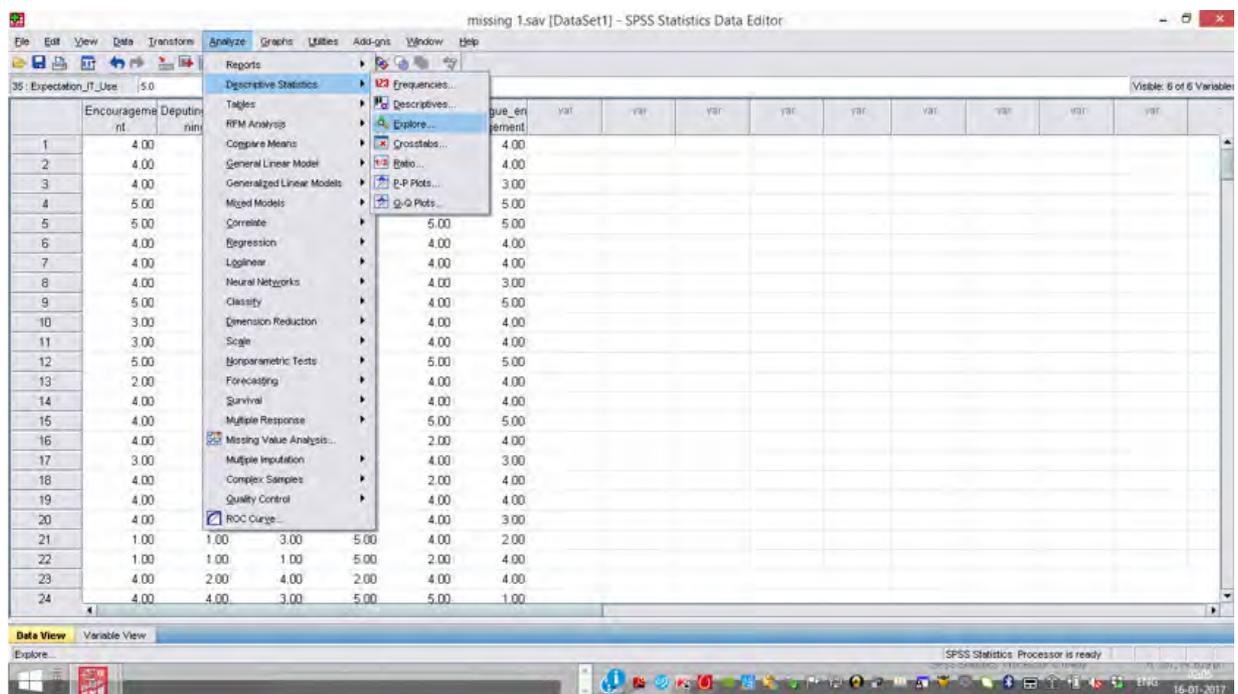
- Any value, which is beyond the range of $-1.5 \times$ Inter Quartile Range to $1.5 \times$ Inter Quartile Range are outliers;
- Any value which is out of range of 5th and 95th percentile can be considered as an outlier;
- Data points, three or more standard deviation away from mean are considered an outlier.
- Based on the above guidelines, the cases 259, 260, 277, 279 are identified as outliers.

(ii) Trimmed mean method

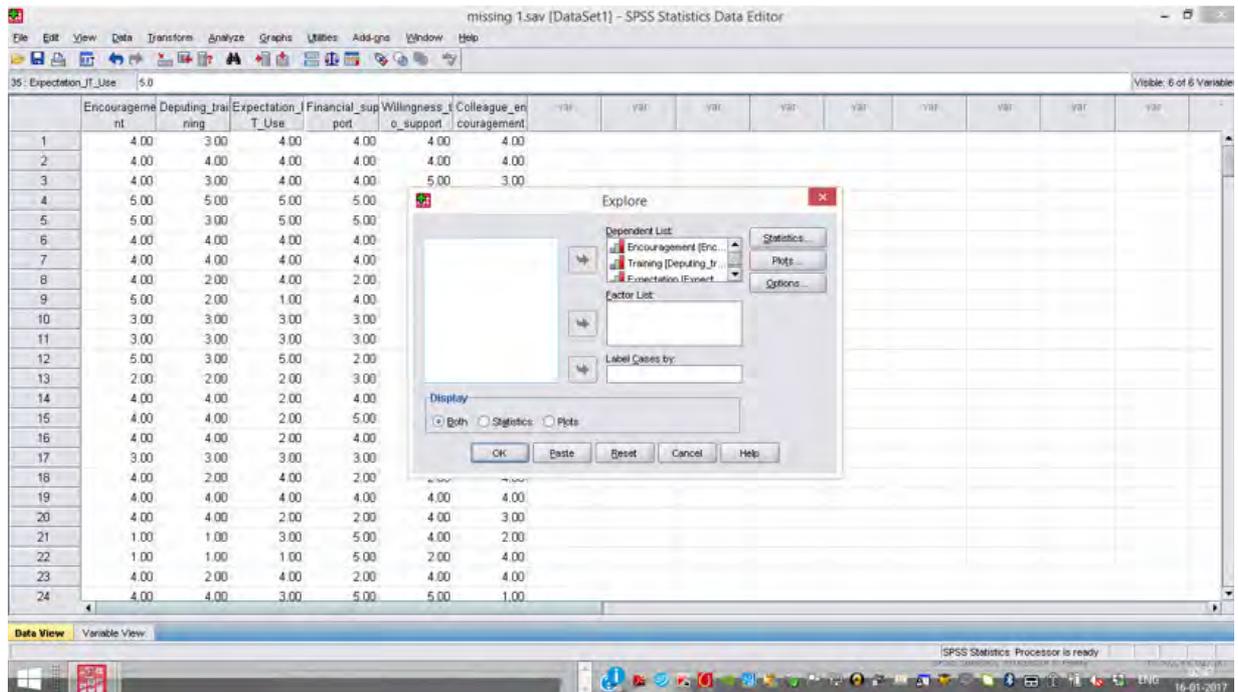
- Used to determine the extent of a problem likely to be created by the outliers;
- The original mean of a variable is compared with the 5% trimmed mean (the new mean calculated after the top and bottom 5 percent of cases are removed from the distribution);
- If both means are similar, it can be concluded that the outlying values are not too different from the distribution, which can be retained (Walfish, 2006).

Trimmed mean method in SPSS

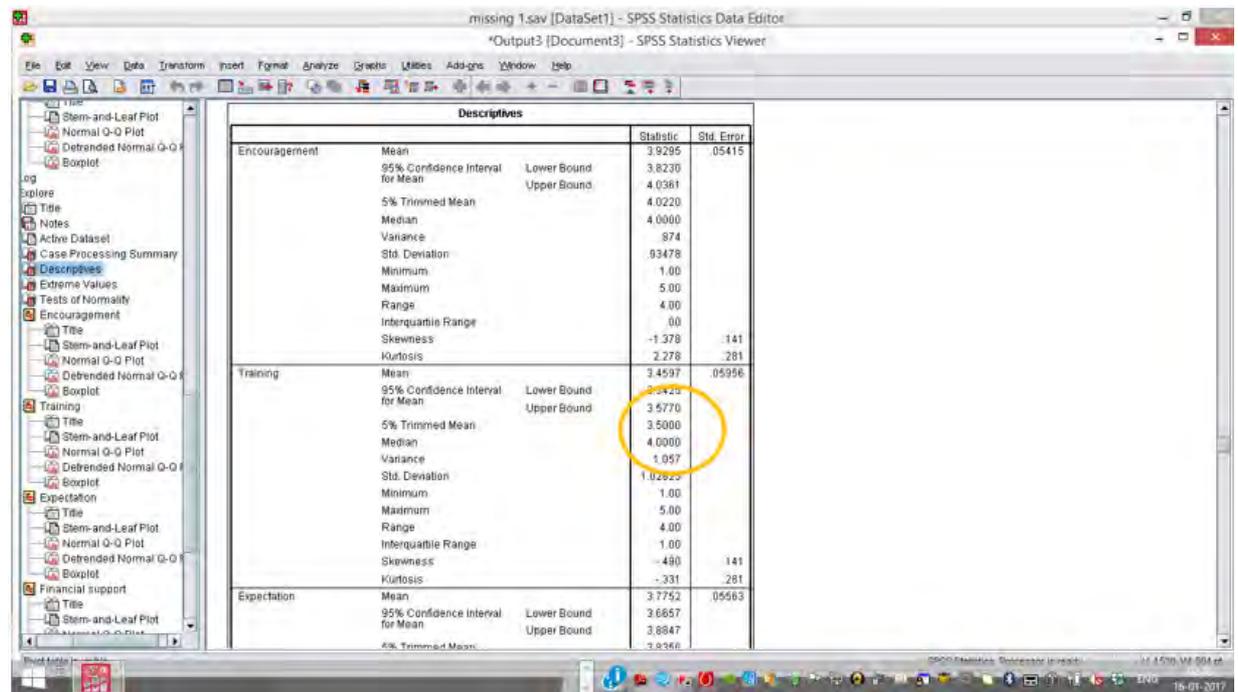
1. Open the SPSS data sheet and select Analyze> Descriptive statistics> Explore



2. Drag the variables into Dependent list box and click on Statistics and check the box outliers and click Continue.



3. In the output document section Descriptive will provide 5% trimmed mean value for each variable.



Since the outliers were identified for Training in the previous section, now delete cases 259, 260, 277, 279 of training variable, and repeat the procedure 1 through 3.

The screenshot shows the SPSS Statistics Data Editor window with the 'Descriptives' table. The table is divided into three sections: Encouragement, Training, and Expectation. The 'Training' section has a yellow circle around the 5% Trimmed Mean value of 3.5113.

Variable	Statistic	Std. Error	
Encouragement	Mean	3.9558	05194
	95% Confidence Interval for Mean	Lower Bound	3.8536
		Upper Bound	4.0580
	5% Trimmed Mean	4.0404	
	Median	4.0000	
	Variance	.793	
	Std. Deviation	.89069	
	Minimum	1.00	
	Maximum	5.00	
	Range	4.00	
Interquartile Range	.00		
Skewness	-1.343	.142	
Kurtosis	2.424	.283	
Training	Mean	3.4796	05836
	95% Confidence Interval for Mean	Lower Bound	3.3041
		Upper Bound	3.5944
	5% Trimmed Mean	3.5113	
	Median	4.0000	
	Variance	1.001	
	Std. Deviation	1.00004	
	Minimum	1.00	
	Maximum	5.00	
	Range	4.00	
Interquartile Range	1.00		
Skewness	-.458	.142	
Kurtosis	-.346	.283	
Expectation	Mean	3.7925	05455
	95% Confidence Interval for Mean	Lower Bound	3.6852
		Upper Bound	3.8999
	5% Trimmed Mean	3.9477	

After removal of outliers, the 5% trimmed mean has increased to 3.5113 which is close to the original mean of 3.500. Thus, these cases are not considered as serious outliers and so retained for the analysis.

(ii) Bivariate and multivariate outlier detection

Bivariate detection

- Required for correlation, regression and other analyses involving two variables;
- Outliers can be identified through Scatter plots.

Multivariate detection

- Required for conducting confirmatory factor analysis in Structural Equation Modeling;
- Methods – Mahalanobis D^2 , Cook's D and Madira's coefficient of multivariate kurtosis (For Structural Equation Modeling);
- Thresholds for D^2 / df – Small sample 2.5; large sample – 4 (Hair Jr. et al., 2006).

Treatment of outliers

(i) **Ignore outliers:** If the global outliers are satisfying 5% trimmed mean criteria, they may be retained.

(ii) Deleting observations

- Delete outlier values only if it is due to data entry error, data processing error, or if outlier observations are very small in number.
- The contextual outliers may be deleted.

(iii) Transforming and binning values

- Transforming variables can also eliminate outliers;
- Natural log of a value reduces the variation caused by extreme values;
- Binning is also a form of variable transformation;
- Decision Tree algorithm allows to deal with outliers well due to binning of variable;
- We can also use the process of assigning weights to different observations.

12.3 Key Points

- Outliers are the observations with a unique combination of characteristics identifiable as distinctly different from other observations in the data sheet.
- Outliers can bias or influence estimates that may be of substantive interest. Moreover they can also impact the basic assumption of regression, ANOVA and other statistical models.
- Various outlier treatment methods, including trimmed mean method along with others, may be used based on the magnitude of the problem caused by outliers.

References

- Hair, Jr., J. F., Black, W. C., Babin, B. J., Anderson, R. E., and Tatham, R. L. (2006.) *Multivariate data analysis*. Sixth edition. Upper Saddle River, NJ: Pearson Prentice Hall.
- Malik, K., Sadawarti, H., and Kalra G S. (2014.) Comparative analysis of outlier detection techniques. *International Journal of Computer Applications* 97(8):12-21.
- Ray, S. (2016.) A comprehensive guide to data exploration. (Retrieved from: <https://www.analyticsvidhya.com/blog/2016/01/guide-data-exploration/#one>).
- Singh, K. and Upadhyaya. S., (2012.) Outlier detection: Applications and techniques. *International Journal of Computer Science Issues* 9:307-323.
- Walfish, S. (2006.) A review of statistical outlier methods. *Pharmaceutical technology*. (Retrieved from <http://www.pharmtech.com/review-statistical-outlier-methods>).



Photo Credit: Sub

13

Data Exploration and Preparation – Testing Statistical Assumptions

P Sethuraman Sivakumar

Objectives

1. Make readers understand the importance of checking and cleaning data prior to statistical analysis
2. Explain various methods and tools for examining the data for statistical assumption prior to analysis

13.1 Introduction

- Testing statistical assumptions for data analysis is an essential step in the data exploration and preparation process.
- As many statistical methods work on specific assumptions, it is necessary to test the data to check that it complies with these distributional requirements.
- This module provides an overview of various normality testing and linearity assessment and data transformation methods.

13.2 Discussion

Testing assumptions for statistical analysis

- The inferential statistical methods are based on certain assumptions representing the requirements of the statistical theory.
- Four major assumptions – normality, linearity and absence or correlated errors.

I. Normality testing

- Many parametric statistics require normally distributed data for performing analysis;
- Type of normality – Univariate (all parametric analysis) and Multivariate (Structural Equation Modeling).

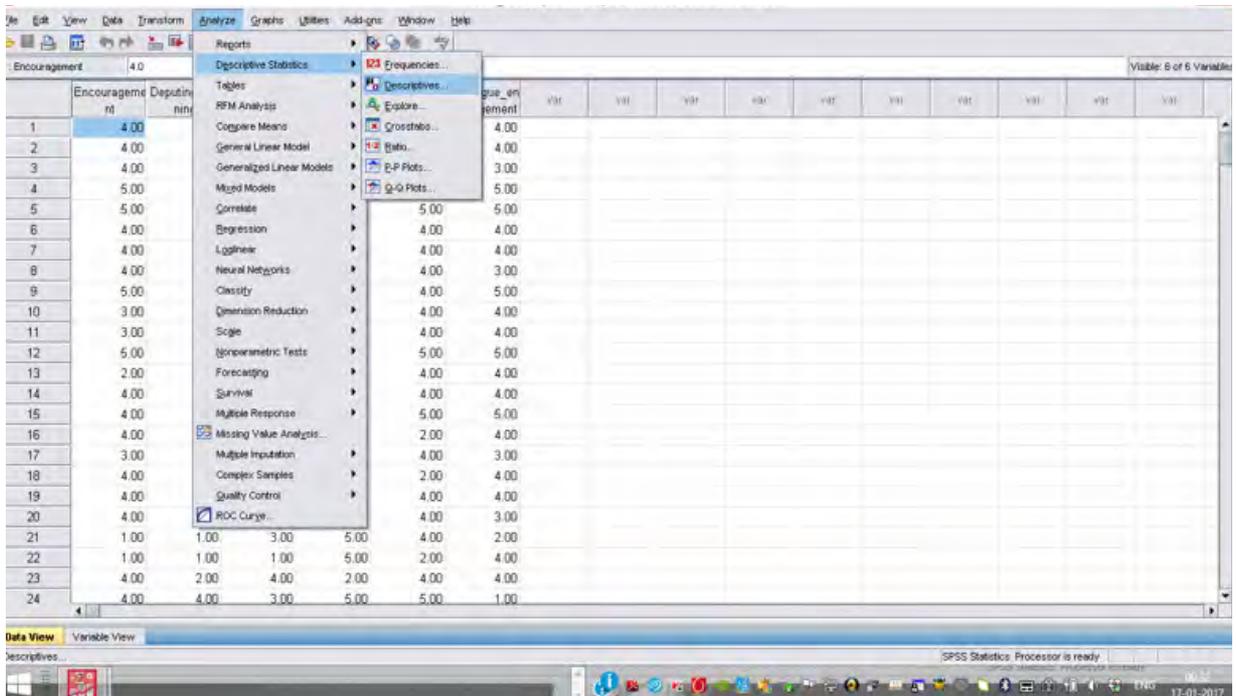
Methods for testing normality

a. Univariate normality

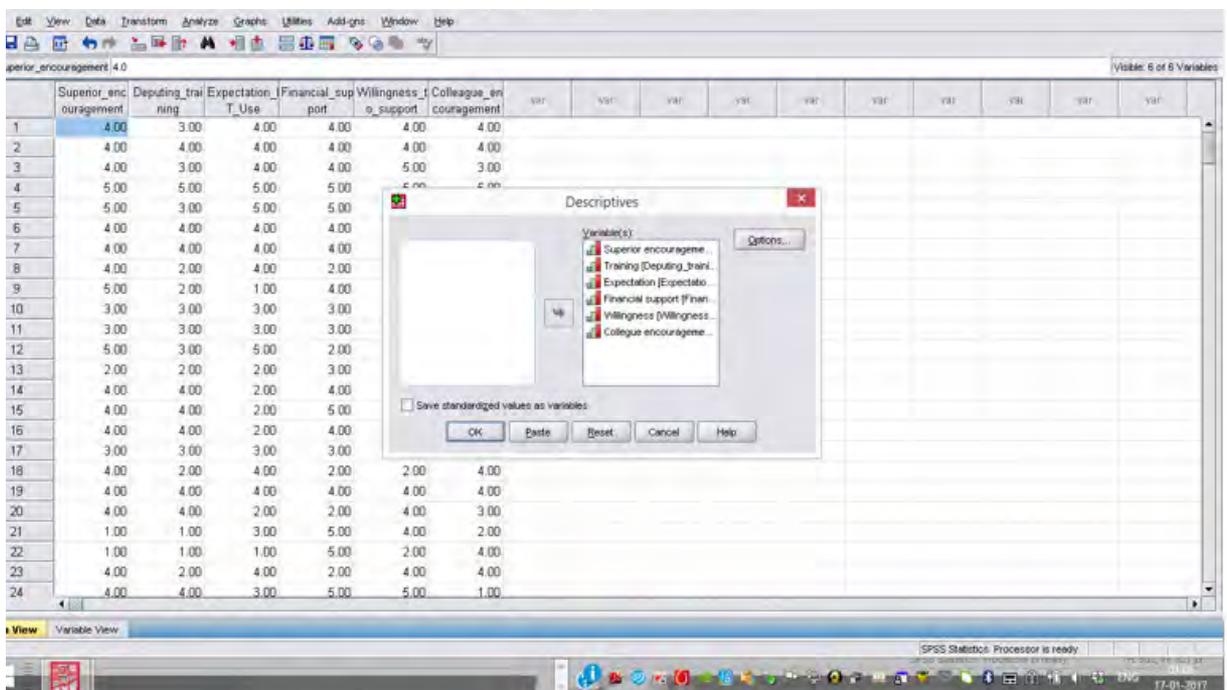
i. Skewness and kurtosis

- If skewness and kurtosis exceed the absolute value of 1, we assume that there is non-normality;
- Procedure for calculating skewness and kurtosis using SPSS:

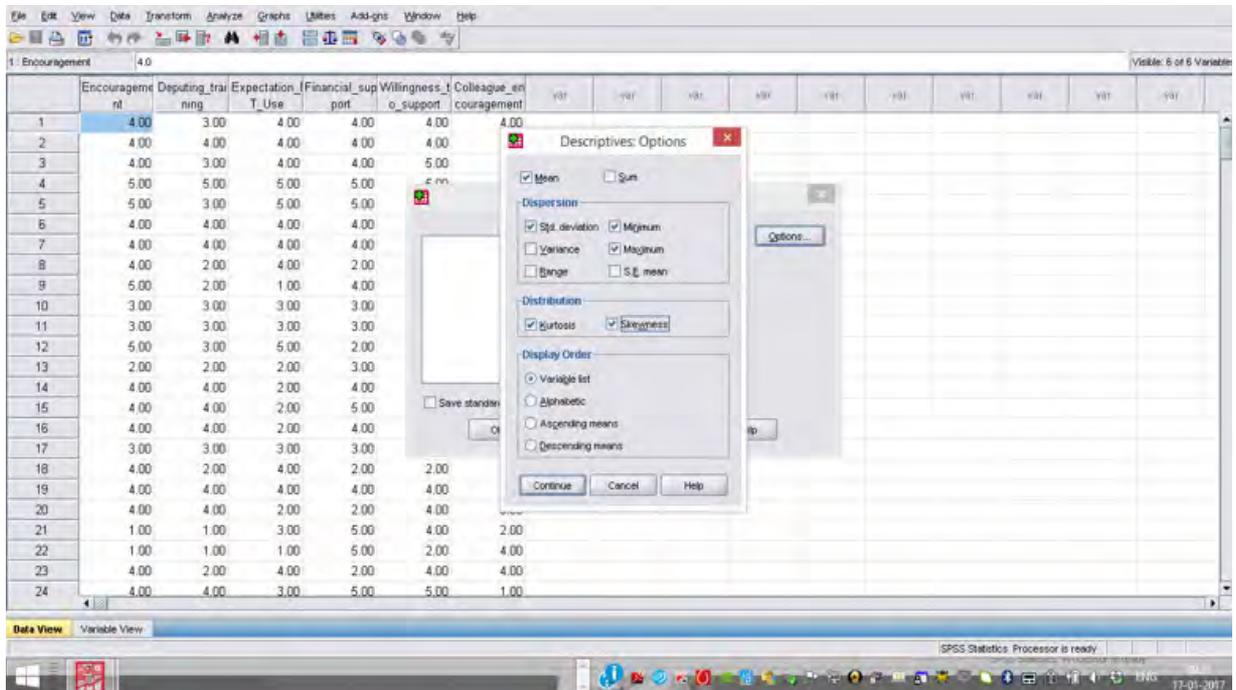
1. Select **Analyse > Descriptive statistics > Descriptives**



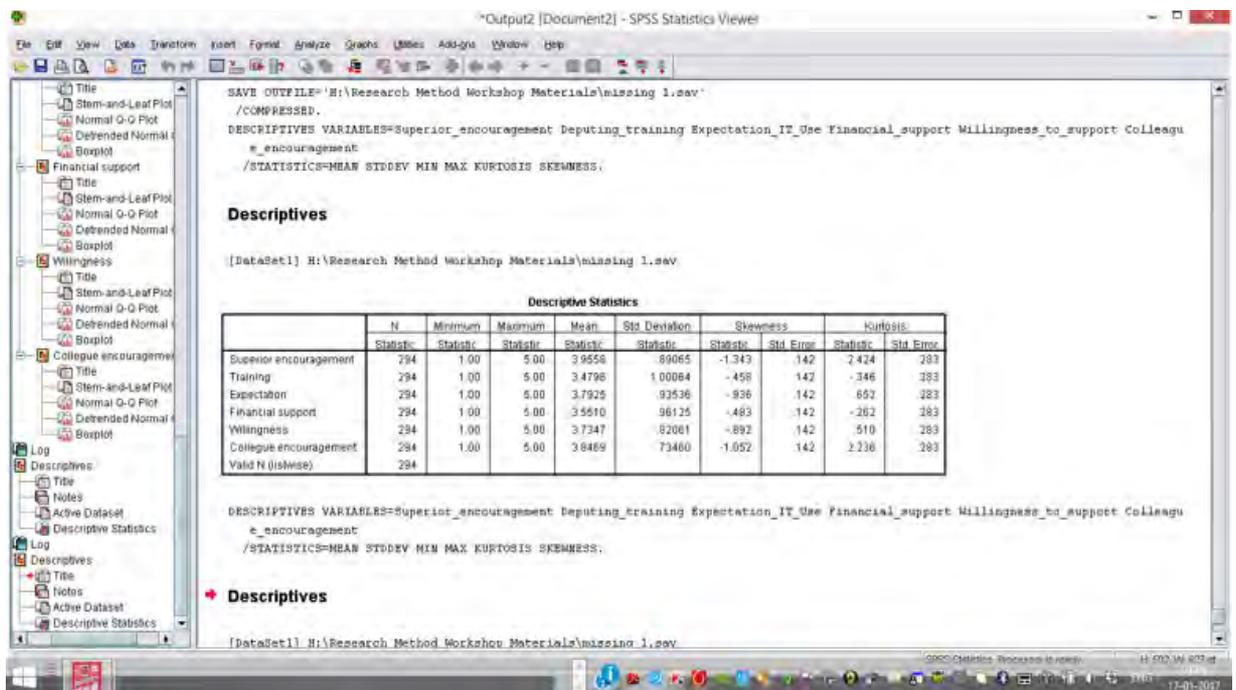
2. Drag all variables into **Variables** box.



3. Click **Options** and check the boxes next to **skewness** and **kurtosis**. Click **Continue**> **OK**.



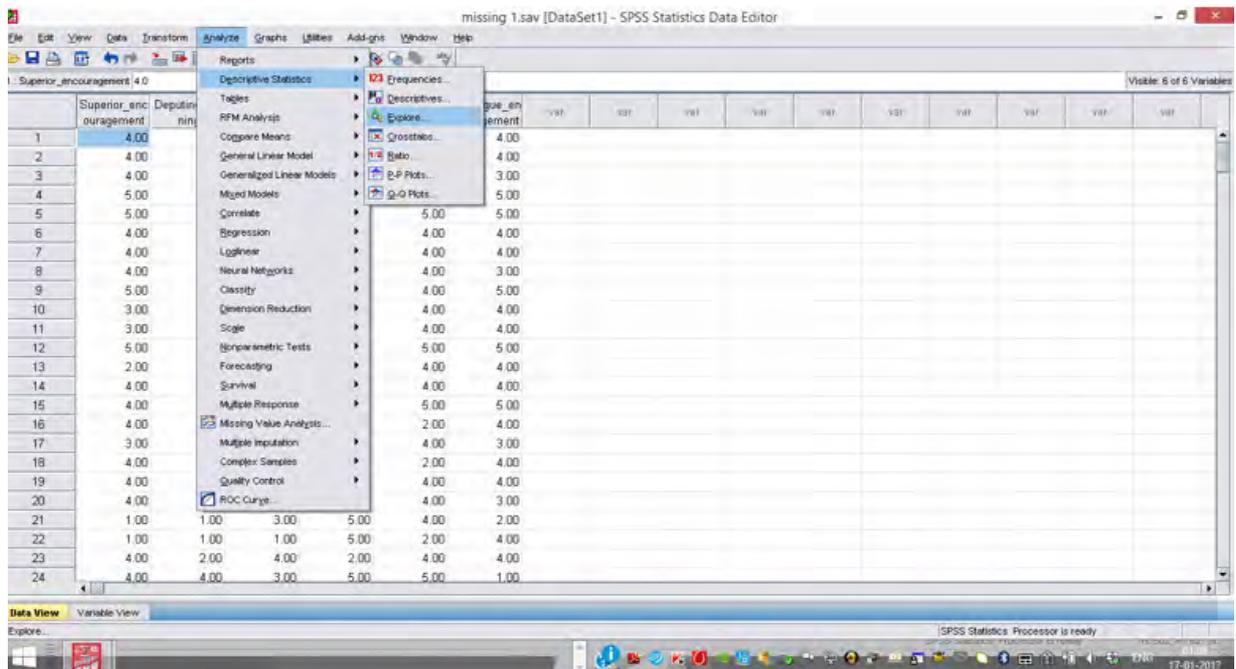
4. From the Output, the skewness and kurtosis values of all the variables may be analysed. The Output table indicates the presence of non-normality in two variables – superior encouragement and Collegial encouragement.



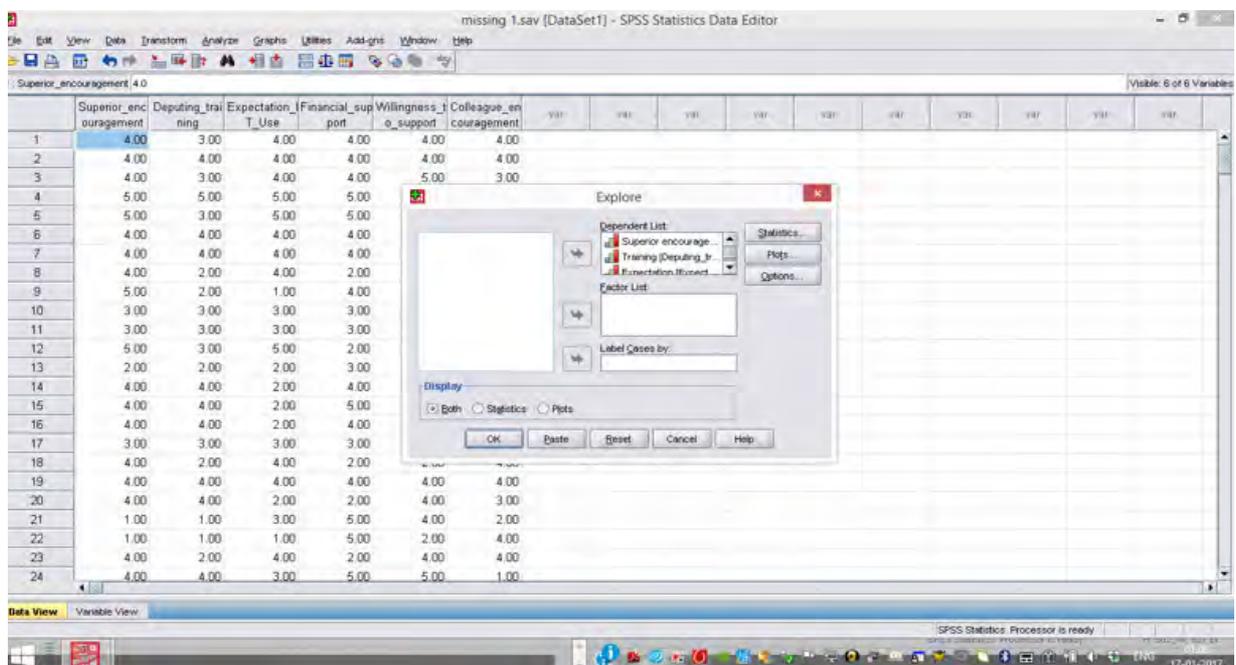
ii. Normal probability plot

- The normal probability plot (Chambers et al., 1983) is a graphical technique for assessing whether or not a data set is approximately normally distributed;
- The data are plotted against a theoretical normal distribution in such a way that the points should form an approximate straight line;
- Departures from this straight line indicate non-normality.
- The SPSS procedure for Normal Probability plot is described below:

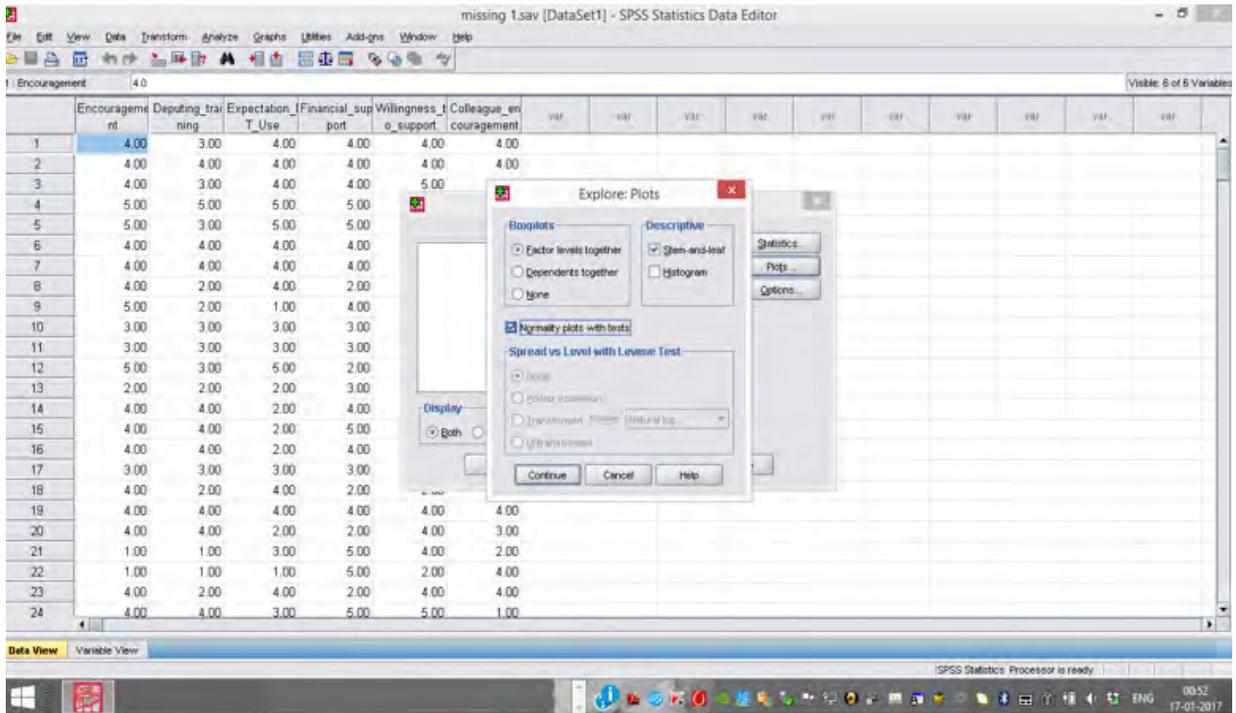
1. Select **Analyze > Descriptive Statistics > Explore**



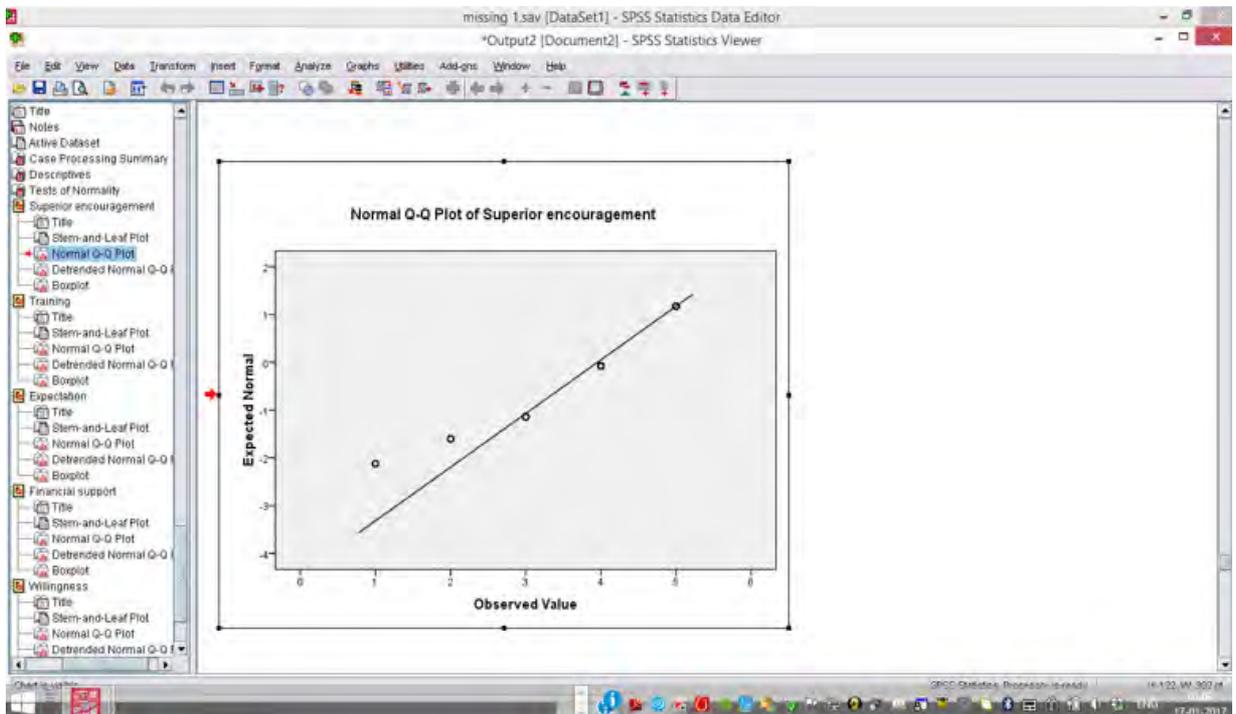
2. Drag all the variables into **Dependent box**.



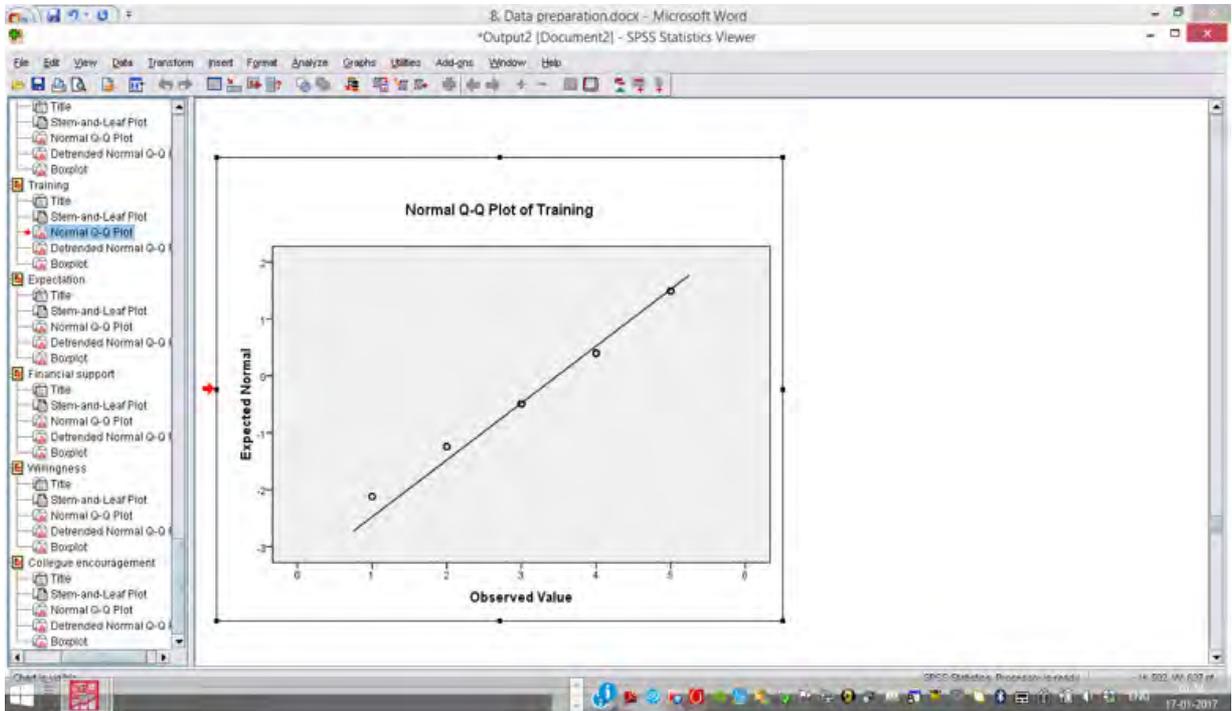
3. Select **Plots** and check **Normality plots with tests**. Click **Continue**> **OK**.



4. In the output pane, select **Normal Q-Q Plot** for each variable



The Normal Q-Q Plot of Superior encouragement indicates presence of non-normality.



The Normal Q-Q Plot of training indicates Normality.

b. Bivariate normality

- Scatter plots are widely used to detect bivariate normality.

c. Multivariate normality

- The multivariate normality of the data is assessed through Mahalanobis Distance (D²) and Mardia's multivariate kurtosis.
- Mahalanobis Distance – A detailed description of calculating Mahalanobis Distance is described in Arifin (2015).
- Mardia's multivariate kurtosis – The Mardia's coefficient can be estimated through SPSS AMOS and a detailed procedure.

II. Linearity

- The variable should be linearly related with other variables.
- Important assumption for correlation-based techniques.
- Used for multiple regression, logistic regression, factor analysis and structural equation modeling.

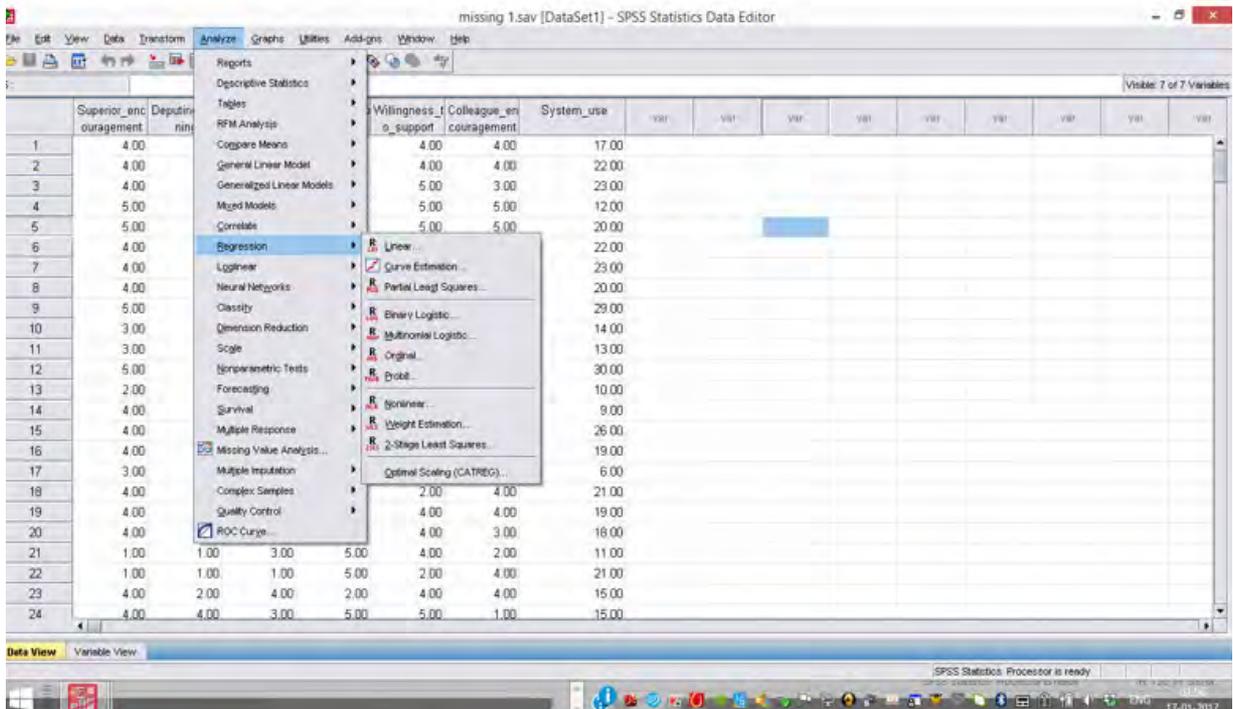
III. Multicollinearity

- Indicates the presence of high correlation between independent variables.
- Undesirable character which causes inaccurate prediction.
- Essential requirement for OLS estimation (Multiple regression).
- Presence of multicollinearity – Tolerance value < 0.2 and VIF > 5.

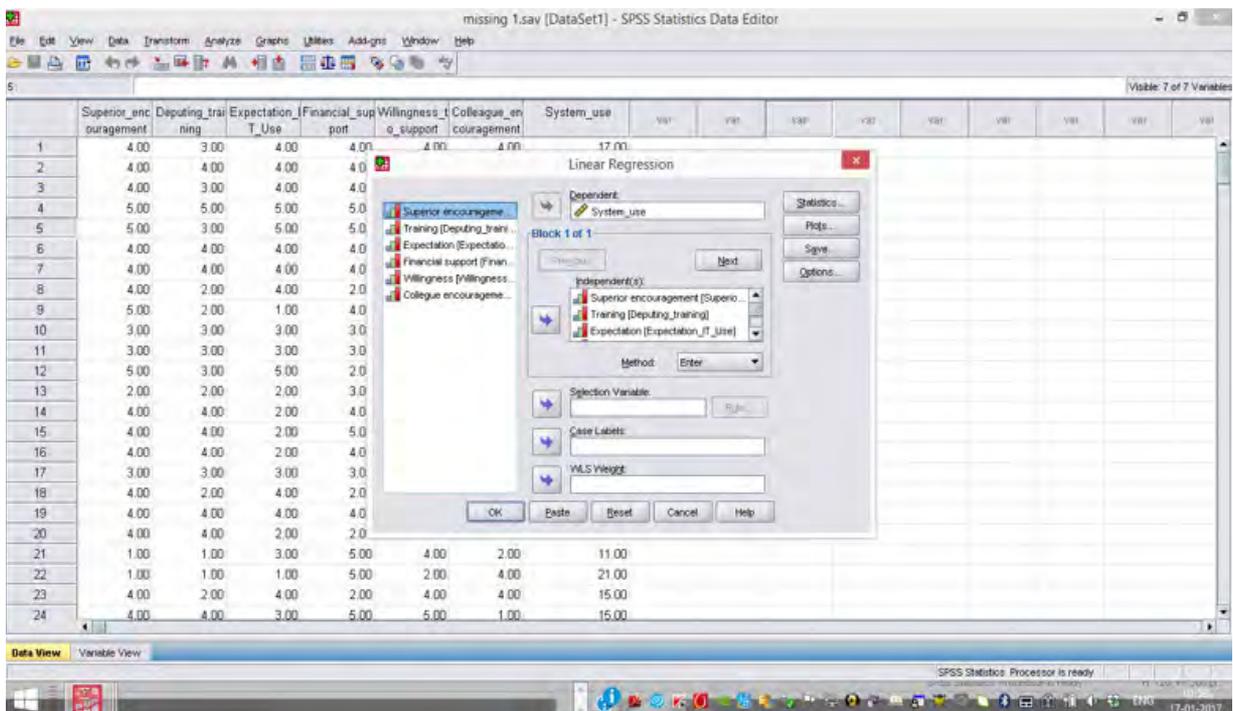
SPSS tests both linearity and multicollinearity through regression analysis.

1. In this example, the Computer use (System use) is added along with other organisational support variables.

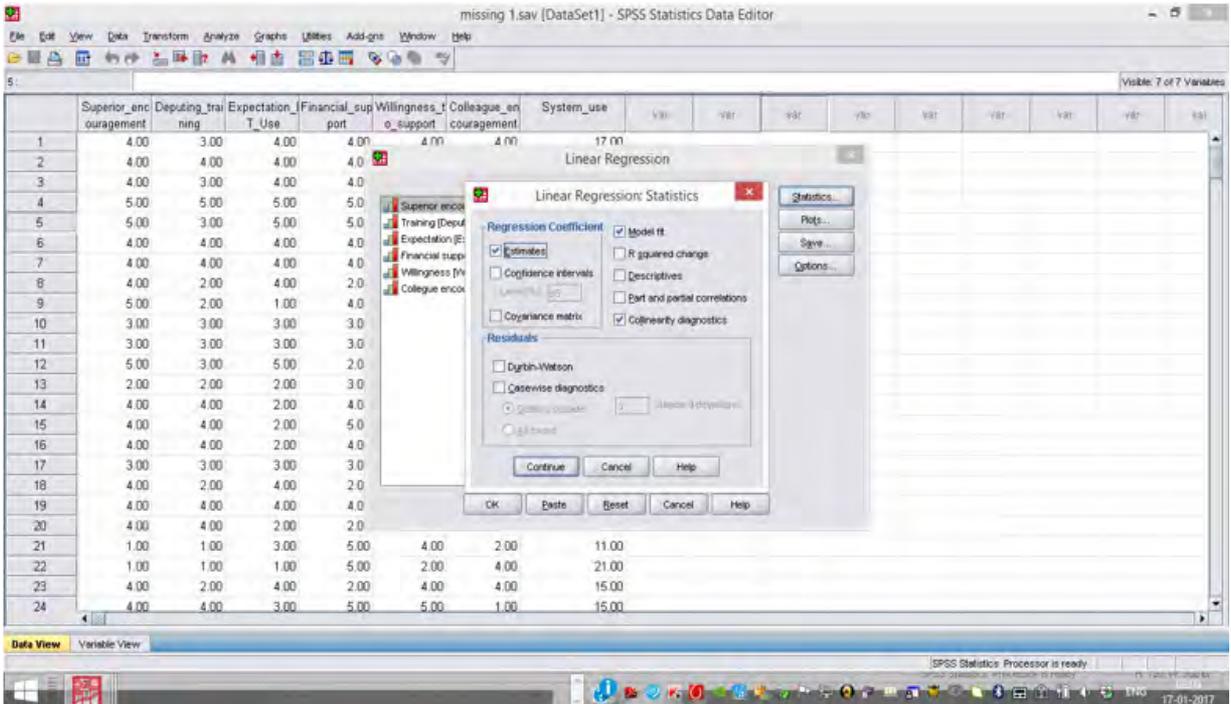
2. For multiple regression analysis, select **Analyze > Regression > Linear**



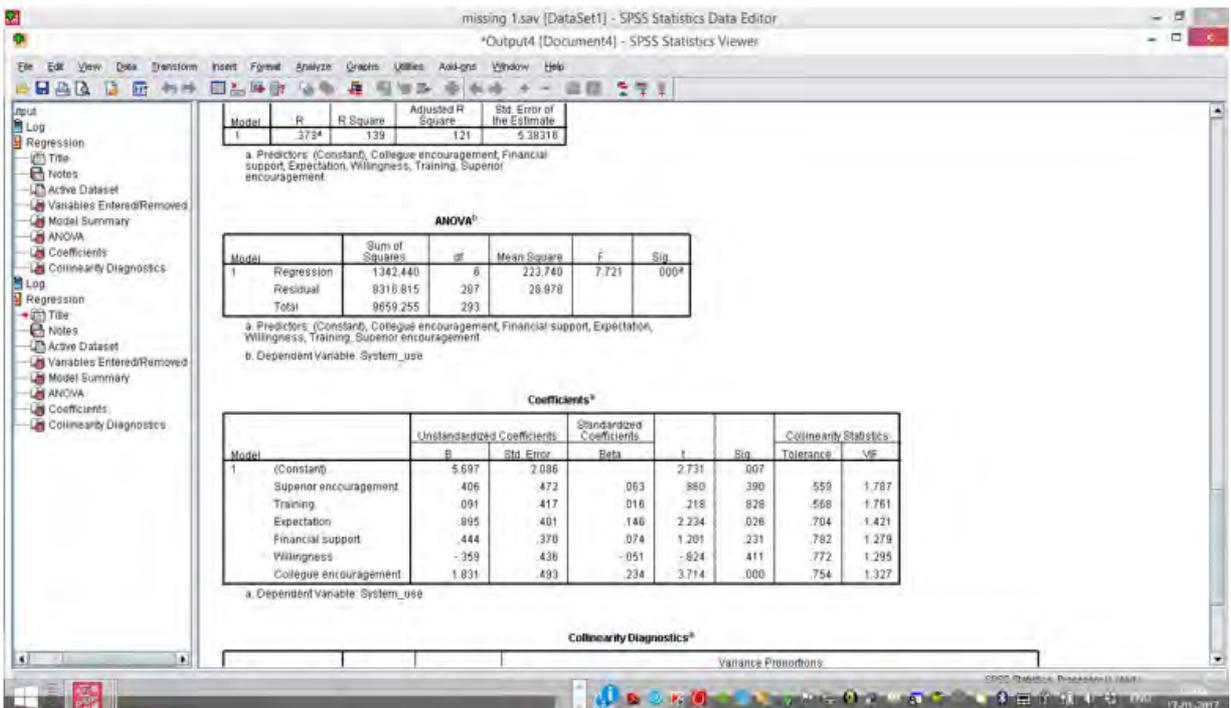
3. Drag the variable 'System use' into **Dependent variable box** and others into **Independents box**.



4. Click **Statistics**, check the **Collinearity Diagnosis**. **Continue**> **OK**



5. The significant ANOVA model and regression coefficients indicate the presence of linearity among variables.



- In this analysis, a significant ANOVA model and the expectation and colleague encouragement factors were linearly related.
- Regarding tolerance levels, a Tolerance value < 0.2 and VIF > 5 indicate presence of

multicollinearity. In this example, all the variables are free from multi-collinearity.

Data transformations

Purpose

- To correct errors in the data;
- To improve the relationship among variables.

Methods of transformations

Non-normality

- Flat distribution – Inverse ($1/X$ or $1/Y$);
- Negatively skewed transformations – Squared or cubed;
- Positive skewness – Logarithm or square root.

13.3 Key Points

- Testing the data to check if it satisfies statistical assumption is an important component of data exploration and preparation phase.
- Univariate and multivariate normalities are necessary assumptions for factor and confirmatory factor analyses.

References

- Chambers, J., Cleveland, W., Kleiner, B., and Tukey, P. (1983.) Graphical methods for data analysis. Wadsworth.
- Arifin, W. N. (2015.) The graphical assessment of multivariate normality using SPSS. Education in Medicine Journal, 7(2):e71-e75.



Photo Credit: Nimisha Mittal

14

Statistical Designs in Data Analysis – Descriptive to Inferential Statistics

A Dhandapani and
P Sethuraman Sivakumar

Objectives

1. Provide an overview of statistical tests available for interpreting the data
2. Provide guidelines for choosing specific tests based on the nature of research

14.1 Introduction

- Statistical techniques offer a wide range of techniques and strategies for classifying, summarising, and drawing inferences from data. Three basic functions of statistics are :
 1. Design – Planning and carrying out of a research study through experimental designs and sampling;
 2. Description – Methods for summarizing data;
 3. Inference – Making predictions or generalizations about a population of individuals or samples.
- In this module, statistical description and inference methods are explained to provide guidelines for researchers to choose specific statistical tests based on research problems.

14.2 Discussion

A. Questions to be answered through statistics

(i) Description of a phenomenon

- Describe a phenomenon by answering questions, such as 'What is going on?' and 'What is this phenomenon?'
- Examples: Status of current fertilizer usage; labour usage for weed management.

(ii) Relationship

- Identifies and analyzes the relationship between two or more variables;
- Starts after describing the phenomenon;
- Examples include: relationship between adoption and extension contact.

(iii) Causation

- One or more variables cause changes in another variable – Cause and effect study;
- Starts after describing a phenomenon and identifies relationships between variables;
- Examples: age and education level cause changes in adoption of a technology.

(iv) Comparisons

- Comparing two or more 'conditions' to test if they are different and identify the 'Best' condition;

Examples:

- o Comparing two forms of fertilizers;
- o Comparison of several varieties at On Farm Trials.

(v) Similarities

Grouping of similar units and dividing dissimilar units;

Check if the variables can be grouped under different categories using a few common characters;

Examples - grouping of varieties and taxonomical classification of insects.

B. Variable Types

Qualitative/ Categorical variables

- Categorical variables take on values that are names or labels;
- Types – Nominal and ordinal;
- Nominal – having 2 or more categories and lack order (Example, Gender);
- Ordinal – Two or more categories and have order (Example, Small, Medium and Large Farmers) but differences are not meaningful;
- Only limited number of statistical methods can be used for categorical variables.

Quantitative/Contiguous variables

The values of a quantitative or continuous variable can be ordered and measured;

1. **Interval** – Quantitative Variables having own scale; lacks '0' and divisions are not meaningful (temperature in Celsius);
2. **Ratio** – Quantitative variables, valid '0' and divisions are meaningful (Yield, Germination, etc.);
3. A few statistical terms are described in Box 14.1.

Box 14.1: A few terms used in statistical analysis

- **Population** – Aggregates of individuals, either living or materials/things;
- **Population parameters** – Constants which describe characteristics of populations;
- **Sample** – A subset of population chosen in a scientific way;
- **Sample size** – The number of units selected for observation in a sample;
- **Variable** – Measurement taken on individuals in the population;
- **Estimator** – Function of sample values which are used to estimate population parameters. Various estimators are possible but estimators with some properties such as Unbiased (roughly indicate the estimator takes true population parameter value on an average), Consistent (the estimator value tends to true value when the sample size is increased to infinity), Efficient (estimator is having less variance than other possible estimators) etc., are preferred for analysis.

C. Types of statistical tests

There are two types of statistical tests:

(i) Descriptive statistics

- Provide descriptions and summarisation of the data, either through numerical calculations or graphs or tables.

Methods:

- Measures of **central tendency** – describe the central position of a frequency distribution for a group of data through mean, median and mode;
- Measures of **spread** – Summarize a group of data by describing how they are spread out.

(ii) Inferential statistics

- Draw inferences and predictions about a population based on a sample of data taken from the population.
- The methods of inferential statistics are given below:

1. Estimation of parameters

- Used to make estimates about population values based on sample data.
- Two types of estimation statistics
 - Confidence intervals – To establish a range that has a known probability of capturing the true population value;
 - Parameter estimation – To make inferences about how well a particular model might describe the relationship between variables in a population; Eg. Multiple regression.

2. Testing of statistical hypotheses

- To make statistical inferences about whether or not the data we gathered support a particular hypothesis;
- T-Test, Chi-Square and ANOVA (analysis of variance).

Choosing a right statistical test

Table 14.1 below provides information on how to choose the appropriate statistical analysis and SAS procedures/R functions for use. Note that many statistical analyses would require several assumptions to be satisfied. One should look for the assumptions of each analysis and validate those assumptions before using the appropriate analysis.

Table 14.1: Guidelines for choosing statistical tests

Population & No. of variables	Type of Dependent Variable	Analysis	SAS Procedure
One Population, one variable (no IV)	Ratio & Interval	Single sample t-test (mean)	TTEST
	Ordinal & Interval	Single sample Median test	UNIVARIATE
	Nominal	Chi-Square Goodness of fit	FREQ

Population & No. of variables	Type of Dependent Variable	Analysis	SAS Procedure
One Population (one IV), Two levels (independent groups)	Ratio & Interval	Two sample t-test / independent t-test (mean)	TTEST
	Ordinal & Interval	Wilcoxon Man Whitney test	UNIVARIATE
One Population, two groups (dependent or matched)/one IV with two levels	Nominal	Chi-Square test	FREQ
	Ratio & Interval	Paired t-test	TTEST
	Ordinal & Interval	Wilcoxon signed rank	UNIVARIATE
One Population, more than 2 groups/One IV with two or more levels	Nominal	McNeamer test	FREQ
	Ratio & Interval	One way ANOVA	GLM(or ANOVA)
	Ordinal & Interval	Kruskal Wallis test	NPAR1WAY
One Population, more than 2 groups (dependent or matched)/One IV with two or more levels	Nominal	Chi-Square test	FREQ
	Ratio & Interval	One way repeated measurements ANOVA	GLM
	Ordinal & Interval	Friedman Test	FREQ
One dependent & two or more IVs	Nominal	Repeated measures Logistic Regression	GENMOD
	Dependent – Ratio or Interval; independent – ordinal/nominal	Factorial ANOVA	GLM
	Dependent – Ratio or Interval; Independent – ordinal/nominal & ratio/interval	Analysis of Covariance/ Multiple Regression	GLM REG
	Dependent – ordinal or nominal; Independent – ordinal/nominal/ratio/interval	Logistic Regression	LOGISTIC
	Ratio/Interval	Correlation Simple Regression	CORR
One dependent and one IV	Ordinal/Interval	Non-parametric Correlation	CORR
	Nominal	Simple logistic regression	LOGISTIC
One or more interval IVs and/or	Interval or ratio	Multiple regression	PROC REG
	One or more categorical IVs	Discriminant analysis	PROC DISCRIM
Two or more related variables	Ratio/Interval	Factor Analysis	FACTOR
Two sets of two or more IVs	Ratio/Interval	Canonical correlation	PROC CANCELL

Note: IV: Independent Variable; All analyses are also available at IBM-SPSS
Source: IDRE-UCLA (2017)

Tests of significance

3. The methods of inference used to support or reject claims based on sample data

Testing of statistical hypotheses

- Statistical hypothesis is a statement about the population;
- Types – Null Hypothesis and alternate hypothesis;
- Various types of hypotheses are explained elsewhere in this manual.

Errors (Type I and Type II)

- When applying a statistical test, two types of errors are committed, known as Type I and Type II error (Fig. 14.1);
- It is not possible to make these errors zero or minimize both of them;
- Usually, one (Type I) error is fixed and Type II error is minimized.



Fig. 14.1: Decisions based on Type I and Type II errors

Significance level

- Probability of committing Type I error is called significance level [Prob (Reject H_0 / H_0 is true) = α]
- α is usually fixed at very small levels such as 0.05 (5%) or 0.01 (1%).

P-Value

- Indicates the exact significance level chosen for the particular statistical test;
- Probability of obtaining the test statistic value if null hypothesis is true;
- Smaller P-Value (< 0.01) indicates evidence against null hypothesis, while larger P-value (> 0.05) indicates otherwise.

Power of a test

- Probability of Rejecting a null hypothesis when it is false [1 - P(committing Type II Error)]
- Power of a test should be high (or Type II error should be minimum).

Critical region

4. Area of the sampling distribution of a statistic that will lead to the rejection of the hypothesis tested when that hypothesis is true (Fig 14.2).

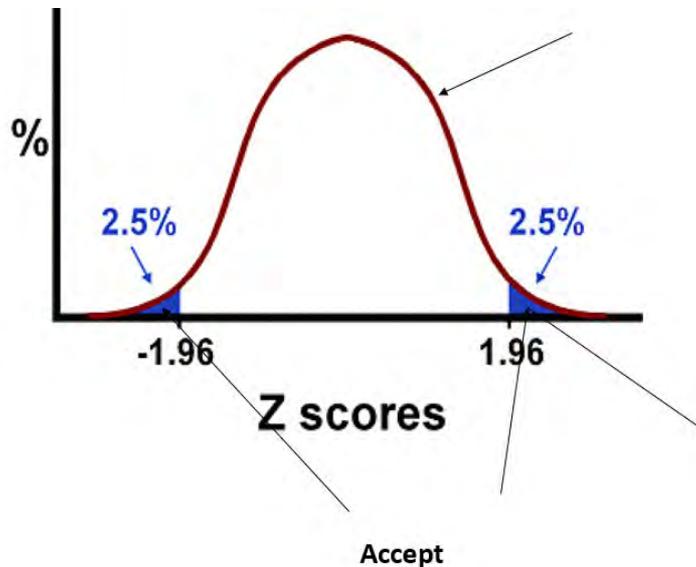


Fig. 14.2: Critical region

Box 14.2: Choosing type of graphs

(i) Line Graph

- Used to track changes in a phenomenon over short and long periods of time as well as to compare changes in specific time periods;
- Suitable to track even small changes over time;
- Example, increase in milk yield between 1996-2000 and 2001-2005.

(ii) Pie Chart

- Used to compare parts of a whole, which do not show changes over time;
- Example, contribution of various sources to farmer's total annual income (agriculture – 56%; service – 20%; others – 24%).

(iii) Bar Graph

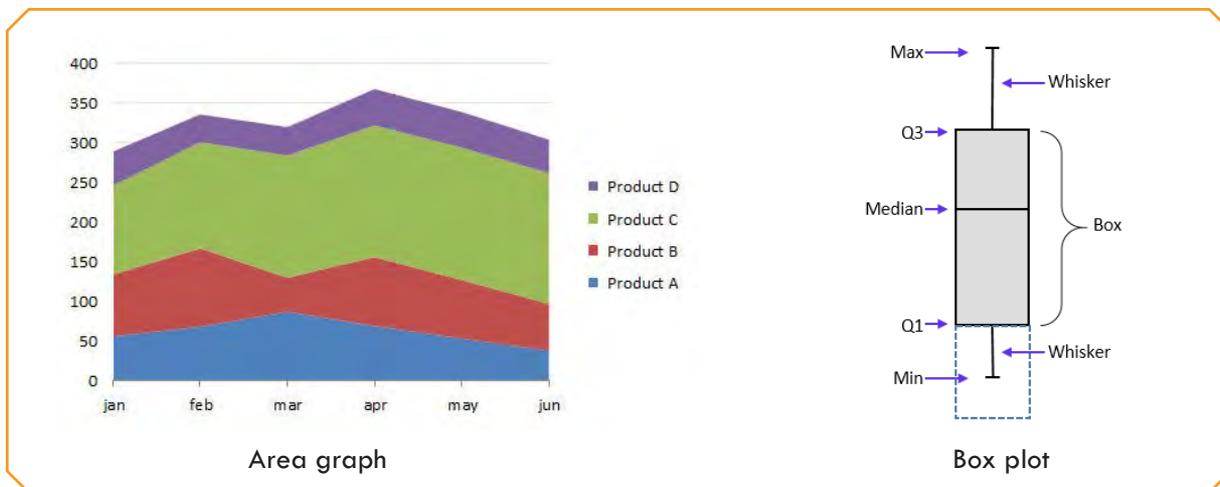
- To compare things between different groups or to track changes over time;
- When trying to measure change over time, bar graphs can be suitable only if the changes are larger;
- Example, year-wise egg production in a village.

(iv) Area Graph

- Used to track changes in two or more related groups over time for one or more groups.

(v) Box Plot/ Box-Whisker Plot

- Useful in presenting several descriptive statistics of quantitative variables;
- Provides five values – Min, Max, Median, Quartile 1 and Quartile 3;
- Used in detecting outliers.



Choosing statistical tests based on research questions

A few examples of research questions and appropriate statistical test (test of significance) are displayed in Table 14.2, while table 14.3 describes choice of techniques based on research or extension purpose.

Table 14.2: Selecting statistical tests based on research questions and hypotheses

Population & No. of variables	Type of Dependent Variable
Male farmers' gain in knowledge from a training programme is higher than in women farmers	Paired t test
Adoption of artificial insemination is higher among Haryana dairy farmers than in Kerala farmers	Independent t test
Farmers' choice to enrol in a specific distance education programme depends on his/her farming system	Chi-square test
Performance of agricultural start-ups depend on enterprising tendency of the entrepreneur	One way ANOVA
Farmers' innovativeness and their risk-taking ability are interrelated	Pearson correlation
Success of a self-help group depends on the cohesion among group members	Regression

Table 14.3: Selecting statistical tests based on purpose of research and extension

Research purpose	Statistical methods
Identifying latent structure in the data, e.g., Summated scale development	Exploratory factor analysis Confirmatory factor analysis
Assessing the consequences of technology adoption in a specific system through longitudinal approach	Structural Equation Modeling
Segmenting farmers into homogenous categories based on their personal and social characters	Cluster analysis

Research purpose	Statistical methods
Identifying 'recommendation domains' for a specific technology or package based on farmer and production system characters	Cluster analysis
Identifying the 'unknown' crop varieties located during the varietal adoption survey by matching varietal attributes to check if it is a 'released variety'	Discriminant analysis
Assessing the agreement among the ratings judges who are evaluating exhibition stalls for awarding 'Best stall' (Inter-rater agreement/reliability)	Intra-class correlation coefficient
Assessment of internal consistency reliability of a scale	Cronbach alpha, item-item correlation; item-total correlation
Identifying the socio-economic and other variables influencing adoption of crop or animal technologies	Logistic regression – binary, ordered and multinomial Probit and tobit regressions
Identifying differences in the ranking pattern of male and female farmers while assessing qualities of improved varieties in the Front Line Demonstration (FLD) or On Farm Testing (OFT)	Spearman rank correlation and Friedman ANOVA

For details on how to use Statistical Package for the Social Science (SPSS) for basic statistical analysis refer annexure A-3.

14.3 Key Points

- Statistical analyses help to describe and summarise data as well as derive inferences about relationships, associations and causality among variables in a phenomenon.
- Both descriptive and inferential statistics are used to get insights into social phenomena.
- Choosing the right statistical analysis depends on research purpose, research questions and hypotheses.

References

- IDRE-UCLA. (2017.) Introduction to SAS. UCLA: Statistical Consulting Group. (Retrieved from: <http://www.ats.ucla.edu/stat/sas/whatstat/>).

15

A Few Guidelines for Data Interpretation and Inference

I S Rao and P Sethuraman Sivakumar

Objectives

1. Provide an overview of interpretation and inference making from the data
2. Explain various approaches for effectively interpreting the data and drawing inferences and conclusions

15.1 Introduction

- Analysis, inference and interpretation are important components of extension research.
- They help the extension researcher to categorise, summarise, and interpret the data with respect to the stated objectives.
- This module provides various approaches for effective interpretation of data and for drawing meaningful inferences about the research questions.

15.2 Discussion

(a) Analysis

- A systematic examination and evaluation of data or information, by breaking it into its component parts in order to uncover their interrelationships (businessdictionary.com).
- Methods of analysis of data are described elsewhere in this manual.

(b) Interpretation

- An Interpretation is an Inference derived from a specific point of view (Meegan, 2015). In extension research, the inference is made from statistical analysis.
- Statistical interpretation is the analysis and explanation of: (i) what is typical and what deviates from the average or norm; (ii) how widely the data varies; and (iii) how it is distributed in relation to the variable(s) being measured (businessdictionary.com).
- The goal for interpretation is to provide further clarification, elucidation, meaning, and understanding, while the goal for inference is to arrive at a logical conclusion and to provide explanation

Wisdom hierarchy: Data–Information–Knowledge–Wisdom Hierarchy (DIKW)

- According to Russell Ackoff (Ackoff, 1989), the content of the human mind can be classified into five categories: data, information, knowledge, understanding and wisdom.

1. **Data:** Raw information presented as symbols (text, numbers and graphics). For example, 52.5 and 5 & 4.
 2. **Information:** The data are processed into a useful form which provides answers to 'who', 'what', 'where', and 'when' questions. For example, Weight – 52.5 kg; Height – 5 ft 4 inches.
 3. **Knowledge:** Generated by applying the data and information to answer 'how' questions. For example, average weight of adults belonging to the 18-25 age group measured through a scale is 52.5 kg.
 4. **Understanding:** The knowledge acquired is processed through understanding the reasons for a particular phenomenon – appreciation through 'why' question. For example, the average weight of adults belonging to the 18-25 age group is 52.5 kg with a height of 5 ft 4 inches, which indicate that they are in a normal weight category (BMI = 19.9).
 5. **Wisdom:** is evaluation of understanding. Developed through application of knowledge and understanding in real life and verifying the outcomes. For example, food consumption and lifestyle factors indicate that adults belonging to the 18-25 age group with a BMI of 19.9 are following a healthy lifestyle which help them to stay within the normal weight category.
- A pyramid structure (Fig. 15.1) explains the structural and functional relationships between data, information, knowledge and wisdom. According to this pyramid, the learning process starts at the base from Data and progresses through the pyramid to the top level, Wisdom.

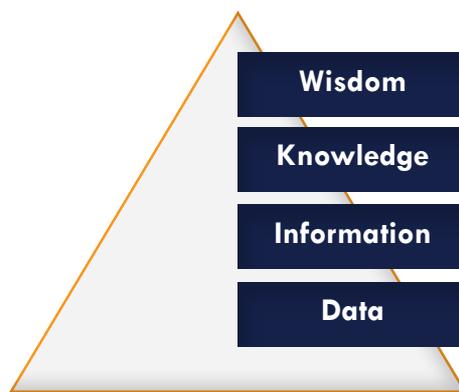


Fig. 15.1: Wisdom pyramid

- The process begins by gathering Data in the first level, which is then processed to form Information at the second level. When this Information is examined or considered it takes the form of Knowledge at the third level, and the creation of Knowledge leads to acquiring Wisdom at the fourth or topmost level (www.click4it.org).

Importance of deriving realistic interpretation

1. Enables the researcher to acquire in-depth knowledge about the concept behind his own findings.
2. The researcher is able to understand his findings and the reasons behind their existence.
3. More understanding and knowledge can be obtained with the help of further research.
4. Provides very good guidance to the studies relating to the research work.

Some Data Interpretation and Analysis Tips (Clip, 2007)

1. Consider the data from various perspectives – Use multiple perspectives of the phenomenon under study;

2. Think beyond the data but do not stray too far from the data – Bring interpretations from outside data using triangulation techniques;
3. Make the link between the data and your interpretations clear – Since extension research uses statistical interpretation, the researcher has to link the data, results and interpretations in a logical way to derive conclusions;
4. Make visible the assumptions and beliefs that influence your interpretation – The assumptions and beliefs of the researcher underlying the results and way of interpretation should be clearly specified;
5. Reflect on your own thinking and reasoning – Using scientific and commonsense approaches for deriving conclusions from the data;
6. Take care not to disregard outlying data or data that seems to be the exception;
7. Data that is surprising, contradictory or puzzling can lead to useful insights.

How to write interpretation from the data

When a table or graph is prepared using the analysis, please follow these steps to interpret data:

Step 1

- Start by explaining the simple and straightforward points about each table or graph.
- Start with a statement about ‘what does it show?’
- Add the numbers from the table or graph to support the statement.
- Examples:

The data displayed in Table 1 indicates that majority of the respondents (75%) preferred IR8 variety of rice followed by Figure 1 indicates that farmers adopted IR8 rice variety on a larger scale (85% of total rice area) than local landraces (15%).

Step 2

- Provide more details and explain.
- For each table or graph, ask and answer the question, ‘Why is it like this?’
- Example:

Majority of the respondents (75%) preferred IR8 variety of rice because they liked its taste (65%) and indicated that it can also tolerate pests and diseases (47%).

Step 3

- Identify similarities, differences, patterns, links and relationships between the groups.
- Identify the magnitude and nature of differences or similarities, patterns and explain why these differences exist.
- Example:

Majority of the adapters of IR8 rice variety are graduates (45%), medium farmers (67%), and possess a paddy harvester (34%). The multiple regression analysis indicated that extension contact significantly influenced adoption of IR8 rice variety ($r=0.59^{**}$).

(c) Inference

- An assumption or conclusion that is rationally and logically made, based on the given facts or circumstances through inductive reasoning (businessdictionary.com).
- The act of passing from statistical sample data to generalizations.

- Depends on extension (i.e., the number of cases under comparison) and intension (i.e., the depth of interpretation that is possible).

Interpretation vs. Inference

- The goal of interpretation is to provide further clarification, elucidation, meaning, and understanding, while the goal for inference is to arrive at a logical conclusion and to provide an explanation.
- Interpretation values the depth and completeness of phenomena and Inference values its strength.

The Ladder of Inference (Argyris, 1982)

- The Ladder of Inference describes our unconscious decision process employed while moving ahead from examining facts to arriving at a decision or action (Table 15.1).

Table 15.1: Levels of the ladder along with associated actions

Level in the ladder	Action taken at each level of the ladder
7. Actions	Derive actions based on the beliefs
6. Beliefs	Develop general beliefs about the situation
5. Conclusions	Draw conclusions about the situation
4. Assumptions	Make assumptions based on researcher's interpretation of the world
3. Interpreted reality (Meanings)	Add meaning to the data generated by observing and selecting reality
2. Select reality	Select data from the reality through observation
1. Observe reality	Observe the reality and sense the data and experiences

- The Ladder of Inference helps to draw better conclusions, or to challenge other people's conclusions based on true facts and reality.

Steps in using the Ladder of Inference (Adapted from Mindtools, 2017)

1. When conventional interpretations are drawn from the data, try to implement the Ladder of Inference.
2. Determine the level in the ladder that the researcher is currently positioned at. It is decided based on the conventional interpretation already made by the researcher. Use Table 15.1 to decide on the researcher's current level of interpretation.
3. When the current level in the ladder is identified, analyze the researcher's reasoning by working back down the ladder. This process will help the researcher to trace the facts and reality that he/she is actually working with.
4. When working down the ladder, the researcher asks 'WHAT is he/she thinking?' and 'WHY'. Based on the answers to these questions, the researcher adjusts the way and mode of reasoning. When the researcher's conventional interpretation is leading to a specific action, he/she should work backwards from the top. In this process, the researcher asks the following questions:

- Why did I choose this action? What are other alternate actions available to me?
 - Name the beliefs that led to the action? Are they well defined?
 - Reason for drawing a specific conclusion – Is it sound enough to answer the research question?
 - What are my assumptions at this level? Are they correct?
 - What data have I chosen to use and why? Have I selected data rigorously?
 - What are the facts in the reality selected by me? Are they relevant? Are any facts left out?
5. Repeat the steps until clarity is achieved.
 6. With the help of the interpretation, the researcher is able to reach a conclusion from the set of gathered data.

15.3 Key Points

- Analysis, inference and interpretation are the important components of extension research.
- Data interpretation and writing follows an inductive approach.
- Interpreting the data iteratively using techniques, such as laddering, can help in deriving accurate inferences.

References

- Ackoff, R.L. (1989.) From data to wisdom. *Journal of Applied Systems Analysis* 16 (1989):3–9.
- Argyris, C, (1982.) Reasoning, learning, and action. San Francisco: Jossey Bass.
- Clip. (2007.) Module 5: Tips for interpreting data. *Communities of learning, inquiry and practice*. (Retrieved from: http://insites.org/insites_archive/CLIP_v1_site/downloads/PDFs/TipsIntrprtData.5G.8-07.pdf).
- Meegan, G. (2015.) The elements of thought – one by one. (Retrieved from: <https://theelementsofthought.org/the-elements-of-thought-one-by-one/interpretation-and-inference/>).
- Mindtools.com (2017.) The Ladder of Inference: Avoiding “Jumping to Conclusions”. (Retrieved from: https://www.mindtools.com/pages/article/newTMC_91.htm).



Photo Credit: Subash S P

16

Focus Group Discussion

A Dhandapani and
P Sethuraman Sivakumar

Objectives

1. Introduce focus group discussion (FGD) as a tool for data collection
2. Explain steps in conducting, coding and analysis of FGD data

16.1 Introduction

- Focus Group Discussion (FGD) is a qualitative method for collecting data in extension;
- It is a qualitative data extraction and perspective documentation tool widely used for exploring explaining, and enumerating social science research. This module presents a field oriented approach for designing and using FGD.

16.2 Discussion

A. Focus group discussion (FGD)

- A research technique that initiates interaction among participants for the purpose of understanding their perspectives on a topic determined by the researcher.
- Emphasis is on the interaction between participants in the group (Morgan, 1997) for generating data and insights on a specific topic.
- Interaction creates a 'synergistic effect' (Stewart et al., 2007) because it allows the participants to respond and build on the reactions of other members in the group.
- Group interaction is able to 'reveal points of agreement, conflict, and uncertainty' (Pfeffer, 2008).
- Researchers can gain understanding by listening to the participants' discussions, challenges and contradictions as it takes place in the group.

B. Kinds of interactions

There are two kinds of interaction in focus groups (Bryman, 2008).

(i) Complementary

- Sharing experiences, concerns and needs;
- Providing a 'framework of understanding' about a topic by social groups.

(ii) Argumentative

- Questioning, disagreeing with, and challenging each other;

- Reveals people's underlying beliefs about a phenomenon;
- Disagreement can offer an opportunity for participants to review and modify their views.

C. Group composition

(a) Homogeneous group Vs. Heterogeneous group

- Homogeneity – Socio-cultural background and personal characteristics of the participants in the topic of discussion, and not their views and attitudes (Morgan, 1997; Peek and Fothergill, 2009).
- The selection of homogeneous and heterogeneous groups depends on the nature of research work and researchers' objectives.
- Similar level of experience in the topic of discussion – Initiate open and lively discussion, and help in gathering a collective view (Ivanoff and Hultberg, 2006).
- Homogeneous
 - Create comfort within the group which enhances a more fluid discussion among the participants;
 - This would make participants feel more open about discussing sensitive information.
- Heterogeneous
 - Useful when the researcher wishes to maximize the possibility of exploring subjects from different perspectives;
 - Assessing community attitudes and beliefs.

(b) Pre-existing group Vs Constructed group

- The pre-existing group has members who are familiar with each other, while the constructed group involves strangers brought together for the focus groups.

Pre-existing groups

- Naturally-occurring group – participants can connect with each other and characterised by natural, free flowing interaction (Kitzinger, 2005);
- Presence of group dynamics will initiate natural interactions (Leask et al., 2001);
- Deeper the levels of disclosure of opinion on the specific topic;
- Useful in focus group research involving specific groups (e.g., tribals) or focusing on examining sensitive issues (e.g., displacement of tribal communities);
- Disadvantage - established hierarchies and patterns of interaction between group members is detrimental for open interaction (Krueger and Casey, 2009).

Constructed groups (Leask et al., 2001)

- They are free from any group conformity or obedience to their leadership which is present in familiar groups;
- Advantage - Open and honest discussion without social desirability and conformity biases;
- Disadvantage - group members are reluctant to talk to strangers.

(c) Sample size

(i) Number of participants in a group

- Determined by the preferences of the research design (nature and size of the social groups required for the design), and the practical constraints of doing research (accessibility to

- representative sample and their willingness to participate) (Conradson, 2005);
- Ideal group size – Six to ten members (Smithson, 2008)
 - Groups with less than four members – less information, domination by a few people;
 - Groups with more than eight members – longer waiting time to get an opportunity to talk;
- Mini-focus groups (Krueger and Casey, 2009)
 - Four to six participants;
 - Greater opportunity and more time for each participant to talk, resulting in more relevant and interesting data;
- Over-recruit participants to manage dropouts.

(ii) Number of groups

- The number of focus group interviews is decided based on the saturation point – an event in the data collection process where generating additional information from participants no longer generates new understanding (Liamputtong, 2009);
- Occurs after 3 to 5 FGD for each topic.

D. How long should a focus group take?

- As a rule of thumb, a focus group should not last longer than two hours. Most focus groups are conducted within one-and-a-half hours (Morgan, 1997).
- There are a number of reasons why a session should not take too long:
 - Long discussions are tiring for participants. If the session continues for too long, people's attention will wane, which will change the dynamics of the group;
 - Participants may run out of ideas and find the discussion boring;
 - The participants may have other engagements.
- A focus group may be conducted in two separate sessions of about 20 minutes each, with a break for refreshments, so as to keep participants active throughout the sessions.

Criteria used to determine success of focus groups (Merton et al., 1990)

The success of focus groups depends on how consistently they are used with the purposes and aims of the research (Stewart et al., 2007). Four criteria for assessing success of focus groups are displayed in Box 16.1.

Box 16.1: Criteria for assessing success of focus groups

Specificity – The extent to which an FGD can extract specific information based on participants' experiences and perspectives.

Range – The extent to which the FGD covers a wide range of relevant issues pertaining to the research question, as well as unexpected or unanticipated issues.

Depth – The extent of depth with which the FGD has captured the participants' perspectives.

Personal context – The extent to which the focus groups can reason out the personal context in which the responses are received.

E. People

Sampling

- Purposive sampling – Selecting ‘information-rich respondents’ for generating relevant data;
- Snowball sampling – For selecting vulnerable or marginalised participants who often are difficult to reach (Bernard 2006; Liamputtong 2007);
- Adequate representation from gender, age, power and clique categories.

(i) Moderator/facilitator

- Initiate and facilitate discussion;
- Goal is to generate the maximum number of different ideas and opinions from as many different people in the time allotted;
- A good moderator needs to:
 - be sensitive to the needs of the participants;
 - be non-judgmental about the responses from the participants;
 - respect the participants;
 - be open-minded;
 - have adequate knowledge about the project;
 - have good listening skills;
 - have good leadership skills;
 - have good observation skills;
 - have patience and flexibility.

(ii) Note-taker

- The note-taker records the key issues emerging in the session and other factors that may be important in the analysis and interpretation of the results;
- Record the participants’ responses, ideas and other proceedings in field notes as well as via audio/ video recording.

F. Use of stimulus materials and activities in focus groups

- These stimulus materials and activities, like video film, movie, cartoons, prompts, games, exercises, and vignettes are used to ‘warmup’ participants at the beginning (Bailey, 2008);
- They also allow the participants to feel more relaxed and comfortable in the group.

G. Focus group location, venue and timing

- Neutral venue which is private, quiet, comfortable, spacious and free from distractions (Halcomb et al., 2007);
- Examples – Community halls, local village office, temples/churches/mosques, farmer’s home, etc.
- A frequently used and accessible venue may be selected as it provides participants with a familiar and comfortable discussion environment.

Method of data collection

In each village, two FGDs should be conducted. In an FGD, a questioning route is used to ask questions sequentially to the group to elicit responses. About 10-12 questions are asked in a systematic and logical way; however, eight questions are ideal. Types of questions asked in FGD are as follows:

1. Engagement questions: introduce participants to each other and make them comfortable with the topic of discussion;
2. Exploration questions: Topic of the discussion;
3. Exit question: Check to see if anything was missed in the discussion.

Guidelines for designing questions

- Short and to the point
- Focused on one dimension each
- Unambiguously worded
- Open-ended or sentence completion types
- Non-threatening or embarrassing
- Worded in a way that they cannot be answered with a simple 'yes' or 'no' answer (use 'why' and 'how; questions instead).

Steps in conducting FGD

Before FGD

1. Select a topic of discussion
2. Prepare the questioning route
3. Select the village, location, nature, and number of participants
4. Visit the village and finalize the participants, place, time and number of FGD
5. Rehearse the questioning route and arrange materials.

During FGD

1. Welcome the participants
2. Explain the purpose and request participants to trust the organizers and participate in an open-minded manner
3. Begin with a simple talk and slowly proceed to the important questions
4. Keep the participants engaged and never allow them to divert from the topic
5. Record the discussions both in note books and using an audio recorder
6. Summarize after completion of discussion
7. Thank participants and appreciate their cooperation.

Analysis of FGD data

At this stage, the summaries of focus groups are prepared to get an overall picture of the aspect under study. Several steps in analyses are given below:

Method 1. Thematic analysis

The data analytical strategies used in thematic analysis can be used for analyzing Focus Group Data.

Method 2: MS EXCEL

1. Consolidate focus group responses

Prepare a master table showing the summaries of the focus groups. Read all the tables and re-code all the answers and give the same label to similar answers. For instance, if one focus group reported 'food

vendor' and another reported 'vegetable vendor', re-code; if judged similar in the local context, these answers can use the same label (e.g. 'food vendor' for both).

2. Arrange the rows for each table separately so that all the responses of one focus group are in the same row

Do this for all the tables (focus groups). At the end, all the tables should be transformed into rows and the final qualitative data table should have as many rows as the number of focus groups.

- a. Pull together all the lines;
- b. Examine the results

- The sources of livelihood in order of importance (use 'Source' and 'Contribution to total income' columns);
- For the most important sources, report the more recurrent inputs and constraints (use 'Inputs' and 'Constraints' columns);
- Identify differences across the country (use 'Village' and 'Area' columns).

A detailed procedure for analyzing FGD data is explained in Eliot and Associates (2005).

16.3 Key Points

- Focus group discussion is an effective tool for collecting data along with insights on a specific topic.
- Focus group discussions are conducted in 3 to 5 numbers per topic with 6-8 participants in each group.
- The researcher, moderator and note-taker are primary people involved in the FGD.
- FGD data can be analysed through thematic analysis and MS EXCEL methods.
- FGD is an essential tool for qualitative extension data collection.

References

- Bailey, A. (2008.) Let's tell you a story: Use of vignettes in focus group discussions on HIV/AIDS among migrants. In P.Liamputtong (ed.), *Doing cross-cultural research: Ethical and methodological perspectives*, (pp. 253–264). Dordrecht: Springer.
- Bernard, H.R. (2006.) *Research methods in anthropology: Qualitative and quantitative approaches*. Fourth edition.
- Bryman, A. (2008.) *Social research methods*. Third edition. New York: Oxford University Press.
- Conradson, D. (2005.) Focus groups. In R. Flowerdew & D. Martin (eds.), *Methods in human geography: A guide for students doing a research project*, (pp. 128–143). Harlow: Pearson Prentice Hall.
- Halcomb, E.J., Gholizadeh, L., DiGiacomo, M., Phillips, J., and Davidson, P.M. (2007.) Literature review: Considerations in undertaking focus group research with culturally and linguistically diverse groups. *Journal of Clinical Nursing*, 16(6):1000–1011.
- Eliot and Associates (2005). *Guidelines for conducting a focus group*. Eliot and Associated. Retrieved from : https://assessment.aas.duke.edu/documents/How_to_Conduct_a_Focus_Group.pdf.
- Ivanoff, S.D., and Hultberg, J. (2006.) Understanding the multiple realities of everyday life: Basic assumptions in focus-group methodology. *Scandinavian Journal of Occupational Therapy* 13(2):125–132.
- Kitzinger, J. (2005.) Focus group research: Using group dynamics to explore perceptions,

- experiences and understandings. In I. Holloway (ed.), *Qualitative research in health care* (pp. 56–70). Maidenhead: Open University Press.
- Krueger, R.A., and Casey, M.A. (2009.) *Focus groups: A practical guide for applied research*. Fourth edition. Thousand Oaks, CA: Sage.
 - Leask, J., Hawe, P., and Chapman, S. (2001.) Focus group composition: A comparison between natural and constructed groups. *Australian and New Zealand Journal of Public Health* 25(2):152–154.
 - Liamputtong, P. (2007.) *Researching the vulnerable: A guide to sensitive research methods*. London: Sage.
 - Liamputtong, P. (2009.) *Qualitative research methods*. Third edition. Melbourne: Oxford University Press.
 - Merton, R.K., Fiske, M., and Kendall, P.L. (1990.) *The focused interview*. Second edition. New York: Free Press.
 - Morgan, D.L. (1988.) *Focus group as qualitative research*. Newbury Park, CA: Sage.
 - Morgan, D.L. (1997.) *Focus groups as qualitative research*. Second edition. Newbury Park, CA: Sage.
 - Peek, L., and Fothergill, A. (2009.) Using focus groups: Lessons from studying daycare centers, 9/11, and Hurricane Katrina. *Qualitative Research* 9(1):31–59.
 - Pfeffer, N. (2008.) What British women say matters to them about donating an aborted fetus to stem cell research: A focus group study. *Social Science and Medicine* 66(12):2544–2554.
 - Smithson, J. (2008.) Focus groups. In P. Alasuutari, L. Bickman & J. Brannen (eds.), *The Sage handbook of social research methods* (pp. 357–430). London: Sage.
 - Stewart, D.W., Shamdasani, P.N., and Rook, D.W. (2007.) *Focus groups: Theory and practice*. Second edition. Thousand Oaks, CA: Sage.



Photo Credit: Subash S P

17

Case Study Research

Subash S P

Objectives

1. Introduce the case study research approach
2. Explain the method and techniques for doing case study research
3. Suggest how to carry out case study research with a rigorous methodology

17.1 Introduction

- Case study research is an ideal methodology for holistic and in-depth analysis of a social phenomenon.
- There are several scientific benefits in doing case studies as they prove to be a favourable breeding ground for insights and farming hypothesis.
- Generally they are often contested as a procedure involving too many subjective decisions, but if they are framed properly with methodological rigour, objective and generalizable results can be obtained.
- In this chapter we briefly discuss the case study approach, types of case studies and methodology for conducting a case study.

17.2 Discussion

17.2.1 What is case study research?

A case study is an empirical inquiry that investigates a contemporary phenomenon in depth and within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident. It is not a data-gathering technique, rather a methodological approach of collection of data (Berg, 2004). Case studies are preferred research methods in the following situations (Dasgupta, 2015):

1. When a phenomenon is broad and complex (e.g., need for, and relevance of, genetically modified crops/animals);
2. Where the existing body of knowledge is insufficient to permit the posing of casual questions (e.g., climate adaptation strategies followed by farmers);
3. When a holistic, in-depth investigation is needed (e.g., farmer suicides in India);
4. When a phenomenon cannot be studied outside the context in which it occurs (e.g., indigenous knowledge of tribals in North Eastern States of India).

17.2.2 Typology of case study research

Typology provides a framework for choosing various forms of case study designs based on the subject and object (analytical framework).

- (i) The subject of the study – the specific case;
- (ii) The object – the analytical frame or theory through which the subject is viewed and which the subject explicates. The object includes purpose, approach and process:
 - Purpose – explaining/evaluating or exploring;
 - Approach – describing or interpreting a phenomenon, building or testing theory;
 - Process – single, multiple or parallel.

The decision tree for choosing the case is displayed in Fig. 17.1 below (Thomas, 2010).

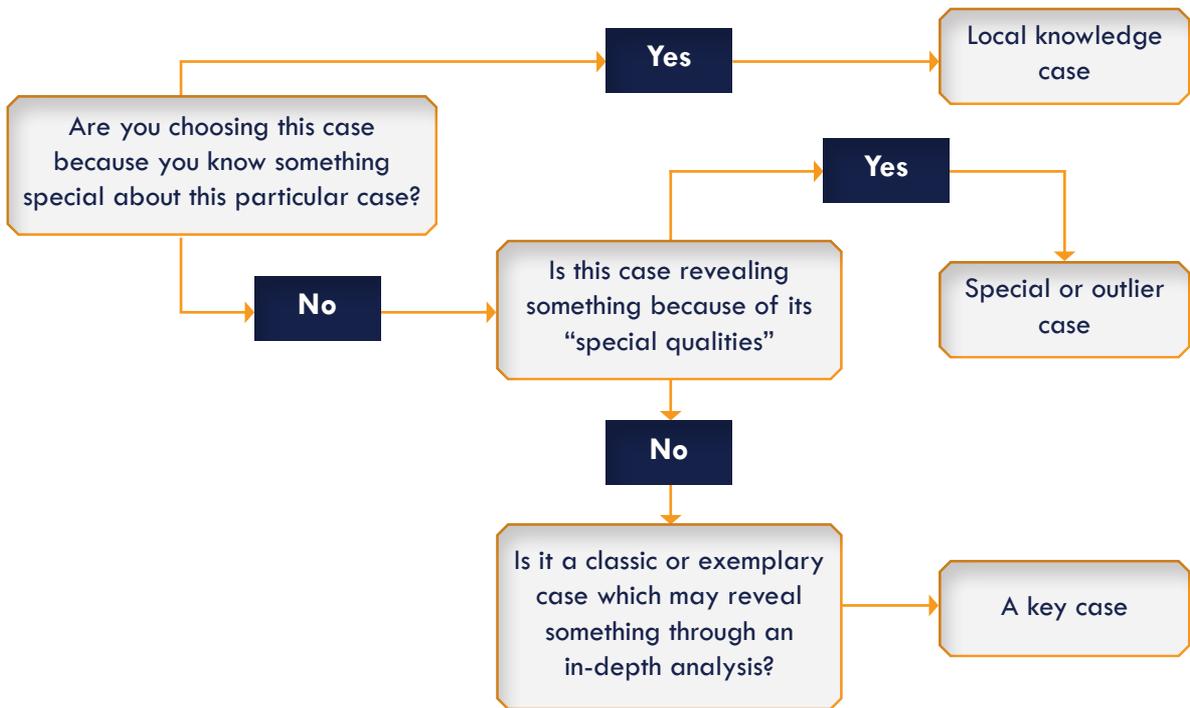


Fig. 17.1: Decision tree for choosing a case

(i) The subject

The subject may be of three types:

(a) Local knowledge case

- An investigator is aware of the existence of a case in his/her vicinity.
- The researcher has limited time and resources and he/she can confine the research only to cases in the institution's proximity
- For example, a local progressive farmer.

(b) Key case

- A classic or exemplary case of a specific phenomenon.
- For example, a successful agripreneur living in a distant place.

(c) Outlier case

- A case which is different from others in its category or class.

- For example, a small dairy farmer who adopted Good Practices in Dairying because of his/her interest.

(ii) Object

(a) Purpose

1. Intrinsic

- To understand some particular trait or characteristic of a specific case (Stake, 1995).
- Purpose – just for understanding a case, with no specific purpose, like theory building or testing, etc.
- Example: a KVK Subject Matter Specialist wishes to develop a case study out of a Technology Assessment Trial conducted in the district.

2. Instrumental

- To probe an issue in order to get insights or draw generalisations.
- The case is of secondary interest but it facilitates an understanding of the phenomenon under study.
- Example: develop a case study on the 'enchantment discontinuance' phenomenon in improved varieties of turmeric.

3. Evaluative

- To check how some intervention or aspect is working in a specific context.
- Example: case study on an On Farm Trial of an improved variety.

4. Explanatory

- Explaining a phenomenon by indicating its elements and relationships.
- Causal studies done with Pattern matching technique (Yin and Moore, 1988).
- Example: studying the group processes which determine the sustainability of Self-Help groups.

5. Exploratory

- To explore a phenomenon of interest.
- To answer questions – what, why, and how.
- Example: case studies about farmers' suicides in Maharashtra, India.

(b) Approach

1. Theory testing or building

- Theory building – develops ideas through exploration, identifies elements and relations and proposes a theory.
- Theory testing – proving the truthfulness in a theory for applying it for a specific purpose.
- Involves decisions to choose appropriate designs, methods, techniques, tools and processes.

2. Illustrative/demonstrative

- Utilizing one or two instances of an event to describe a situation.
- Illustrates and provides metaphors by which the learner can get 'inside' a problem, think about it

and connect it with characters.

- Enable learners to share their experiences and relate with the case study.
- Example: a few illustrative case studies on alternative agriculture can be found in <https://www.nap.edu/read/12832/chapter/10>

3. Interpretative

- In-depth understanding of a particular phenomenon.
- Ethnographic approach.

4. Experimental

- To test the influence of a specific intervention on human behaviour or plant or animal yield under controlled conditions.
- Example –on-farm experiments , pre- and post-knowledge tests.

(c) Process

(i) Single or multiple (Thomas and Myers, 2015)

- Single –classic case type with focus on a single case.
- Multiple cases – a cross-case analysis where key focus is on the nature of the difference between one case and another.

(ii) The boundary and the shape (Thomas and Myers, 2015).

1. Single

Single studies, containing no element of comparison will take essentially three forms, wherein personal or systemic features of the subject are bound by time in some shape or form.

- Retrospective – collection of data relating to a past phenomenon, situation, person, or event, or studying it in its historical integrity.
- Snapshot – the case is being examined in one defined period of time: a current incident; event; a day in the life of a person.
- Diachronic – shows change over time at specific time intervals.

2. Multiple

Multiple cases involve comparison between cases or between elements of a single case –

- Nested – Analysis between elements of one single case;
- Parallel – Comparing two cases happening at the same time;
- Sequential – Cases occurring in a sequence.
- Example: can be a Self-help group, Farmer Producer Groups, Organisational studies.

Steps in designing a case study

- Choosing the right case study research design is a tricky process.
- Thomas (2010) has outlined a framework for choosing a right case study method based on the subject, purpose, approach and process (Fig 17.2).

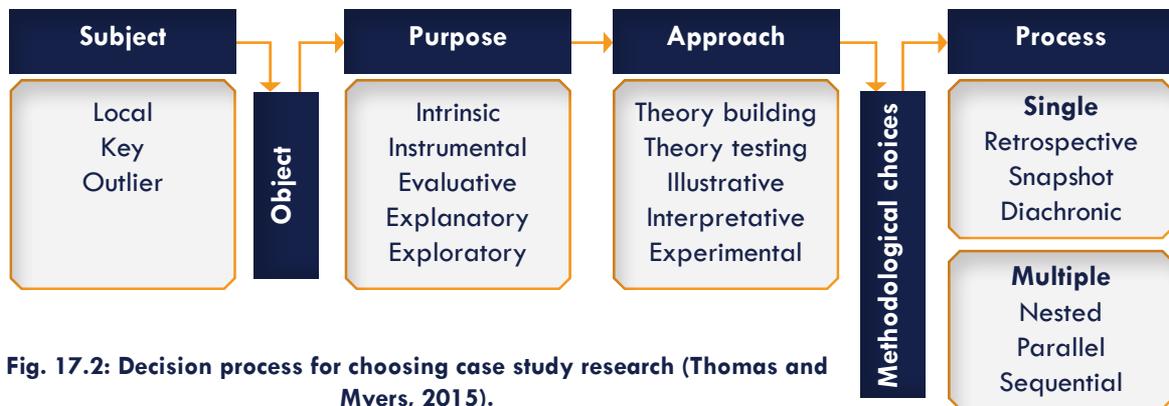


Fig. 17.2: Decision process for choosing case study research (Thomas and Myers, 2015).

A researcher can choose the right approach by using the following steps:

1. Decide on the subject/unit of study – local knowledge case, key case or outlier case;
2. When the subject is decided, the researcher has to choose the object or analytical framework for the case based on:
 - What is the purpose of the case study? Intrinsic/instrumental/evaluative/explanatory or exploratory;
 - What is the approach required for conducting the case study? Theory building/testing/illustrative or demonstrative/interpretative/experimental;
 - What is the process required for the work?
 - Single case – retrospective/snapshot/diachronic;
 - Multiple – nested/parallel/sequential.

Example

The research problem – developing a case study on farmers’ suicides in Maharashtra (Fig. 17.3).

Steps

1. Unit of the study – key cases of farmers/families who committed suicide;
2. Analytical framework/object
 - Purpose – Instrumental case study (because the phenomenon of suicide is more important than the farmers);
 - Approach – Interpretative (in-depth analysis of the causes and other aspects of suicides);
 - Process – Multiple cases with nested approach for several cases

Since suicide is an outcome of several inter-related elements of the farmers life, a nested approach is ideal.

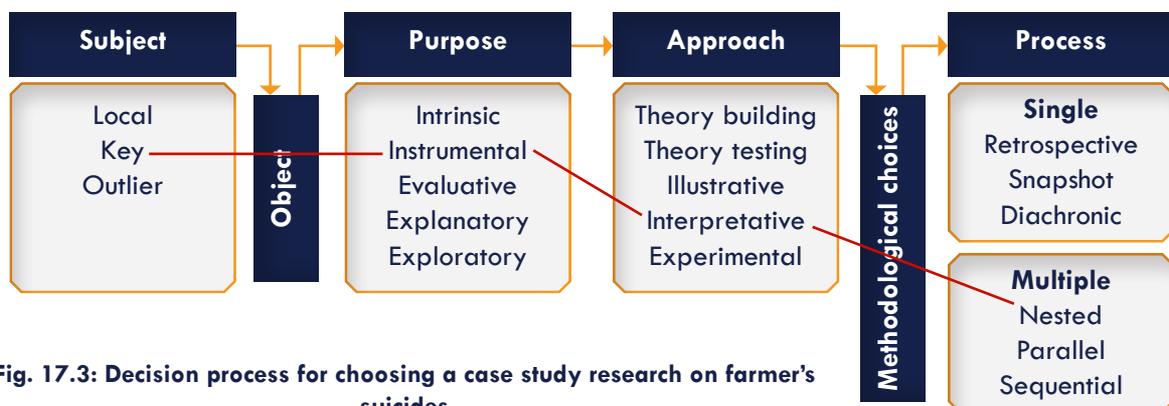


Fig. 17.3: Decision process for choosing a case study research on farmer’s suicides

Box 17.1: Points to be considered while doing an individual case study

- How broad an area of social life will be covered;
- Research question and nature of the research problem should be defined prior to the study;
- You may opt for a single lengthy interview or several interviews and it could be supplemented with field notes, direct observation, journals, etc.;
- Subject's relationship with others is to be recorded, as there exists an interconnectedness between the individual and his/her social life. Therefore, interviews of others related to the subject must also be carried out.

Source: Berg, (2004).

17.2.3 Data collection for case study

Data for use in the case study approach could include life histories, documents, oral histories, in-depth interviews, participant observation, and other data gathered through various techniques. Interview (see Box 17.2), documents, ethnography, survey data and observations are common methods of collecting data. Other methods, such as personal documents, autobiographies, etc., can also be used for case study research. Elizabeth Roberson's (1998) exploration of the letters of Eli Landers, a young Confederate soldier (in Watt and Lott, 2002) is a good case in point.

Box 17.2: Skills required for a researcher for conducting an interview for case study research

Yin (1998) had propose five basic research skills while doing a case study:

- Inquiring mind;
- Listening, observation and sensing;
- Adaptability and flexibility;
- Thorough understanding of the issues;
- Unbiased interpretation of data.

Read Yin (1998) for further details.

Analysis in case study research

The data collected using the case study approach could be analysed by quantitative and qualitative techniques. The selection of the technique depends on the paradigm of analysis used for the study. The basic four paradigms of scientific enquiry and the research strategy is given in Table 17.1.

Table 17.1: Approaches to scientific knowledge

No.	Paradigm	Research Strategy
1	Positivism	Induction
2	Post-positivism/realism	Deduction
3	Critical realism	Retroduction
4	Interpretativism	Abduction

IGNOU teaching material on Research Methodology

Methodological rigour

Methodological rigour could be tested using validity (internal, construct, external) and reliability. The ways in which they are tested for positivistic and non-positivist paradigms are given below (Table 17.2).

Table 17.2: Positivist framework

Internal validity	Construct validity	External validity	Reliability
Research framework explicitly derived from literature	Data triangulation	Cross-case analysis (multiple or nested case)	Case study protocol (how was the case study conducted)
Pattern matching	Review of transcripts and drafts by peers	Rationale for case study selection (appropriate with regard to research question)	Case study database
Theory triangulation	Review of transcripts & drafts by key informants	Details of case study contexts	Organisation's/ individual's actual name
	Clear chain of evidence, explanation of data collection & data analysis		

Source: Baker (2010)

For non-positivistic framework (see Lincoln and Guba, 1990):

1. Credibility – confidence in the ‘truth’ of the findings;
2. Transferability – showing that the findings have applicability in other contexts;
3. Dependability – showing that the findings are consistent and could be repeated;
4. Confirmability – a degree of neutrality or the extent to which the findings of a study are shaped by the respondents and not on researcher bias, motivation, or interest.

17.2.4 Building theories from case study research

As discussed before, case studies could help in developing theories (see Box 17.3). Eisenhardt (1989) had detailed steps in which a theory could be built from a case study. The heading of the steps are provided below. For details please refer to the article.

1. Resultant theory is often novel, testable, empirically valid;
2. There should be a defined research question before starting the study;
3. Selecting case: Polar types;
4. Crafting instrument and protocols;
5. Entering the field: overlapping data analysis;
6. Analysing within-case data;
7. Searching for cross-case patterns;
8. Shaping hypothesis: refining, building evidence;
9. Enfolded literature;
10. Reaching closure: theoretical saturation.

Box 17.3: Example of building theories from case study research

Ostrom's (1998) seminal work on 'Coping with tragedy of commons' was developed based on case studies. Please access the article through the link below:

<http://dlc.dlib.indiana.edu/dlc/bitstream/handle/10535/5876/Coping%20with%20tragedies%20of%20the%20commons.pdf>

17.2.5 Qualitative content analysis in case study research

The qualitative data collected for the case could be analysed using content analysis. The steps in content analysis are as follows:

1. Identify research questions;
2. Determine analytic categories (sociological constructs);
3. Read through the data and establish grounded categories;
4. Determine systematic (objective) criteria of selection for sorting data chunks into the analytic and grounded categories;
5. Begin sorting the data into various categories;
6. Look for pattern;
7. Literature review, explanation drawing relationship.

Please refer to any standard book on content analysis (Berger, 2004) or YouTube tutorials for content analysis.

17.3 Key Points

- Case study could be used as a methodological approach for holistic and in-depth analysis.
- Objectivity and generalization could be drawn from a case study using appropriate methodology.
- There are different types of case studies.
- The case study should be designed keeping in mind the objective as well as the research paradigm of the study.
- The case study could be analysed using quantitative and qualitative techniques depending on the type of data collected.
- For qualitative data analysis content analysis could be used.
- Methodological rigour could be ensured by checking validity and reliability of the study.

References

- Berg, B.L. (2004) *Qualitative research methods*. United Kingdom: Pearson Education Inc.
- Dasgupta, M. (2015) Exploring the relevance of case study research. *Vision* 19(2):147–160.
- Eisenhardt, K. M. (1989) Building theories from case study research. *Academy of Management Review* 14(4):532–550.(Available at: <http://doi.org/10.5465/AMR.1989.4308385>).
- Feola, G., Lerner, A. M., Jain, M., Montefrio, M. J. F., and Nicholas, K. A. (2015) Researching farmer behaviour in climate change adaptation and sustainable agriculture: Lessons learned from five case studies. *Journal of Rural Studies* 39:74–84 .(Available at: <http://doi.org/10.1016/j.jrurstud.2015.03.009>)
- Jensen, J.L., and Rodgers, R. (2001) Cumulating the intellectual gold of case study research. *Public Administration Review* 61(2):236-246.

- Hall, A., Rasheed Sulaiman, V., Clark, N., and Yoganand, B. (2003) From measuring impact to learning institutional lessons: An innovation systems perspective on improving the management of international agricultural research. *Agricultural Systems* 78(2):213–241.(Available at: [http://doi.org/10.1016/S0308-521X\(03\)00127-6](http://doi.org/10.1016/S0308-521X(03)00127-6)).
- Lincoln, Y.S., and Guba, E.G. (1990) Judging the quality of case study reports. *Qual Stud Educ.* 3:53-59.
- Stake, R.E. (1995) *The art of case study research: Perspective in practice.* London: Sage.
- Subash, S.P. (2015) *Farm innovation and outscaling.* ICAR-National Academy of Agricultural research Management, Hyderabad.
- Thomas, G. (2010) *How to do your case study: A guide for students and researchers.* Sage.
- Thomas, G, and Myers, K (2015) *The Anatomy of the Case Study.* London, Sage.
- Watt, S., and Lott, C. (2002) Eli Landers: Letters of a confederate soldier. *Social Education* 66(2):122-130.
- Yin, R.K. (1998) The abridged version of case study research: Design and methods. In L. Bickman & D.J. Rog, *Handbook of applied social research methods.* Thousand Oaks, CA: Sage.
- Yin, R and Morre, G. (1988). The use of advanced technologies in special education. *Journal of Learning Disabilities* 20(1):60.



Photo Credit: Anthony Christopher

18

Human Subjects Research- Ethical Practices in Data Collection

P Sethuraman Sivakumar

Objectives

1. Provide an overview of human subjects research
2. Describe the human subjects research guidelines followed in India

18.1 Introduction

Human Subjects Research is conducted using human beings as subjects and involves interaction and intervention. After considering past experiences in conducting human experiments wherein human rights were violated sometimes, ethical guidelines were formulated and implemented globally by various agencies. Though extension research as such is yet to consider and implement human subject research guidelines, it is important to endorse them at the earliest so as to conduct ethically sound research. This module provides an overview of aspects to be taken into consideration when conducting research involving human subjects.

18.2 Discussion

Human subjects research

A systematic, scientific investigation designed to develop or contribute to generalizable knowledge by involving human beings as research subjects. The researcher obtains data through intervention or interaction with the individual or any identifiable information (IRB, 2017).

- Direct interaction – by collecting data through personal interviews, administering psychological tests, collecting blood/saliva samples, etc.;
- Indirect interaction – analysing secondary data obtained through other researches involving human beings as subjects;
- Intervention - includes both physical procedures for data gathering (EEG recordings, etc.) and manipulations of the subject or the subject's environment that are performed for research purposes;
- Private information – data on subject's psychological character, demographic and personally identifiable information (e.g., name, address, email, telephone number);

A few examples of human subject research conducted in extension are given in Table 18.1.

Table 18.1: A few examples of human subject research in extension

Example	Reason for considering as Human Subject Research
A personal or telephonic interview of farmers about their media utilisation behaviour	Systematic investigation with interaction between the investigator and the subjects
Collecting personal and contact details of farmers from district or local Agricultural Officers for choosing samples	These details are identifiable personal information about the subjects
Administering an attitude scale to assess the farmers' attitude towards GMO crops	Attitude is a psychological character of the subject and private information

Need for conducting ethically sound Human Subjects Research in extension

Most extension research involves interaction with respondents; and the researcher also collects identifiable personal information from participants. Though there are no strict guidelines for conducting ethically sound extension research, it is necessary to follow them for the following reasons:

- Many international journals demand evidence of conducting ethically sound research either through Institute Ethics Committee approval or statement from authors that the research was conducted using ethically sound procedures.
- When human subjects provide personally identifiable information, it is necessary to protect them from misuse by others. For example, sharing of phone numbers or email of respondents with others for any other purpose is an unethical practice and the researcher is liable for legal action if the respondent files a case of misuse or abuse of personal information.
- When research involves technology assessment and refinement, researchers use farmers' land and other resources for testing technologies. Furthermore, conducting pesticide trials on a farmer's field may sometimes lead to unexpected side effects, such as animal toxicity, after-effects on other plants, etc. In case of sensory evaluation, food poisoning due to unforeseen circumstances will make the researcher vulnerable to legal action. If consent is obtained from the participant, it will be easier to handle these issues.

History

(i) The Nuremberg Code (1947)

- During Second World War (1939-45), the German medical practitioners conducted experiments on human participants without their consent, which led to permanent impairment of the subjects and their faculties, or even death sometimes.
- The Nuremberg Code (1947) – was the first International Statement on the ethics of medical research using human subjects, which highlighted the fundamental need for obtaining consent from subjects.

(ii) Declaration of Helsinki (1964)

- The World Medical Association has formulated general principles and specific guidelines on use of human subjects in medical research.
- Institute Review Board – is in place to review Human Subject Research applications.

(iii) The Belmont Report

Based on the Tuskegee Syphilis Study (1932-1972) –

- In 1932 the US Government misled 623 African-Americans into participating in a study of untreated syphilis;
- The subjects participated in the study unknowingly and their ailment was untreated till 1972.
- There are three fundamental ethical principles for using any human subjects for research:
 1. Respect for persons – protecting the autonomy of all respondents;
 2. Beneficence – do not harm subjects;
 3. Justice – fair distribution of costs as well as benefits to potential research participants.

(iv) Indian Council of Medical Research (ICMR) Guidelines

- In 1982, ICMR issued a 'Policy Statement on Ethical Considerations involved in Research on Human Subjects' for the benefit of all those involved in clinical research in India.
- Regulatory document is available – on 'Ethical Guidelines for Biomedical Research on Human Participants' (http://www.icmr.nic.in/ethical_guidelines.pdf)

Ethical review procedure for research involving human subjects

- Mandatory that all proposals on biomedical research involving human participants should be cleared by an appropriately constituted Institutional Ethics Committee (IEC), (also referred to as Institutional Review Board (IRB), Ethics Review Board (ERB) and Research Ethics Board (REB) in other countries), to safeguard the welfare and rights of the participants.
- Regulations are practiced only by medical institutions and they are yet to take root in social sciences involving human subjects.

Institutional Ethics Committee (IEC)

Responsibilities

- Ensure a competent review of all ethical aspects of the project proposals received by it in an objective manner.
- Provide advice to researchers on all aspects of the welfare and safety of the research participants after ensuring the scientific soundness of the proposed research through the appropriate Scientific Review Committee.

Composition

- The number of persons in an Ethics Committee – 8 to 12 members.
- Composition
 1. Chairperson – Must be from outside the institution;
 2. One-two persons from basic medical science area;
 3. One-two clinicians from various institutes;
 4. One legal expert or retired judge;
 5. One social scientist/representative of a non-governmental voluntary agency;
 6. One philosopher/ethicist/theologian;
 7. One lay person from the community;
 8. Member Secretary – from the same institution.

Review procedures

1. Apply through research guide or Head of Department (HoD);
2. Submit copies of all the survey instruments including questionnaires, psychological tests or scales, prior informed consent, debriefing form, introduction letter from guide/Dean/Director requesting to participate in the work, any personal communication by email or letter between researcher and participant, cards used for data collection and PRA instruments;
3. Approval of IEC is necessary for conducting research.

Types of review procedures

1. Exemption from review

YES

Research on educational practices, such as instructional strategies or effectiveness of, or the comparison among instructional techniques, curricula, or classroom management methods.

NO

1. When research on use of educational tests, survey or interview procedures, or observation of public behaviour can identify the human participant directly or through identifiers, and the disclosure of information outside research could subject the participant to the risk of civil or criminal or financial liability or psycho-social harm;
2. When interviews involve a direct approach or access to private papers.

2. Expedited Review

- Minor deviations from originally approved research during the period of approval (usually of one year duration);
- Revised proposal previously approved through full review by the IEC or continuing review of approved proposals where there is no additional risk, or activity is limited to data analysis.

3. Full Review

- Collection of data from voice, video, digital, or image recordings made for research purposes;
- Research on individual or group characteristics or behaviour not limited to research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behaviour or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

Tools used for human subject protection

(i) Informed consent of participants

- The investigator must obtain the informed consent of the prospective participant prior to data collection;
- Informed consent protects the individual's freedom of choice and it shows respect for the individual's autonomy and is given voluntarily to participate or not in research;
- The respondent should be briefed about the research through a document called 'Informed Consent Form with Participant/Patient Information Sheet';

- Various elements of informed consent are displayed in Box 18.1;
- A sample informed consent form is displayed in Fig. 18.1.

<p style="text-align: center;">Behavioral Consent Form</p> <p>You are invited to participate in a survey conducted as a part of research project on "Enhancing food security and sustainable livelihoods in the North-Eastern India through tuber crops technologies" implemented by the Central Tuber Crops Research Institute, Trivandrum, Kerala. The purpose of the survey is to assess the how various psychological factors like eating attitudes, body image and fat teasing by peers etc influence the eating habits.</p> <p>This survey packet contains three standard psychological tests i.e. Eating Attitudes Test (EAT-26), Body Shape Questionnaire and Fat Talk scale. The survey questionnaires can be completed in 15 minutes time. The questionnaires used in this survey are universally accepted instruments for measuring the specified psychological attribute and contain questionnaires that are standardised to represent a particular concept.</p> <p>Though our survey won't pose any risks for participants, you may feel minimal levels of fatigue while responding to questionnaires. However, this should not be greater than the level of stress you might experience during routine daily experience. Since the purpose of this survey is to assess various psychological attributes like eating attitudes and body image, few questions may appear to be sensitive. However, these tests are universally accepted measures of psychological attributes, that won't be harmful to the respondents.</p>	<p>Your participation in the survey is voluntary and if you wish, you can stop at any time. If you wish to know the result of the survey, we will be happy to share them with you whenever it is completed. The records of this study will be kept private and confidential.</p> <p>Statement of Consent: I have read the above information. I have asked questions and have received answers. I consent to participate in the study.</p> <p>_____</p> <p>Signature of the participant with Date</p> <p style="text-align: right;">_____</p> <p style="text-align: right;">Signature of Investigator with Date</p>
--	---

Fig. 18.1: A sample informed consent form

(ii) Compensation for participants

Participants are compensated for their time through: (i) cash payment; (ii) gifts; (iii) a chance to participate in a lucky draw, etc.

Box 18.1: Elements of the informed consent sheet

1. Nature and purpose of study stating it as research;
2. Duration of participation and number of participants;
3. Procedures to be followed;
4. Investigations, if any, to be performed;
5. Foreseeable risks and discomforts adequately described and whether project involves more than minimal risk;
6. Benefits to participant, community or medical profession as may be applicable;
7. Policy on compensation;
8. Availability of medical treatment for any injuries or risk management;
9. Steps taken for ensuring confidentiality;
10. No loss of benefits on withdrawal during the survey;
11. Benefit sharing in the event of commercialization (e.g. ITK);
12. Contact details of Principal Investigator (PI) or local PI/Co-PI in multicentric studies for acquiring more information related to the research, or in case of injury;
13. Contact details of Chairman of the IEC for appeal against violation of rights;
14. Voluntary participation;
15. Storage period and procedures for storing research data.

18.3 Key Points

- Human Subject Research is a research work conducted using human subjects as respondents;
- Involves intervention and interaction with human participants for gathering personal identifiable information;
- Extension research involving interaction with participants are subject to human subject research regulations;
- The Institute Ethics Committee is the authority for reviewing research procedures in advance;
- Informed consent and compensation are key elements in conducting ethically sound human subjects research.

References

- Institute Review Board. (2017.) Non-human subject research. Salt Lake City, Utah: The University of Utah. (Retrieved from: <http://irb.utah.edu/guidelines/human-subject-research.php>).

19

Modeling Adoption of Agricultural Technologies

Sivaramane N

Objectives

1. Introduce Modeling adoption of agricultural technologies
2. Explain the multinomial and ordinal logit models
3. Suggest future work on adoption modelling

19.1 Introduction

- Measuring agricultural technology adoption is a multi-faceted process involving several choices and decisions to be made based on the research questions, resource availability, expertise of the researchers, etc.
- In general, adoption is assessed at the individual, institutional and societal levels. Various theories and processes of measuring adoption are discussed elsewhere (Sivakumar and Sulaiman, 2015).
- The most popular method employed in adoption study is Logit model. Logit model is a limited dependent variable model where the dependent variable is binary. This model is used in identifying the factors influencing adoption. However, this model is not applicable in cases where the dependent variable has multiple options or where it is of ordinal type.
- In this module, two econometric approaches, namely, multinomial logit and ordinal logit model employed in modeling adoption, are discussed in detail.

19.2 Discussion

- Multinomial model is an extension of the binary logistic regression model where the dependent variable is nominal and has multiple options.
- A farmer has the option of selecting a particular variety of crop for this season, or he/she takes the decision to dispose farm produce in an accessible market. Here the dependent variable is discrete and one choice is selected ultimately by the farmer based on several factors taken up for study.
- In case the dependent variable of the regression model is a rank or ordinal variable, ordinal logit/probit model will be more appropriate. For instance, the category of adoption of a particular technology (early adopters, late adopters and non-adopters) as a function of socio-economic variable falls under ordinal logit/probit category. The models can be selected based on the nature of dependent variable (Box. 19.1).

Box 19.1: Suitability of econometric models for adoption

Adoption as categorical dependent variable

- Response categories – Adopted/Non-adopted; Adopted/Partially Adopted; Choice of variety;
- While logit model is a non-parametric model, it enjoys wider acceptability in social science research;
- Probit which is based on normal distribution is widely used for experimental data.

Adoption as a composite dependent variable

- Composite variable – Adoption Index
- Censored regression model, such as Tobit, can be used

A. Multinomial logit model

This model is similar to logit model except that the dependent variable has more categories. In this case, the independent variable adoption has qualitative categories which are not ordered. Examples are:

- Choice of varieties sown;
- Selection of brands of pesticides, fertilizers, etc.;
- Choice of occupations;
- Choice of brands of soap, television, newspaper, etc.;
- Choice of color of shirt, bag, cycle, vehicles, etc., within the brands.

Logit model generally requires more samples or data points unlike linear regression model. Multinomial logit model also requires a relatively large data set. The multinomial model generates $j-1$ sets of parameter estimates, comparing different levels of the dependent variable to a reference level. This makes the model considerably more complex, but also much more flexible. The general form of the model is

$$P\left(Y_i = \frac{1}{X_i}\right) = \frac{1}{1 + \sum_{j=2}^J \exp(X_i \beta_j)} \text{ for } m=1$$
$$P\left(Y_i = \frac{m}{X_i}\right) = \frac{\exp(X_i \beta_j)}{1 + \sum_{j=2}^J \exp(X_i \beta_j)} \text{ for } m>1$$

1. Ordinal logistic regression

The ordinal logit model is very similar to the multinomial model except that the values or classes are ordered. Variables containing scores, ranks or ratings can be used in ordinal logit model. Examples are:

- Different levels of adoption of varieties – fully adopted, partially adopted, non-adoption;
- Rating of nine varieties of rice crop based on its adaptation to climate change.

The ratings in ordinal logit are modeled as

$$\text{Rating}=1, \log \frac{P_1}{P_2+P_3+\dots+P_9}$$

$$\text{Rating}=1 \text{ or } 2, \log \frac{P_1+P_2}{P_3+\dots+P_9} \dots\dots$$

$$\text{Rating}=1 \text{ or } 2 \text{ or } \dots \text{ or } 8, \log \frac{P_1+P_2+\dots+P_8}{P_9}$$

For example: The following data set containing the preference ratings of varieties of rice (Table 19.1) chosen in case of an extreme event (drought).

Table 19.1: Varieties of rice crops suitable for climate change; An option survey of farmers' selected particular option

	Ratings 1	Ratings 2	Ratings 3	Ratings 4	Ratings 5	Ratings 6	Ratings 7	Ratings 8	Ratings 9
Variety 1	0	0	1	7	8	8	19	8	1
Variety 2	6	9	12	11	7	6	1	0	0
Variety 3	1	1	6	8	23	7	5	1	0
Variety 4	0	0	0	1	3	7	14	16	11

Note: The figures in the table indicate frequency of farmers' selected particular option which is used as input for fitting ordinal logistic model.

2. Interpretation of results

(i) Odds ratio

An odds ratio (OR) is a measure of association between an exposure and an outcome. The OR represents the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure.

$$\text{OR} = \exp(b)$$

Where b_i are estimates of logistic model. Odds ratio describes the odds of selecting an option over the reference option.

(ii) Pseudo Rsquare

This measure is used to compare among the models. SPSS outputs three Rsquare measures, viz., Cox and Snell, Nagelkerke, and McFadden. All these tests are used for comparisons, however, varying in the degree of rejection of null hypotheses.

(iii) Reference category

Reference category is used for comparison with other selected categories. Care should be taken to choose the reference category so that the logistic estimates are positive which helps in proper interpretation.

(iv) Receiver Operating Characteristic curve (ROC)

ROC curve is a graphical plot drawn by plotting the true positive rate (known as sensitivity) against the false positive rate (1 - specificity) at various threshold settings. This analysis provides tools to select possibly optimal models and to discard suboptimal ones.

B. Multinomial logit using SPSS

The model dataset mlogit.sav, which is provided along with SPSS software is used in this example. The same can be downloaded from University of California, Los Angeles website. This data set explains an experiment involving 200 high school students and their scores on various tests, including a video game and a puzzle. The outcome measure in this analysis is the students' favorite flavor of ice cream – vanilla, chocolate or strawberry. The screenshot below shows the type of variable used in multinomial logit model. The dependent variable (ice cream) is measured as nominal, and similarly the gender variable (FEMALE).

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	id	String	9	0		None	None	8	Left	Nominal	Input
2	female	Numeric	9	0	gender	{0, male}...	None	8	Right	Nominal	Input
3	ice_cream	Numeric	10	0	favorite flavor of ...	{1, chocolat...	None	8	Right	Nominal	Target
4	video	Numeric	9	0	score on video ...	None	None	8	Right	Scale	Input
5	puzzle	Numeric	9	0	score on puzzl...	None	None	8	Right	Scale	Input
6											
7											
8											

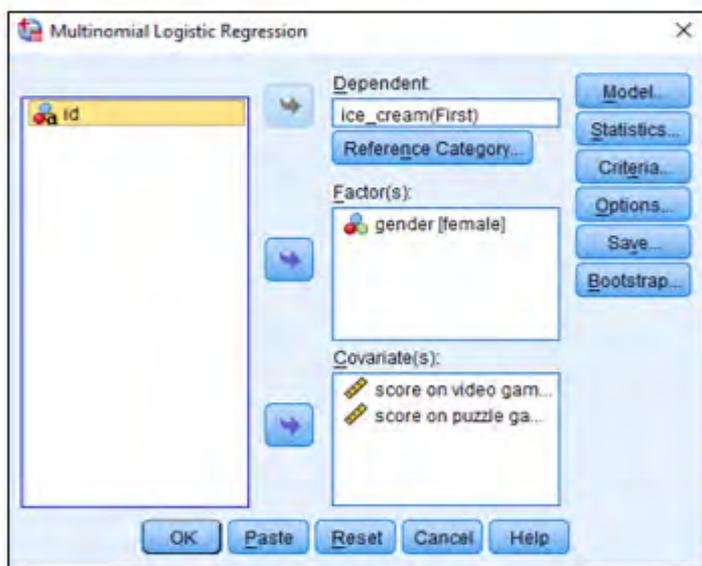
For employing multinomial logit model, select the appropriate option: ANALYZE > REGRESSION > MULTINOMIAL LOGISTIC ... (refer to the screenshot below).

The screenshot shows the SPSS Analyze menu with the following options:

- Reports
- Descriptive Statistics
- Tables
- Compare Means
- General Linear Model
- Generalized Linear Models
- Mixed Models
- Correlate
- Regression**
 - Automatic Linear Modeling...
 - Linear...
 - Curve Estimation...
 - Partial Least Squares...
 - Binary Logistic...
 - Multinomial Logistic...**
 - Ordinal...
 - Probit...
 - Nonlinear...
 - Weight Estimation...
 - 2-Stage Least Squares...
 - Optimal Scaling (CATREG)...
- Loglinear
- Neural Networks
- Classify
- Dimension Reduction
- Scale
- Nonparametric Tests
- Forecasting
- Survival
- Multiple Response
- Missing Value Analysis...
- Multiple Imputation
- Complex Samples
- Quality Control
- ROC Curve...

The background shows a data table with columns 'id', 'female', and 'puzzle'.

Ice cream is chosen as the dependent variable and chocolate is used as the reference category by specifying the option FIRST. Though it is not mandatory, the reference category is chosen in such a way that the odds-ratio is greater than one.



The output is given below:

Pseudo R-Square

Cox and Snell .153
Nagelkerke .174
McFadden .079

Parameter Estimates

Favorite flavor of ice cream ^a		B	Std. Error	Wald	df	Sig.	Exp (B)
Vanilla	Intercept						
	video	.024	.021	1.262	1	.261	1.024
	puzzle	.039	.020	3.978	1	.046	1.040
	[female=0]	.817	.391	4.362	1	.037	2.263
	[female=1]	0 ^b	.	.	0	.	.
Strawberry	Intercept	-6.819	1.442	22.351	1	.000	
	video	.046	.025	3.430	1	.064	1.048
	puzzle	.082	.024	11.816	1	.001	1.085
	[female=0]	.849	.448	3.592	1	.058	2.338
	[female=1]	0 ^b	.	.	0	.	.

^aThe reference category is chocolate.

^bThis parameter is set to zero because it is redundant.

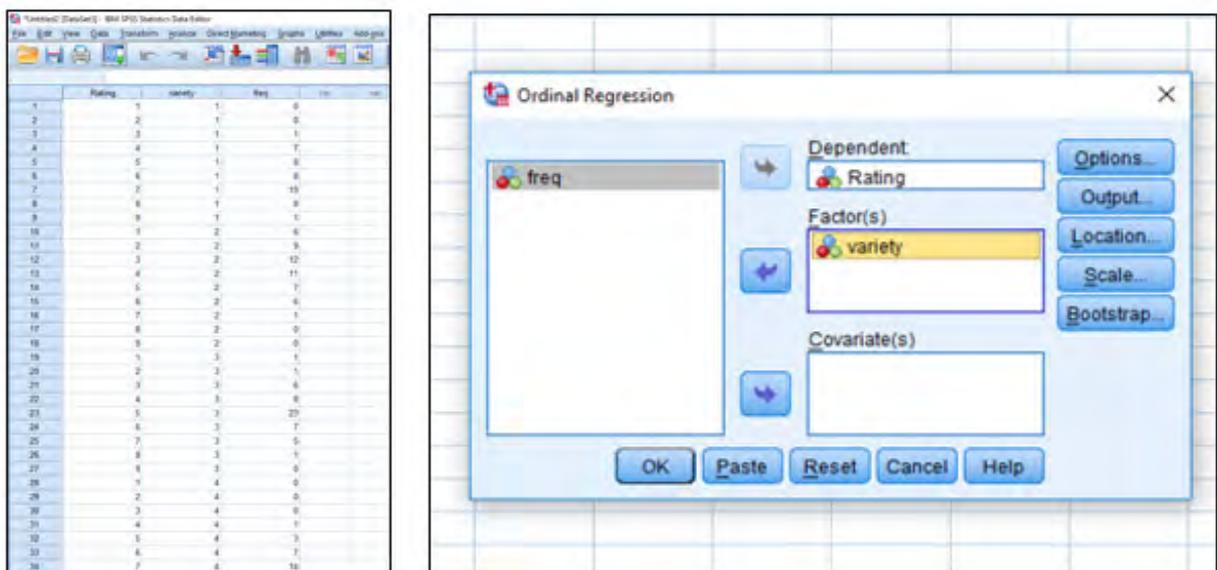
Pseudo R-square should be interpreted with caution and at least it should not be used as a measure of capture of variation in Y as the case in linear regression. It may be used only for comparison among models. The higher the pseudo R-square, the better the fit is. The third flavor, chocolate, is taken as the reference category by selecting the FIRST option in SPSS. See the dialog box where FIRST option is

specified next to Ice cream under dependent variable. The parameter estimates showed that in vanilla equation, both the variables video and puzzle are insignificant. While in strawberry equation, the choice of the ice cream is dependent on video and puzzle scores. This reveals that as the score of video and puzzle increases, strawberry flavor is preferred over chocolate flavor. Exp(B) shows the odds-ratio of respective variables. Female=0 indicates category 'MALE'.

A detailed discussion on this example is given in IDRE-UCLA Webpage (<http://www.ats.ucla.edu/stat/spss/output/mlogit.htm>)

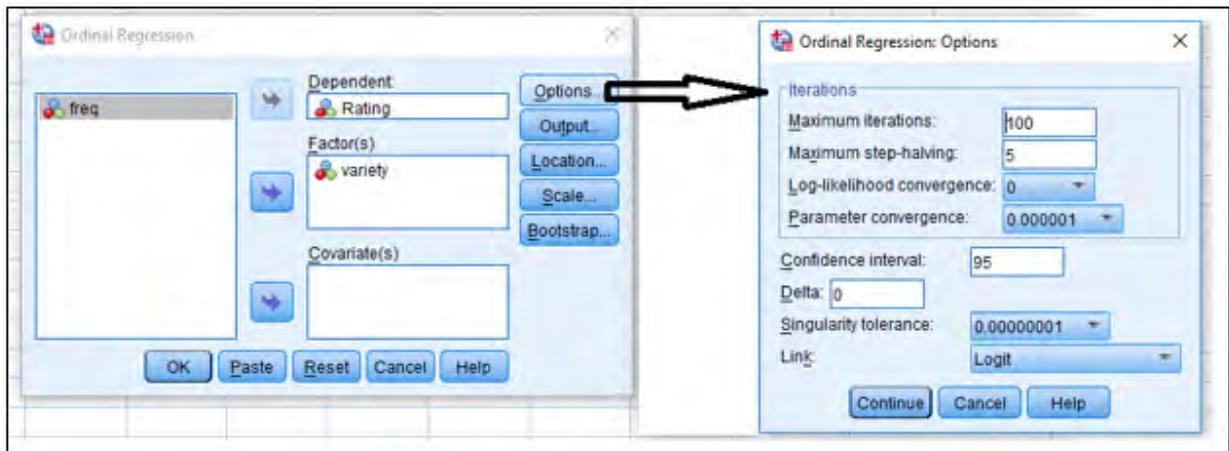
C. Ordinal logit using SPSS

The model dataset orice.sav is used for this demonstration. The dataset is fed into SPSS in the format as shown in the screenshot. Then to call Ordinal Logit dialog box, click ANALYZE > REGRESSION > ORDINAL.



The variable VARIETY is nominal, hence included under FACTOR(S). Don't forget to select the option LOGIT link after clicking OPTIONS... as shown in the figure below.

		Estimate	Std. Error	Wald	df	Sig.
Threshold	[Rating = 1]	-7.080	.562	158.486	1	.000
	[Rating = 2]	-6.025	.475	160.551	1	.000
	[Rating = 3]	-4.925	.427	132.948	1	.000
	[Rating = 4]	-3.857	.390	97.709	1	.000
	[Rating = 5]	-2.521	.343	53.971	1	.000
	[Rating = 6]	-1.569	.309	25.838	1	.000
	[Rating = 7]	-.067	.266	.063	1	.801
	[Rating = 8]	1.493	.331	20.344	1	.000
Location	[variety=1]	-1.613	.378	18.226	1	.000
	[variety=2]	-4.965	.474	109.645	1	.000
	[variety=3]	-3.323	.425	61.094	1	.000
	[variety=4]	0 ^a	.	.	0	.



The output of ordinal logistic model is given below.

Parameter estimates

Link function: **Logit**.

a. This parameter is set to zero because it is redundant.

Here the rating 9 is taken as the reference category, if rating=1 was the reference category, then the coefficients would have been positive and easy to interpret. There is no option of specifying the reference category in SPSS dialog box other than FIRST and LAST. It may either be done through syntax mode or by reversing the rating scale. Finally, the ordinal logit yields $n-1$ intercepts where n =number of ratings. The odds ratio may be computed by taking exponential of the estimate.

19.3 Key Points

- Measuring adoption of a technology is a multi-faceted process, which involves various methods of analysis and employing various analytical tools as envisaged by researchers/managers/extension functionaries.
- In advanced modelling, for identifying the factors influencing the adoption of a technology, models such as binary logistic modeling or logit, multinomial logit, and ordinal logit models are more pertinent.
- For the analysis where adoption variable is binary, say, adopted or not adopted, logit model is used. However, very little literature is available where advanced models such as multinomial or ordinal model are employed in adoption studies even though they are more pertinent. For instance, in cases where the dependent model is having more choices or categories, such as a few varietal options for farmers, choice from a few brands of pesticides available, etc., the multinomial model is appropriate.
- The multinomial model generates $j-1$ sets of parameter estimates, comparing different levels of the dependent variable to a reference level. This makes the model considerably more complex, but also much more flexible.
- The ordinal logit model is very much similar to the multinomial model except that the values or classes are ordered. Variables containing scores, ranks or ratings can be used in ordinal logit model.
- In this module, an illustration is provided where Rating of nine varieties of rice crop based on its adaptation to climate change were used to fit ordered logit model. Measures, such as Pseudo Rsquare and rationale in selecting the reference category, are discussed in this module along with reference about testing the fit using ROC curve.

- The screenshot of SPSS showing the steps involved in performing multinomial logit model and ordinal logit model is provided in the module.

References

- Agresti, A. (1996.) An introduction to categorical data analysis. New York, NY: John Wiley and Sons.
- Agresti, A. (2002.) Categorical data analysis. Second edition. New York, NY: John Wiley & Sons.
- Belsley, D. A. (1991.) Conditioning diagnostics: Collinearity and weak data in regression. New York, NY: Wiley.
- Burnham, K. P. and Anderson, D. R. (2002.) Model selection and multimodel inference. New York, NY: Springer.
- Doss C. (2006.) Analyzing technology adoption using micro studies: limitations, challenges, and opportunities for improvement. *Agri. Econ.* 34:207–219.
- Greene, W. (2000.) *Econometric Analysis* Fourth edition. New York, NY: Wiley.
- Hosmer, D. W. and Lemeshow, S. (2000.) *Applied logistic regression*. Second edition. New York, NY: John Wiley & Sons.
- Long, J. S. (1997.) *Regression models of categorical and limited dependent variables*. Thousand Oaks, CA: Sage.
- Madsen, M. (1976.) Statistical analysis of multiple contingency tables; Two examples. *Scand. J. Statist.* 3:97-106.
- SAS Institute, Inc. (2004.) *SAS/STAT 9.1 user's guide*. Cary, NC: SAS Institute Inc.
- Stokes, M. E., Davis, C. S., and Koch, G. G. (2000.) *Categorical data analysis using the SAS system*. Cary, NC: SAS Institute Inc.

20

Scale Development in Extension Research

P Sethuraman Sivakumar

Objectives

1. Introduce various scale development approaches followed in extension research
2. Explain the scale construction approaches in the classical measurement paradigm

20.1 Introduction

- Scale development is an integral part of any empirical social research. The scales help the researcher in measuring socio-psychological phenomenon in a simple and accurate manner.
- Traditionally, scale development research employed a classical measurement paradigm for isolating and quantifying socio-psychological phenomenon.
- Recent advances in the measurement paradigm has brought in various tools and techniques, which drastically changed measurement systems.
- This module describes various paradigms in scale development research, and associated methods, techniques and tools employed in developing sound scales.

20.2 Discussion

(1) Measurement

- Assessment, estimation, observation, evaluation, appraisal or judgment of an event/ phenomenon.
- Performed through assignment of numerals to objects or events according to rules (Stevens, 1951).
- Identifying the constructs (latent constructs - variable that is not directly observable), items which define the latent constructs and specifying the relationship between the constructs is a prerequisite for deciding on the measurement methods.

(2) Levels of measurement

Refers to the relationship among the values that are assigned to the attributes for a variable. The four levels of measuring social variables include:

- Nominal – Classifying objects into mutually exclusive categories based on a property;
- Ordinal – Ordering objects based on the degree of possessing a property;
- Interval – Measuring attribute distances among objects;
- Ratio – Measuring actual level of possessing a property.

(3) Approaches to measurement

Salzberger and Koller (2013) has identified four measurement approaches, such as (i) classical test theory; (ii) formative models; (iii) the C–OAR–SE approach; and (iv) item response theory.

(i) Classical test theory

- Theory of testing based on the idea that a person's observed or obtained score on a scale/test is the sum of a true score and an error score;
- Separates error score from true scores and improves the reliability of the measures;
- Assumes a reflective model – In case of latent variable models, it assumes that causality flows from the latent variable to items. In other words, it assumed that the construct defined the items;
- Three measurement models – Parallel, tau-equivalent and congeneric measures.

(a) Parallel measures

- If the researcher applies the same scale or test on subjects multiple times at different time intervals, the item score still remains the same;
- The measures have equal true score variances and equal error variances.

(b) Tau equivalent measures

- A scale or test contains items which produce same true variances on multiple administration, but have different error scores.

(c) Congeneric measures

- If the researcher applies the same scale or test on subjects multiple times at different time intervals, it produces different true and error scores for each attempt;
- Popular measurement model of classical test theory.

(ii) Formative or index models

- In case of latent variable models, the formative model assumes that causality of a phenomenon flows from items or variables to the latent variable (Edwards and Bagozzi, 2000);
- Assumes the items or variables which define the construct;
- Index is a summary of multiple measures and these models are not measuring the construct per se.

(iii) The C–OAR–SE approach

- C-OAR-SE stands for construct definition–object representation–attribute classification–rater entity identification–selection of item type and answer scale–enumeration and scoring rule (Rossiter, 2010);
- The classical test approaches merely model the relationship between measures and scores, relying excessively on statistics to assess validity, while ignoring the construct and disregarding content validity;
- In C-OAR-SE, the construct is defined by including the object to which the attribute refers (e.g., extension agency), and the rater entity (e.g., farmers), as well as the attribute itself (e.g., service quality).

(iv) Item response theory (IRT)

- Based on the relationship between individuals' performances on a test item and the test takers'

- levels of performance on an overall measure of the ability that item was designed to measure;
- Focuses on individual responses to particular items rather than aggregate statistics, as followed in classical test theory.

(4) Scale

Scale is a collection of items, the responses to which are scored and combined to yield a scale score. Scale scores can be categorized according to level of measurement.

(5) Scale development strategies

There are various strategies used for constructing scales. Scale development strategies are broadly classified as

- Deductive; and
- Empirical (Kaplan and Saccuzzo, 2005).

Deductive strategies comprise (i) logical-content, and (ii) theoretical approach; whereas empirical strategies are comprised of (i) stimulus-centred; (ii) subject-centred; and (iii) response scale strategies (Dawis, 1987).

(a) Deductive strategies

Deductive strategies employ reasoning and deductive logic to determine the meaning of a scale response.

(i) Logical-content strategy

- Scale developers use their judgements to logically deduce the content for selecting items, which are obviously related to the characteristic being measured;
- Relies on face validity of the items;
- Also called rational/content/intuitive strategy.

(ii) Theoretical strategy

- Uses extension theory to determine the content of the scale items;
- Uses a deductive approach to select items which are representing a theory;
- Develops a homogeneous scale using standard statistical procedures, such as item analysis, etc.

Both the theoretical, rational, and logical strategies are no longer popular methods in scale development.

(b) Empirical strategies

Empirical strategies rely on data collection and statistical analyses of item responses for selecting suitable items.

(i) Stimulus-centred or judgement strategies

- Judges the degree of item differences along the measurement dimension;
- Types – Turnstone equally appearing interval scale, Q sort method and paired comparison methods.

(ii) Subject-centred or individual difference strategies

- Measures individual differences among respondents with respect to a particular attribute or aspect;
- Examples – Likert scales, factor analysis, semantic differential scales.

(iii) Response scale strategy

- The scale score are assigned both to stimuli (items) and subjects (respondents);
- Examples – Guttman scalogram, Rasch scale and external criterion methods.

(A) Stimulus-centred or judgement strategies

The Turnstone's scaling methods (Thurston and Chave, 1929)

Louis Leon Thurstone was one of the first, and widely popular, scaling theorists, who invented three different methods for developing a unidimensional scale:

- (a) method of equal-appearing intervals;
- (b) method of successive intervals;
- (c) method of paired comparisons.

These three methods differed in how the scale values for items were constructed, but in all three cases the resulting scale was rated the same way by respondents. This section focuses only on equal-appearing interval method, and the other two methods are described elsewhere (Edwards and Gonzalez, 1993; Ray and Mondal, 2014).

(a) Equal-appearing interval method

1. Generating potential scale items – About 80 to 100 items or variables representing different dimensions of the construct, (e.g., attitude of farmers to GMO crops), are drafted.

2. Rating scale items – About 20 to 30 expert judges are asked to rate each statement on a 1 to 11 point scale in terms of how much each statement indicates a favourable attitude towards GMO crops (1= Extremely unfavourable attitude towards GMO to 11= an extremely favourable attitude toward GMO).

3. Computing scale values – For each statement, the median and the inter-quartile range are computed. The median is the value above and below which 50% of the ratings fall. The first quartile is the value below which 25% of the cases fall and the third quartile is the value above which 75% of the cases fall. The interquartile range is the difference between the third and first quartiles, or Q3 minus Q1. To facilitate the final selection of items for the scale, sort the table of medians and Interquartile Range in ascending order by Median, and as descending order by Interquartile Range (Table 20.1).

Table 20.1: Analysis of judge's rating of items

Item or variable number	Median	Q1	Q3	Interquartile range (Q3-Q1)
24	1	1	2.5	1.5
12	1	1	2	2
21	2	1	5	4

4. Selecting the final scale items – An item or variable with a higher median (6+) and a small inter-quarter range would show stronger agreement of favourability among the judges and may be selected for the final scale. Items with high inter-quartile range are deleted. The average median value of each selected item represents its scale value.

5. Administering the scale – The selected items are randomly placed in a questionnaire with an agreement response type (1 = Agree, 0 = Disagree) and the respondents are asked to indicate their endorsement of each scale statement/item by indicating whether they agree or disagree with the statement. The respondents' total scale scores are computed by taking the average of scale values that a specific respondent has agreed with.
A typical Thurstone scale contains 22 items.

(b) Q-sort method (Stephenson, 1953)

Used for systematically studying the subjectivity, viewpoints, opinion, beliefs and attitude of a person about a specific phenomenon.

Steps

1. Definition of domain of a particular issue (e.g. attitude towards SRI technology);
2. Development of a pool of statements (Concourse) through personal interviews, focus groups and secondary data;
3. After deciding on statements, the categories and sub-categories of the topic are developed. This procedure helps the researcher to select the statements which represent all the dimensions of the topic (e.g., feasibility, cost and returns, adaptability to farming system, etc);
4. Selection of participants representing different perspectives about the topic of study (P- sets);
5. From the concourse, several sets of statements are randomly selected for providing to each respondent. To ensure content validity, all the sample statements are usually reviewed by domain experts and tested in one or more pilot studies;
6. Each participant is supplied with a set of statements (Q-sets) which are randomly selected from concourse and asked to rank them in agreement continuum (most disagreed, neutral, and most disagreed) on a symmetrical Q-grid. The statements are usually written on separate cards and respondents are asked to place them in an appropriate place in the Q-grid;
7. Q-Grid is the answer sheet used in Q-methodology, which forces the respondents to place cards into the shape of a quasi-normal distribution. There are fewer statements that can be placed at the extreme ends and more that are allowed to go into the middle area. The middle represents the grey zone, or almost neutral, reaction. A sample Q grid for 32 statements is given in Figure 20.1.

Response	Most disagreed					Neutral					Most agreed
Score	-5	-4	-3	-2	-1	0	1	2	3	4	5
Statements	19	8	9	21	15	27	31	1	6	32	3
		30	1	14	28	16	4	18	12	7	
			17	11	20	13	26	22	5		
				23	29	2					
					24						
					10						

Fig. 20.1: An example of Q- Grid

8. The data are analysed through correlation and by-person factor analysis;
9. A detailed description of Q-methodology can be found in Valenta and Wigger (1997) and Coogan and Herrington (2011).

(c) Rank-order methods

(i) Paired comparison method (Guilford, 1954): Based on Thurstone’s law of comparative judgement. The stimuli (e.g., person, object, event, state, or condition) are presented in pairs in all possible combinations with every other stimuli. The respondent’s task is to select one stimulus from each pair on the basis of the scaling dimension, that is, the basis of comparison. From the number of times each stimulus is chosen, the stimuli can be rank-ordered with more precise information than if all of the stimuli were just rank-ordered in the usual way. A detailed description of this method is provided by Edwards (1957).

(B) Subject-centered strategies

(i) The Likert method of summated ratings (Likert, 1932): In the Likert method, a large number of items representing a subject domain are selected and rated on the 5-point bipolar agreement continuum (1 = Strongly Disagree to 5 = Strongly Agree) by a large group of respondents (> 100). Items which have high reliability (Cronbach alpha > 0.60) and discriminating ability are selected for the final scale. The scale scores are obtained by summing the item scores for the selected items. A detailed description of this method is provided in Edwards (1957).

(ii) Factor analysis: is a data reduction technique in which a large number of variables are reduced to a smaller set of factors without much loss of information. It can be used to select items for Likert-type scales. A detailed discussion of factor analysis is given in Module 21.

(iii) Semantic differential (Osgood et al., 1957)

- A special type of rating scale which measures the connotative meaning of objects, events, and concepts, to derive the respondents’ attitude towards them;
- Uses several rating dimensions for rating the same item or stimulus object;
- The rating dimensions are bipolar, anchored at both ends with contrasting adjectives, with a 7-point rating continuum;
- Usually measures three dimensions, such as (i) Evaluation - ‘good-bad’, (ii) potency - ‘powerful-weak,’ and (iii) activity – ‘active – passive’;
- For example, a farmer’s attitude towards the service quality of Krishi Vigyan Kendra (KVK) can be measured using the following Semantic Differential scale:

	-3	-2	-1	0	1	2	3	
Innovative								Conservative
Helpful								Unsympathetic
Useful								Not useful
Timely								Late
Good								Bad
Reliable								Unreliable

- A detailed account on using Semantic Differential in attitude research can be found in <http://www.indiana.edu/~socpsy/papers/AttMeasure/attitude.html>

C. Response scale strategies

(i) Guttman's scalogram analysis or cumulative scale (Guttman, 1944)

- A Guttman scale presents a number of items to which the person is requested to agree or not agree;
- Includes a set of items or statements so that a respondent who agrees with any specific question in the list will also agree with all previous questions;
- Helps to predict item responses from the total score for the respondent (Abdi, 2010);
- Used for grouping respondents based on their perceived wealth through wealth ranking (Guest, 2000).

Steps (Changing minds, 2017)

1. Generate a list of possible statements on a specific topic, about 80 – 100 numbers;
2. Get a set of judges to score the statements with a Yes or No, depending on whether they agree or disagree with them (judges' favourableness towards each item);
3. Draw up a table with the respondents in rows and statements in columns, showing whether they answered Yes or No;
4. Sort the columns so the statement with the most Yes's is on the left;
5. Sort the rows so the respondent with the most Yes's is at the top;
6. Select a set of questions that have the least set of 'holes' (No's between 'Yes's);
7. Estimate the scale value for each statement using statistical analysis;
8. After selecting the final items, present the items and ask the respondent to check items with which they agree;
9. The respondents' scale score is computed by summing the scale values of items they agree with.

(ii) The Rasch method

- A model for analysing categorical data, which postulates that item response which is a function of two parameters – an item parameter and a person parameter;
- The probability of the respondent choosing a specified response (e.g., right/wrong answer) is modelled as follows:
 - For any item, the respondent with higher ability should have a higher probability of getting the item right than a person with lower ability;
 - Given any person, an item of lower difficulty should be solved (gotten right) with a higher probability than would an item of higher difficulty.
- A detailed account of Rasch analysis can be found at <http://www.rasch-analysis.com/>

(iii) External criterion methods: Heterogeneous and multi-dimensional scales. The three most-used methods of item selection are:

- (a) the group difference method;
- (b) the item validity method; and
- (c) the multiple regression method.

(a) Group difference method: Items for the scale are selected by administering them to two groups:

- a high and low criterion group; or
- a criterion group and a control.

Items with a larger mean difference are selected for inclusion in the scale. The size of the difference can

be used to give differential weights to items and response choices.

(b) Item validity method: Comparable to the Likert scale method, except that instead of the Likert's item-total-score correlation, this method uses the correlation between item score and criterion score as the basis for item selection.

(c) Multiple regression method: The criterion variable is regressed on selected items so as to identify the item weights. Items are added to the equation one-by-one and the variance explained by each item is used to select the specific item for the final scale. High item correlation with criterion variability and its low correlation with other items are used as selection criteria for inclusion in the final scale.

20.3 Key Points

- Measurement is the assessment, estimation, observation, evaluation, appraisal or judgment of an event/phenomenon.
- Four measurement approaches, such as (i) classical test theory; (ii) formative models; (iii) The C-OAR-SE approach; and (iv) item response theory are widely used in social science research.
- Scale development strategies are broadly classified as: (i) deductive; and (ii) empirical. Deductive strategies are comprised of (i) logical-content and (ii) theoretical approach; whereas empirical strategies are comprised of (i) stimulus-centred; (ii) subject centred; and (iii) response scale strategies.

References

- Abdi, H. (2010.) Guttman scaling. In Salkind, N. J., ed., *Encyclopaedia of Research Design*. SAGE Publications.
- Changing minds. (2017). Guttman scale. Retrieved from http://changingminds.org/explanations/research/measurement/guttman_scale.htm
- Coogan, J., and Herrington, N. (2011.) Q Methodology: an overview. *Research in Secondary Teacher Education* 1(2):24–28.
- Dawis, R. V. (1987.) Scale construction. *Journal of Counselling Psychology* 34, 481-489.
- Edwards, A. L., and Gonzalez, R. (1993.) Simplified successive interval scaling. *Applied Psychological Measurement* 17, 21-27.
- Edwards, J. R., and Bagozzi, R. P. (2000.) On the nature and direction of relationships between constructs and measures. *Psychological Methods*, 5(2):155–174.
- Guest, G. (2000.) Using Guttman scaling to rank wealth: Integrating quantitative and qualitative data. *Field Methods* 12:346-357.
- Guilford, J. P. (1954.) *Psychometric methods*. Second edition. New York: McGraw-Hill.
- Guttman, L. (1944.) A basis for scaling qualitative data. *American Sociological Review* 9:139-150.
- Kaplan, R.M., and Saccuzzo, D.P. (2005.) *Psychological testing: Principles, applications, and issues*, Sixth edition. Belmont, CA: Thompson Wadsworth.
- Likert, R. (1932.) A technique for the measurement of attitudes. *Archives of Psychology* No. 140.
- Osgood, C. E., Suci, C. J., and Tannenbaum, P. H. (1957.) *The measurement of meaning*. Urbana: University of Illinois Press.
- Rossiter, J. R. (2002.) The C-OAR-SE procedure for scale development in marketing. *International Journal of Research in Marketing* 19:305–335.
- Salzberger, T., and Koller, M. (2013.) Towards a new paradigm of measurement in marketing. *Journal of Business Research* 66(9):1307-1317.
- Stephenson, W. (1953.) *The study of behavior*. Chicago: University of Chicago Press.
- Stevens, S. S. (1951.) Mathematics, measurement, and psychophysics. In: S. S. Stevens (ed.), *Handbook of experimental psychology*. New York: Wiley. Pp. 1 -49.

- Thurstone, L. L., and Chave, E. J. (1929.) *The measurement of social attitudes*. Chicago: University of Chicago Press.
- Valenta, A., and Wigger, U. (1997.) Q-methodology: Definition and application in health care informatics. *Journal of American Medical Informatics Association* 4(6):501–510.



Photo Credit: S Ranadive

21

Developing Summated Scales - Latent Factor Identification

P Sethuraman Sivakumar

Objectives

1. Provide an overview of summated scales
2. Describe steps in developing summated rating scales

21.1 Introduction

- In the classical test approach, survey research design is the common approach employed for developing summated scales to measure social and psychological attributes.
- Summated scales are developed by combining several individual variables into a single composite measure. Summated scales are popular in extension research since they (i) provide means of reducing measurement error; and (ii) represent multiple aspects in a single measure.
- A number of authors have recommended different steps in constructing summated scales (Churchill, 1979; Anastasi, 1988; Hinkin et al., 1997; DeVellis, 2003).
- Considering the requirements of extension science, a modified approach for developing summated scales is developed and displayed in Figure 21.1.

21.2 Discussion

Steps in developing summated scales

Step 1: Define the construct clearly

- Use existing theory and research to guide research in order to provide a sound conceptual foundation;
- Define the construct and items clearly and precisely.

Step 2: Item generation

- Content of the scale – Select the items that reflect the scale's purpose, i.e., all the items in the scale should reflect the latent variable/construct underlying them;
- A large number of items, about 80-100, which represent different indicators and latent construct under study, are selected. Adequate internal consistency reliability can be obtained with four or five items per scale (Harvey et al., 1985).

Step 3: Content adequacy assessment

- A psychometric property which exists when the content of a measure covers a representative sample of the domain to be measured (Anastasi, 1982);
- Assessing the scale content adequacy prior to final questionnaire development using domain specific experts is a necessary step to ensure adequate construct validity;
- Various methods for assessing content adequacy are described in Annexure A-4.

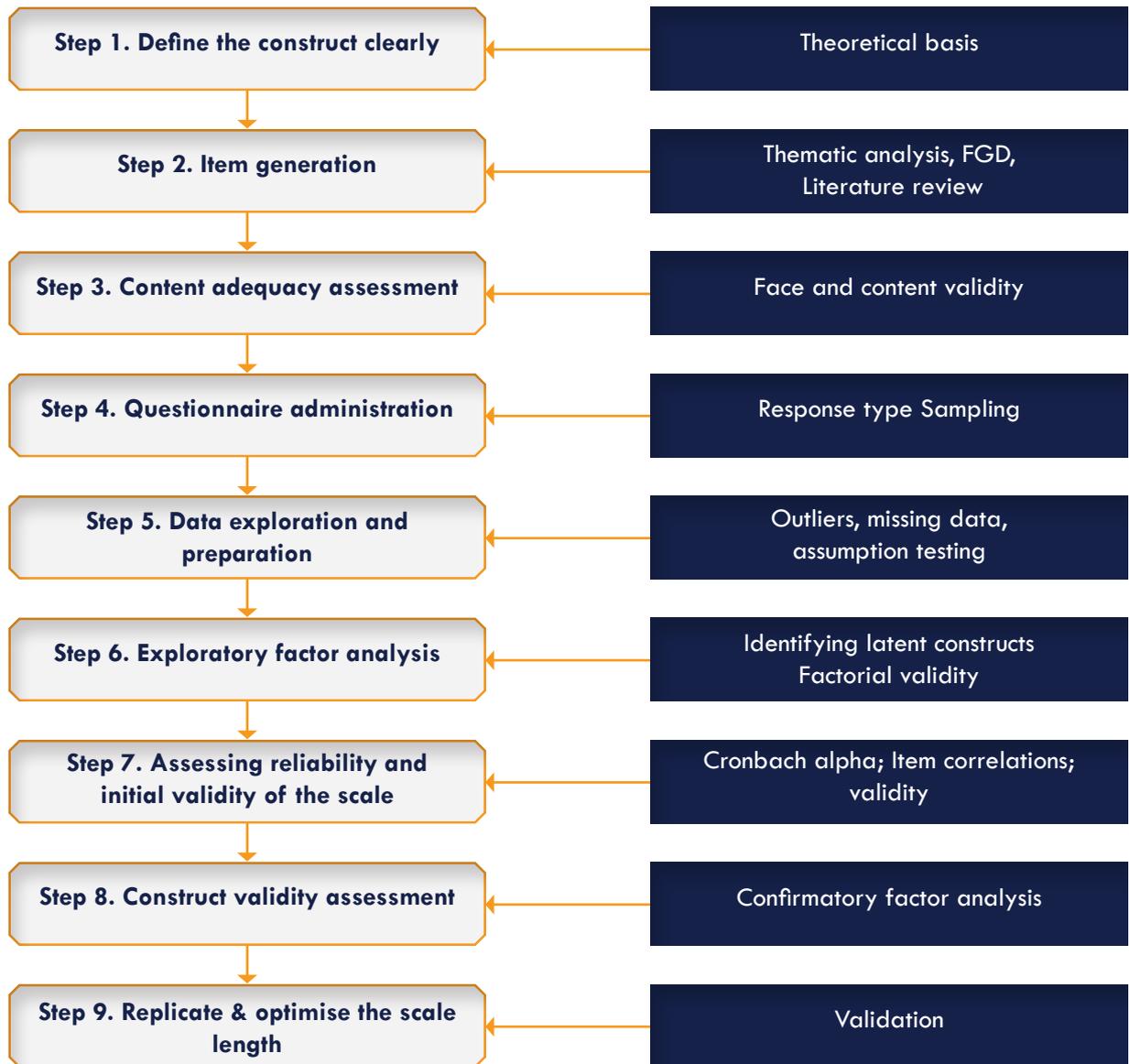


Fig. 21.1: Modern scale development process (Adapted from Churchill, 1979; Hinkin et al., 1997 and DeVellis, 2003)

Step 4: Questionnaire administration

After content adequacy assessment, the selected items are administered to a sample for developing a measurement model which explains their reliability and validity. It has two stages – item scaling and sampling.

(i) Item scaling

- The selected items are arranged in a questionnaire with a response format developed, based on the objectives and domain of the study;
- Likert scales are the most commonly used in survey research for developing scales. Likert scales with five- or seven-point scales can generate adequate variance for examining the relationships among items and scales, and create adequate internal consistency reliability (Lissitz and Green, 1975).

(ii) Sampling

- Optimum sample size for developing sound scales is estimated based on item to respondent ratio. It indicates the minimum number of respondents required for each item or statement in the initial scale;
- Ideal item to respondent ratio ranges for exploratory/confirmatory factor analysis is 1:5;
- For exploratory factor analysis, a minimum of 150 samples are required if the item intercorrelations are reasonably strong (> 0.3) (Guadagnoli and Velicer, 1988);
- For confirmatory factor analysis, a minimum sample size of 100 is required (Bollen, 1989).

Step 5: Data exploration and preparation

Various steps in preparing data for analysis are described in Modules 11, 12 and 13.

Step 6: Conducting Exploratory Factor analysis

- Factor analysis is a statistical method used to describe variability among observed and correlated variables by deriving a small number of latent or unobserved variables or factors;
- Factors – Unobserved constructs represented by highly correlated variables (items);
- Uses correlations among many variables to sort related variables into groups called factors;
- Used only with metric or continuous variables – interval and ratio scales. For categorical or non-metric variables including binary, nominal and ordinal levels, Boolean factor analysis is used.

Objectives of factor analysis

There are two ways of specifying objectives of factor analysis.

(i) Based on unit of analysis (Variable or respondent)

(a) R type factor analysis

- Unit of analysis – attributes of items or variables;
- Analysing a set of variables to identify the dimensions that are latent (not easily observable);
- Widely used form of factor analysis.

(b) Q type factor analysis

- Unit of analysis – attributes of persons;
- Analysing individual attributes for grouping them into different categories, similar to cluster analysis.

(ii) Based on manipulation of data

(a) Summarising

- To describe the data in a smaller number of concepts than original variables;
- Example, constructing a scale using item analysis.

(b) Data reduction

Use the factor loadings for identifying variables for subsequent analysis with other techniques or for making estimates of factors themselves. For example, after factor analysis, the factor scores can be used for further estimation through multiple regression, etc., (e.g., Germplasm characterisation).

Categories of factor analysis

There are two main categories of factor analysis: (i) Exploratory Factor Analysis (EFA); and (ii) Confirmatory Factor Analysis (CFA). While EFA is used to identify latent variables among the items or variables, CFA examines the factor structure identified through EFA to develop a sound measurement model for the scale.

(1) Exploratory Factor Analysis

What is EFA

- Used to identify latent factors that explain the covariation among a set of measured variables based on correlation among items;
- Explores how many factors exist among a set of items or variables and the degree to which these items or variables are related to the factors;
- Helps to estimate construct validity of a scale at the initial stages by answering the following questions (Tabachnick and Fidell, 2001):
 1. How many factors or constructs underlie the set of items?
 2. What are the defining features or dimensions of the factors or constructs that underlie the set of items?

1. Checking assumptions of EFA

(i) Conceptual assumptions

- The observed patterns in the items or variables are conceptually valid – based on sound theory;
- Samples should be homogeneous.

(ii) Statistical assumptions (Hair Jr. et al., 2006)

- Significant intercorrelation among items or variable ($r > 0.30$);
- Bartlett's test of sphericity – Significant ($p < 0.05$);
- Kaiser – Meyer – Olkin (KMO) Measure of Sampling Adequacy – 0.6 to 1;
- Linearity and normality.

2. Deriving factors and assessing overall fit

A. Selecting the factor extraction method

The method of identifying the factors that best characterize a set of variables.

Types of factor extraction in SPSS (IBM, 2017)

a) Principal components

- Create uncorrelated linear combinations of the unobserved variables;

- First component has maximum variance followed by smaller portions distributed across extracted components;
- Used when correlation matrix is singular (Determinant is zero – can be identified by checking ‘correlation matrix box’).

b) Unweighted or ordinary least-squares

- Minimize the residual between the observed and the reproduced correlation matrix;
- Used when the assumptions about the distribution of observed variables are violated, e.g., non-normality.

(c) Generalize least-squares

Used when the researcher wants factors to fit highly unique variables, with weak factor loadings.

(d) Maximum likelihood

- Generate the solution that is most likely to produce the correlation matrix;
- Requires data which has multivariate normality.

(e) Alpha factoring

Considers variables as a sample drawn directly from the population; maximizes internal consistency reliability Cronbach alpha.

f) Principal axis factoring

Similar to Principal component analysis (PCA), but it accounts for common variance in the data.

B. Exploratory factor analysis vs Principal components analysis

- The choice of using EFA or PCA depends on two criteria: Objectives of factor analysis; and Amount of prior knowledge about the variance in the variables (Table 21.1);
- Total variance in data = Common Variance among variables + Unique variance of each variable + Error variance.

Table 21.1: Difference between exploratory factor analysis and principal component analysis

Aspect	Exploratory factor analysis	Principal Component Analysis
Objective	Identification of latent constructs in the data	Data reduction for predicting the phenomenon
Nature of factors extracted	Correlated factors	Uncorrelated factors
Variance extracted	Considers only common variance (unique variance is not considered) – Common + error variances	Considers total variance (Common + Unique + error)

Aspect	Exploratory factor analysis	Principal Component Analysis
Application in social sciences	Suitable for scale construction in social sciences as it extracts correlated factors and considers only common variance	Not suitable for developing scales in social science research because: <ul style="list-style-type: none"> • Most socio-psychological latent factors are inherently correlated; • Social science scales are unidimensional or multi-dimensional, which are defined by common variance.

C. Criteria for number of factors to extract

(i) Latent root or Kaiser's eigen value criterion

- Any individual factor should account for the variance of at least single variable if it is to be retained for analysis;
- Factors with eigen value > 1 are selected (Kaiser, 1960);
- Reliable when the number of variables is between 20-50;
- Can be applied only to Principal Components extraction method. Since other methods like Principal Axis Factoring and Maximum Likelihood Methods assume that the variance of a variable is less than 1, these criteria are not applicable (Russell, 2002).

(ii) Percentage of cumulative variance extracted criterion

Acceptable values of cumulative variance extracted – Natural sciences: 90%, Social sciences: 60% (Hair Jr., et al., 2006).

(iii) Cattell's Scree plot criterion

A graph of the relationship between the relative magnitude of the eigenvalues and the number of factors. The point at which the curve begins to straighten out indicates the maximum number of factors to extract (Cattell, 1966).

(iv) Horn's Parallel Analysis

- Parallel analysis generates eigenvalues from a random set of data based on the same number of variables and the same number of cases;
- The randomly generated eigenvalues are plotted on a scree plot along with the actual eigenvalues;
- The factors with actual eigenvalues higher than the random eigenvalues are retained with the logic being that a factor that explains more variance than chance is meaningful (Horn, 1965).
- Most effective method for choosing the number of factors to retain in the analysis;
- SPSS Syntax for conducting Parallel analysis is available in <https://people.ok.ubc.ca/briocconn/nfactors/nfactors.html>

The researcher can use any of the above criteria to decide on the number of factors to be retained.

D. Rotation and interpretation of the factors

1. Estimate the factor matrix

- Examining the relationship between factors and variables through factor loadings – I structure or pattern coefficients;

- Initial unrotated matrix is computed, containing factor loadings of the variables;
- Factor loadings – Correlations of each variable with factor, with high loadings making the variable representative of the factor;
- Don't interpret this unrotated matrix.

2. Factor rotation

- To attain more interpretable factors, the researcher applies a rotation of the factor axes;
- Factor rotation refers to changing the reference points for the variables;
- A rotated factor solution provides a different set of structure coefficients than the unrotated solution. The rotated structure coefficients are used for interpretation.

Types of factor rotation (Hair Jr. et al., 2006)

(I) Orthogonal rotation

- The factors are extracted so that their axes are maintained at 90 degrees;
- Each factor is independent of (orthogonal to) all other factors (no correlation between factors) – Popular rotation in factor analysis;
- Used when the goal is to reduce data into a small number of variables or to develop a set of uncorrelated measures for subsequent analysis by other multivariate methods.

Types

- VARIMAX – Finds the pattern of structure coefficients that maximizes their variance;
- QUARTIMAX – Rotates initial factors in such a way that a variable load is high on one factor and less on others;
- EQUIMAX – Rarely used.

(II) Oblique rotation

- Factor rotation is computed so that the extracted factors are correlated;
- Useful for extracting theoretically meaningful constructs.

Types

- DIRECT OBLIMIN – Produces either high or low factor correlations;
- PROMAX – Best suited for Maximum Likelihood factor extraction method.

3. Factor interpretation and re-specification

(i) Examining factor matrix of loadings (Hair Jr. et al., 2006)

- Factor loadings are indicated through structure (e.g., Orthogonal rotation) and pattern (e.g., Oblique rotation) coefficients.
 - Structure coefficients – Zero-order correlations between factor and variables; displayed in the Rotated Component Matrix;
 - Pattern coefficients – Regression coefficients which indicate unique loads of factors into variables; displayed in the Pattern Matrix.
- In case of oblique rotation, the factors are correlated and produce large structure coefficients. Therefore the pattern coefficients are used for interpretation.

(ii) Identify significant loading(s) for each factor

Method 1: Practical significance (Hair Jr. et al., 2006)

- Factor loading should not exceed 1;
- Interpretation of structure/pattern coefficients - ± 0.30 to ± 0.40 – Minimum level of interpretation; >0.50 – Practically significant; and >0.70 – Indicates well defined structure.

Method 2: Statistical significance

- To achieve a power level of 80 percent at 5% level of significance, the sample sizes displayed in Table 21.2 may be followed (Hair Jr. et al., 2006).

Table 21.2: Statistical significance of factor loadings based on sample size

Factor loading	Minimum size	Sample
0.30	350	
0.40	250	
0.45	200	
0.50	120	
0.55	100	
0.60	85	
0.65	70	
0.75	50	

(iii) Assess the commonalities of the variables

- If any variable is found not significantly loaded on identified factors, examine the commonalities. When a variable's commonality falls below 0.50, it may be deleted (Hair Jr. et al., 2006).

(iv) Re-specify the factor model if needed

Problems in factor analysis

- A variable has no significant loadings;
- Even with a significant loading, the commonalities are low (<0.50);
- A variable has cross-loading (significant loading in more than one factor).

Remedies

- Ignore the problematic variables;
- Employ an alternative rotation method;
- Decrease or increase the factors retained;
- Modify the type of factor model used (principal component/common factor).

(v) Label the factors

- When the significant factor loadings are detected, the researcher provides a name for the extracted factors based on the nature of variables loading on them;

- In case of structure coefficients, the factor names are decided based on the magnitude of the coefficients of the variables. The stronger the correlation, the variable is likely to play a major component of the factor name;
- The pattern coefficients will not help in deciding the factor names and the researcher chooses the names of the factors using theory.

(vi) Validation of factor analysis

- Assessing the degree of generalisability of the results to the population and potential influence of individual cases on overall results.

Confirmatory factor analysis

- Validate the model using a split sample or a separate sample;
- Confirmatory factor analysis (CFA) through structural equation modeling – SPSS AMOS, LISREL or EQS. A detailed description of CFA is described elsewhere in another module.

Step 7: Assessing reliability and preliminary validity of the scale

(a) Reliability – Degree of consistency between multiple measurements of a variable.

Types

(i) Test-retest – Consistency of responses of same individuals on a specific measure at two different points of time.

(ii) Internal consistency – All the individual items of a scale measure the same construct, and exhibit substantial inter-correlation ($r > 0.30$) (Hair Jr. et al., 2006).

How to assess reliability of a scale (Hair Jr. et al., 2006)

- If the item-to-total and inter-item correlation coefficients exceed 0.50 and 0.30 respectively, the scale is considered reliable;
- Cronbach alpha (Internal consistency reliability) coefficient – Ideal alpha value for exploratory research – 0.60; Hypothesis testing – 0.70;
- Indicator reliability is measured through commonality for each variable (>0.50);
- Confirmatory factor analysis – composite reliability and variance extracted.

(b) Validity – The extent to which a scale or measure accurately represents the concept or construct of interest.

Types

(i) Content validity – Discussed in the ‘Content adequacy’ section of this module.

(ii) Convergent validity (Hair Jr. et al., 2006)

- Degree to which two measures of the same concept are correlated;
- High correlations ($r > 0.60$) indicate that the scale is measuring the intended concept;
- E.g., using two different personality measures with the same population and checking for the correlation between two results.

(iii) Discriminant validity

- Degree to which conceptually similar measures or scales are distinct;
- Low correlation between the new scale and an established scale which measures a similar concept ($r < 0.30$);
- To assess the discriminant validity, the researcher intentionally includes a conceptually similar scale in the questionnaire in order to calculate discriminant validity at a later stage;
- For e.g., when a researcher is developing a scale to measure innovativeness of farmers, he/she may include an established scale to measure Risk-taking ability since Risk-taking ability is related to innovativeness, but is not the same. After developing a preliminary scale of Innovativeness through EFA, the researcher can correlate the scores with Risk-taking ability ratings of the same sample to establish the discriminant validity of Innovativeness scale.

(iv) Nomological validity

- Degree to which the summated scale makes an accurate prediction of other concepts in the model;
- For example, Innovativeness is the predictor of respondents' attitude towards Genetically Modified Crops. If the Innovativeness scale scores can accurately predict attitude scores, the measure is nomologically valid.

Calculating summated scale scores

- When the factor scores are obtained, factors labelled, and their reliability and validity assessed, the variables are pooled under their respective factors;
- The variable scores under each factor are summed and averaged to obtain the value for the respective factor;
- Negative statements are reverse scored during summation;
- This step is followed only when the researcher wishes to stop with EFA without optimising the scale length through Confirmatory Factor analysis.

21.3 Key Points

- Summated scales are developed by combining several individual variables into a single composite measure. Summated scales are popular in extension research since (i) they provide the means of reducing measurement error; and (ii) represent multiple aspects in a single measure.
- The summated scales are constructed using a ten-step approach starting from construct identification to validation.
- Exploratory factor analysis plays a significant role in initial factor identification.

References

- Anastasi, A. (1988.) Psychological testing. Sixth edition. New York: Macmillan.
- Bollen, K.A. (1989.) Structural equations with latent variables. New York: John Wiley & Sons.
- Cattell, R. B. (1966.) The scree test for the number of factors. *Multivariate Behavioral Research*, 1:245-276.
- Churchill, G.A. (1979.) A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, 16:64-73.
- DeVellis, R. F. (2003.) Scale development: Theory and applications. Second edition. Thousand Oaks, CA: Sage.
- Guadagnoli, E. and Velicer, W.F. (1988.) Relation of sample size to the stability of component patterns. *Psychological Bulletin* 103:265-275.

- Hair, Jr., J. F., Black, W. C., Babin, B. J., Anderson, R. E., and Tatham, R. L. (2006.) *Multivariate data analysis*. Sixth edition. Upper Saddle River, NJ: Pearson Prentice Hall.
- Harvey, R.J., Billings, R.S. and Nilan, K.J. (1985.) Confirmatory factor analysis of the job diagnostic survey: Good news and bad news. *Journal of Applied Psychology*, 70:461-468.
- Hinkin, T. R., Tracey, J. B., and Enz, C. A. (1997.) Scale construction: Developing reliable and valid measurement instruments. *Journal of Hospitality & Tourism Research* 21(1):100-120.
- Horn, J. L.. (1965.) A rationale and test for the number of factors in factor analysis. *Psychometrika* 30:179–185.
- IBM. (2017.) Factor analysis extraction. IBM Knowledge Centre. (Retrieved from: https://www.ibm.com/support/knowledgecenter/en/SSLVMB_23.0.0/spss/base/idh_fact_ext.html)
- Kaiser, H.F. (1960.) The application of electronic computers to factor analysis. *Educational and Psychological Measurement* 20:141-151.
- Lissitz, R. W., and Green, S. B. (1975.) The effect of the number of scale points on reliability: A Monte Carlo approach. *The Journal of Applied Psychology* 60:10-13.
- Russell, D. W. (2002.) In search of underlying dimensions: The use (and abuse) of factor analysis. *Personality and Social Psychology Bulletin* 28:1629–1646.
- Tabachnick, B. G., and Fidell, L. S. (2001.) *Using multivariate statistics*. Fourth edition. New York: Harper & Row.



Photo Credit: Subash SP

22

Developing Summated Scales – Structural Equation Modeling

P Sethuraman Sivakumar

Objectives

1. Provide an overview of Structural Equation Modeling for scale development
2. Describe various steps involved in Confirmatory Factor Analysis for constructing summated scales

22.1 Introduction

- Structural Equation Modeling (SEM) is a multivariate technique combining aspects of factor analysis and multiple regression that enables the researcher to simultaneously examine a series of interrelated dependence relationships among the measured variables and latent constructs as well as between several latent constructs (Hair Jr. et al., 2006).
- Structural equation modeling is also known as 'causal modeling' or 'analysis of covariance structures'.

22.2 Discussion

Components

Measurement model: To assess the factor (latent variable – item) structure identified through factor analysis for summated scale development through Confirmatory Factor Analysis (CFA);

Structural model: To assess the causal relationships among latent and indicator variables as well as among latent variables – Structural/path modeling for testing hypotheses about relationships among constructs; While CFA operates with latent variables identified through factor analysis, structural modeling works only with observed/measured variables.

Example

If a researcher wishes to develop a scale to measure women's empowerment, he/she can identify its dimensions through summated scale development steps described in the module. After identifying various dimensions of empowerment – economic, socio-cultural, familial, and psychological –the researcher can employ these items for determining these constructs in SEM.

1. Development of measurement model which confirms the factor-item relationships identified through factor analysis;
2. Testing hypotheses related to women empowerment through structural model.

3. However, since this module deals with only summated scale development only measurement model development through confirmatory factor analysis is described.

Confirmatory factor analysis

- Confirmatory factor analysis (CFA) is a statistical technique used to verify the factor structure of a set of observed variables;
- Helps to identify items suitable for each factor and also to estimate construct validity of the scale;
- As a part of summated scale development, CFA is the 8th step (Fig. 21.1 in Module 21) which follows reliability and validity estimation;
- The purpose of CFA is twofold – (i) it confirms a hypothesized factor structure; and (ii) it is used as a validity procedure in the measurement model.

Difference between EFA and CFA

- In EFA the data determines the factor structure
- Statistical objective – Exact variance.
- In CFA a theoretical factor structure is specified and tested for its fit with the observed covariances among the items in the factors
- Statistical objective – reproduce covariance matrix.

Step 1: Testing Assumptions of CFA

- Detection and management of Univariate outliers – Described in Module 12;
- Checking univariate normality – Described in Module 13;
- Checking multivariate normality – Described in Module 13;
- Sample size.

CFA is sensitive to sample size. The minimum sample sizes estimated from the number of constructs and commonalities are given in Table 22.1.

Table 22.1: Minimum sample size requirements (Hair Jr. et al., 2006)

Number of constructs	No. of items/construct	Minimum sample size
Five or less	> 3	100
Seven or less	> 3	150
Large number of constructs (>7)	< 3	500

Step 2: Model specification – Creating a visual diagram of the measurement model

- This model is a hypothesized structure of latent (factor) and indicator (items) variables derived from EFA. The CFA helps us to test and validate this hypothesized structure;
- The SEM-CFA model can specify latent variables, indicators, directional (factor loadings of indicators on latent variables), and non-directional (correlation among latent factors);
- Various elements and types of relationships among factors and variables are explained in Table 22.2.

Directionality

- Model specification requires the researcher to describe the pattern of directional and non-

directional relationships among the variables;

- Classical test theory assumes that all models are reflective – All the directions of arrows should originate from latent variable and move towards the items. Directional effects are represented by regression coefficients. All directional effects (between latent variables, latent to indicator variables, residual in indicator variable) are considered as parameters. Non-directional arrows are specified between latent factors and are essentially covariances (correlation between latent factors) (Hair Jr. et al., 2006);
- In a reflective model, minimum of three items/indicators are required for each factor or latent variable. It indicates that the researcher needs at least three statements for each dimension of the construct or variable.

Model identification

- Model identification is an important aspect of CFA Modeling (Hair Jr. et al., 2006);
- For each free parameter, it is necessary that at least one algebraic solution is possible expressing that parameter as a function of the observed variances and covariances;
- If at least one solution is available, the parameter is identified. If not, the parameter is unidentified;
- To correct an under-identified model, the model must be changed or the parameter changed to a fixed value. To do this, the researcher adds a regression weight of one for every directional path of the first indicator variable of the latent factor. However, IBM-SPSS-AMOS creates this fixed value by default;
- To obtain accurate estimates, an over-identified model is required.

Other aspects

- When the model is specified, along with all elements, the researcher assumes that the model is true and expects a reasonable level of Goodness-of-fit with the hypothesized model;
- The example of computer use behaviour explained in Module 5 (Thematic analysis) is depicted as a visual diagram in Figure 22.1;
- Perceived usefulness (per_use), perceived ease of use (per_eou) and organisational support (OS) are depicted along with items measuring each latent construct, e-indicator measurement error;
- SEM CFA assumes a linear relationship among the indicators of measured variables. In case of non-linear relationships, the Kenny-Judd model (Kenny and Judd, 1984) can be used.

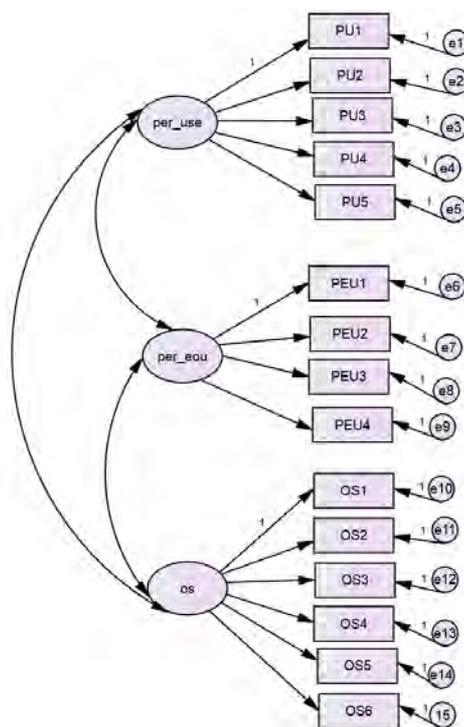
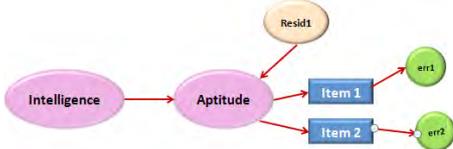
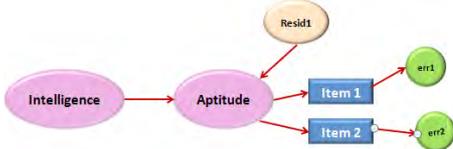


Fig. 22.1: Visual diagram of the Computer Use Behaviour constructs or latent factors and their items

Table 22.2: Elements and relationships of factors and variables in CFA (Hair Jr. et al., 2006)

Terminology	Alternate name	Meaning	Symbol
Latent variable	Factor, construct	Unobserved hypothetical variable	 
Indicator	Measured or manifest variable	Observed variable	 
Factor loading	Path loading	Correlation between latent variable and indicator	
Non-directional association	Covariance, correlation	Correlation between two latent variables	
Indicator error	Predictor error, measurement error	Error in indicator that is not accounted for by latent variable; Indicator error is also considered a latent variable (err1, err2)	
Latent variable error	Residual error	Error in prediction of endogenous variables from exogenous variable (Resid1)	
Explained variance		Percentage of variance in dependent latent variable accounted for by predictor(s)	
Independent variable	Exogenous variable, predictor	Cause fluctuations in another latent variable	
Dependent variable	Endogenous variable criterion	Variable that is predicted by other latent variables or indicators	
Set parameter	Constrained parameter; Fixed path	Parameter that is set at a constant and not estimated. Parameters fixed at 1.0 reflect an expected 1:1 association between variables. Parameters set at 0 reflect the assumption that no relationship exists.	

Step 3: Working with Software

Software available for analyzing SEM/CFA

- IBM-SPSS-AMOS (Analysis of Moment Structures)
- LISREL (Linear Structural Relationships)

- SAS – PROC CALIS Procedure (Covariance Analysis and Linear Structural Equations)
- EQS (Equations)

Among them, IBM SPSS AMOS is user-friendly and popular in SEM research. A step-by-step description for using SPSS AMOS for Confirmatory Factor Analysis is given in Annexure A-5.

22.3 Key Points

- Structural equation model is widely used in Summated Scale Development with multidimensional constructs.
- The factors extracted and their indicators from factor analysis are analysed through Confirmatory Factor Analysis to test their factor structure.
- The CFA provides coefficients and estimates for assessing non-normality, goodness-of-fit of model, and estimates of validity and reliability of the scale.

References

- Bagozzi, R. P. and Y. Yi. (1988.) On the evaluation of structural equation models. *Journal of the Academy of Marketing Science* 16 (1):74-94.
- Bentler, P.M. and Bonnet, D.C. (1980.) Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin* 88 (3):588-606.
- Bollen, K.A. (1989.) *Structural equations with latent variables*. New York: John Wiley and Sons.
- Fornell C.R. and D.F. Larcker. (1981.) Structural equation models with unobservable variables and measurement error. *Journal of Marketing Research* 18 (1):39-50.
- Hair, Jr., J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006.) *Multivariate data analysis*. Sixth edition. Upper Saddle River, NJ: Pearson Prentice Hall.
- Hu, L.T. and Bentler, P.M. (1999.) Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6 (1):1-55.
- Joreskog K., and Sorbom, D., (1996.) *LISREL 8: User's Reference Guide*. Chicago, IL: Scientific Software International, Inc. Pp. 271–274.
- Kelloway, K. E. (1998.) *Using LISREL for structural equation modeling: A researcher's guide*. Thousand Oaks: Sage.
- Kenny, D. A., & Judd, C. M. (1984.) Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin* 96:201-210.
- Mardia, K.V. (1970.) Measures of multivariate skewness and kurtosis with applications. *Biometrika* 57:519–530.
- McDonald, R.P. and Ho, M.H.R. (2002.) Principles and practice in reporting statistical equation analyses. *Psychological Methods* 7 (1):64-82.
- Miles, J. and Shevlin, M. (2007.) A time and a place for incremental fit indices. *Personality and Individual Differences* 42 (5):869-74.
- Steiger, J.H. (2007.) Understanding the limitations of global fit assessment in structural equation modeling. *Personality and Individual Differences* 42 (5):893-98.
- Tabachnick, B.G. and Fidell, L.S. (2007.) *Using multivariate statistics*. Fifth edition. New York: Allyn and Bacon.
- Wheaton, B., Muthen, B., Alwin, D., F., and Summers, G. (1977.) Assessing reliability and stability in panel models. *Sociological Methodology* 8 (1):84-136.



Photo Credit: Madhu S

23

Methods for Evaluation of Extension Programmes

P V K Sasidhar and B S Sontakki

Objectives

1. Understand the meaning of evaluation and differentiate between appraisal, monitoring, evaluation, and impact assessment
2. Discuss different types, tools, techniques, designs, approaches and challenges of evaluation

23.1 Introduction

The demand for evaluation of extension programmes/projects is growing as funding agencies and stakeholders want feedback from the programme/project planners on:

- What did you do with the money?
- Why should we continue to fund extension programmes/projects?
- Are the programmes effective?
- How will you improve or terminate ineffective programmes/projects?

Therefore evaluation of extension programmes/projects is essential in order to answer the above-mentioned questions along with corroborative evidence.

23.2 Discussion

23.2.1 What is evaluation?

Evaluation is a systematic collection and analysis of information about the characteristics and outcomes of a programme/project as a basis of judgment to improve its effectiveness and/or to inform decisions about current or future programming (USAID, 2011). Evaluation is one of the essential core competencies for all extension professionals in order to establish accountability of extension programmes/projects.

23.2.2 Steps in evaluation

Various steps in designing and executing an evaluation study are displayed in Box 23.1.

Box 23.1: Steps in programme evaluation

1. Writing an evaluation proposal with:
 - Need for evaluation – economic/technical/social reasons
 - Research/evaluation questions
 - Objectives
 - Evaluation framework (select a suitable framework)
 - Bennett's Hierarchy
 - Logical Framework Analysis etc
 - Budget
2. Designing and approval of an evaluation survey instrument
3. Coding, pre-testing, re-coding in statistical software and data collection
4. Data entry and analysis
5. Evaluation report writing and communicating findings – Short reports/Longer reports

Example: See the following links for an evaluation proposal and an evaluation report with survey instrument on 'Integrated contract broiler farming: An evaluation case study in India'

<http://www.meas-extension.org/meas-offers/case-studies/contract-farming>
<http://dev.meas.illinois.edu/wp-content/uploads/2015/04/MEAS-EVAL-2015-Broiler-India-long-Sasidhar-Suvedi-June-2015.pdf>

23.2.3 Appraisal vs. Monitoring vs. Evaluation vs. Impact Assessment

Though they are often interchangeably used, following distinctions exist between them:

- **Appraisal:** It is a critical examination of a programme/project proposal, normally before implementation and funding with respect to economic viability, technical feasibility, and/social desirability. It is basically a planning and project formulation activity guided by evaluation findings of similar programmes/projects that have already been implemented.
- **Monitoring:** It is a continuous process that starts and ends with a programme/project which is required for immediate use and mid-course correction. It is usually done by implementing personnel who cover all aspects of the programme/project for its correction or management. Monitoring is a symptomatic and early warning system.
- **Evaluation:** It is an in-depth one-shot operation at a point of time (usually at completion or mid-way in the programme) undertaken for future planning/replication/expansion. It is a learning and diagnostic process usually done by an outside agency covering a sample.
- **Impact Assessment:** Building on appraisal, monitoring and evaluation, the focus of an impact assessment is on longer-term and wider-ranging changes, beyond the immediate results of the programme or project.

23.2.4 What are we evaluating?

In general, evaluation is conducted to assess the progress, outputs, outcomes and impact of an extension programme. A few commonly used analytical aspects described by Dale (2004) are given below:

- i. **Relevance:** To what extent has the programme or project addressed, or is addressing, problems of high priority, as viewed by actual and potential stakeholders, especially beneficiaries?

- ii. **Effectiveness:** To what extent are the planned outputs, expected changes, intended effects (immediate and effect objectives) and intended impact (development objective) being, or have been, produced or achieved?
- iii. **Impact:** The overall consequences of the programme or project for the intended beneficiaries and any other people;
- iv. **Efficiency:** This means the amount of outputs created and their quality in relation to the resources (capital and human efforts) invested;
- v. **Sustainability:** This means the maintenance or augmentation of positive achievements induced by the evaluated programme or project (or any component);
- vi. **Replicability:** This means the feasibility of repeating the particular programme or project, or parts of it, in another context, i.e., at a later time, in other areas, for other groups of people, by other organisations, etc.

23.2.5 Types of Evaluation

Evaluations are of the following types based on the programme phase(s) (planning, implementation, and conclusion) at which evaluation is conducted:

- i. **Baseline Evaluation:** Needs assessment is a form of baseline evaluation to find out the target group and their perceived needs/expectations from the programme. If collected, this will also establish baseline data to compare programme results later. Example: Scope for animal welfare education in open and distance learning: findings from a needs assessment study in India (Sasidhar and Jayasimha, 2015).
- ii. **Formative Evaluation:** It is undertaken during the programme implementation stage to determine whether the programme is going as per plan, and changes, if any, are required to meet the objectives. It is also termed as process/mid-term/concurrent evaluation. Generally, formative evaluations are undertaken in long-term programmes/projects for cross checking/corrective measures. Example: Formative evaluation of Kisan Call Centres in Tamil Nadu (Karthikeyan et al., 2006).
- iii. **Summative Evaluation:** It is undertaken once the programme achieves a stable state of operation or towards the end of a programme to find out its results, effectiveness, impact and further course of action. Summative evaluation findings help extension functionaries/funding agencies to make decisions on programme continuation, modifications, further expansion, reduction, or closure. If summative evaluation is done immediately after completion of a project it is called terminal/outcome evaluation. If it is done some time after completion of a project, it is called ex-post evaluation. Example: Integrated contract broiler farming: An evaluation case study in India (Sasidhar and Suvedi, 2015).
- iv. **Follow-up Evaluation:** It is undertaken long after completion of the programme to see whether there are any long-term changes among beneficiaries. When follow-up evaluations are repeated at set time intervals to study the long term benefits, sustainability of results and outcomes, they are called longitudinal evaluations. Example: Evaluation of a distance education radio farm school programme in India: Implications for scaling up (Sasidhar et al., 2011).

Various evaluation techniques used at different programme stages are displayed in Box 23.2.

Box 23.2: Evaluation tools and techniques by programme stage

Programme Stage	Evaluation type	Evaluation questions	Evaluation tools and techniques
Planning stage	Needs assessment	<ul style="list-style-type: none"> What are the felt and unfelt needs of beneficiaries? 	Survey Focus Group Discussion
	Feasibility study	<ul style="list-style-type: none"> Do they fit into the extension programme mission? 	Observation Content analysis of records
	Baseline study	<ul style="list-style-type: none"> Can the extension programme address these needs? Is the extension programme feasible (socio-economic & environmental)? 	Economic analysis B:C ratio
Implementation stage	Formative evaluation	<ul style="list-style-type: none"> Is the programme meeting its objectives or intended outcomes? Are beneficiaries satisfied with the programme? 	Annual monitoring reports Technology adoption patterns Knowledge, Attitude, Skill and Aspirations (KASA) change Satisfaction surveys
Conclusion stage	Summative evaluation	<ul style="list-style-type: none"> Are needs addressed? Have the desired outcomes been achieved? What is the cost effectiveness of the programme? 	Pre- & post-programme data Cohort studies Panel studies Economic analysis

23.2.6 Evaluation Designs

- i. **Pre vs. Post-programme Evaluation:** It is undertaken through comparison of present situation against situation prior to the programme. It suffers from 'memory or recall bias' of the respondents. To overcome this, it is always advisable to have a 'baseline' or 'benchmark' study and a 'post-programme' study. Unfortunately, in most programmes baseline studies are not undertaken or sometimes the available data is inadequate.
- ii. **With-Without Programme Evaluation:** A comparable sample of beneficiaries and non-beneficiaries are to be selected within the programme area. Alternately, a comparable adjoining area where programme has not been implemented can also be studied along with the programme area. Example: Comparative evaluation of KVK (with beneficiaries and non-beneficiaries).

A combination of the above two designs is most appropriate. Various evaluation designs used in impact assessment are described in Module 24.

23.2.7 Evaluation Data Collection Methods

The different methods to collect evaluation data (Dale, 2004; Bagnol, 2007 and 2014; Suvedi and Kaplowitz, 2016) are given below:

(a) Conventional Methods

(i) Document examination

- In general, evaluators may draw information from existing documents of various kinds/secondary sources, such as plan documents, monitoring reports, progress reports, etc.;
- Used both at the beginning and at the end of the programme.

(ii) Evaluator's observation and measurement

(a) Informal and formal observation

- Examining physical facilities, work done by some people, and organisational performance;
- Used as a supplementary method to other methods.

(b) Direct measurement

- Measuring physical changes and changes in livelihoods and socio-economic structures;
- Uses various quantitative, qualitative and mixed methods for collecting and interpreting data.

Quantitative Methods

- Measures a limited number of predetermined outcomes and are appropriate for judging effects, attributing cause, comparing or ranking, classifying and generalizing results;
- Example: Structured interview, questionnaire survey, standardized key informant interviewing.

Qualitative Methods

- Through descriptions, they help in understanding programme context and problems, clarifying relationship between programme objectives and implementation, identifying unintended programme consequences and conducting in-depth impact analysis;
- Focus group discussion, participant observation, in-depth case study, oral histories, community stories.

Mixed Methods

- Combine qualitative and quantitative methods to complement each other;
- Appropriate for understanding complex social phenomena allowing plural viewpoints in a programme.

(c) Participatory and collective analysis

- They involve stakeholders in programme evaluation through partnerships and dialogues, taking into account the aspects that are often left out by evaluators and can only be identified by stakeholders themselves.

A few basic methods are given in Box 23.3, and advanced methods in Box 23.4.

Box 23.3: Basic participatory and collective evaluation methods

1. **Meetings**– Arranging and conducting meetings with stakeholders, like project beneficiaries, development agents, etc., and proceedings are used as evaluation reports. The main disadvantage is that it involves only little interaction and in-depth assessment is not possible;
2. **Informal Group Discussion**– the evaluator comes into an unplanned group setting in which a relevant discussion takes place, coordinated by the evaluator;
3. **Facilitated Group Discussion**– Group discussions may also be planned and arranged, and will normally be moderated by the evaluator. It is highly interactive and is an in-depth discussion on the effect of the extension programme;
4. **Workshop-based Participatory Analysis** – The beneficiaries and development agencies

Box 23.3: Basic participatory and collective evaluation methods

jointly explore problems, plan or evaluate, in a workshop setting. Participatory rural appraisal (PRA), Participatory learning methods (PLM), Participatory assessment monitoring and evaluation (PAME) are a few examples of this approach;

- 5. Collective Brainstorming**– Intensive and open-minded communication event that a group of persons agree to embark on in a specific situation. Useful method for analysing problems, which occur suddenly and require an immediate solution.

Source: Dale (2004)

Box 23.4: Participatory evaluation of disease control through vaccination campaigns

The KYEEMA Foundation implemented the 'Regional Newcastle disease control project' in Malawi, Mozambique, Tanzania and Zambia with the support of AusAID. In January 2012, a participatory evaluation was carried out in three villages of Malawi to evaluate the impact of vaccination campaigns. All the male and female farmers interviewed had chickens vaccinated three times by community vaccinators in March, July and November 2011. The first question asked was: Since the first vaccination did the number of chickens in the flock increase, stay the same, or decrease? Each participant was asked to respond by placing a stone on one of the three possible answers written on a flip chart on the ground. The answers were used to generate the table below.

Response	Village 1		Village 2		Village 3	
	12 Women	10 Men	7 Women	8 Men	16 Women	7 Men
Increased	11 (92%)	8 (80%)	7 (100%)	6 (75%)	12 (75%)	7(100%)
Stayed the same	1 (8%)	0	0	0	2(12.5%)	0
Decreased	0	2(20%)	0	2 (25%)	2 (12.5%)	0
Total	12 (100%)	10 (100%)	7 (100%)	8 (100%)	16 (100%)	7(100%)

The same exercise also included the following questions:

- Since the 1st vaccination did the number of birds that died increase, stay the same, or decrease?
- Since the 1st vaccination did the number of chickens sold increase, stay the same, or decrease?
- Since the 1st vaccination did the number of chickens consumed increase, stay the same, or decrease?

To evaluate the increase in size of household flocks since vaccination in 2011, the participants were asked to state the number of chickens they had in January 2010, and later the number of chickens they had in January 2012. By analysing the median and the average or calculating the average percentage increase per household, it is possible to see the evolution in flock size. Similarly, it is possible to evaluate the number of chickens sold and consumed.

Source: Bagnol (2014)

23.2.8 Evaluation Approaches

Though several approaches are available, Logical Framework Approach (Box 23.5) and Bennett's Hierarchy Approach (Bennett, 1979) (Box 23.6) provides a comprehensive framework to evaluate inputs, outputs, and outcomes of an extension programme. Both these approaches provide a process by examining the chain of means (what actions we take in an extension programme) and ends (the result of these actions) through different levels.

Box 23.5: Logical Framework Analysis evaluation of Kisan Call Centers in Tamil Nadu

	Inputs	Output		Outcome
		Activities	Participants	
Situation Mission Objectives Assumptions Environmental factors	1. Level I, II and III centre staff	1. Instant reply to farmers at level I		Short term
	2. Telephone, ICT, and agricultural technology	2. Escalating the unanswered calls to level II and III from level I office	Farmers	
	3. Dedicated line with a toll free number identified by MoA, Gol	3. Documentation and reporting to DAC by level III centre		
	4. IVRS	4. Reply to farmers within 72 hours	Scientists	Medium term
	5. Fund by DAC	through post/e-mail through off-line mode by level III centre		
	6. Knowledge Management System	5. M&E by level III		Long term
	7. Service timings: 6:00 AM to 10:00 PM	6. Database using KMS	Level I officers	
	8. Documentation of agricultural technologies and FAQs	7. Training to level I officials		
	9. Capacity building	8. Documenting FAQs		
		9. Mass media advertisements		
Outcomes				
	Short Term	Medium Term	Long Term	
	1. Awareness about KCC	1. Adoption of the advice recommended by KCC	1. Overall development of agriculture through ToT	
	2. Participation in KCC	2. Gratification with the services of KCC	2. Reduction of gap between research extension and farmer linkage	
		3. Information sharing about call centre number (1800-180-1551) and advice given by KCC		

Box 23.6: Bennett's hierarchy applied in evaluation of contract broiler farming (CBF) and non- contract broiler farming (NCBF) systems

Evaluation hierarchy	Measurement	Indicators
Level 7: End results	Socio-economic changes and impacts	<ul style="list-style-type: none"> • SWOT parameters • FGD on: selection of contract farmers; terms and conditions applicable in CBF
Level 6: Practice change	Technical advice adoption	<ul style="list-style-type: none"> • Non-adoption, discontinuation, partial adoption and full adoption of technical advices
Level 5: KASA	Farmers' perceptions	<ul style="list-style-type: none"> • Perceptions on inputs (chicks, feed, medicines and EAS) and outputs (broiler birds, manure value and payment system)
Level 4: Reactions	Farmers' feedback	<ul style="list-style-type: none"> • Factors of motivation to do CBF and NCBF • Reasons to change integrator(s) or input providers in the past two years
Level 3: Outputs	Technical and economic performance Technical and economic performance	<ul style="list-style-type: none"> • Broiler birds (flock size, mortality number, birds sold, sale age, sales rate and bird lifting days) • Productivity (mortality percentage, birds sold, feed consumption and body weight) • Efficiency (FCR, sale age, weight gain/day) • Economics of inputs and outputs • EAS (frequency of information from various sources)
Level 2: Activities	Activities in CBF and NCBF	<ul style="list-style-type: none"> • Physical and human resource activities in CBF and NCBF
Level 1: Inputs	Investments and demographics	<ul style="list-style-type: none"> • Fixed and variable costs • Age, gender, education, social category, family and size, poultry occupation and experience

Source: Sasidhar and Suvedi (2015)

23.2.9 Challenges in Extension Programme Evaluation

Results of the extension programme, especially KASA changes and long term outcomes are often intangible to measure. Suvedi and Vander Stoep (2014) outlined three major challenges in evaluating extension programmes:

- Frequent use of descriptive, one-shot case study for impact evaluation and lack of precise and straight assessments of cause-and-effect outcomes of programme. Extension needs to establish a scientific knowledge base and generate evidence on its work. Example: Do farmers who attend extension trainings adopt new technologies earlier than those who do not?
- It is difficult to establish a control or comparison group, maintain uniform treatments and measure long-term impacts of the extension programme.
- Non-availability/inadequate baseline and monitoring data on impact indicators to see expanded impacts over time.

23.3 Key Points

- Evaluation has been neglected and mistakenly seen as a fault-finding mechanism.
- It is to be understood that evaluation is a learning process and our past experiences guide us to a better future.
- Every extension professional should acquire evaluation as a core competency to assess the impact of their programmes.
- Choosing evaluation designs, methods and tools, suitable for each programme stage will help to conduct the process smoothly by creating sound data.

References

- Bagnol, B. (2007.) Participatory tools for assessment and monitoring of poultry raising activities and animal disease control. HPAI Communication Workshop, Bangkok: FAO (available at: www.participatoryepidemiology.info/userfiles/Participatory%20Tools_9_03_08.pdf).
- Bagnol, B. (2014.) Conducting participatory monitoring and evaluation. Pages 81-85 in FAO, Decision tools for family poultry development. FAO Animal Production and Health Guidelines, No. 16. Rome, Italy: FAO (available at: <http://www.fao.org/docrep/019/i3542e/i3542e.pdf>)
- Bennett, C.F. (1979.) Analyzing impacts of extension programs. Washington, D.C., USA: U.S. Department of Agriculture.
- Dale R, (2004.) Evaluating Development Programmes and Projects. New Delhi, India: Sage Publications
- Karthikeyan, C., Vijayaraghavan, K. and Lavanya, P. (2007.) Formative evaluation of Kisan Call Centres. Tamil Nadu. Indian Journal of Extension Education, 43(1&2):20-25.
- Sasidhar, P. V. K., Suvedi, M., Vijayaraghavan, K., Singh, B., and Babu, S. (2011.) Evaluation of a distance education radio farm school programme in India: implications for scaling up. Outlook on Agriculture 40(1):89-96. DOI: 10.5367/oa.2011.0033
- Sasidhar, P.V.K. and Jayasimha, N.G. (2015.) Scope for animal welfare education in open and distance learning: findings from a needs assessment study in India. Scientific and Technical Review of OIE, 34(3):689-97 (available at: <http://web.oie.int/boutique/extrait/02sasidhar689697.pdf>)
- Sasidhar, P.V.K. and Suvedi, M. (2015.) Integrated contract broiler farming: An evaluation case study in India. Urbana, IL: USAID-MEAS. www.meas.illinois.edu (available at: <http://dev.meas.illinois.edu/wp-content/uploads/2015/04/MEAS-EVAL-2015-Broiler-India-long-Sasidhar-Suvedi-June-2015.pdf>)
- Suvedi, M., and Vander Stoep, G. (2014.) Improving the monitoring and evaluation of agricultural extension services. Discussion paper. East Lansing, Michigan, USA: Department of Community Sustainability, Michigan State University.
- Suvedi, M. and Kaplowitz, M.D. (2016.) Process skills and competency tools – what every extension worker should know – Core Competency Handbook. Urbana, IL: USAID-MEAS (available at: <http://www.meas-extension.org/meas-offers/training>).
- USAID. (2011.) Evaluation policy. Washington, D.C., USA: Bureau for Policy and Planning.



Photo Credit: Nimisha Mittal

24

Methods of Analysing Impact of Agricultural Technologies

Sivaramane N

Objectives

1. Introduce advanced methods of assessing impact of agricultural technologies
2. Explain the quantitative methods applied in assessing impact
3. Suggest future work for improving impact assessment

24.1 Introduction

- Assessing effects of an intervention on the user group is a crucial component for assessing the utility of a technological intervention.
- According to the World Bank, impact assessment is intended to determine more broadly whether the program had the desired effects (both positive and negative) on individuals, households, and institutions, and whether those effects are attributable to the program intervention (Baker, 2000).
- Impact is measured in terms of positive and negative, primary and secondary long-term effects produced by an intervention, directly or indirectly, intended or unintended.

24.2 Discussion

24.2.1 Importance

In National Agricultural Research and Education System (NARES), the impact assessment is used

- For assessing the socio-economic effects of an extension, technology or communication intervention;
- In the identification of problems associated with technologies and processes in the user system;
- For developing research and development priorities;
- For improving accountability for resources and efforts invested in the intervention.

24.2.2 Impact assessment framework

Impact assessment (IA) is an integral part of programme planning, implementation and evaluation. IA, along with processes like monitoring and evaluation, are employed to ensure progress of the work according to the programme objectives and to verify if results obtained are as desired by the implementing agency. IA measures the achievement of project milestones, outputs, outcomes and their impact on the targeted population. This process is pictorially depicted in Fig. 24.1.

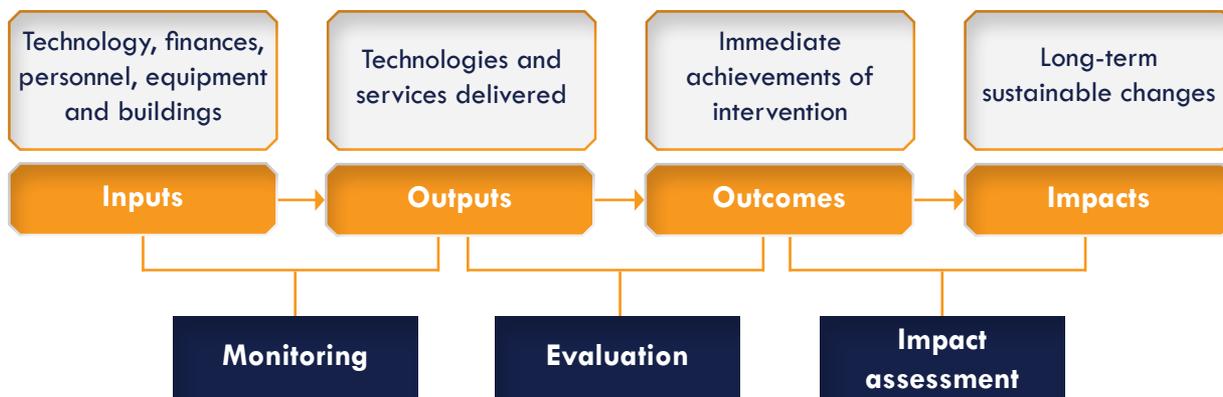


Fig. 24.1: Impact assessment framework (Adapted from Devos Vaughan and Associates - <http://devosvaughan.com/services-programme-review.html>)

Definitions of various elements in the impact assessment framework are displayed in Box 24.1.

Box 24.1: Definitions of elements in the Impact Assessment Framework

Inputs – All the resources that contribute to the production and delivery of outputs. Inputs include finances, personnel, equipment and buildings;

Outputs – The final products, or goods and services produced for delivery;

Outcomes – The medium-term results for specific beneficiaries that are a consequence of achieving specific outputs. Outcomes should relate clearly to an institution’s strategic goals and objectives set out in its plans. Outcomes are ‘what we wish to achieve’;

Impacts – The long-term results of achieving specific outcomes, such as reducing poverty and creating jobs.

The quality of impact assessment depends on the soundness of the programme implementation process. For effective impact evaluation, the questions below need to be answered at every stage of technology or extension intervention.

(i) Planning

WHAT	realistic objectives and specific outcomes?
WHO	target/beneficiary group?
WHEN	duration?
WHY	assumptions?
SO WHAT	continuity and sustainability?

(ii) Implementation

HOW	methods? processes?
-----	---------------------

(iii) Evaluation

WHAT/WHY	type of assessments?
HOW/WHEN	methods in data collection?
WHO	community? stakeholders?
SO WHAT	continuity and sustainability?

(iv) Report

WHERE/WHEN/HOW	publishable platforms, e.g. social media, traditional media, print?
WHO	funding bodies and stakeholders? feedback to community?
WHY	benefits?

24.2.3 Indicators for impact assessment

- Indicators are common tools to assess the performance and then the impact of the technology/extension/communication interventions. They measure the accomplishment of the project's goals and targets;
- Indicators should be defined during the formulation stage, but they often need to be specified in greater detail during implementation;
- Indicators should measure intangible as well as tangible changes, particularly in projects that value factors such as personal and social development;
- The procedure for selecting good impact indicators is described in Box 24.2.

Box 24.2: Procedure for selecting impact indicators

1. Decide the level of impact assessment (output/outcome/impact);
2. Identify a limited number of relevant indicators that adequately measure the selected impact level;
3. Select the data sources available and the type of data collection needed for each indicator;
4. Construct a matrix listing the indicators, identifying their importance for selected impact level (high/low), ease of obtaining data on the indicator (easy/feasible but effort/difficult levels) and the cost of data collection;
5. Prioritize indicators by importance, the ease of obtaining data, and cost and select a set of indicators using weighted matrix ranking/rating method;
6. Group selected indicators by source of data to determine the set of sources, which can provide data on clusters of indicators;
7. Make a final selection of a group of indicators and decide on a data collection plan in light of available logistical, human and financial resources, and time.

(Adapted from Badioli, 2011)

Indicators for assessing impacts of technology or extension interventions

What should be measured

- Technical aspects: physical input-output of goods and services;
- Institutional aspects: organizational and managerial aspects, including customs, tenure, local organizations, and cultural setting;
- Socio-cultural aspects: broader social implications, resource and income distribution, and employment opportunities;
- Commercial aspects: business and financial, securing supplies, and market demand;
- Economic aspects: economic efficiency, costs and benefits;
- Environmental aspects: biological and physical effects.

Social and behavioural indicators

- Changes in awareness, knowledge, skills level of beneficiaries;

- Increases in the number of people reached;
- Policy changes;
- Changes in behaviour, e.g., adoption;
- Changes in community capacity;
- Changes in organisational capacity (skills, structures, resources);
- Increases in service usage.

Socio-cultural indicators

- Food security;
- Poverty reduction;
- Status of women improved;
- Distribution of benefits across gender and geographical locations;
- Changes in resource allocation;
- Changes in cash requirement;
- Changes in labour distribution;
- Nutritional implications.

Technology level indicators

- Adoption of improved technology – symbolic and actual adoption;
- Horizontal impact – increase in area under improved variety/breed;
- Vertical impact – increase in productivity of improved variety/breed;
- Reduction in cost of production;
- Risk reduction;
- Increase in annual income/economic capacity;
- Jobs created.

Environmental impact measures

- Effect on soil erosion and degradation, silting, compact soil, soil contamination, water contamination;
- Changes in hydrological regimes;
- Effects on biodiversity, air pollution, greenhouse gases.

Institutional impact measures

- Changes in organizational structure;
- Change in the number of scientists;
- Change in composition of the research team;
- Multidisciplinary approaches and improvements;
- Changes in funding allocated to the program;
- Changes in public and private sector participation;
- New techniques or methods.

24.2.4 Approaches for impact assessment

Broadly, impact assessment approaches can be classified into quantitative, qualitative, participatory, and mixed-methods. Quantitative methods focus on generating quantitative data on the impact indicators, while qualitative methods are concerned with impacts within local socio-cultural and institutional context. Participatory methods include an exploratory assessment of impacts wherein users and mixed-methods combine both qualitative and quantitative approaches so as to provide a comprehensive view of the impacts within the socio-cultural milieu.

Quantitative impact assessment approaches

- Focus is on assessing the degree and extent of the impacts quantitatively. Some degree of quantification may be necessary in all impact assessments, in order to evaluate the success of the intervention and the extent of any adverse effects;
- Largely depends on micro-economic approaches following econometric models;
- Involving baseline studies – The precise identification of baseline conditions, definition of objectives, target setting, rigorous performance evaluation and outcome measurement;
- Costly and limited scope – Limited in the types of impacts which can be accurately measured and may pose difficulties for inference of cause and effect.

Types

a. Based on time period of assessment

Impact assessment may take place before approval of an intervention (ex ante), after completion (ex post), or at any stage in between.

Ex ante assessment

- Forecasts potential impacts as part of the planning, design and approval of an intervention;
- Example – Assessment of adoption of Genetically Modified Rice variety.

Ex post assessment

- Identifies actual impacts during and after implementation, to enable corrective action to be taken if necessary, and to provide information for improving the design of future interventions;
- Example – Knowledge gain of trainees from a multimedia CD on SRI method of cultivating rice.

b. Based on research design

Impact assessment may be conducted using various designs – Experimental, quasi-experimental and non-experimental designs.

1. Experimental approaches

(i) Randomized evaluations

- In randomized evaluations, the programme benefits are extended to a randomly selected treatment group (beneficiaries), while keeping an identical group as control;
- The progress of the treatment and control groups on selected impact indicators is tracked over time (Khandker et al., 2010).
- Randomised evaluations are used when
 - The eligible population is large enough to deliver the programme;
 - A program needs to be gradually phased in until it covers the entire eligible population.
- The major advantage of this approach is its ability to avoid bias in selecting respondents.

(ii) Pre-test/Post-test with random assignment to intervention or comparison groups

- In these randomized experiments, study subjects are randomly assigned to a group that receives the technological intervention (study or treatment group) or a comparison group that does not receive

- the intervention (control or non-treatment group);
- Data for each group are collected before and after the intervention;
- At the end of the experiment, differences between the intervention and comparison groups can be attributed directly to the effect of the intervention, if the sample is large enough;
- Used in small samples consisting of less than 30 persons per group;
- Statistical methods are selected based on the research question
 - o Whether the mean change in the outcome from before and after the treatment differed between the two groups – Repeated measures ANOVA;
 - o Whether the post-test means, adjusted for pre-test scores, differ between the two groups – Analysis of Covariance (ANCOVA).

(iii) Post-test only randomised experiment

- Two groups are randomly assigned the subjects and treatment conditions;
- Data collected only after the intervention;
- Statistical tests – Regression, t test, ANOVA.

2. Quasi-experimental approaches

(i) Pre-test/Post-test with non-random assignment to intervention or control groups

- In this design, data are collected before and after the intervention;
- Assigning subjects to the intervention and comparison groups is non-random;
- Comparison groups in the quasi-experimental design can be identified through propensity score matching;
- Propensity score matching – The control population (non-beneficiaries) is selected by ‘matching’ them with the actual beneficiaries on a few observable characteristics (personal and socio-economic attributes);
- The matched control groups can be selected either before project implementation (prospective studies) or afterwards (retrospective studies);
- Approaches for matching the beneficiaries and non-beneficiaries – nearest-neighbour (NN) matching, caliper and radius matching, stratification and interval matching, kernel matching, and local linear matching (LLM) (Khandker et al., 2010);
- Statistical analysis – Reliability-corrected ANCOVA.

(ii) Two group post-test only with non-random assignment

- Data are collected only after the program has ended among participants who had received the intervention and among non-participants;
- Matching participants and non-participants with similar characteristics and accounting for any relevant differences are especially important in the post-test only design to isolate effects of the intervention.

(iii) Double difference or difference-in-differences (DID) methods

- Estimates the effect of a specific intervention by comparing the changes in outcomes over time between a population that is enrolled in a program (the intervention group) and a population that is not (the control group);
- Can be applied in both experimental and quasi-experimental designs and requires baseline and follow-up data from the same treatment and control group;
- A baseline survey is conducted prior to the intervention to assess the outcome indicators with both beneficiaries (treatment) and non-beneficiaries (control or comparison), which are compared to

- estimate the differences (Khandker et al., 2010);
- After the intervention, the survey is repeated to assess the differences in treatment and control groups;
- The mean difference between the 'after' and 'before' values of the outcome indicators for each of the treatment and comparison groups is calculated followed by the difference between these two mean differences. The second difference i.e., difference in the difference is the estimate of the impact of the program;
- Statistical methods – Repeated measures ANOVA, Repeated measures linear regression analysis, propensity score matching;
- Useful technique to use when randomization on the individual level is not possible.

(iv) Regression discontinuity design

- A pre-test – post-test comparison method design that elicits the causal effects of interventions by assigning a cut-off or threshold above or below which an intervention is assigned;
- In this design, the participants are assigned to intervention or comparison groups solely on the basis of a cut-off score on a pre-intervention measure.
- The average treatment effect is estimated by comparing observations lying closely on either side of the threshold;
- Used in those conditions in which randomization is unfeasible and the researcher is interested in targeting an intervention or treatment to those who most need or deserve it;
- Two types of regression discontinuity designs:
 - Sharp method which assigns a discrete cut-off point for both beneficiaries and non-beneficiaries and compares means of treatment effects to assess the impact;
 - Fuzzy method was used in instances where a few eligible people are excluded from the intervention or became ineligible due to other reasons. The fuzzy version is the widely used method.

3. Non-experimental designs (MLE, 2013)

- These designs have only an intervention group without any control;
- Weakest designs for impact assessment;
- Used under (i) limited resource condition, (ii) researchers are unable to create a comparison group, (iii) when an intervention covers the entire population.

(i) Pre-test/post-test designs

- The researcher measures pre- and post-intervention changes in the specific phenomenon (outcome indicator) between subjects;
- When changes occur in outcome indicators among the intervention participants, they cannot attribute all these changes to the intervention using this design alone because there is no comparison group.

(ii) Time-series designs

- The changes in outcome indicator over time is estimated to determine trends;
- The data is collected multiple times before and after the intervention to analyse trends before and after.

(iii) Longitudinal study

- The researcher records repeated measures of the same variables from the same people;
- A panel design is a special type of longitudinal design in which evaluators track a smaller group of

people at multiple points in time and record their experiences in great detail.

(iv) Post-test only design

- Researchers observe the intervention group at one point in time after the intervention, focusing particularly on comparing responses of sub-groups based on such characteristics as age, sex, ethnicity, education or level of exposure to the intervention.

Methods to strengthen non-experimental impact assessment

(a) Measure participants' level of exposure to the program: Measuring the participants' initial level of exposure to the intervention aspect will help to offset selection bias;

(b) Collect data from the same participants over time using a panel or longitudinal design: Individuals serve as their own control;

(c) Instrumental variable (IV) methods.

- A statistical estimation method to be used with non-experimental design of impact analysis;
- Unobservable biases in sample selection are minimised by including a new variable in the analysis;
- The new variable is called as 'instrumental variable' since it increases the probability that a person be selected as beneficiary in the study, but is demonstrated to have no influence on the outcome of the intervention;
- For example, 'geographical variation' may be included as an instrumental variable when assessing the impact of poverty alleviation programmes.

Economic impact assessment

A. Partial Budgeting Technique (PBT): This is another useful technique for studying the economic impact of a small scale intervention or a single technology adoption. For example, intervention in terms of drip irrigation in tomato cultivation can be studied using this approach. This technique has four components:

1. Increase in income;
2. Reduction or elimination of costs;
3. Increase in costs;
4. Reduction or elimination of income.

$$\text{Profit/Loss} = [\text{Added returns} - \text{Added costs}] + [\text{Reduced costs} - \text{Reduced returns}]$$

Features of this method

- The technique is simple and easy to learn;
- It examines only net changes in costs and benefits, therefore it is effective for assessing economic viability of single intervention technologies;
- It requires less data than whole farm budgeting since fixed costs are not examined;
- It allows early conclusions about the adaptability of the new technology.

Partial Budget for Drip Irrigation intervention where flood irrigation was practiced:

A. Added Income:

Additional income – 25000

(Difference in revenue earned)

B. Added cost:

Labour cost – 700

Drip system cost – 40000

(Seasonally apportioned)

Total added cost – 20700

D. Reduced cost:

Labour cost – 2500

Water cost (imputed) – 2000

Total reduced cost – 4500

C. Reduced income: Nil

A-B-C+D = 25000+4500-20700-0 = 8800

B. Net Present Value (NPV)

The NPV of an investment is a simple criterion for deciding whether or not to undertake an investment. NPV answers the question of how much cash an investor would need to have today as a substitute for making the investment. If the net present value is positive, the investment is worth taking on because doing so is essentially the same as receiving a cash payment equal to the net present value. If the net present value is negative, making the investment today is equivalent to giving up some cash today and the investment should be rejected. If the projected return on an investment is identical to the selected discount rate, the NPV=0 and the investor is indifferent with respect to making the investment. NPV is the present value of all benefit discounted at the appropriate discount rate, minus the present value of all costs discounted at the same rate. Symbolically it can be mentioned that:

$$NPV = \sum_0^t \frac{(B_t - C_t)}{(1 + r)^t}$$

Where: B_t is the benefit at time t
 C_t is the cost at time t
r is the discount rate
t refers to the time period

C. Benefit-Cost Ratio (BCR)

The benefit-cost ratio takes the times series data on benefits and costs used to construct NPV and organizes them in a ratio form rather than as an absolute value. Alternatively, the BCR can be defined as the ratio of the discounted benefits to the discounted costs of an investment with reference to the same point in time.

$$BCR = \frac{\sum B_t(1 + r)^t}{\sum C_t(1 + r)^t}$$

D. Internal Rate of Return (IRR)

The IRR reveals the rate of growth of capital invested in the business. For purposes of analysing the economic impacts of R&D, IRR is called as private rate of return (PRR) when the return to a single company’s (the innovator’s) R&D investment is being studied, or the social rate of return (SRR) when industry-wide or economy-wide rates of return are estimated. IRR is the rate at which NPV=0 and it can be easily computed using GOAL SEEK option under WHAT-IF ANALYSIS of MS EXCEL. Otherwise, the following formula may be used

$$IRR = r_1 + \frac{NPV_1}{NPV_1 - NPV_2} (r_2 - r_1)$$

r_1 = lower discount rate chosen
 r_2 = higher discount rate chosen

NPV1 = NPV at r_1
NPV2 = NPV at r_2

Cost concepts

To measure the profitability of new technology, the following cost concepts can be used

- Cost A1 = All actual expenses in cash and kind incurred in production;
- Cost A2 = Cost A1 + Rent paid for leased in land;
- Cost B1 = Cost A1 + Interest on value of owned capital assets;
- Cost B2 = Cost B1 + Rental value of owned land and rent paid for leased in land;
- Cost C1 = Cost B1 + Imputed value of family labour;
- Cost C2 = Cost B2 + Imputed value of family labour;
- Cost C3 = Cost C2 + 10% of Cost C2 on account of managerial functions performed by the farmer.

Other important tools

- Adoption rate/index/quotient
- Factors influencing the adoption of a technology (logit/probit)
- Consumer surplus model
- Dummy variable regression models
- Structural change in time series models
- Input-output model/SAM
- Competitiveness index (for exports)

Participatory approaches

- Participatory Impact Assessment (PIA) is an extension of Participatory Rural Appraisal (PRA) and involves the adaptation of participatory tools combined with more conventional statistical approaches specifically to measure the impact of technology and extension interventions on people's lives;
- Consists of a flexible methodology that can be adapted to local conditions;
- Acknowledges local people or project clients as experts by emphasizing the involvement of project participants and community members in assessing project impact;
- Participatory impact assessment answers the following questions:
 - What changes have there been in the community since the start of the project?
 - Which of these changes are attributable to the project?
 - What difference have these changes made to people's lives?
- Most extension research employs a qualitative participatory approach to assess the 'perceptual impacts' using participatory methods, such as Participatory Rural Appraisal, Focus groups, case studies, participant observation, etc.;
- This approach plays an important role in impact evaluation by providing information useful to understand the processes behind observed results and assess changes in people's perceptions of their well-being;
- Participatory methods can be used to improve the quality of survey-based quantitative evaluations by helping generate evaluation hypothesis and strengthening the design of survey questionnaires and expanding or clarifying quantitative evaluation findings (Khandker et al., 2010);
- A recent approach for participatory impact assessment is MAPP, which is described in Box 24.3.

Box 24.3: Method for Impact Assessment of Poverty Alleviation Projects, 1998

An innovative impact assessment approach developed by Ms. Susanne Neubert of German Development Institute, Germany, called MAPP (Method for Impact Assessment of Poverty

Box 24.3: Method for Impact Assessment of Poverty Alleviation Projects, 1998

Alleviation Projects, 1998), combines a quantitative approach with participatory assessment to derive tangible results in order to address the needs of managers and policy makers. In this method, impact is assessed through a series of workshops with stakeholder representatives. It has wide applications to analyse complex development goals like poverty reduction, democratization, good governance, economic and sustainable development.

A detailed description of various impact assessment methods used in socio-economic research can be found at <http://are.berkeley.edu/~sadoulet/papers/deJanvryetal2011.pdf>

A few applied impact assessment methods commonly found in the literature on impact study are presented in Table 24.1. However, a right mix of both qualitative and quantitative techniques are required to study the impact.

Table 24.1: Impact Assessment Methods and Techniques

Impact type	Method	Technique
Intermediate impact • Institutional changes • Changes in the enabling environment	Survey, monitoring	Simple comparison/trend analysis
Direct product of research	Effectiveness analysis using logical framework	Simple comparison: target vs. actual
Economic impact	Econometric approach, surplus approach	Production function, total factor productivity, index number methods, and derivatives
Socio-cultural impact	Socio-economic survey/ adoption survey	Comparison over time
Environmental impact	Environmental impact assessment	Various • Qualitative • Quantitative

24.3 Key Points

- Impact assessment (IA) is crucial for judging the utility of an intervention.
- In NARES, impact assessment was studied to assess socio-economic effects, problems with technologies and process, prioritization of research projects and to justify the costs sunk into research.
- The IA framework includes inputs, outputs, outcome and impacts. Identification of criteria and indicators greatly affects IA studies.
- Indicators, which are developed specifically for technical, institutional, socio-cultural, commercial, economic and environmental aspects, have to be spelt out clearly and objectively, so that it is amenable for precise quantification. Both quantitative and qualitative methods are used in IA studies.
- Although the quantitative approach is preferred for better results, yet the qualitative approach is a sinequa non in many instances. IA is classified in many ways, viz., ex-ante and ex-post; experimental,

quasi-experimental, and non-experimental.

- The popularly adopted economic impact assessment techniques include partial budgeting techniques such as Net present value, benefit-cost ratio and internal rate of return. Other techniques such as economic surplus model, regression models, structural change in time series data, input-output model, etc. were used often for economic-impact assessment.

References

- Badioli, F. (2011.) Indicators of Impact. Paper presented at the Training Workshop on Sustainability & Embedding. Euromed Heritage. (Retrieved from: www.euromedheritage.net/euroshared/doc/Results%20indicators_FBADIOLI.pdf)
- Baker, J. L. (2000.) Evaluating the impact of development projects on poverty: A handbook for practitioners. Washington D.C.: World Bank.
- Khandker, R. S., Koolwal, G.B. and Samad, S.A. (2010.) Handbook on Impact Evaluation: Quantitative Methods and Practices. Washington, D.C.: World Bank.
- Manyong, V.M., B. Douthwaite, O. Coulibaly, and J.D.H. Keatinge. (2001.) Participatory impact assessment at the International Institute of Tropical Agriculture: functions and mechanisms. Pages 69–74 (Annex) in: The future of impact assessment in the CGIAR: needs, constraints and options. Proceedings of a workshop organized by the Standing Panel on Impact Assessment of the Technical Advisory Committee, 3–5 May 2000. Rome, Italy: FAO.
- MLE (2013.) Types of evaluation designs. Measurement, Learning and Evaluation Project. (Retrieved from: <https://www.urbanreproductivehealth.org/toolkits/measuring-success/types-evaluation-designs>)

25

Publishing Extension Research

Rasheed Sulaiman V

Objectives

1. Introduce the importance of publishing extension research
2. Explain aspects related to preparation of good manuscripts for publishing in journals
3. Introduce different non-journal formats for publishing research outputs

25.1 Introduction

Publishing sound research findings is vital for the progress of science. Publishing allows free exchange of information and allows one's research findings to be added to the literature on a given subject and will also help in building new theories and developing new policies. It also ensures and promotes accountability of researchers to the investments made by the government or shareholders who have funded it through taxes or contributions. Apart from these, publishing also helps the researcher in gaining feedback on his/her work, improving career prospects, securing new research funding, and gaining credibility as an expert on a specific research domain or topic.

As an applied social science that addresses the needs of stakeholders in the rural setting, the science and practice of extension is highly dependent on research in extension. To remain relevant, extension needs sound methods and models to help its professionals to address emerging problems in agriculture and rural development. However, due to several reasons, there is very little research that happens in the field of extension in India. Most of the current research in extension in India is from students' thesis work and due to the time and resource limitations of student's research, most of the findings and recommendations coming out of these have very limited policy relevance. Only very little of these is published in international and/or multi-disciplinary journals.

"Though extension scientists in the NARS can potentially engage in useful research that can influence the technology development process, they are mostly engaged in organizing training programmes, conducting events, dealing with visitors to the institute and handling documentation responsibilities" (Gowda et al., 2014). Though many extension practitioners are involved in extension work, they rarely publish research articles – may be due to their lack of interest in such publications, or their capacity to write articles based on their work experience – in different types of publications, which is a big loss to the extension profession.

25.2 Discussion

Academic/Journal Publishing in Extension

According to Rao et al. (2014), academic publishing in extension has suffered mainly due to the limitations in quantity and quality of extension research. The quality of extension journals has also declined due to:

- a. Lack of willingness and capacity of extension researchers to take up quality research to produce quality output;
- b. Lack of capacity to prepare quality research papers (drafting skills/articulation/language);
- c. Lack of serious review of submitted papers by editors/editorial committee members/reviewers (especially those in charge of Indian journals on extension).

The first and foremost measure to address these problems is to enhance the quality of extension research. This depends on the methodological rigour and ability to produce results that can be generalised across similar socio-economic systems.

Secondly, extension researchers should reach out to other relevant social science and general agricultural journals having higher rating and professional standing to publish their quality research work. Though there are a number of non-extension journals (related to social science and general agriculture), which also publish extension articles, many researchers are not aware of these. To address this issue, the AESA has compiled and published a list of journals which are relevant for extension researchers (AESA, 2014).

Thirdly, one should understand the requirements of different journals and the format/style of the journal before actually writing the paper. "Every journal has its audience, history, format, and type of complexity, be it methodologically or conceptually. It is best to know your target journal upfront and specifically write for that, keeping all expectations of that journal in mind" (Mulder, 2016).

What do editors and reviewers look for from journal articles in extension?

Based on my personal experience as a reviewer of papers of a few social science journals, I base my review of extension papers mainly on the following aspects:

- a. Relevance of core results and conclusions to the specific focus of the journal (scope and objectives of the journal), extension policy, extension practice (beyond the immediate community of one's research);
- b. Novelty of the research problem: A novel research problem which addresses a global issue;
- c. Conceptual framework used linked to the quality of the review of literature (findings from the latest papers relevant to the topic you are discussing, different dimensions related to the topic);
- d. Rigour in methodology: Appropriateness of research methods to the nature of the problem studied/experimented, including data analysis techniques;
- e. Extent of adherence to the journal's specifications (Author Guidelines); and
- f. Clarity of presentation: Structure, coherence (logical flow of ideas), use of correct grammar and spelling.

Though journals do differ on the reviewer guidelines, I think the above aspects are perhaps common across all journals. Reviewers make recommendations to the Editor to publish, or to publish with minor/major revisions suggested, or to reject the paper entirely. Generally, reviewers explain in detail the reasons for revision and how the paper could be revised. Sometimes the same reviewer also goes

through the revised version to make sure that the quality of the paper has improved by following his/her comments. It should be noted here that reputable journals publish only about less than 20% of the manuscripts it receives. There are several reasons for rejection of research manuscripts (Box 25.1).

Box 25.1: Common reasons for rejection of research articles

A recent check of reasons for recommending rejections of manuscripts to the Journal of Agricultural Education and Extension (JAEE) by reviewers shows the following:

- A lack of a conceptual or theoretical framework;
- The language which is used is difficult to understand;
- An insufficient or incomplete problem statement (which does not address a general practical or scientific problem, or does not contribute to the advancement of agricultural education and extension theory development);
- The presentation of inaccurate or inconsistent data;
- An outdated review of literature, which does not link with current mainstream research in the journals which are listed in the Web of Knowledge;
- A sample which is too small or even biased;
- A sampling method which is inappropriate or insufficiently described (the research methods are not documented in a way that informs the readers on how to repeat the study);
- The treatment of an unimportant or irrelevant topic (or a theme which is not interesting for the JAEE at worldwide level);
- Conclusions which are not related to the research question.

Source: Mulder, M (2016)

“Inexperienced scholars should understand a few key aspects of the revision process. First, it is important to address the revisions diligently; second, it is imperative to address all the comments received from the reviewers and avoid oversights; third, the resubmission of the revised manuscript must happen by the deadline provided by the journal; fourth, the revision process might comprise multiple rounds” (Shaikh, 2016). One need not necessarily agree with all the comments of the reviewer, but should respond politely to all the comments in detail, clearly indicating how one has addressed the comments in the revised version. In case of disagreement, provide details, including publications to support your argument.

Though publication of research findings in academic journals is highly important, there are several other forms of publishing that are equally important. The next section discusses some of these publishing formats.

Beyond Research Journals

In addition to publishing research results in peer-reviewed journals, books, working papers and conference proceedings which are accessed mainly by fellow researchers in extension, extension researchers can, and should, publish in other formats which many others mostly access. For instance, policy makers in extension, especially senior bureaucrats in the Ministry of Agriculture or Rural Development is keen to learn from the researchers’ experience with field level implementation of a programme for which he/she is responsible. Similarly, the Members of Parliament, other senior officials representing different organisations/enterprises in the agri-food innovation system might be interested in knowing about the performance of extension and how it could be improved. But many members of this diverse audience often do not access research journals to know about these aspects due to lack of time and/or low aptitude for reading relatively dense and complex presentations of findings in research journals.

Another major audience for evidence-based good practices on extension is extension practitioners who are interested to learn from the good practices in extension followed by other practitioners. Practitioners are also keen to learn from best practices, new frameworks, and tools that could be applied in the field. Extension researchers should also meet the knowledge requirements of these wider set of actors too. Some of the relevant forms of publishing non-journal articles are discussed below.

Policy Briefs: A policy brief is a stand-alone document, usually a 4-page document that presents the findings and recommendations from research to a policy relevant audience. It should offer policy makers specific recommendations on what needs to be done and how and what are the implications for such action. For instance, one could write a policy brief on the performance of Agricultural Technology Management Agency (ATMA), Krishi Vigyan Kendras (KVKs), Kisan Call Centres (KCC) etc., based on research on these different organisations/schemes where there is a lot of public interest in improving their performance. Policy briefs are written in a different style and it would be useful to read the writing and presentation style used in policy briefs. Some of the good practices include: “breaking the text into short paragraphs, using boxes, subheading and bullet points as long as the text still makes sense and considering use of diagrams and tables to save on words” (French-Constant, 2014).

Blogs: Social media has over the past one decade become one of the most important sources of information and advice for those who have access to the internet. Increasingly, blogs are used for science communication. This is especially true for many of us who gain new insights and experience on the job which are relevant to a wider audience, but which can't be communicated as a journal paper. Blogs provide you with the right platform to share your views, opinions, and the lessons you have learnt through the internet. You could either start your own blog on the internet or publish your write-up as blogs on other platforms that publish blogs. For instance, since 2013, the AESA has been publishing blogs on several topics related to extension – ranging from extension research, use of ICTs, varied extension approaches, livestock extension, extension policy, etc., – from researchers, practitioners and policy makers (<http://www.aesa-gfras.net/blogs.php>).



Blogs are increasingly used as a 'teaser' to promote more detailed research outputs, such as policy papers and journal outputs, by several researchers currently.

Good Practices: The essence of identifying and sharing good practices is to learn from others and to re-use knowledge. It is defined as anything that has been tried and shown to work in some way — whether fully or in part but with at least some evidence of effectiveness — and that may have implications for practice at any level elsewhere. Three possible levels of good practice flow from this: promising practices, demonstrated practices, and replicated practices (Serrat, 2008). GFRAS has been publishing global good practice notes on several extension approaches and methods in easy-to-understand formats (<http://www.g-fras.org/en/ggp-home.html>).



AESA has been publishing good practices based on experiences from practitioners who have experimented with successful approaches in extension delivery (<http://www.aesa-gfras.net/goodpractices.php>).

Toolkits, Manuals and Guides and Tools: There is considerable appetite for practical hands-on 'How to do' toolkits, manuals and guides among project design and implementation teams within extension organisations. For instance, project managers would like to have guidance on how to design demand-led programmes for farm women, support adaptation to climate change, farmer organisation development, linking farmers to markets, etc. Usually these are developed based on the analysis of good practices followed by different projects. These have to be written step-wise with specific cases and illustrations. IFAD is developing toolkits to build the capacity of project design and implementation teams on various thematic areas, to support the scaling up agenda, and support policy dialogue. These modular toolkits are composed of three different documents: a 'Teaser', a 'How To Do' Notes and 'Lessons Learned', serving different purposes and audiences (https://www.ifad.org/en/topic/tags/knowledge_notes/7001030). Trainers in extension often need Manuals/Facilitator Guides on specific topics to support the organisation of trainings/workshops on different topics.

25.3 Key Points

- Publishing research outputs having high quality and impact as journal articles and in various other formats is important for the growth of extension discipline, and also for enhancing its contribution to extension policy and practice.
- Publishing scholarly research in high impact journals depends not only on the quality of one's research (scientific/policy relevance, use of appropriate methods of data collection and analysis and quality interpretation, etc.), but also on selecting the right journal for one's research topic and the level of one's writing. Apart from these one also needs a certain amount of luck for publishing peer-reviewed articles.
- Researchers need to master the skills of writing for different formats (policy briefs, blogs, good practices, manuals, guides, etc.), including writing for social media.
- There are no shortcuts to improve one's writing skills other than reading and writing. Language editors can only marginally improve one's presentation style and correct grammar and spelling, but they cannot improve the overall quality of the paper.

- Finally, one needs passion for the topic of one's research and for writing if one wants to publish quality research work.

References

- AESA. (2014.) (Available at: <http://www.aesagfras.net/Resources/file/Where%20we%20can%20publish%20extension%20research%20-%20Final%20Note.pdf>)
- French-Constant, Laura. (2014.) How to plan, write and communicate an effective policy brief. Three steps to success. (Available at: <http://www.researchtoaction.org/wp-content/uploads/2014/10/PBWeekLauraFCfinal.pdf>)
- Gowda C M.J., Dixit S, Burman R and Ananth P. N. (2014.) Extension research and technology development. Blog 29, Agricultural Extension in South Asia. (Available at: http://aesa-gfras.net/Resources/file/FINALM_J_Chandre%20Gowda-13-FEB.pdf)
- Mulder, Martin. (2016.) Editorial: Publishing in the JAEE. Journal of Agricultural Education and Extension 22(1):3-6. (Available at: <http://www.tandfonline.com/doi/full/10.1080/1389224X.2016.1140973>)
- Rao, S V N, Natchimuthu, K and Ramkumar, S. (2014.) Scientific publishing in extension: Are we doing enough and are we doing well? (Available at: http://www.aesa-gfras.net/admin/kcfinder/upload/files/Blog_40.pdf)
- Serrat, Olivier. (2008.) Identifying and sharing good practices. Knowledge Solutions, ADB. (Available at: <https://www.adb.org/sites/default/files/publication/27598/identifying-sharing-good-practices.pdf>)
- Shaikh, Aijaz A. (2016.) 7 steps to publishing in a scientific journal. (Available at: <https://www.elsevier.com/connect/7-steps-to-publishing-in-a-scientific-journal>)

26

Multi-Disciplinary Extension Research

P Sethuraman Sivakumar

Objectives

1. Provide an overview of multidisciplinary extension research
2. Recommend multi-disciplinary research areas to extension scientists

26.1 Introduction

- Extension research is inherently multi-disciplinary in nature, which focuses on devising strategies to study, solve and predict complex agricultural problems.
- Despite its multi-disciplinary nature, much of the extension research focus is on just the extension subject, without developing strategies to solve field issues. This module provides an overview of multi-disciplinary research and suggests various options to bring in a composite perspective.

26.2 Discussion

Discipline

- A discipline is a branch of knowledge, instruction, or learning;
- A discipline is held together by a shared epistemology, i.e., assumptions about the nature of knowledge and acceptable ways of generating or accumulating knowledge (Choi and Pak, 2006).

Collaborations

- The research can be conducted collaboratively with the involvement of two or more disciplines;
- Different type of collaborations are as follows:
 - Multi-disciplinarity – Draws on knowledge from different disciplines but stays within the boundaries of those fields (NSERC, 2012);
 - Inter-disciplinarity – Analyses, synthesizes and harmonizes links between disciplines into a coordinated and coherent whole (Choi and Pak, 2006);
 - Trans-disciplinarity – Integrates the natural, social and health sciences in a humanities context, and in so doing transcends each of their traditional boundaries (Soskolne, 2000).

Differences among multi-disciplinary, inter-disciplinary and trans-disciplinary collaborations are displayed in Table 26.1.

Table 26.1: Differences between various types of research collaborations (Choi and Pak, 2006)

Aspect	Multi-disciplinary	Inter-disciplinary	Trans-disciplinary
Discipline boundaries	Working with several disciplines within their boundaries	Working between several disciplines by cutting across their boundaries	Working across and beyond several disciplines
Number of disciplines involved	Involves more than two disciplines	Involves two disciplines	Involves scientists from relevant disciplines (> 2), as well as stakeholders, non-scientists, and non-academic participants
Mode of work	Members from different disciplines working independently (parallelly or sequentially) on different aspects of a common problem	Members from different disciplines working jointly on the same problem or aspect of the common problem	Members from different disciplines working together on the same aspect of the common problem
Goals and research questions	Individual goals and each discipline has its own research questions	Shared goals and research questions	Shared goals and research questions and shared skills
Roles of scientists	Participants have separate but inter-related roles	Participants have common roles	Common and dynamic roles
Type of collaboration	Additive – Each discipline adds its contributions, no interaction	Integrative – Analyses, synthesises, harmonizes, and links to a coordinated and coherent whole	Holistic/ Transcendental – subordinates the disciplines and looks at the dynamic of the whole system
Nature of outcome	The outcome is the sum of the individual parts	The outcome is more than the sum of the individual parts	Multiplicity of individual contributions

Need for multi-disciplinary research for extension

(i) To resolve a real world problem

- The Indian agricultural systems are complex and farmers operate in complex, diverse and risk-prone environments;
- The farmers' problems are often caused by the interplay of natural, biological, chemical, economic and psychological issues;
- Understanding and solving these problems require expertise from various biological and social disciplines;
- For example, depletion of ground water – When farmers are suffering due to lowering of ground water levels, multi-disciplinary research can be undertaken to understand the problem, assess the current and future effects, and suggest strategies. This research work calls for a soil and water conservation engineer (to assess the reasons for ground water depletion and suggest technical measures), an agronomist/horticulturist (to suggest water-efficient cropping systems), an economist (to estimate the cost economics) and an extensionist (to understand the community processes regulating digging of bore wells/tube wells as well as to put up an awareness campaign on sustainable water use).

(ii) To provide different perspectives on a problem

- Experts from different disciplines assess the problem and provide various viewpoints to solve the problem;
- For example, a farmer's suicide can be assessed through multiple perspectives involving an economist (studying indebtedness – causes and effects), an agronomist/horticulturist (assessing the role of cropping systems), and an extensionist (gauging the psychological state of the farmer and his/her family leading to the suicide). These perspectives can be compared and critically analysed to identify the primary cause of suicide and to suggest preventive or corrective measures.

(iii) To develop an objective and accurate perspective of a problem or phenomenon

- An agricultural problem or phenomenon is caused by interplay of various forces. A multi-disciplinary team of experts will assess the problem or phenomenon in an objective way and provide accurate results;
- For example, when farmers adopt local landraces and ignore improved crop varieties, a multi-disciplinary team can provide an objective and accurate view on why this happens. The extension scientist can assess the adoption of crop varieties and provide a subjective view of the reasons for non-adoption. If the farmers indicate low yield and poor taste of improved varieties as reasons, the agronomist can assess the reasons for low yield while the food technologist can study the biochemical and textural qualities which cause low acceptability by the farmers. These results will thus provide an objective and accurate description of the reasons for non-adoption of improved varieties.

(iv) To improve the quality of research outputs

- In general, extension research outputs include new knowledge, field tools and techniques, extension models and policy recommendations, which are delivered in the form of research papers, technical reports, policy briefs, multimedia programmes, etc.;
- Since extension is a social science which deals with all types of technologies and processes, it offers technological solutions to field problems and facilitates its effective implementation. To improve the quality of research outputs and service delivery, extension fully depends on allied agricultural disciplines without which the field problems may not be solved;
- Multi-disciplinary collaboration will help extension scientists to publish agricultural problems or phenomenon-oriented research papers in eminent international journals, besides improving field practice. When extension scientists are part of a technology development process, they will be able to share the credit with other agricultural scientists;
- For example, a plant breeding research programme aimed at developing a drought-resistant rice variety is implemented through farmers' participatory research. After identifying elite lines, the breeder conducts on-farm trials in various locations where the extension scientist helps him to identify farmer-preferred varieties. Since the extension scientist is a part of the breeding work, his/her name will be included in the variety development team.

(v) To secure external funding for research projects

- Most of the funding agencies in India and abroad are offering problem-oriented financing, where the priority issues are solved through multi-disciplinary teams;
- Funding agencies look for proposals either in consortium or twinning mode, public-private partnership, and group funding with multidisciplinary teams;
- Single discipline projects are seldom preferred by funding agencies.

A few examples of multi-disciplinary research themes and the role of extension scientists are displayed in Table 26.2.

Table 26.2: Multidisciplinary research work and the role of extension scientists

Research theme	Objective	Role of extension scientist	Sample research papers
Participatory varietal development	To develop plant varieties with farmer preferred attributes	<ul style="list-style-type: none"> To conduct farmer participatory varietal evaluation during On-Farm Trial 	Witcombe et al., (1996); Thapa et al. (2009)
Development of conceptual technological profiles, such as variety, food product, pesticide, etc	To design a technology, such as variety, food product based on farmers'/ consumers' preferences before starting the actual research	<ul style="list-style-type: none"> Designing a conceptual technology profile through conjoint analysis Developing farmer-preferred variety profiles by integrating specific attribute preferences stated by farmers. Utility values of each varietal attribute indicate the importance of that specific attribute for farmers Developing a food product profile (a conceptual food product) by integrating the product attributes (taste, colour, texture,), packaging (bottle, tetra pack) and price attributes 	Crop variety design - Baidu-Forson et al. (1997) Cattle traits – Tano et al. (2003) Fish products – Özayan et al. (1998)
Development of user-friendly agricultural implements and other machineries	To develop user-friendly (including gender specific) agricultural implements through farmer participatory research (ergonomics)	<ul style="list-style-type: none"> Agricultural implements can be designed and pilot-tested using ergonomic process for their efficiency by involving prospective users Will assess the difficulty and efficiency levels in operating the equipments through video recording, motion monitors, wrist systems and analytical software (e.g., JET) 	Agricultural hand tools – Wibowo and Soni (2016). Milking machine – Pinzke et al. (2001)
Development of strategies for enhancing climate adaptation capacities of the farmers	To develop farming systems for enhancing climate adaptation capacity of the farmers	<ul style="list-style-type: none"> Documentation of indigenous climate adaptation strategies of the farmers and assess their rationality of r refinement and large scale application 	Nkomwa et al. (2014) Bola et al. (2013)
Developing watersheds for sustainable crop production	To assess and create adequate irrigation facilities for farmers	<ul style="list-style-type: none"> To assess the utility of the irrigation service as developed by an Institution, based on farmers' perspective 	Ghosh et al. (2005).
Development of a food product	To develop a food product from agricultural produce or animal products	<ul style="list-style-type: none"> To assess consumer acceptability of a food product through sensory evaluation. 	Sivakumar et al. (2010)

Research theme	Objective	Role of extension scientist	Sample research papers
Crop and animal production systems assessment for planning technology interventions	To assess crop and animal production systems for problems and opportunities to develop technology interventions	<ul style="list-style-type: none"> • Production system diagnosis through Participatory Rural Appraisal • Documentation of existing crop and animal production systems 	Motivation for pesticide use through Means-end analysis – Lagerkvist et al. (2012)

26.3 Key Points

- Multi-disciplinary extension is an emerging research tradition in extension, where various disciplines work together for common goals without crossing their boundaries.
- Multi-disciplinary extension research is additive in nature, where individual contributions are summed up to achieve the common goal.

References

- Baidu-Forson, J., Ntare, B., and Waliyar, F. (1997.) Utilizing conjoint analysis to design modern crop varieties: empirical example for groundnut in Niger. *Agricultural Economics* 16:219–226.
- Bola, G., Mabiza, C., Goldin, J., Kujinga, K., Nhapi, I., Makurira, H., and Mashauri, D. (2013.) Coping with droughts and floods: A case study of Kanyemba, Mbire District, Zimbabwe. *Physics and Chemistry of the Earth* 67–69:180–186.
- Choi B. C. K., and Pak A. W. P. (2006.) Multidisciplinarity, interdisciplinarity, and transdisciplinarity in health research, services, education and policy. 1. Definitions, objectives, and evidence of effectiveness. *Clinical and Investigative Medicine*, 29:351-364.
- Ghosh S., Singh R., and Kundu, D. (2005.) Evaluation of irrigation-service utility from the perspective of farmers. *Water Resour Manag* 19:467–482.
- Lagerkvist, C.J., Ngigi, M.W., Okello J.J., and Karanja, N. (2012.) Means-end chain approach to understanding farmers’ motivations for pesticide use in leafy vegetables: The case of kale in peri-urban Nairobi, Kenya. *Crop Protection Journal* 39:72-80.
- Natural Sciences and Engineering Research Council of Canada. (2012.) Guidelines for the preparation and review of applications in interdisciplinary research. Ottawa: NSERC. (Retrieved from: http://www.nserc-crsng.gc.ca/NSERC-CRSNG/Politiques-Politiques/prepInterdiscip-prepInterdiscip_eng.asp).
- Nkomwa, E.C., M. K. Joshua., C. Ngongondo, M. Monjerezi and F. Chipungu. (2014.) Assessing indigenous knowledge systems and climate change adaptation strategies in agriculture: A case study of Chagaka village, Chikhwawa, Southern Malawi. *Journal of Physics and Chemistry of the Earth* 67-69:164-172.
- Özayan, A., Wes Harrison, R., and Meyers, S.P. (1998.) An analysis of buyer preferences for new food products derived from Louisiana’s undersized crawfish. Bulletin No. 864. Louisiana State University Agricultural Center. (Retrieved from: <http://www.lsuagcenter.com/nr/rdonlyres/a137c8e7-8d92-477c-8620-f44e57e5f692/3920/b864craw.pdf>).
- Pinzke, S, Stal, M, Hansson, G. (2001.) Physical workload on upper extremities in various operations during machine milking. *Ann Agric Environ Med*. 8:63–70.
- Sivakumar, P.S., Panda, S.H., Ray, R.C., Naskar, S.K., and Bharathi, L.K. (2010.) Consumer acceptance of lactic acid fermented sweet potato pickle. *Journal of Sensory Studies* 25:706-719.

- Soskolne, C. (2000.) Transdisciplinary approaches for public health. *Epidemiology* 11:S122.
- Tano, K., Kamuanga, M., Faminow, M.D., and Swallow, B. (2003.) Using conjoint analysis to estimate farmers' preferences for cattle traits in West Africa. *Journal of Ecological Economics* 45:393-407.
- Thapa, D.B., Sharma, R.C., Mudwari, A., Ortiz-Ferrara, G., Sharma, S., Basnet, R.K., Witcombe, J.R., Virk, D.S. and Joshi, K.D. (2009.) Identifying superior wheat cultivars in participatory research on resource poor farms. *Field Crops Research* 112:124-130.
- Wibowo, R.K.K. and Soni, P. (2016.) Farmers' injuries, discomfort and its use in design of agricultural hand tools: A case study from East Java, Indonesia. *Agriculture and Agricultural Science Procedia* 9:323-327.
- Witcombe, J.R., Joshi, A., Joshi, K.D., and Sthapit, B.R. (1996.) Farmers' participatory crop improvement. 1. Varietal selection and breeding methods and their impacts on biodiversity. *Experimental Agriculture* 32:445-460.

Final Remarks: Where do we go from here?

*Rasheed Sulaiman V, P Sethuraman Sivakumar, B S Sontakki, Saravanan R and
Nimisha Mittal*

Strengthening research and evaluation in extension will continue to remain important always if we have to enhance the contribution of extension to agricultural and rural development. As a field-oriented discipline, extension grows with every bit of quality research work that

- brings new insights into behavioural theories;
- develops sound field extension methods and tools; and
- supports research/extension/innovation management and policy.

Without applying the most relevant and up-to-date social science research methods and improving the methodological rigour in research, extension as a discipline will fail to grow or remain relevant to the rapidly evolving demands of its clients. Extension research also needs to answer the questions posed by its practitioners – the wide range of extension and advisory service providers – on improving the quality and effectiveness of advisory support for innovation.

This manual is an output of a collaborative effort by four organisations and one network. We hope it will be of immense value to extension researchers and practitioners alike, and more so to young researchers and research scholars who are keen to do research in extension. Other social science researchers working on agricultural and rural development may also find the different modules useful. In this manual we have tried our best to present most of the good practices in research and evaluation along the different stages in the research cycle, up to publications. We ventured into developing this manual on the firm belief that a manual like this is absolutely necessary to strengthen extension research, and thereby the extension profession in India. Interactions with our professional colleagues in other countries in the region (South Asia) have also convinced us on the urgency of organising a workshop on this theme and then developing a manual based on that experience.

We also realise that a manual like this can best be described as ‘a work in progress’, as there has been a rapid evolution in social science research methods globally. This trend is likely to continue, and therefore new versions of this manual will have to be developed to support young researchers – at least once in 3-5 years. But we also recognise that a manual like this alone won’t be able to make a big difference in addressing the current research capacity gaps among extension researchers. In other words, a manual like this is necessary but not sufficient to strengthen capacities in this area.

So where do we go from here?

Of course, we will try to make this manual available at University/Department libraries by mailing a limited number of printed hard copies of the manual. We will definitely post an electronic version of this manual online for free download and promote the same regionally and globally through the AESA and GFRAS networks. We will also try to make this document available to the faculty involved in

teaching research methods in extension, and will also share the slides used for presentations at the India workshop to help others to adapt and use them suitably in their own situations.

But the ultimate objective of bringing out this manual can be achieved only when more researchers receive trainings that use this manual as one of the important background documents or as a guide for trainers/facilitators. We have integrated research expertise in our workshop and in this manual from organisations working on diverse aspects of extension, namely ICAR-NAARM (Research and extension management), ICAR-CTCRI (Multi-disciplinary extension research), AESA/CRISP (Networking and policy research in extension) as well as MANAGE (Extension training and management).

- As a follow up, ICAR-NAARM has already included this area (Good Practice in Extension Research and Evaluation) as one of the important topics for training in its Annual Training Calendar for 2017-18.
- The extension group at ICAR-CTCRI is currently expanding its research focus to new areas, such as technology commercialisation and techno-entrepreneurship development, as well as longitudinal extension field research to bring 'futuristic insights' into the extension discipline.
- As a policy research centre working on extension and innovation for more than a decade, CRISP will continue to lead policy relevant research in this area and influence extension policy.
- AESA, through its country level extension networks in South Asia and through its links to GFRAS, will be promoting this work regionally and globally and will also try to raise resources to organise similar workshops elsewhere.
- MANAGE has already integrated a one-week module on 'Good Practices in Extension Research and Evaluation' in the proposed three-month 'MANAGE Internship Programme for the PG students in extension education' starting from April 1, 2017 to June 30, 2017.

As a consortium, we are currently planning to organise a series of topic-focused (for example, psychometric scaling, qualitative extension research, etc.) and sector-specific (research modules for KVKs, ICAR-ATARI, state extension departments) programmes in the coming years. This training module will also be expanded and customised to provide need-based training for extension scientists working in various regions of India.

We hope other educational and training centres engaged in education and research in India and elsewhere will also start using and adapting this manual for organising face-to-face training for young researchers in the coming days.

We would be highly obliged and grateful to receive feedback on this manual. We firmly believe that this humble beginning will usher in more collaborative efforts towards improving the relevance, quality, and image of extension as a profession.

Annexure A-1

Guidelines for Designing Web Surveys

A well-designed web survey questionnaire looks professional, is easy to understand by the respondents, and motivates them to proceed to subsequent sections in order to complete the survey. Dillman (2000) and Sue and Ritter (2011) have identified a set of design principles that can be applied for designing Web-based questionnaires.

1. Welcome Screen

- A welcome screen is the first page of the survey which introduces the readers to the topic.
- A good welcome screen is motivational, emphasizes the ease of responding, and instructs respondents on the action needed for proceeding to the next page.
- The welcome screen provides an opportunity to describe or reiterate the purpose of the survey, explains how the respondent was selected for participation, discusses the conditions of anonymity and confidentiality, and explains how to collect or redeem incentives, if applicable.

2. Access Control

- Access control provides a personal identification number or password for limiting access to only the people in the sample.
- Essential to restrict the survey only to the selected sample and eliminate unintended participants from filling the survey.

3. First Question

- The first question of the survey sets the tone; it is therefore essential that the first question be short, simple, and, if possible, fun for respondents.
- Good first questions are closed-ended with radio buttons or check boxes for responses.

4. Conventional Format

- Present each question in a conventional format similar to that normally used on self-administered paper questionnaires.
- Includes numbering questions, left justifying text, and presenting response options either to the right of, or directly below, the questions to which they refer.

5. Colour

- Colour can enhance the appearance of the survey, assist with navigation, and motivate the

respondent.

- Colour comprehension is associated with culture. A few colour associations of adults in India are as follows: (<http://www.webdesignerdepot.com/>; <http://www.sensationalcolor.com/>)
 - Red - Purity, Fertility
 - Blue - Strength
 - Yellow - Commerce
 - Saffron - Sacred
 - Purple - Wealth
 - Pink - Feminine
 - Black - Evil, negativity, and inertia
 - Green – Nature

6. Appearance

- Compatibility with browser affects respondent's viewing.
- Appearance of questionnaire at different screen resolutions may be tested before hosting the survey.

7. Directions and Instructions for Filling-up

- Instructions for completing the questionnaire should always be included.
- When writing directions, avoid jargon and do not use abbreviations without first giving their expanded form.
- Instructions might address some or all of the following questions:
 1. Does the respondent have to answer all the questions?
 2. Can the respondent select only one answer for the question or more than one?
 3. How does the respondent move to the next question?
 4. Is there a time limit for completing the questionnaire?
 5. Can the respondent skip a question and return to it later?
 6. How does the respondent change an answer?
 7. Does the respondent need to single or double click on the answer?
 8. Can the respondent begin the survey and return to it later?
 9. If the respondent returns to complete the survey later, will he or she have to start all over again?

8. Formats for Response Types

- Response types include: radio buttons, check boxes, drop-down menus, rank order matrices, constant sum, and open-ended text boxes.
- It is important to maintain consistency in terms of font type and size, width of response categories, and colors used throughout the questionnaire.
- A few response types and their suitability for a specific research question are displayed in Table 10.3 in Module 10.

9. Requiring Answers

- If the respondents are participating voluntarily in the survey, please avoid compelling the respondents to provide an answer to each question before being allowed to answer subsequent ones.

10. One page Vs Multi-page Questionnaire

- Multi-page questionnaire – Respondents focus on one question at a time, perhaps mitigating order effects.
- Order effects occur if respondents' answers to particular questions are influenced by previously recorded answers.

11. Double and Triple Banking

- If the number of response options exceeds what will fit in two or three columns on one screen, consider other options such as drop-down menus.

12. Navigation Guides

- Provide navigation aids, such as progress bars, next button, etc.
- Since the respondents' participation is voluntary, provide Exit Survey button in all pages.

13. Font Type and Text Size

- Arial and Times New Roman are similar in terms of legibility and preference, and 14-point fonts are preferred by children and older adults.

14. Images, Graphs, and Charts

- Images, graphs, and charts should be used sparingly when creating online surveys.

15. Motion, Sound, and Links

- Motion, sound, and links will increase the time to download the survey which may be avoided.

16. Ensuring that Participants Respond Only Once

- Assigning a unique identifier to each respondent is the most effective way to prevent people from entering multiple responses.
- Compare host names or Internet protocol (IP) addresses of submissions.

References

- Dillman, D. A. (2000). *Mail and Internet Surveys: The Tailored Design Method* (2nd ed.). New York: Wiley.
- Sue, V.M. and L.A. Ritter. (2011). *Conducting online surveys*. Sage Publications, Thousand Oaks, CA.

Annexure A-2

Advanced methods of data collection

(a) Simulations

Simulations are models of real world processes, systems or events developed with the help of computer software using relevant indicators depending on the objectives. They are interactive and inquiry based, supported with learning materials to make data collection more accurate. Simulations can help in modelling demographic change or behaviour, and as a tool for formalizing theory (Rose et al., 2015). In extension, simulations can be used for action research, capacity development, mapping flow of information, understanding farming communities, planning research methodology and developing explanations for adoption and diffusion phenomena.

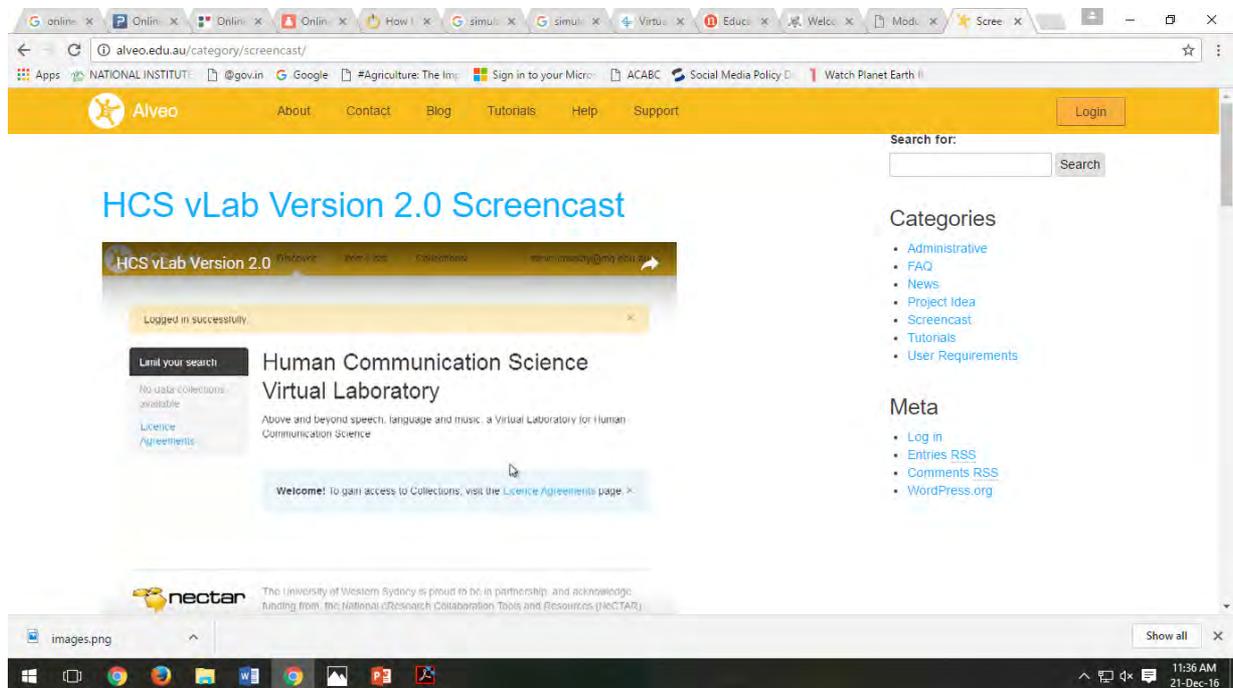
The screenshot shows a web browser displaying the Open University Press website. The page features the book cover for 'Simulation for the Social Scientist' (Second Edition) by Nigel Gilbert and Klaus G. Troitzsch. The cover is red with a blue and white pattern of small figures. The text on the page includes the publisher's name 'OPEN UNIVERSITY PRESS McGraw-Hill Education', the date 'February 2005', and ISBNs for paperback (0 335 21600 5) and hardback (0 335 21601 3) editions. A list of bullet points highlights key topics: 'What can computer simulation contribute to the social sciences?', 'Which of the many approaches to simulation would be best for my social science project?', and 'How do I design, carry out and analyse the results from a computer simulation?'. A paragraph discusses the growing interest in social simulation due to advances in hardware and software. The authors' names, 'Nigel Gilbert' and 'Klaus G. Troitzsch', are prominently displayed. The browser's address bar shows the URL 'cress.soc.surrey.ac.uk/s4ss/index.html'.

(b) Virtual laboratories

Virtual laboratories are online platforms where research activities can be carried out with the help of online software tools, with discipline-specific research environments integrating databases, computational tools and models. They also provide a platform for scientists across the world to collaborate efficiently

and provide remote access. Virtual laboratories are the next step of MOOCs, which limits itself to videos and lectures, while virtual labs go beyond to help put the ideas into practice so as to gain practical skill as well (Waldrop, 2013; Intersect Australia, 2016).

Virtual labs can be used in extension for skill development of extension personnel, for distance learning, and technology transfer on aspects such as pest and disease diagnosis, micro-irrigation, laboratory techniques, etc. Data on participant performance is also available for better planning.

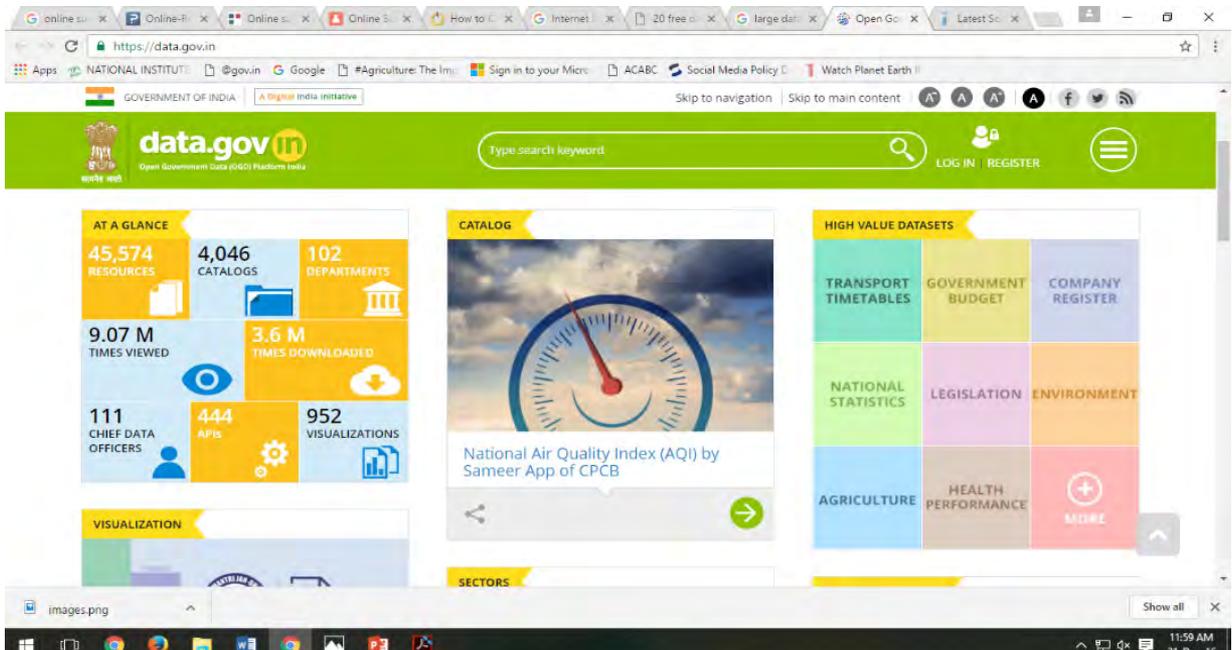


(c) Large data sets

Large data sets are important for social science research that is related to administrative records, public opinion surveys, economic and social data, and studies compiled by other researchers. Large data sets are generally available in machine-readable format (file layout maps, technical notes, questionnaires, reports, and errata).

Globally, some important resources found are: Sociosite (www.sociosite.net/databases.php), Cornell Institute for Social and Economic research (www.ciser.cornell.edu/ASPs/datasource.asp). In India, data.gov.in (<https://data.gov.in/>), Indiastat.com (<http://www.indiastat.com/default.aspx>), Census of India (<http://censusindia.gov.in/>) are a few major websites. Big data sets often entail cleaning up the data into a usable format and visualization of data. Tools like Data Wrangler, Open Refine, Tabula, NVD3 can be used for data clean up while Chart.js, processing.js, Timeflow, etc., are useful tools for data visualization (Suda, 2013).

Working with large data sets is gaining more popularity in extension research for mapping development activities, identifying research areas, and planning methodologies for further research. Data cleaning and visualization tools can be very helpful for getting better insights into the process.



(d) Audience Response Systems (ARS)

Audience response system helps in collecting data electronically in group settings while maintaining confidentiality; with an added advantage of reduced errors and time required for data management. ARS allow immediate interaction electronically through a remote control used by audience members to respond to questions. Responses are recorded automatically in an electronic receiver that displays individual response without disclosing identity (McCarter and Caza, 2009; Gamito et al., 2009). Some examples are polleverywhere.com and polldaddy.com. Extension researchers often work with large groups of respondents and on topics that might be sensitive in nature, to which, getting a proper response can become difficult. With ARS, opinions of large groups can be collected, collated, and presented in a very short time maintaining complete anonymity. Also, questions of a sensitive nature can be dealt with comfortably with this tool.

References

- Gamito, E.J., Burhansstipanov, L., Krebs, L.U., Bemis, L., and Bradley, A. (2009.) The use of an electronic Audience response System for data collection. *Journal of Cancer Education*, 20, 80-86.
- Intersect Australia. (2016.) Virtual laboratories and eResearch tools. Module 2. (Retrieved from: <http://intersect.org.au/docs/attachments/Module2.pdf>).
- Mccarter, M.W. and Caza, A. (2009.) Audience response systems as a data collection method in organizational research. *Journal of Management and Organization*, 15, 122-131.
- Rose, S., Spinks, N. and Canhoto, A.I. (2015.) Simulation based research. Chapter-6. *Management research: Applying the principles*. (Retrieved from: http://documents.routledgeinteractive.s3.amazonaws.com/9780415628129/Chapter%206%20-%20Simulations%20final_edited.pdf).
- Suda, B. (2013). 20 free data visualization tools. (Retrieved from: <http://www.creativebloq.com/infographic/20-free-data-visualisation-tools-5133780>).
- Waldrop, M.M. (2013.) Education online: The virtual lab. (Retrieved from: <http://www.nature.com/news/education-online-the-virtual-lab-1.13383>).

Annexure A-3

Using SPSS for basic statistical analysis

A. t tests for assessing mean differences

I. Independent samples t test

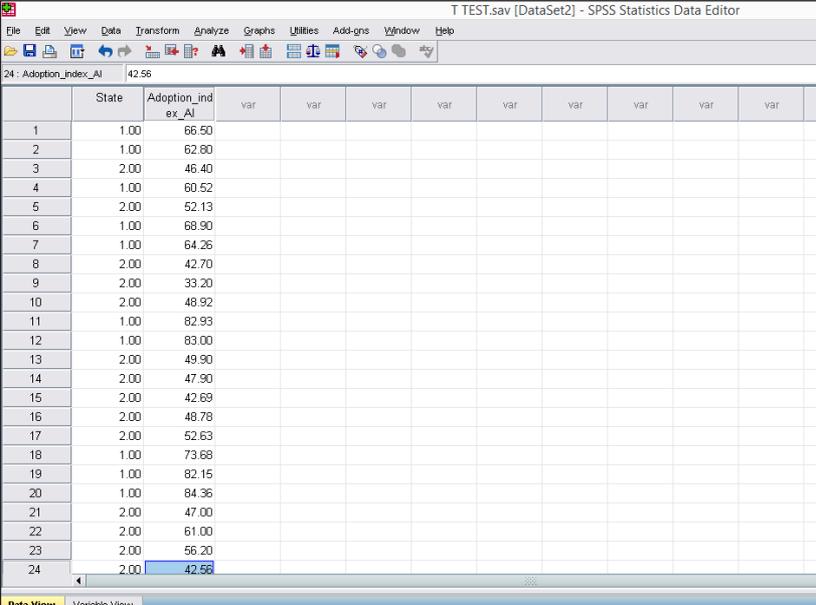
- Purpose – To test the mean difference of a specific attribute of two unrelated or independent samples;
- Assumption
 - Both variables are normally distributed;
 - The variances of the dependent variable in the two populations are equal.

Example

Adoption of Artificial Insemination technology is significantly higher among Haryana dairy farmers as compared to Kerala farmers.

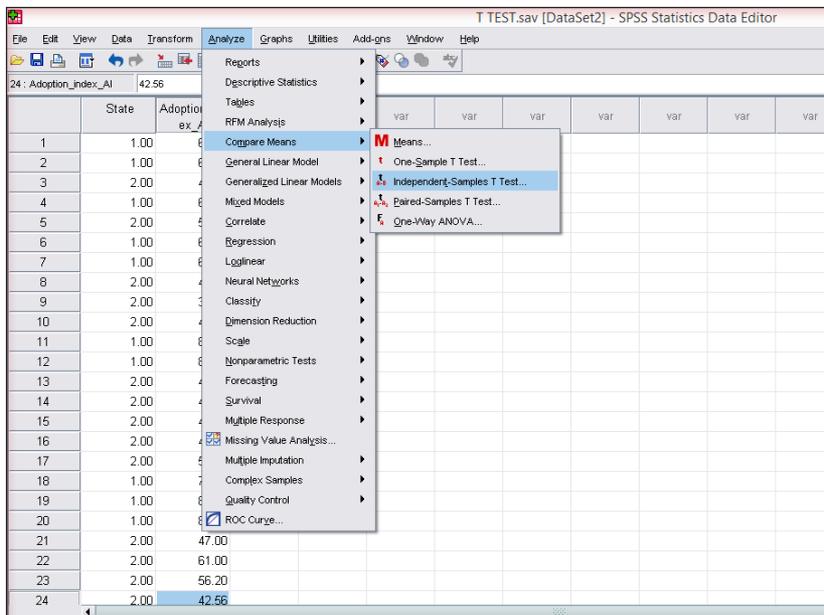
SPSS Analytical procedure

1. Enter the data in the spreadsheet as follows:

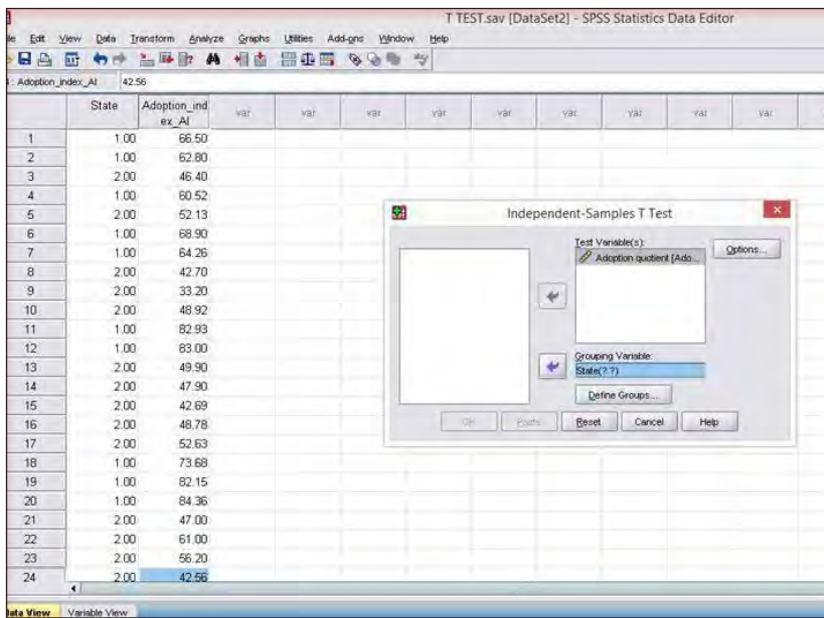


	State	Adoption_index_AI	var								
1	1.00	66.50									
2	1.00	62.80									
3	2.00	46.40									
4	1.00	60.52									
5	2.00	52.13									
6	1.00	68.90									
7	1.00	64.26									
8	2.00	42.70									
9	2.00	33.20									
10	2.00	48.92									
11	1.00	82.93									
12	1.00	83.00									
13	2.00	49.90									
14	2.00	47.90									
15	2.00	42.69									
16	2.00	48.78									
17	2.00	52.63									
18	1.00	73.68									
19	1.00	82.15									
20	1.00	84.36									
21	2.00	47.00									
22	2.00	61.00									
23	2.00	56.20									
24	2.00	42.66									

2. Click on Analyze → Compare means → Independent Samples T Test

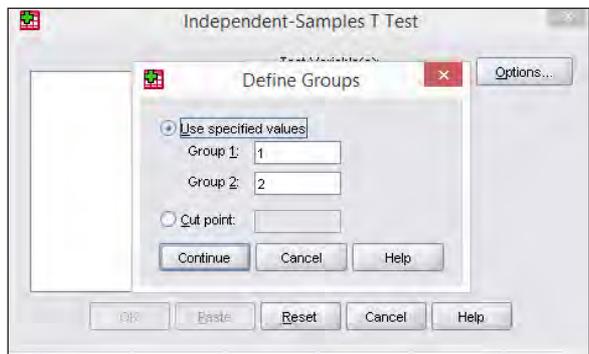


3. Move Adoption Quotient to the **Test Variable(s)** (dependent) box and move state to the **Grouping** (independent) **Variable** box



4. Next click on Define Groups

- Type 1 (for Haryana) in the Group 1 box and 2 (for Kerala) in the Group 2 box;
- This will enable us to compare Haryana and Kerala farmers with the dependent variable (Adoption quotient).



5. Click on **Continue** then on **OK**.

Output : Independent Samples t Test

T-Test

[Descriptive] > [Popovopk Method Modeling Network] > [Newel] > [workup 24-10-2016] > [Pavizal, brokay 16-12-2016] > [Dr. Dhesakopani] > [Data Analysis] > T TEST.

State	N	Mean	Std. Deviation	Std. Error Mean
Adoption quotient - Haryana	11	68.9864	15.9327	4.76713
Kerala	24	48.0007	6.72703	1.10787

Pair of means for Haryana and Kerala farmers. The Haryana farmers' adoption quotient (Mean = 68.9864) is higher than Kerala (Mean = 48.0007).

	Levene's Test for Equality of Variances		t-Test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
Adoption quotient -	4.044	.056	4.495	23	.000	20.98585	4.91912	11.32602	30.64447
			4.119	12.999	.001	20.98585	5.09545	9.96420	32.00710
								Lower	Upper

Indicates whether the variances of farmers' adoption quotient scores are equal. Since p = 0.056, the variances are equal and the "t" statistics displayed in the "Equal variances assumed" row can be interpreted.

The mean difference of adoption quotient between Haryana and Kerala is 20.98585. The "t" value of 4.495 is significant a level (p=0.000). This indicates significant difference between adoption quotient scores of both states.

Calculation of effect size

Reporting effect size of all statistical analysis is essential as it helps future researchers to assess the strength of relationships among test variables. For the independent test, the effect size (d) can be estimated using the following formula:

$$d = \frac{M_a - M_b}{SD_{pooled}}$$

Where d – effect size; $M_a - M_b$ – Mean difference and SD_{pooled} – Pooled standard deviation
 Formula for estimating pooled standard deviation

$$SD_{pooled} = \sqrt{\frac{(SD_1^2 + SD_2^2)}{2}}$$

Where SD_{pooled} – Pooled standard deviation; SD_1 = standard deviation for group 1 ;
 SD_2 = standard deviation for group 2

The SD_{pooled} is estimated at 4.89 and d value is 4.29

Box A-3.1: How to write the output of independent samples "t" test

Results

Table A-3.1 below shows that the Haryana farmers were significantly different from their counterparts in Kerala with regard to adopting the Artificial Insemination technology for dairy cattle. An examination of group means indicate that Haryana farmers' adoption of artificial insemination technology ($M = 68.99$) is significantly higher than that of their counterparts in Kerala ($M = 48.00$). The effect size d is estimated at 1.42, which is a very large size for effects in the behavioral sciences (Sawilowsky, 2009).

Box A-3.1: How to write the output of independent samples “t” test

Table A-3.1: Comparison of adoption of artificial insemination between dairy farmers of Haryana and Kerala states

State	Mean Adoption quotient	Standard Deviation	t	p	Effect size (d)
Haryana	68.99	16.81	4.50	.000	1.42
Kerala	48.00	6.73			

II. Paired sample t test

- Purpose – Used to assess the differences in a specific attribute or aspect before and after an intervention (repeated measures);
- Used in the single group quasi-experimental study (pre-, post-, or before-after experiments) in which the same assessment is used as the pre-test before the intervention, and as the post-test after the intervention.

Assumptions

1. The independent variable is dichotomous and its levels are paired. (e.g., pre-post, etc.);
2. The dependent variable is normally distributed in the two conditions.

Example

- The training programme is effective in increasing knowledge level of the farmers on horticultural nursery management.

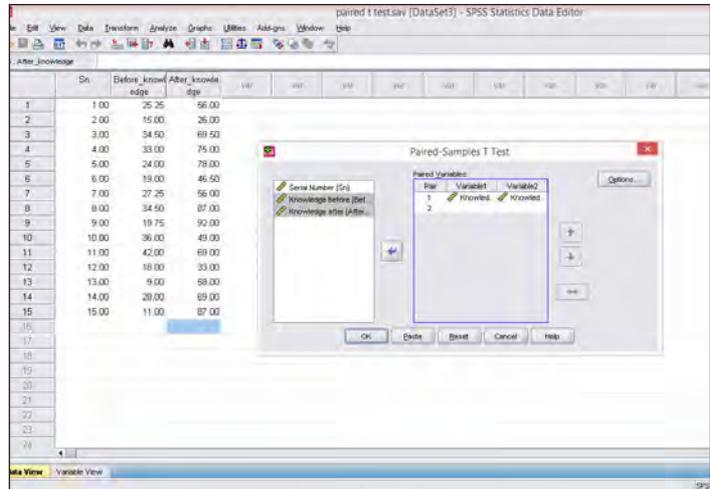
SPSS Analytical procedure

1. Enter the data in the spreadsheet as follows:

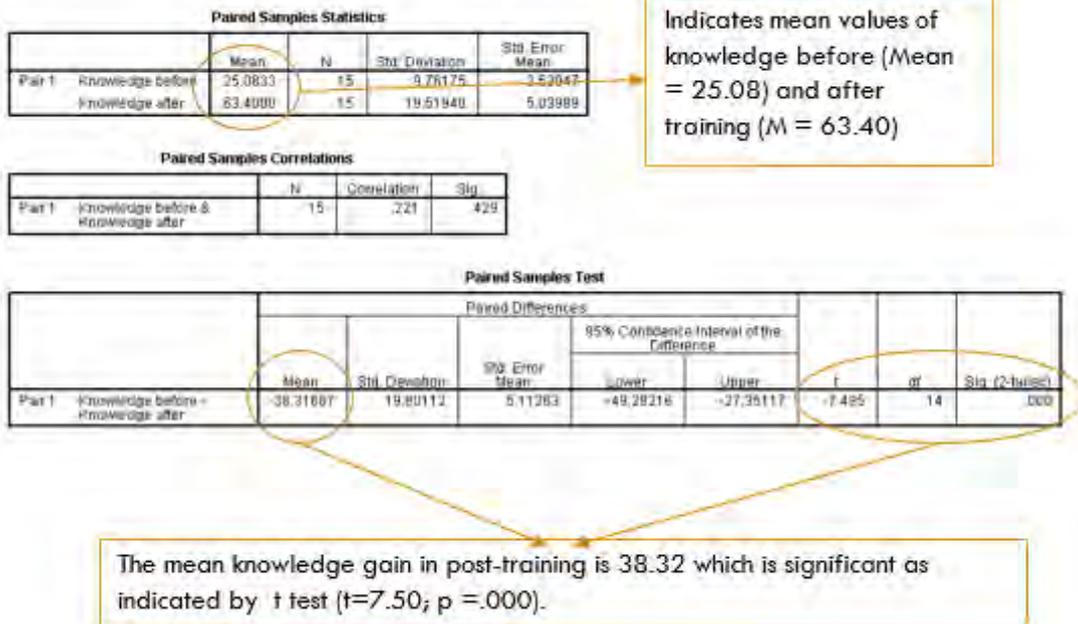
The screenshot shows the SPSS Data View window for a file named 'paired t test.sav'. The data is organized into a table with the following columns: 'Sn', 'Before_knowledge', and 'After_knowledge'. The rows represent 15 individual farmers, with their scores for knowledge before and after a training program. The 'Sn' column contains sequential numbers from 1 to 15. The 'Before_knowledge' column contains values ranging from 9.00 to 15.00. The 'After_knowledge' column contains values ranging from 26.00 to 89.00. The window title bar indicates the file is 'paired t test.sav [DataSet3] - SPSS Sta'.

Sn	Before_knowledge	After_knowledge
1	1.00	25.25
2	2.00	15.00
3	3.00	34.50
4	4.00	33.00
5	5.00	24.00
6	6.00	19.00
7	7.00	27.25
8	8.00	34.50
9	9.00	19.75
10	10.00	36.00
11	11.00	42.00
12	12.00	18.00
13	13.00	9.00
14	14.00	28.00
15	15.00	11.00

2. Select Analyze → Compare Means → Paired Samples T Test
 - Move both of the variables, Before_ knowledge and After_ knowledge, to the Paired Variables: box A-3.2;
 - Click on OK.



Output: Paired samples t test



Box A-3.2: How to write the output of paired samples “t” test

Results

The paired “t” test indicated that the farmers’ knowledge level has significantly increased after the training programme (t = 7.50; p=.000) (Table A-3.2). The effect size d is estimated at 2.49, which is a huge effect in the behavioral sciences (Sawilowsky, 2009).

Box A-3.2: How to write the output of paired samples “t” test

Table A-3.2: Comparison of adoption of artificial insemination between dairy farmers of Haryana and Kerala states

State	Mean Knowledge level	Standard Deviation	t	p	Effect size (d)
Before training	28.08	9.76	7.50	.000	2.49
After training	63.40	19.52			

B. Chi-square, Phi and Cramer's V tests

- Purpose – To determine if a statistically significant relationship exists between the two dichotomous or nominal variables;
- Requires a relatively large sample size (> 50) and a relatively even split of the subjects among the levels or categories;
- Since Chi-square test indicates only if the relationship between two nominal variables is statistically significant, it does not tell the strength of the relationship. Phi and Cramer's V tests are used along with Chi-square to estimate the strength of the association between two nominal variables (Morgan et al., 2011). These coefficients are essentially the estimators of effect size.

Assumptions for the Use of Chi-square, Phi and Cramer's V tests

- All the data are treated as nominal, even if ordered;
- If the expected frequencies are less than 5, the Fisher's exact test is used. Chi-square will provide accurate estimates only if at least 80% of the expected frequencies of all variables are over 5 (Morgan et al., 2011).

Example

Do male and female farmers differ on their extension participation?

Variables:

- Gender: Male – 1; Female – 2;
- Extension participation: Participated – 1; Not participated – 0.

SPSS Analytical procedure

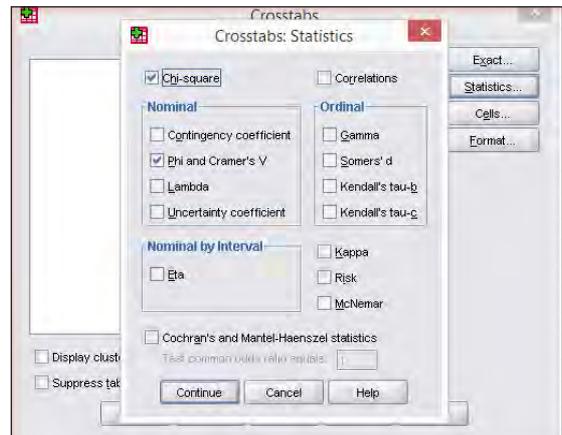
1. Enter the data in the spreadsheet as follows:

	Gender	Extension_participation
1	1.00	0.00
2	2.00	0.00
3	2.00	0.00
4	2.00	0.00
5	1.00	1.00
6	1.00	1.00
7	2.00	1.00
8	1.00	1.00
9	1.00	1.00
10	2.00	0.00
11	1.00	1.00
12	1.00	0.00
13	1.00	0.00
14	2.00	1.00
15	1.00	1.00
16	1.00	1.00
17	1.00	1.00
18	1.00	1.00
19	2.00	0.00
20	2.00	0.00

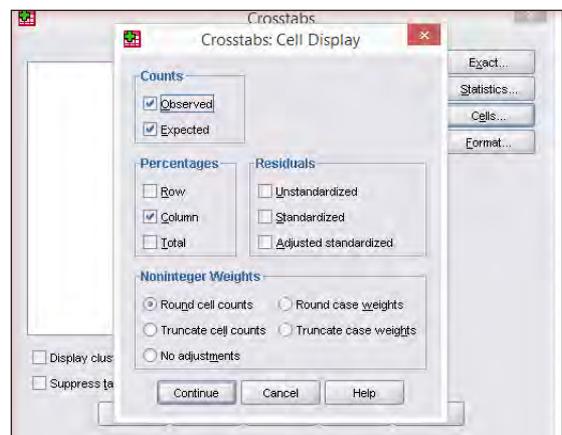
- Click on **Analyze** → **Descriptive Statistics** → **Crosstabs**. Move **Extension participation** to the Rows box using the arrow key and put **gender** in the **Columns** box.



- Next, click on **Statistics** and Check **Chi-square** and Phi and Cramer's V. Click on **Continue**.



- Once you return to the Crosstabs menu, click on **Cells**. Ensure that **Observed** is checked; check **Expected** under **Counts**, and **Column** under **Percentages**. It helps the interpretation if the total of the percentages of students for each level of the presumed independent variable (gender) adds up to 100%. As gender is the column variable, the Column is checked. Click on **Continue**, then on **OK**.



Output: Chi-square test

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Extension participation * Gender	20	100.0%	0	0%	20	100.0%

Extension participation * Gender Crosstabulation

			Gender		Total
			Male	Female	
Extension participation	Not participated	Count	0	6	6
		Expected Count	0.4	3.6	4.0
		% within Gender	25.0%	75.0%	85.0%
Participated	Count	6	5	11	
	Expected Count	6.6	4.4	11.0	
	% within Gender	75.0%	20.0%	35.0%	
Total	Count	12	8	20	
	Expected Count	12.0	8.0	20.0	
	% within Gender	100.0%	100.0%	100.0%	

Since the expected counts in 50% categories are less than 5, Fisher's exact test is appropriate.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4.848 ^a	1	.028		
Continuity Correction ^b	3.039	1	.081		
Likelihood Ratio	5.032	1	.025		
Fisher's Exact Test				.065	.040
Linear-by-Linear Association	4.606	1	.032		
N of Valid Cases	20				

Fisher's exact test is non-significant.

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 3.60.
 b. Computed only for a 2x2 table

Since the expected counts are less than 5 in two cells, Fisher's exact test is appropriate.

Symmetric Measures

	Value	Approx. Sig.
Nominal by Nominal - Phi	.492	.028
-Cramer's V	.492	.028
N of Valid Cases	20	

Since the values are computed for 2X2 table, Phi coefficient is selected to measure the effect size.

Box A-3.3: How to write the output of Chi-square test

Results

To investigate male and female farmers in terms of their participation or non-participation in extension programmes, a Chi-square analysis was conducted. The assumptions were met and the results are displayed in Table A-3.3.

Table A-3.3: Chi-square Analysis of extension participation or non-participation among male and female farmers

Variable	N	Gender		X ²	Fishers exact significance -two tailed (p)
		Male	Female		
Extension participation				4.848*	0.065
Participated	11	9	2		
Not participated	9	3	6		
Total		12	8		

* Significant at 5% level

Results displayed in the Table A-3.3 indicate that the Pearson Chi-square was significant. Since the expected counts are less than 5 in two cells, Fisher's exact test was used to check the significance. The non-significant Fisher's exact test indicate that we cannot be certain that males and females are different in terms of 'participating' or 'not-participating' in extension programmes.

C. One-Way (or Single Factor) ANOVA

- Purpose – Compare the means of the samples or groups in order to make inferences about the population means;
- Used for one independent variable with three or more levels. If more than one independent variable is involved, factorial ANOVA or higher versions are appropriate.

Assumptions

- Observations are independent;
- Variances on the dependent variable are equal across groups;
- The dependent variable is normally distributed for each group.

Example

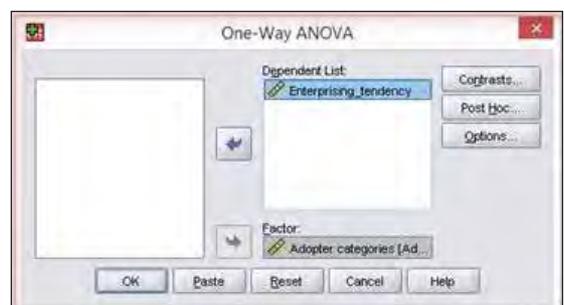
Are there any differences among grape farmers' adopter categories (fully adopted, partially adopted or non-adopted) with regard to their enterprising tendency?

SPSS Analytical procedure

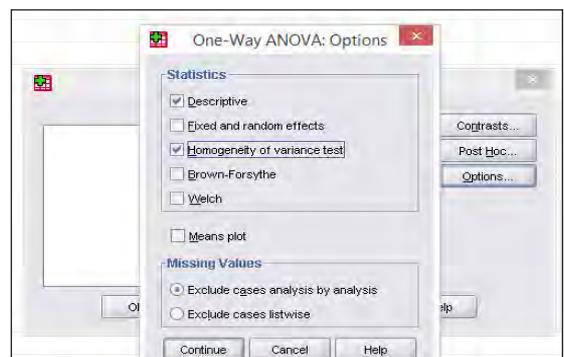
1. Enter the data in the spreadsheet as follows:

	Adopter_cate gories	Enterprising_t endency	VAR00001	VAR	VAR	VAR	VAR	VAR	VAR
1	1.00	24.00							
2	2.00	32.00							
3	1.00	18.00							
4	3.00	45.00							
5	3.00	52.00							
6	3.00	36.00							
7	2.00	31.00							
8	2.00	38.00							
9	1.00	18.00							
10	1.00	15.00							
11	1.00	16.00							
12	2.00	31.00							
13	2.00	42.00							
14	3.00	56.00							
15	1.00	12.00							
16	3.00	65.00							
17	3.00	51.00							
18	3.00	49.00							
19	3.00	62.00							
20	3.00	54.00							
21									
22									
23									
24									

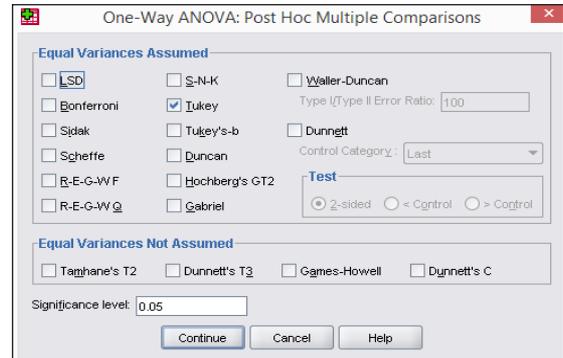
2. Click **Analyze** → **Compare Means** → **One-Way ANOVA**. Move Enterprising tendency into the Dependent List box and Click on Adopter categories and move it to the Factor (independent variable) box.



3. Click on **Options**. Under **Statistics**, choose **Descriptive** and **Homogeneity of variance test** (To check the assumption that variances were equal).



- Then, in the main dialogue box, click on **Post Hoc**. Check **TUKEY** because it is the appropriate test when group variances are equal (ANOVA assumption). **Continue** and then **OK** to run the analysis.



Output of One-way ANOVA test

Descriptives

Enterprising_tendency	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Not adopted	6	17.1867	4.02078	1.84148	12.9471	21.3862	12.00	24.00
Partially adopted	5	34.8000	4.88991	2.22261	29.6290	40.9710	31.00	42.00
Fully adopted	9	52.2222	8.68588	2.89529	45.5457	58.8988	36.00	65.00
Total	20	37.3500	16.62362	3.71715	29.5699	45.1301	12.00	65.00

Test of Homogeneity of Variances

Enterprising_tendency	Levene Statistic	df1	df2	Sig.
	1.255	2	17	.310

ANOVA

Enterprising_tendency	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4467.361	2	2233.681	48.485	.000
Within Groups	783.189	17	46.070		
Total	5250.550	19			

Comparison of means for each adopter level

The non-significant Levene statistic indicates that variances in all three groups are the same and data is suitable for ANOVA (Testing of assumptions)

The between-group differences in the enterprising tendency is significant

Post Hoc Tests

Multiple Comparisons

Enterprising_tendency
Tukey HSD

(I) Adopter categories	(J) Adopter categories	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Not adopted	Partially adopted	-17.63333*	4.11003	.001	-28.1770	-7.0897
	Fully adopted	-35.05556*	3.57732	.000	-44.2327	-25.8785
Partially adopted	Not adopted	17.63333*	4.11003	.001	7.0897	28.1770
	Fully adopted	-17.42222*	3.78588	.001	-27.1343	-7.7101
Fully adopted	Not adopted	35.05556*	3.57732	.000	25.8785	44.2327
	Partially adopted	17.42222*	3.78588	.001	7.7101	27.1343

*. The mean difference is significant at the 0.05 level.

Mean differences in enterprising tendency among adopter categories are significantly different

Box A-3.4: How to write the output of One-way ANOVA

Results

A statistically significant difference was found in the enterprising tendency of three adopted categories of grapes technologies [$F(2,17) = 48.49, p=.000$](Table A-3.4).

Table A-3.4: One way ANOVA on mean differences in enterprising tendency among adopter categories of grapes technologies

Grape technology adopter categories	n	Mean	SD	F	P
Fully adopted	9	52.22	8.69	48.49**	.000
Partially adopted	5	34.80	4.97		
Not adopted	6	17.17	4.02		

**Significant at 1% level

Results displayed in the Table A-3.4 indicate that the farmers who fully adopted improved grape technologies had highest enterprising tendency levels ($M=52.22$; $SD = 8.69$), followed by partially adopted farmers ($M=34.80$; $SD=4.97$), than non-adopters ($M=17.17$; $SD=4.02$). Post-hoc Tukey's HSD tests indicated that all categories of adopted differed significantly from others ($p<0.01$; Tukey's HSD).

References

- Cohen, J. (1988.) Statistical power analysis for the behavioral sciences. Routledge.
- Morgan, G. A., Leech, N. L., Gloeckner, and Barrett, K. C. (2011.) IBM SPSS for introductory statistics: Use and interpretation. Fourth edition. New York: Routledge.
- Sawilowsky, S. (2009.) New effect size rules of thumb. Journal of Modern Applied Statistical Methods. 8 (2):467–474.

Annexure A-4

Guidelines for avoiding biases in selecting items

When selecting the items, the following aspects may be considered to avoid biases:

1. Avoid redundant items.
2. Avoid exceptionally long items.
3. Keep the level of reading difficulty appropriate for potential respondents. Using Fry's readability graph (Fry, 1977) will help to customise the reading level to the respondent's capacity.
4. Avoid 'double-barrelled' items that convey two or more ideas at the same time. For example, consider an item that asks the respondent to agree or disagree with the statement, "I am adopting nano-pesticides because I am always innovative in doing things". There are two different statements here with which the respondent could agree: "I am adopting nano-pesticides" and "I am always innovative in doing things."
5. Use a mixture of positively- and negatively-worded items to avoid acquiescence or affirmation biases
 - Acquiescence or affirmation bias – Tendency of respondents to agree with all items irrespective of their content.
6. Avoid ambiguous pronoun references – such as a pronoun with multiple, possible antecedents. For example, a statement is worded as "I always seek their advice in choosing improved tomato varieties for my farm". In this statement, the word 'their' is an ambiguous pronoun as it may indicate multiple people – fellow farmer, local leader, extension agent, etc.
7. Avoid mixing items that assess behaviours with items that assess affective responses to or outcomes of behaviours. For example, in the examination of supervisory behaviour: "My supervisor treats me fairly", should not be included in a scale with the outcome statement "I feel committed to my supervisor".

Source: Harrison and McLaughlin (1993); DeVellis (2003)

Methods for assessing content adequacy

1. Judges' rating

Providing statements or items to experts along with construct definitions and asking them to match items with a corresponding definition.

2. Lawshe's Content Validity Ratio (Lawshe, 1975)

- Basis – The content experts' agreement on the 'essentiality' of the items or statement in measuring a specific construct.
- Each of the subject matter expert raters in the judging panel is asked a question for each statement

or item: “Is the skill or knowledge measured by this item ‘essential,’ ‘useful, but not essential,’ or ‘not necessary’ to the performance of the construct?”.

- Items deemed ‘essential’ by a critical number of panel members are then included within the final instrument, with items that fail to achieve this critical level discarded.
- If more than half of the panelists indicate that an item is essential, that item has at least some content validity. Greater levels of content validity accrue to an item when a larger numbers of panelists agree that a particular item is essential.
- Using these assumptions, Lawshe developed a formula termed the Content Validity Ratio (CVR) to compute agreement level for each item.

$$CVR = (ne - n/2)/(n/2)$$

Where ne = number of panellists indicating ‘essential’; N = total number of panellists.

- The CVR values range between -1 (perfect disagreement) and +1 (perfect agreement) with CVR values above zero indicating that over half of the panel members agree that an item is essential.
- After calculating CVR for each item, this value is compared with the minimum value as prescribed in the CVR table (Table A-4.1). If the calculated CVR exceeds the CVR critical value given in the table, the item is retained.

Table A-4.1: Minimum Values of CVR critical for One Tailed Test (p = .05)

Panel size	CVR critical
5	1.000
6	1.000
7	1.000
8	0.750
9	0.778
10	0.800
11	0.636
12	0.667
13	0.538
14	0.571
15	0.600
20	0.500
25	0.440
30	0.333
35	0.314
40	0.300

Source: Ayre and Scally (2013)

3. Item-objective congruence index

The item-objective congruence procedure was proposed by Rovinelli and Hambleton (1977) as a criterion-referenced item analysis. This procedure estimates an index of the validity of an item based on the ratings of two or more content specialists. A detailed procedure for estimating item-objective congruence is described in Sivakumar and Sulaiman (2015).

4. Q sort methodology

- The Q sort method described elsewhere in this manual can also be used to identify items corresponding to given constructs.
- Various experts or stakeholders can sort the items into constructs based on their view points.
- A modified form of Q sort combined with factor analysis developed by Schriesheim et al. (1993) is a sound alternative method for content adequacy assessment.

5. Analysis of Variance method (Hinkin et al., 1997)

- Combination of expert rating and ANOVA method;
- It can be conducted with a relatively small sample and has a very low cost – both in terms of time and money.

Steps

1. An item pool with a relatively small number of items (<100) is generated.
2. The items are sorted into four different versions of the same questionnaire by rearranging questions and instructions in each questionnaire.
3. The experts are provided with the questionnaires, sequentially one-by-one, and asked to rate the item's consistency with each of the identified dimensions/constructs, using a 5-point response scale (1 = not at all consistent to 5 = completely consistent).
4. The mean scores of all items in the identified dimension/construct is calculated and subjected to ANOVA in order to identify if the specific item has a significantly higher score compared to others on a particular dimension.
5. The items with significantly higher score on a specific dimension is sorted into it.

References

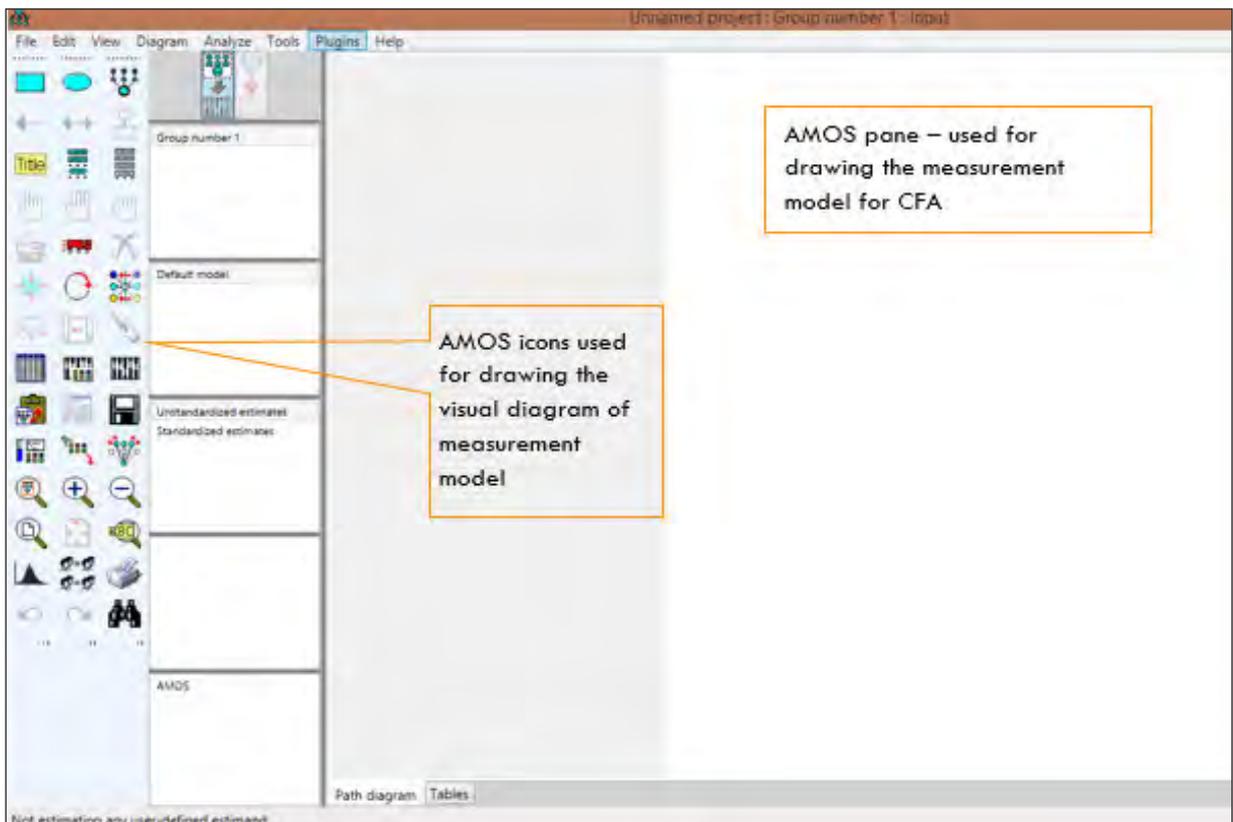
- Ayre, C., and A. J. Scally. (2014.) Critical values for Lawshe's content validity ratio revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development*, 47(1), 79–86.
- DeVellis, R. F. (2003.) *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: Sage.
- Fry E. (1977.) 'Fry's readability graph: clarifications, validity, and extensions to level 17'. *Journal of Reading*, 21,242-252.
- Harrison, D.A. and McLaughlin, M.E. (1993.) Cognitive processes in self-report responses: Tests of item context effects in work attitude measures. *Journal of Applied Psychology*, 78,129-140.
- Hinkin, T. R., Tracey, J. B., and Enz, C. A. (1997.) Scale construction: Developing reliable and valid measurement instruments. *Journal of Hospitality & Tourism Research*, 21(1), 100-120.
- Lawshe, C.H. (1975.) A quantitative approach to content validity. *Personnel Psychology*, 28, 563–575
- Rovinelli, R. J., and Hambleton, R. K. (1977.) On the use of content specialists in the assessment of criterion-referenced test item validity. *Dutch Journal of Educational Research*, 2, 49-60.
- Schriesheim, C.A., Powers, K.J., Scandura, T.A., Gardiner, C.C. and Lankau, M.J. (1993.) Improving construct measurement in management research: Comments and a quantitative approach for assessing the theoretical adequacy of paper-and-pencil survey-type instruments. *Journal of Management*, 19, 385-417.
- Sivakumar, P.S., and Rasheed Sulaiman V. (2015.) Extension Research in India-Current Status and Future Strategies. AESA Working Paper 2. Agricultural Extension in South Asia. (Available at: <http://www.aesa-gfras.net/admin/kcfinder/upload/files/Final%20one%20WP2.pdf>).

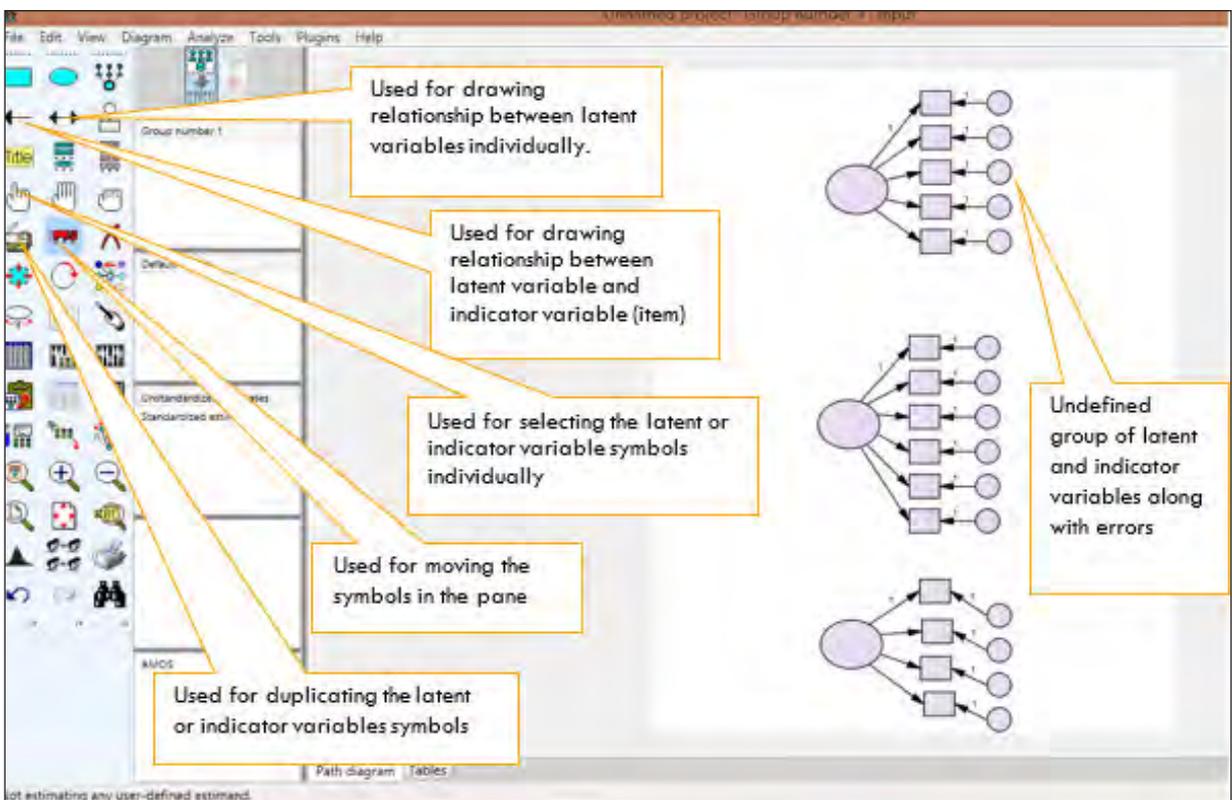
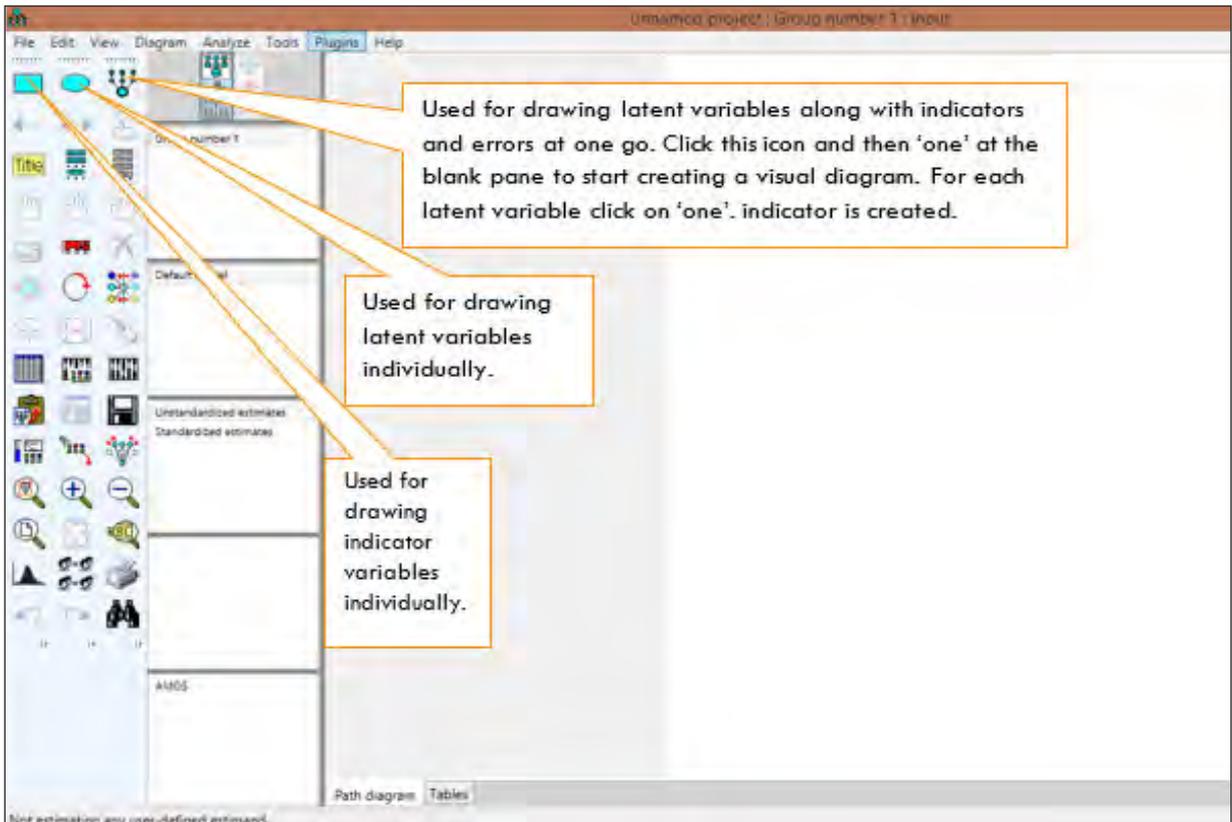
Annexure A-5

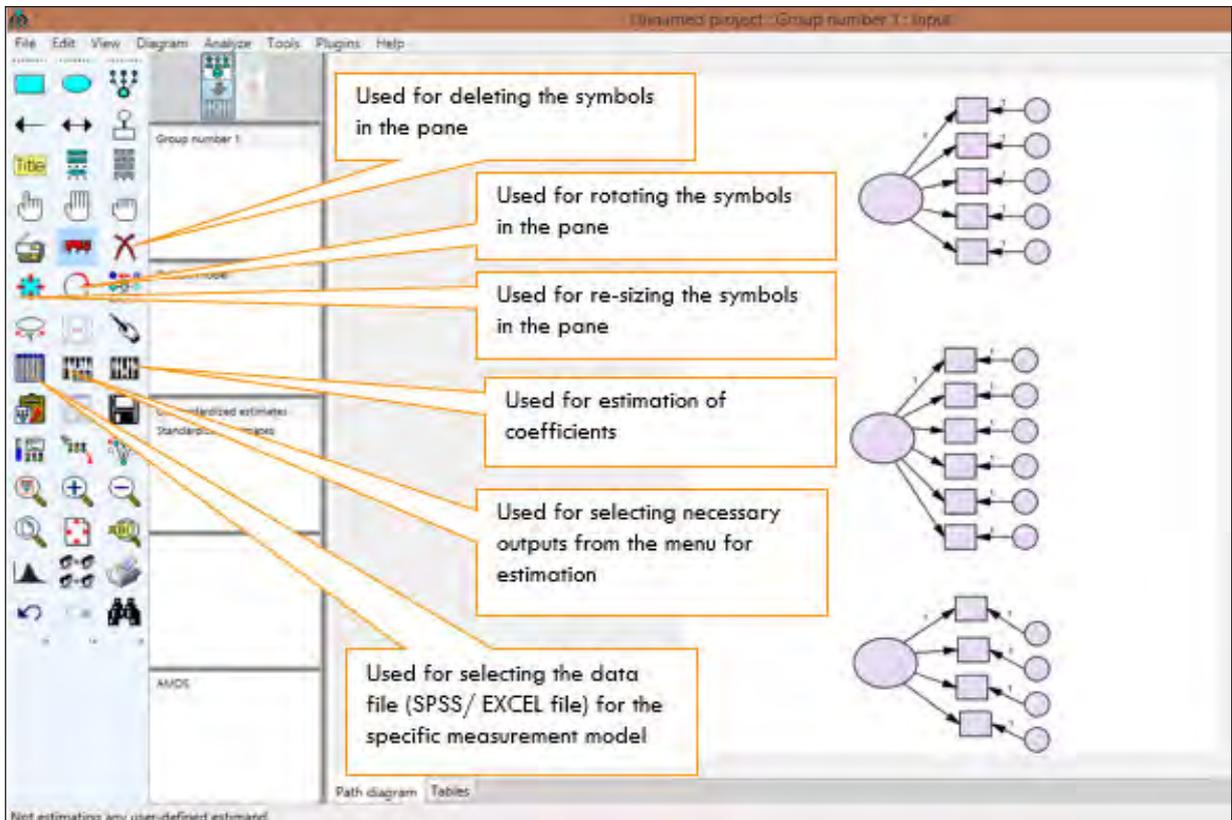
IBM SPSS AMOS

- Easy-to-use program for visual SEM;
- Helps to quickly specify, view, and modify the model graphically using simple drawing tools;
- Prints a high quality image of the final model;
- All Visual diagrams are created using AMOS GRAPHICS.

IBM SPSS AMOS Interface

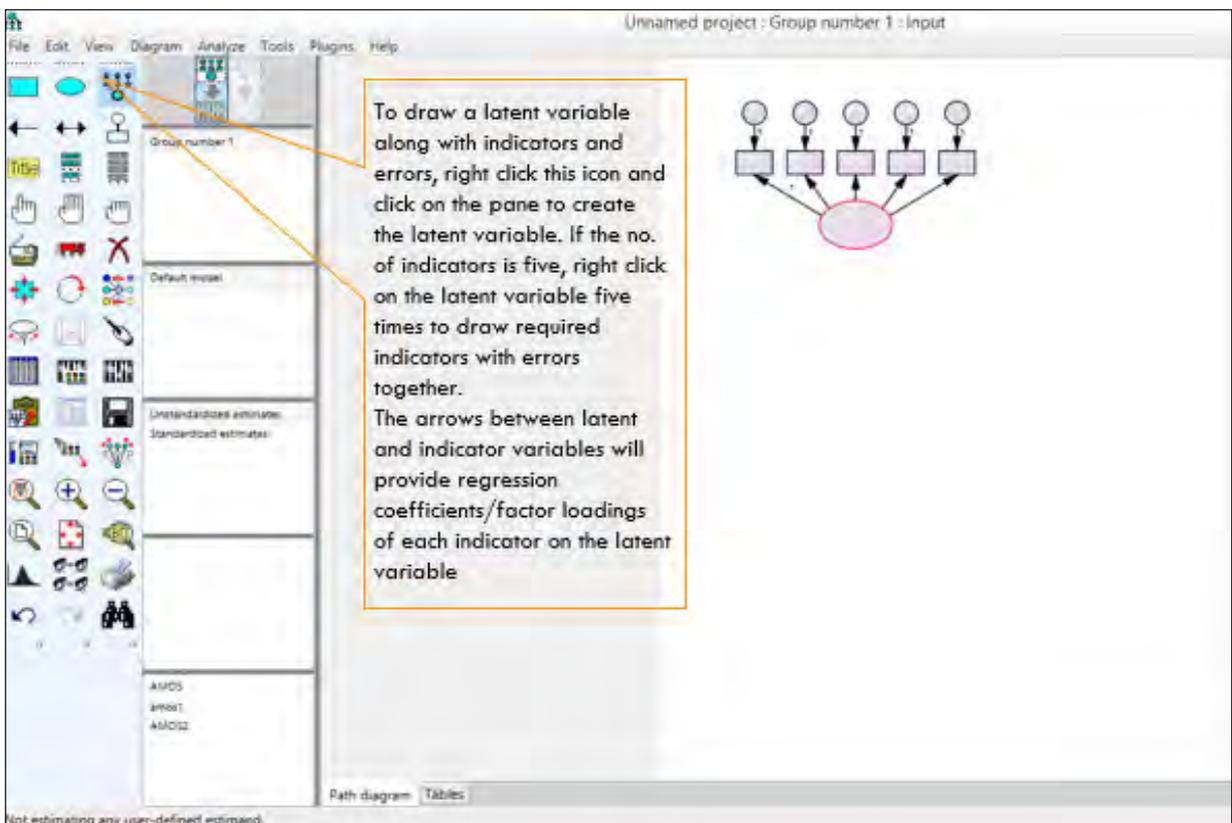


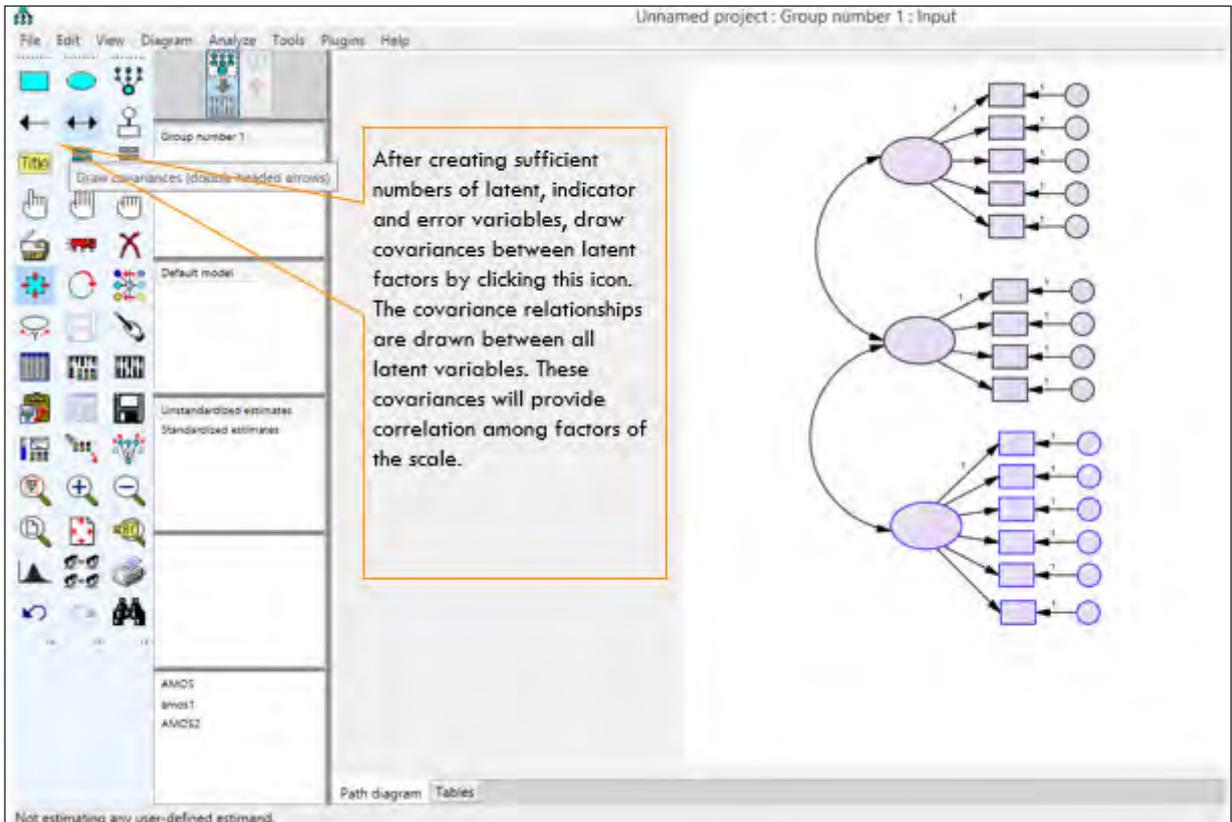




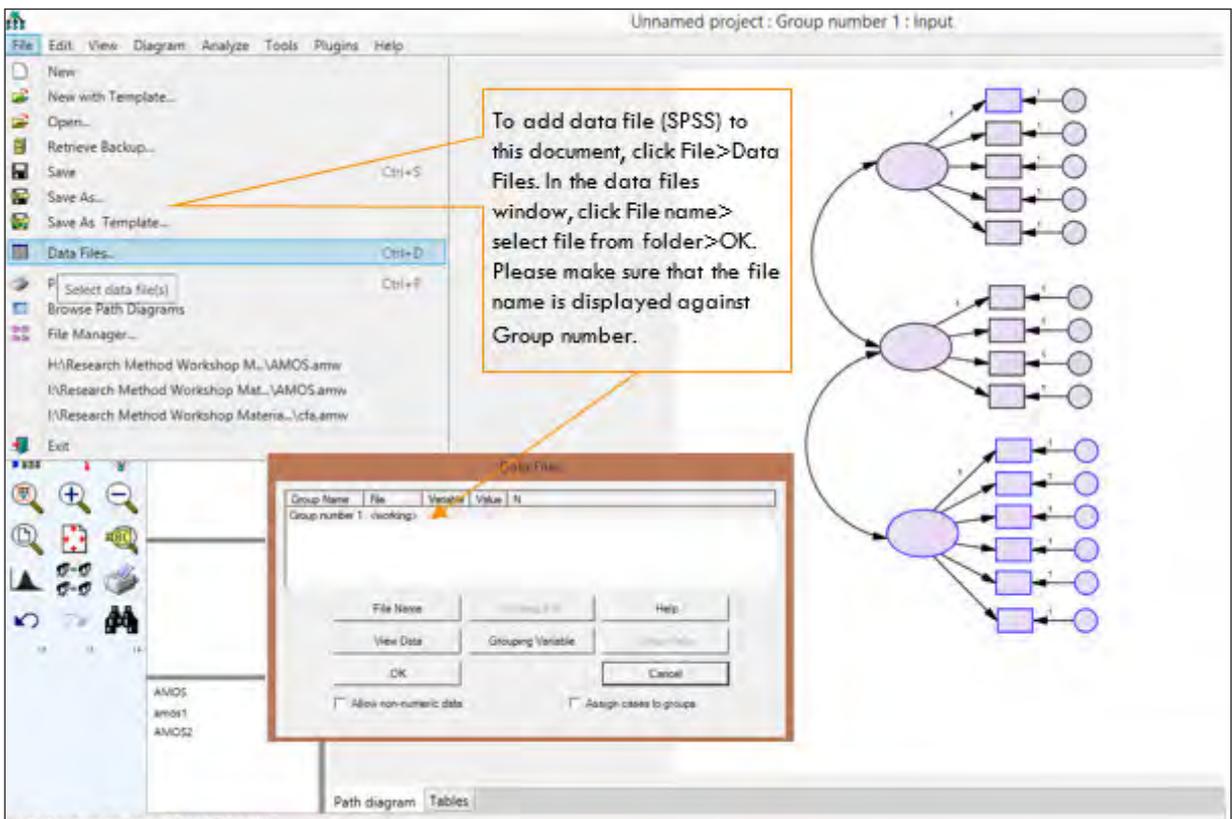
Steps in using IBM SPSS AMOS for conducting CFA

1. Draw measurement model using AMOS Graphics

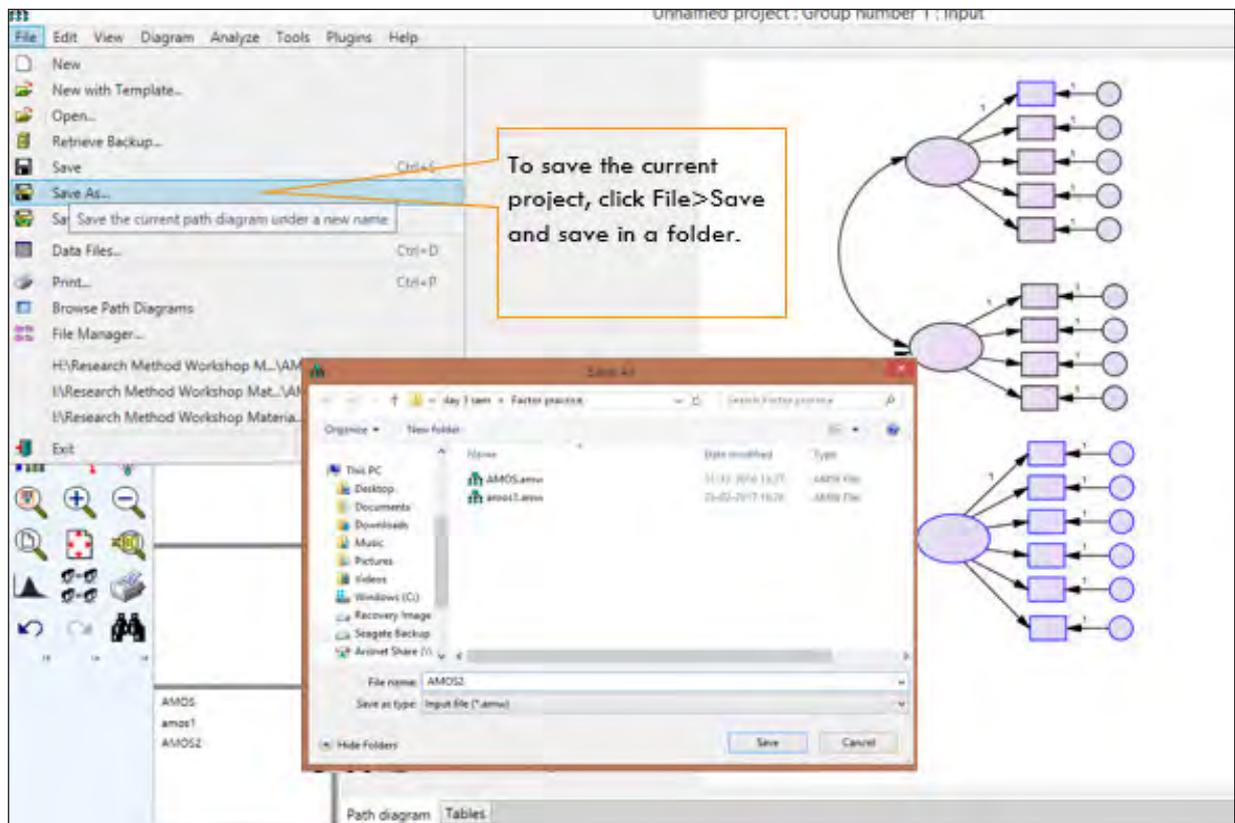




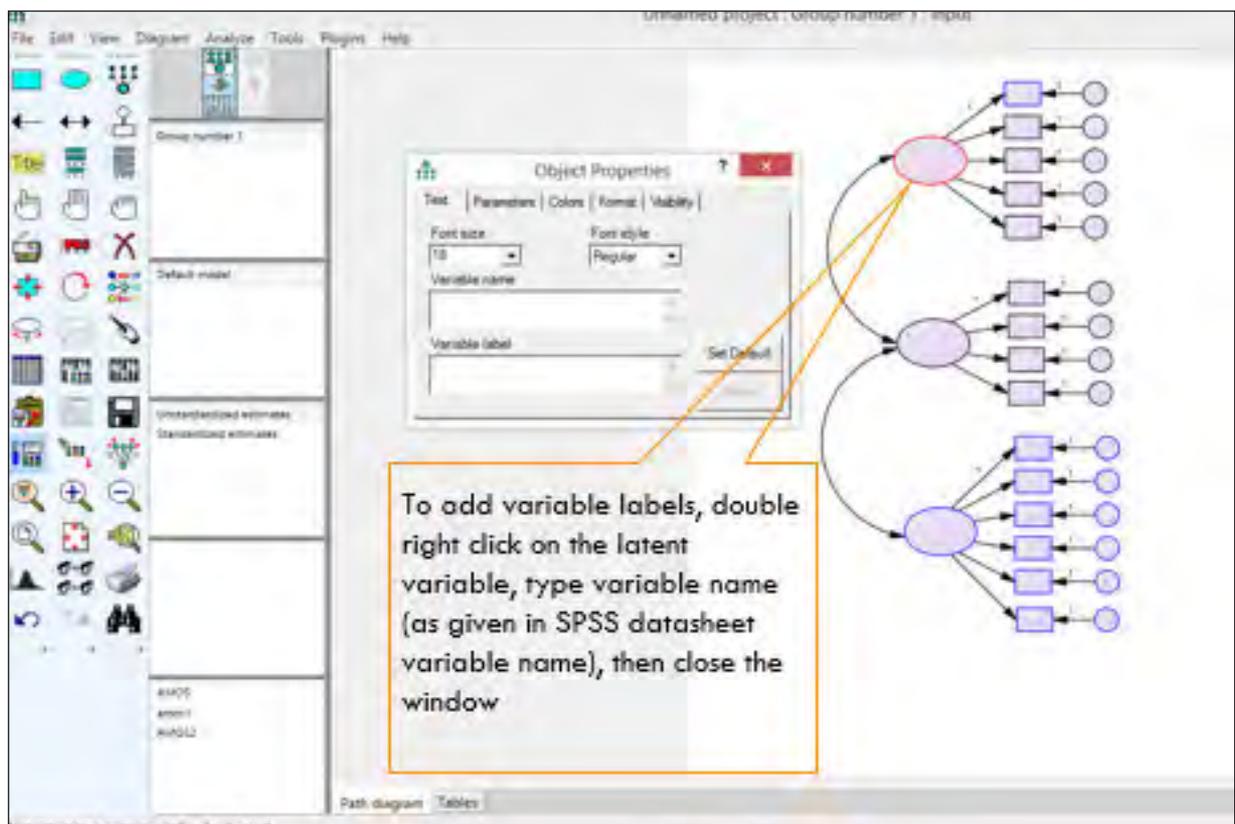
2. Add data file to the AMOS project



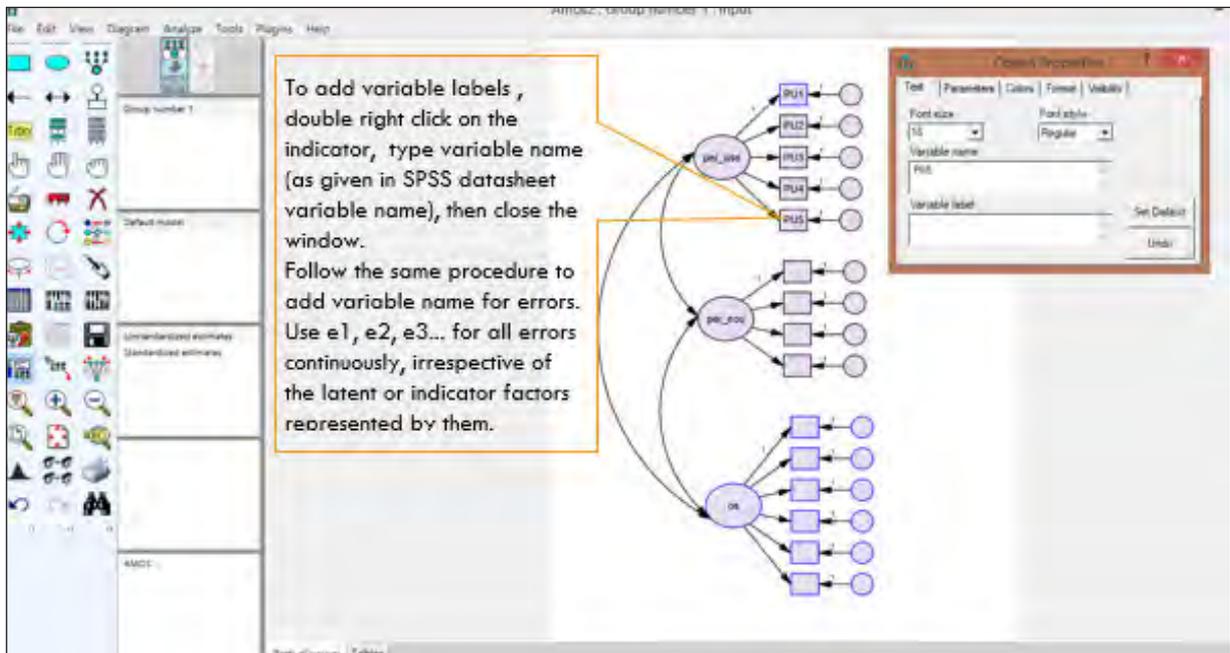
3. Saving the current project



4. Labelling the latent variables



5. Labelling the indicator and error variables



The screenshot shows the AMOS software interface. On the left is a toolbar with various icons. The main area displays a path diagram with three latent variables (ovals) labeled 'psi_1', 'psi_2', and 'psi_3'. 'psi_1' is connected to five indicator variables (squares) labeled 'PI1' through 'PI5'. 'psi_2' is connected to three indicator variables labeled 'PI6' through 'PI8'. 'psi_3' is connected to three indicator variables labeled 'PI9' through 'PI11'. A text box with an orange border is overlaid on the diagram, containing the following text:

To add variable labels, double right click on the indicator, type variable name (as given in SPSS datasheet variable name), then close the window. Follow the same procedure to add variable name for errors. Use e1, e2, e3... for all errors continuously, irrespective of the latent or indicator factors represented by them.

On the right, a 'Change Properties' dialog box is open, showing fields for 'Variable name' and 'Variable label'. The 'Variable name' field contains 'PI1' and the 'Variable label' field is empty. The dialog has tabs for 'Text', 'Parameters', 'Colors', 'Format', and 'Visibility'.

6. Choosing the estimation procedure and outputs

- Mathematical algorithm that will be used to identify estimates for each free parameter (Hair Jr. et al., 2006).

a. Maximum likelihood method

- Provides 'most likely' parameter values to achieve best fit of the model;
- Sensitive to multivariate normality;
- Widely used in CFA.

b. Weighted least squares

c. Generalised least squares

d. Asymptotically distribution free (ADF)

- Insensitive to non-normality;
- Requires a large sample.

AMOS procedures for analysis

Click this icon to choose outputs and coefficients for analysis

In the Analysis properties window, click on the Maximum Likelihood method

Click the Output box and choose standardised estimates, squared multiple correlations and tests for normality and outliers and close the window

AMOS Output

Viewing text output of all estimates

AMOS output provides normality testing coefficients and two sets of estimates for assessing the model.

1. Goodness-of-fit indices – The test of the proposed model is same as the theoretical model;
2. Reliability and validity of the scale.

1. Tests of Multivariate normality

Mardia's multivariate kurtosis

- The Mardia's (1970) coefficient measures the multivariate kurtosis which indicate the presence or absence of multivariate normality;
- Interpretation – Very small multivariate kurtosis values (e.g., less than 1.00) are considered negligible while values ranging from one to ten often indicate moderate non-normality. Values that exceed ten indicate severe non-normality (Mardia, 1970).

1. Goodness-of-Fit indices

(i) Absolute fit indices

- Determine how well an a priori model fits the sample data (McDonald and Ho, 2002);
- Indicates how well the proposed theory fits the data;
- Absolute fit indices – Chi-Squared test, RMSEA, GFI, AGFI, the RMR and the SRMR (Table A-5.1).

(ii) Incremental fit indices

- Incremental fit indices, also known as comparative (Miles and Shevlin, 2007) or relative fit indices (McDonald and Ho, 2002), are a group of indices that do not use the chi-square in its raw form but compare the chi-square value to a baseline model;
- For these models the null hypothesis is that all variables are uncorrelated (McDonald and Ho, 2002);
- Indices – TLI, NFI, IFI, CFI, RFI, AGFI.

(iii) Parsimonious fit measures

- Diagnoses whether model fit is due to over fitting the data with too many coefficients;
- Indices – Normed Chi-square, PGFI, PNFI.

The cut-off values for the Goodness-of-fit coefficients are displayed in Table A-5.1. As a thumb rule, if three or more goodness-of-fit measures are in the acceptable range, the model is accepted.

Table A-5.1: Criteria for model fit assessment, item validity and reliability

Test	Guideline	Reference	
Absolute fit measures:			
Likelihood ratio Chi-square (χ^2)	H0: $\Sigma = \Sigma(\theta)$ HA: $\Sigma = \Sigma\alpha$	Small, Insignificant χ^2 ($p > 0.05$)	Joreskog (1996)
Root Mean Square Residual (RMR)	Average residuals between observed and estimated input matrices.	RMR < 0.07; <0.03 – excellent fit	Steiger (2007)
Goodness-of-Fit Index (GFI)	Overall degree of fit of the squared residuals from prediction compared with the actual data). Less influenced by sample size and normality	GFI > 0.95	Tabachnick and Fidell (2007)

Test		Guideline	Reference
Root mean square residual (RMSEA) and standardised root mean square residual (SRMR)	The RMR and the SRMR are the square root of the difference between the residuals of the sample covariance matrix and the hypothesised covariance model.	RMSEA \leq 0.06 or SRMR \leq 0.09	Tabachnick and Fidell (2007); Hu and Bentler (1999)
Incremental or Comparative fit measures			
Tucker-Lewis Index (TLI) or NNFI	A comparative index between the proposed and the null model	TLI > 0.90	Hair et al. (1998)
Normed Fit Index (NFI)	A relative comparison of the proposed model to the null model. [$\chi^2_{\text{null}} - \chi^2_{\text{proposed}}$]/ χ^2_{null}	NFI > 0.90	Bentler and Bonnet (1980)
Incremental Fit Index (IFI)	$(\chi^2_{\text{indep}} - \chi^2_{\text{model}}) / (\chi^2_{\text{indep}} - df_{\text{model}})$	IFI = Higher values close to 1	Kelloway (1998)
Comparative Fit Index (CFI)	Estimated based on non-central χ^2 distribution. Calculated by $1 - [(\chi^2_{\text{model}} - df_{\text{model}}) / (\chi^2_{\text{indep}} - df_{\text{indep}})]$	CFI > 0.90	Kelloway (1998)
Relative Fit Index (RFI)	$(\chi^2_{\text{indep}} - \chi^2_{\text{model}}) - [df_{\text{indep}} - (df_{\text{model}}/n)] / [\chi^2_{\text{indep}} - (df_{\text{indep}}/n)]$	RFI > 0.90	Kelloway (1998)
Adjusted Goodness-of-Fit Index (AGFI)	Goodness-of-fit adjusted by degrees of freedom (df).	AGFI > 0.90	Hair et al. (1998)
Parsimonious fit measures			
Normed chi-square	χ^2/df <ul style="list-style-type: none"> • Alternate to Chi-square • Used when Chi-square of the model is significant 	2.0 to 5.0	Wheaton et al. (1977); Tabachnick and Fidell (2007).
Parsimonious Goodness-of-Fit Index (PGFI)	$1 - (P/N) \times \text{GFI}$. P = Number of estimated parameters in the model, N = Number of data points. Adjusts GFI for the degrees of freedom in the model.	PGFI = Higher values close to 1	Kelloway (1998)
Parsimonious Normed Fit Index (PNFI)	$(df_{\text{model}}/df_{\text{indep}}) \times \text{NFI}$. Adjusts NFI for model parsimony	PNFI = Higher values close to 1	Kelloway (1998)

2. Reliability and validity

Convergent validity

Assessed through standardized loadings (λ), composite reliabilities (squared multiple correlations), and variance extracted by each latent factor.

Cut-off values

- Standardised regression coefficient for each indicator variable : $\lambda > 0.6$ (Bollen, 1989);
- Composite reliability – squared multiple correlation for each indicator variable : $R^2 \geq 0.50$ (Bagozzi and Yi, 1988);
- Variance extracted by each latent factor – ≥ 0.50 (Fornell and Larcker, 1981).

Composite reliability

- Squared multiple correlations (R^2) are used as a measure of reliability of each Indicator variable;
- Used to assess the amount of variation in latent variables explained by predictors;
- $R^2 > 0.5$ (Bagozzi and Yi, 1988).

AMOS – Viewing Text outputs

The screenshot displays the AMOS software interface. The 'View' menu is open, and the 'Text Output' option is highlighted. A callout box with an orange border points to this option, containing the text: "Click the 'Text Output' to view the estimates of the analysis". The main workspace shows a path diagram with three latent variables: 'per_use', 'per_eou', and 'os'. 'per_use' is measured by indicators PU1 through PU5. 'per_eou' is measured by indicators PEU1 through PEU4. 'os' is measured by indicators OS1 through OS6. Each indicator has an associated error term (e1 through e15). The diagram shows standardized path coefficients and R-squared values for each latent variable. The status bar at the bottom indicates 'Not estimating any user-defined estimand.'

(i) Normality testing - Mardia's multivariate kurtosis

Amos2 : Group number 1 : OK: Default model

Amos Output

Assessment of normality (Group number 1)

Variable	min	max	skew	c.r.	kurtosis
OS6	1.000	5.000	-1.042	-7.342	2.088
OS5	1.000	5.000	-.870	-6.130	.388
OS4	1.000	5.000	-.502	-3.537	-.253
OS3	1.000	5.000	-.959	-6.761	.679
OS2	1.000	5.000	-.496	-3.496	-.341
OS1	1.000	5.000	-1.371	-9.662	2.220
PEU4	2.000	5.000	-.883	-6.225	1.521
PEU3	2.000	5.000	-.545	-3.844	.638
PEU2	2.000	5.000	-.911	-6.417	1.637
PEU1	1.000	5.000	-1.010	-7.115	1.799
PU5	1.000	5.000	-1.219	-8.592	3.525
PU4	1.000	5.000	-1.006	-7.090	2.241
PU3	1.000	5.000	-1.394	-9.821	3.302
PU2	1.000	5.000	-.903	-6.362	1.925
PU1	2.000	5.000	-.805	-5.670	1.105
Multivariate					83.021

Mardia's multivariate kurtosis value exceeds 10, which is an indicator of severe non-normality in the data. In this case, the Chi-square value will be inflated. To correct this, bootstrapping can be used.

(ii) Assessing Goodness-of-fit measures

Amos2 : Group number 1 : OK: Default model

Amos Output

Notes for Model (Default model)

Computation of degrees of freedom (Default model)

Number of distinct sample moments: 120
 Number of distinct parameters to be estimated: 33
 Degrees of freedom (120 - 33): 87

Result (Default model)

Minimum was achieved
 Chi-square = 179.941
 Degrees of freedom = 87
 Probability level = .000

Click the 'Notes of model' to view Chi-square Goodness-of-fit

The probability higher than 0.05 indicates the model is significantly different from theoretical model. Ideally, the Chi-square should be non-significant. However, we can look for other

AMOS Output

Model Fit Summary

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	33	179.941	87	.000	2.068
Saturated model	120	.000			
Independence model		105	.000		15.609

CMIN

Model	RMR	GFI	AGFI	PGFI
Default model	.035	.926	.898	.672
Saturated model	.000	1.000		
Independence model	.175	.444	.365	.389

Baseline Comparisons

Model	NFI	RFI	IFI	TLI	CFI
Default model	.890	.867	.940	.927	.926
Saturated model	1.000	1.000	1.000	1.000	1.000
Independence model	.000	.000	.000	.000	.000

Parimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.829	.738	.778
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	92.941	58.340	135.309
Saturated model	.000	.000	.000
Independence model	1533.996	1406.615	1668.761

FMN

Model	FMN	F0	LO 90	HI 90
Default model	.606	.313	.196	.000
Saturated model	.000	.000	.000	.000
Independence model	5.519	5.165	4.422	5.619

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.060	.048	.072	.091
Independence model	.222	.212	.231	.000

P value is significant indicating the model is not fitting the data. However, normed Chi-square CMIN/DF is less than 5 indicating an acceptable value

RMR ≤ 0.05; IFI & PGFI = 0.90; IFI - Close to one indicates good fit of the model.

AMOS Output

Parimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.829	.738	.778
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	92.941	58.340	135.309
Saturated model	.000	.000	.000
Independence model	1533.996	1406.615	1668.761

FMN

Model	FMN	F0	LO 90	HI 90
Default model	.606	.313	.196	.000
Saturated model	.000	.000	.000	.000
Independence model	5.519	5.165	4.422	5.619

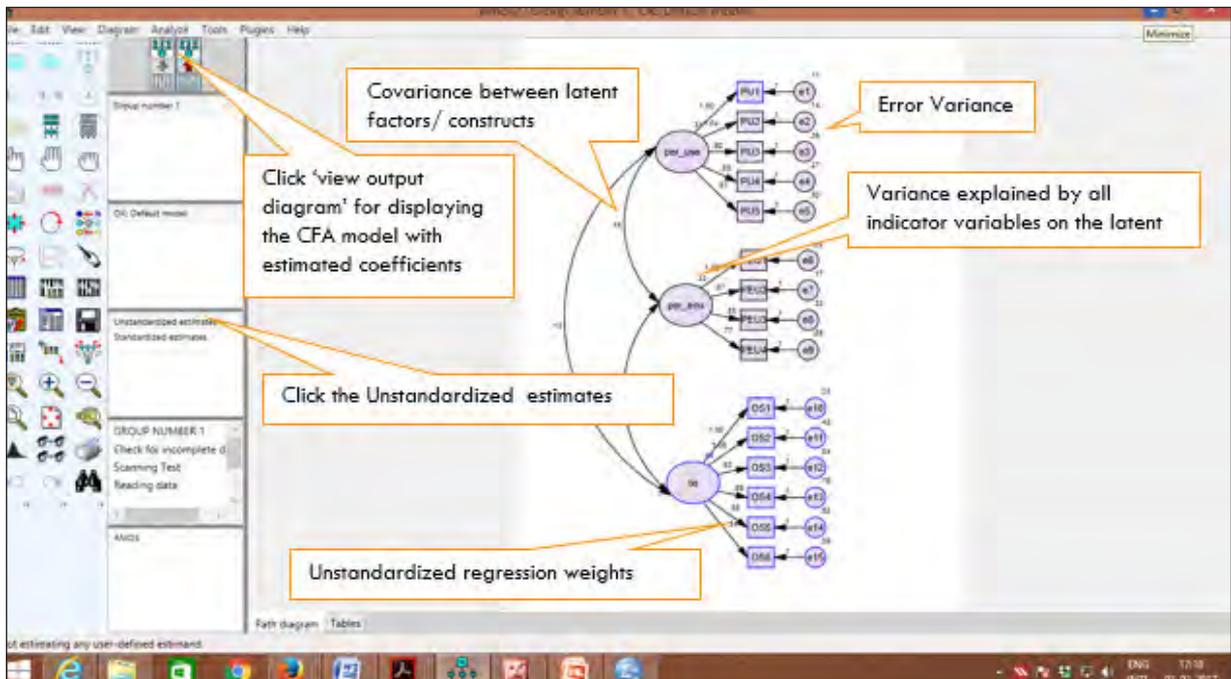
RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.060	.048	.072	.091
Independence model	.222	.212	.231	.000

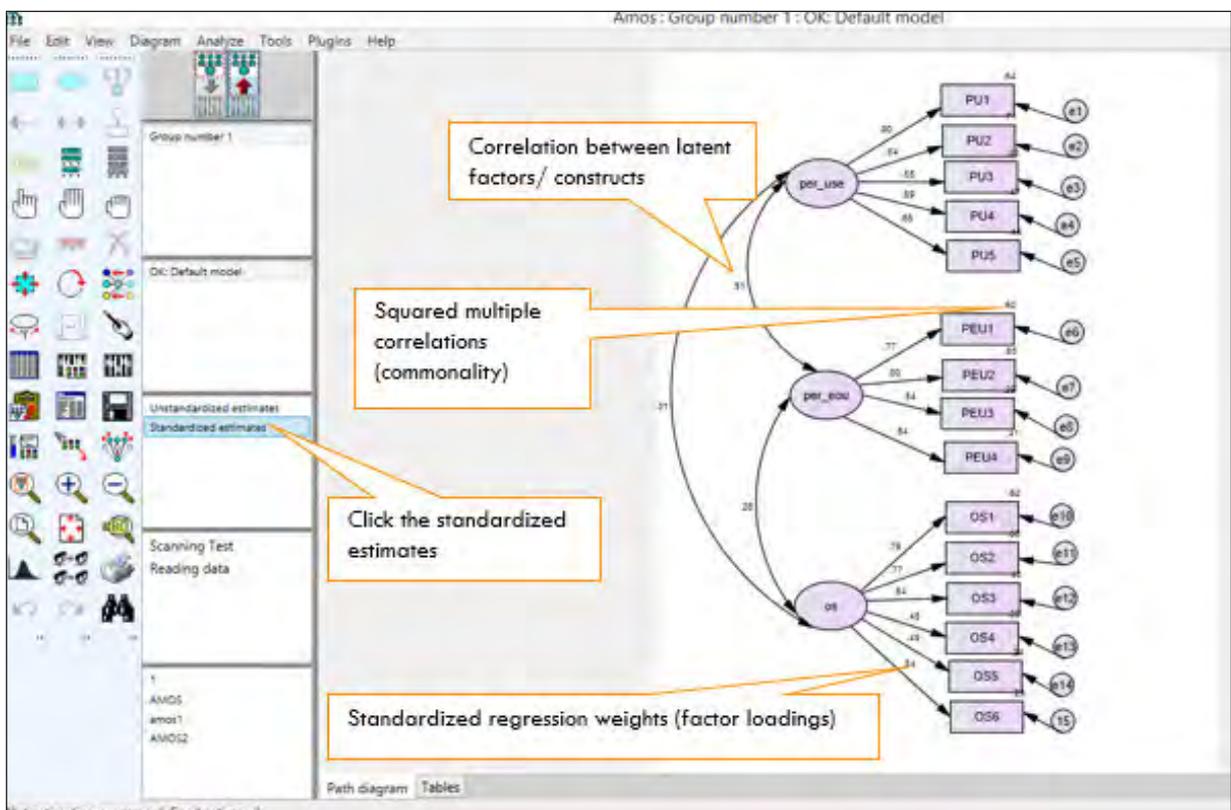
RMSEA = 0.60 indicates good fit of

(iii) Reliability and validity

Unstandardized estimates (In CFA model)

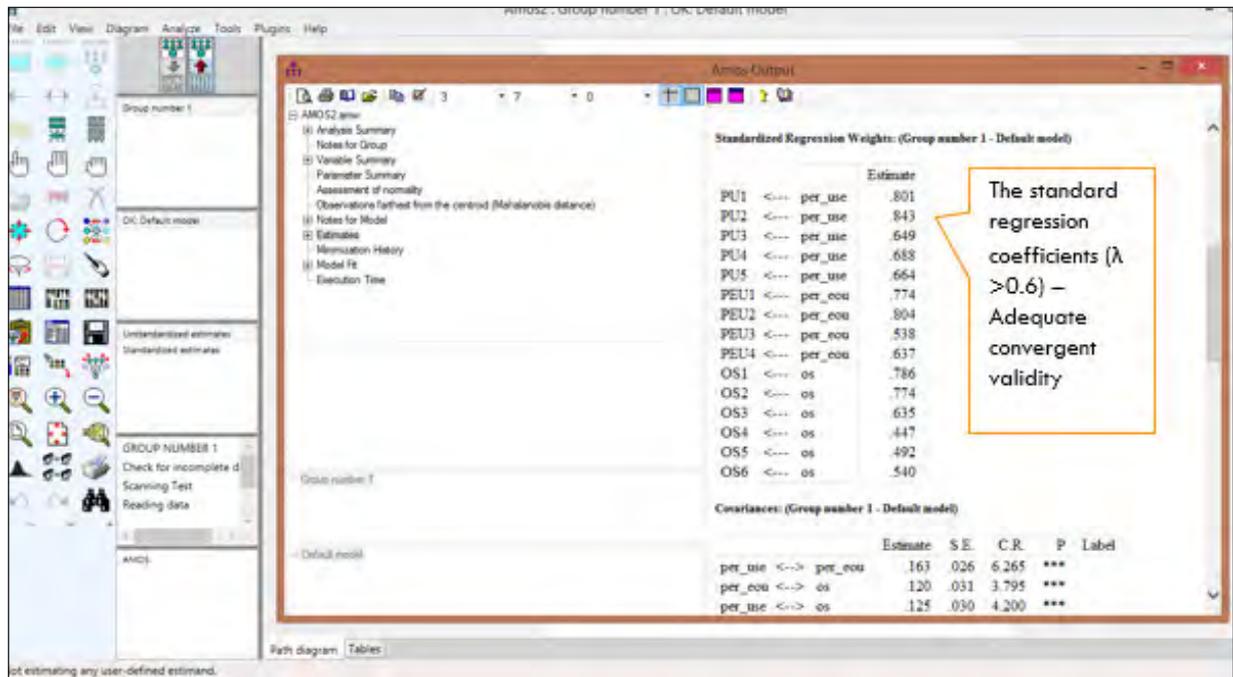


Standardised estimates (In CFA model)

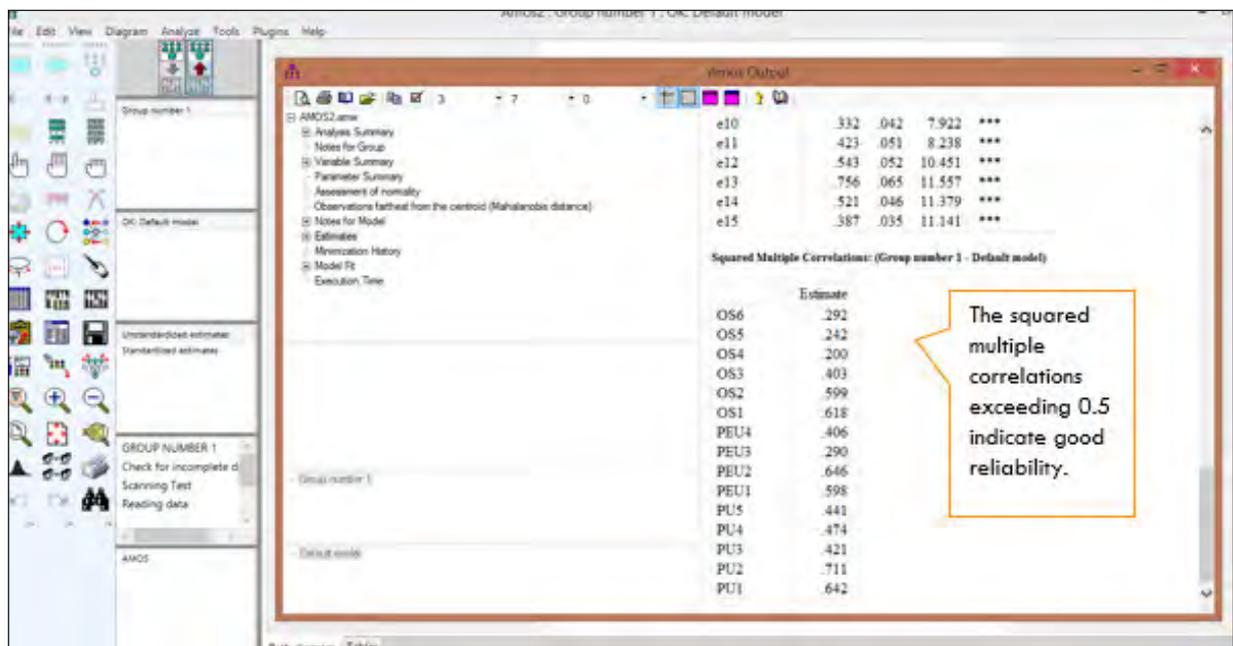


Validity and reliability assessment (From Text Output)

Standardised regression coefficients (λ)



Composite reliabilities



Variance of latent factors

AMOS Output window showing Variance estimates for latent factors (e1-e15). The table below summarizes the data from the screenshot:

Label	Estimate	S.E.	C.R.	P
per_bse	.308	.039	7.586	***
per_eou	.327	.046	7.107	***
os	.558	0	***	***
e1	.172	.019	7.521	***
e2	.135	.018	7.521	***
e3	.362	.033	10.594	***
e4	.170	.026	10.576	***
e5	.296	.027	10.779	***
e6	.219	.027	7.992	***
e7	.169	.024	7.135	***
e8	.334	.030	11.113	***
e9	.284	.027	10.356	***
e10	.332	.042	7.922	***
e11	.423	.051	8.238	***
e12	.543	.052	10.451	***
e13	.756	.065	11.557	***
e14	.521	.046	11.379	***
e15	.387	.035	11.141	***

Squared Multiple Correlations: (Group number 1 - Default model)

Estimate
OS6
.292

Callout text: Variance extracted for latent factors from indicators; the variance exceeding 0.5 represent adequate convergent validity

A detailed account of similar work by this author is published in the Journal of Agricultural Education and Extension.* This publication provides a detailed account of modelling, estimation and interpretation. The method of writing the outputs from AMOS estimation can be inferred from this paper.

*Sivakumar, P.S., Parasar, B., Das, R.N., and Anantharaman, M. (2014.) Determinants of computer utilization by extension personnel: A structural equations approach. The Journal of Agricultural Education and Extension, 20(2):191-212.



Photo Credit: Ravi Marella

About the Editors and Authors



Bharat S. Sontakki is Head, Extension Systems Management Division at the ICAR-NAARM, Hyderabad, with 25 years of research, teaching and outreach experience. He has completed 25 research projects and has organized over 200 training programmes. Apart from being the author of over 100 publications – 50 research articles, 40 conference papers, 10 edited books and proceedings, and 15 book chapters – he has also lent policy support by serving in high-level committees, such as the Working Groups of the Planning Commission (Government of India). Furthermore, he has acted as Organizing Secretary of international and national seminars/workshops. (bharatss@naarm.ernet.in).



Dhandapani A is Principal Scientist (Statistics/Computer Applications) at the ICAR-NAARM, Hyderabad. Having more than 20 years of experience at ICAR, Dr. Dhandapani is currently focusing on implementation of Research Data Management Policy of ICAR. Besides research publications in national and international journals, Dr. Dhandapani has two patents in his name for information systems. Before joining ICAR-NAARM in 2007, he worked at the ICAR-National Center for Integrated Pest Management (ICAR-NCIPM), New Delhi. During his tenure at ICAR-NCIPM, Dr. Dhandapani was involved in statistical analysis and development of Information systems for Pest Management, and he has also designed e-Pest Surveillance for monitoring of pests (dhandapani@naarm.ernet.in).



Nimisha Mittal is Program Manager, Centre for Research on Innovation and Science Policy (CRISP), Hyderabad. A postgraduate from the Indian Institute of Forest Management (IIFM) Bhopal, she previously worked with Udyogini and Development Research Services (DRS), New Delhi. She has 16 years of extensive experience of research in broad-based fields, such as livelihoods, education, environment and community-based natural resource management with special focus on rural women. Apart from her involvement in all these endeavors, she also conducts and leads research, evaluation, and workshops (nimisha61@gmail.com).



Rasheed Sulaiman V is Director, Centre for Research on Innovation and Science Policy (CRISP), Hyderabad. Previously he served as Senior Scientist (Agricultural Extension) at the ICAR-National Institute of Agricultural Economics and Policy Research (NIAEPR), New Delhi. His research interests include agricultural extension, agricultural innovation systems, and policy analysis. He was a member of the XIth Plan Working Group on Agricultural Extension, constituted by the Planning Commission, Government of India (2011-12).

Dr Rasheed co-ordinates the Agricultural Extension in South Asia (AESAs) Network and is currently the Chair of the Steering Committee of the Global Forum for Rural Advisory Services (GFRAS) (rasheed.sulaiman@gmail.com).



Saravanan Raj is Director (Agricultural Extension), National Institute of Agricultural Extension Management (MANAGE), Hyderabad, India. He heads the Centre for Agricultural Extension Innovations and Reforms. Prior to this, he worked as an Associate Professor (Extension Education and Rural Sociology) in the Central Agricultural University, India and implemented research and extension programmes among tribal farmers in the most disadvantaged regions of Northeast India for nearly 14 years. His main research interests include e-Extension, social media, private extension services, institutional innovations and extension reforms (saravanan.raj@manage.gov.in).



P.V.K. Sasidhar is Director, School of Extension and Development Studies, IGNOU, New Delhi, where he coordinates distance education programmes in Extension Education, Development Studies and Urban Development. He was a Norman Borlaug Fellow at Tuskegee University, USA (2008), and USAID Fellow at Michigan State University, USA (2015). His research interests centre primarily on evaluation and impact assessment studies; and he has more than 200 research publications to his credit. His recent works include: 'Integrated contract broiler farming - An evaluation case study in India' and 'Assessment of core competencies of livestock extension professionals in India' for which he collaborated with Michigan State University (pvksasidhar@ignou.ac.in).



P. Sethuraman Sivakumar is Senior Scientist, ICAR-Central Tuber Crops Research Institute, Thiruvananthapuram, India. His specialization is on consumer behaviour and entrepreneurship development, and he has published over 30 papers in both international and national journals. He was a Fulbright-Nehru Postdoctoral Fellow at the Department of Psychology, Florida State University, Tallahassee, USA, during 2013-2014. Notably, he was a recipient of Outreach Lecture Fellowship of International Institute of Education, Washington DC. Recognising his efforts to promote multi-cultural understanding among US Communities, Florida State University recognised him as a 'Global Ambassador' in 2014 (sethu_73@yahoo.com).



N. Sivaramane is Senior Scientist at ICAR-NAARM, Hyderabad. He has over 15 years of experience in research, teaching and training, and has served previously as Scientist at the ICAR-Indian Agricultural Statistics Research Institute, New Delhi. He has been involved in teaching various courses to post-graduate and doctoral students in the disciplines of Agricultural Economics and Agricultural Statistics at the ICAR-Indian Agricultural Research Institute, New Delhi. His areas of specialization include Market Research, Econometrics, Agricultural Statistics, Impact Evaluation and Agricultural Marketing. He is engaged in research and training in his field of specialization and has several research articles and books to his credit (sivaramane@yahoo.com).



Sreenivasa Rao I is Professor & University Head at the Department of Agricultural Extension, College of Agriculture, Professor Jayashankar Telangana State Agricultural University, (PJ TSAU), Hyderabad, India. He is currently engaged in teaching extension courses for undergraduate, and post-graduate students and is serving as a research guide for PhD students. His publications include 62 scholarly articles in reputed peer reviewed international and national journals, 20 popular articles and two book chapters. His earlier stints include working as Head, Krishi Vigyan Kendra, Vizianagaram; and as Deputy Director of Extension at PJ TSAU. He serves as a guest faculty at many institutions and training centres (illuris@gmail.com).



Subash S.P. is Scientist (Agricultural Economics), at ICAR-NAARM, Hyderabad, India. He started his career with the International Maize and Wheat Improvement Center (CIMMYT) and was associated with several projects, such as 'Adoption of improved varieties' and 'Impact assessment of Community based seed producers in mid-hills of Nepal', 'Heat tolerant maize for Asia (HTMA)', 'Strengthening impact assessment in the CGIAR system', and 'Abiotic stress tolerant maize for Asia (ATMA)'. He has worked in projects across South and South East Asia, and his core areas of research are behavioural economics and institutional economics (subashspar@gmail.com).



Suchiradipta Bhattacharjee is MANAGE Fellow at the Centre for Agricultural Extension Innovations and Reforms, National Institute of Agricultural Extension Management (MANAGE), Hyderabad, India. Her specialization is in Agricultural Extension. Her research interests include: Agricultural Innovation Systems (AIS), ICT4Ag, social media, mExtension, and rural development (suchiradipta@hotmail.com).



Thammi Raju D is Principal Scientist, ICAR-NAARM, Hyderabad, India. He has served previously as Professor & University Head of Veterinary and Animal Husbandry Extension Education at Sri Venkateswara Veterinary University, Tirupati. He has more than 80 publications to his credit, and his areas of interest include extension research, pedagogy in higher education, teaching methodologies, extension and communication management, ICTs in livestock extension, expert systems, education information systems, and online education including MOOCs (Massive Open Online Education). He was trained at the School of Agriculture, Policy and Development, University of Reading, UK and the Borlaug Institute of International Agriculture, Texas A&M University (dtraju@naarm.ernet.in).



P. Venkatesan is Senior Scientist, Division of Extension Systems Management, ICAR-NAARM, Hyderabad, India. He has specialized in Agricultural Extension and has more than 15 years of experience as a developmental scientist, with more than 25 papers published in national and international journals; and has also guided a number of post-graduate students. He specializes in the area of indigenous technical knowledge, participatory technology development, participatory impact monitoring and assessment, tribal development, technology assessment and refinement, and socioeconomic and environmental impact assessment (venkatesan@naarm.ernet.in).

Manual on good practices in extension research and evaluation

There has been an increasing concern on the declining quality and contribution of extension research and the limited influence extension research has on improving extension practice, designing new programmes and influencing policies.

Reorienting the extension research is the need of the hour which needs a coordinated approach by integrating “state-of-the-art” methods from other sciences to improve the utility and visibility of the extension research outcomes.

This manual is intended to assist young researches and PhD students conducting quality research and evaluation in extension. It has been compiled by the resource persons who participated in the Workshop on “Good Practices in Extension Research and Evaluation” jointly organised by the ICAR-National Academy of Agricultural Research Management (NAARM), Centre for Research on Innovation and Science Policy (CRISP), Agricultural Extension in South Asia (AESAs), ICAR-Central Tuber Crops Research Institute (CTCRI) and the National Institute of Agricultural Extension Management (MANAGE) at NAARM, Hyderabad (India) during 29 November-2 December 2016.



AESA Secretariat: Centre for Research on
Innovation and Science Policy (CRISP),
Road No.10, Banjara Hills, Hyderabad 500034, India
www.aesa-gfras.net
Email: aesanetwork@gmail.com



This work has been primarily
funded by the Global forum
for Rural Advisory Services
(GFRAS)