RESEARCH METHODS IN I/O PSYCHOLOGY

Objectives

Understand Empirical Research Cycle

Knowledge of Research Methods

Conceptual Understanding of Basic Statistics
Goals of the scientific method

- Description
- Explanation
- Prediction
Characteristics of the scientific method

- Objective
- Control
- Use of operational definitions
- Replication
- Generalizability
The Empirical Research Cycle

Research process — summarized as 5-step sequence

1. Statement of the problem
2. Design of research study
3. Measurement of variables
4. Analysis of data
5. Conclusions from research
1. **Statement of the problem**

- deductive [theory to data]
  - Theory → Hypothesis → Observation → Confirmation

- inductive [data to theory]
  - Observation → Pattern → Tentative Hypothesis → Theory
The Empirical Research Cycle

Research process — summarized as 5-step sequence

1. Statement of the problem
2. Design of research study
3. Measurement of variables
4. Analysis of data
5. Conclusions from research
2. Research design

A. Naturalness of the research setting

- lab or field
- "naturalness" or "artificiality" of the setting
- field research typically employs a real-life setting, while lab research is more contrived and artificial
B. Degree of control

- confounding and extraneous variables
- manipulation—this is reflective of a high degree of control
- research designs that permit manipulation are technically referred to as "experiments"
### Design of the Research Study

#### Setting

<table>
<thead>
<tr>
<th>Artificial</th>
<th>Natural</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CONTROL</strong></td>
<td><strong>CONTROL</strong></td>
</tr>
<tr>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>lab experiment</td>
<td>field experiment</td>
</tr>
<tr>
<td>observation (expt)</td>
<td>quasi-experiment</td>
</tr>
<tr>
<td></td>
<td>correlational dsgn</td>
</tr>
<tr>
<td></td>
<td>survey (questionnaire)</td>
</tr>
<tr>
<td></td>
<td>observation (non-expt)</td>
</tr>
</tbody>
</table>
Experimental Research

• **experiment**
  - a research method in which the investigator manipulates a variable under carefully controlled conditions and observes whether changes occur in a second variable
  - experiments are used in order to detect *cause-and-effect* relationships

• **conditions that make a true experiment**
  - manipulation of independent variables
  - random assignment into experimental conditions (experimental conditions & control)
Experimental and Control Groups

- **experimental group**
  participants who receive some special treatment in regard to the independent variable

- **control group**
  participants who do not receive the special treatment given to the experimental group
Experimental and Control Groups

• the logic of the experimental approach
  if the two groups are identical in all respects except for the variation created by the manipulation of independent variable, then any differences between the experimental and control groups on the dependent variable must be due to the manipulation of the independent variable
Effectiveness of the After-Action Review (AAR) as a Training Protocol

Sample

Experimental
AAR

Control
No AAR

Measure DV
AAR has higher scores
Advantages and Disadvantages of Experimental Research

• + permit cause-and-effect conclusions
• + can be either field- or lab-based
• - lab experiments tend to be artificial
• - cannot be used to explore some research questions
Design of the Research Study

Non-Experimental — Quasi-Experiments

- participants must be and are selected for different conditions from pre-existing groups
- levels of the IV are/may be selected from pre-existing values and not created through manipulation by the researcher
- unlike true experimental designs where participants are randomly assigned to experimental and control groups, with quasi-experimental designs they are NOT
- quasi-experiments DO NOT permit the researcher to control the assignment of participants to conditions or groups
Greenberg: Employee Theft and Underpayment Inequity

<table>
<thead>
<tr>
<th>Plant 1</th>
<th>Plant 2</th>
<th>Plant 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>Condition 1</td>
<td>Condition 2</td>
</tr>
<tr>
<td>No cut in pay</td>
<td>Inadequate explanation</td>
<td>Adequate explanation</td>
</tr>
</tbody>
</table>

DV
Employee theft
Greenberg’s Results

Design of the Research Study

Mean Percentage of Employment Theft

Time Period Relative to Pay Cut

Before During After

- Inadequate explanation
- Adequate explanation
- Control
Non-Experimental — Observational

- careful, usually prolonged, observation of behavior without intervening directly with the research participants
  - no manipulation by researcher
  - no random assignment
- often referred to as ex post facto designs
Design of the Research Study

Non-Experimental — Correlational Designs

• measure two or more variables and attempt to determine the degree of relationship between them

• characterized by:
  ▪ no manipulation
  ▪ low control
  ▪ no causal inferences

• do not confuse correlational designs with correlations
  – can have one without the other—that is, one does not automatically imply the other
Non-Experimental — Survey (questionnaire [?]) Research

- measurement and assessment of opinions, attitudes, and other descriptive phenomenon usually by means of questionnaires and sampling methods
Advantages and Disadvantages of Descriptive Non-Experimental Designs

• + give researchers a way to explore questions that experimental approaches cannot investigate

• - do not permit cause-and-effect conclusions
Design of the Research Study

- Primary vs. secondary (meta-analysis) research methods
validity

- a key (maybe THE key) criterion in the evaluation of any piece of research or test (measure)
- the appropriateness of inferences drawn from data
  - data = results of research study ⇒ research validity
  - data = test scores ⇒ test and measurement validity
Design of the Research Study

Research Validity

- A conclusion based on research is valid when it corresponds to the actual or true state of the world.
- Facets of research validity:
  - Internal
  - External
  - Statistical-conclusion
  - Construct
- Threats to research validity
- Controls
Design of the Research Study

• **internal validity**—is the extent to which we can infer that a relationship between two variables is causal or that absence of a relationship implies absence of cause. The extent to which observed relationship obtained from research design/study is real or artifactual.

<table>
<thead>
<tr>
<th>Threats</th>
<th>Controls?</th>
</tr>
</thead>
<tbody>
<tr>
<td>• history</td>
<td>• mortality</td>
</tr>
<tr>
<td>• maturation</td>
<td>• selection</td>
</tr>
<tr>
<td>• testing</td>
<td>• regression</td>
</tr>
</tbody>
</table>
Design of the Research Study

- **external validity**—is the inference that presumed causal relationships can be generalized to and across alternate measures of cause and effect, and across different types, persons, settings, and times. That is, how generalizable are the findings?

<table>
<thead>
<tr>
<th>Threats</th>
<th>Controls?</th>
</tr>
</thead>
<tbody>
<tr>
<td>● population validity</td>
<td></td>
</tr>
<tr>
<td>● ecological validity</td>
<td></td>
</tr>
<tr>
<td>● temporal validity</td>
<td></td>
</tr>
</tbody>
</table>
Design of the Research Study

- statistical conclusion validity/inference—appropriateness of inferences (or conclusions) made from data as a result (or function) of conclusions drawn from statistical analysis. That is, are the IV and DV statistically related?

**Threats**

- low statistical power
- violation of test assumptions
- poor test reliability

**Controls?**
• **construct validity**—has to do with labels that can be placed on what is being observed and the extent to which said labels are theoretically relevant. It is a question of whether the research results support the theory underlying the research. That is, is there another theory that could adequately explain the same results?

<table>
<thead>
<tr>
<th>Threats</th>
<th>Controls?</th>
</tr>
</thead>
<tbody>
<tr>
<td>● loose connection between theory and study</td>
<td></td>
</tr>
<tr>
<td>● evaluation apprehension</td>
<td></td>
</tr>
</tbody>
</table>
| ● experimenter expectancies (“good-subject” response) | }
• good research design or study
  – free from threats
  – no alternative explanations
  – permits robust conclusions about relationships between study’s variables
The Empirical Research Cycle

Research process — summarized as 5-step sequence

- Statement of the problem
- Design of research study
- Measurement of variables
- Analysis of data
- Conclusions from research
3. Variables and Their Measurement

• What is a variable?
  ▪ some property of an object, phenomenon, or event whose measurement can take on two or more values
• Types of variables
  ▪ independent/dependent
  ▪ predictor/ criterion
  ▪ continuous/discrete (artificially and naturally discrete)
  ▪ quantitative/qualitative (categorical)
Measurement of Variables

• Levels of measurement

Tests, measurement devices, or systems are tools used to assess a person's score or status on a variable.

The five basic levels of measurement:

• labels
• nominal
• ordinal
• interval
• ratio
• Characteristics of good measurement—a good test or measurement system should be:
  ▪ reliable
  ▪ valid
  ▪ objective
  ▪ standardized
Psychometric properties (of test scores)

- reliability of scores
- validity of scores
Reliability

• consistency or stability of scores/measure
• consistency of scores obtained by the same person when examined with the same test (or equivalent forms) on different occasions, time, places, etc.
• measurement error!!
• reliability, like validity, is based on correlations
Methods for assessing the reliability of test scores

- test-retest reliability (temporal consistency/stability)
  - involves the repeated administration of the same test to the same sample—issues?
• alternate-form reliability (temporal consistency/stability, and inter-form consistency/equivalence)
  ▪ a measure of the extent to which 2 separate forms of the same test are equivalent
  – issues?
• split-half, odd-even (or random split) reliability (internal consistency)
  ▪ the primary issue here is one of obtaining comparable halves
  – issues?
  – Spearman-Brown correction
Measurement of Variables

- Coefficient alpha (Cronbach’s alpha) (inter-item [and also internal] consistency)
  - these are measures of inter-item consistency, (i.e., the consistency of responses to all items on the test)
  - b/c average of all possible split-halves is equal to the coefficient alpha, they are also considered to be measures of internal consistency
  - issues??
- issues
  - indication of the extent to which each item on the test measures the same thing as every other item on the test
  - the more homogeneous the domain (test), the higher the inter-item consistency
• scorer or inter-rater reliability or agreement
  ▪ extent to which 2 or more raters are consistent or agree
  ▪ also extent to which same rater is consistent over time (intra-rater reliability)
    – issues??
Summary of $r_{xx}$

- $r_{xx} = \text{property of scores and not tests}$
- choice of specific $r_{xx}$ metric depends on:
  - facet of consistency of interest
  - construct of interest
  - resources available
- good level of $r_{xx}$?
  - $\geq .80$ for applied decision making
  - $\geq .70$ (or .65?) for research purposes
- $r_{xx} = \text{a necessary but not sufficient condition to justify/support specified test use}$
Validity

• appropriateness of inferences drawn from scores

• for specified use, it is an assessment of WHAT the test measures and HOW WELL it does so
Strategies for assessing test validity (i.e., validation techniques/strategies)

- criterion-related
- content-related
- construct-related
- meta-analysis & validity generalization
- synthetic validity/j-coefficients
- face validity
Measurement of Variables

![Diagram showing the relationship between predictor measures, criterion measures, predictor constructs, and performance domains.]
Measurement of Variables

<table>
<thead>
<tr>
<th>Predictor Measure</th>
<th>Criterion Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

- 1: predictor measure to criterion measure
- 2: predictor construct domain to predictor measure
- 3: predictor construct domain to performance domain
- 4: criterion measure to performance domain
- 5: criterion measure to predictor measure
Criterion-related validity

- effectiveness of a test in predicting an individual's behavior in specific situations
  - test score is intended as an indicator or predictor of some other behavior (that typically will not be observed until some future date)
  - test or predictor score is correlated with a criterion (i.e., a direct and independent measure of that which the test is designed to predict)
Criterion-related validity

- criterion-related, validity coefficients can range from \(-1.0\) to \(+1.0\)
- absolute value is used to compare different validity coefficients
• criterion-related validation designs
  ▪ predictive
  ▪ concurrent
  ▪ postdictive
Measurement of Variables

<table>
<thead>
<tr>
<th>Predictor Measure</th>
<th>Criterion Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance Construct Domain</td>
<td>Performance Domain</td>
</tr>
</tbody>
</table>

Arrows indicate relationships and measurement processes.
Content-related validity

- effectiveness of a test in predicting an individual's behavior in specific situations
  - adequacy with which a specified content domain is sampled
  - degree to which a predictor covers a representative sample of the behavior being assessed (e.g., knowledge-based tests such as classroom tests)
Content-related validity

- systematic examination of test content to determine whether it covers a representative sample of the behavior domain being measured
- typically rational and nonempirical, in contrast to criterion-related validity which is empirical
- content domain to be tested should be fully described in advance in very specific terms
Measurement of Variables

Diagram:

1. Predictor Measure → Criterion Measure
2. Predictor Construct Domain ↔ Predictor Measure
3. Predictor Construct Domain ↔ Performance Domain
4. Performance Domain ↔ Criterion Measure
5. Predictor Measure ↔ Performance Domain
Measurement of Variables

<table>
<thead>
<tr>
<th>Measure</th>
<th>Criterion Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. 1</td>
<td>B. 2</td>
</tr>
<tr>
<td>C. 4</td>
<td>D. 5</td>
</tr>
</tbody>
</table>

Diagram:
- Predictor Measure
- Criterion Measure
- Predictor Construct Domain
- Performance Domain

Arrows indicate relationships:
- 1 from Predictor Construct Domain to Criterion Measure
- 2 from Predictor Measure to Predictor Construct Domain
- 3 from Performance Domain to Predictor Construct Domain
- 4 from Criterion Measure to Performance Domain
- 5 from Predictor Construct Domain to Predictor Measure
Construct-related validity

- extent to which the test may be said to measure a theoretical construct or trait
  - **convergent validity**—different measures of the same construct should be correlated or related to each other
  - **discriminant validity**—different measures of different constructs should not be correlated or related to each other
### Measurement of Variables

#### Multi-Trait/ Multi-Method Matrix (MTMM) approach

**TRAIT**

<table>
<thead>
<tr>
<th>Trait</th>
<th>Paper-and-pencil</th>
<th>Assessment Center</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conscientiousness</td>
<td>NEO-FFI</td>
<td>Assessor Ratings</td>
</tr>
<tr>
<td>Cognitive ability</td>
<td>Wesman Personnel Classification Test</td>
<td>Assessor Ratings</td>
</tr>
</tbody>
</table>

A, B, C, D represent different methods or instruments used for measurement.
• **face validity**

If you had to be tree, what type of tree would you be and why?

You have to be at work by 7:30 a.m. It is typically a 20-minute drive to work. You leave your apartment at 7:00 a.m., go down to start your car, and it does not start. What would you do and why?
Summary of $r_{xx}$ and $r_{xy}$

- A test can be reliable but not valid.
- However, a test cannot be valid but not reliable—a test that does not correlate with itself cannot be expected to correlate with anything else.
- Thus, reliability is a necessary but not sufficient condition for validity.
  - Reliability sets the upper limit of validity.
    - Square root of a test’s reliability is the upper limit of its validity.
Summary of $r_{xx}$ and $r_{xy}$

- test with unknown reliability and validity is to be avoided
- reliabilities cannot be negative; but validity coefficients can be negative
- $r_{xx}$ = property of test scores not tests
- $r_{xy}$ = use of the test and not an inherent property of the test
Summary of $r_{xx}$ and $r_{xy}$

Interpreting the magnitude of correlations and effect sizes: MUST BE CONTEXT-SPECIFIC!!

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r$</td>
<td>$d$</td>
<td>$r$</td>
</tr>
<tr>
<td>small</td>
<td>.10</td>
<td>.20</td>
<td>&lt; .20</td>
</tr>
<tr>
<td>medium</td>
<td>.30</td>
<td>.50</td>
<td>.20 to .30</td>
</tr>
<tr>
<td>large</td>
<td>.50</td>
<td>.80</td>
<td>&gt; .30</td>
</tr>
</tbody>
</table>

- reliability
  - $\geq .80$ for applied decision making
  - $\geq .70$ for research purposes
Summary of $r_{xx}$ and $r_{xy}$

Interpreting the magnitude of correlations and effect sizes:

**MUST BE CONTEXT-SPECIFIC!!**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r$</td>
<td>$d$</td>
</tr>
<tr>
<td>small</td>
<td>.10</td>
<td>.20</td>
</tr>
<tr>
<td>medium</td>
<td>.30</td>
<td>.50</td>
</tr>
<tr>
<td>large</td>
<td>.50</td>
<td>.80</td>
</tr>
</tbody>
</table>

- reliability
  - ≥ .80 for applied decision making
  - ≥ .70 for research purposes
The Empirical Research Cycle

Research process — summarized as 5-step sequence

- Statement of the problem
- Design of research study
- Measurement of variables
- Analysis of data
- Conclusions from research
4. Statistical (Empirical) Analyses of Data

- Statistical tests are procedures that are used to:
  - describe data
  - analyze relationships between variables (i.e., make inferences)
## Analysis of Data

### Statistical Tests and Procedures

<table>
<thead>
<tr>
<th>Descriptive Statistics</th>
<th>Inferential Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Measures of Central Tendency</strong></td>
<td><strong>Measures of Variability (Dispersion)</strong></td>
</tr>
<tr>
<td>mean</td>
<td>variance</td>
</tr>
<tr>
<td>median</td>
<td>standard deviation</td>
</tr>
<tr>
<td>mode</td>
<td>percentiles</td>
</tr>
<tr>
<td>range</td>
<td>min</td>
</tr>
</tbody>
</table>
Normal Curve

Analysis of Data

-3SD -2SD -1 SD Mean +1 SD +2 SD +3 SD

68% 68% 95% 99%

68%

95%

99%
The Concept of Correlation

- permit investigators to see whether there is a link or association between the variables of interest
- do not permit cause-and-effect conclusions
The Concept of Correlation

- a correlation exists when two variables are related to one another
- two aspects of a correlation coefficient
  - direction - positive or negative
  - strength - range of coefficients is from -1.00 to +1.00
Positive Correlation

high school GPA

college GPA
Analysis of Data

Negative Correlation

age

hair on head
Positive Correlation

Cognitive ability

Job performance
Analysis of Data

Negative Correlation

Role ambiguity

Job satisfaction
Analysis of Data

Strong Positive Correlation

(a) Strong positive correlation

\[ \text{(11 and 21)} \]
Analysis of Data

Strong Negative Correlation

![Graph showing a strong negative correlation between hours practicing on the piano and errors during recital. The graph depicts a descending trend, indicating that as hours of practice increase, errors decrease significantly.](image-url)
Height and Weight; $r = .71$
Analysis of Data

$r = -.09$

![Scatter plot showing the relationship between shoe size and GPA. The correlation coefficient is $r = -.09$. The points on the scatter plot suggest a weak negative correlation between the two variables.]
Interpreting Correlations

- correlation and prediction
- correlation and causation
  - X may cause Y
  - Y may cause X
  - Z causes both X and Y
- squaring the coefficient tells you how much of the variability in X is associated with the variability in Y
Analysis of Data

Prediction

![Graph showing the relationship between cognitive ability and job performance. The graph includes a scatter plot with a linear regression line, indicating a positive correlation between the two variables.](attachment:image.png)
Variance Explained

Cognitive ability

Job performance
Summary

• correlation coefficients
  ▪ the degree of linear relationship (association) between 2 variables
  ▪ $r$
  ▪ -1.0 to +1.0
  ▪ correlation does not mean causality
    • causality determined through laboratory or statistical control
  ▪ statistical vs. practical significance
The Empirical Research Cycle

Research process — summarized as 5-step sequence

1. Statement of the problem
2. Design of research study
3. Measurement of variables
4. Analysis of data
5. Conclusions from research
5. Conclusions

- theoretical and applied implications
- limitations
  - generalizability
  - size and representativeness of sample
  - research method, protocol, and paradigm
- suggestions for future research
Summary

• good research study
  – free from threats
  – no alternative explanations
  – permits robust conclusions about relationships between study’s variables
Ethics in Research

- Truth in reporting
- Treatment of research participants

- Additional professional and legal issues associated with I/O practice
  - e.g., obligations to mgmt vs. employees