EXPERIMENTAL RESEARCH In CS, IS and IT

Zainal A. Hasibuan, PhD
(zhasibua@cs.ui.ac.id)

STUDY PROGRAM:
MAGISTER INFORMATION TECHNOLOGY,
COMPUTER SCIENCE FACULTY,
UNIVERSITY OF INDONESIA
Research Methods
(upon which methodologies can be built)

Types of Research Methods

- Correlational
- Causal-Comparative
- Case Study
- Experimental
- Ethnographic*
- Survey*
- Action Research
- Historical
Experimental Research
Experimental Group Designs

Three Types

Experimental ∥ Quasi-experimental ∥ Non-experimental
What is Experimental Design?

• An experimental design:
  “Is the traditional approach to conducting quantitative research”

Sumber: Creswell, J.C. 2005
Experimental Design

• Treatment, control, comparison
• do something to subjects (guinea pigs?)
• who are randomly selected and randomly assigned to groups
• for the purpose of determining the cause of an effect (difference between groups)
Characteristics of Experiments

- Random assignment
- Control over extraneous variables
- Manipulation of the treatment conditions
- Outcome measures
- Group comparisons
- Threats to Validity
Random Assignment

• **Random assignment** is the process of assigning individuals to the treatments.

• The random assignment of individuals to groups (or condition within a group) distinguishes a rigorous’ “true” experiment from an adequate, but less-than-rigorous, “quasi-experiment”.
Experimental Research

- **Experiment**: The investigation of the relationship between two or more variables by deliberately producing a change in one variable in a situation and observing the effects of that change on other aspect of the situation.

- **Cause-and-effect**
Experimental Research

- **Experimental manipulation**: Change that an experimenter deliberately produces in a situation
- **Treatment**: the manipulation implemented by experimenter
- **Experimental group**: any group receiving a treatment in an experiment
Experimental Research

• In an *observational study*, measurements of variables of interest are observed and recorded, without controlling any factor that might influence their values.

• An *experiment*, on the other hand, deliberately imposes some treatment on individuals in order to observe their responses.

• In principle, only experiments can give good evidence for causation.
Steps In Conducting Experimental Research

- Step 1. Decide if an experiment addresses your research problem
- Step 2. Form hypotheses to test Cause-and-effect Relationships
- Step 3. Select an Experimental Unit and Identify Study participants
- Step 4. Select an Experimental Treatment and Introduce it
- Step 5. Choose a Type of Experimental design
- Step 6. Conduct the Experimental
- Step 7. Organize and Analyze the Data
- Step 8. Develop an experimental Research Report

Sumber: Creswell, J.C. 2005
The Road to Experimental Research
Types of Experimental Designs

• Between-Group Designs
  – True experiments (pre-and posttest, posttest only)
  – Quasi-experiments (pre-and posttest, posttest only)
  – Factorial Design

• Within-Group or Individual Designs
  – Time series experiments (interrupted, equivalent)
  – Repeated measures experiments
  – Single-subject experiments
## Types of Experimental Design

<table>
<thead>
<tr>
<th></th>
<th>True Experiment</th>
<th>Quasi Experiment</th>
<th>Factorial</th>
<th>Time Series</th>
<th>Repeated Measures</th>
<th>Single Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random assignment?</td>
<td>Yes</td>
<td>No</td>
<td>May be Used</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Number of groups/individuals compared?</td>
<td>Two or more</td>
<td>Two or more</td>
<td>Two or more</td>
<td>One group</td>
<td>One group</td>
<td>One individual studied at a time</td>
</tr>
<tr>
<td>Number of interventions used?</td>
<td>One or more interventions</td>
<td>One or more interventions</td>
<td>Two or more interventions</td>
<td>One or more interventions</td>
<td>Two or more interventions</td>
<td>One or more interventions</td>
</tr>
<tr>
<td>Number of times the dependent variables measured/observed?</td>
<td>Once</td>
<td>Once</td>
<td>Once</td>
<td>After each intervention</td>
<td>After each intervention</td>
<td>Multiple points</td>
</tr>
<tr>
<td>Controls typically used?</td>
<td>Pretest matching, blocking, covariates</td>
<td>Pretest matching, blocking, covariates</td>
<td>Pretest matching, blocking, covariates</td>
<td>Group becomes its own control</td>
<td>Covariates</td>
<td>Individuals become their own controls</td>
</tr>
</tbody>
</table>
### Types of Between-Group Designs

**True Experimental Designs**

#### Pre-and Posttest Design

<table>
<thead>
<tr>
<th>Random assignment</th>
<th>Control Group</th>
<th>Pretest</th>
<th>No Treatment</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random assignment</td>
<td>Experimental Group</td>
<td>Pretest</td>
<td>Experimental Treatment</td>
<td>Posttest</td>
</tr>
</tbody>
</table>

#### Posttest-Only Design

<table>
<thead>
<tr>
<th>Random assignment</th>
<th>Control Group</th>
<th>No Treatment</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random assignment</td>
<td>Experimental Group</td>
<td>Experimental Treatment</td>
<td>Posttest</td>
</tr>
</tbody>
</table>
## Types of Between-Group Designs
*(Quasi Experimental Designs)*

### Pre-and Posttest Design

<table>
<thead>
<tr>
<th>Time</th>
<th>Select Control Group</th>
<th>Select Experimental Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest</td>
<td>Pretest</td>
<td>Pretest</td>
</tr>
<tr>
<td>No Treatment</td>
<td>No Treatment</td>
<td>Experimental Treatment</td>
</tr>
<tr>
<td>Posttest</td>
<td>Posttest</td>
<td>Posttest</td>
</tr>
</tbody>
</table>

### Posttest-Only Design

<table>
<thead>
<tr>
<th>Time</th>
<th>Select Control Group</th>
<th>Select Experimental Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Treatment</td>
<td>Posttest</td>
<td>Posttest</td>
</tr>
<tr>
<td>Experimental</td>
<td>Posttest</td>
<td>Posttest</td>
</tr>
<tr>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Experiment Example

• *New* “instant breakfast” product.
• To assess its nutritional quality, researchers in the lab feed 30 newly weaned male white rats and measure their weight gains over a 28-day period.
• They randomly select 15 rats and feed them using the new product.
• The other 15 rats receive a standard diet.
Design of Experiments

• Experimental units: individuals on which the experiment is done, also called subjects when the units are human beings.
  – The rats

• Treatment: the specific experimental condition applied to the units.
  – “instant breakfast” diet

• Factors: the explanatory variables, which often have levels.
  – the diet
Principles of Experimental Design

• Control
  – Researcher decides which subjects are assigned to the treatment group

• Randomization
  – Impartial and objective

• Replication
  – Reduces chance variation in the results and can help achieve statistical significance
Experimental Research

• **Experimental Research**: research that allows for the causes of behaviour to be determined

• **Experiment**: a carefully regulated procedure where one or more factors are deliberately manipulated and all other factors are held constant.

• Cause-effect relationship occurs if:
  – The cause is correlated with the effect.
  – The cause occurred before the effect.
  – We can rule out other plausible explanations of the causal relationship

• **Example**: Learning effectiveness of e-Learning system
Experimental Research: Factors

• Independent Variable (IV): factor that is manipulated

• Dependent Variable (DV): factor that is measured

• Experimental condition: subjects that are manipulated

• Control condition: subjects that are not manipulated

• Confounding variable: an extraneous variable that should be controlled, but is not. Can lead to false/spurious conclusions!

• Anecdote:
  – Day 1: drink water + beer. Result? Drunk!
  – Conclusion?
Examples of Experimental Research Design

- **Problem**: A variety of peer-to-peer (P2P) systems for sharing documents are currently available. The challenging is how to design a routing strategy that leads the user finding the documents needed.

- **Research Question**: What mechanisms to effective and efficient keyword-based searching for documents in unstructured P2P system?

- **Purpose**: To explore the efficiency and effectiveness of keyword-based searching of documents in P2P system by proposing reinforcement learning mechanism.

- **The efficiency is measured in term of response time, and effectiveness is measured in term of relevant documents.**
Experimental Research Methodology (Adopted from Nizar, 2007)

• Problem analysis
  – Comparing several learning algorithm
  – Choose the algorithm

• Simulation set up
  – Technological setting
    • Running in PC Pentium 4
  – Experimental setting
    • Apply simple keyword-based routing on…….
    • Control the searching process
    • Run several times

• Data Analysis
  – Record the amount of time used and the number of document relevant
  – Tabulated the data

• Data Interpretation
  – Based on the chart
  – Based on efficiency and effectiveness
Another Example: The Use of Maximal Frequent Sequences to Improve Document Ranking in Information Retrieval System for Indonesian Language (Dwi Astuti, 2006)

- **Background and Problem**
  - The amount of textual information available through the World Wide Web has increased dramatically in recent years
  - Web users need effective search mechanism in order to find useful information from the enormous quantities of available text data
  - Very often, users are precision-oriented, they prefer a small set of documents containing a good proportion of useful documents to a large set of documents that contains a lot of useful information, but a fair amount of irrelevant information as well. Thus, there has been a growing interest in high-precision IR systems in recent times.
Experimental Research Methodology, Astuti (2006)

- **Data Sets**
  - The data sets used in this experiment were comprised of 2 Indonesian corpuses, a news corpus and a scientific corpus.
  - The news corpus is a collection of on-line news from famous Indonesian newspapers, consists of 3000 documents, and is formatted in the TREC-like format. The corpus comes with 20 set of queries and their relevancies [ASI04].

- **Programming Language and Libraries**
  - We implemented most of the method and algorithm in Python, a high-level, interpreted object-oriented programming language well-known within the UNIX and internet community (http://www.python.org).
Experimental Research Methodology, Astuti (2006)

• **Collection Preprocessing**
  – The reindexed collection was then subjected to word filtering to remove punctuation, one-letter words, numbers, word repetition, and about 250 stop words.
  
  – The filtered collection then underwent stemming using an Indonesian stemmer [ASI03].

• **Term-Frequency Processing**
  – Use a tf-normalized version of the tfc (term frequency cosine) term-weighted components with the following formula [SAL88]:

\[
 tf \ idf_w = \frac{tf_w \cdot \log \frac{N}{n_w}}{\max (tf) \cdot \sqrt{\sum_{w_i \in W} (tf_{w_i} \cdot \log \frac{N}{n_{w_i}})}}
\]
Experimental Research Methodology, Astuti (2006)

• Applying The Process of MFS Discovery
  – Pair Discovery
  – MFS Discovery
  – MFS Indexing

• Evaluation of Retrieval System
  – Boolean Term Frequency Model
  – Term Frequency Cosine Model
  – MFS Model
  – Hybrid (Term Frequency Cosine/MFS) Model
  – Evaluation Strategy
The Experimental Results

• **The Process of MFS Discovery**
  - The results of searching frequent word pairs (2-grams) between stemmed and non-stemmed version of each collection (figure 1 and 2).
  - The results of MFS discovery for different threshold were shown in figure 3 and 4. Because of space limitation, we cannot provide all results from different combination of threshold, parameter g and collection version (stemmed and non-stemmed).
Figure 1. Result Of Word Pair Searching For News Collection
Figure 2. Results Of Word Pair Searching For Scientific Collection

![Graph showing results of word pair searching for scientific collection. The graph compares non-stemming and stemming methods. The x-axis represents the number of documents, and the y-axis represents the number of word pairs. The graph shows a decreasing trend in the number of word pairs as the number of documents increases. The non-stemming method generally has more word pairs than the stemming method.]
Discussion and Interpretation

- The Discovery of MFS
  - The process of discovering MFS for each collection started with the search of frequent word pairs (or 2-gram)
  - Using a bigger parameter $g$ results in more pairs being constructed
  - The search for frequent pairs proved to be a major challenge for BsdDB component of the retrieval system
  - The discovery of MFS proved to be CPU and memory intensive process
  - Smaller threshold spreads the distribution of MFS
  - Smaller threshold increases the processing time substantially
Discussion and Interpretation

• **Evaluation of Retrieval System**
  – *Term frequency method with tfc (term frequency cosine) scoring scheme gave a respectable results*.

  – *A modified tfc with correction factor (tfc2) improves the precision of the term frequency cosine model for the scientific collection*.

  – *More MFS representing document usually improves the precision of the retrieval algorithm*.

  – *Stemming improves the precision of all retrieval models*. 
Conclusion and Suggestion

• An improved method for utilizing MFS in calculating the rank of documents should be formulated.

• To be practical in the real-world application, a new algorithm for doing incremental MFS discovery needs to be developed.
Quasi Experimental Research
Quasi-Experimental Designs

• A research design in which an experimental procedure is applied but all extraneous variables are not controlled
Characteristics of Quasi-Experimental Research

- There is a control or comparison group
- Intact groups are used
- The treatment is randomly assigned to groups.
Nonequivalent Control Group Design

• A quasi – experimental design in which the results obtained from nonequivalent experimental and control groups are compared
Quasi-Experimental Research

Parametric Tests

Statistical Analysis: The *t* Test

*For testing the significance of difference between two sample means*

Basic Assumptions

1- *Scores form an interval or ratio scale*
2- *Scores are normally distributed*
3- *Score variances for the populations under study are equal (SD=SD)*
Quasi-Experimental (Cont’d)

Analysis of Variance (ANOVA)

*Comparison of two or more group means*

Multivariate Analysis of Variance (MANOVA)

*Statistical technique for determining whether groups differ on more than one dependent variable.*

Basic Assumptions

1-*Scores form an interval or ratio scale*
2-*Scores are normally distributed*
3-*Score variances for the populations under study are equal (SD=SD)*
Quasi-Experimental (Cont’d)

Nonparametric Tests

Nonparametric statistics tests statistical significance that do not rely on any assumptions about shape or variance of population scores.

Used with measures that yield categorical or rank scores, or do not have equal intervals. Nonparametric tests are less powerful, they require larger samples to yield the same level statistical significance.

1-**The Chi-Square Test** = used to determine whether research data in the form of frequency counts are distributed differently for different samples.
2- The Mann-Whitney U test is used to determine whether the distributions of scores of two independent samples differ significantly from each other.

3- The Wilcoxon signed rank test is used to determine whether the distributions of scores of two samples differ significantly from each other when the scores of the samples are correlated.
Quasi-Experimental (Cont’d)

Nonparametric Tests (Cont’d)

4- **The Kruskal-Wallis test** = If more than two groups of subjects are to be compared, a nonparametric one-way analysis of variance (Kruskal-Wallis) can be used.
Quasi-Experimental Research

• Quasi-experimental research
  – Almost but not quite real experiments
  – No manipulation of the variables (so no IV)
  – Compare groups biased on naturally occurring variables

• Two types of natural variables
  – Subject variable: Characteristics that vary between participants, but cannot be manipulated
  – Time variable: Comparing individuals at different points in time (age 3 and 6)

• One-shot post-test, no control group
• Example: The impact of marketing strategy
Diagramming Research

• To illustrate research designs, a number of symbols are used
  – $X_1 =$ Treatment
  – $X_2 =$ Control Group
  – O = Observation (pretest or posttest)
  – R = Random Assignment
A Sample Research Design

- Single-Group Pretest-Treatment-Posttest Design

\[ R \ O \ X_{1} \ O \]

This means subjects are randomly assigned to a group, which is then given a pretest, then there is a treatment, then there is a posttest.
Classification of Research Design (Causal Comparative)

One-group pretest-posttest design

$O_1 \rightarrow X \rightarrow O_2$

Nonequivalent control group

Group 1: $O_1 \rightarrow X \rightarrow O_2$

Group 2: $O_3 \rightarrow O_4$

Equivalent time-samples design

$X_1 \rightarrow O_1 \rightarrow X_2 \rightarrow O_2$
Research design with more power (time series)

- Pre-test post-test

\[ \begin{array}{cccccc}
O_1 & O_2 & O_3 & \times & O_4 & O_5 & O_6 \\
\end{array} \]

- Pre-test post-test with control group

\[ \begin{array}{cccccc}
O_1 & O_2 & O_3 & \times & O_4 & O_5 & O_6 \\
O_1 & O_2 & O_3 & O_4 & O_5 & O_6 \\
\end{array} \]
Changes to Look For

- No effect
- Change in the rate or slope
- Change in the intercept
Non-Experimental Research
Research Designs by Similarities

Experimental & Quasi-experimental
- Involves Researcher Intervention

Non-experimental
- Examines phenomena as they exist
  Descriptive, Causal-Comparative, and Correlational
Descriptive Research

• **Purpose**
  – To describe the way things are
  – Or “what is”

• **Many of the methods used, can also be used for correlational research**
  – Difference is the purpose
    • describing v. examining a relation

• **Two main types**
  – Surveys
  – Observations
Basic of Descriptive Research

- **Objective:** Describe market characteristics or functions

- **Characteristics:**
  - Marked by the prior formulation of specific hypotheses
  - Preplanned and structured design

- **Methods:** Secondary data
  - Surveys
  - Panels
  - Observation and other data
Causal-Comparative Research

The Purpose

Purpose of explaining educational phenomena through the study of cause-and-effect relationships. The presumed cause is called the independent variable and the presumed effect is called the dependent variable. Designs where the researcher does not manipulate the independent variable are called ex post facto research.
Causal-Comparative Research

Ex Post Facto = Causal-Comparative Research

• Explores possible causes and effects
• The independent variable is not manipulated, it has already been applied
• Focuses first on the effect, then attempts to determine what caused the observed effect.
• Seeks to explain differences between two groups that have occurred
• Example: Why are IT multinational companies more innovative than local firms?
Correlational Designs

The Purpose

To discover relationships between variables through the use of correlational statistics. Involves correlating data on two or more variables for each individual in a sample and computing a correlation coefficient.

Two major purposes:
1-To explore causal relationships between variables;
2-To predict scores on one variable from research participants’ scores on other variables.
Correlation Research Design

Advantages

1- Enables researchers to analyze the relationships among a large number of variables in a single study.
2- They provide information concerning the degree of the relationship between the variables being studied.

Parametric Test

Pearson r statistical procedure

Basic Assumptions

1- Scores form an interval or ratio scale
2- Scores are normally distributed
3- Score variances for the populations under study are equal (SD=SD)
**Scattergrams Representing Different Degrees and Directions of Correlation between Two Variables**

Positive correlation \( (r=0.99) \)

Negative correlation \( (r=-0.73) \)

- Grade point
- I.Q.
- Computer use
- Age
Conclusion
Question?