CSIT940
CS Research Methodology
Spring Session 2024

Experiment Design
(and theory base for research)

This set of slides was
edited from previous slides
used for this subject. --LeongHW

Highlights of this Lecture

- **Theory base for Research**

- Project Design

- Experimental Design

Theory base for Research

- **Definition of theory**
- Elements of a Good Theory
- Types of Theory Building
- How to choose/develop theory for your research

Bacharach, S. B. (1989). “Organizational Theories: Some Criteria for Evaluation,” Academy of Management Review (14:4), 496-515.

Definition of theory

Definition - Theory

A scientific theory is a system of constructs (or concepts) and propositions (relationships between those constructs) that collectively presents a logical, systematic, and coherent explanation of a phenomenon of interest within some assumptions and boundary conditions (Bacharach 1989).

Theories explain things and not simply predict.

Explanations require causation

Causation require:

correlations between two constructs,
temporal precedence (the cause must precede the effect in time),
rejection of alternative hypotheses (through testing)

Theoretical explanations must be able to generalize across many situations using a few variables

Building blocks of theory

Four building blocks of theory[1]:

Constructors: What concepts are important for explaining a phenonmena?

Propositions: How are the concepts related to each other?

Logic: Why are they connected to each other?

Assumptions / Boundary conditions: What circumstances under which the relationship is true?

Bacharach, S. B. (1989). "Organizational Theories: Some Criteria for Evaluation," Academy of Management Review (14:4), 496-515.

Building blocks of theory

Constructs are concepts specified at a high level of abstraction and chosen specifically to explain the phenomenon

* uni-dimensional (single concept e.g. length, height, weight, etc.)

* multi-dimensional (employ multiple underlying concepts e.g. personality, culture, advancement, texture, etc)

* defined such that there is clarity on how to measure and at what level of analysis (individual, group, organizational, etc.)

* Variables are measurable representations of abstract constructs

independent - explains other variables

dependent - explained by other variables

mediating - explained by independent variables while also explaining dependent variable

moderating -influence the relationship between independent and dependent variables

Building blocks of theory

Propositions are associations postulated between constructs based on deductive logic

* stated in declarative form and ideally indicating a cause-effect relationship

* may be conjectural but must be testable and rejectable if unsupported by empirical observations

* stated at empirical level as relationship betw variables (hypothesis)
  eg: we hypothesize that the life span of an automobile tyre depends on certain variables:

$$Y = \alpha X + \beta G + \gamma Z$$

where *Y* is life span of tyre (years); *X* is age of driver (years); *G* is km per year (km/annum) and *Z* is engine capacity of car (litres)

Building blocks of theory

Logic provides the basis for justifying the propositions as postulated

connects the theoretical constructs provides meaning and relevance to the relationships between constructs

Building blocks of theory

Theories are constrained by assumptions and boundary conditions that govern their applicability

* values - economic, political, etc.

* time - epoch, age group

* space - which region of the world, individual, groups, etc.

How to recognize a good theory

(1) Logical consistency

The theoretical constructs, propositions, boundary conditions, and assumptions must be **logically consistent**.

How to recognize a good theory

(2) Explanatory power

The theory must explain or predict reality to a large measurable extent and better than competing theories. It must be possible to quantitatively evaluate the explanatory power.

How to recognize a good theory

(3) Falsifiability

- A theory must be stated in a way that it can be disproven. Theories that cannot be tested or falsified are not scientific theories and any such knowledge is not scientific knowledge. A theory that is specified in imprecise terms or whose concepts are not accurately measurable cannot be tested, and is therefore not scientific.

How to recognize a good theory

(4) Parsimony

Parsimony examines how much of a phenomenon is explained with how few variables.

When there are multiple explanations of a phenomenon, the simplest or logically most economical explanation should be favoured.

This concept is called parsimony or **"Occam's razor"**.

Parsimony aims to go for the simplest explanation possible that will fully explain the phenomenon.

Parsimony prevents the adoption of overly complex or outlandish theories.

Approaches to theory building[2]

Grounded theory building

inductively based on observed patterns of events or behaviors

dependent on the observational and interpretive abilities of the researcher

care required to ensure theory is not subjective and non-confirmable

What happens if we use a machine learning approach to learn the model from data?

---

[2]Steinfield, C.W. and Fulk, J. (1990). "The Theory Imperative," in Organizations and Communications Technology, J. Fulk and C. W. Steinfield (eds.), Newbury Park, CA: Sage Publications.

Approaches to theory building[2]

Bottom-up conceptual analysis

inductively identify different sets of predictors relevant to the phenomenon of interest using a predefined framework

relies heavily on the inductive abilities of the researcher

interpretation may be biased by prior knowledge of researcher

a framework might be helpful in guiding the process and thinking through the development

---

[2]Steinfield, C.W. and Fulk, J. (1990). "The Theory Imperative," in Organizations and Communications Technology, J. Fulk and C. W. Steinfield (eds.), Newbury Park, CA: Sage Publications.

Approaches to theory building[2]

Extension or modification of existing theories to explain a new context

deductive approach leveraging large body of work on theories

concepts, propositions, and/or boundary conditions of the old theory may be retained or modified to fit new context

---

[2]Steinfield, C.W. and Fulk, J. (1990). "The Theory Imperative," in Organizations and Communications Technology, J. Fulk and C. W. Steinfield (eds.), Newbury Park, CA: Sage Publications.

Approaches to theory building[2]

Application of existing theories in entirely new contexts by drawing upon the structural similarities between the two contexts

deductive approach involving reasoning by analogy

[2]Steinfield, C.W. and Fulk, J. (1990). "The Theory Imperative," in Organizations and Communications Technology, J. Fulk and C. W. Steinfield (eds.), Newbury Park, CA: Sage Publications.

Choosing or developing a theory

Understand the problem and define it precisely

Identify the constructs and associated variables that underpin the problem (question)

Identify how the variables interplay

Check what others are doing about related problem - what model or theory?

Determine if what exists applies to your problem (question) with or without modification

Determine if there are prarallels to your problem in other fields that could help build your thory or model

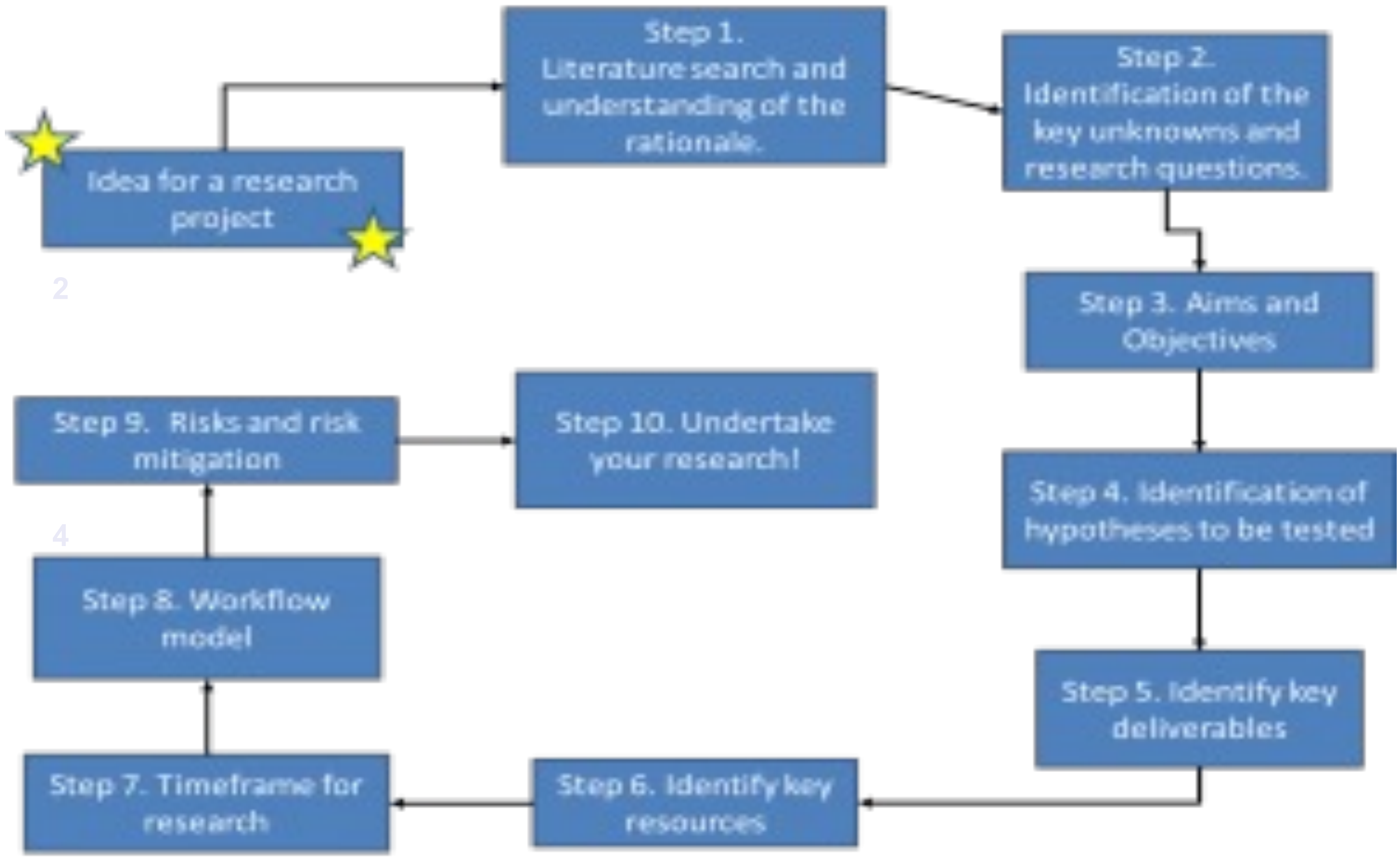
Select the theory or model most appropriate for your problem

Design your research method - data collection, analysis, evaluation, etc.

Highlights of this Lecture

- Theory base for Research
- **Research/Project Design**
- Experimental Design

A Good Research Design in 10 Steps



Step 1. Why are you doing this research?

Decide WHAT you are doing, WHY you are doing it.
Many people become so immersed in their project, they cannot see the wood for the trees, and assume that everybody knows why their work is important.

EXPLAIN (esp to non-specialists) WHY the work is important.
Your examiner will check if you understand the relevance.

Whether you are writing a proposal or an introduction to your thesis, you should **start with the rationale**.

Grants are awarded competitively on the basis of their relevance and importance. Papers are only published if they are relevant. The rationale is probably one of the most important parts of your research design, and you skim over it or ignore it at your peril.

Step 2. Identify the key unknowns

Once you have decided on the broad area of your research project, and you have established a good rationale and reason for undertaking it, **you need to read up on previous work. What are the key unknowns and key research questions?** What gaps can my research project fill? Write a 'Wider Justification' where you explore previous work, but where you identify gaps in knowledge.

The best projects are **novel, but not too novel** (which is risky and difficult).

Avoid 'me-too' science. Don't just jump on the bandwagon because everyone else is doing it.

Your project must have clearly defined research questions.

Equally, large, speculative science is unlikely to get funded.

Step 2. Identify the key unknowns

If you are doing a PhD dissertation or MSc thesis,
you should expect to spend a considerable amount of time at the start of your research project reading the available literature.

Make a list of research questions and key unknowns as you come across them. This is a vital step to becoming an expert in your area.

Finally, this stage is imperative to make sure that your research has not already been done!

Step 3. What is your aim and what are your objectives?

Once you have worked out a rationale for your research, you need to decide on an aim.

This is the most important part of your research design, and it should address
2 the key unknowns identified in Step 2 above.

3 Ideally, you should be able to express your aim in one sentence, e.g.,
Aim: to reconstruct the glacial history of the NE Antarctic Peninsula on centennial to millennial timescales.

4

Your objectives should help you to achieve your aim. You can identify, typically, 3-5 objectives that will each bring you a step closer to your achieving your aim.
5 Ideally, each objective should be associated with research questions so that you are always trying to achieve something new and original. This will also keep your research focused and on the right lines.

Step 3. What is your aim and what are your objectives?

Good aims and objectives usually have the following characteristics:

1

* Specific, achievable and feasible
* Clear sense of deliverables
* 2 Specific, clear, over-arching research question
* Realistic about methods and timescale available

3

Use words like

Compare, determine, characterise, explain, quantify, interpret, measure

4

Poor aims and objectives typically have the following characteristics:

Vague, broad, unspecified titles

5 No hypothesis or research question

Overambitious and not realistically achievable

Step 4. What hypotheses are you testing?

As scientists, we need to test hypotheses.
These hypotheses should be indentified by your analysis of previous work and key unknowns.

We work within a scientific Research Programme:
this means that there are key things that we hold to be true
(evolution, basic processes of glacier movement,etc),
and also areas continuously under development and being questioned
(details of past glacier history).

You need to write one or two hypotheses that you will **test**.

A good scientist should attempt to falsify her hypotheses. Your hypotheses should be based on the literature, your identification of the key knowns and unknowns, and should move the science forward.

I am still editing these slides.

--LeongHW

Step 5. Identify the key deliverables.

What are the key outcomes and deliverables of your research going to be? The deliverables should use words such as, *understanding, quantification, conceptual, process, analysis, characterisation, determination.* For example:

An improved understanding of process XX

Glacier velocity maps

Process-based conceptual models of process YY

Quantification of ZZ

Analysis of glacier fluctuations over time (glacier outlines submitted to GLIMS)

These deliverables should enable you to test your hypotheses and achieve your aim. They should be specific and achievable, and help you achieve closure with your 'Big Question'.

Step 6. Identify key resources

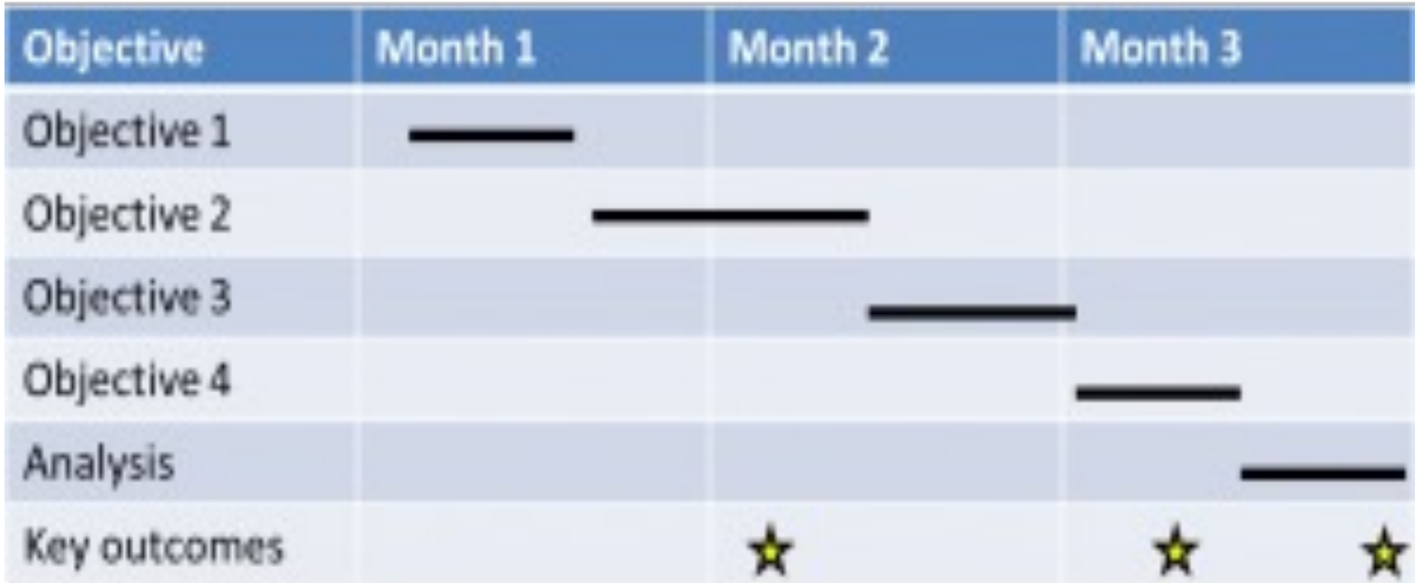
What resources will you need to complete this research
1 project? Will you need to do fieldwork, and if so, for how
long? Will you need any specific computer resources,
2 packages, programmes, remotely sensed images,
computer codes?

3

4

5

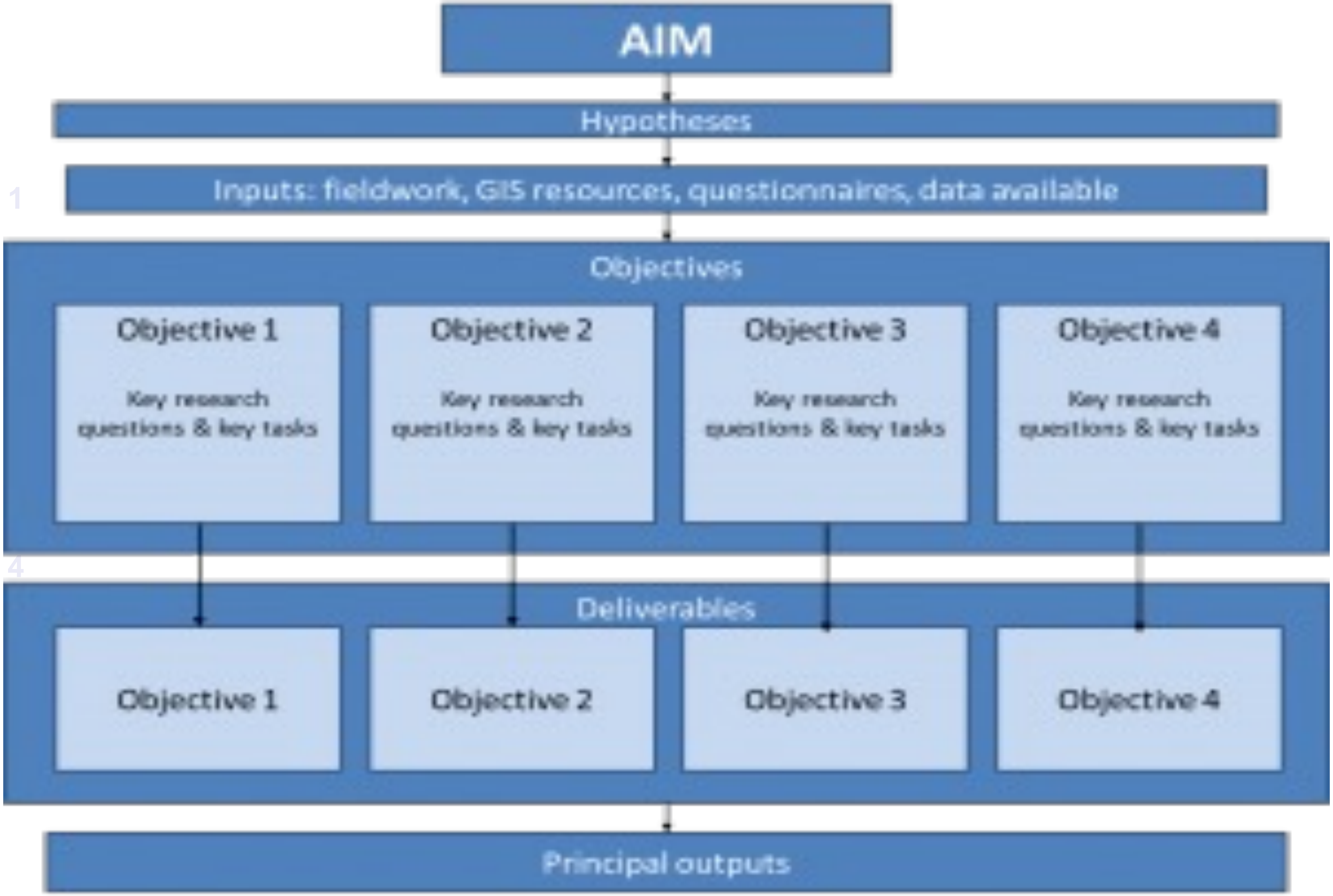
Step 7. Timeframe for research



Simplistic gantt chart (they do not need to be complicated). You should identify the key outputs in your caption. Needless to say, they should match the identified project deliverables.

Your research design should incorporate a realistic assessment of the time commitments for each objective. Write a Gantt Chart (you can just do this in Excel or on paper!) that outlines each objective and the amount of time you have available. Work out in detail how much time each objective will take you, and be realistic about whether it is achievable in your time available. Typically, students underestimate the amount of time a specific objective will take them, so be cautious in your estimation.

Step 8. Draw up a work-flow model



Step 8. Draw up a work-flow model

Once you have worked through the steps above, you are ready to put it together into a coherent workflow model. I think these should be included in all dissertations and grant applications, as they clearly set out how the different objectives fit together.

Write the aim at the top, and then the hypotheses beneath it. Include your resources or inputs below this. Then, each in a separate box, outline each objective and the key deliverables associated with this objective. Finally, at the bottom, give your end result; e.g., hypotheses accepted or refuted; a general model or process XX; quantification of YY. This makes it very clear how your research project will fit together, what you will achieve, and how it fits with your aims and hypotheses.

The workflow model should illustrate your ***deliverables*** and thus provide closure for your ‘Biq Question’.

Step 9. Risks and risk mitigation

If you are planning on doing fieldwork, you will need to do a risk assessment and clearly identify hazards and how you will mitigate
1 or prevent them. But you should also be aware of more general risks; do you have the relevant knowledge? Are the resources that
2 you need available? Will the costs change?

Risks could include:
3 * Unreliable exchange rates
* Wildlife hazards (polar bears?)
4 * Weather (hot / cold / blizzards)
* Environmental hazards and disposal of waste
5 * Equipment failure
* Not being able to obtain key datasets
* Access to field areas

Step 9. Risks and risk mitigation

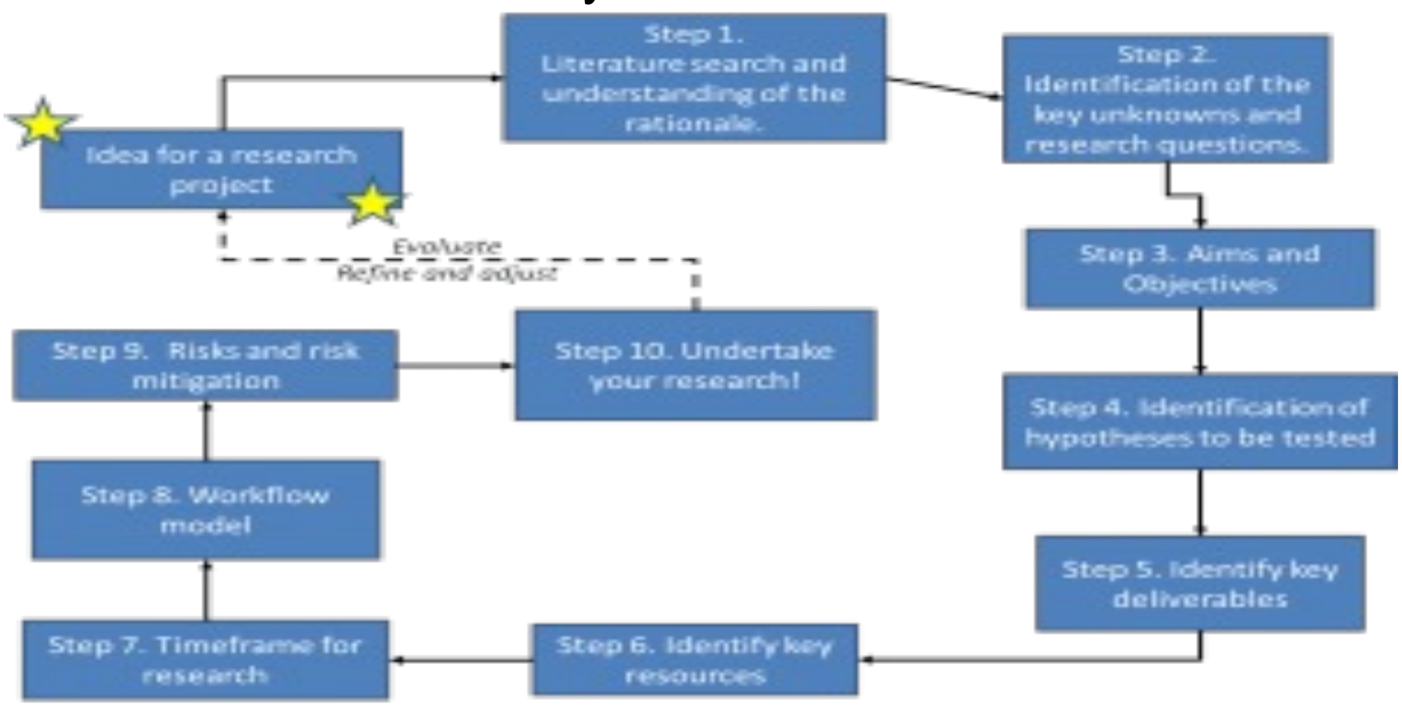
Adding my own:

Do you know the area?
Do you know the methods used?
Do you have the data?
Do you have the resources? Equipment, $$$,
Do you have enough time?
Etc, etc,

How to mitigate these risk?

Step 10. Undertake your research!

Now that you have spent some time carefully planning your research, you are ready to begin. You are going to research something that is important, interesting, and something that you will enjoy. You have written a good research plan, and know that your work is relevant to society and other scientists. Have fun!



Highlights of this Lecture

- Theory base for Research
- Research/Project Design
- **Experimental Design**

Experimental Design

1. What is experiment design?
2. Elements of good experiment design
3. Improving Internal and External Validity
4. Measurement of Constructs
5. Levels of measurement (Rating scales)

Definition of experiment design

Definition - Experiment Design

A comprehensive plan for data collection in an empirical research project.

It must specify the following processes: Data collection

Instrument development Sampling

Data collection

Data collection methods:

Positivist- including laboratory experiments, survey research, and aimed at theory or hypothesis testing

Interpretive- including action research, ethnography, and aimed at theory building

Positivist method is a deductive approach starting with a theory and aims to test theoretical postulates using empirical data

Interpretive method is inductive in nature and starts with data; hoping to derive theory about phenomenon of interest from observed data

Data collected can be quantitative or qualitative quantitative data involve numeric scores, metrics, etc. qualitative data includes interviews, observations, etc.

Attributes of a experiment design

(1) Internal validity

This is equivalent to causality and examines whether the observed change in a dependent variable is indeed caused by a corresponding change in hypothesized independent variable, and not by variables extraneous to the research context.

Causality requires:

Covariation of cause and effect Temporal precedence

Absence of plausible alternative explanation (or spurious correlation

Attributes of a experiment

(2) External validity

It is important that **experiment** design affords generalizability. Observed associations made on samples (data collected) should be generalizable to the population (population validity), or to other people, organizations, contexts, or time (ecological validity).

Attributes of a experiment design

(3) Construct validity

We want a chosen measurement scale to measure the theoretical construct that it is intended to measure very well (or accurately).

How do we measure empathy, resistance to change, and organizational learning?

Construct validity can be assessed in positivist research using correlational or factor analysis of pilot test data.

Attributes of a experiment design

(4) Statistical conclusion validity

It is expected that the conclusions derived using a statistical procedure are valid.

Methods

Five methods can be used to improve internal and external validity:

Manipulation- researcher manipulates the independent variables in one or more levels (called "treatments"), and compares the effects of the treatments against a control group where subjects do not receive the treatment.

Elimination- researcher eliminates extraneous variables by holding them constant across treatments. For example by restricting the study to a single gender, a single socio-economic status or small-to-medium enterprises.

Methods

Five methods can be used to improve internal and external validity:

Inclusion- the researcher considers the role of extraneous variables by including them in the research design and separately estimating their effects on the dependent variable. Factorial designs can be used in this case.

Statistical control- extraneous variables are measured and used as covariates during the statistical testing process.

Randomization- researcher uses random sampling to cancel out the effects of extraneous variables.

Random selection - select sample randomly from a population. Random assignment - subjects selected in a non-random manner are randomly assigned to treatment groups.

Randomization assures external validity by allowing inferences drawn from the sample to be generalized to the population from which the sample is drawn.

Example Research Designs

Experimental studies: intended to test cause-effect relationships (hypotheses) in a tightly controlled setting.

Separate the cause from the effect in time, administering the cause to one group of subjects (the “treatment group”) but not to another group (“control group”), and observing how the mean effects vary between subjects in these two groups.

Experimental design has strong internal validity due to its ability to isolate, control, and intensively examine a small number of variables;

Weakness lies in limited external generalizability.

Example Research Designs

Field surveys: non-experimental designs that do not control for or manipulate independent variables or treatments, but measure these variables and test their effects using statistical methods.

Capture snapshots of practices, beliefs, or situations from a random sample of subjects in field settings through a survey questionnaire or less frequently, through a structured interview.

Cross-sectional field surveys: independent and dependent variables are measured at the same point in time (e.g., using a single questionnaire)

Longitudinal field surveys: dependent variables are measured at a later point in time than the independent variables.

Field surveys have strong external validity because data is collected in field settings.

Non-temporal nature makes internal validity (cause-effect relationships) difficult to infer.

Example Research Designs

Secondary data analysis: analysis of data that has previously been collected and tabulated by other sources.

Effective means of research where primary data collection is too costly or infeasible, and secondary data is available at a level of analysis suitable for answering the researcher’s questions.

Limitations include:

uncertainty about data collection method - not systematic or in scientific manner;

collection for a presumably different purpose and not properly address research question;

uncertainty about internal validity if the temporal precedence between cause and effect is unclear.

Example Research Designs

Case research: in-depth investigation of a problem in one or more real-life settings (case sites) over an extended period of time.

Data collection method - combination of interviews, personal observations, and internal or external documents.

Method can discover a wide variety of social, cultural, and political factors potentially related to the phenomenon of interest that may not be known in advance.

Weakness lies in dependence of interpretation on observational and integrative ability of the researcher.

Example Research Designs

Focus group research: involves bringing in a small group of subjects (typically 6 to 10 people) at one location, and having them discuss a phenomenon of interest for a period of 1.5 to 2 hours.

Moderated and led by a trained facilitator, who sets the agenda and poses an initial set of questions for participants;

Builds holistic understanding of the problem situation based on comments and experiences of participants

Internal validity cannot be established due to lack of controls

Findings may not be generalized to other settings because of small sample size.

Example Research Designs

Action research: introduces interventions or “actions” into phenomena and observes the effect; characterized by simultaneous problem solving and insight generation.
Researcher is usually a consultant or an organizational member embedded within a social context.

Choice of actions must be based on theory, which should explain why and how such actions may cause the desired change.

Initial theory is validated by the extent to which the chosen action successfully solves the target problem.

Generalizability of findings is often restricted to the context where the study was conducted.

Example Research Designs

Ethnography: an interpretive research design inspired by anthropology.

Emphasizes research phenomenon must be studied within the context of its culture.

Researcher is deeply immersed in a certain culture over an extended period of time (8 months to 2 years); engaging, observing, and recording the daily life of the studied culture

Theorizes about the evolution and behaviors in that culture.

Conceptualization

Conceptualization: mental process by which fuzzy and imprecise constructs (concepts) and their constituent components are defined in concrete and precise terms.

Conceptualization process is important because of the imprecision, vagueness, and ambiguity of many social science constructs.

Unidimensional constructs have a single underlying dimension - person’s age or height.

Multidimensional constructs consist of two or more underlying dimensions.

Example: conceptualize a person’s English proficiency as consisting of 4 dimensions - listening, writing, reading and speaking. English proficiency is then considered as a multidimensional construct.

Each underlying dimension is measurable.

In science or computer science precise definition is given to concepts in order to be able to measure accurately.

Operationalization

Operationalization: the process of developing indicators or items for measuring these constructs. *For example the construct "Reaction time" may be operationalized by indcator that measures the time interval elapsed in performing an action.*

Allows us to examine the closeness amongst the indicators as an assessment of their accuracy (i.e. reliability)

Combination of indicators at the empirical level representing a given construct is called a variable.

(i) independent, (ii) dependent, (iii) mediating, or (iv) moderating. Indicator - reflective or formative.

Indicator may have several attributes (or levels) and each attribute represent a value.

Example: a "gender" variable may have two attributes: male or female. Example: a customer satisfaction scale may be constructed to represent five attributes: "strongly dissatisfied", "somewhat dissatisfied", "neutral", "somewhat satisfied" and "strongly satisfied".

Attribute values - quantitative (numeric) or qualitative (non- numeric).

Levels of Measurement

Levels of measurement: the values an indicator can take.

Rating scalesis another term used for levels.

Psychologist Stevens (1946) defined four generic types of rating scales for scientific measurements:

Nominal Ordinal Interval Ratio

Nominal

Nominal(or categorical) scales measure categorical data - gender, industry type, season, etc.

Indicators or variables have mutually exclusive attributes.

Measure of central tendency ismode;

Permissible statistics is chi-square and frequency distribution;

One-to-one transformation:-

$Spring = 1;\ Summer = 2;\ Autumn = 3;\ Winter = 4;$

Ordinal

Ordinalscales measure rank-ordered data - first, second, third, etc. - based on actual values.

Example: *“strongly agree”*, *“somewhat agree”*, *“neutral”*, *“somewhat disagree”*, *“strongly disagree”*.
Actual or relative values hidden.

Measure of central tendency aremodeandmedian;

Permissible statistics are percentile and non-parametric analysis.

Rank-preserving transformation :- monotonically increasing function.

Interval

Intervalscales measure rank-ordered data where measured values are equidistant from adjacent attributes.

Allows us to examine “how much more” is one attribute when compared to another - not possible with ordinal or nominal scales.

Example: Temperature scale where difference between 10^oC and 40^oC is the same as that between 20^oC and 50^oC

Measures of central tendency includemode,median;mean; Measures of dispersion include standard deviation and range.

Permissible statistics are percentile; non-parametric; chi-square and frequency distribution; regression; correlation; analysis of variance.

Allowed scale transformation are positive linear.

Ratio

Ratioscales have all the qualities of nominal, ordinal and interval scales as well as a “true zero”.

Named a “ratio” because the ratio of two points on the measures are meaningful and interpretable.

The “zero” implies that the underlying construct is not present.

Example: Mass, Electric charge, Age, Firm size. Measures of central

tendency includemode,median;mean,
geometric mean,harmonic mean; Measures of dispersion include standard deviation and range.

All statistical methods are allowed.

Allowed scale transformation are positive similar (logarithmic or multiplicative).

Binary scale

Binaryscales are nominal scales consisting of binary items that assume one of two possible values.

Have you ever written a letter to a public official	Yes	No
Have you ever signed a political petition	Yes	No
Have you ever donated money to a political cause	Yes	No
Have you ever donated money to a candidate running for public office	Yes	No
Have you ever written a political letter to the editor of a newspaper or magazine	Yes	No
Have you ever persuaded someone to change his/her voting plans	Yes	No

Figure : A six-item binary scale ; political activism measurement [1].

Can you find other examples? Discussion!

Likert scale

Likertscales are used for measuring ordinal data.

	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
I feel good about my job	1	2	3	4	5
I get along well with others at work	1	2	3	4	5
I'm proud of my relationship with my supervisor at work	1	2	3	4	5
I can tell that other people at work are glad to have me there	1	2	3	4	5
I can tell that my coworkers respect me	1	2	3	4	5
I feel like I make a useful contribution at work	1	2	3	4	5

Figure : Likert scale measuring employment self-esteem [1].

Can you find other examples? Discussion.

Semantic differential scale

Semantic differentialscale is composite or multi-item in nature and used to capture opnions or feelings toward a single statement. Pairs of adjectives framed as polar opposites are used.

How would you rate your opinions on national health insurance?						
	Very much	Somewhat	Neither	Somewhat	Very much	
Good	□	□	□	□	□	Bad
Useful	□	□	□	□	□	Useless
Caring	□	□	□	□	□	Uncaring
Interesting	□	□	□	□	□	Boring

Figure : Semantic differential scale measuring attitude toward national health insurance [1].

Can you find other examples? Discussion.

Guttman scale

Guttman scale is composite and uses a series of items in increasing order of intensity of the construct of interest; from least to most intense.

Each item has a weight that is used in forming an aggregate score; weighted combination of each response.

How will you rate your opinions on the following statements about immigrants?		
Do you mind immigrants being citizens of your country	Yes	No
Do you mind immigrants living in your own neighborhood	Yes	No
Would you mind living next door to an immigrant	Yes	No
Would you mind having an immigrant as your close friend	Yes	No
Would you mind if someone in your family married an immigrant	Yes	No

Figure : Five-item Guttman scale measuring attitude toward immigrants [1].

Can you find other examples? Discussion.


Thank you!