Lecture Topic 2 :

Research Design, Problem Formulation, Research Gap, Research Hypothesis, Research Questions and Research Objectives, Operational Framework & Guide to write Research Proposal

SOURCES OF SLIDES:

1) DR. NURULHUDA FIRDAUS
2) ALLAHYARHARMAH DR. ROSMAH ALI

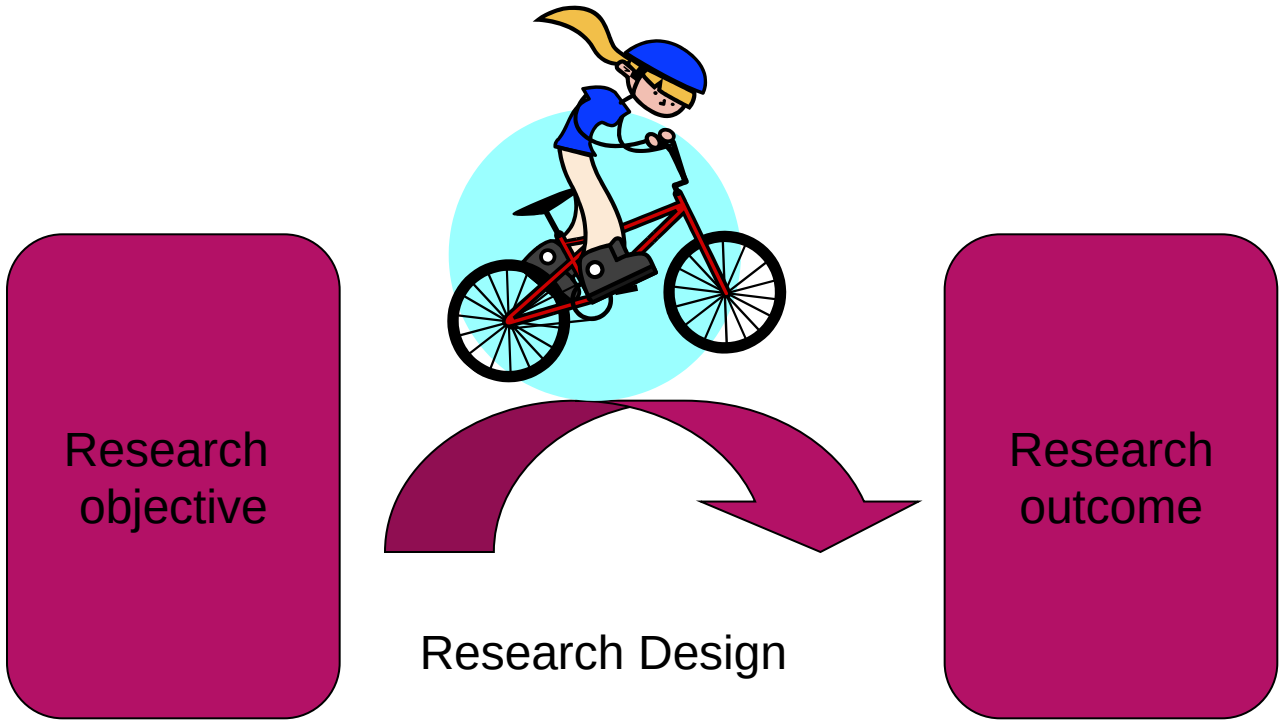
RESEARCH DESIGN: What it is?

- A research design is a plan for an entire research study to be conducted
  - provides a road map for the entire study
- It involves specifying:
  - Research topics: how do you decide the topic!
  - philosophical assumptions (problems, objectives etc..)
  - research method
  - data collection techniques you will use
  - data analysis and,
  - findings

RESEARCH DESIGN: What it is?

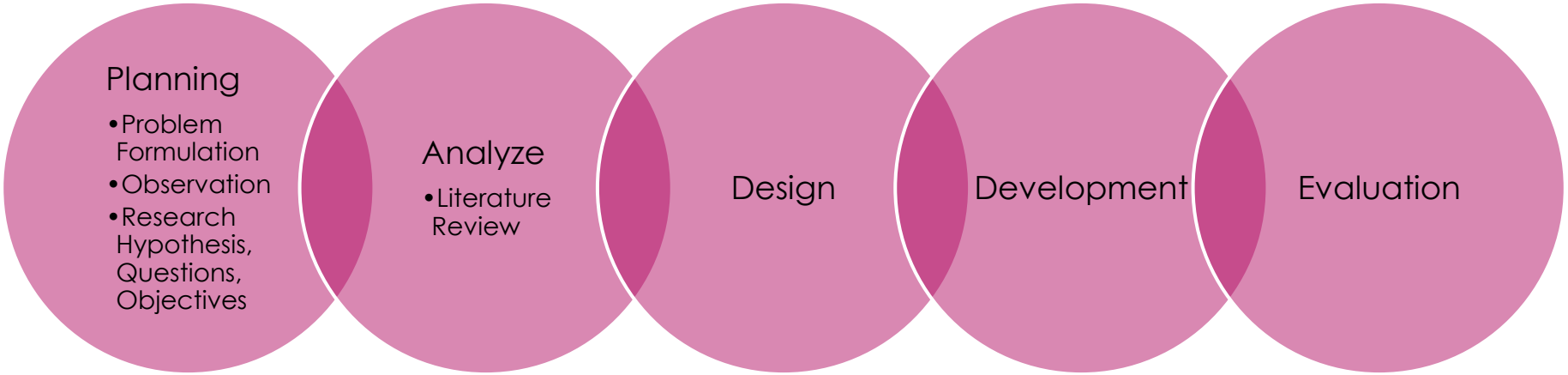
- Is the strategy for a study and the plan by which the strategy is to be carried out.
- Issues relating to purpose of study (exploratory, descriptive, hypothesis testing), its location (setting), type of investigation, time horizon, unit of analysis.
- It specifies the methods and procedures for the collection, measurement and analysis of data.

RESEARCH DESIGN: What it is?

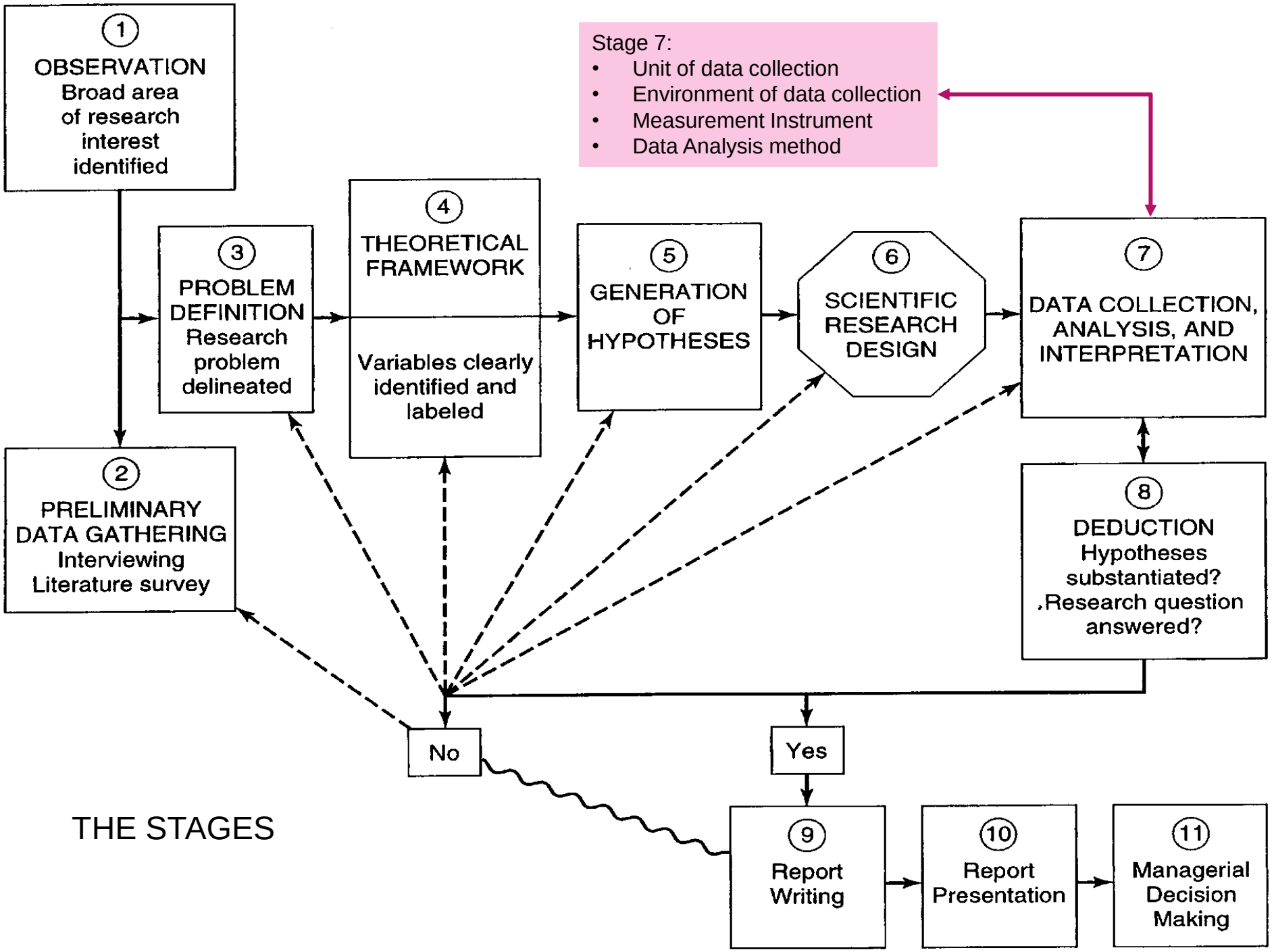


THE STAGES

Notes: Not Limited to



Notes: This is a typical research design, it may varies depending on the study


Stage 7:
• Unit of data collection
• Environment of data collection
• Measurement Instrument
• Data Analysis method
1
OBSERVATION
Broad area of research interest identified
2
PRELIMINARY DATA GATHERING
Interviewing
Literature survey
3
PROBLEM DEFINITION
Research problem delineated
4
THEORETICAL FRAMEWORK
Variables clearly identified and labeled
5
GENERATION OF HYPOTHESES
6
SCIENTIFIC RESEARCH DESIGN
7
DATA COLLECTION, ANALYSIS, AND INTERPRETATION
8
DEDUCTION
Hypotheses substantiated?
Research question answered?
No
Yes
9
Report Writing
10
Report Presentation
11
Managerial Decision Making
THE STAGES

CHOOSING THE RESEARCH TOPIC

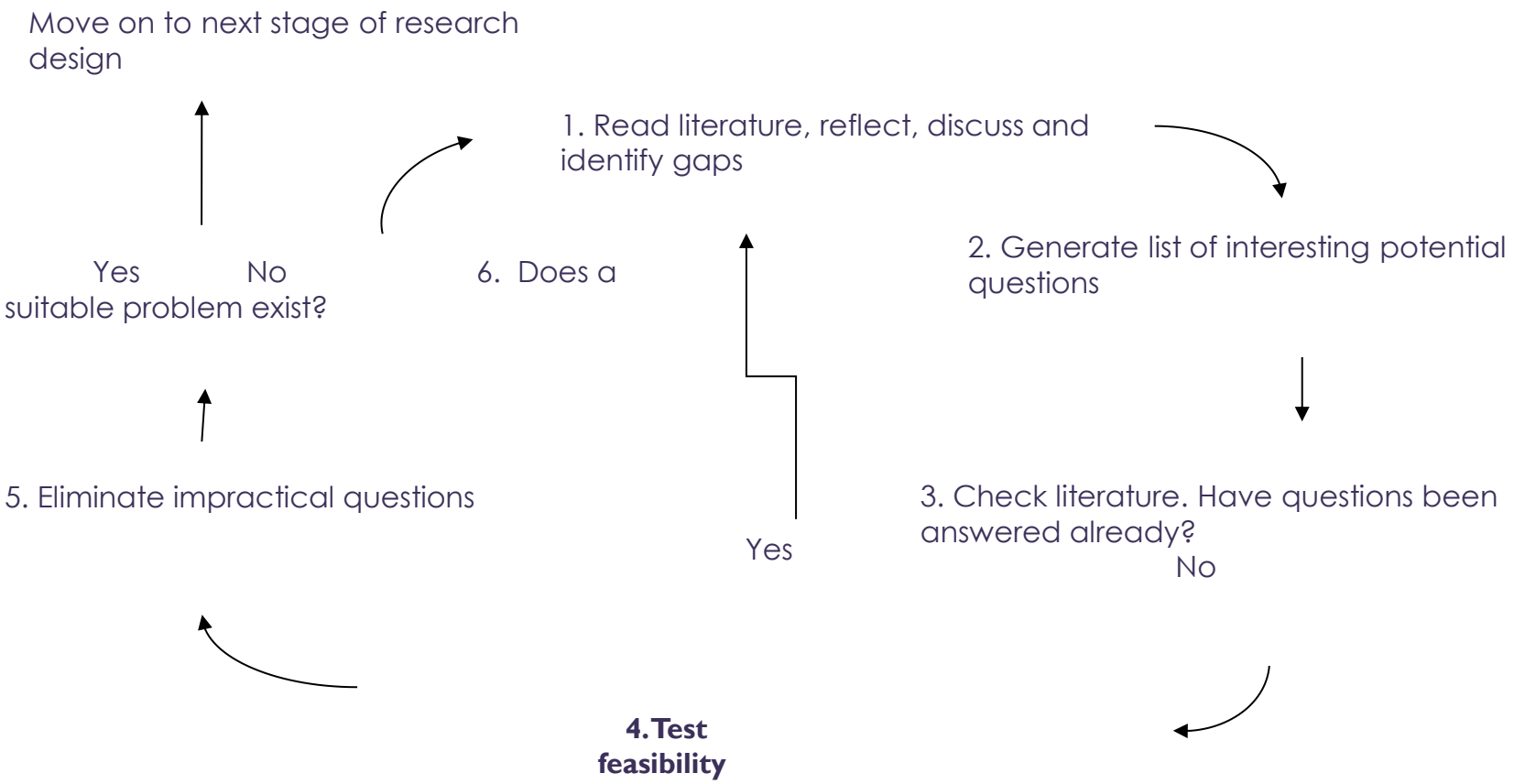
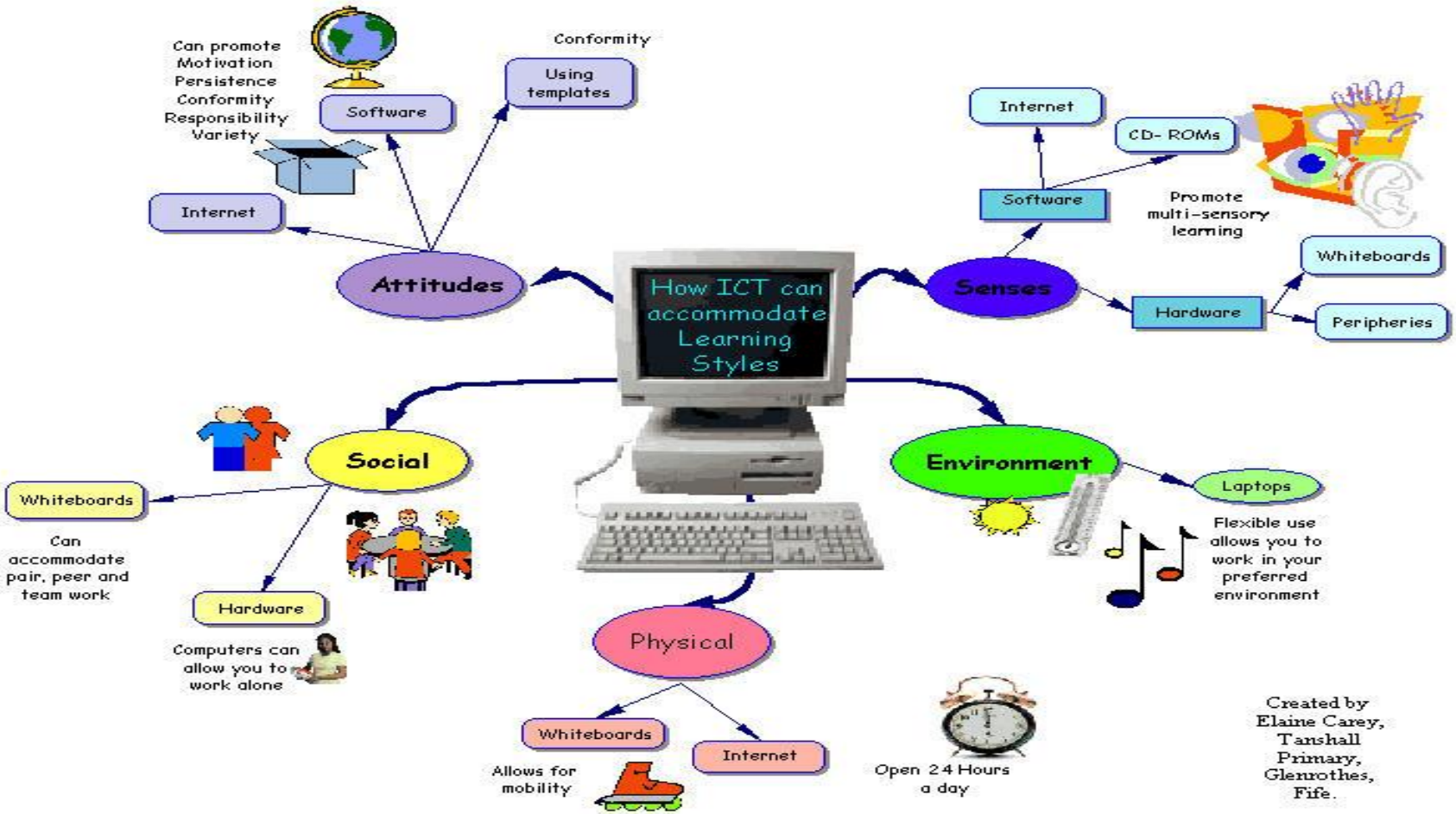


Figure 3.1. How to identify a research problem (adapted from Collis & Hussey, 2003)

CONCEPT MAP OF POTENTIAL RESEARCH TOPICS



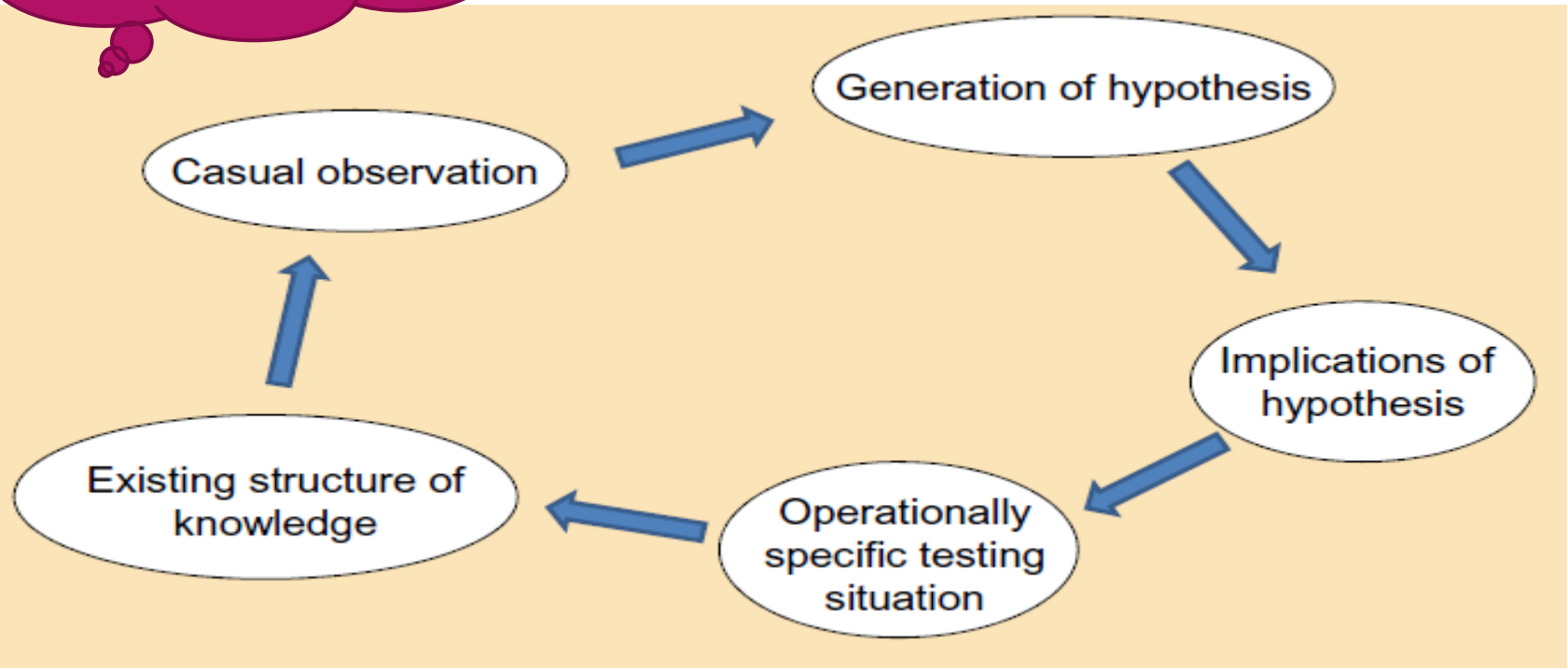
THE PROBLEM FORMULATION

PHILOSOPHICAL ASSUMPTIONS: THE PROBLEM FORMULATION

- Problem Formulation:
  - The Objectives
    - To impart the skills on how to select and formulate the research problems
  - Why need to formulate the problems?
    - Defining a research problem : to determine current knowledge about a problem, its causes, and possible solutions
    - Clarifying the problem: making clear exactly what we need to find out through research
    - We need to know as much as possible about a topic before we can formulate a researchable problem

PHILOSOPHICAL ASSUMPTIONS: THE PROBLEM FORMULATION

Research become cyclic and dynamic

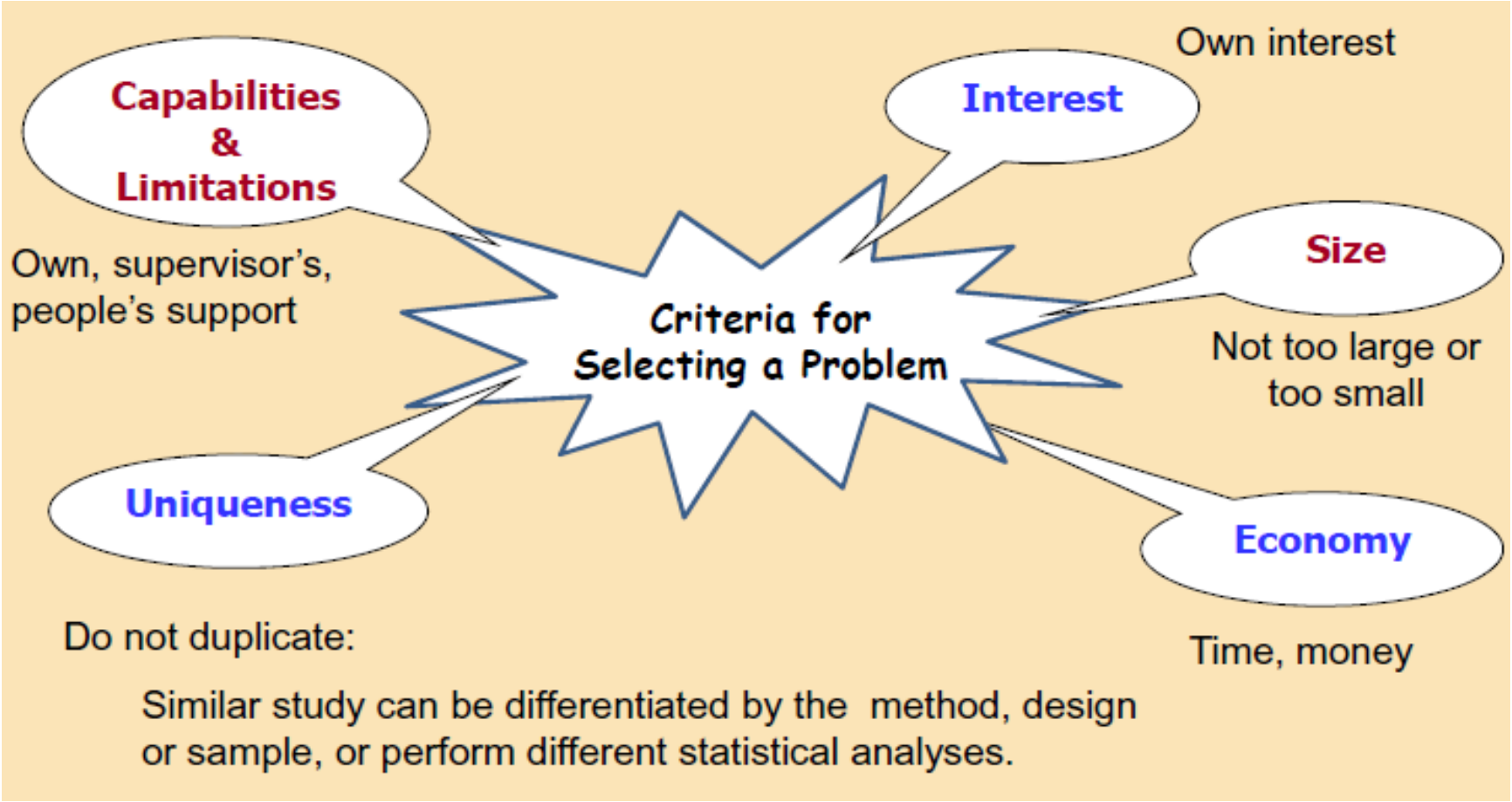


Scientific method of acquiring knowledge of problem solving

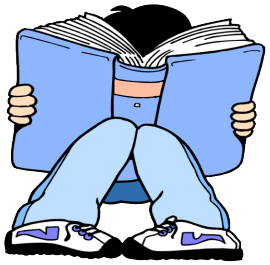
PHILOSOPHICAL ASSUMPTIONS: THE PROBLEM FORMULATION

- Part of the important features in research proposal:
  - Problem Formulation and Solution to Problem:
    - Problem statement
    - Purpose statement
    - Problem organized into research questions or hypotheses or objectives [check your field of study!]
    - Basic assumptions and limitations of study

THE PROBLEM FORMULATION: Criteria for Selecting a Problem



THE PROBLEM FORMULATION: Sources for Selecting a Problem



Reading

Academic Experience

Exposure to Field Situations

Research

Daily Experience

Brainstorming

Consultations

Intuition

THE PROBLEM FORMULATION: How to Select the Problem

- Personal Practical Experience
  - From past studies, look for
    - √ Solution already found in parallel situation
    - √ Incomplete solution
    - √ Result unclear, doubtful, debatable, etc.
- Sources: books, journals, periodicals, annual reports….

Need a Critical Study of the literature, i.e to:
– Define the problem
– Limit the problem area
– Avoid unnecessary repetition
– Search for new approaches
– Recommend suitable methods

THE PROBLEM FORMULATION: How to Select the Problem

- Consultation / interaction with others
  - Ask experts in area of interest – industries, universities, etc

- Study current developments and trends
  - Examine the theoretical structure of the field
  - Explore areas of dissatisfaction, public interest
  - journal, magazine
- Brainstorming

THE PROBLEM FORMULATION: GRILLING THE PROBLEM

THE PROBLEM FORMULATION: GRILLING THE PROBLEM

- Is the idea viable?
- Is it practicable?
- The time factor?
- Has it been done before?
- What result is expected?
- What do colleagues think?
- Will a statistician be needed?
- What will you personally do? (individual's role in the study)
- Think! If you do not solve the problem who will suffer the most.

THE PROBLEM FORMULATION:
Asking question approach – six question that should be in your research

Who- People

What*– Problems, things, idea

Where- Places

When- Past, present, future

Why *- Causes, reasons, results, condition

How* - Methods, techniques, mechanism

* Most common

DOs and DON'Ts in SELECTING a RESEARCH PROBLEM

a. Research problem cannot be borrowed; the research has to find his own problem;
b. Guide can only help to choose a topic/subject
c. Right question must be addressed:
   - Having a topic to read about is different from having a problem to solve.
     - leads to aimless and endless gathering of data
     - no way of knowing when we have enough
     - leads to a struggle to decide what to include in report
d. Be objective: have unbiased and unattached approach
e. Hanging loose: be uncommitted before selection
f. Keep alternatives: have more than one problem to ponder

DOs and DON'Ts in SELECTING a RESEARCH PROBLEM

g. Interact with experts and practitioners
h. Avoid superficial and obvious problem
i. Avoid overdone subjects and controversial subjects
j. Avoid too narrow or too broad/vague problem
k. Have a preliminary study (examine methodology, etc.)
l. Problems should suit your interest, competence and ability
m. Identify gaps through literature survey

DOs and DON'Ts in SELECTING a RESEARCH PROBLEM

n. Check availability of required data and cooperation of people concerned

o. Problem should be novel, significant and useful to practitioners; utility of the expected findings should be judged

p. Spend lot of time in writing and note taking to understand

q. Make preliminary outline:
   - Disagree with what is read
   - Draw diagrams to connect disparate facts
   - Summarise sources
   - Record random thoughts
   - Discard later if necessary

r. Start writing at the very beginning as you go to encourage critical thinking, to understand sources better and draft more effectively.

THE PROBLEM FORMULATION

Useful Advice

a) Begin early thinking to save time and later panic
b) Ask for Help: discuss with others
c) Look for problems as you read: gap, error, misunderstanding, contradictions, inconsistencies, incomplete explanations, do more than just pointing out
d) Look for the problem that your claim solves; work backward to formulate a better, more interesting problem than the one that is started.

Booth, Wayne C, et.al. *The craft of research*. 2 ed. Chicago: The University of Chicago Press, 2003. (http://www.scribd.com/doc/966557/Research-Methodology-Part-2-Selecting-Formulating-Research-Problem-)

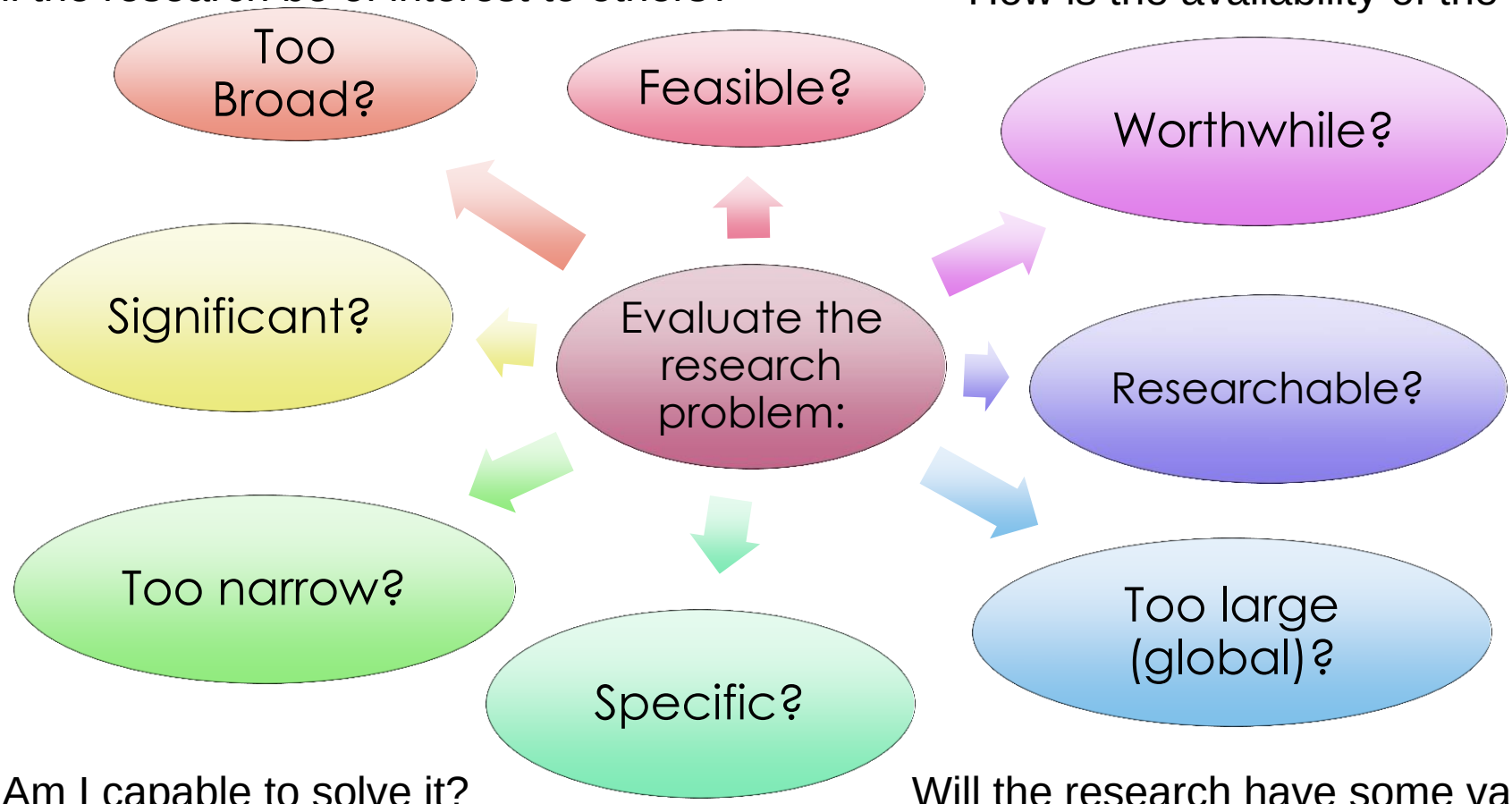
THE PROBLEM FORMULATION

1. Stating problem in a general way and **develop a title**
2. Understanding the nature of problem and build a **conceptual model**
3. Surveying available **literature** and past studies
4. Developing ideas through discussion – experience survey (setting **investigative questions**)
5. **Rephrasing the problem:**
   - Objectives and/or hypothesis
   - Title
   - Terms and concepts
   - Assumptions and postulates
   - Significance and value
   - Suitability in terms of ability, time, money, data, etc.
   - Scope and limitations
     - Time and space-coordinates
     - Unit of analysis
     - Environmental conditions

EVALUATING THE RESEARCH PROBLEM

Will the research be of interest to others?

How is the availability of the data?



Am I capable to solve it?

Will the research have some value?

Characteristic of good topic

- Interesting – keep the researcher interested in it throughout the research process
- Researchable – can be investigated through the collection and analysis of data
- Significant – contributes to the improvement and understanding of educational theory and practice
- Manageable – fits the level of researcher's level of research skills, needed resources, and time restrictions
- Ethical – does not embarrass or harm participants

STATEMENT OF RESEARCH PROBLEM

- Can be written in a form of:
  - Research Question
    - E.g: "What effect has the introduction of information security awareness course had on the perception of the computer science students towards information security?"
  - Problem Statement
    - E.g: "This study is designed to measure the introduction of information security awareness course had on the perception of the computer science students towards information security."
  - Statement of purpose
    - E.g: "The purpose of this study is to investigate the impact of the introduction of information security awareness course had on the perception of the computer science students towards information security"

RESEARCH HYPOTHESIS (RH)

HYPOTHESES DEFINITIONS

A hypothesis

- Is stated in a declarative form
- Posits a relationship between variables
- Is brief and to the point
- Is testable

Example: "Two variations of one basic algorithm (Extended AISEC or Original AISEC) produce result that are not statistically significantly different (i.e, they don't have the same median)".

Example: Graduate students who read the text in research methods will score higher on their comprehensive exams than graduate students who did not read their research methods text.

IMPORTANCE OF HYPOTHESES

- Hypotheses:
  - Direct our observations
    - Identifies the variables examined and data to be collected
  - Describe a relationship among variables
    - Can state that as one variable increases, the other will decrease; as one variables increases, the other will increase, and so on.
  - Refer to populations
    - Hypotheses help researchers infer that results of a sample will translate to a population

TYPES OF HYPOTHESES

- Research Hypotheses
  - A statement of the relationship among two or more variables or groups.
    - Example: Graduate students who read the text in research methods will score higher on their comprehensive exams than graduate students who did not read their research methods text.
- Statistical Hypotheses
  - mathematical, or logical statements that help researchers interpret the results of research
  - consist of the ***Null Hypothesis*** (H_0), the hypothesis of no difference and the ***Alternative Hypothesis*** (H_1 or H_A) which is similar in form to the research hypothesis.

TYPES OF HYPOTHESES: Statistical Hypotheses

- Remember, *and this is important*:
  - **The null hypothesis always implies that there is no relation or statistical difference between variables or groups**
  - **The alternative hypothesis implies that there is a meaningful relationship among variables or groups**

DEVELOPING RESEARCH QUESTIONS (RQ)

(I know what general area, but
I'm not sure of my research question?)

THE IMPORTANCE OF GOOD QUESTIONS

A good research question:

- Defines the investigation
- Sets boundaries
- Provides direction

O'Leary, Z. (2004) The Essential Guide to Doing Research. London: Sage

CYCLES OF RESEARCH QUESTION DEVELOPMENT

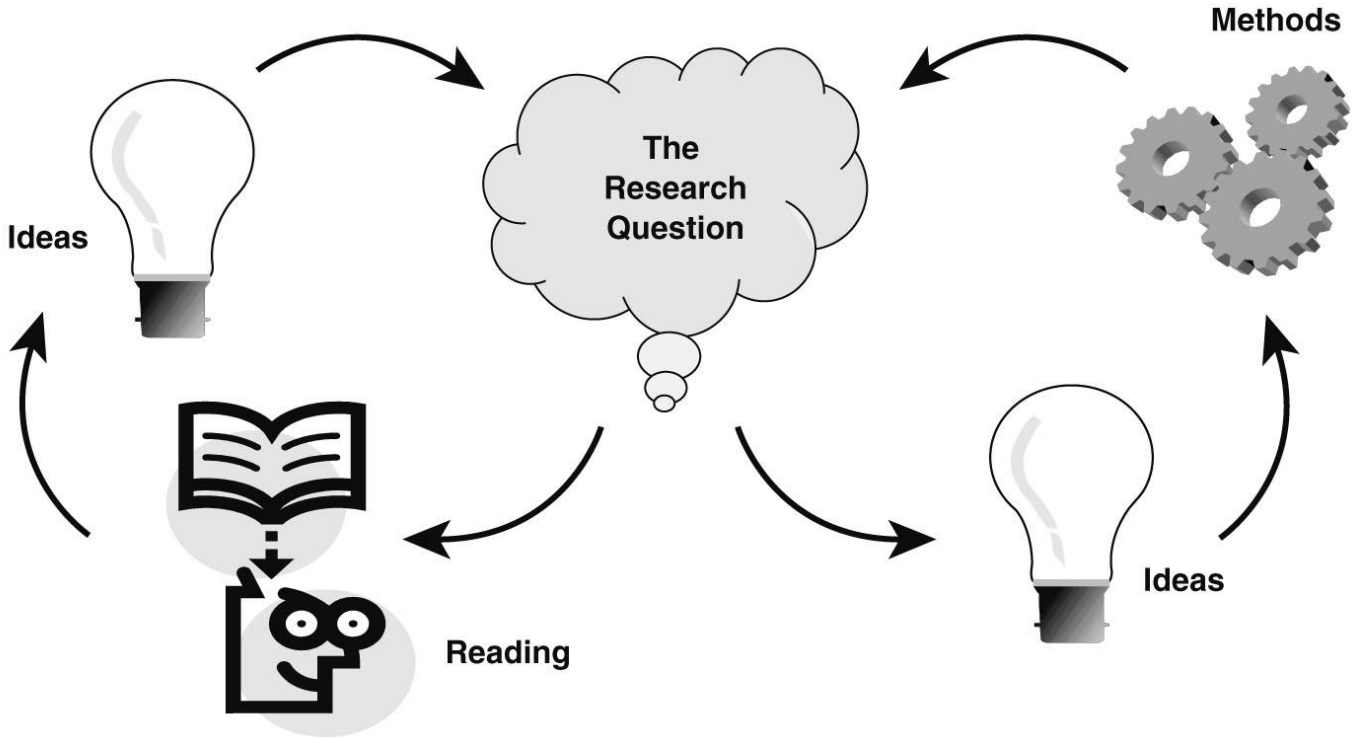


FIGURE 3.4 CYCLES OF RESEARCH QUESTION DEVELOPMENT

O'Leary, Z. (2004) The Essential Guide to Doing Research. London: Sage

Good Question Checklist

✓Is the question right for me?

- Will the question hold my interest?
- Can I manage any potential biases/subjectivities I may have?

O'Leary, Z. (2004) The Essential Guide to Doing Research. London: Sage

Good Question Checklist

- ✓ Is the question right for the field?

  - Will the findings be considered significant?
  - Will it make a contribution?

O'Leary, Z. (2004) The Essential Guide to Doing Research. London: Sage

Good Question Checklist

✓Is the question well articulated?

- Are the terms well-defined?
- Are there any unchecked assumptions?

O'Leary, Z. (2004) The Essential Guide to Doing Research. London: Sage

Good Question Checklist

✓Is the question doable?

- Can information be collected in an attempt to answer the question?
- Do I have the skills and expertise necessary to access this information? If not, can the skills be developed?
- Will I be able to get it all done within my time constraints?
- Are costs likely to exceed my budget?
- Are there any potential ethics problems?

O'Leary, Z. (2004) The Essential Guide to Doing Research. London: Sage

Example of RQ

MAIN RQ

- Can an AIS algorithm be developed that derives and maintains diversity

SUB RQ

1. What are the immunological properties that can address the problem of profile adaptation?
2. What is an effective mechanism to use for the representation of profile and information items?
3. How can we build an AIS that has the ability to adapt to changes in a user profile given multiple interests and radical changes in interests?
4. What is the baseline on which the performance of the AIS should be evaluated?

From RH TO RQ

RH

- Security risks will causes poor computer system's performance.

RQ

- What are the security risks for networked information systems?

From RH TO RQ

RH

"Scalability and speedup of large scale data integration solutions provide information actionable in near-real time online space weather analytic".

Main RQ:

"What are the mechanism to achieve scalability and speedup of large scale near-real time big data integration for space weather analytic and how it will be implemented through online analytic?"

Sub RQ:

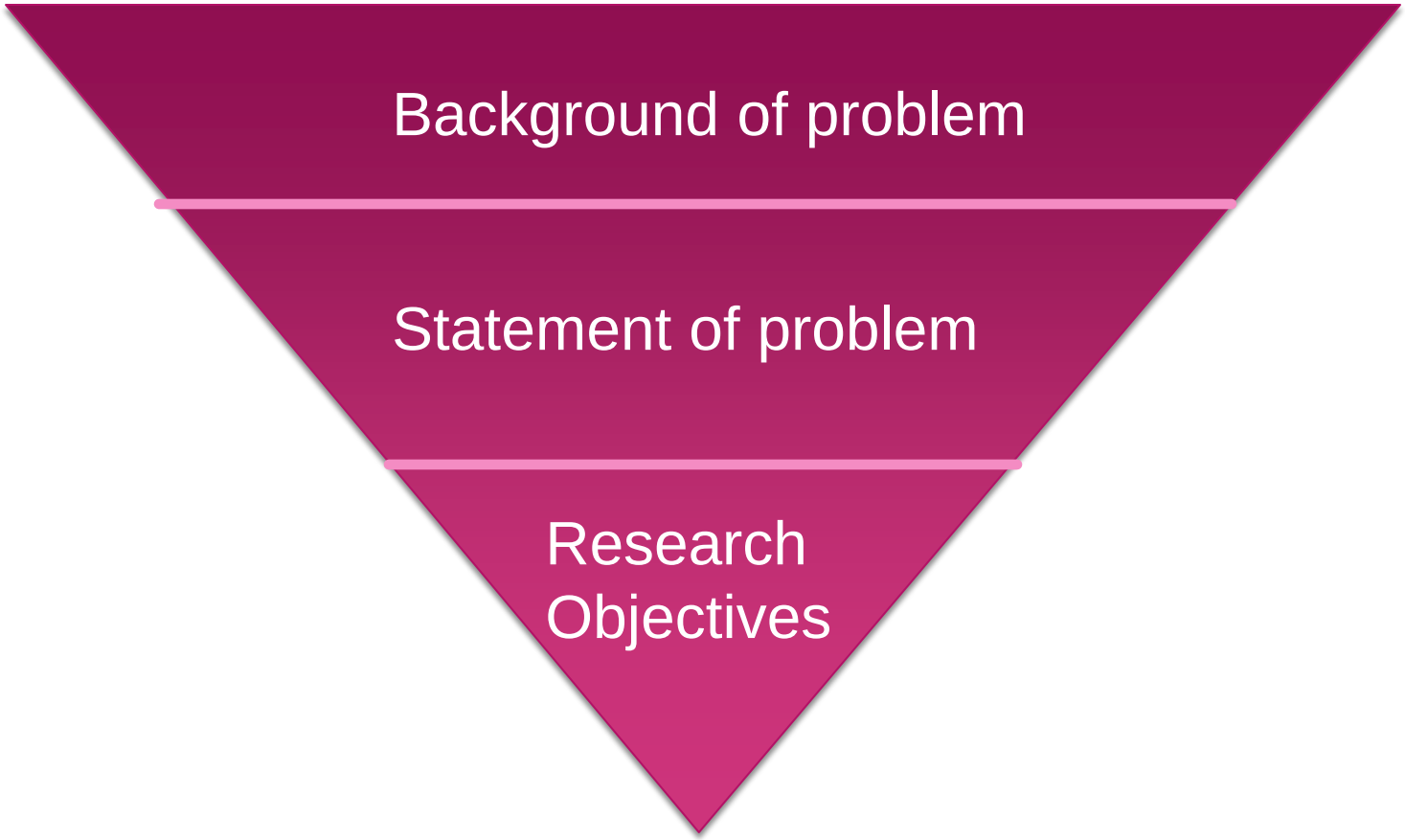
1.What are the architecture for online space weather prediction?
2.What are the components to derive big data integration algorithm for online analytic of space weather data?
3.How to integrate large scale space weather data input that have the features of heterogeneous, real time change data capture (CDC) and distribution, data transformation, data quality, data validation, and metadata management?
4.How the processes and analysis will be done in reaping the benefits of online space weather data analytic for large scale data?

RESEARCH OBJECTIVES (RO)

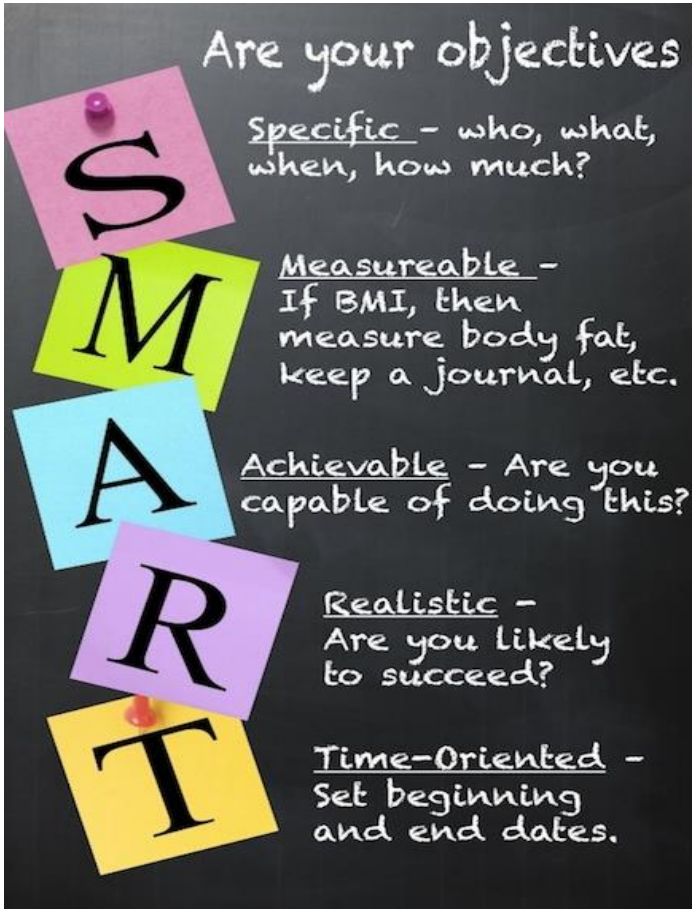
RESEARCH OBJECTIVES (RO)

- Clearly defined RO is important as it enlighten the way in which the researcher has to proceed.
- RO: is a clear, concise, declarative statement which provides direction to investigate the variables.
- Generally, RO focus on the way to measures the variables, such as to identify or describe them
- Objectives should be closely related to the statement of the problem
- Objectives is a purpose that can be reasonably achieved within the expected time frame and with the available resources

IDENTIFYING RESEARCH OBJECTIVES



CHARACTERISTICS OF RESEARCH OBJECTIVE



SMART Objectives

- **S**pecific: Be precise about what you are going to achieve
- **M**easurable: Quantify the objectives
- **A**chievable: Align with the needs of the target audience
- **R**ealistic: Do you have the resources to make the objective happen?
- **T**ime-Specific: State when you will achieve the objective

SMART: Specific Objectives

Specific: Be precise about what you are going to achieve

- Specify target
- Specify intended outcome
- Concrete
- Detailed
- Focused
- Well-defined
- Straightforward
- Action-oriented

When setting objectives that are **S**pecific, ask the following questions:

- What am I going to do? Use action verbs such as *develop, execute, conduct, build*.
- Why is it important to do this task?
- Who is going to be involved?
- When do I want this task to be completed?
- How am I going to do this task?

sMART: Measurable Objectives

- Help you know you are making progress toward completing the objectives.
- Will have tangible evidence of completion of the objective.

When setting objectives that are measurable, ask the following questions:

- How will I know when this objective has been achieved?
- What measurements can I use?
- What milestones can I use to track progress toward completion?

SM**A**RT: Achievable Objectives

- Achievable objectives are those that you can actually accomplish (something you can realistically do within the time frame set)—not an aspiration or vision.

SMA**R**T: Realistic Objectives

- Realistic relates to the relationship between the objective and the overall goals of the study/research

- A realistic goal can answer yes to these questions:
  - Does this seem worthwhile?
  - Is this the right time?
  - Does this match our other efforts/needs?
  - Is it possible to complete this objective?
  - Is it applicable in the current socio- economic environment?

SMART: Time-Specific Objectives

- Time-oriented objectives are those that have deadlines for completion. The time frames create a sense of urgency and lead to action.
- The deadlines, just as with overall objectives, must be achievable and realistic.
- For a complex objective, break it into small parts and set a deadline for completion of each phase.

When setting objectives that are time-specific, ask the following questions:

- What is the earliest—yet achievable and realistic—date for this objective to be started and to be completed?
- Have I included these dates in the statement of the objective?
- Are there other projects/objectives that must be completed first, or are there others that are reliant upon the completion of this objective?

Examples of research objectives

- Use action verb such as to describe, to evaluate, to develop

Examples:

- To identify the security risks for networked information systems
- To develop a risk evaluation tools for networked information systems
- To test the evaluation tool for networked information systems

From RQ to RO

RQ

- What are the security risks for networked information systems?
- What risk evaluation tools can be developed for networked information systems?
- How to test the evaluation tool for networked information systems?

RO

- To identify the security risks for networked information systems
- To develop a risk evaluation tools for networked information systems
- To test the evaluation tool for networked information systems

From RQ to RO

RQ

- What are the architecture for online space weather prediction?
- What are the components to derive big data integration algorithm for online analytic of space weather data?
- How to integrate large scale space weather data input that have the features of heterogeneous, real time change data capture (CDC) and distribution, data transformation, data quality, data validation, and metadata management?
- How the processes and analysis will be done in reaping the benefits of online space weather data analytic for large scale data?

RO

- To investigate the architecture of online space weather data analytic and data abstraction needs in deriving big data integration algorithms
- To design data integration model of large scale near-real time data inputs for space weather data.
- To develop big data integration algorithm for near-real time data using space weather data input.
- To evaluate the accuracy and timely performance of the proposed data integration algorithm.

Alignment of RO, RQ/RH and Method

Research Objectives (RO)?	Research Question (RQ) and /or Research Hypotheses (RH)?	How to answer the RQ or test the RH?

Example of Chapter 1

PROJECT BASED

RESEARCH BASED

SAMPLE OF RESEARCH PROPOSAL

KEMENTERIAN PENDIDIKAN TINGGI
MINISTRY OF HIGHER EDUCATION

APPLICATION FORM
FUNDAMENTAL RESEARCH GRANT SCHEME (FRGS)
Skim Geran Penyelidikan Fundamental
(Pindaan 1/2012)

JABATAN PENGAJIAN TINGGI
KEMENTERIAN PENDIDIKAN TINGGI

A. Application Details	
Application ID	144580-165190
A(i). Selected Grant	FRGS 2016-1
A(ii). Title Of Proposed Research Project	Improving Data Integration Algorithm for Online Analytics in Large Scale Malware Data using Big Data Approaches
A(iii). Keyword	Master Data Management (MDM), Data Integration, Online Analytics, Large Scale Data, Big Data Security Analytics

Research Title

Research Keyword

SAMPLE OF RESEARCH PROPOSAL

1. Problem Statement

In this era of terabyte data to Exabyte data and now towards yottabyte data to process, it forms the characteristics of volume (data at rest), velocity (data in motion-fastest streaming data), variety (data in many forms-structured, unstructured, text, multimedia) and veracity (data in doubt-uncertainty due to data inconsistency). The large availability of data potentially enables more powerful analysis and unexpected outcomes. Thus, data needs to be processed and condensed into more connected forms in order to be useful for event comprehension and decision making. This can be done in the form of data integration. Data integration means combining information from various sources into something useful. It's about efficiently managing data and making it available to

Introduction of The problem background

those who need it. However, with the large scale, online activities and various format of data, data integration solutions on online analytic must include heterogeneous, real time change data capture (CDC) and distribution, data transformation, data quality, data validation and metadata management. In addition, the algorithm for data integration should also taking into account the data scalability and time performance on online analytic of large scale data. Adapting these to online analytic is a challenging problem, mainly because of the fundamental design choices that favor throughput over latency. The purpose of this research is to provide the algorithm for data integration of large scale data for online analytic. Large scale data which will be the focus of this study is near and real time malware data, an example of Cyber-attacks type instead of others, namely, zero day attacks and advanced persistent threats (APT). The rapid growth of the Internet has brought with it an exponential increase in the type and frequency of cyber-attacks. Analyzing logs, network packets, and system events for forensics based on traditional tools had rigid and defined schema. In today security analytic, it need to go beyond logs, events and alerts but more types of information, including email and social activity, full-packet and Domain Name System captures, business process data, external threat feeds, malware information, network flows and anomalies. Thus, from the outcome of this fundamental study on online analytic of large scale data integration, it will have more reliable and available, and provide guarantees that queries are processed to completion. Therefore, it will improve research on security analytic which can make it possible to predict and stop cyber-attacks.

The problem/challenges

Purpose of the study

Significance of the study

SAMPLE OF RESEARCH PROPOSAL

2. Hypothesis

"Scalability and speedup of large scale data integration solutions provide information actionable in near-real time online security analytic".

Should have connection

3. Research Questions

The main research question derives from this study is
"What are the mechanism to achieve scalability and speedup of large scale near-real time big data integration for security analytic and how it will be implemented through online analytic?"

From the main research question, sub research questions are identified as below:
1.What are the architecture for online security analytic?
2.What are the components to derive ETL (extract transform load) and ELT (extract load transfer) data integration algorithm for online analytic of large scale malware data?
3.How to integrate large scale malware data input that have the features of heterogeneous, real time change data capture (CDC) and distribution, data transformation, data quality, data validation, and metadata management?
4.How the processes and analysis will be done in reaping the benefits of online security analytic for large scale data?

SAMPLE OF RESEARCH PROPOSAL

From the main research question, sub research questions are identified as below:

1.What are the architecture for online security analytic?

2.What are the components to derive ETL (extract transform load) and ELT (extract load transfer) data integration algorithm for online analytic of large scale malware data?

3.How to integrate large scale malware data input that have the features of heterogeneous, real time change data capture (CDC) and distribution, data transformation, data quality, data validation, and metadata management?

4.How the processes and analysis will be done in reaping the benefits of online security analytic for large scale data?

(c) Objective (s) of the Research

This study embark to the following objectives:

1. To investigate the architecture of online security analytic and data abstraction needs in deriving ETL and ELT data integration algorithms.
2. To improve data integration frameworks of large scale near-real time data inputs for online security analytic.
3. To propose an integration algorithm for ETL and ELT near-real time data using large scale malware data input for security analytic.
4. To evaluate the accuracy and timely performance of the proposed data integration algorithm.

SAMPLE OF RESEARCH PROPOSAL

(c) Objective (s) of the Research

This study embark to the following objectives:

1. To investigate the architecture of online security analytic and data abstraction needs in deriving ETL and ELT data integration algorithms. **S M**
2. To improve data integration frameworks of large scale near-real time data inputs for online security analytic. **S R**
3. To propose an integration algorithm for ETL and ELT near-real time data using large scale malware data input for security analytic. **A T**
4. To evaluate the accuracy and timely performance of the proposed data integration algorithm. **M T**

Project Objective

Following are the two objectives for the study:

(1) To ensure a profitable partnership between Celcom with Google Inc., Celcom has to develop a Direct Operator Billing (DOB) with Google which the transaction is a "commission based fees" NOT "revenue share" basis.

(2) To test on the system readiness directly to the Google Play Store on Android based with the actual Celcom subscriber.

Scope

Commented [hf7R6]: utk problem statement, perlu masukkan elemen purpose project iaitu software testing..

Commented [hf8]: suggest:
1.to study .... (software testing on apps)--- reflex to literature review
2. to identify test cases for testing activities --- reflex to software testing documentation and testing activities
3. to create software testing documentation based on standard...
4. to conduct testing on the system....

END OF PART 2