

Assessment Methods and Strategies

Types of Reliability

Type	What it Is	How You Do It	What It Looks Like
Test-Retest	Measure of stability	Give same test twice to same people	$r_{\text{time 1 with time 2}}$
Parallel Forms	A measure of equivalence	Give 2 different forms to the same people	$r_{\text{form 1 with form 2}}$
Inter-Item	Measure of each item consistency	Statistical analysis of each item with scale	
Inter-Rater	Measure of agreement	Have 2 people rate behaviors; evaluate agreement	$\frac{\# \text{ of agreements}}{\# \text{ of total observations}}$

Types of Validity

Face validity

Measure appears to assess a construct

Content validity

Content reflects domain of interest

Criterion validity

Concurrent validity - Correlation with another measure

Predictive validity - Correlation with future performance

Types of Validity

Convergent validity

Measures assess similar constructs

Discriminant validity

Measures assess different constructs

Convergent and Discriminant Validity of a Depression Measure

Jay and John (2004)

Table 1

Content Validity: The 33 CPI-D Items Classified Into the DSM-IV Depression Symptom Categories in Study 1

Symptom category	No.	CPI-D item
Sad or empty mood	245	Feels happy most of the time (R)
	416	Not quite as happy as others
	419	Life often has no meaning
	353	Not understood by others
	133	Feels as good as ever (R)
	398	Handed a raw deal in life
Diminished interest or pleasure in things	021	Life is full of interesting things (R)
	156	Hardly ever excited or thrilled
	280	Enjoys different kinds of play (R)
	124	Not likely to speak to others first
	070	Crosses the street to avoid meeting others
	161	Sometimes cross without good reason
Feelings of worthlessness or guilt	339	Thinks is no good at all
	279	Often gets disgusted with self
	147	Feels useless at times
	390	Has not lived the right kind of life
	311	Cannot do anything well
	257	Feels has done something wrong or wicked
	050	As capable and smart as most others (R)
	369	Has more regrets than others do
Suicidality or hopelessness about life	365	Future seems hopeless
	259	Life is worthwhile (R)
	015	Feels as if something bad is about to happen
	299	Does not care what happens
Diminished concentration	054	Hard to keep mind on task or job
	456	Has trouble concentrating
	099	Cannot keep mind on one thing
	013	Slow in making up mind
Fatigue	238	Sometimes can't get going
	426	Gets tired easily
Changes in appetite	400	Has a good appetite (R)
Changes in sleeping	459	Sleep is fitful and disturbed
	135	Wakes up fresh and rested (R)

Note. California Psychological Inventory (CPI) items are abbreviated and paraphrased and are included to illustrate the item content of the CPI-D. The number to the left of each item is the item number on the Form 462 version of the CPI (Gough, 1987). Item numbers on the Form 434 version of the CPI (Gough & Bradley, 1996) are the same except for Items 135 and 419, which are not included; Item 456, which is numbered Item 362; and Item 459, which is numbered Item 402. CPI-D = California Psychological Inventory Depressive Symptom scale; *DSM-IV* = *Diagnostic and Statistical Manual of Mental Disorders—Fourth Edition*; R = reverse keyed.

Table 4

Alpha Reliability (on the Diagonal), Convergent and Discriminant Validity Correlations, and Descriptive Statistics in the Three Samples in Study 2

Sample and measure	CPI-D	BDI	CES-D	<i>M</i>	<i>SD</i>	Skewness	Kurtosis
Sample A							
CPI-D	(.88)			8.5	6.2	0.91	0.44
BDI	.78	(.87)		6.1	6.2	1.60	3.00
CES-D	.69	.69	(.88)	14.6	9.4	0.92	0.63
Sample B							
CPI-D	(.88)			8.8	6.3	0.76	-0.10
CES-D	.68		(.88)	12.8	8.8	0.90	0.75
Sample C							
CPI-D	(.90)			8.4	5.7	0.65	-0.15
BDI	.81	(.90)		8.5	7.9	1.30	1.40
CES-D	.79	.76	(.92)	17.1	10.5	0.79	0.09
Mean convergent r^a	.76	.76	.73				
STAI discriminant r^b	.55		.64	16.6	10.3	0.52	-0.18

Note. All correlations were significant at $p < .01$. Alpha reliabilities are in parentheses on the diagonal. CPI-D = California Psychological Inventory Depressive Symptom scale; BDI = Beck Depression Inventory; CES-D = Center for Epidemiological Studies Depression Scale; STAI = State-Trait Anxiety Inventory.

^a Mean of the convergent correlations across Samples A, B, and C using Fisher's r -to- z transformations.

^b STAI correlations are discriminant correlations of CPI-D or CES-D with STAI Anxiety in Sample B.

Construct Validity

- Measures should assess relevant aspects of the construct of interest
 - A measure's name does not necessarily reflect the construct that it measures
- Existing measures
 - Extent to which the measure reflects the construct
 - Popularity of a measure does not ensure its construct validity
- More than one measure of the construct should be included

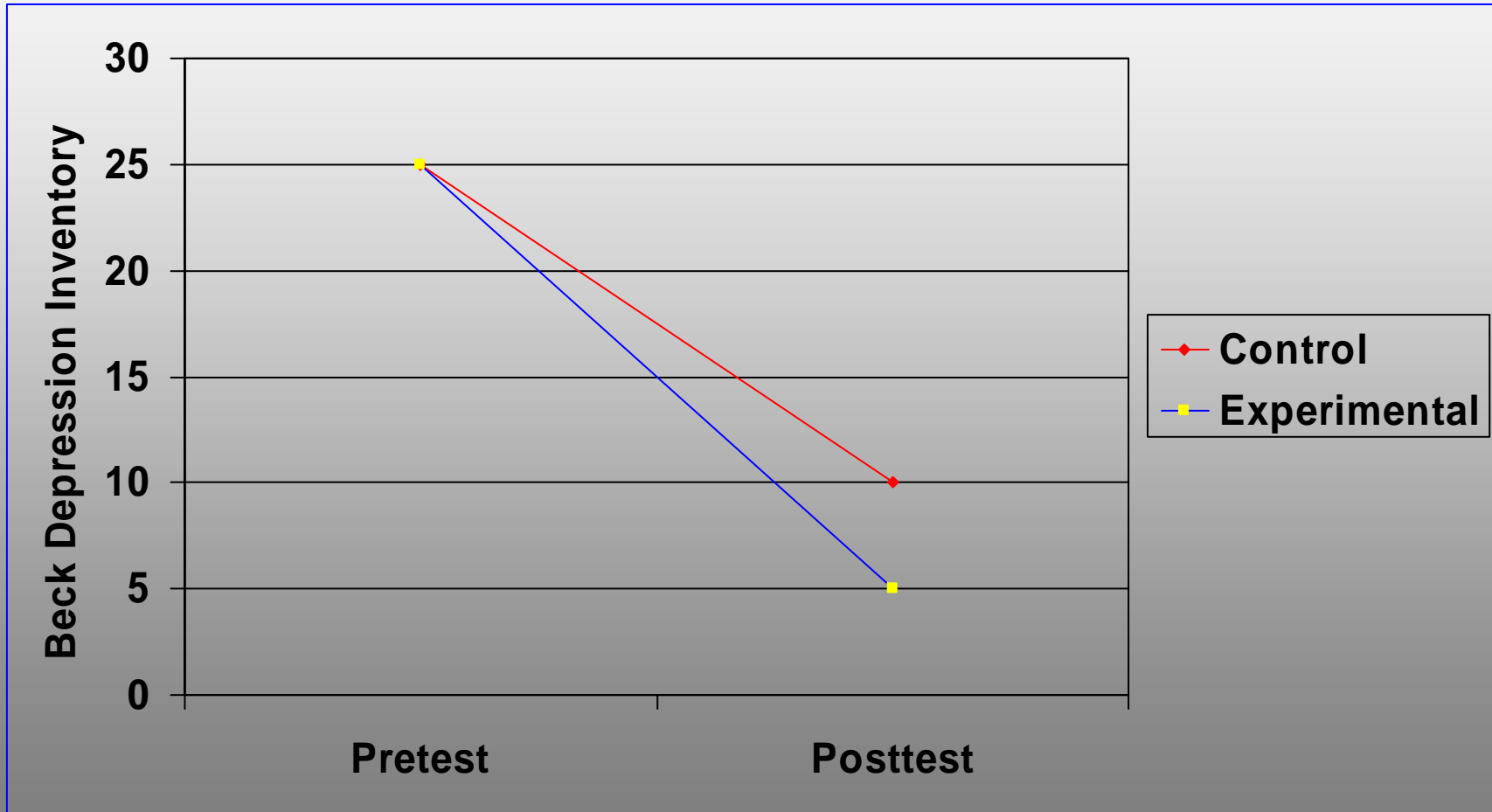
Reliability and Validity in a Particular Research Context

- If you are using an existing measure in a context in which it was not validated, you need to determine its psychometric properties in this context
 - Internal consistency
 - Criterion validity
- If you modify a measure (e.g., add or delete items), it becomes a new measure

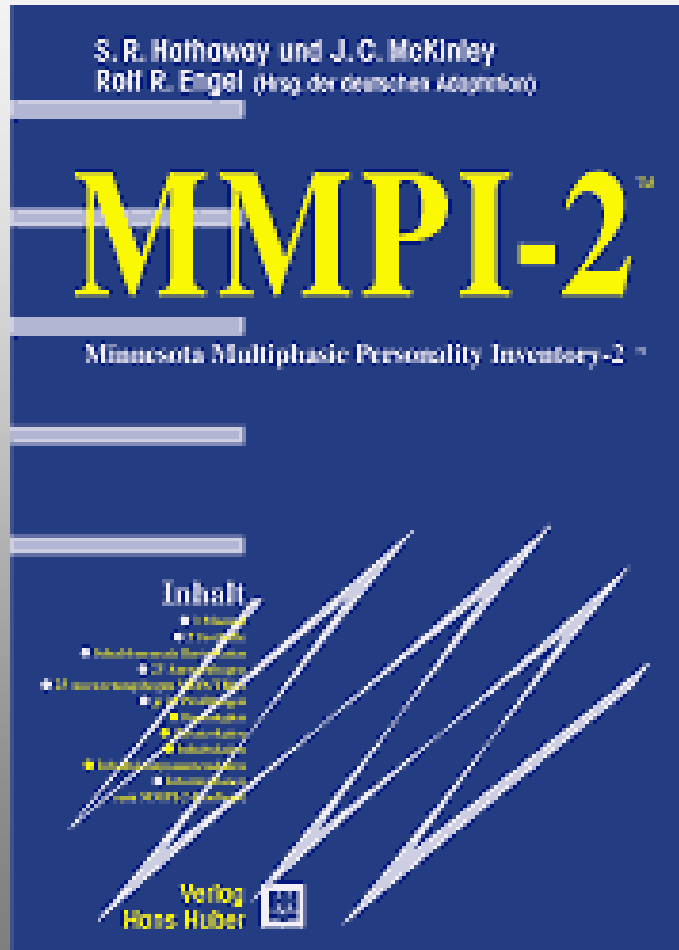
Measurement Sensitivity

- Can the measure reflect change as a result of an experimental manipulation or group differences?
 - Can a test be too reliable?
- Large range of possible responses necessary to detect change
 - 5-item test vs. 50-item test
- If participants' responses are skewed, change can occur primarily in one direction

Skewed Responses



Test Development (Haynes, Nelson, & Blaine, 1999)



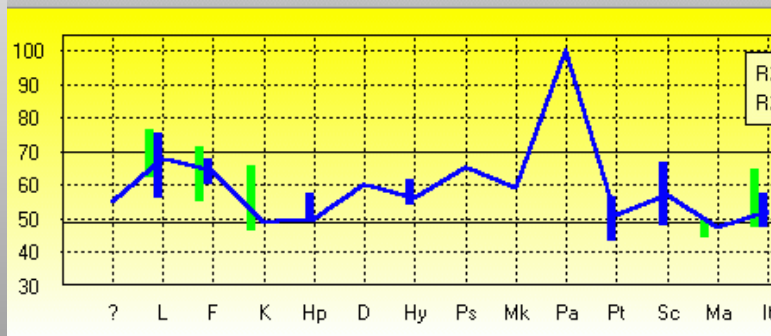
- Specify construct
- Specify contexts for use of test
- Specify the intended functions of the test

Test development



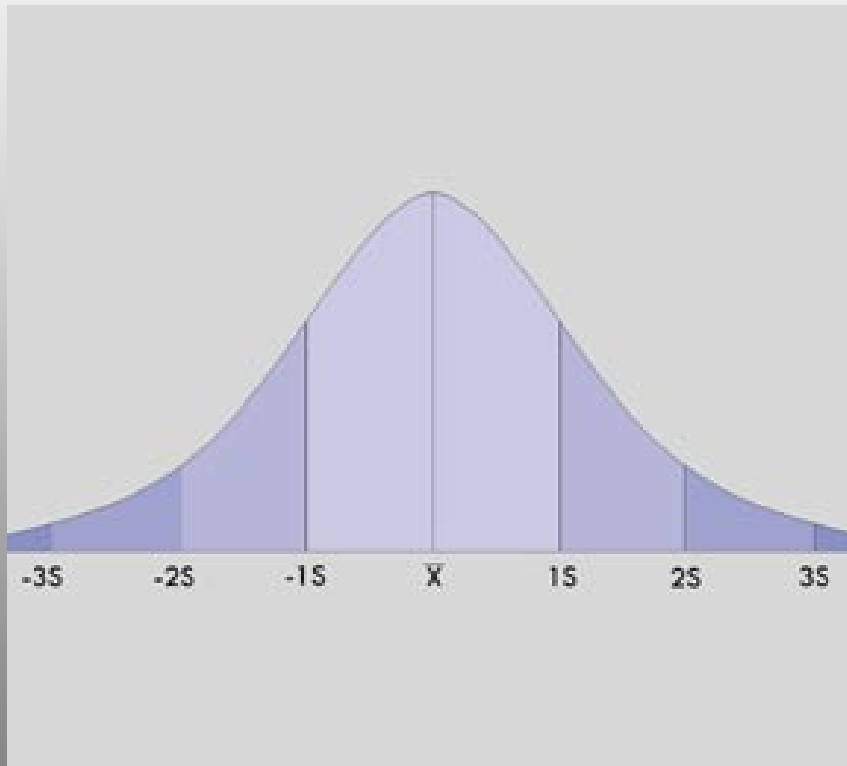
- Generate and select items based on:
 - Rational deduction
 - Clinical experience
 - Theories
 - Empirical literature
 - Existing tests
 - Suggestions by experts
 - Population sampling

Test development



- Match items to facets and dimensions
- Establish quantitative parameters of the test (e.g., response formats, scales)
- Develop instructions
- Expert review
- Population review
- Rereview

Test development



- Pilot test
- Develop representative norms
- Evaluate validity

Assessment of Hypochondriasis

Longley et al. (2005)



What Questions Should Be Included on a Measure of Hypochondriasis?



Table 1

Development of the Multidimensional Inventory of Hypochondriacal Traits (MIHT): Factors and Item Loadings

Factor and item	Item loadings on factors			
	I	II	III	IV
I. Cognitive component: Hypochondriacal alienation				
1. Others do not seem sympathetic to my health problems.	.75	.05	.11	.12
2. I wish others took my health complaints more seriously.	.72	.12	.15	.19
3. I get upset about the way others respond to my illness.	.70	.12	.20	.14
4. Sometimes others do not seem very concerned about my health complaints.	.69	.20	.10	.28
5. The more I talk about my health problems the less others seem to listen.	.64	.06	.13	.12
6. Few people seem to take my health concerns as seriously as I do.	.61	.18	.09	.27
7. People seem unconvinced my symptoms are signs of illness.	.55	.14	.14	.11
II. Behavioral component: Hypochondriacal reassurance				
8. I turn to others for support when I do not feel well.	.01	.72	.17	.11
9. I like to be reassured when I feel sick.	.17	.71	.11	.06
10. If my symptoms worry me, I appreciate sympathy from others.	.18	.69	.08	.09
11. When I am hurt or ill, I like to have someone help me.	.05	.66	.17	.20
12. When I feel physical pain, I let others know.	.05	.56	.18	.17
13. It is important that others care about my health complaints.	.22	.54	.08	.30
14. When I was ill as a child, I liked to have my parents fuss over me.	.20	.51	.13	.11
15. Telling people about my health problems makes me feel better.	.27	.48	.15	.22
III. Perceptual component: Hypochondriacal absorption				
16. I am aware of my body position.	.15	.11	.75	.06
17. I am usually aware of how I feel physically.	.07	.22	.64	.07
18. I am aware of physical sensation.	.20	.13	.61	.05
19. Even when I listen to a lecture or talk, I am alert to how my body feels.	.21	-.02	.59	.17
20. I notice how clothes feel against my body.	.23	.03	.59	.14
21. When lying in bed at night, I am often aware of my body.	.12	.02	.57	.13
22. Generally, I am sensitive to changes in my body.	.03	.12	.51	.19
23. I keep close track of what is happening to me physically.	-.07	.14	.51	.33
24. I am aware of how my body feels after a big meal.	.04	.29	.46	.07
IV. Affective component: Hypochondriacal worry				
25. I worry a lot about my health.	.23	.15	.20	.75
26. When I experience pain, I fear I may be ill.	.20	.33	.17	.58
27. Reading articles about disease makes me worry about my health.	.15	.11	.21	.57
28. If I notice a skin blemish I worry it might lead to something serious.	.12	.09	.07	.53
29. I am concerned with the possibility of being diagnosed with a serious disease.	.26	.27	.15	.52
30. I worry about the physical problems of getting older.	.16	.13	.29	.42
31. I try to avoid things that make me think of illness or death.	.25	.21	.07	.41

Note. $N = 312$, Student 4 sample. Numbers in boldface are item loadings of .40 or greater, on that factor.

Table 2

Descriptive Statistics Across Three Samples for the Multidimensional Inventory of Hypochondriacal Traits (MIHT) and the Big Five Inventory (BFI) Neuroticism Scales

Scale	Student samples ^a (<i>N</i> = 754)			Couples sample (<i>N</i> = 172)			Medical patients sample (<i>N</i> = 120)		
	<i>M</i>	<i>SD</i>	α	<i>M</i>	<i>SD</i>	α	<i>M</i>	<i>SD</i>	α
MIHT scales									
Cognitive	15.12 _a	5.46	.89	13.12 _b	4.65	.88	15.09 _a	5.80	.87
Behavioral	25.51 _a	5.52	.82	22.92 _b	5.74	.83	22.76 _b	6.69	.86
Perceptual	27.60 _a	6.20	.83	26.83 _a	5.76	.81	31.49 _b	6.29	.86
Affective	17.32 _a	5.25	.80	14.84 _b	4.49	.82	17.45 _a	5.27	.80
BFI scale									
Neuroticism	23.91 _a	6.25	.86	21.51 _b	6.47	.85	24.20 _a	6.23	.82

Note. Means in the same row that do not share subscripts differ at $p < .05$ (two-tailed).

^a Student samples are the combined Student 4 (*N* = 312) and Student 5 (*N* = 442) samples.

Table 3

Correlations of the Multidimensional Inventory of Hypochondriacal Traits (MIHT) Scales With the Whitely Index (WI), Somatosensory Amplification Scale (SSAS), and the Illness Attitude Scales (IAS)/Health Anxiety Questionnaire (HAQ)

Scale	MIHT scales			
	Cognitive	Behavioral	Perceptual	Affective
WI				
Cognitive	.60	.27	.28	.55
Affective	.50	.43	.49	.72
SSAS				
Perceptual	.41	.34	.53	.42
IAS/HAQ				
Cognitive	.51	.32	.32	.57
Behavioral	-.12	.07	.12	.07
Perceptual	.40	.28	.43	.42
Affective	.42	.41	.30	.73

Note. $N = 312$, Student 4 sample. Boldface type indicates convergent correlations between corresponding scales.

Table 6*Correlations of the Multidimensional Inventory of Hypochondriacal Traits (MIHT) and the Big Five Inventory (BFI)*

Scale	MIHT scale			
	Cognitive	Behavioral	Perceptual	Affective
BFI				
Neuroticism	.37	.27	.20	.31
Extroversion	-.04	.16	.10	.08
Openness to Experience	-.07	-.15	-.04	-.16
Conscientiousness	-.21	.16	-.10	-.05
Agreeableness	-.27	.07	-.06	-.08

Note. $N = 312$, Student 4 sample. The correlations between the MIHT and the BFI Neuroticism scale are shown in boldface type.

Table 7

Correlations of the Four Multidimensional Inventory of Hypochondriacal Traits (MIHT) Scales and the Big Five Inventory (BFI) Neuroticism Scale With Retrospective Health Care Utilization Indexes and Cross-Sectional Self-Reported Somatic Symptoms

Measure	MIHT scale				BFI scale
	Cognitive	Behavioral	Perceptual	Affective	Neuroticism
Health care utilization indexes^a					
No. of clinics visited	.29	.12	.06	-.07	.19
No. of appointments	.24	.10	.05	.06	.16
No. of diagnoses	.08	-.03	-.04	-.06	.09
No. of medical tests	-.02	-.03	.20	-.07	.03
No. of medications prescribed	.14	.03	.06	-.07	.22
No. of medical personnel	.07	.03	-.01	-.07	.18
Somatic Symptom Inventory					
Self-reported somatic symptoms	.25	.18	.08	-.09	.48

Note. $N = 120$, Patient sample. Moderate correlations (i.e., $\geq |.20|$) are shown indicated in boldface type.

^a All health care utilization indexes were based on medical chart data from the previous 12 months.

Methods of Assessment

- Global ratings
- Self-report inventories
- Projective techniques
- Behavioral observations
- Psychobiological measures

Global Ratings: DSM Global Assessment of Functioning Scale

- 91-100: Superior functioning in a wide range of activities, life's problems never seem to get out of hand, is sought by others because of his or her many positive qualities. No symptoms.

Global Ratings: DSM Global Assessment of Functioning Scale

- 81-90: Absent or minimal symptoms (e.g., mild anxiety before an exam); good functioning in all areas, interested and involved in a wide range of activities, socially effective, generally satisfied with life, no more than everyday problems or concerns (e.g., an occasional argument with family members).

Global Ratings: DSM Global Assessment of Functioning Scale

- 71-80: If symptoms are present, they are transient and expectable reactions to psychosocial stressors (e.g., difficulty concentrating after family argument); no more than slight impairment in social, occupational, or school functioning (e.g., temporarily falling behind in schoolwork).

Global Ratings: DSM Global Assessment of Functioning Scale

- 61-70: Some mild symptoms (e.g., depressed mood and mild insomnia) OR some difficulty in social, occupational, or school functioning (e.g., occasional truancy, or theft within the household), but generally functioning pretty well, has some meaningful interpersonal relationships.

Global Ratings: DSM Global Assessment of Functioning Scale

- 51-60: Moderate symptoms (e.g., flat affect and circumstantial speech, occasional panic attacks) OR moderate difficulty in social, occupational, or school functioning (e.g., few friends, conflicts with peers or co-workers).

Global Ratings: DSM Global Assessment of Functioning Scale

- 41-50: Serious symptoms (e.g., suicidal ideation, severe obsessional rituals, frequent shoplifting) OR any serious impairment in social, occupational, or school functioning (e.g., no friends, unable to keep a job).

Global Ratings: DSM Global Assessment of Functioning Scale

- 31-40: Some impairment in reality testing or communication (e.g., speech is at times illogical, obscure, or irrelevant) OR major impairment in several areas, such as work or school, family relations, judgment, thinking, or mood (e.g., depressed man avoids friends, neglects family, and is unable to work; child frequently beats up younger children, is defiant at home, and is failing at school).

Global Ratings: DSM Global Assessment of Functioning Scale

- 21-30: Behavior is considerably influenced by delusions or hallucinations OR serious impairment in communication or judgment (e.g., sometimes incoherent, acts grossly inappropriately, suicidal preoccupation) OR inability to function in almost all areas (e.g., stays in bed all day; no job, home, or friends).

Global Ratings: DSM Global Assessment of Functioning Scale

- 11-20: Some danger of hurting self or others (e.g., suicide attempts without clear expectation of death, frequently violent, manic excitement) OR occasionally fails to maintain minimal personal hygiene (e.g., smears feces) OR gross impairment in communication (e.g., largely incoherent or mute).

Global Ratings: DSM Global Assessment of Functioning Scale

- 1-10: Persistent danger of severely hurting self or others (e.g., recurrent violence) OR persistent inability to maintain personal hygiene OR serious suicidal act with clear expectation of death.

Global Ratings: DSM Global Assessment of Functioning Scale

- Current rating
- Highest during past year

Case 1: Bad Voices

- This woman was referred to the psychiatric emergency department by a psychologist who was treating her in an anxiety disorders clinic. After telling her psychologist that she heard voices telling her to kill herself and then assuring him that she would not act on the voices, she skipped her next appointment. Her psychologist called her to say that if she did not voluntarily come to the emergency department for an evaluation, he would send the police for her.
- Interviewed in the emergency department, she was at times angry and insistent that she did not like to talk about her problems and that the psychologist would not believe her or help her anyway. This attitude alternated with flirtatious and seductive behavior.
- She first saw a psychologist 3 years ago. At that time, she began to hear a voice telling her that she was a bad person and that she should kill herself. She would not say exactly what it told her to do, but she reportedly drank nail polish remover in a suicide attempt. At that time, she remained in the emergency department for 2 days and received an unknown medication that reportedly helped quiet the voices. She did not return for an outpatient appointment after discharge and continued having intermittent periods of auditory hallucinations over the next 3 years with some periods lasting for months at a time. For example, often when she was near a window, a voice would tell her to jump out, and when she walked near traffic, it would tell her to walk in front of a car.

Case 1: Bad Voices

- She reports that she continued to function well after that first episode, finishing high school. About 2 months ago, she began to have trouble sleeping and felt “nervous.” It was at this time that she responded to an ad for the anxiety clinic. She was evaluated and given Haldol, an antipsychotic. She claims that there was no change in the voices at that time, and only the insomnia and anxiety were new. She specifically denied depressed mood or anhedonia or any change in her appetite but did report that she was often tearful and lonely, and sometimes ruminated about “bad things,” such as an attempted rape of her at age 14. Despite these symptoms, she continued working part-time as a salesperson in a department store.
- She says she did not keep her follow-up appointment at the anxiety clinic because the Haldol was making her stiff and nauseous and was not helping her symptoms. She denies wanting to kill herself, and cited how hard she was working to raise her children as evidence that she would not “leave them that way.” She did not understand why her behavior had alarmed the psychologist.
- She denied alcohol or drug use, and a toxicology screen for various drugs was negative. Physical examination and routine laboratory tests were also normal. She had stopped the Haldol on her own 2 days before the interview.

Case 1: Bad Voices

- Following the interview, there was disagreement among the staff about whether to let the client leave. It was finally decided to keep her overnight, until her mother could be seen the following day. When told she was to stay in the emergency department, she replied angrily, yet somewhat coyly: “Go ahead. You’ll have to let me out sooner or later, but I don’t have to talk to you if I don’t want to.”
- When the mother was interviewed the following morning, she said she did not see a recent change in her daughter. She did not feel that she would hurt herself but agreed to stay with her for a few days and make sure she went for follow-up appointments. In the family meeting, she complained that her mother was unresponsive and did not help her enough. However, she again denied depression and said she enjoyed her job and her children. About the voices, she said that over time she had learned how to ignore them and that they did not bother her as much as they had at first. She agreed to outpatient treatment provided the therapist was a woman.

Case 2: Under Surveillance

- This man was brought into the emergency department by the police for striking an elderly woman in his apartment building. His chief complaint is, “That damn bitch. She and the rest of them deserved more than that for what they put me through.
- He has been continuously ill since the age of 17. During his first year of college, he gradually became more and more convinced that his classmates were making fun of him. He noticed that they would snort and sneeze whenever he entered the class room. When a woman he was dating broke off the relationship with him, he believed that she had been “replaced” by a look-alike. He called the police and asked for their help to solve the “kidnapping.” His academic performance in school declined dramatically, and he was asked to leave and seek psychiatric care.
- He got a job as a bank teller, which he held for 7 months. However, he was getting an increasing number of distracting “signals” from co-workers, and he became more and more suspicious and withdrawn. It was at this time that he first reported hearing voices. He was eventually fired, and soon thereafter was hospitalized for the first time. He has not worked since.

Case 2: Under Surveillance

- He has been hospitalized 3 times, the longest stay being 3 months. However, in the past five years, he has been hospitalized only once for 3 weeks. During the hospitalizations he has received various antipsychotic drugs. Although outpatient medication has been prescribed, he usually stops taking it shortly after leaving the hospital. Aside from twice-yearly lunch meeting and his contacts with mental health workers, he is totally isolated socially. He lives on his own and manages his own financial affairs, including a modest inheritance. He reads the *Wall Street Journal* daily. He cooks and cleans for himself.

Case 2: Under Surveillance

- He maintains that his apartment is the center of a large communication system that involves all three major television networks, his neighbors, and apparently hundreds of “actors” in his neighborhood. There are secret cameras in his apartment that carefully monitor all his activities. When he is watching television, many of his minor actions (e.g., going to the bathroom) are soon directly commented on by the announcer. Whenever he goes outside, the “actors” have all been warned to keep him under surveillance. Everyone on the street watches him. His neighbors operate two different “machines”; one is responsible for all of his voices, except the “joker.” He is not certain who controls this voice, which “visits” him only occasionally and is very funny. The other voices, which he hears many times each day, are generated by this machine, which he sometimes thinks is directly rung by the neighbor whom he attacked. For example, when he is going over his investments, these “harassing” voices constantly tell him which stocks to buy. The other machine he calls “the dream machine.” This machine puts erotic dreams into his head.

Case 2: Under Surveillance

- He describes other unusual experiences. For example he recently went to a shoe store 30 miles from his house in the hope of getting some shoes that wouldn't be "altered." However, he soon found out that, like the rest of the shoes he buys, special nails had been put into the bottom of the shoes to annoy him. He was amazed that his decision concerning which shoe store to go to must have been known to his "harassers" before he himself knew it, so that they had time to get the altered shoes made up especially for him. He realizes that great effort and "millions of dollars" are involved in keeping him under surveillance. He sometimes thinks this is all part of a large experiment to discover the secret of his "superior intelligence."
- At the interview, he is well groomed, and his speech is coherent and goal-directed. His affect is, at most, only mildly blunted. He was initially angry at being brought in by the police. After several weeks of treatment with an antipsychotic drug failed to control his psychotic symptoms, he was transferred to a long-stay facility with the plan to arrange a structured living situation for him.

Limitations of Global Ratings

- Reliability of single items
- Construct validity – do the ratings measure what they are supposed to measure?
- Sensitivity

Self Report Inventories: Beck Depression Inventory

- 0 I do not feel sad.
- 1 I feel sad.
- 2 I am sad all the time and can't snap out of it.
- 3 I am so sad or unhappy that I can't stand it.

Self Report Inventories: Beck Depression Inventory

- 0 I don't feel I am worse than anybody else.
- 1 I am critical of myself for my weaknesses or mistakes.
- 2 I blame myself all the time for my faults.
- 3 I blame myself for everything bad that happens.

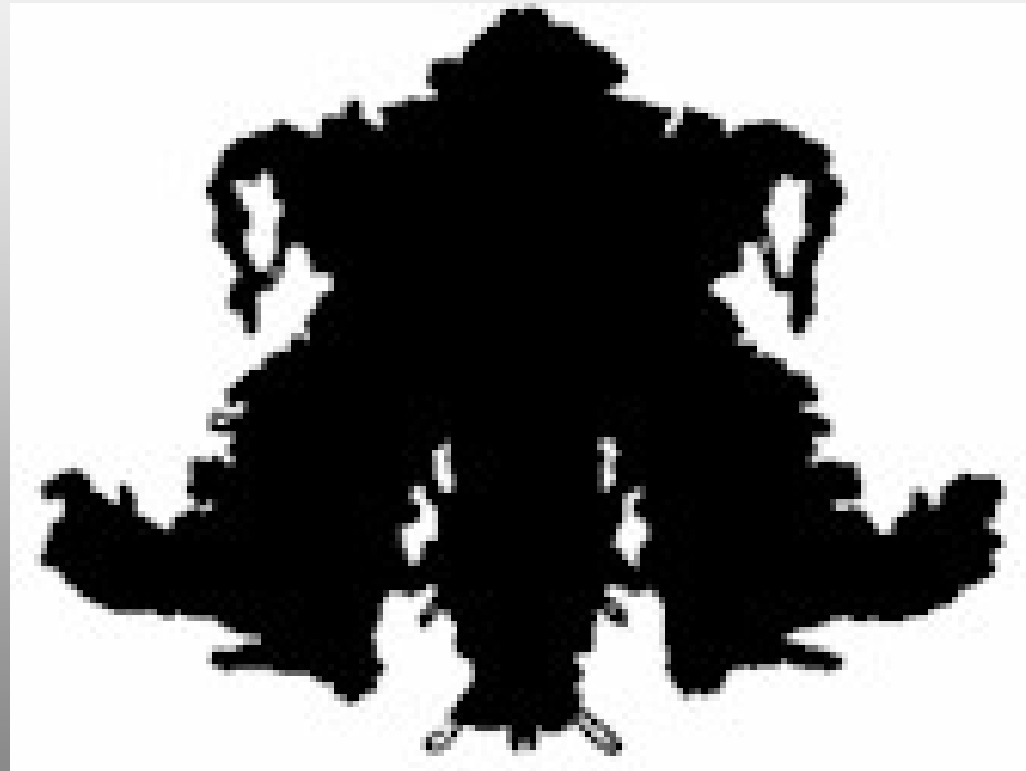
Self Report Inventories: Beck Depression Inventory

- 0 I don't have any thoughts of killing myself.
- 1 I have thoughts of killing myself, but I would not carry them out.
- 2 I would like to kill myself.
- 3 I would kill myself if I had the chance.

Limitations of Self-Report Inventories

- Response bias

Projective Techniques: Rorschach Inkblot Method



Projective Techniques: Rorschach Inkblot Method



Limitations of Projective Techniques

- Reliability of scoring
- Validity of interpretations
- “We believe there is no scientific basis to support the continued widespread use of the Rorschach in clinical, legal, forensic, and occupational settings” (Hunsley & Bailey, 2001)

Behavioral Observations: Peer Interaction Task (Dishion et al.)

- Plan an activity together (e.g., party)
- Solve 4 problems that have occurred within the past month:
 - A problem for the study boy/girl related to getting along with parents
 - A problem for the study boy/girl related to getting along with peers
 - A problem for the friend related to getting along with parents
 - A problems for the friend related to getting along with peers

Behavioral Observations: Peer Interaction Task (Dishion et al.)

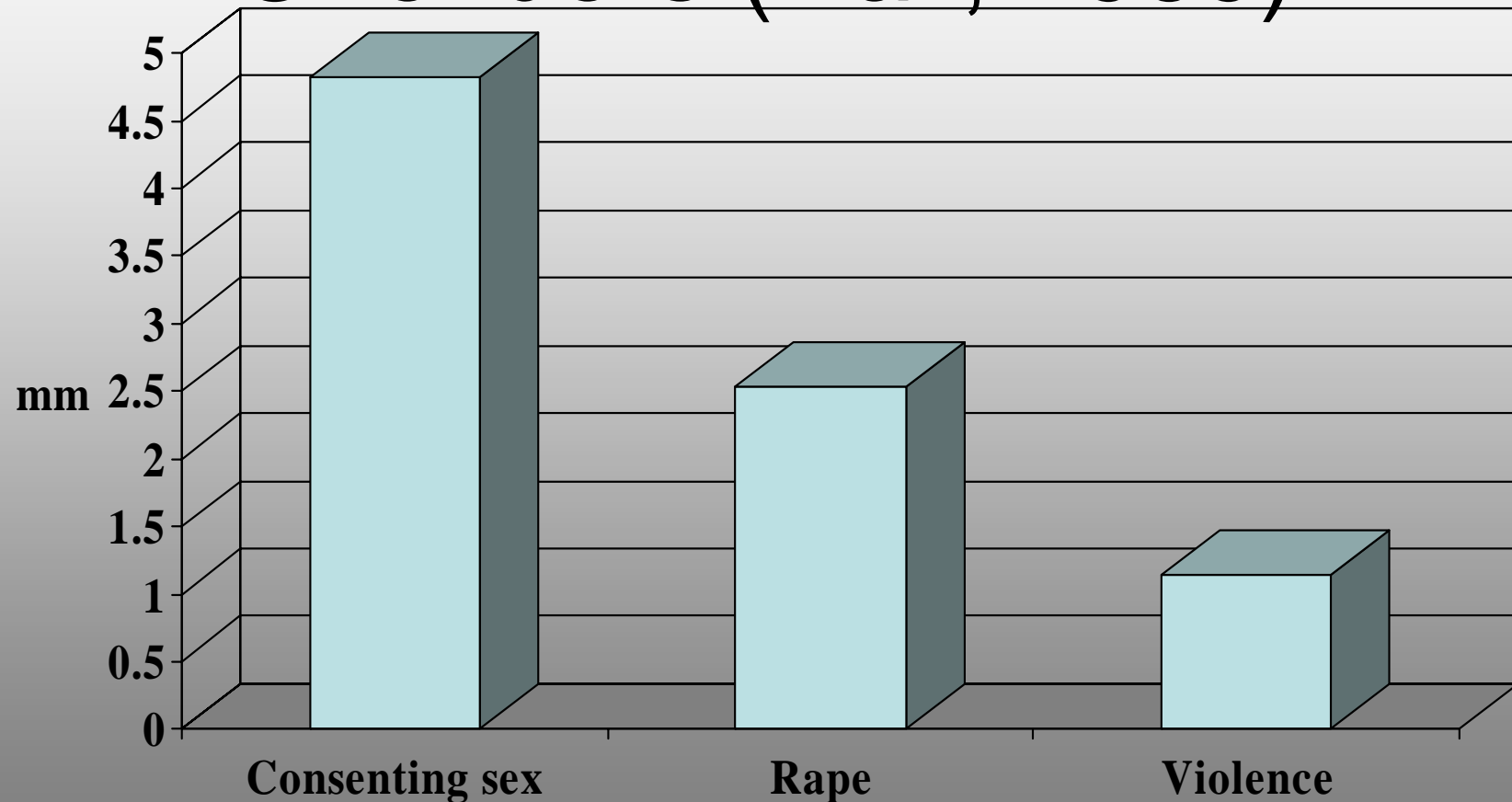
- Coding
 - Rule-breaking (verbal behavior and gestures that contain some violation of legal and/or conventional norms of conduct)
 - Normative (anything not coded as rule-breaking)

Limitations of Observational Methods

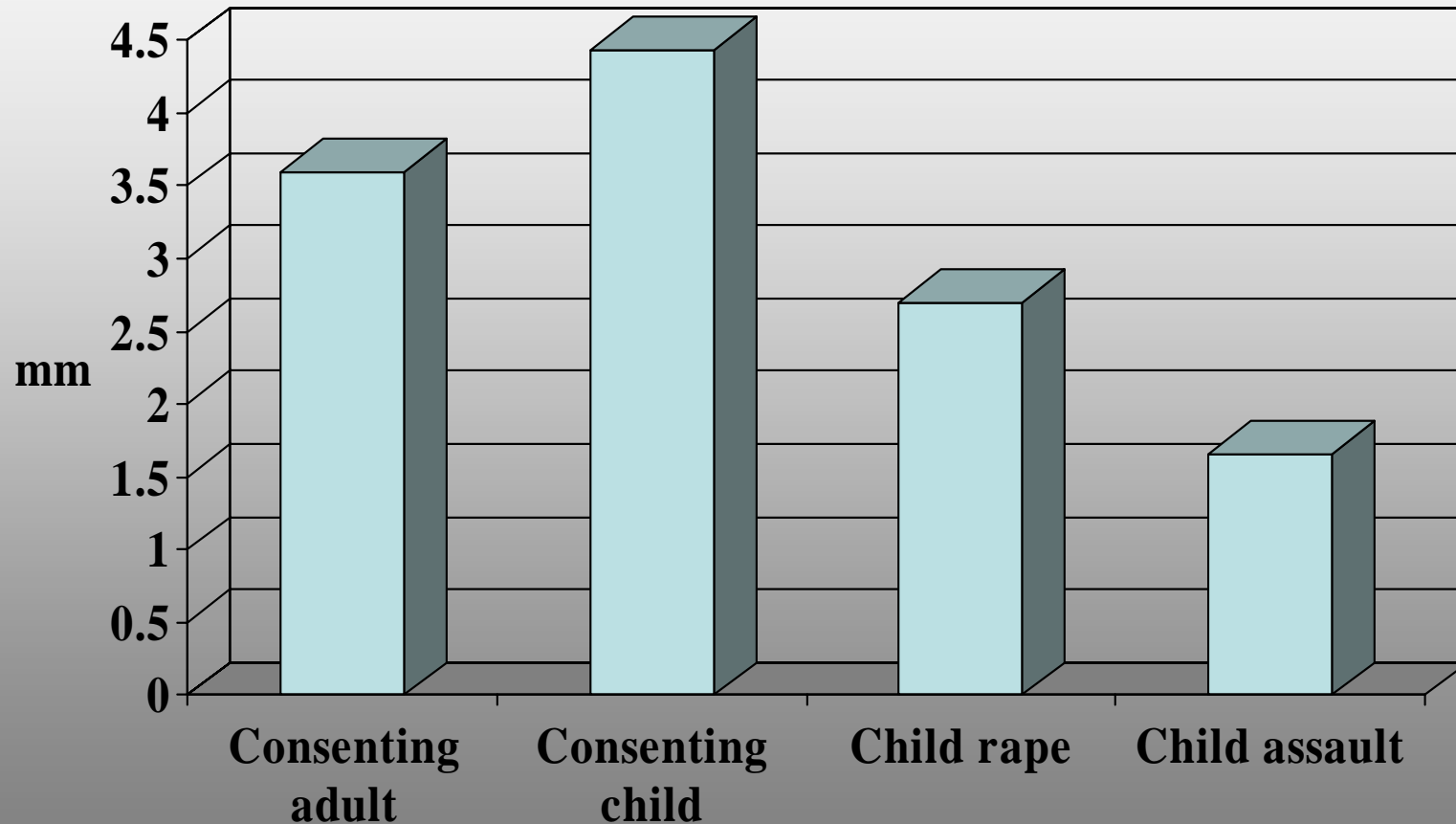
- Sample of behavior
- Contrived situation



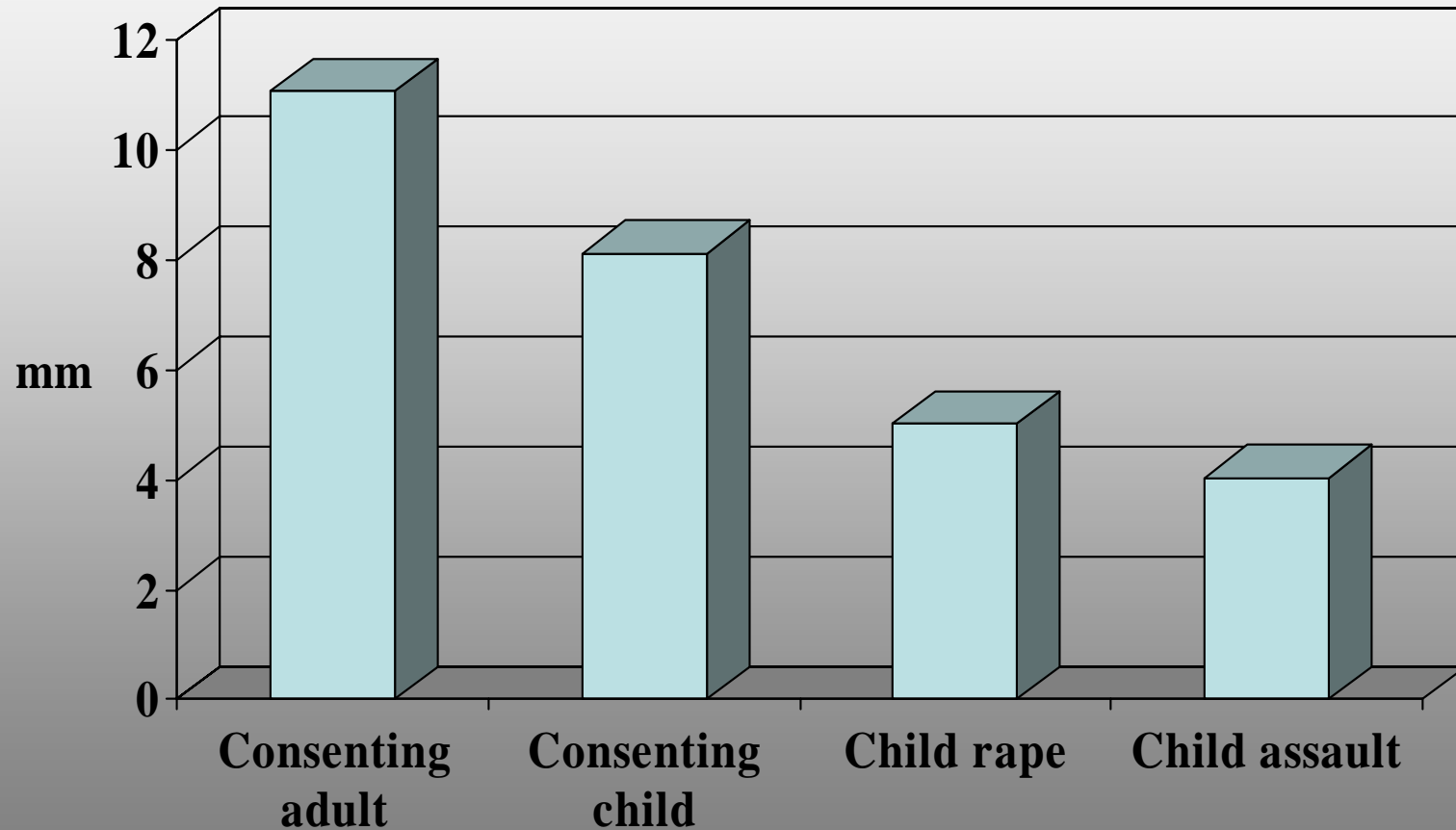
Psychophysiological Measures: Sexual Arousal Among Sexual Offenders (Hall, 1989)



Sexual Arousal Among Sexual Offenders (Hall et al., 1988)



Sexual Arousal Among Normal Men (Hall, 1995)



Limitations of Psychobiological Measures

- Physiological responses do not necessarily correspond to psychological states

Unobtrusiveness and Reactivity of Measures



Anonymity may
decrease response
bias

Filler or buffer items to
disguise the purpose
of the test

Computerized
assessment may be
less reactive

Unobtrusive measures

- Techniques
 - Observation in naturalistic settings
 - How is observation accomplished?
 - Frequency of target behavior
 - Lack of standardization
 - Observation in contrived situations - Stimulus is presented
 - Participant suspiciousness
 - Archival records (e.g., criminal records)
 - Selectivity in what is recorded
 - Physical traces (e.g., library vandalism)
- Ethical issues
 - Informed consent

Multitrait Multimethod Matrix

- Method variance

- Correlations will be high within one method (e.g., self-report, behavioral observations)
- Correlations of the same trait across methods support the validity of the measurement (concurrent validity)

- Trait variance

- Different constructs should not be correlated (discriminant validity)
- Correlation may suggest shared trait variance

Development and Validation of an Observational Coding System for Emotional Overinvolvement

Fredman et al. (2004)



Observational Coding

- Patients with obsessive-compulsive disorder or panic disorder with agoraphobia discussed with relatives the 2 most problematic issues in their relationship
- Interaction is videotaped and reaction of relatives is coded
 - Intrusiveness ($r = .92$)
 - Excessive self-sacrifice ($r = .91$)
 - Exaggerated emotional response ($r = .73$)
- KPI coding
 - Criticism ($r = .88$)

Interview

- Camberwell Family Interview - Interview with relative in patient's absence
 - Emotional overinvolvement
 - Criticism

Self-Report

- Relative Reaction Questionnaire
 - Facilitation - the extent to which the relative criticized the patient in an attempt to help him or her
 - Guilt/Responsibility - the extent to which the relative felt guilty or responsible for the patient's illness

Table 2

Multitrait–Multimethod Matrix: Parametric Correlations of Three Behavioral Observation Scales and Expressed Emotion Measures of Convergent and Discriminant Validity

Measure	INT	ESS	EER	CFI-EOI	RRQ-Facilitation	RRQ-Guilt/Resp.	KPI-Criticism	CFI-Criticism	<i>M</i>	<i>SD</i>
Behavioral INT	(.92)	.26*	.50**	.27**	.33**	.19	.16 ^a	.06	1.44	0.90
Behavioral ESS		(.91)	-.02	.14	.20	.18 ^b	.00	.07	1.40	0.82
Behavioral EER			(.73)	.35**	.23*	.13	.10	.11	1.23	0.63
CFI EOI				(.91)	.15	.11	-.02	.02	0.80	1.13
RRQ Facilitation					(.44)	.27**	.11	-.02	4.44	1.78
RRQ Guilt/Responsibility						(.71)	-.10	-.01	1.75	1.45
KPI criticism							(.88)	.36**	4.87	4.90
CFI criticism								(.81)	6.32	7.28

Note. As a result of missing data for the RRQ items, sample sizes ranged from 91 to 95. Reliability coefficients are on the diagonal. CFI EOI, KPI criticism, and CFI criticism were log-transformed to enhance the normality of the distributions for the purpose of conducting the parametric correlations, but means and SDs (standard deviations) are for the nontransformed variables. INT = intrusiveness; ESS = excessive self-sacrifice; EER = exaggerated emotional response; CFI = Camberwell Family Interview; EOI = emotional overinvolvement; RRQ = Relative's Reactions Questionnaire; KPI = Kategoriensystem für Partnerschaftliche Interaktion.

^a The nonparametric correlation (.17) was statistically significant ($p < .05$). ^b The nonparametric correlation (.18) was statistically significant ($p < .05$).

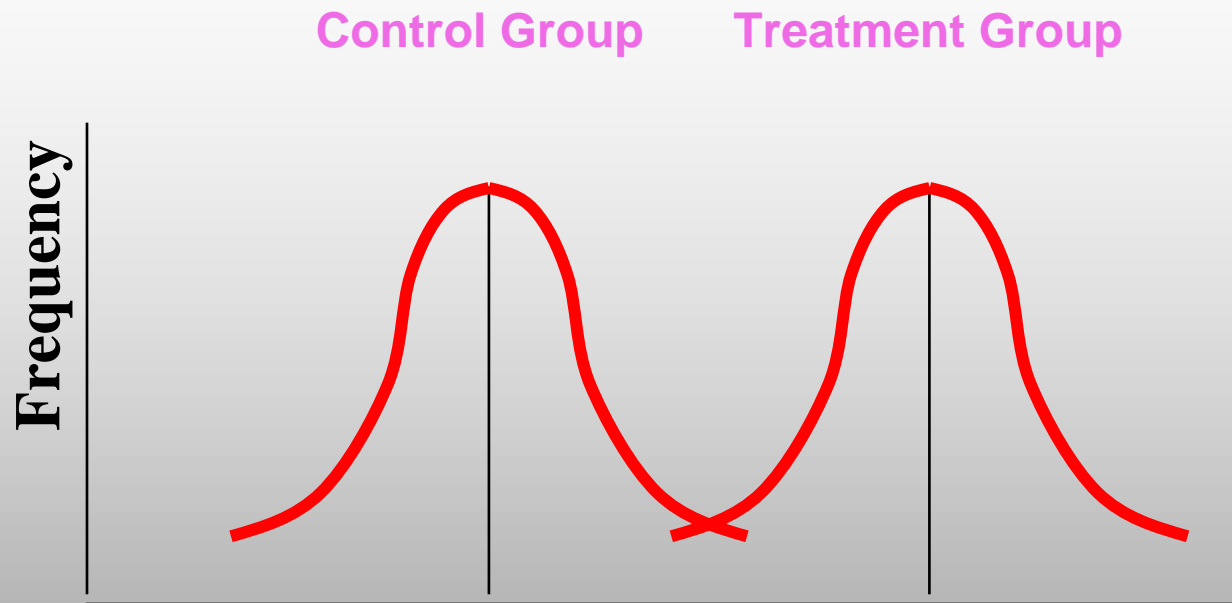
* $p < .05$. ** $p < .01$.

Assessment and Evaluation of Interventions

Assessing the Clinical Significance of Changes

- Practical value
 - Most changes are a matter of degree rather than complete elimination
- Clinical significance
 - No longer meets diagnostic criteria
 - 2 SDs from the mean of the dysfunctional sample

Clinical Significance: Nonoverlapping distributions



CBT for Panic Disorder

Kenardy et al. (2003)

- Clinically significant change = 1.96 SD above dysfunctional group
- Computer augmentation
 - Programmed to signal the participant five times daily—at 8 a.m., 11 a.m., 4 p.m., 7 p.m., and 9 p.m.—to prompt the practice of the therapy components.
 - The computer program included a self-statement module, a breathing control module, and a new exposure module incorporating both situational exposure and interoceptive exposure.
 - The exposure included goal setting and specification of exposure tasks. In the case of interoceptive exposure, specific tasks, such as hyperventilation, were presented in relation to concern about particular salient somatic symptoms, such as difficulty breathing.

Table 2
Percentage of Participants Meeting Conjoint Criteria for Clinically Significant Change

Variable	WL	CBT6	CBT6-CA	CBT12	$\chi^2(7)$
Body Sensations Questionnaire					
Posttreatment	6.1	38.2	45.7 _a	51.4 _a	28.42, $p < .001$
6-month follow-up	—	32.1	30.0	55.9	
Agoraphobic Cognitions Questionnaire					
Posttreatment	20.6	50.0	64.7 _a	55.3 _a	20.54, $p = .005$
6-month follow-up	—	60.7	48.3	54.3	
State-Trait Anxiety Inventory-Trait					
Posttreatment	5.9	48.5	60.0 _a	60.5 _a	43.87, $p < .001$
6-month follow-up	—	48.1	46.7	60.0	
Mobility Inventory					
Alone					
Posttreatment	2.9	26.5	54.5 _a	60.5 _a	39.85, $p < .001$
6-month follow-up	—	39.3	44.8	60.0	
Accompanied					
Posttreatment	2.9	32.4	36.4 _a	47.4 _a	25.57, $p = .001$
6-month follow-up	—	39.3	41.4	48.6	
Fear Questionnaire					
Agoraphobia					
Posttreatment	14.7	38.2	37.1	43.2	10.28, $p = .173$
6-month follow-up	—	39.3	36.7	47.1	
Social Phobia					
Posttreatment	8.8	20.6	28.6	32.4	6.92, $p = .44$
6-month follow-up	—	27.9	30.0	47.1	
Blood/injury					
Posttreatment	8.8	44.1	42.9 _a	45.9 _a	19.66, $p = .006$
6-month follow-up	—	65.5	28.7	38.7	

Note. Conjoint criteria represents the percentage of individuals who reliably improved (reliable change index > 1.96) and whose mean placed them closer to the mean of the functional population than the mean of the dysfunctional population. Entries with subscripts differ significantly from WL at $p < .05$. WL = wait list; CBT6 = 6 sessions of therapist-delivered cognitive-behavioral therapy; CBT6-CA = CBT6 with computer augmentation; CBT12 = 12 sessions of therapist-delivered cognitive-behavioral therapy. Dash indicates data were not available.

Comparison Methods

- Normative
 - Who is the normative group?
 - Are peer norms appropriate?
 - How do others view the problem?
 - Social impact measures
 - e.g., school referrals, arrest rates, hospitalizations



Comparison Methods

- Ipsative
 - Does impairment still exist?
 - Is the change clinically significant?
 - Does the person feel better?
 - A small change may make a big difference



Value of Treatment

- Disseminability
 - Extension to new settings
- Cost
 - Comprehensiveness
 - Cost of treatment vs. no treatment
- Acceptability



Table 1

Average Program Expenditures, Number of Hours Used by Patients and Cost Per Clinical Hour for Participants in the Fals-Stewart et al. (1996) Study

Variable	Total
Program expenditures	
Labor (salaries and fringe benefits)	\$510,720
Contracted services ^a	\$107,290
Rent and maintenance	\$41,970
Other ^b	\$59,650
Total average annual costs	\$719,630
Average number of hours used by patients per year ^c	17,615
Average cost per hour of service	\$40.85

Note. All dollar figures were converted to 1992 U.S. dollar equivalents using the Consumer Price Index.

^a Contracted services included laboratory analysis and fees to pay physicians, accountants, and attorneys. ^b Examples include utilities, telephone, office supplies, repairs, training, and other miscellaneous costs.

^c This is the mean number of hours used per year for the years 1991 through 1993.

Table 2

Aggregate Costs of Providing Individual, Conjoint, and Group Therapy for Husbands in the Behavioral Couples Therapy and Individual-Based Treatment Conditions

Variable	BCT (<i>n</i> = 40)		IBT (<i>n</i> = 40)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
No. of sessions attended				
Individual	19.7	3.4	32.7	3.4
Group	13.1	5.4	9.8	4.9
Conjoint	10.1	2.3	0.0	0.0
Total cost	\$1,372.72	341.2	\$1,359.94	314.7

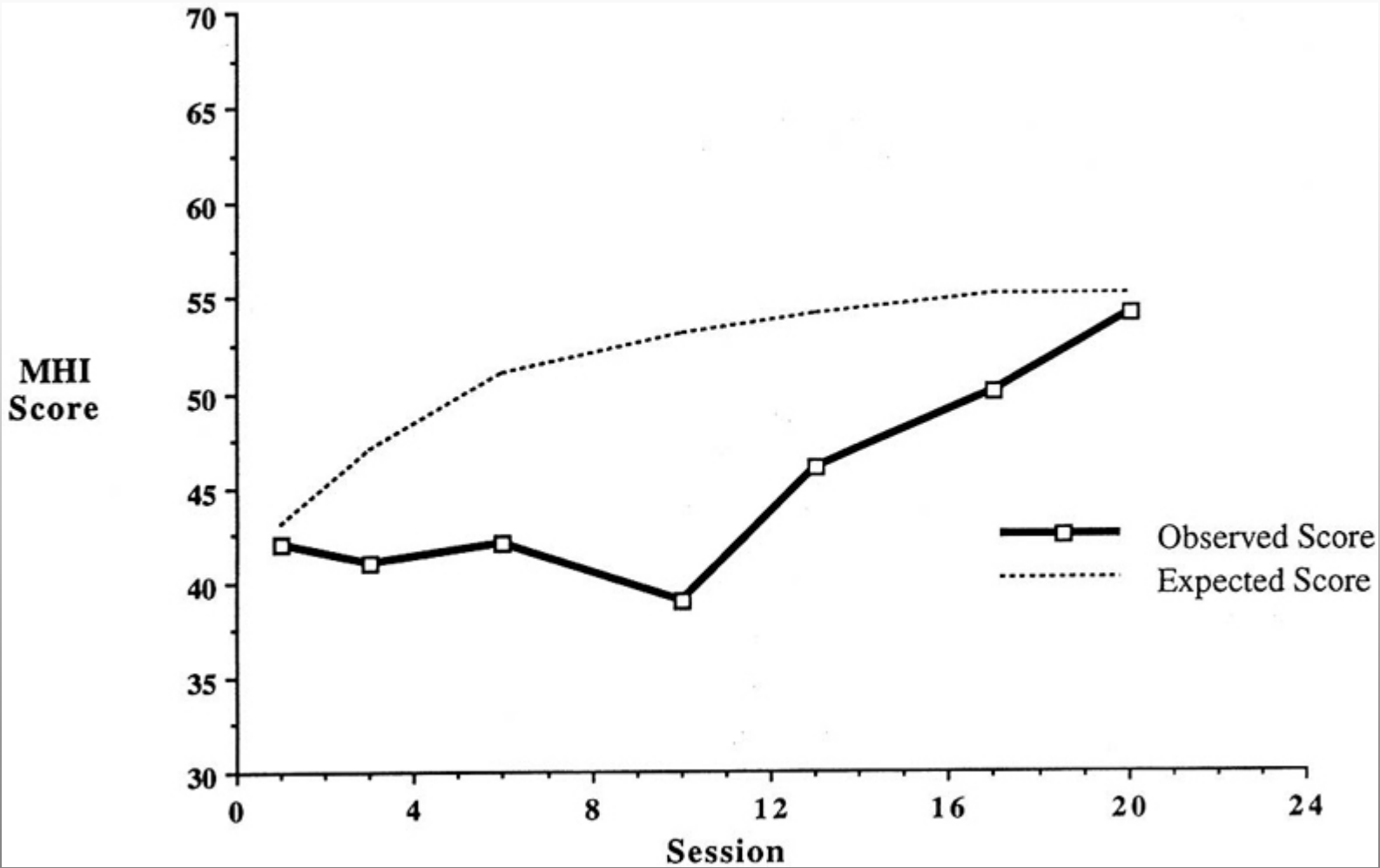
Note. To calculate costs of providing treatment, we multiplied the number of individual and conjoint sessions by \$40.85 (i.e., the cost of providing 1 hr of clinical services); group sessions were multiplied by \$11.56 (i.e., the cost of providing 1.5 hr of clinical service to an average group of 5.3 patients). The costs of providing these types of interventions were summed to derive a total cost of the treatment. BCT = behavioral couples therapy plus individual-based treatment. IBT = individual-based treatment only.

Assessment During the Course of Treatment

- Pre-post assessment most common
- Assessment throughout the course of treatment
 - May reveal the mechanisms of change (e.g., therapeutic alliance)
 - May reveal when changes occur

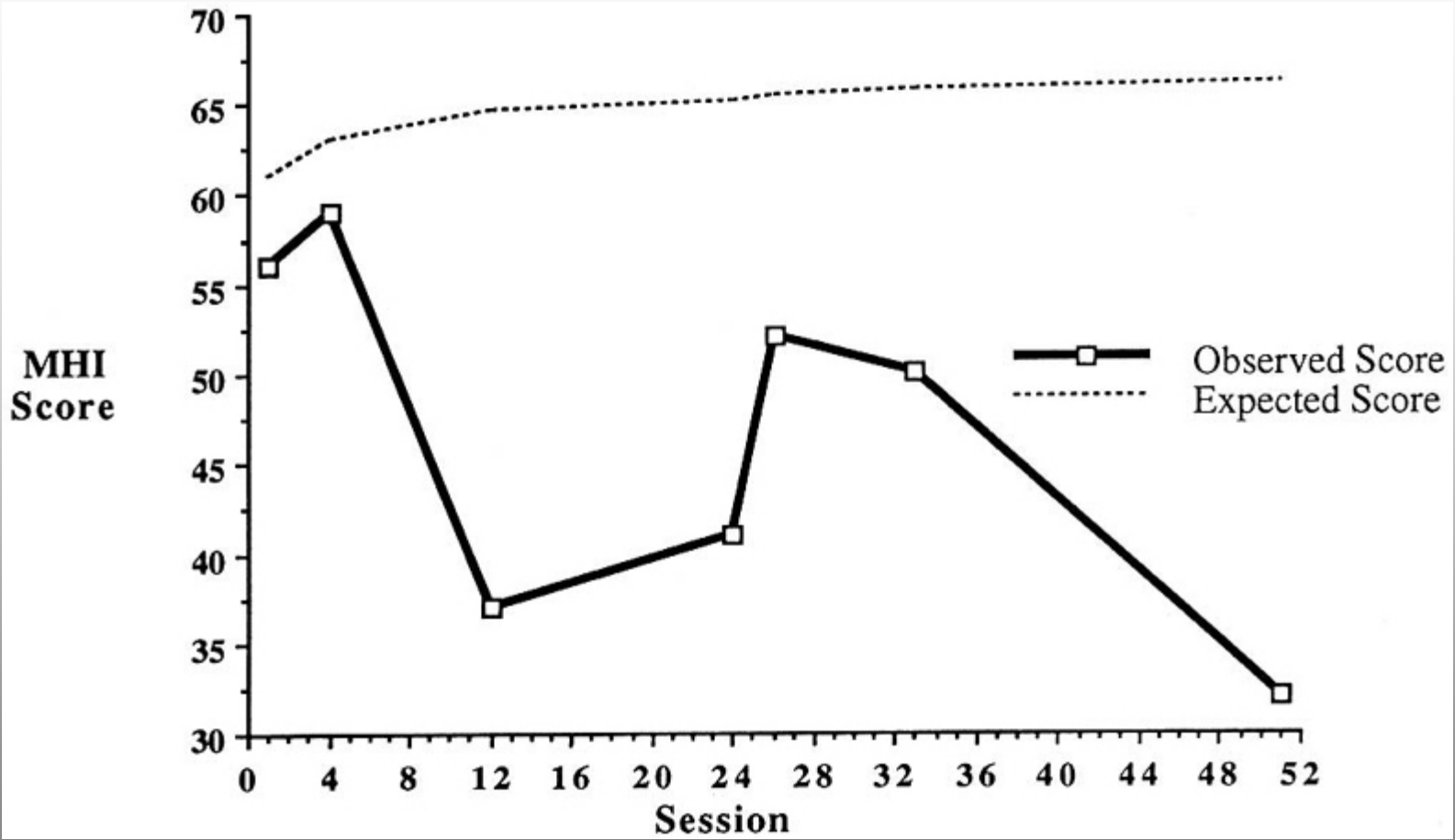
Patient responding to psychotherapy as expected after a change in treatment intervention (Leon et al., 1999)

- The patient was a 32-year-old, employed, married woman presenting with depression and alcohol abuse after the death of her father. Treatment initially focused on her family and social history; however, this approach proved unsuccessful and after 10 sessions she was responding below her expected trajectory. A “solution-focused” approach was implemented in Session 11, and by Session 20 the patient had improved according to expectations.



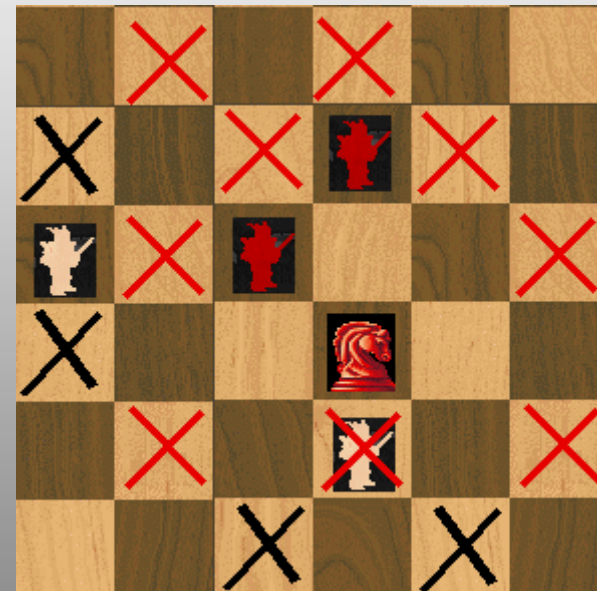
Patient responding below expectations in psychotherapy

- The patient was a 47-year-old, unemployed man presenting with anxiety due to a lack of career goals and inconsistent social support. He entered treatment with little expectation for improvement. Unfortunately, he was unable to achieve even the modest progress and outcome predicted for him.

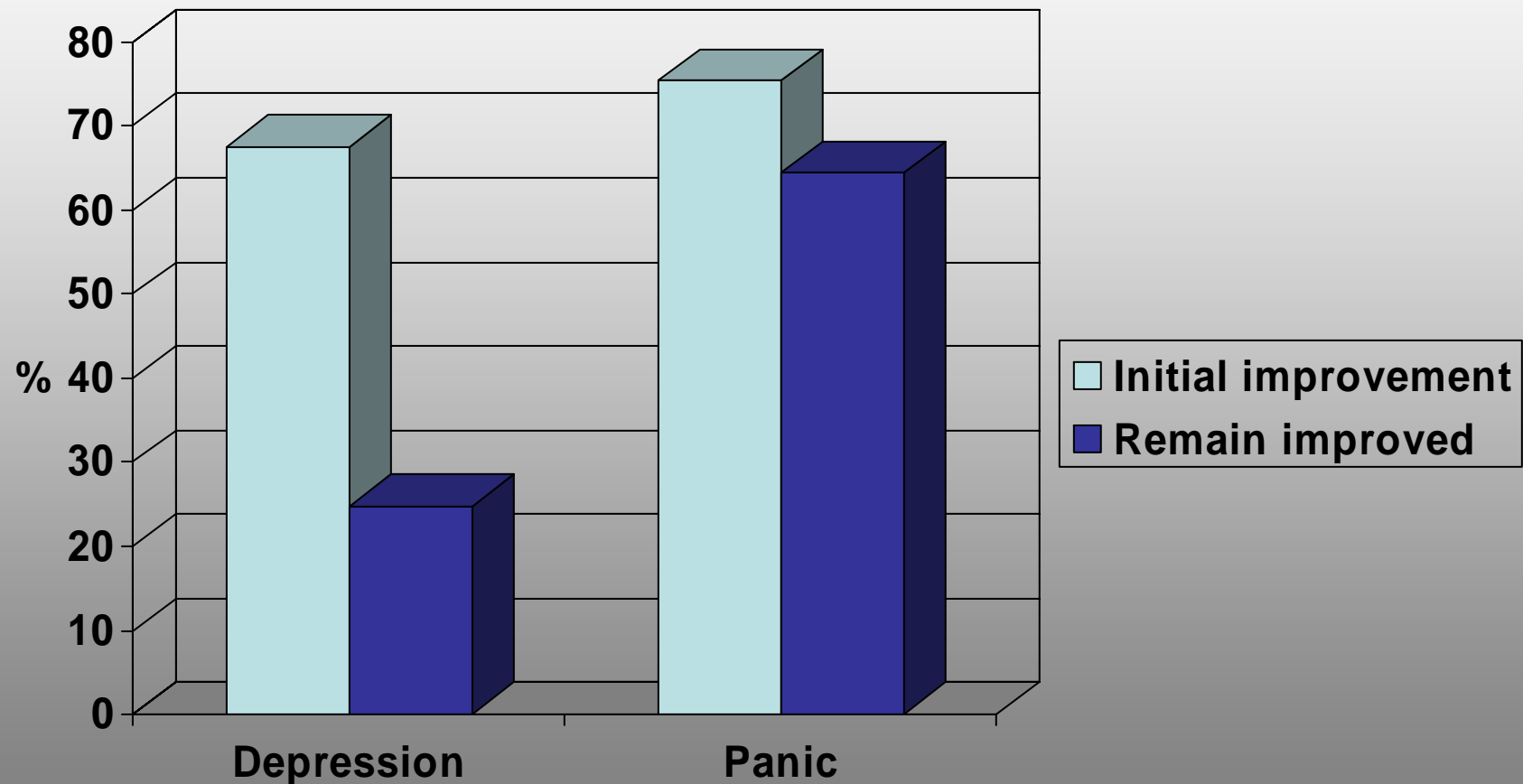


Follow-up Assessment

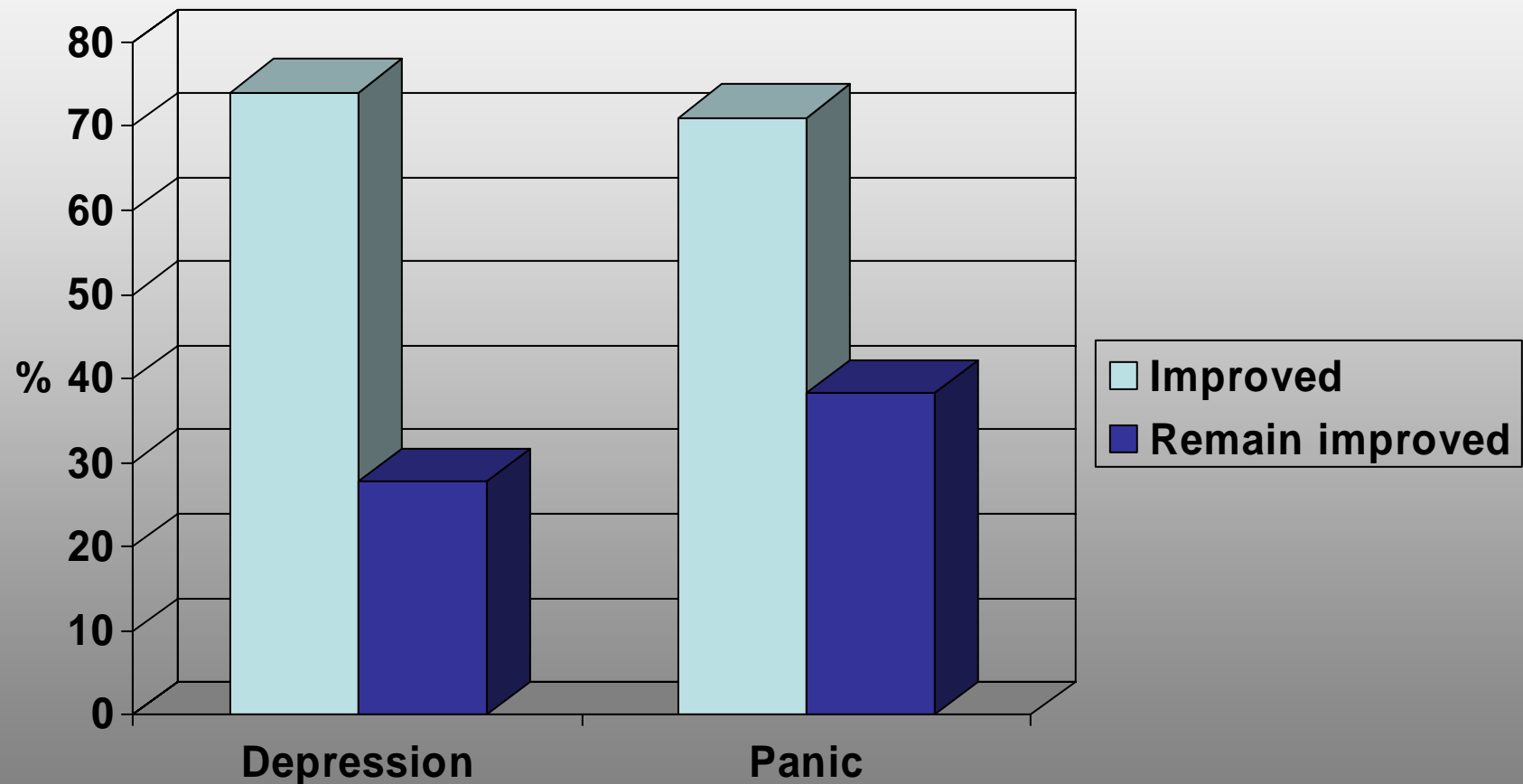
- What are the lasting effects of treatment?
- Attrition may prevent accurate follow-up assessment
 - Ongoing contact with participants may prevent attrition



Treatment Effects 12-18 Months After Termination (Westen & Morrison, 2001)



Treatment Effects 24+ Months After Termination (Westen & Morrison, 2001)



Statistical Methods of Data Evaluation

Statistical Significance

- $p < .05 = 5$ times by chance out of 100
- Statistical significance is a function of sample size
 - Large samples make trivial differences significant
- Statistical significance does not address the strength or importance of an effect



Alternatives or supplements to significance tests

- Strength of effect – r , r^2 , R , R^2
 - Small effect size, $r = .1$
 - Medium effect size, $r = .3$
 - Large effect size, $r = .5$
- Effect size is not equivalent to importance

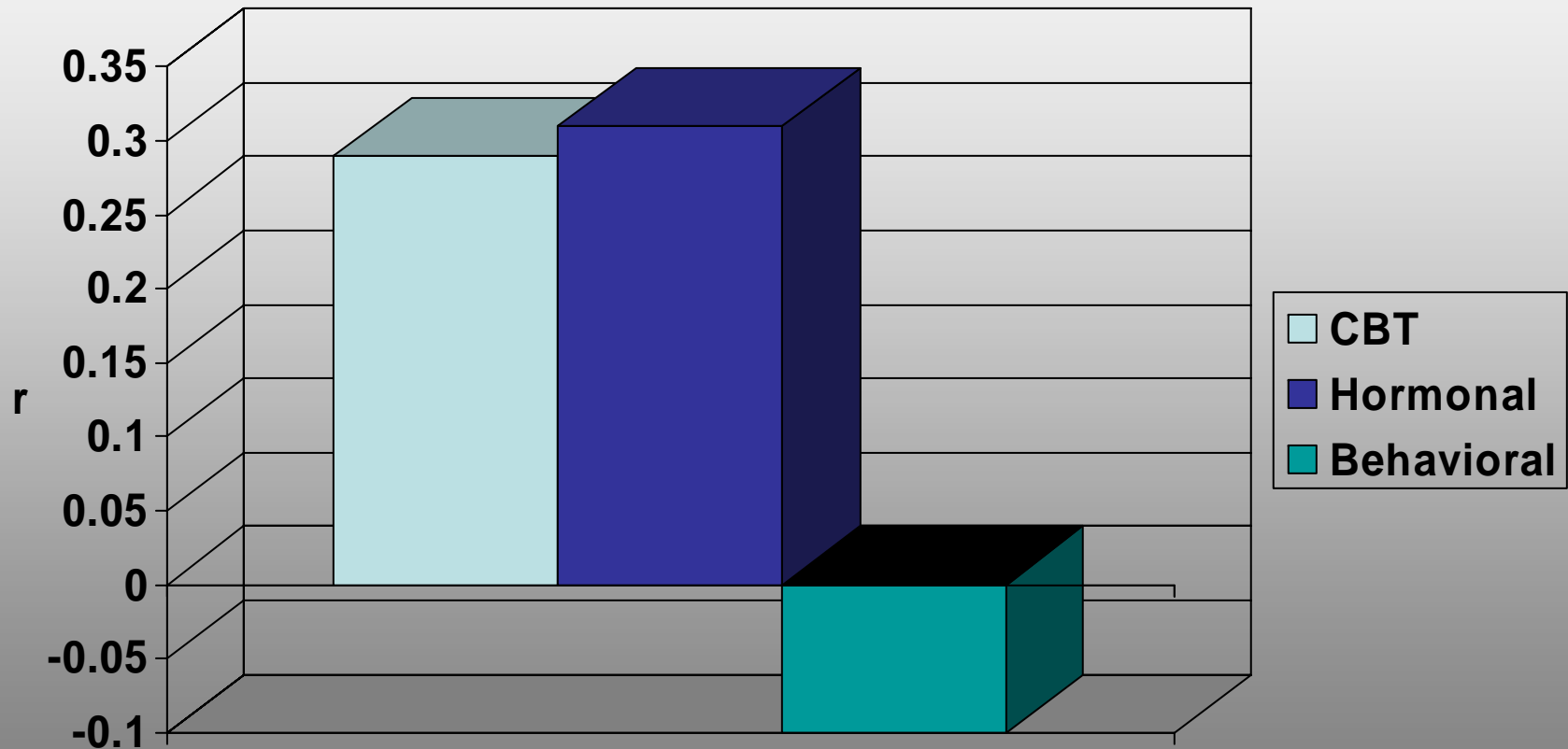
Alternatives or supplements to significance tests

- Confidence intervals – likelihood that the effect size falls within a particular range
 - e.g., $ES = .3 \pm .05$, 95% confident that the effect size is within $r = .25$ to $.35$

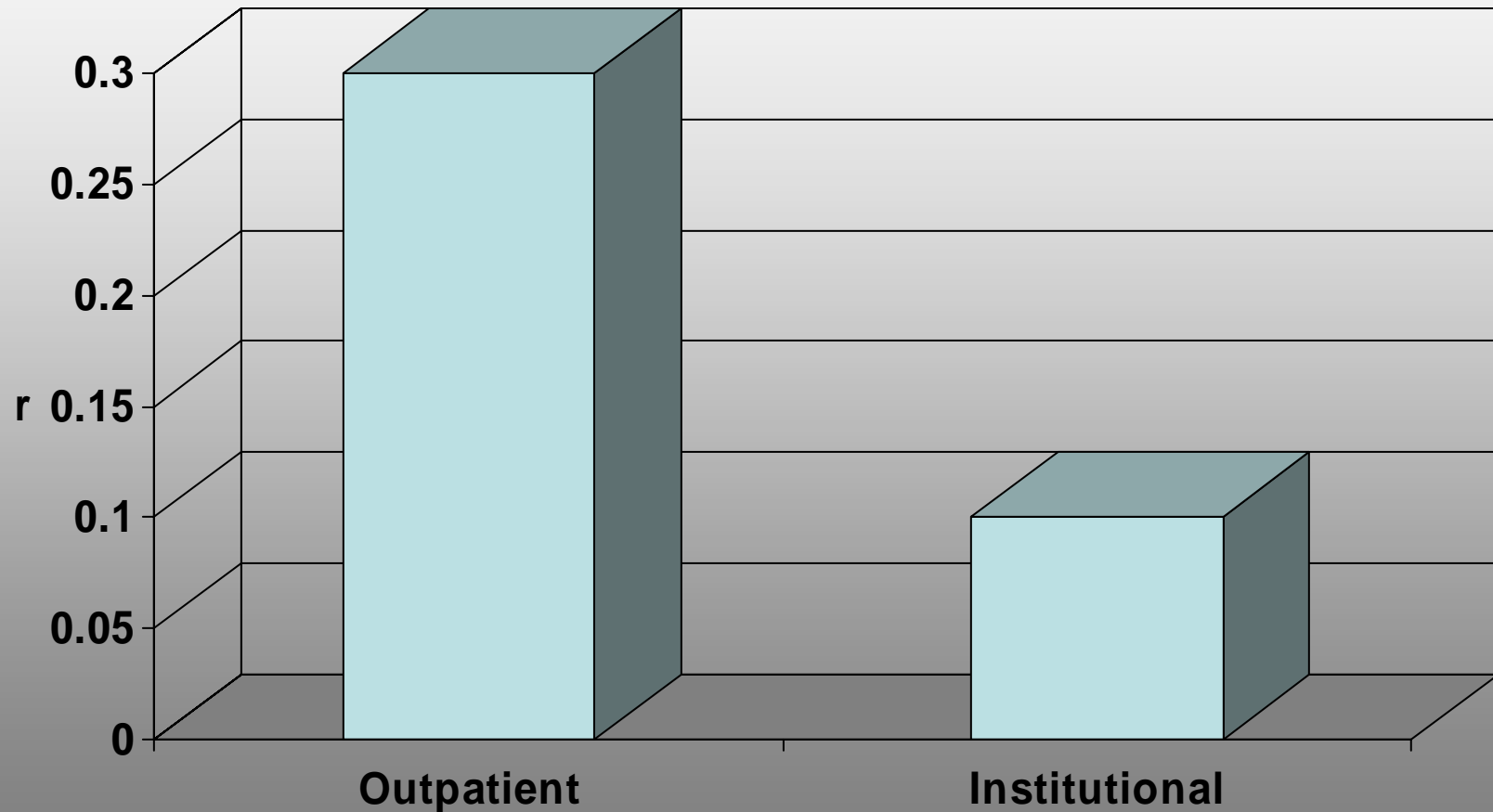
Alternatives or supplements to significance tests

- Meta-analyses
 - Effect sizes of multiple studies are combined
 - Studies need similar measures for inclusion
 - Quality of studies may vary
 - Individual studies may be weighted (e.g., by N , quality)
 - Moderators can be analyzed (e.g., groups for which effects are strongest)

Meta-analysis of 12 sex offender treatment studies (Hall, 1995)



Moderator of Treatment Effects (Hall, 1995)

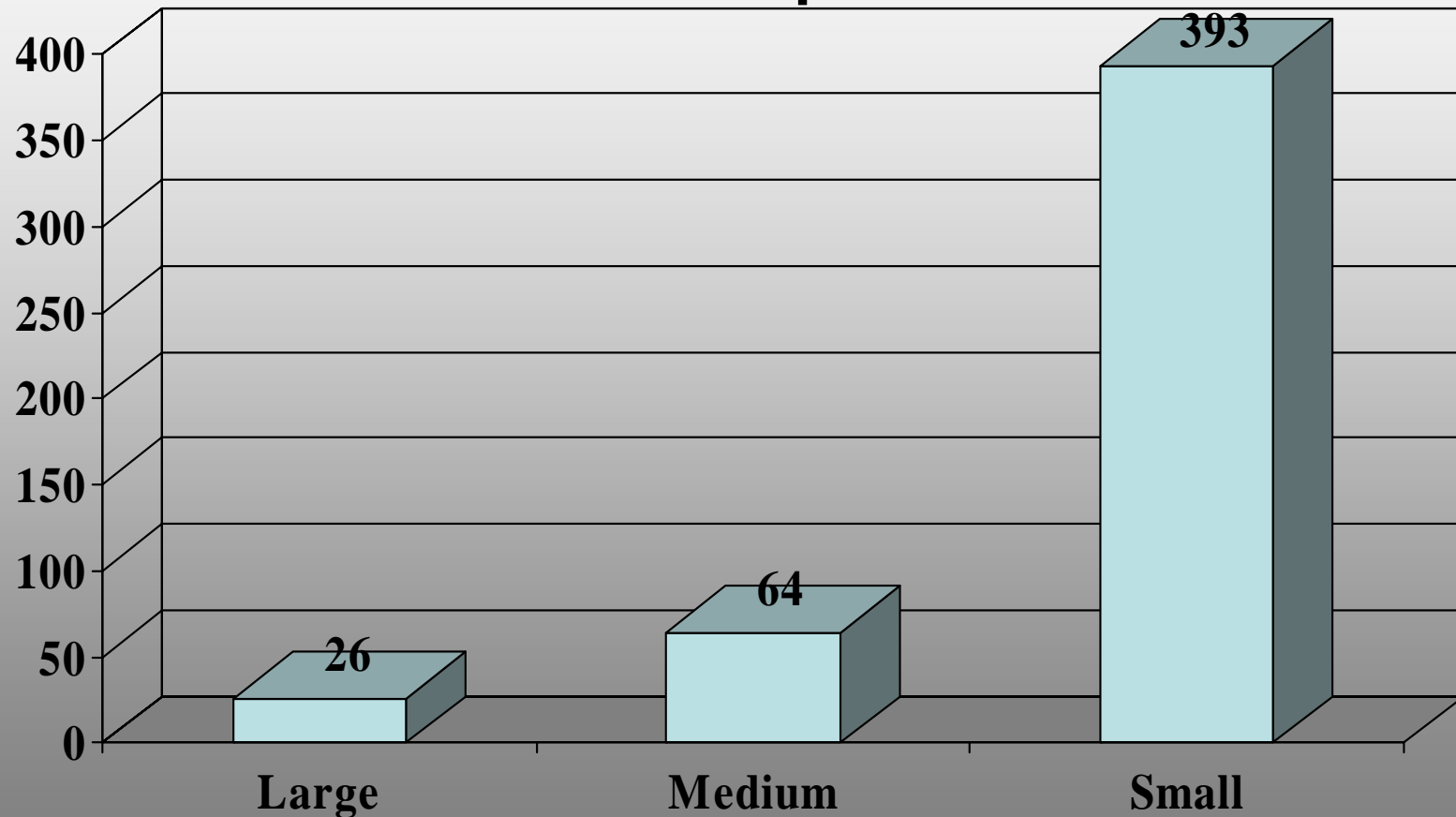


Power



- .80 – chance of detecting a difference is 4 out of 5 if there is a real difference
- Most psychotherapy studies do not have sufficient power to detect small ($r = .1$) and medium ($r = .3$) effect sizes with power = .80

Participants Per Group Necessary to Detect Effect Sizes for an ANOVA w/ 2 Groups, Power = .80



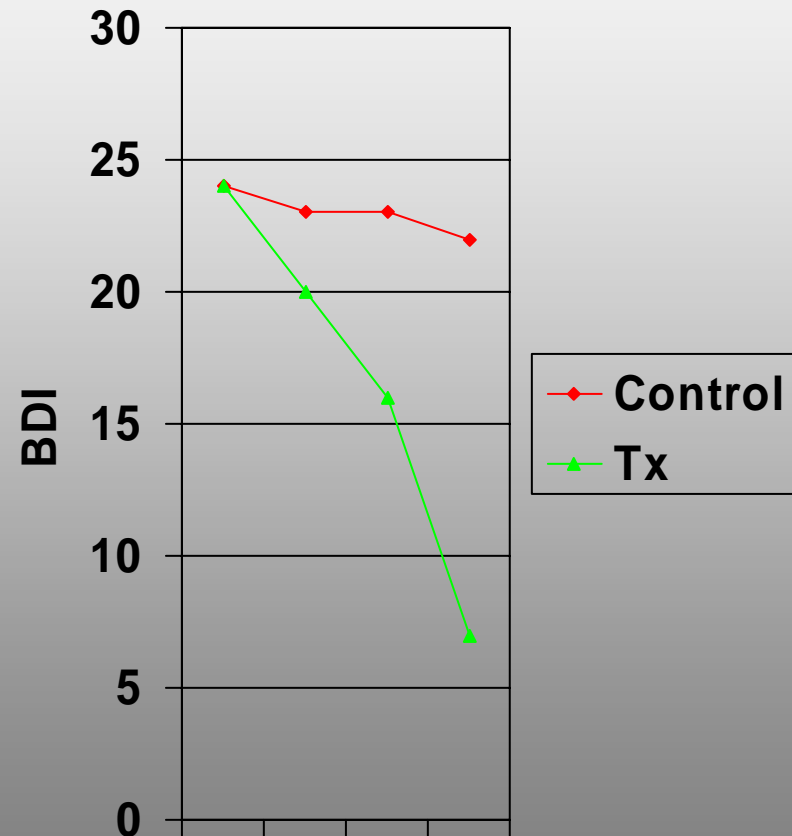
Increasing Power

- Increase N



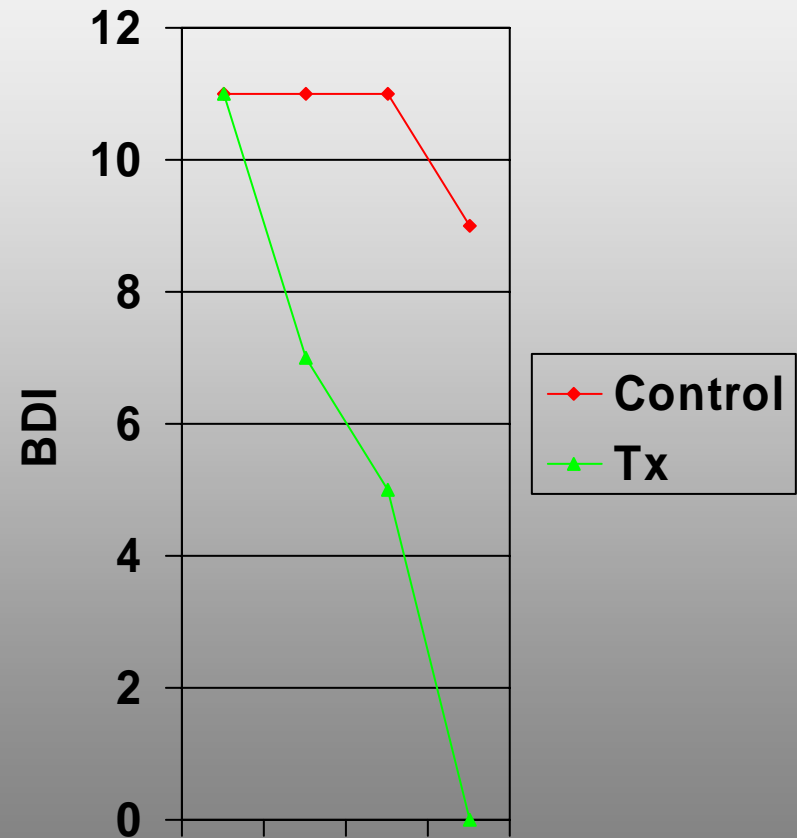
Increasing Power

- Extreme conditions (e.g., severe psychopathology) may increase effects



Increasing Power

Potent interventions
may increase effects



Increasing Power

- Use a pretest to account for within-subject variability and reduce error term
 - Without repeated measures: $ES = (m_1 - m_2) / s$
 - With repeated measures: $ES = (m_1 - m_2) / s \sqrt{1 - r^2}$

Increasing Power

- Decrease variability (error)
 - Heterogeneity of participants, inconsistent procedures, unreliable measures



Multiple Comparisons

- Bonferroni adjustment - Divide alpha ($p = .05$) by the number of comparisons
 - For 10 comparisons, $p = .005$ ($.05/10$)
- Multivariate analyses – several measures in a single data analysis
 - Alpha adjustments for post hoc tests

Objections to Statistical Significance Testing

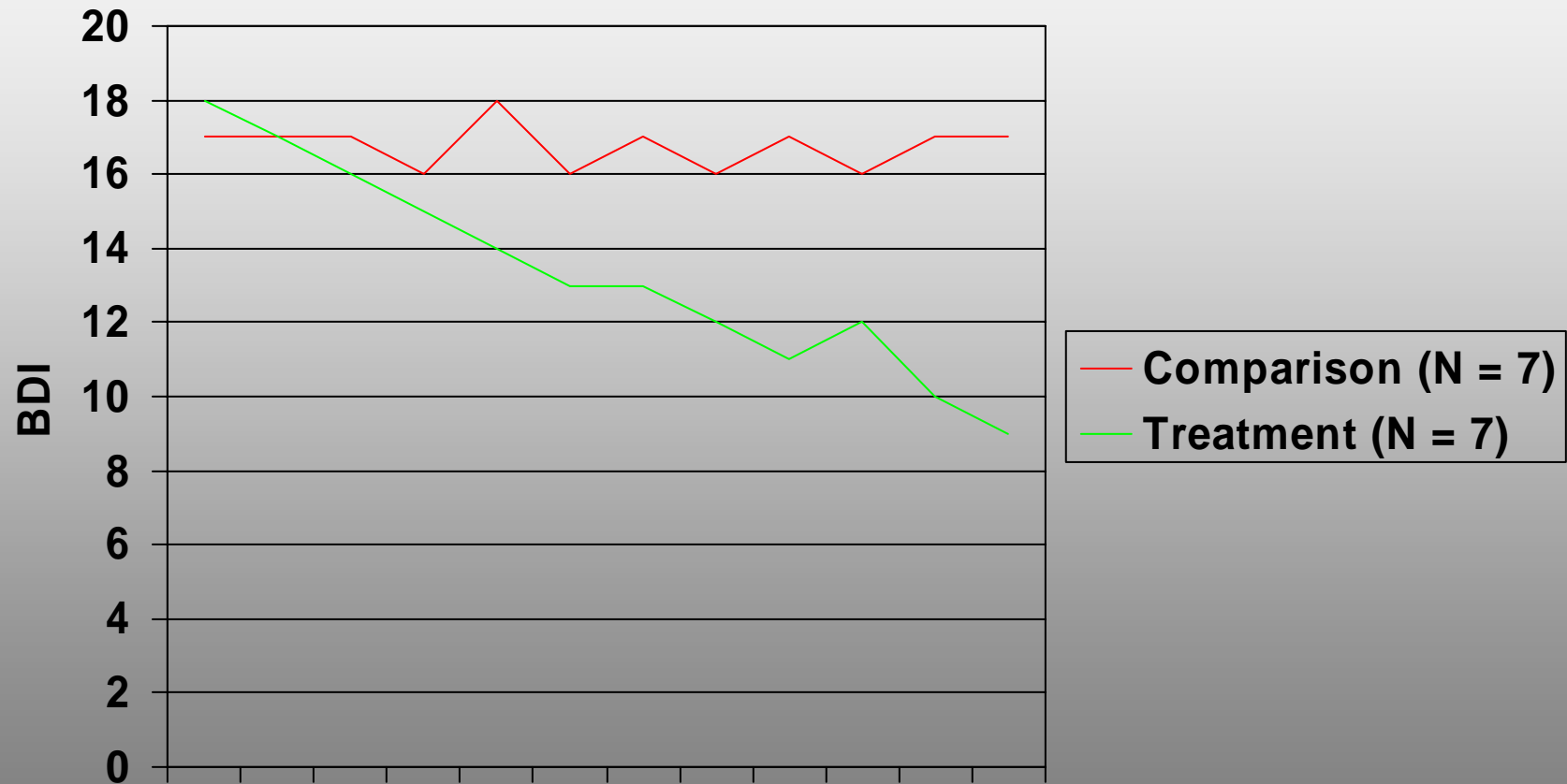


- All or none decision making ($p < .05$)
 - Studies without statistically significant differences are unlikely to be published
- Null hypothesis is rarely true
- Significance is a measure of N
 - Many psychology studies have low power
 - The same results in 2 studies may be statistically significant or nonsignificant
- Statistical significance says nothing about the importance of the effects

Objections to Statistical Significance Testing

- p values refer to the likelihood that the finding would be obtained by chance if a large number of tests were run
 - They do not reflect the likelihood that the findings are true
 - A lower p value ($p < .0001$) is not a stronger effect than a higher p value ($p < .05$)
- A failure to detect a statistically significant difference does not necessarily mean that there is no difference

No Differences: Low Power

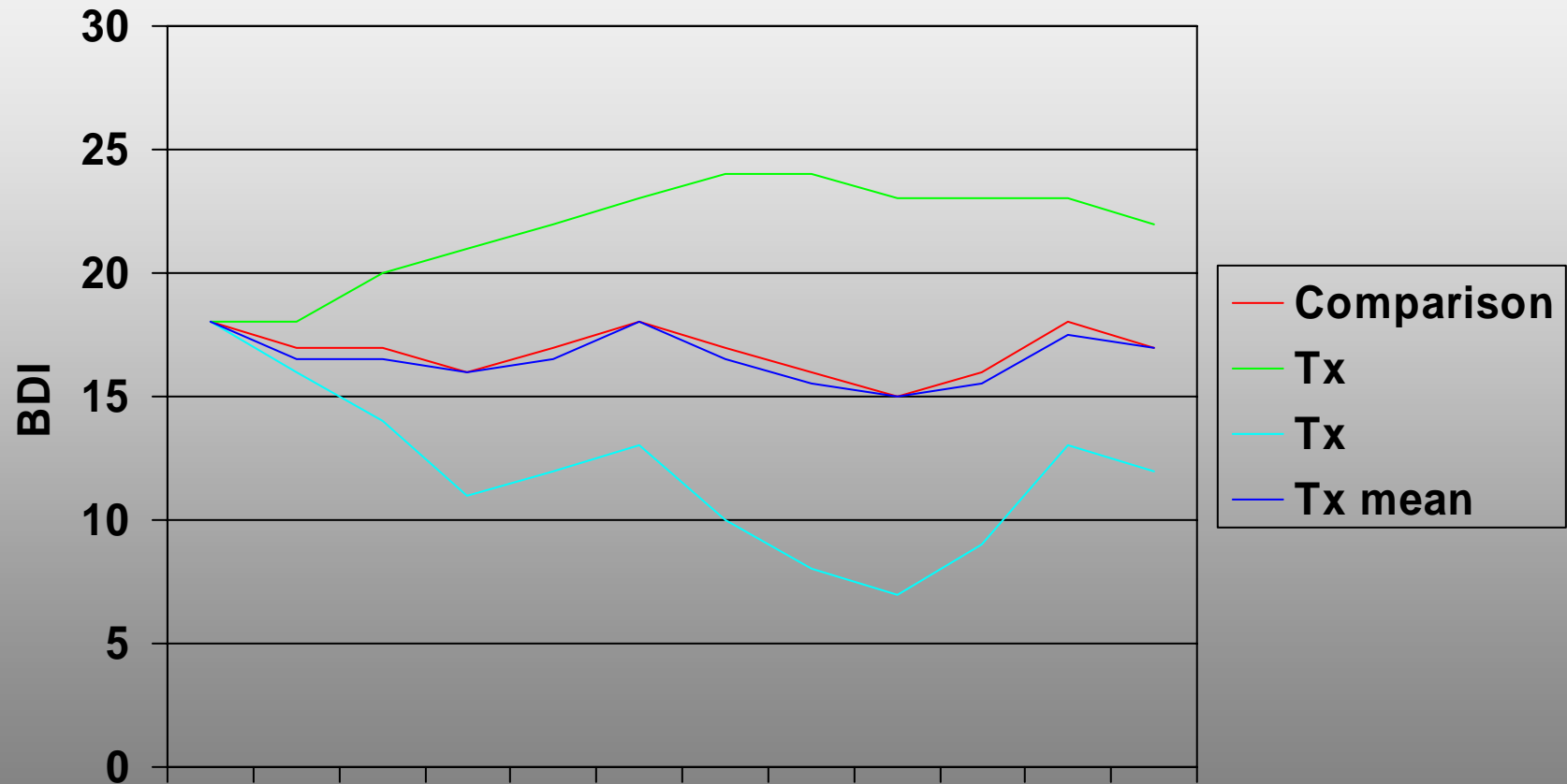


No Differences: Manipulation Failure

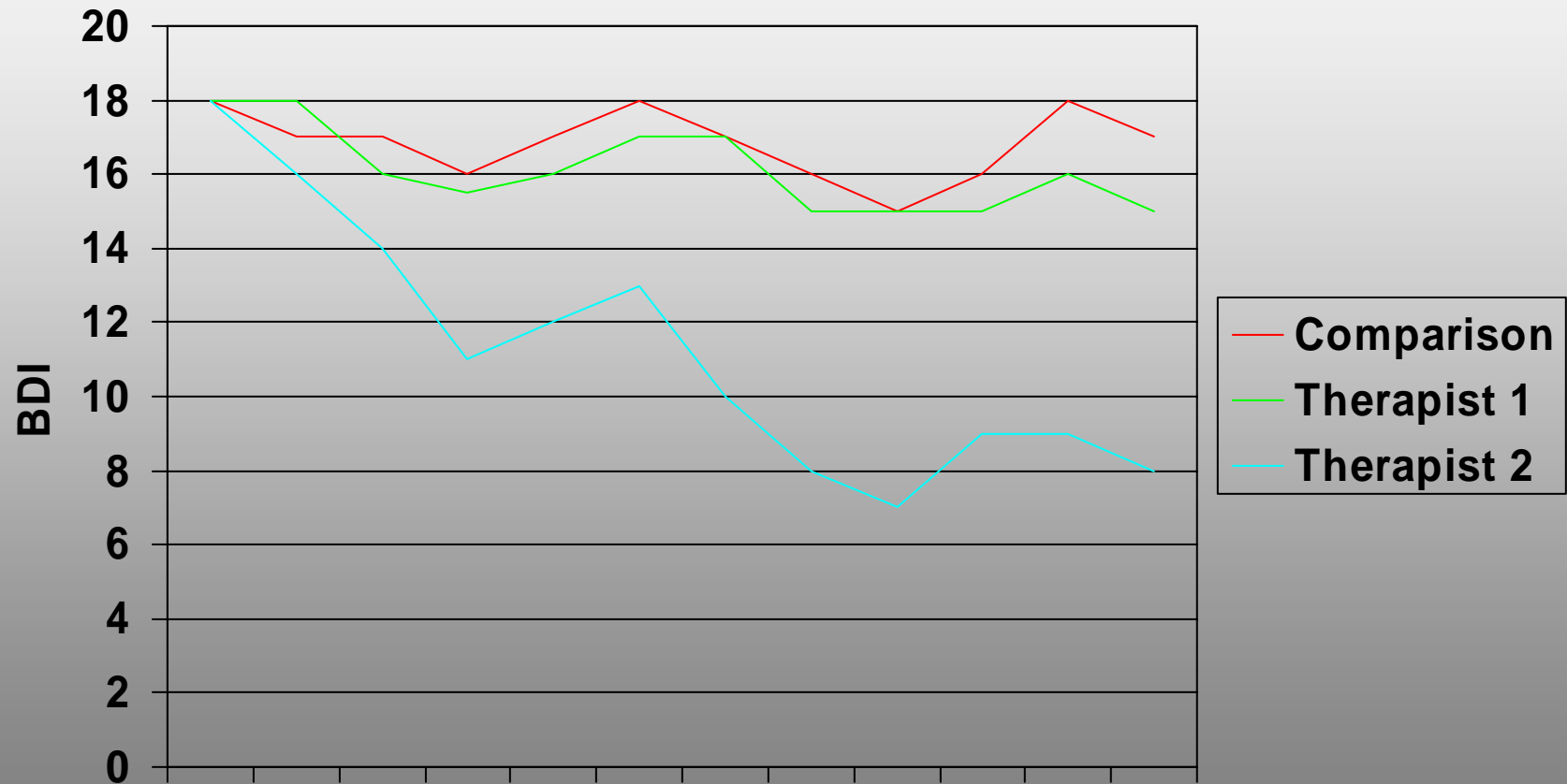
- Was the manipulation successful?
 - e.g., Diffusion of treatment – Treatment as usual group received elements of experimental treatment



No Differences: Participant Heterogeneity



No Differences: Inconsistent Procedures



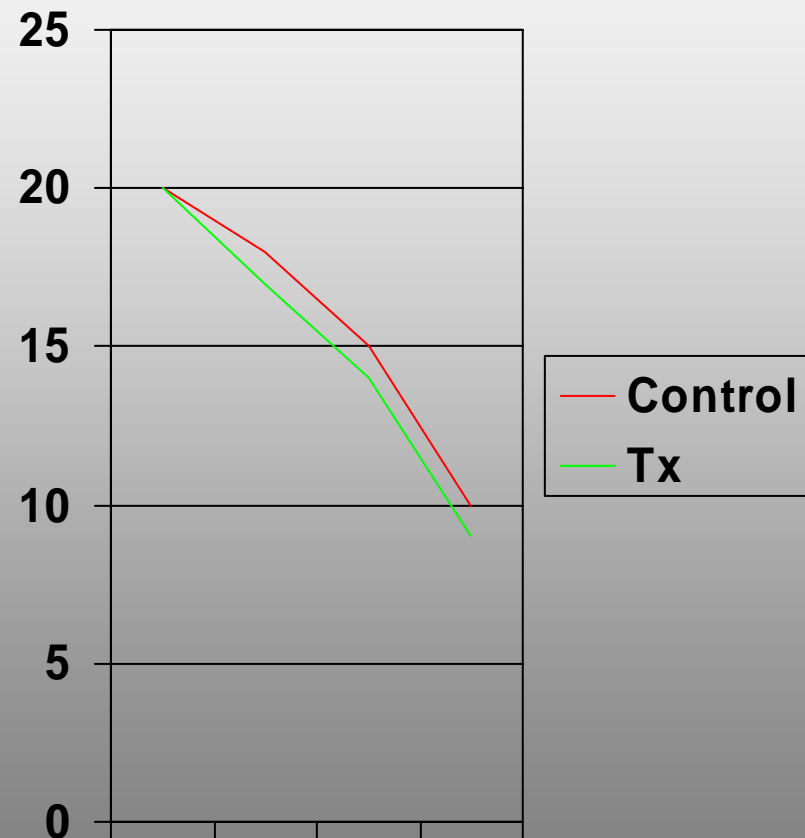
No Differences: Error

- Unreliable measures

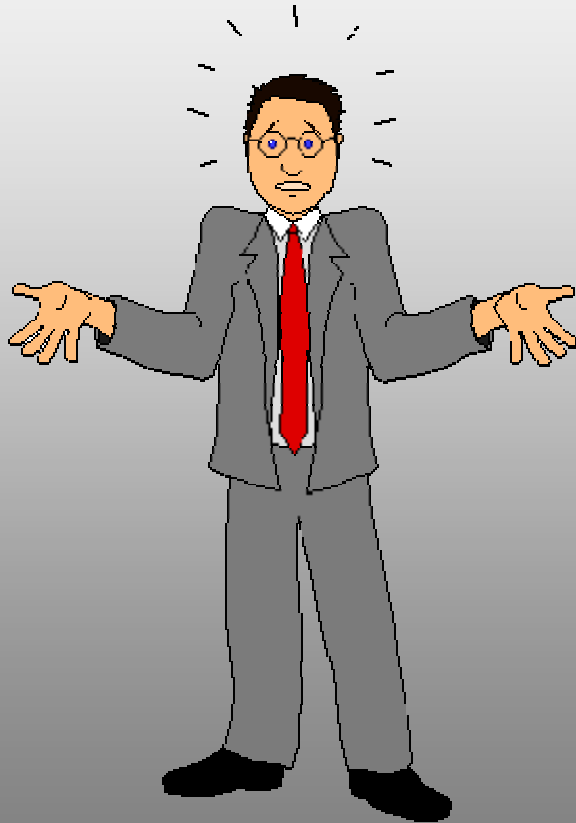


No Differences: Competing Influences

- Competing influences
 - Maturation
 - Statistical regression



No Differences



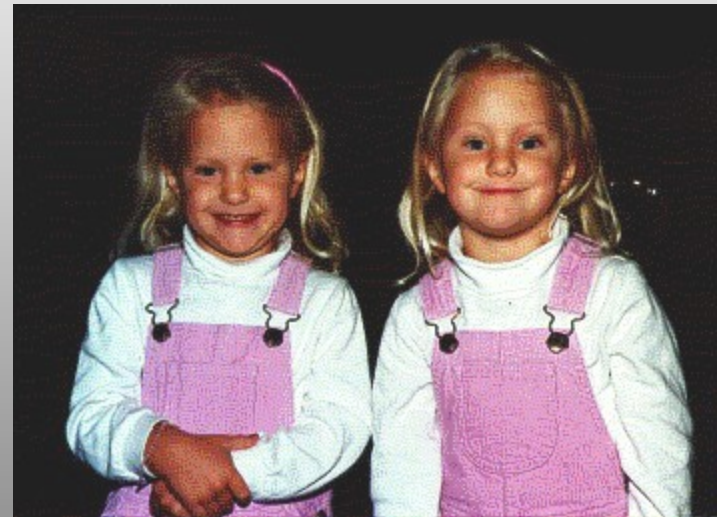
- May be informative if replicated
- Demonstrate the limits to the generalizability of the findings

Interpretation of Results

- Overinterpretation
 - “Highly significant effect” has no statistical meaning
 - Nonsignificant trends generally should not be interpreted
 - “My results are not significant” – tests vs. results
 - “Predicts” - Can be used only if design warrants
 - “Fishing” – post hoc vs. a priori interpretations
 - “Implications” – how far are they from the actual results?

Replication

- Direct replication



Replication

- Systematic replication
 - varying features
 - Are there other IVs that may influence the DV?



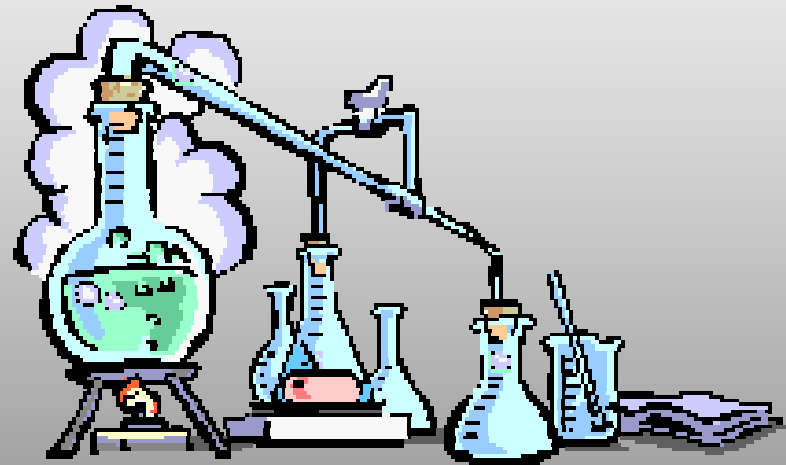
Replication

- Replication by experimenters other than the original one



Replication

- Replication in non-laboratory contexts
 - Efficacy vs. effectiveness research



Ethical Issues and Guidelines for Research

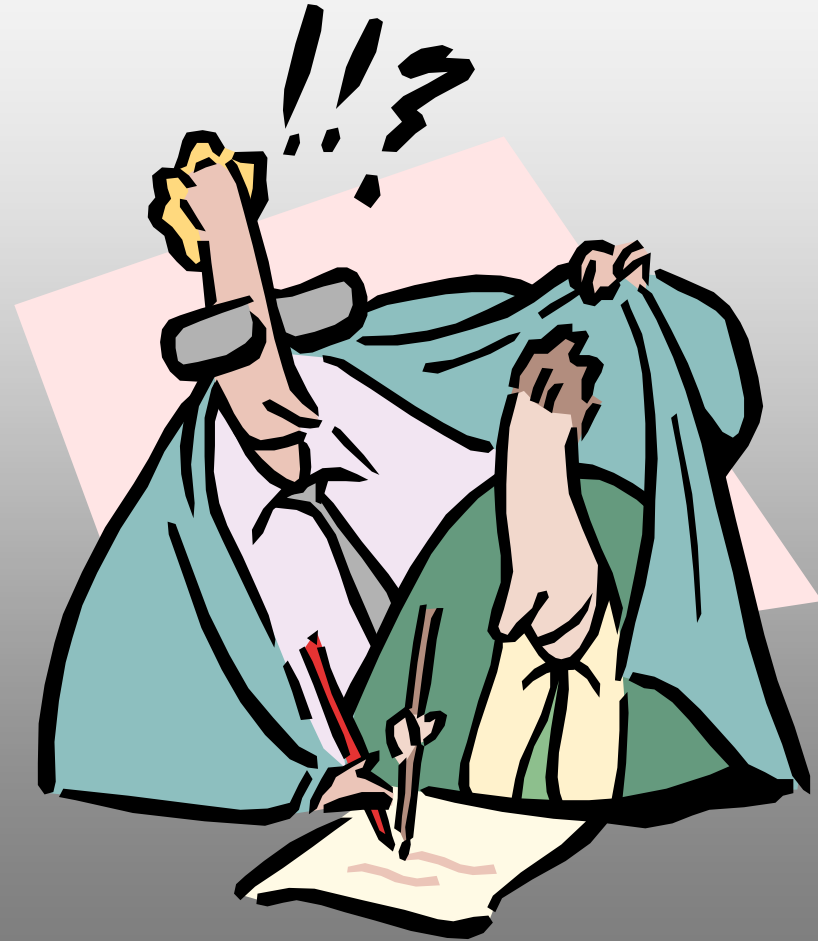
Investigator

- Power disparity between investigator and participant
 - “subject” vs. “participant”



Deception

- Do the benefits outweigh potential harm?
- Can less deceptive methods be used?
- How aversive is the deception?



Debriefing

- Participants may believe that the debriefing is part of the experiment and continued deception
- Debriefing may not completely erase the false impressions of the deception



Delayed Debriefing



- Is potential participant pool contamination a legitimate justification for delayed debriefing?

Invasion of Privacy

- Anonymity
 - Identity is not revealed
- Confidentiality
 - No disclosure to 3rd party
- Third party informants
 - Consent of participant about whom information is gathered



Informed Consent

- Capability of providing informed consent
- Risks explained
 - Does the participant understand the risks?
- Participation without coercion
 - Is offering a relatively large sum of money (e.g., \$40) coercive?
- Opportunity to withdraw self or data



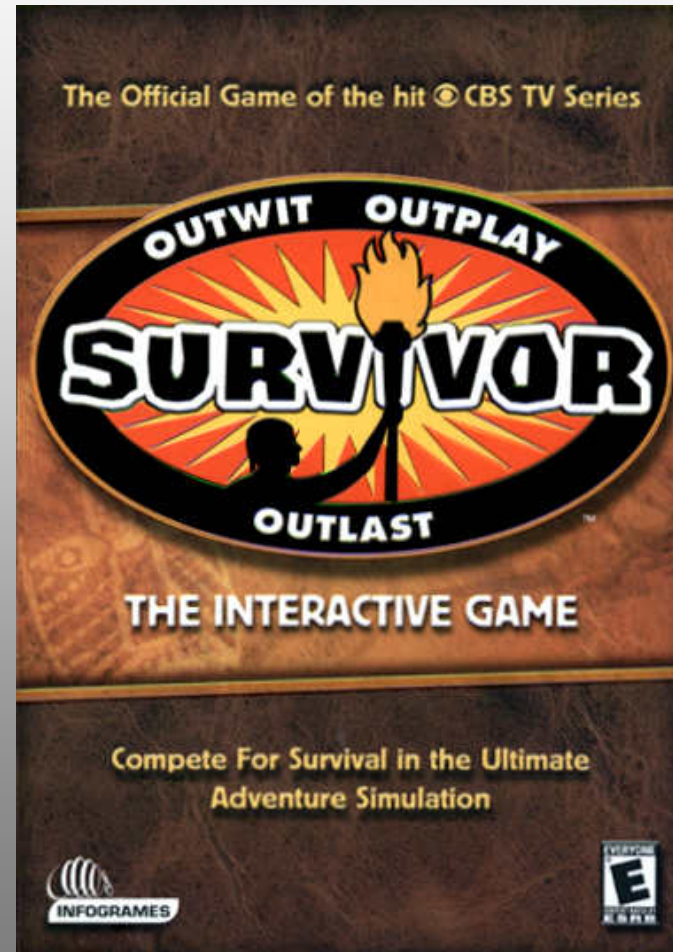
Intervention Research Issues



- Current status of the treatment provided should be disclosed, which may result in:
 - Skepticism
 - Self-selection

Intervention Research Issues

- No treatment control
 - Can participants survive no treatment?
 - Waiting list control



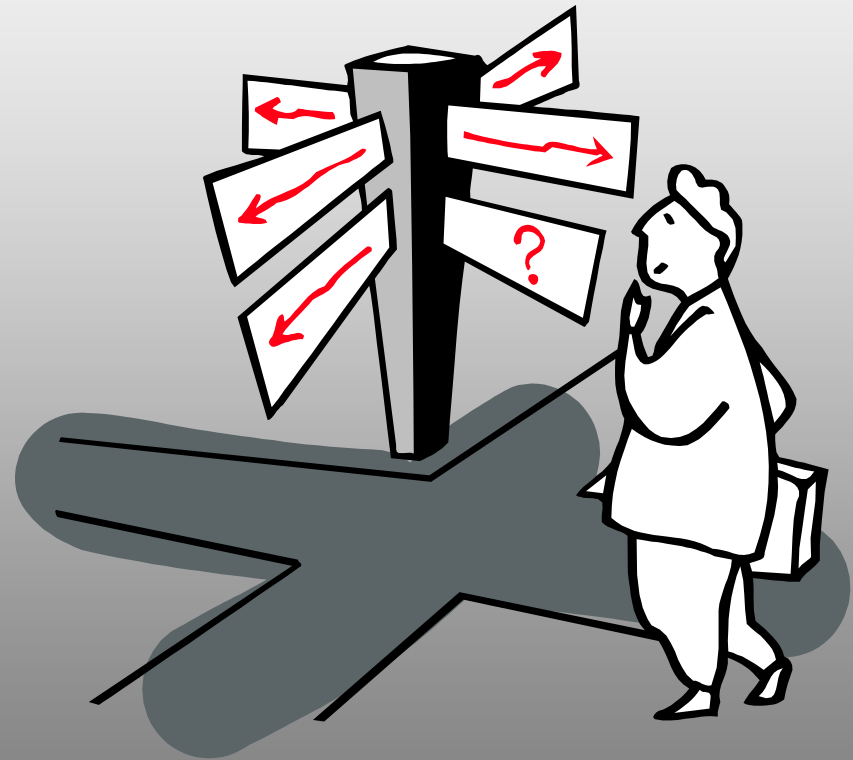
Intervention Research Issues

- Comparison treatments
 - Is the treatment effective?
 - Will the participants become skeptical of treatment?
 - Participants could be offered effective treatment later



Fraud in Science

- Error – honest mistakes
- Fraud – intentional efforts to deceive and misrepresent
- Plagiarism



Authorship (APA, 1992)

- Development of the design
- Writes or prepares portions of the manuscript
- Integrates or brings together theoretical perspectives
- Develops novel conceptual views



Authorship (APA, 1992)

- Designs or develops the measures
- Makes key decisions about the data analyses
- Interprets the results
- A student is usually listed as the principal author on articles resulting from theses or dissertations

What Does Not Constitute Authorship (APA, 1992)

- Collecting or entering data
- Recruiting participants

Communication of Findings and Publication

Communication of Findings



- Description
 - Provide details on study design that allow the evaluation of its adequacy
 - Provide details that would allow replication
- Explanation
 - Rationale for decisions in study (e.g., measures, interventions)
- Contextualization
 - How does this study fit into the context of other studies?

Main Sections of the Research Article

- Title
 - Should succinctly describe the focus of the study
- Abstract (100-120 words)
 - Briefly describe the study
 - Do not include vague statements (e.g., the implications are discussed)

Main Sections of the Research Article

- Introduction (2-4 pp.)
 - Rationale, theory, and objectives
 - Moves from general to specific
 - Convey major issues
 - Do not review literature study-by-study
 - What is the context for this study?
 - Final paragraph should include the purpose of the study and hypotheses
 - Hypotheses should logically follow from the Introduction

Main Sections of the Research Article

- Method(s)
 - Who was studied, why, and how?
 - Rationale for methodological decisions
 - Existing methods are not necessarily useful
 - Do existing methods adequately measure the construct of interest?
 - What are control groups supposed to control?
 - Why were these measures selected?
 - Brief description of each measure
 - Psychometric properties should be included (e.g., reliability, validity)

Main Sections of the Research Article

- Results
 - Begin with descriptive data
 - How do the analyses address the hypotheses of the study?
 - It may be helpful to organize the results into sections with subheadings

Main Sections of the Research Article

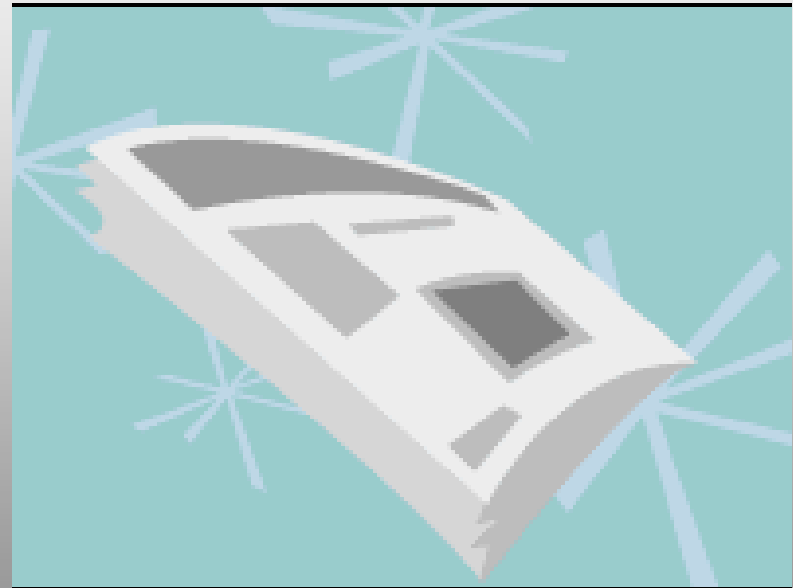
- Discussion
 - Interpretations and conclusions
 - Overview of findings
 - Were the hypotheses supported?
 - Interpretations should correspond with actual methods
 - Integration with prior theory and research
 - Limitations
 - Future directions

Publication Process

- Manuscript prepared in APA Style
- Selection of journal for submission
 - Relevance
 - Prestige
 - Likelihood of acceptance
 - Likely reviewers
- Review process
 - Editor sends manuscript to one or more expert reviewers
 - Substance and methods of the study are reviewed
 - Editor evaluates the manuscript and reviews

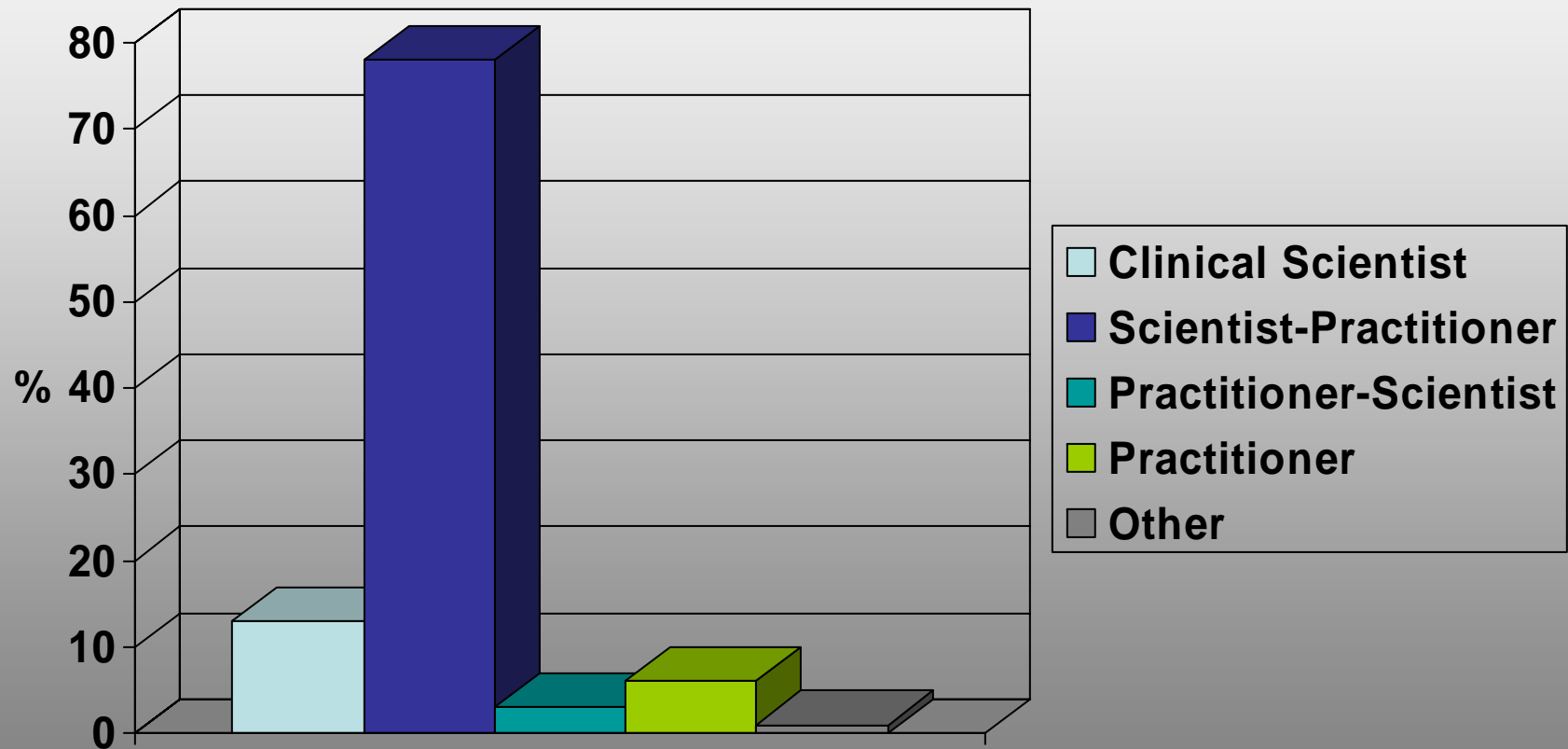
Publication Process

- Editorial decisions
 - Accept pending revisions
 - Reject, but resubmission invited
 - Extensive revision
 - Reject and will not be reconsidered
 - Most common decision (70-90% of submissions)
 - Design flaws
 - Relative importance of the study



Graduate School in Clinical Psychology

Graduate School in Clinical Psychology: 98 Training Programs



The Academy of Psychological Clinical Science <http://psych.arizona.edu/apcs>

- The Academy's broad mission is to advance clinical science. "Clinical science" is defined as a psychological science directed at the promotion of adaptive functioning; at the assessment, understanding, amelioration, and prevention of human problems in behavior, affect, cognition or health; and at the application of knowledge in ways consistent with scientific evidence. The Academy's emphasis on the term "science" underscores its commitment to empirical approaches to evaluating the validity and utility of testable hypotheses and to advancing knowledge by this method.

The Academy of Psychological Clinical Science

- The Academy sees the development and application of clinical science as ongoing and dynamic processes, and is committed to facilitating the evolution of clinical science. Toward that end, it has established five specific goals:
- Training: To foster the training of students for careers in clinical science research, who skillfully will produce and apply scientific knowledge.
- Research and Theory: To advance the full range of clinical science research and theory and their integration with other relevant sciences.

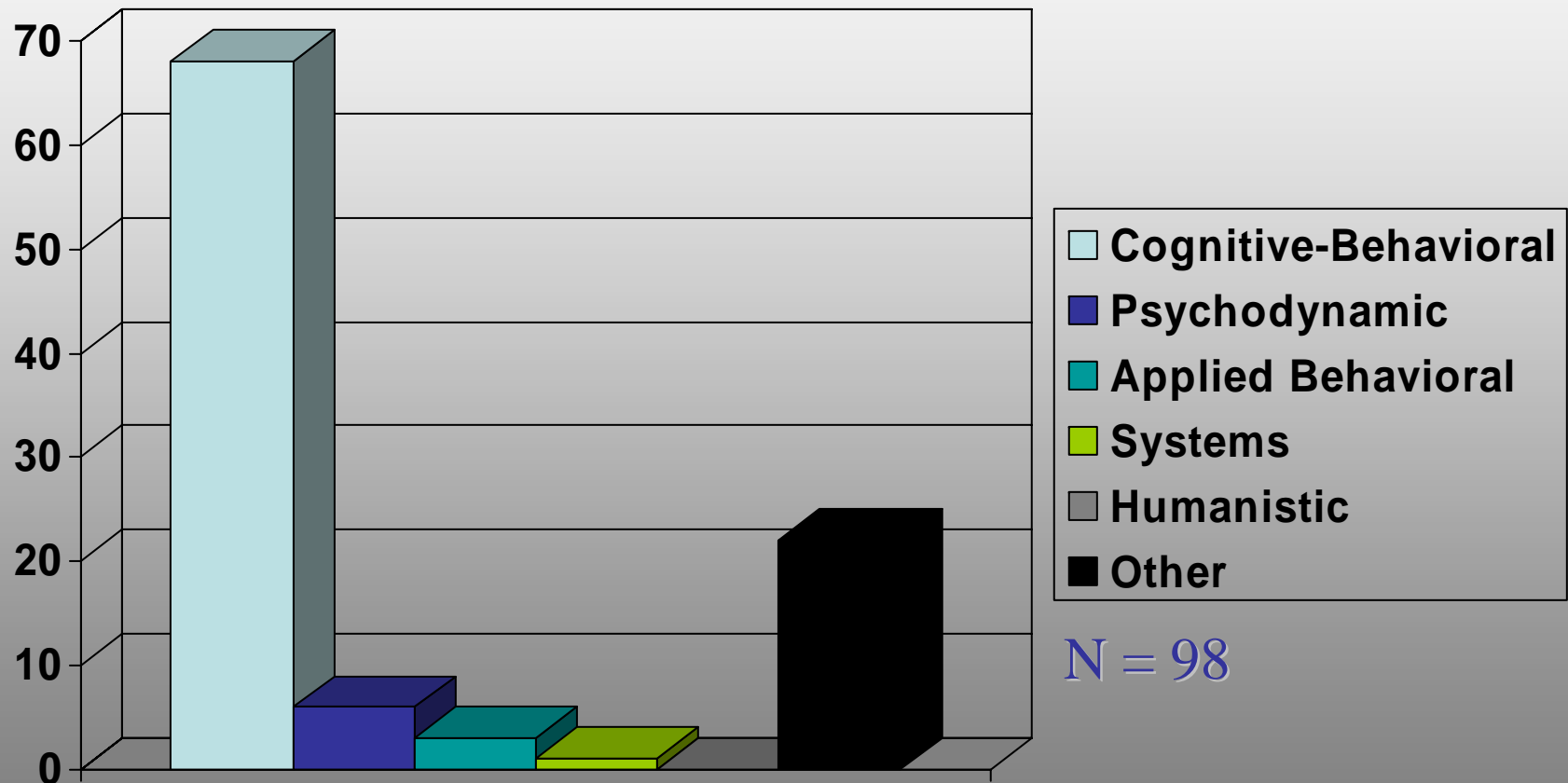
The Academy of Psychological Clinical Science

- Resources and Opportunities: To foster the development of, and access to, resources and opportunities for training, research, funding, and careers in clinical science.
- Application: To foster the broad application of clinical science to human problems in responsible and innovative ways.
- Dissemination: To foster the timely dissemination of clinical science to policy-making groups, psychologists and other scientists, practitioners, and consumers.

Clinical Science Graduate Programs

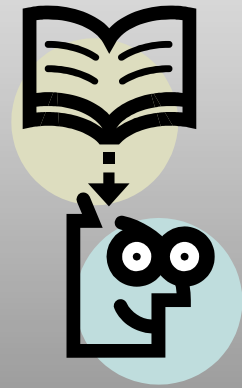
- University of Arizona
- Arizona State University
- Boston University
- University of California- Berkeley
- University of California- Los Angeles
- University of California - San Diego and San Diego State University
- University of Delaware
- University of Denver (Department of Psychology)
- Duke University
- Emory University
- Florida State University
- University of Hawaii
- University of Illinois- Urbana Champaign
- Indiana University
- University of Iowa
- University of Kentucky
- McGill University
- University of Memphis
- University of Miami (Health Psychology)
- University of Minnesota
- University of Missouri
- University of Nevada - Reno
- Ohio State University
- University of Oregon
- University of Pennsylvania
- Pennsylvania State University
- University of Pittsburgh
- Purdue University
- Rutgers University
- San Diego State University and UCSD (joint program)
- University of Southern California
- University of South Florida
- University of Texas
- State University of New York - Binghamton
- State University of New York - Stony Brook
- University of Toronto
- Vanderbilt University
- Virginia Tech
- University of Virginia (Department of Psychology)
- University of Washington
- University of Wisconsin
- Yale University

Graduate School in Clinical Psychology: Theoretical Orientation

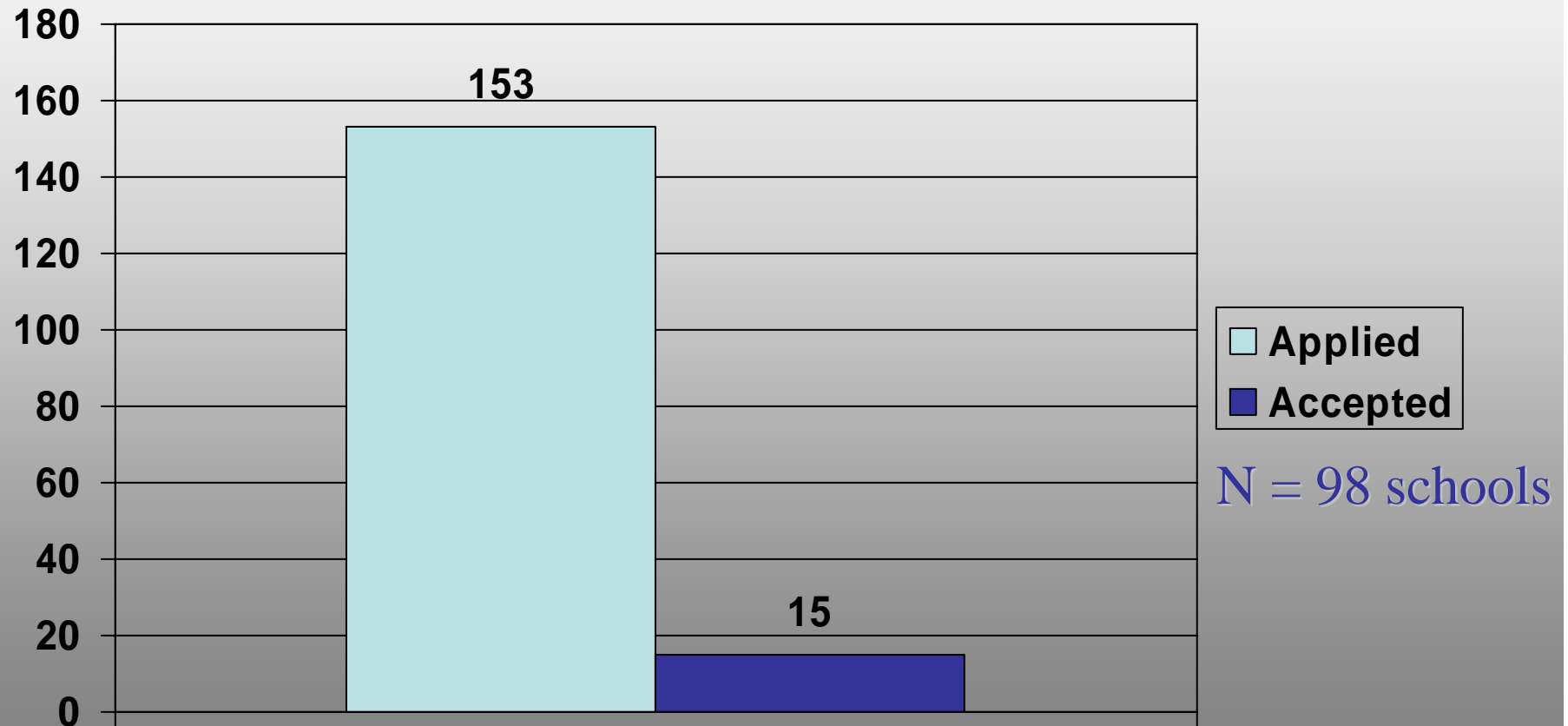


Graduate School in Clinical Psychology: Admissions

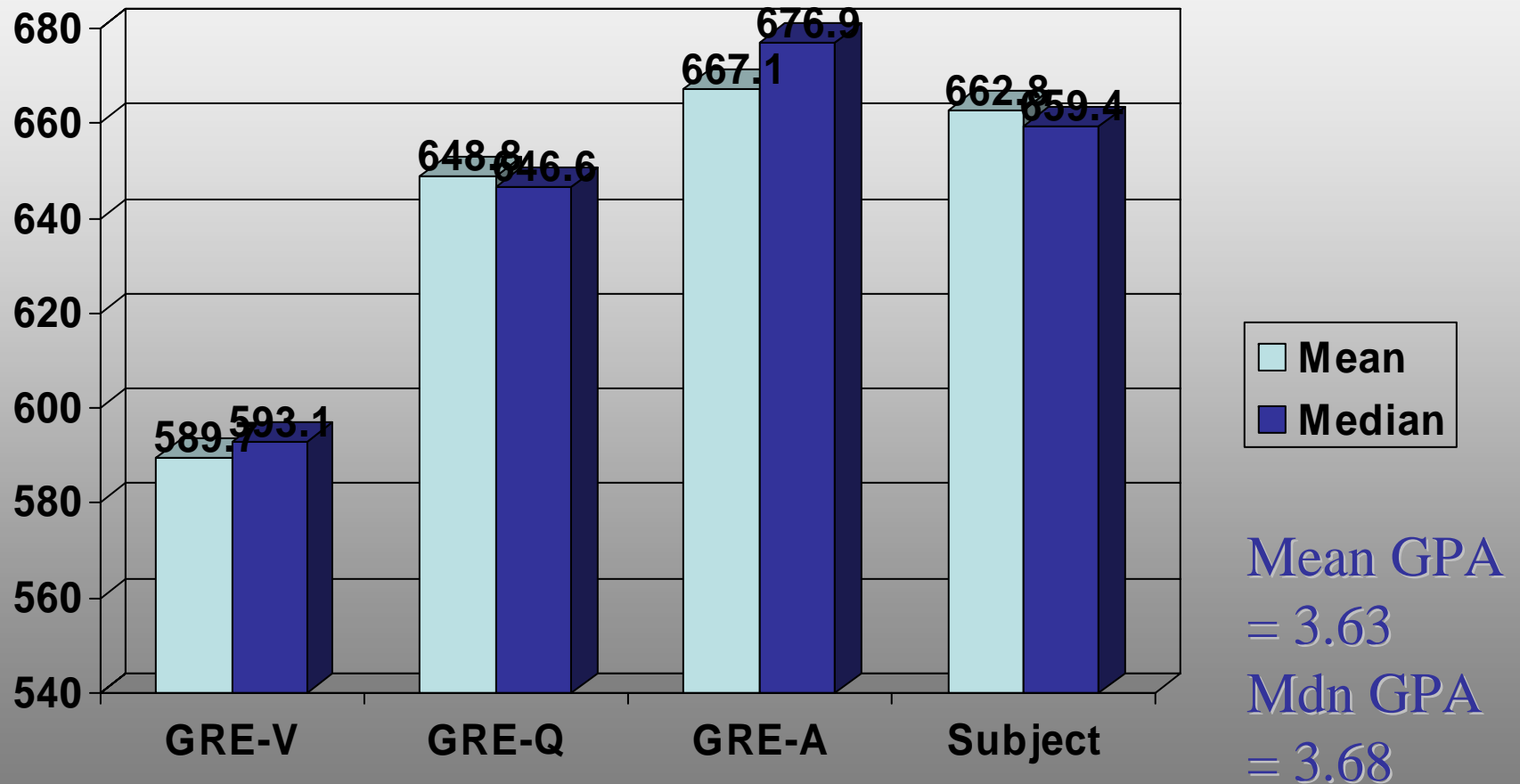
- Admission criteria
 - GRE scores
 - Grades, institution
 - Letters
 - Research experience



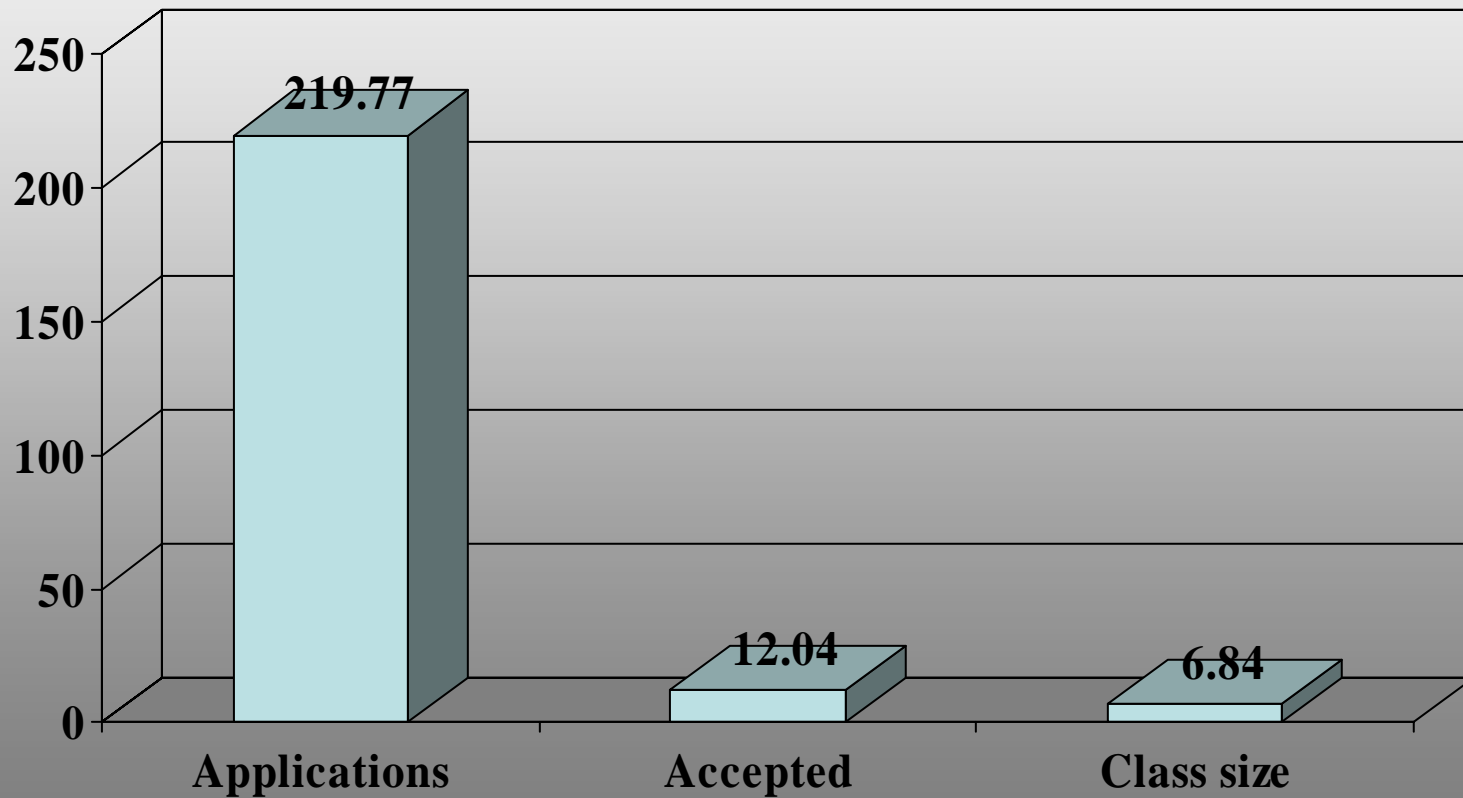
Graduate School in Clinical Psychology: Mean Applications/Acceptances (2000)



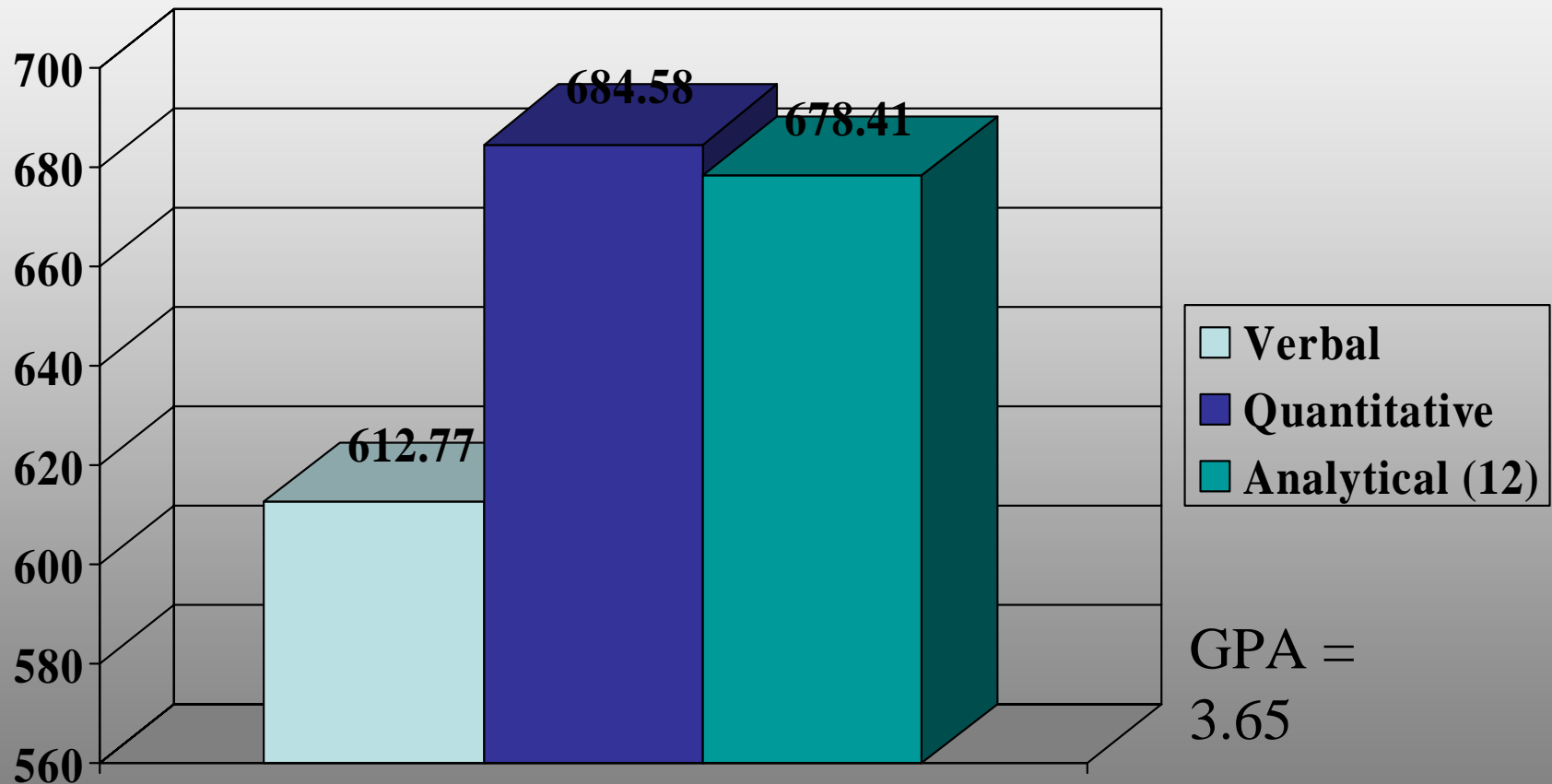
Graduate School in Clinical Psychology: Entry Credentials (2000, N = 98)



Academy Program Means (N = 22)



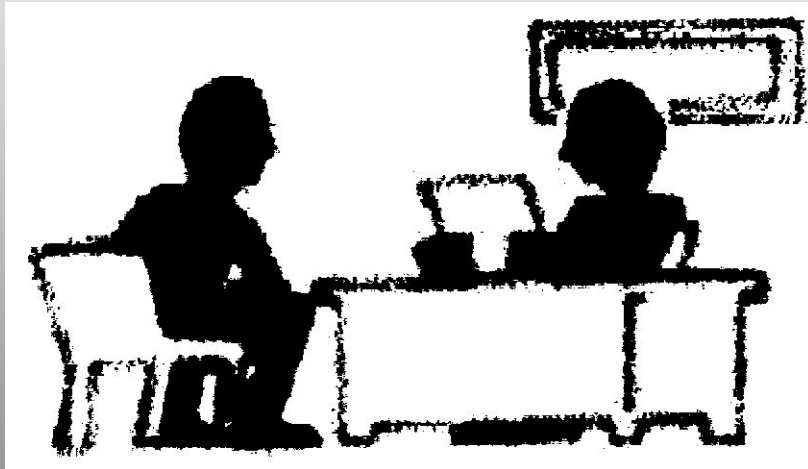
Academy Program GRE Means for Applicants Accepted



Graduate School in Clinical Psychology: Admissions

- Short list
 - Matching interests
- Should I work with a junior faculty member or a more senior one?
 - Junior faculty members may have fewer resources (e.g., funding, experience)
 - But junior faculty may be more “hungry”
 - Senior faculty may have more resources
 - But senior faculty may offer fewer opportunities for innovative work

Graduate School in Clinical Psychology: Interview



- Interview expense is money well spent
 - Some programs provide partial support
- Be prepared
 - Read and know the work of faculty
 - Envision yourself in the setting
- Interviews are a 2-way process

Graduate School in Clinical Psychology: Support

- 98% of students in graduate programs in clinical psychology received financial support during first year
 - Direct support = \$10,042
 - Other support (e.g., tuition, insurance) = \$8,540
 - 73% provide 4+ years of support
- Amount and length of support do not necessarily correspond to the quality of training
 - Short-term vs. long-term gains



Graduate School in Clinical Psychology: Success



- Being in graduate school is not like being an undergraduate
 - Getting good grades and completing required work are not sufficient
- Success = identifying and developing one's own area of research

Professional Organizations

- American Psychological Association
- Association for the Advancement of Behavior Therapy
- American Psychological Society

Relevant Journals

- General
 - American Psychologist
 - Psychological Bulletin
 - Psychological Science
- Clinical
 - Journal of Consulting and Clinical Psychology
 - Journal of Abnormal Psychology
 - Clinical Psychology: Science and Practice
 - Behavior Therapy