

# Validity 2

## Protecting Internal Validity, Describing External Validity, and the In-Basket Example

---

---

---

---

---

---

---

---

### Protecting Internal Validity

1. Design 1: Pretest-posttest control group design

Random assignment is always your first choice for internal validity.

R	O1		O2
	Time →		
R	O3	X	O4

R = random assignment; O = measurement or observation; X = treatment.  
 Controls for History, Maturation, Testing & Regression. Checks on Mortality but not Instrumentation. Pretest allows view up front.

---

---

---

---

---

---

---

---

### Protecting Internal Validity (2)

- Solomon Four-Group Design

R	O1		O2
R	O3	X	O4
R			O5
R		X	O6

Design 2

World's best design for control (but maybe not analysis).  
 Same as pretest-posttest, but 2 more groups with no pretest.  
 Last 2 groups will not have Testing confound.

---

---

---

---

---

---

---

---

### Protecting Internal Validity (3)

- Randomized posttest only design.

R			O5	Design 3
R		X	O6	

Because of random assignment, this is a good design, even though there is no pretest.

Design 1	R	O1		O2	Design 2
	R	O3	X	O4	
Design 3	R			O5	Design 2
	R		X	O6	

---

---

---

---

---

---

---

---

### Review

- Describe the 3 good designs for protecting internal validity
  - Randomized pretest-posttest
  - Solomon Four Group
  - Randomized Posttest Only
- How are the three related?

---

---

---

---

---

---

---

---

### External Validity

- Here we want to generalize from our study to something beyond it. Our study is bound by people, place and time. Our desire is not.
- Population Generalization. People. Ex - college sophomores; normals for therapy study. Solve by using other populations.

---

---

---

---

---

---

---

---

**External Validity (2)**

- Environmental Generalization. Place, situation, context. Ex – management in the laboratory, e.g., performance appraisal. Fix context depending on the problem – e.g., consequence of managerial decision.
- Temporal Generalization. Time. Ex – effects of clothes on hiring decisions (fashion changes). Solve by replication.

---

---

---

---

---

---

---

---

**Review**

- Provide an example of the three types of external validity (generalizability) concerns
  - People
  - Context
  - Time

---

---

---

---

---

---

---

---

**In-Basket Study**

- Background
- Assessment Centers – select or develop managers. Used by business & psychology.
- Use multiple exercises, traits, assessors
- In-basket is a simulation of a manager’s desk. Collection of items to be addressed. Measures traits like planning & organizing, perceptiveness, delegation, decision making and leadership.

---

---

---

---

---

---

---

---

### Study Purpose

- Can we train better performance on an in-basket with minimal effort & time?
- Do scores on the in-basket refer to enduring individual differences or to short-term differences in experience?
- How reliable are scores obtained from the in-basket?

---

---

---

---

---

---

---

---

### Method

- Participants
  - Business students at USF (both u.g & MBA; N=88). Some with managerial experience; all moving toward higher mgmt.
- Training
  - Workbook. Self-paced. Based on another in-basket. Focused on:
    - Skim & set priorities
    - Make connections among related items
    - Delegate properly

---

---

---

---

---

---

---

---

### Method (2)

- Materials (aka apparatus)
- In-baskets
  - 2 forms (A & B)
  - Development of forms
  - Scoring Keys
  - Scorers

---

---

---

---

---

---

---

---

### Method (3)

- Design (2 forms, A and B of in-basket)

Time 1	1 month	Time 2	
G1: take A		Train	Take B
G2: take B		Train	Take A
G3: take A		Take B	Train
G4: take B		Take A	Train

Note. Everyone got trained, but we still got to test for training effects.

---

---

---

---

---

---

---

---

---

---

### Results

- Reliability and correlations (form A)

1. Org & Plan	1	2	3	4	5
	(.71)	( interjudge reliability)			
2. Leadership	.68	(.88)			
3. Perceptive	.29	.45	(.80)		
4. Decisions	.51	.68	.52	(.82)	
5. Delegation	.26	.15	.07	.23	(.83)
Mean	3.2	2.7	1.5	2.6	2.5
S.D.	2.21	2.63	3.39	3.25	2.44
Alpha (reliability)	.38	.44	.46	.49	.72

1. Interjudge reliability is high.
2. Internal consistency is low.
3. Correlations among scales are moderate.

---

---

---

---

---

---

---

---

---

---

### Results (2)

- Correlations between baskets (A & B)

	1	2	3	4	5
1. Org & Plan	(.38)	same dimensions			
2. Leadership	.38	(.21)			
3. Perceptive	.43	.24	(.43)		
4. Decisions	.54	.11	.42	(.30)	
5. Delegation	.08	-.19	.04	-.12	(.26)

Alternate forms reliability is low. Matters in practice.

---

---

---

---

---

---

---

---

---

---

## Results (3)

- ANOVAs for Training Outcomes

Dimension	MS	F	Posttest Means	
1. Org & plan	3.98	3.19	.78	.24
2. Leadership	1.58	1.59	.91	.58
3. Perceptiveness	7.14	8.60*	.83	.11
4. Decision making	2.01	3.89	.42	.04
5. Delegation	6.33	4.46*	.78	.11

Posttest scores standardized on pretest M, SD. N=58; df = 1, 56. \*  $p < .05$ .

---

---

---

---

---

---

---

---

## Discussion

- Reliability problems. Altho judges agree, low alpha and alt forms. Probs in using scores to make decisions about people.
- Workbook affected scores on 2 of 3 dimensions that should have been affected. Brief training can affect scores 1 month later. Scores at least in part reflect experience rather than ind diffs.

---

---

---

---

---

---

---

---