Ch. 13: Clinical Testing

- Overview of Strategies / Designs
  - Deductive
    - Logical/Content
    - Theoretical
  - Empirical
    - Criterion group
    - Factor analytic
- Frequently used tests

Clinical and Counseling Settings

- Hospital
  - Inpatient
    - acute
    - long-term
  - Outpatient
- School
  - Counseling (educational)
  - Counseling (psychological)
- Vocational
- Motivational/Performance

Types of Tests

- IQ & Achievement
- Personality
  - structured ("objective")
  - unstructured ("projective")
- Clinical
  - Diagnostic
    - Depression, Anxiety, etc.
  - Symptom checklists
- Neuropsychological

Design Theories

- Deductive (aka "top down")
  - Use reason, clinical experience and common sense to choose test items that are face-valid to the construct being assessed.
- Empirical (aka "Data-driven")
  - Look for patterns in large groups of data
  - Let the data tell us what the natural groupings are
  - Don't assume face validity or response style

Logical-Content

- Deductive
  - Logical-Content
    - aka "Content approach", "Intuitive approach", "Rational approach"
    - include face-valid questions about the topic being measured, such as "did you wet the bed last night?"
    - assumes that test-takers answer consistently and honestly
    - simple, and simplistic

Logical-Content Tests

- Woodsworth Personal Data Sheet (1920)
  - Developed in WWI to identify soldiers who would fail in combat
  - 116 Yes/No questions -- all face valid.
    - "Do you drink a fifth of whiskey a day"?
    - "Do you frequently daydream?"
- Mooney Problem Checklist (1950)
  - yes/no to many problem items "I'm having trouble with money"
  - still in use today.

Criticism : Logical-Content

- Major assumptions:
  - Test subjects are being honest, w/o intent to deceive
  - Items have single objective interpretation
    - e.g. "I never drink too much alcohol"
- In most cases, these assumptions are flawed.
- Criticisms basically sunk this design in the late 1940s
- Exception: Mooney Problem checklist (non-confrontational situation? someone is seeking help?

Criterion-Group

- Empirical
  - Criterion Group
    - aka "Contrasted Group", "External strategy"...
    - Give test items to a group that has some disease, disorder or diagnosis (e.g. "Schizophrenics")
    - Compare (contrast) with performance in an "normal" group
    - Items which strongly distinguish the groups are kept.
    - Cross-validate results with other group(s)

Cross-Validation

- Method of giving a more fair, accurate estimate of reliability by essentially "repeating the study" with new subjects
- Use the prediction equations (formulas) on a "fresh" sample
- Statistically, this is a guard against "over fitting" your data in the first experiment.
- The more the cross-validation groups differ in demographics, the more you can claim generalizability.

Cross-Validation Example

- Researcher #1 does a study in their hospital showing that Schizophrenic patients say "Yes" to the question "I'm afraid of the color blue" at a much higher rate than control subjects.
- This test item can diagnosis Schizophrenia?
- Researcher #2 attempts to cross-validate this study at a different hospital with a different set of patients and controls, and does not find the same effect.
- Why?

Criterion-Group Tests

- Minnesota Multiphasic Personality Inventory-2 (MMPI-2)
  - Most popular / well researched test
  - 2nd revision; re-revised norms in 2003
  - 10 clinical scales
  - Focused more on psychopathology
  - good psychometrics
- California Psychological Inventory - 3 (CPI-3)
  - 20 scales
  - focused more on health than illness
  - psychometrics are fair, not great

MMPI Design

- Goal was to detect abnormal psychology
- Used eight abnormal criterion groups in a hospital setting.
- Test items which contrasted criterion group with normal group, AND which passed cross-validation at $P<.05$ level were kept.
- Test items were grouped into scales, the scales were named after the criterion group
- Additional scales added: Mf and Si

MMPI-2

- 567 true/false items
- About 2 hours to administer
- Item content appears to be simplistic face valid statements "I like mechanics magazines" but are not necessarily scored in a face-valid way.
- Each item contributes to one or more Clinical Scales or Content Scales
- Additional Validity Scales attempt to correct for response styles, faking bad or good, etc.
- Scales are normed to T-Scores (mean 50, sd 10)

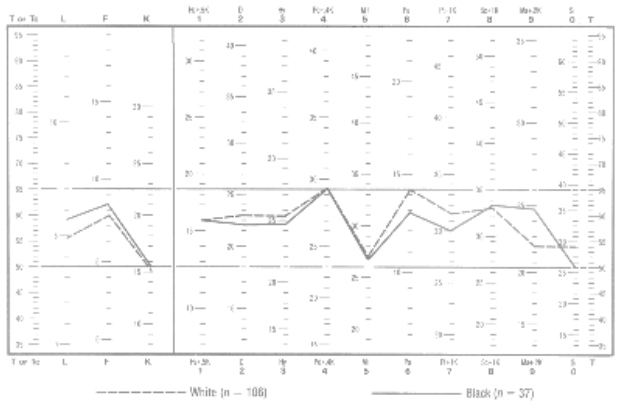
MMPI Clinical Scales

#	Scale	Name	Interpretation
1	Hs	Hypocondriasis	physical complaints
2	D	Depression	depression
3	Hy	Hysteria	immaturity
4	Pd	Psychopathic Deviate	authority conflict
5	Mf	Masculinity-Femininity	stereotypic m/f interests
6	Pa	Paranoia	suspicion, hostility
7	Pt	Psychasthenia	anxiety
8	Sc	Schizophrenia	alienation, withdrawal
9	Hy	Hypomania	elevated mood & energy
0	Si	Social Introversion	introversion, shyness

MMPI Validity Scales

#	Scale	Name	Interpretation
1	L	Lie scale	naive attempt to fake good
2	F	F scale	attempt to fake bad
3	K	K scale	defensiveness
4	VRIN	Variable response inconsistency	random responding
5	TRIN	True response inconsistency	"yes" bias

MMPI Scales



MMPI Extensions

- Because the test collects so much data (567 items) it's possible to score the items in many different ways.
- Many authors have introduced novel scoring methods for the MMPI
- Some of these are actually just face-valid content measures.
- Others are targeted towards specific clinical groups or disorders (alcoholism, etc.)

MMPI Validity

- Thousands of studies have been done on the MMPI supporting its Construct validity
- Used and researched in a very wide range of subjects, settings, disorders.

MMPI Psychometrics

- Overall reliability is good. Not as high as the best IQ tests.
- Median test-retest coefficients range from .50 to .90, average .80s
- Scales are poorly designed -- many items contribute score to more than one scale. Thus the scales are highly intercorrelated.
- Keyed poorly (e.g. all items on the L scale are keyed false) so sensitive to response style
- Scores are affected by demographics (age, gender, IQ, and to a lesser extent, ethnicity)

Criticism : Criterion-Group

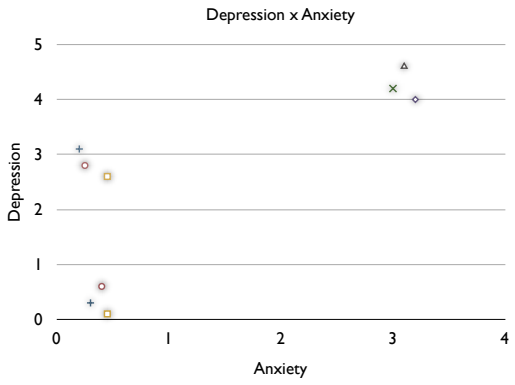
- Major assumptions:
  - People/Patients can be grouped, groups show consistent features and are independent
  - Face validity of response is not important, but responses should be consistent
  - Scales can predict membership in criterion group
- Flaws:
  - Criterion group assumes a lot about patterns of behavior. Neglects possible commonalities across groups.

Factor-Analysis

- FA of the MMPI-2 suggests there may be only 2 major factors! ("positive affectivity" and "negative affectivity")

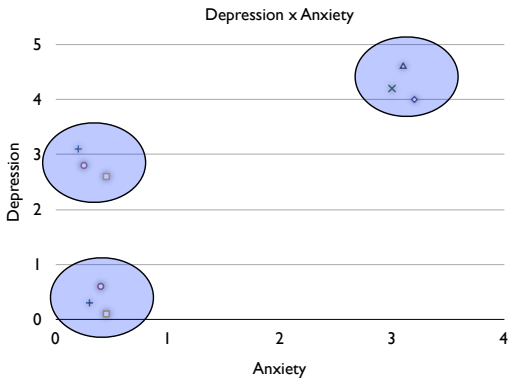
Factor Analysis

Multivariate data often appears to form natural groups



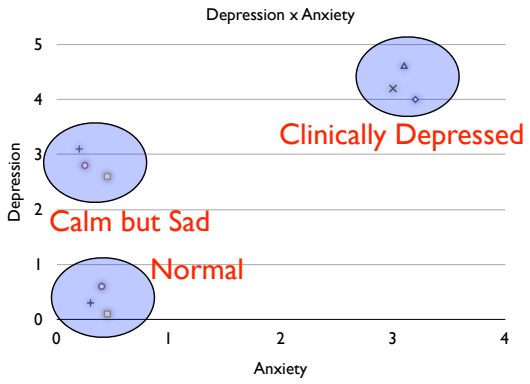
Factor Analysis

"Natural" groupings in data are determined statistically



Factor Analysis

Groups are then named using logical/ content analysis



Factor-Analytic

- Empirical
  - Factor-Analytic
    - Rather than assuming face-validity (logical-content strategy) or identifiable groups (criterion-group strategy) we look for natural groups and patterns in the data
    - Assume that groups seen in the data are not random or accidental, but causal
    - Examine the groups for commonality, reduce extraneous variables
    - Examine the Content, name the factors

Factor-Analytic : 16PF

- Cattell's 16 Personality Factor Questionnaire (16PF)
  - Started with 4504 traits from the dictionary, narrowed it down to 171 traits.
  - Gave these 171 test items to college students
  - Factor Analysis reduced these to 16 distinct factors (with 4 uber-factors)
  - Extensive norms for age, gender, reading level, etc.
  - Similar test available for adolescents and children

16PF Factors

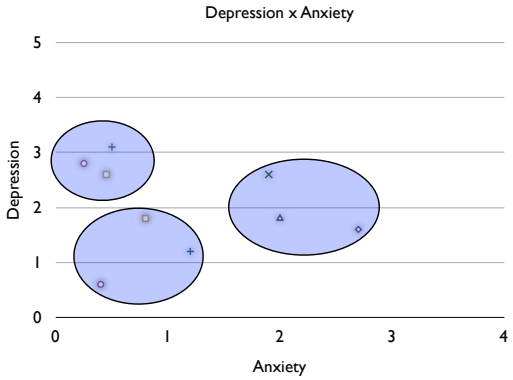
Scale	Low	High
A	cool, reserved	warm, outgoing
B	concrete, dull	abstract, bright
C	affected by feelings	emotionally stable
E	submissive, humble	dominant, assertive
F	sober, restrained	enthusiastic, spontaneous
G	expedient, indulgent	conscientious, conforming
H	shy, timid	bold, venturesome
I	tough-minded	tender-minded, sensitive
L	trusting, easy going	suspicious, skeptical
M	practical, down to earth	imaginative, absent-minded
N	forthright, open	shrewd, calculating
etc…	etc…	

Criticism : Factor-Analytic

- Major assumptions:
  - Data groupings (factors) found are stable
  - Factors can be examined for content and named
- Flaws:
  - The factor analysis process is arbitrary (more than one factor solution possible, no clear way to decide proper # of factors) : 16 or 4?
  - Factors are just collections of data -- naming them requires a face-valid examination of data, which is problematic.

Factor Analysis

of groups, and group boundaries can be arbitrary



Factor Analysis

3 groups? Or only two?