Introducing Key Features of the Rorschach Performance Assessment System (RPAS)

Gregory J. Meyer, Donald J. Viglione, Joni L. Mihura, Robert E. Erard, & Philip Erdberg

Society for Personality Assessment

San Jose, CA; March 26, 2010

Why the RPAS?

- For years most of us were on Exner's Rorschach Research Council (RRC) where we generated, gathered, and reviewed data to focus on the continual refinement of the CS
- When John died February 20, 2006, he left the CS to his family, but with no plans for the RRC or any other individuals to continue the CS
- After cordial discussions over several years the Exner family affirmed their original decision that no changes would be made to the CS
  - This would best honor John's legacy

So, why the RPAS?

- We believe that Rorschach-based assessment needs to be able to evolve
  - Especially important with health-care and academic environments that are increasingly evidence-based
- Goal is to continue to solidify the empirical foundation for using the Rorschach in practice
- Focus on its unique contribution to an assessment
  - Provides an in-vivo sample of problem solving behavior
  - Illustrates what people do, not what they say
  - Hence "Performance Assessment"
- Today is a progress report on our group's efforts
  - Aided by the help of many others
  - Mindful of the fear we all share of fracturing what is now a relatively small and unified base of users

A Review of the Validity Research as a Basis for Variable Selection

Joni Mihura, Gregory Meyer, George Bombel, & Nicolae Dumitrascu

Project Process

- Needed to know, for myself, the empirical basis for the Rorschach to continue teaching it
- Initially started with all of the Rorschach empirical literature, but it was overwhelming and reduced to ≥ 1974 CS
- After the Exner family said the CS could not change, the review eventually became part of the variable selection for a new system

Background for Project

- As a Rorschacher, I initially relied on meta-analysis showing Rorschach validity is on par with MMPI validity *(Hiller et al., 1999)*
  - MMPI: *r* = .30 and Rorschach: *r* = .29
- However, there were no conclusions for individual scales nor were most CS variables included in this meta-analysis

A Closer Look at Hiller et al.'s Meta-analysis

- Total of 34 Rorschach studies
  - But only 14 studies included CS variables
  - On the next slide, the CS variables included in the 14 studies were…
    - Non-CS variables in red
    - Effect size per study in parentheses (*r*)

1. Afr, Lambda, Grandiosity, Splitting (.15)
2. X+%, X-%, Xu% (.43)
3. P, F+% (.65)
4. P, F+%, Sum6 (.08)
5. FQ+, FQo, FQu, FQ-, X-% (.25)
6. X+%, X-% (.44)
7. X-%, W (.43)
8. m (.47)
9. DL, DQ (.39)
10. X+% (.35)
11. Egocentricity Index (-.05)
12. Space (.00)
13. Reflections (.06)
14. WSum6, X-%, DEPI, F+%, D (.91)

CS Variables in Hiller et al.

- This oft-cited meta-analysis supporting the Rorschach only included 15 of the 69 variables in the lower Portion of Structural Summary
- And some did not have good support
  - Space (*r* = .00)
  - Egocentricity (*r* = -.05)
  - Reflections (*r* = .06)

Purpose of Present Project

- An initial attempt to organize the Rorschach CS validity literature for the individual variables in the lower portion of the Structural Summary using a systematic approach to the review
- Differs from review in Vol. 1:
  - Systematic review
  - No unpublished studies
  - Generally casts a broader net of published literature:
    - For example, Vol. 1 reports data from only 5 of the 48 studies for the Egocentricity Index reported here.
  - Effect sizes reported across studies using the same metric (Pearson's *r*)

Search Strategy

- Databases: PsycINFO+MEDLINE
- Keyword: Rorschach
- Limit: English language; Article; Year ≥ 1974
- Final: 2,276 abstracts
  - 69 variables in lower Structural Summary
  - CS Scoring
  - Explicit or implicit hypotheses
  - Samples compared to CS norms recomputed to compare to int'l reference sample
    - Makes a big difference for reality testing variables

Final Product

- 233 CS studies included
  - Hiller et al.'s meta-analysis = 14 CS studies
- 2,565 findings
  - Average = 11 per study
- Reliability of including articles
  - % agreement = 98%; ICC = .90

Study Findings

- Wide range of support per variable
- De-emphasize self-report
- Strength of evidence ratings conducted
  - Good support
  - Some support
  - Not enough good research or inconclusive
  - Existing research does not support the scale
- Will show summary of 69 variables' ratings
  - Also meta-analytic summary for 37 comparisons
    - e.g., Relation of CDI to depression across 4 studies

PTI (Good support)

- Summary effect sizes for psychotic diagnosis *(ES are Weighted [Wgt] r's per study)*
  - Psychotic dx vs. Nonpatients = .76; N=160
  - Psychotic pts vs. Nonpsychotic pts = .57; N=345
  - Schizophrenia vs. Patients including psychosis = .23; N=149
- Plus many other findings

EA (Good support)

RPAS

Related to other measures in ways that would be expected to suggest psychological resources

- Asperger's < Other psychological problems
- ADHD < Controls
  - Positively related to
    - IQ
    - Executive functioning
    - Education
    - Dynamic capacity (observer ratings)
    - Ego strength (observer ratings)
    - Being selected for insight therapy

SumC' (Some support)

- No relation to depression in five studies (Wgt *r* = .07; N=419)
- But related to events that could result in unexpressed negative emotion
  - Children who had been sexually abused
  - Children whose parents had divorced
- Related to fMRI measure of negative mood during negative feedback

M- (Some support)

- Inconsistent relationship to psychotic diagnoses (Wgt *r* = .11; N=939)
- Better for interpersonal disturbances (Wgt *r* = .40; N=194)
  - Psychopathic > nonpsychopathic violent offender
  - Violent offense > student
  - Pedophile > non-sexual offense
  - Cluster A + Borderline PD > International Reference Sample

Space response
(Existing research does not support the scale)

- Not related to:
  - Self-reported anger or cynicism
  - Observer ratings of aggression
  - TAT Hostility Scale
- No good CS studies of oppositionality
- Instead, *positively* related to creativity (verbal more than figural) and the Rey organization score

Egocentricity Index
(Existing research does not support the scale)

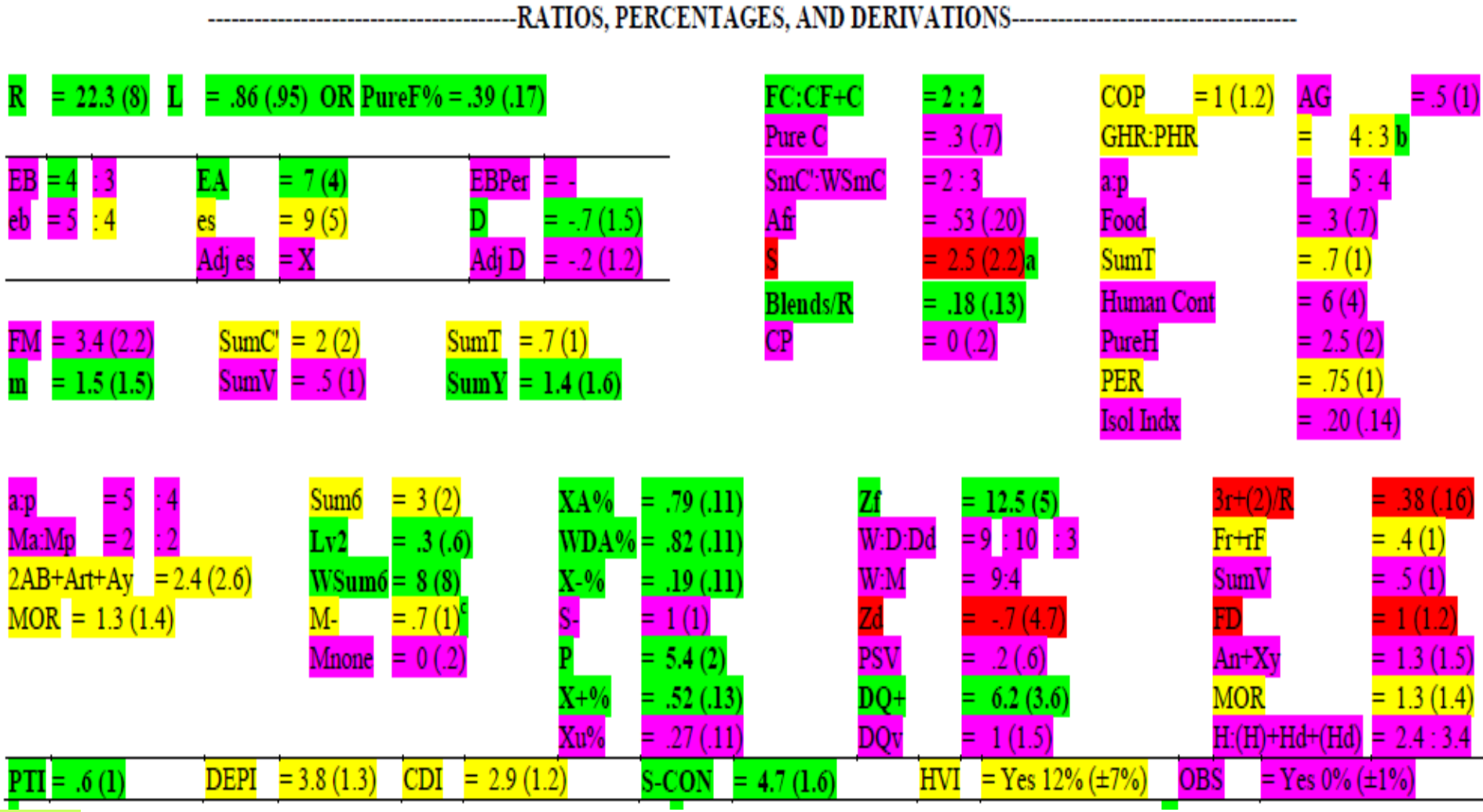
- Not related to depressive diagnosis in 5 studies (Wgt *r* = -.01; N=518)
- Inconsistently related to psychopathy
- Higher for narcissists in 2 of 3 studies and lower for suicidal patients, so maybe it only works at the extremes?

Egocentricity Index: Compare to Vol. 1 Review

- Overlap: Cited articles (5/48); findings (2/108)
- Vol. 1 reported findings from 6 other articles but
  - 1 finding not hypothesized so not in my review
  - Vol. 1 reports significant findings for children and adolescents, but article did not find significant for adolescents (nor females)
  - In Vol. 1, 1 citation included significant post-hoc findings but not the non-significant hypothesized finding (the ns finding is in my review)
  - 3 articles in Vol. 1 did not contain the cited data
- Of the 2 findings that overlapped, for the Vol. 1 citations:
  - 1 MMPI-2 finding not replicated in 2 other studies
  - 1 self-report self-esteem finding not replicated in $\geq$ 3 other studies

Overview of Findings

Good Some None/Mixed Negative



Conclusions

- CS variable constructs with the most support
  - Reality testing and thinking disturbance
  - Complexity and ego strength variables
- CS variable constructs with least support
  - Self and other variables; affective experience
- Nature of CS variables with most support
  - Closest link between the response process and the interpretation

Examples of reasons why the response process fits interpretation but there is low validity in the literature …

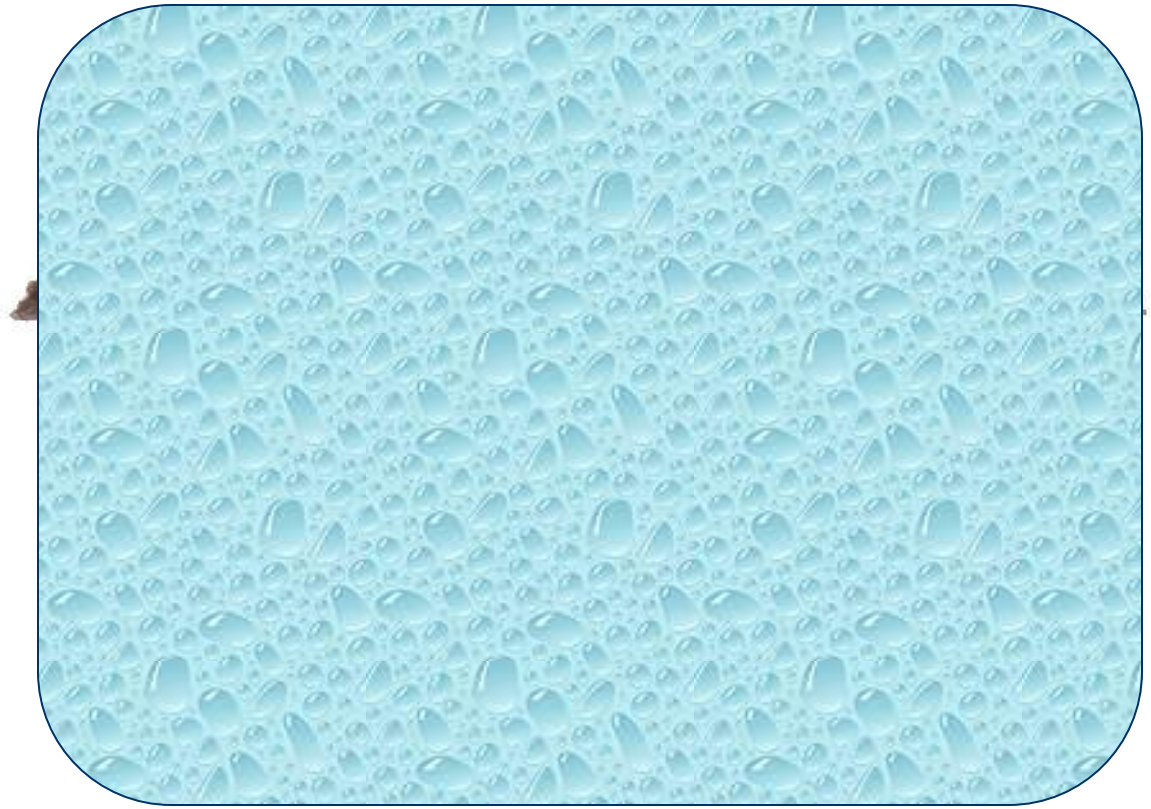
- Space response
  - Previous systems distinguished between
    - Primary/main and secondary/additional
    - Primary Space responses involve reversal of figure-ground
    - Secondary space is when Space is integrated with other blot features (and gets Z score)
  - CS Vol. 1 includes no studies examining *CS-scored* space responses, most recent study = 1955

Space Reversal = Figure Ground Reversal

Card VII: “Upside down the white part looks like a lamp”

Space Reversal vs. Integrated

Card I: “A fox; here’s his eyes and mouth” (Integrated)
Card I: “4 ghosts dancing in the darkness” (Reversal & Integrated)



Space Reversal

- Response process fits interpretation
- Good pre-CS research support for Space reversal oppositional interpretation, not as much for anger
- Space integration responses 4X more than Space reversal *(Dumitrascu, Mihura, Meyer, 2010)*
  - Likely why CS findings are negative
  - Match response process with interpretation better
- ➢ Research suggests modification of CS Space scoring

Contributions to RPAS Variable Selection

- Levels of inclusion
  - Page 1 = RPAS emphasis
  - Page 2 = Hypothesis generating
  - “Page 0” = Not included
- Levels of contribution
  - Strong: Supported in validity review
  - Some: Link between response process and interpretation
  - Some: Clinician survey findings
  - Limit redundancy

Clinician Survey

- Asked experienced Rorschach users what works given their clinical experience
  - Input from more than 200 experienced clinicians
  - Pairwise agreement was limited
    - *M r* = .08
  - But the aggregated judgments of the clinicians were strongly correlated with validity review findings
    - *r* = .39
- So clinician ratings provided another source of guidance

Final selection of RPAS variables

- Also included other non-CS variables with good validity support
  - ROD, MOA, AgC and AgPast, TC/R
- ➢ Will see a sample of the final(ish) selection of variables in later presentations in this symposium

Research on the Range of R as a Foundation for R-Optimized Administration

Donald J. Viglione

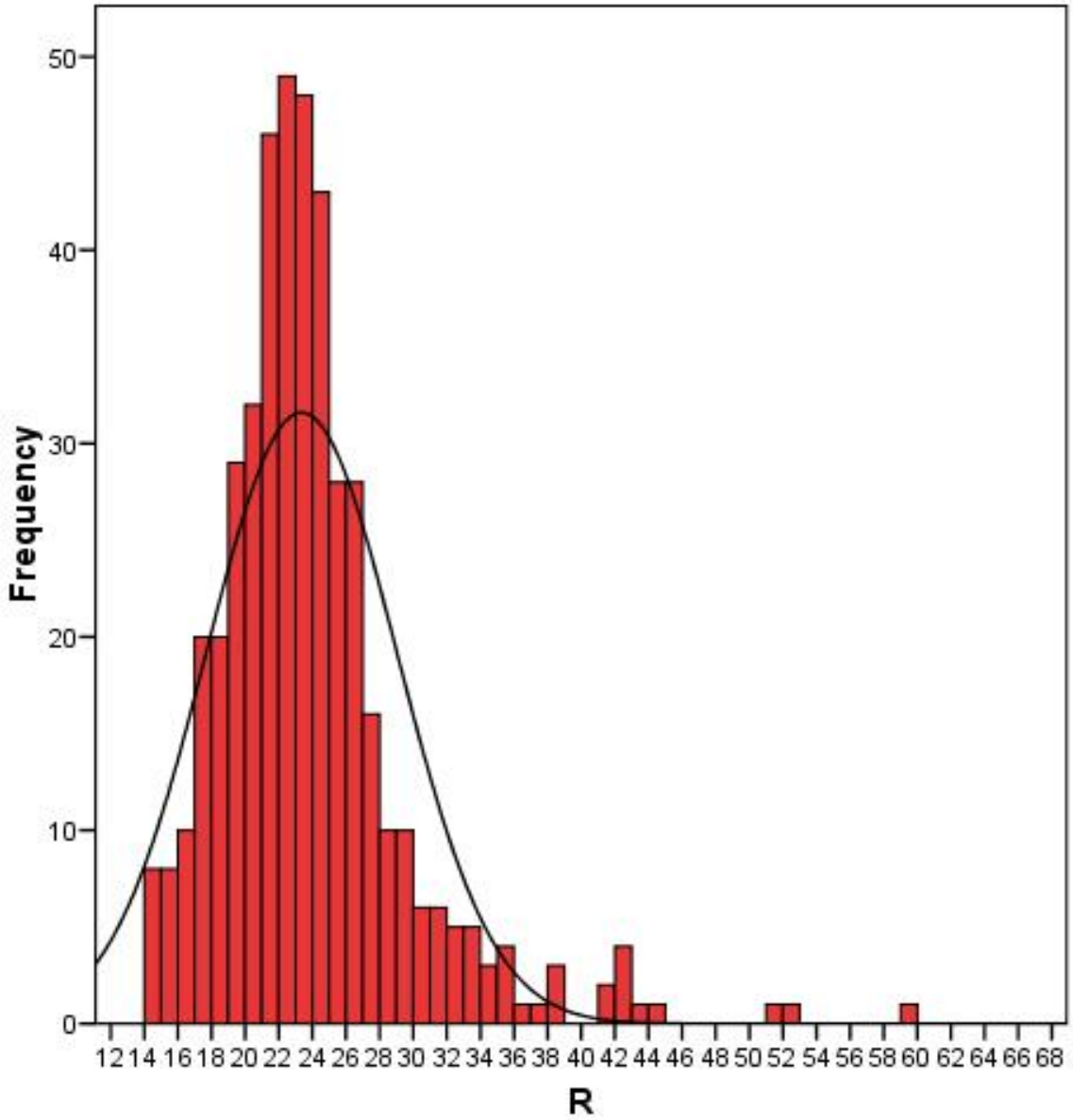
SPA – San Jose

March 26, 2010

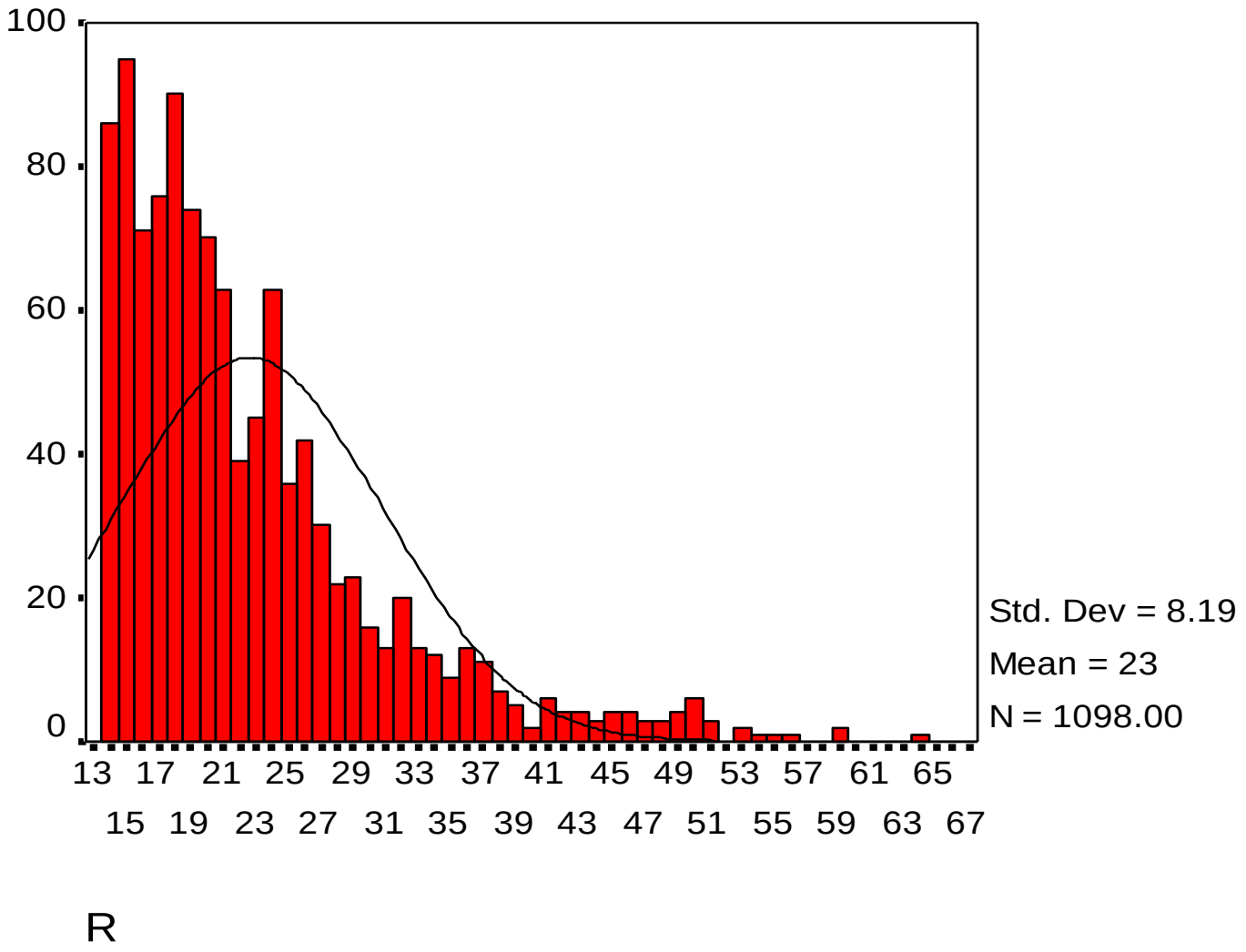
Problems with R

- Evidence suggests R varies considerably across records with many variables correlated with R (Meyer & Viglione, 2008; Viglione & Hilsenroth 2001; Viglione & Meyer 2008)
- R is a nuisance variable or confound -- increases error variance and reduces interpretive accuracy and research progress
- 1974, Vol. I, Exner rejected control on R, getting less variation on R than some other researchers

Distribution of R in the CS 450



Distribution of R in a Subgroup of the International Samples; N = 1098



Research vs. CS estimates of R

CS Original N = 600, Exner (2003)		Exner (2007); N = 450	
M	SD	M	SD
22.67	4.23	23.36	5.68
International Norms (Meyer et al., 2007)		Shaffer et al. (2007), N = 248	
M	SD	M	SD
22.31	7.91	20.48	7.19

Examples of Impact of Corr. with R

Corr. with R	Selected Variable	Low R n = 493	Optimal R n = 619	High R n = 230
		R=14 – 17 Mean = 15.4	R = 18 - 27 Mean = 21.7	R > 27 Mean = 35.1
Very Strong (>.6)	Dd	1.6	2.8	7.1
Strong (>.5)	es	6.1	8.9	14.5
Moderate (>.4)	S	1.7	2.7	4.7
Weak (>.3)	Y	1.1	1.6	2.7
Minimal (>.15)	GHR-PHR	0.6	0.4	-1.3
Virtually = 0	CDI	3.2	2.9	2.8

Interpretive Issues
Short Records -- R < 18

- Likely underestimate problems?
- Less data = conclusions more tentative
- Risk of missing salient personality characteristics
- General factor suppressing all scores?
  - Lack of task engagement, cooperation
- = Utility & Cost-Benefit:
  - Less interpretive yield per record

Long Records – R > 27 or > 35?

- Interpretive yield good -- learn about person *but overestimate problems*?
- Utility & Cost-Benefit: Great cost in scoring, administration confusion, interpretive uncertainty, effort, & time

Other Issues with R

- Managing R is Confusing
  - Re-Administration Procedure when R < 14
  - Card IV, R less than 5 -- intervene
- Problems with re-admin records - vignette
  - responses get mixed up
  - irritation/confusion
  - some studies have modes = 14

Dean, Viglione, Perry, & Meyer, 2007 First Study to Restrain Variation of R

- Endorsed by Rorschach Research Council
- Used preliminary version of R-Optimized ("Alternative")
- Prompted if 1 R on a card up to 3 times, except on Cards V and IX
- 61 inpatient offenders in a forensic psychiatric facility
- Administered Rorschach
  - predictors -- EII-2, PTI, SCZI
- Psychosis/Thought Disorder Criterion Combined --
  - (1) Anderson TLC
  - (2) Chapman Magical Ideation Scale
  - (3) Delusion/Hallucination from SADS

Table 4

Sample Specific Correlations between Rorschach Predictors and the Thought Disorder Summary Scale (TDSz) and the Magnitude of Difference

Scale	All (N = 61)	Alternative (n = 31)	Standard (n = 30)	Difference z
EII-2	.47***	.56**	.38*	0.86
SCZI	.40**	.53**	.26	1.20
PTI	.38**	.44**	.33*	0.48

Note. Alternative = Alternative administration group;
Standard = Standard administration group;
EII-2 = Ego Impairment Index with outlier changed;
SCZI = Schizophrenia Index;
PTI = Perceptual Thinking Index.
* p < .05, ** p <.01, *** p < .001

Dean, Viglione, Perry & Meyer, 2007

Goals of RPAS altered administration procedures

- Utility
  - To provide a simple and easy to remember system for managing R.
    - Reduce demand on Examiners
    - Easier to remember and learn
  - To eliminate exceedingly long testing sessions

- Increase reliability to increase validity
  - To increase the probability of optimal length records
  - R = 18 - 27?
  - Best match for normative data – “sweet spot”
  - Less error introduced by R

2nd Study
"Alternative Administration"
Viglione, Meyer, Converse, & Jordan
(under review)

- Structured to procure Two to Four responses per card.

- Card I – X with only 1 R -- "***Prompt***"
  - Encourage S with a ***prompt*** like one that might be commonly used on Card I on a CS administration.

- Card I – X after 4 R's -- "***Pull***"
  - Goal to eliminate extremely long records
  - Subject is requested to return the card
    - e.g., "OK, let's do the next one." Or, "Thanks! That's all I need on this one."
    - Don't really "pull" the card
  - To preserve rapport, need not take card back after 4th, should be written down but not scored.

Alternative Admin Study #2

- Purpose – To examine how prompts and pull affect productivity

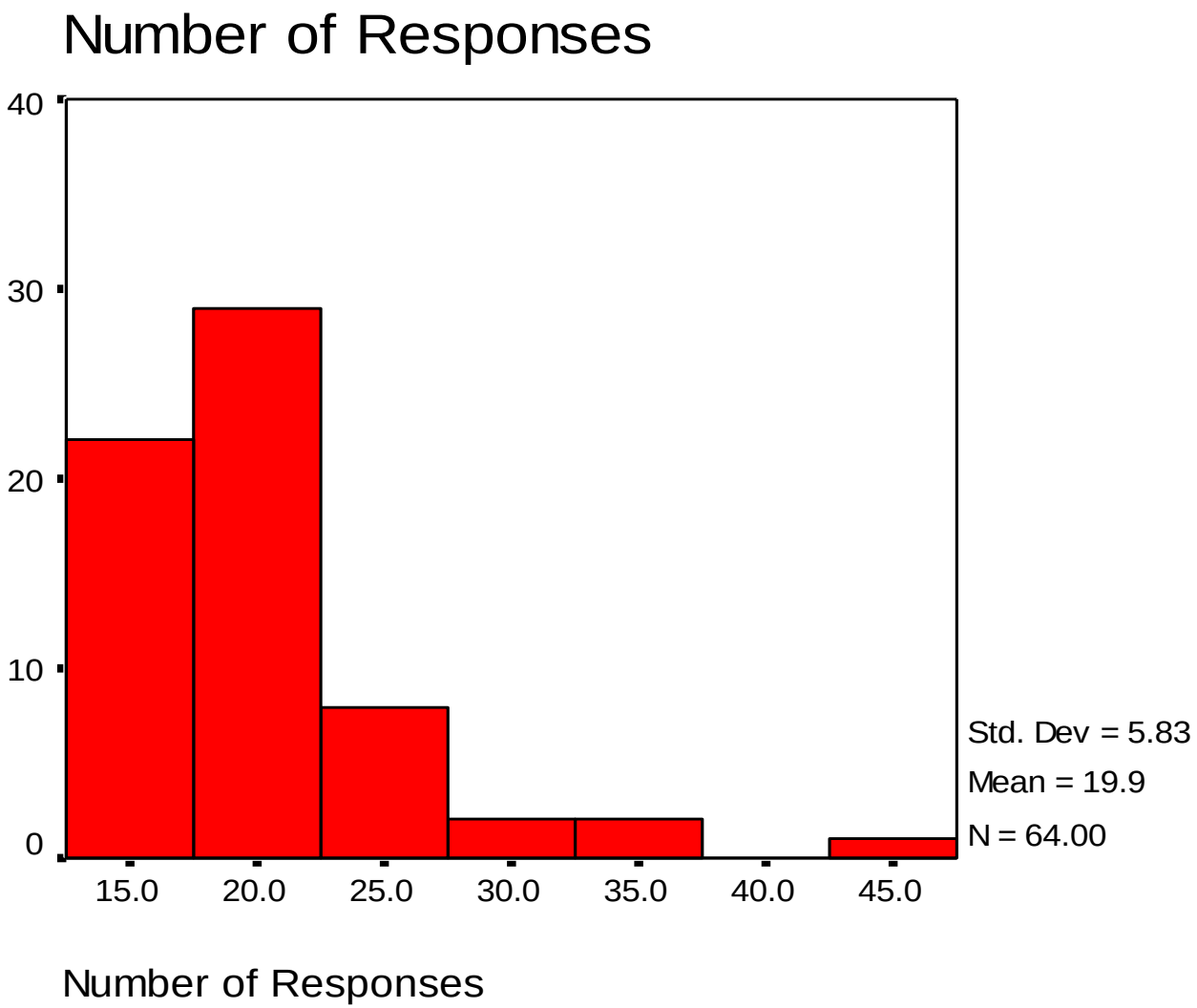
- With Alternative Admin
- Expected fewer short records and more records in optimal interpretive range
  - 18 - 27 R (less certain re cutoff for high range)

Comparison Student Examiner Sample AIU and Loyola Baltimore

Viglione, Rivera, Saltman (submitted)

Homework = Short Records

=Examiner Variability



Method

- About 40 Examiners - 2nd Year graduate students or higher enrolled in personality assessment course.

- Mean number of previous admins for E's was 3.2
  - no difference between groups

Method

- Subjects were volunteer adults (& some children) recruited by students for purposes of practicing Rorschach administration and coding.

- All records that deviated from proper administration guidelines were excluded from the sample (i.e. forgetting to prompt on R-Optimized, prompting after Card 1 on CS, etc.)

Method

N = 116 Rorschach protocols

- n = 54 CS
- n = 62 R-Optimized

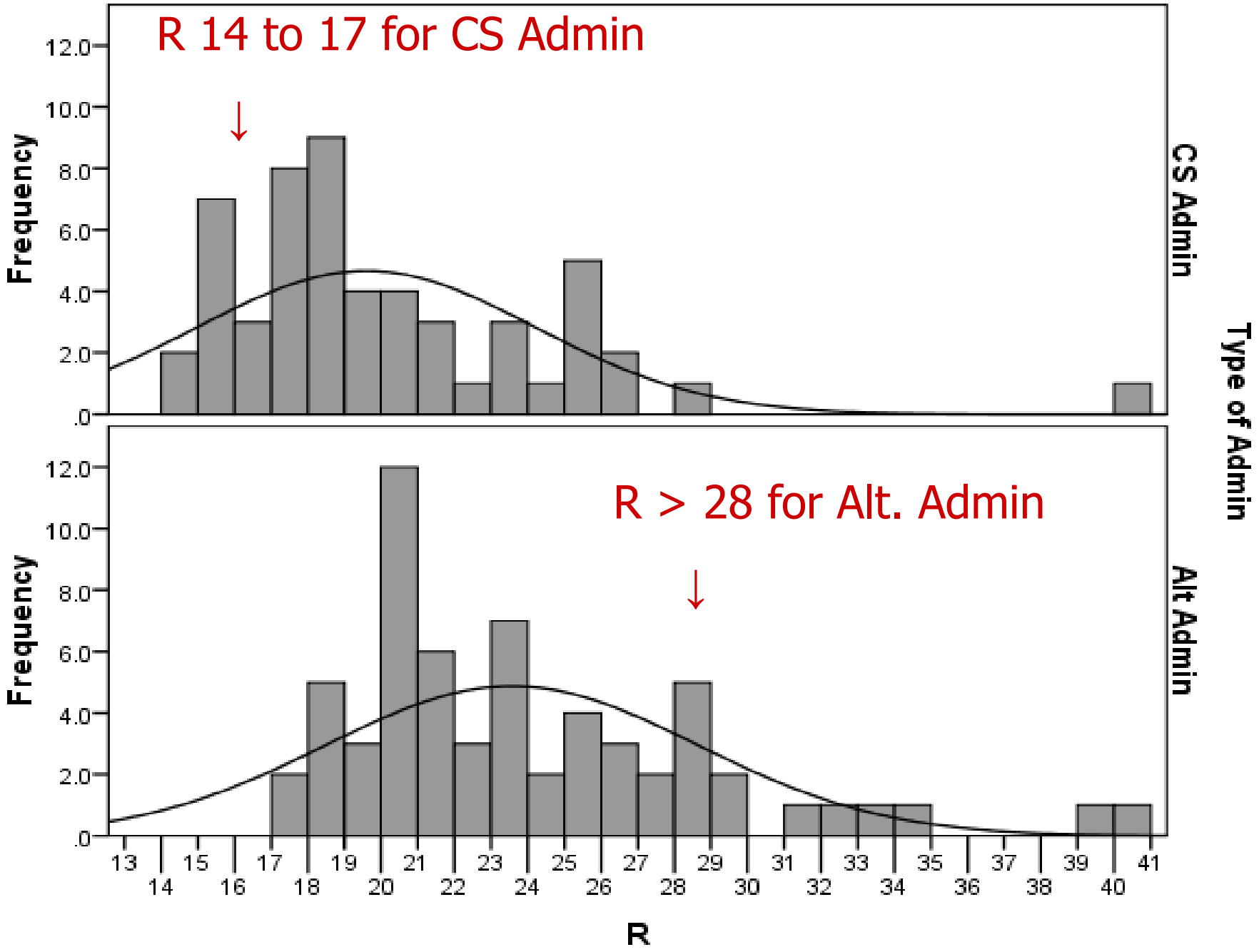
Design

- Between Group Contrast

- Examiners were randomly assigned to either a CS or R-Optimized Administration and told to alternate admin style in subsequent records

R and Form%/Lambda—Complexity & Engagement

	M	*SD*	Med
Alternative (n=62)			
R	**23.56**	5.03	22.5
Lambda	0.99	0.85	0.82
Form%	0.43	0.18	0.45
CS (n=54)			
R	**19.59**	4.60	18.0
Lambda	1.11	2.16	0.64
Form%	0.41	0.18	0.39


R 14 to 17 for CS Admin
R > 28 for Alt. Admin
Frequency
12.0
10.0
8.0
6.0
4.0
2.0
.0
CS Admin
Alt Admin
Type of Admin
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
R

Optimal Range of R?
Yellow "sweet spot"

R Range		CS	Alt Admin
Low	#	9	2
R < 18	%	37%	3%
Optimal	#	26	35
17<R< 28	%	59%	77%
High	#	1	9
> 27	%	4%	21%

p < .005

Prompts, Pulls, Card with 1R

	Admin	M	SD	Min	Max
Total Prompts	Alt.	2.10	1.89	0	8
	CS	0.43	0.50	0	1
Total Pulls	Alt.	0.61	1.81	0	10
	CS	0	0	0	0
Cards with one R	Alt.	0.47	0.88	0	4
	CS	2.72	1.85	0	7

15 of 115 Variables Significantly Different (Notable Variables)

Variable	Admin	*M*	*SD*	Cohen's d
D more	Alt.	9.40	5.04	0.66
	CS	6.20	4.76	
DQo more	Alt.	15.58	5.72	0.71
	CS	11.74	5.09	
F more	Alt.	10.52	5.47	0.47
	CS	8.20	4.28	
A more	Alt.	8.03	3.44	0.43
	CS	6.72	2.64	
Complexity/R less	Alt.	3.28	0.81	0.40
		3.67	1.10	
CF more	Alt.	1.71	1.41	0.47
	CS	1.11	1.14	
FC+CF+C+Cn	Alt.	4.39	2.47	0.51
more	CS	3.24	1.98	
WSumC more	Alt.	3.33	1.87	0.48
	CS	2.48	1.68	
An more	Alt.	1.21	1.51	0.59
	CS	0.48	0.48	
Xy more	Alt.	0.11	0.32	0.38
	CS	0.02	0.14	
Critical Content	Alt.	5.52	4.23	0.38
more	CS	4.07	3.40	

Conclusion Alternative Admin Study

- R-Optimized - successful in eliminating short, non-useful records
- --would lead to better reliability and utility
- Does produce some longer records
  - But no torturous long records, never > 40
  - Problem addressed later
- Range of R – more similar to CS expectations
- *Potential for less variation across sites, countries, more uniformity, less Examiner variation*

Altered Administration Study #3: The R-Optimized Method of Administration

Jenny Evans; Devon MacDermott, Donald Viglione, Greg Meyer, to be presented at this meeting

The "R-Optimized" Method of Administration

- *Devised to lower top end, fewer records with 28 and more R's*
- *Ask for "two,.. maybe three responses"*
- *Kept*
  - *Prompt for 2*
  - *Pull after 4*

Response Frequency for R-Opt vs CS

Response Frequency for CS vs. R-Opt

	<18	18 ≤ R ≤ 27	>27
R-Opt	2	38	4
CS	25	35	4

R and Form% for CS vs. R-Opt

Score	Admin.	*M*	*SD*	*Mdn*
R	R-Opt	22.14	3.49	21.00
	CS	19.61	4.53	18.00
Form% **(Lambda)**	R-Opt	0.42	0.19	0.37
	CS	0.41	0.20	0.38

Prompts and Pulls

	Admin	M	SD
# of Prompts	R-Opt	.93	1.516
	CS	.17	.420
# of Pulls	R-Opt	.14	.632
	CS	.02	.125

Only 3 of 55 RPAS Variables Differ ($p < .05$)

Less than expected by chance (but also limited power)

Variable	Admin.	M	SD	Cohen's d
Dd	ROpt	4.55	2.85	0.58
	CS	2.92	2.77	
X-%	ROpt	0.23	0.12	0.36
	CS	0.19	0.10	
X+%	ROpt	0.48	0.13	0.62
	CS	0.57	0.16	

Conclusion R-Optimized Study

- R-Optimized - successful in eliminating short & long non-useful records
- Very little effect on variables—impact shared across all?
  - should produce better reliability and utility
- Range of R – more similar to CS expectations
- With smaller SD, which turn should reduce any possible error effect associated with variability in R
- *Potential for less Examiner Variation across sites, countries, more uniformity*
- *Distribution limits effect of R on data and suggests we should adopt R-Opt to RPAS*

Developing Norms and Standard Scores for Interpretation

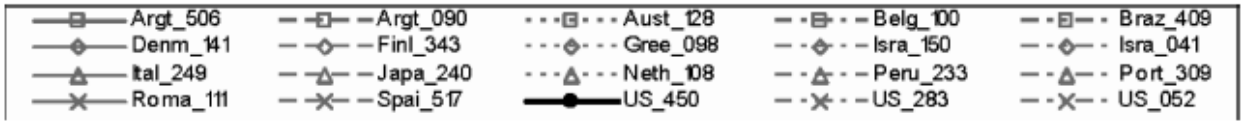
Gregory J. Meyer

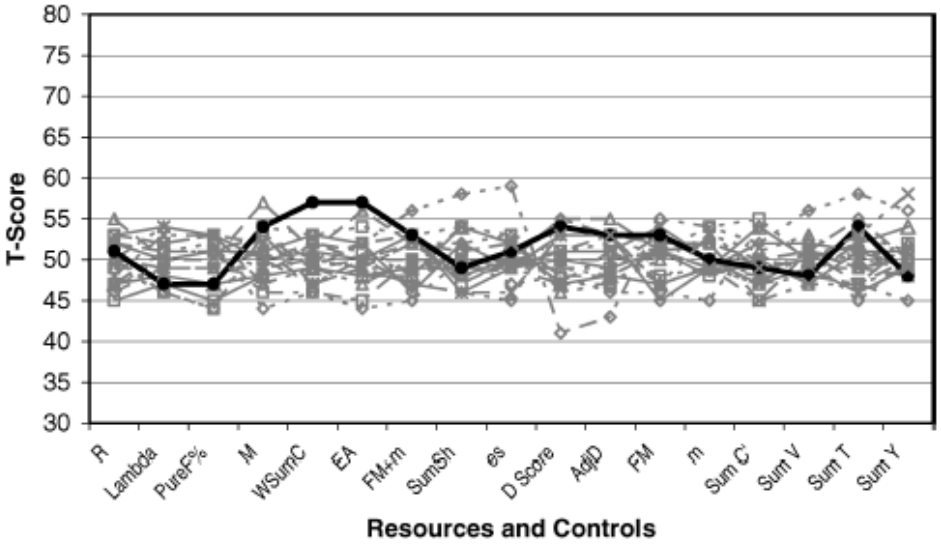
CS International Reference Samples Project

- December 2007, *JPA Special Supplement*
  - Shaffer, Erdberg, Meyer et al. (220 pp)
  - 39 Samples
    - 20 from adults
    - 19 from children and adolescents
  - Countries included:
    - Argentina, Australia, Belgium, Brazil, Denmark, Finland, Greece, Israel, Italy, Japan, The Netherlands, Peru, Portugal, Romania, Spain, and the US
    - Adult French data (Sultan et al., 2004) included also

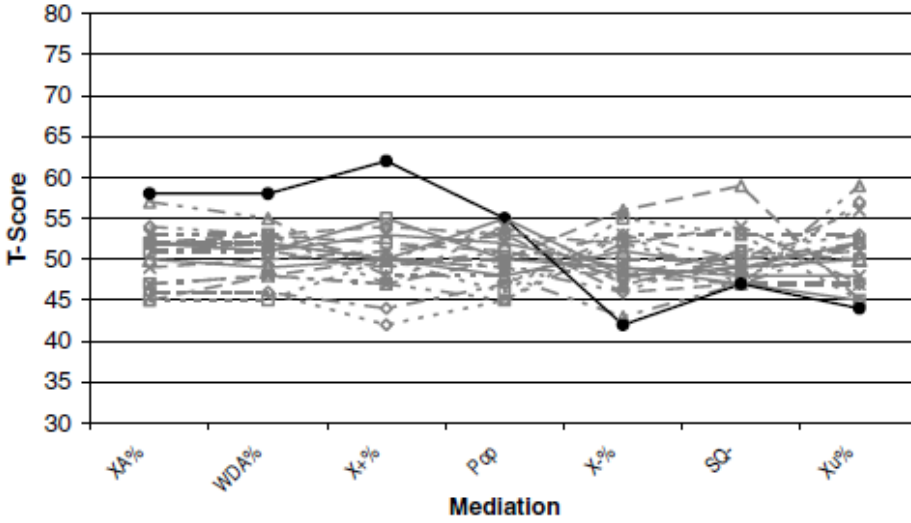
CS International Reference Samples Project

- International samples have great variability in
  - Recruitment strategies
  - Participant selection
  - Language
  - Examiner training
  - Examination context
  - Culture
- With consistent data, they thus provide great generalizability across these dimensions
- We created Composite International Norms
  - Pooled the *M*s and *SD*s to create T-Scores
    - Ideally would use percentile transformations
  - Example T-Score graphs on next slide

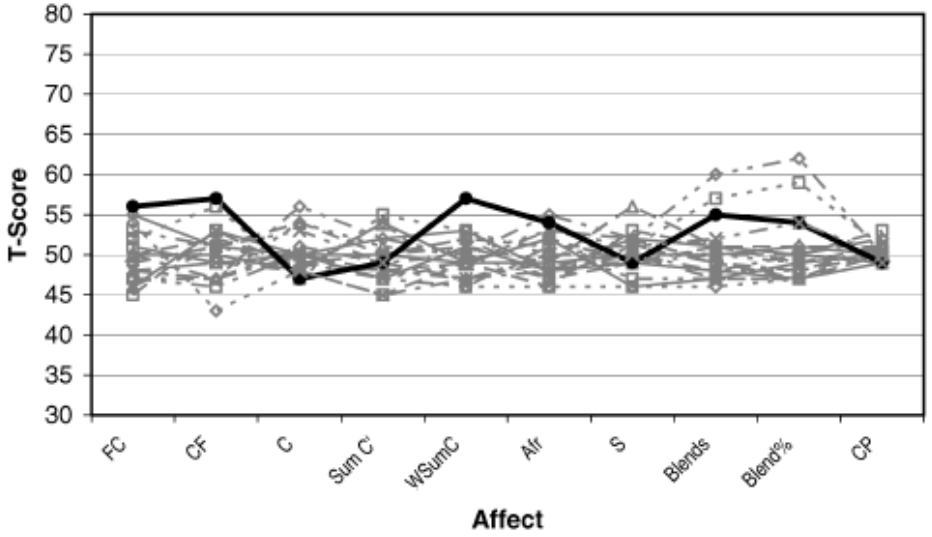




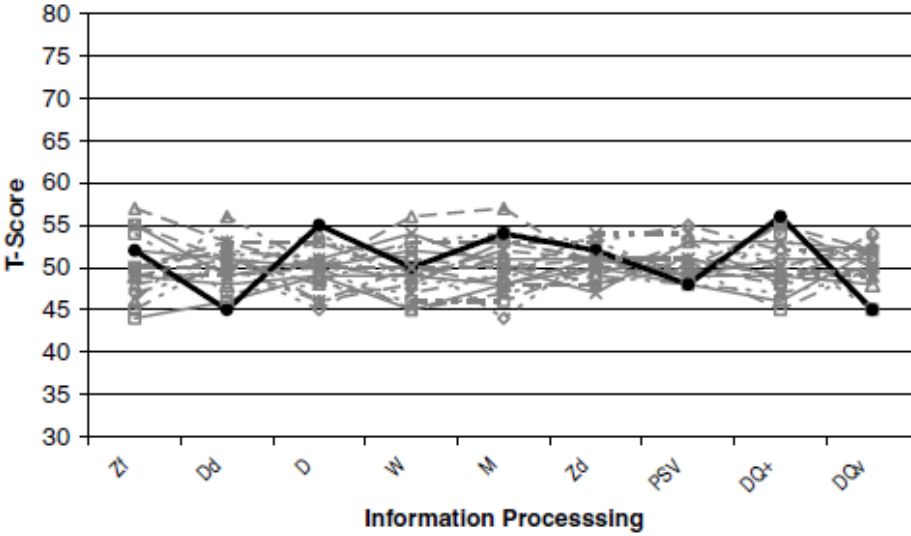
(A)



(E)



(B)



(F)

Summary: Adult Samples

- Basic consistency of scores
  - Across cultures, languages, examiners, exclusion criteria, and recruitment strategies
  - i.e., People look pretty similar overall
- The general consistency supports combining scores

The Moderating Role of Quality?

- But there are a number of reasons why one might have concerns about adopting the normative data from international reference samples
  - e.g., they combine across all the features mentioned before
    - Recruitment strategies
    - Participant selection
    - Language
    - Examiner training
    - Examination context
    - Culture
- Combining data over all samples may inappropriately mix high quality information with information of less optimal quality

The Moderating Role of Quality?

- Shaffer, Erdberg, and Meyer independently rated the overall quality of each adult sample
- Variables considered:

Type of Sample	% of College Students
# of Examiners	Use of Student Examiners
Sample Size	Findings on Examiner Effects
Admin. Quality Control	Scoring Quality Control

M # of Protocols Given per Examiner Before Start of Study

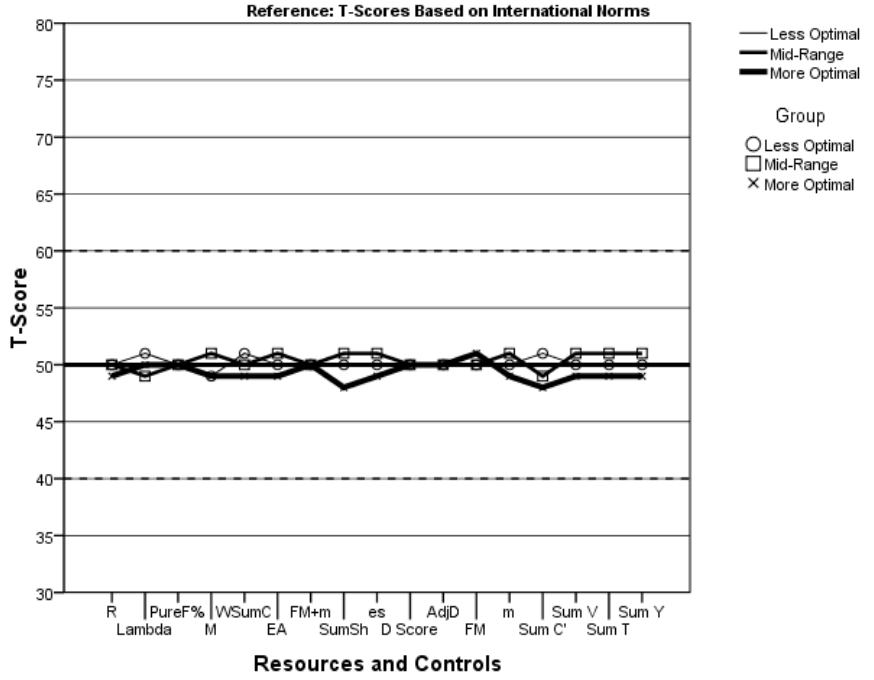
- Then 2 hours of discussion to reach consensus
  - 3 quality categories:
    - Less Optimal, Mid-Range, More Optimal

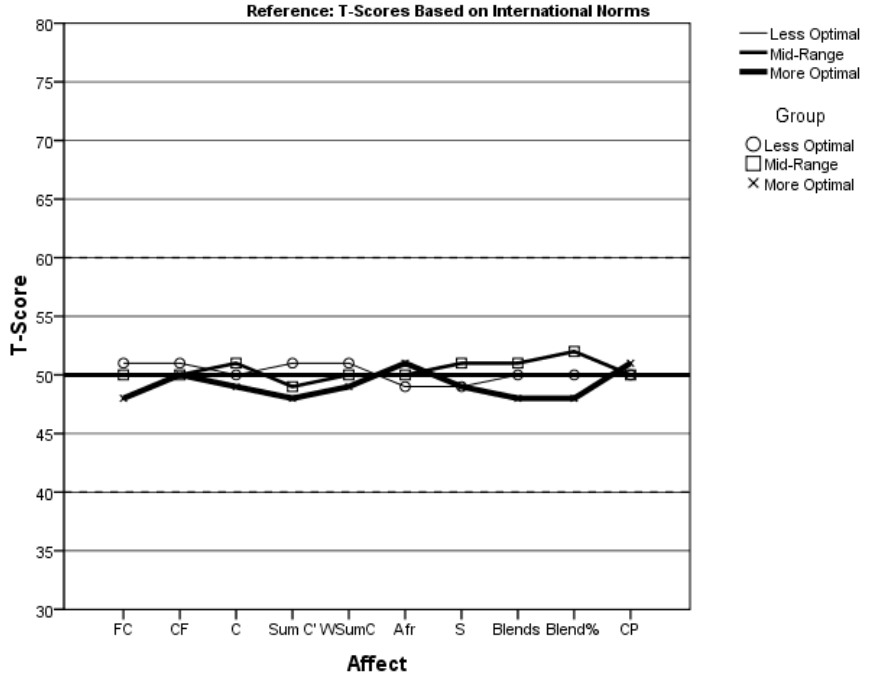
The Moderating Role of Quality?

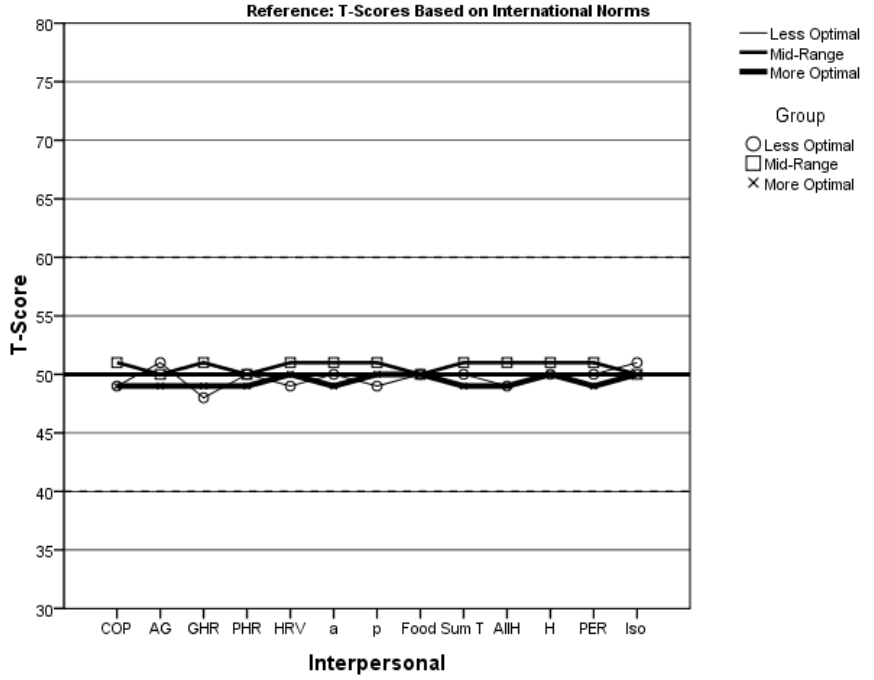
- Less optimal samples (5):
  - Just one examiner
  - Examiners with no previous experience
    - i.e., Contributed the 1st protocols they ever administered
- Mid-range (12, including all US samples)
  - Incomplete info on examiners and quality control
- More optimal samples (4):
  - Experienced examiners
  - Used and described ongoing quality control efforts

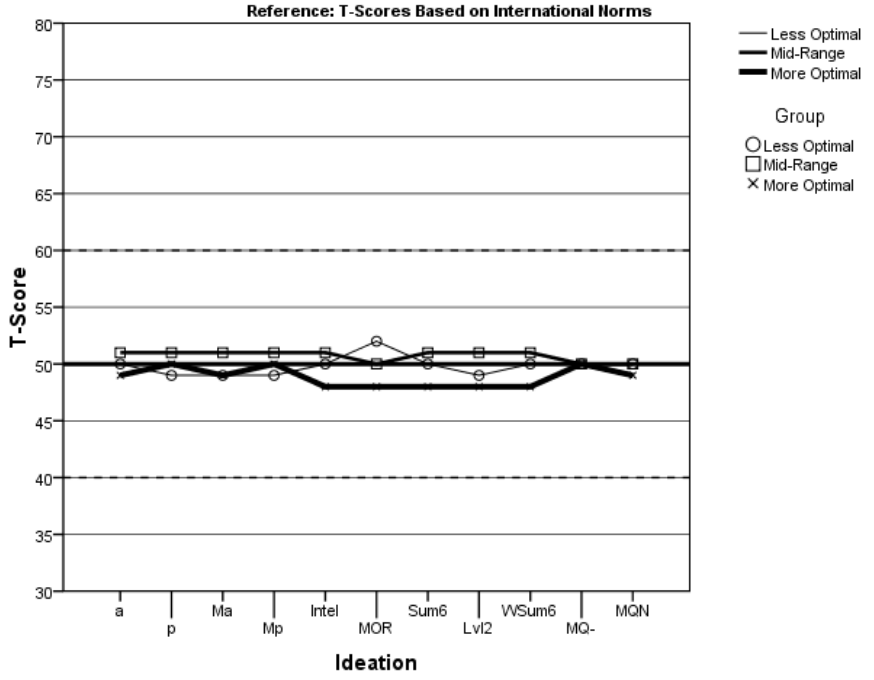
Plot of Scores by Sample Quality

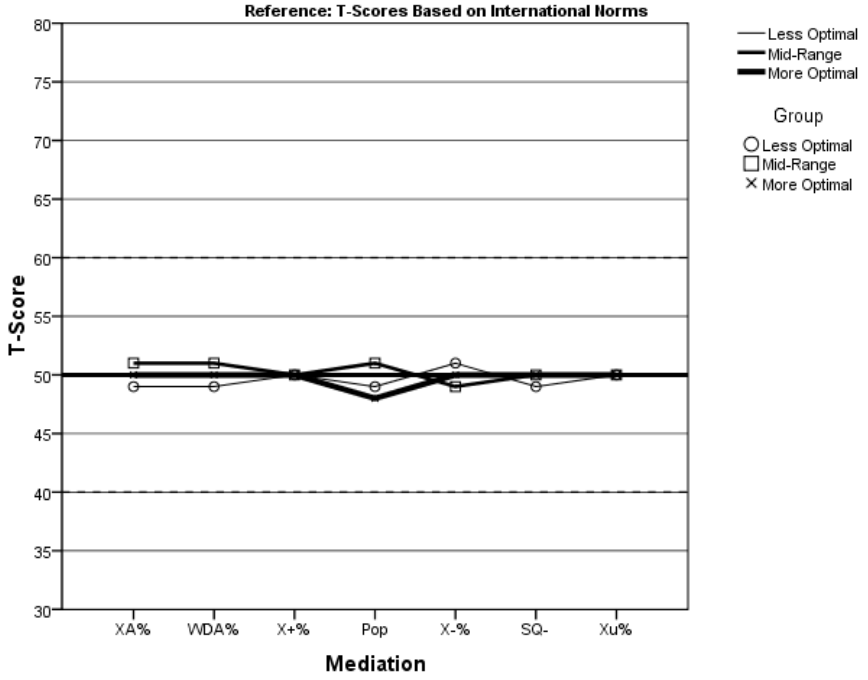
- Created average scores for the 3 quality groups and graphed them as T-Scores using the M and SD from the International Reference Values
- Key question:
  - To what extent do the T-Scores deviate from the expected value of 50 as a function of overall quality?
  - e.g.,
    - Does less optimal group have elevated Lambda?
    - Does more optimal group have more color or higher EA?
    - Does less optimal group look more disturbed?
- Next slides show results
  - As before, they give overall impressions; not specifics
  - Focus on deviation of lines from value of 50

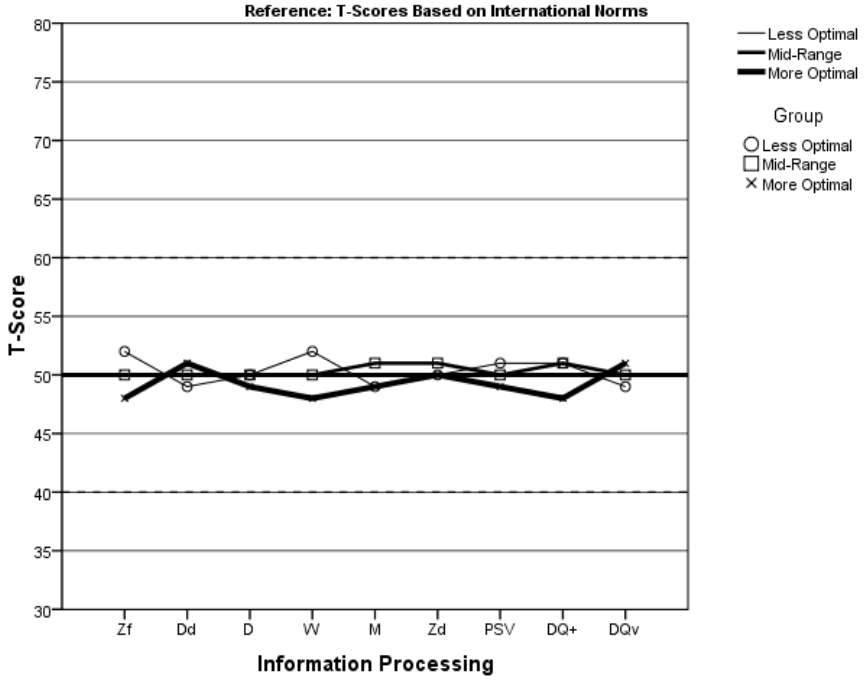

Reference: T-Scores Based on International Norms
T-Score
80
75
70
65
60
55
50
45
40
35
30
R
Lambda
PureF%
M
WSumC
EA
FM+m
SumSh
es
D Score
AdjD
FM
m
Sum C'
Sum V
Sum T
Sum Y
Resources and Controls
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal

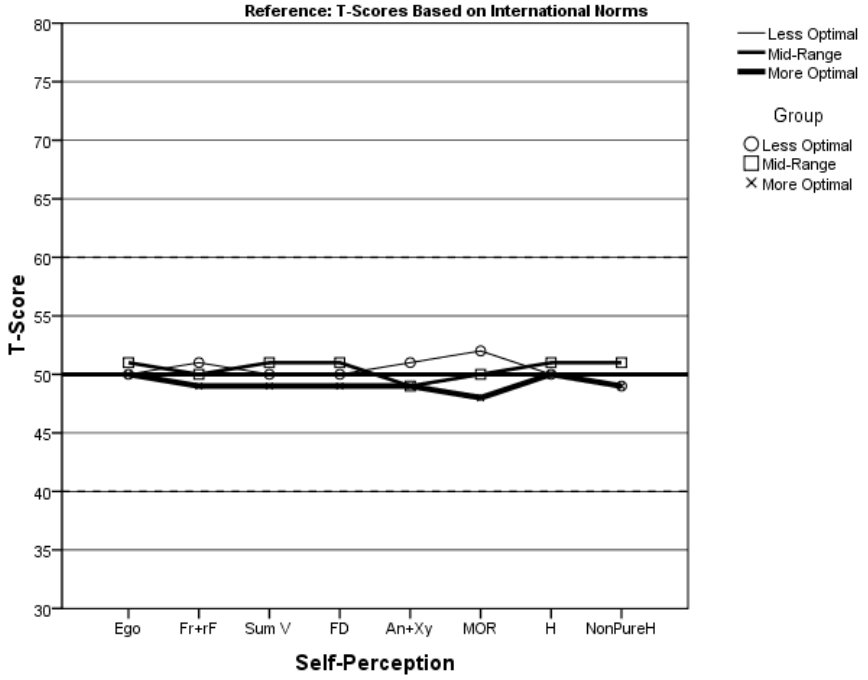

Reference: T-Scores Based on International Norms
T-Score
80
75
70
65
60
55
50
45
40
35
30
FC
CF
C
Sum C'
WSumC
Afr
S
Blends
Blend%
CP
Affect
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal

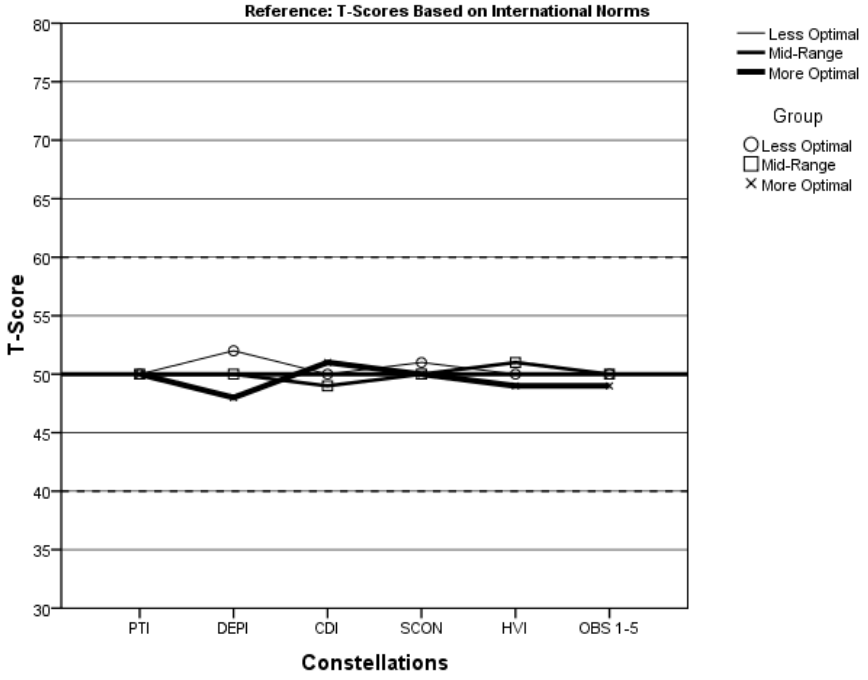

Reference: T-Scores Based on International Norms
T-Score
80
75
70
65
60
55
50
45
40
35
30
COP
AG
GHR
PHR
HRV
a
p
Food
Sum T
AllH
H
PER
Isol
Interpersonal
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal


Reference: T-Scores Based on International Norms
T-Score
80
75
70
65
60
55
50
45
40
35
30
a
p
Ma
Mp
Intel
MOR
Sum6
Lvl2
WSum6
MQ-
MQN
Ideation
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal


Reference: T-Scores Based on International Norms
T-Score
80
75
70
65
60
55
50
45
40
35
30
XA%
WDA%
X+%
Pop
X-%
SQ-
Xu%
Mediation
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal


Reference: T-Scores Based on International Norms
T-Score
80
75
70
65
60
55
50
45
40
35
30
Zf
Dd
D
W
M
Zd
PSV
DQ+
DQv
Information Processing
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal


Reference: T-Scores Based on International Norms
T-Score
80
75
70
65
60
55
50
45
40
35
30
Ego
Fr+rF
Sum V
FD
An+Xy
MOR
H
NonPureH
Self-Perception
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal


Reference: T-Scores Based on International Norms
T-Score
80
75
70
65
60
55
50
45
40
35
30
PTI
DEPI
CDI
SCON
HVI
OBS 1-5
Constellations
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal

Summary – Sample Quality

- The 3 quality-based groups are very similar
  - Deviations from *M* of 50 were trivial
  - Largest differences were 4 T-score points

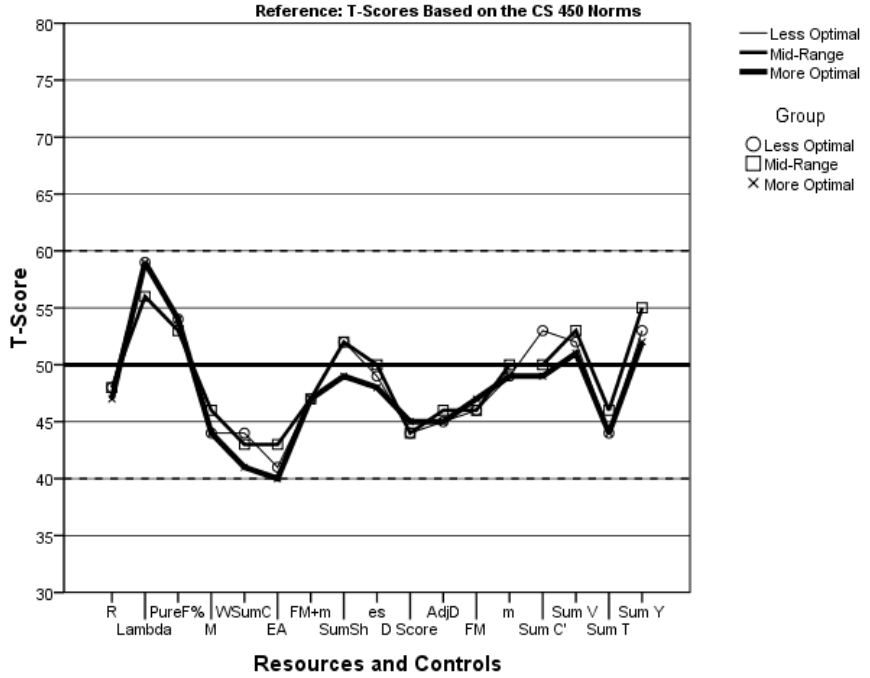
	Less Optimal	More Optimal
• W	52	48
• Zf	52	48
• MOR	52	48
• DEPI	52	48

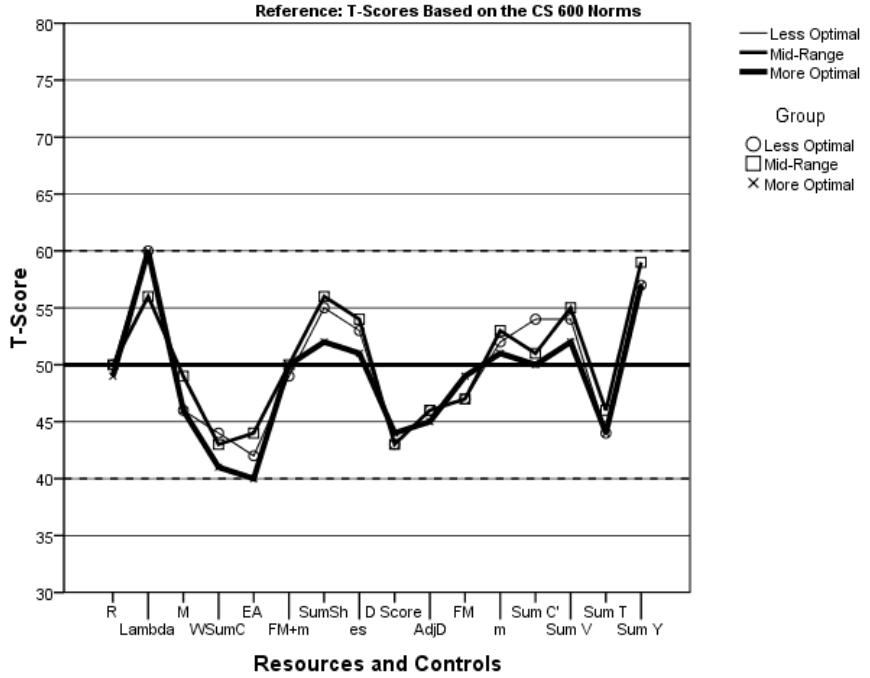
Summary – Sample Quality

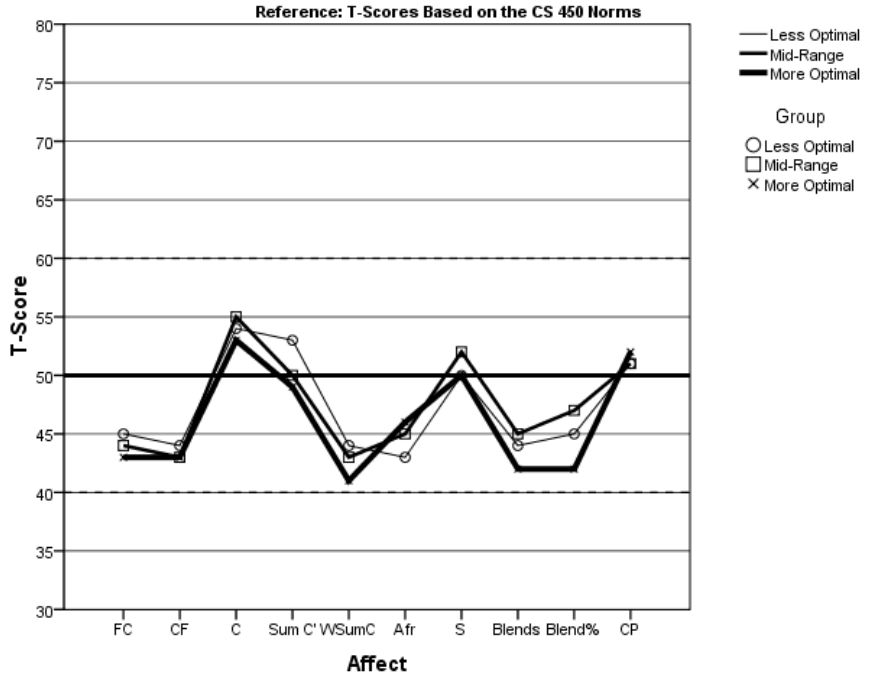
- Differences in overall sample quality do not lead to normative differences in scores
  - Less optimal studies do not produce protocols with less complexity or more pathology
  - But cannot rule out complex interactions
- And certainly training is important
  - Likely that variation in administration and scoring conventions "cancel out" across samples
  - More uniform training should decrease variability

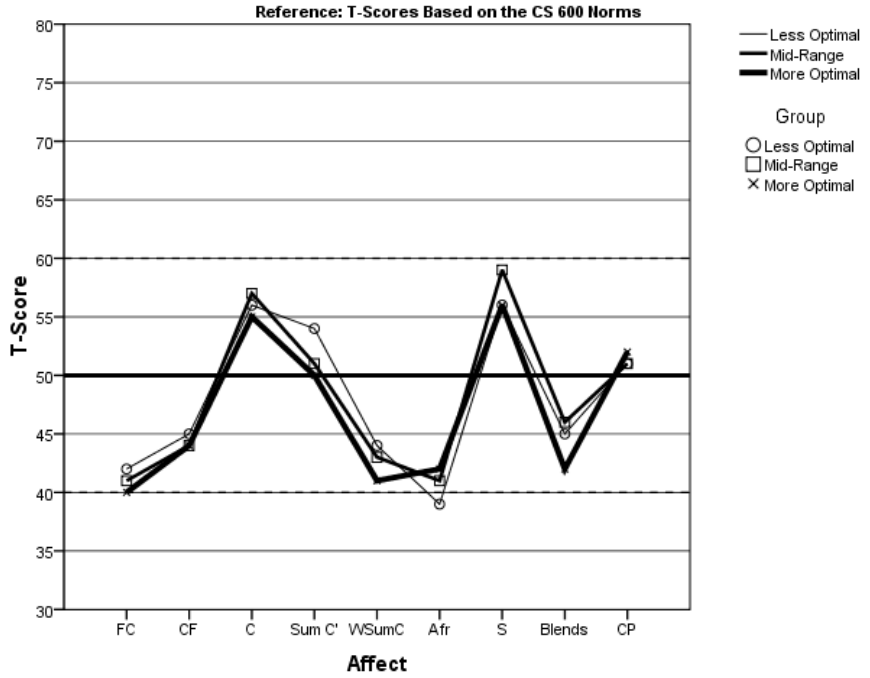
But Why Switch to New Norms?

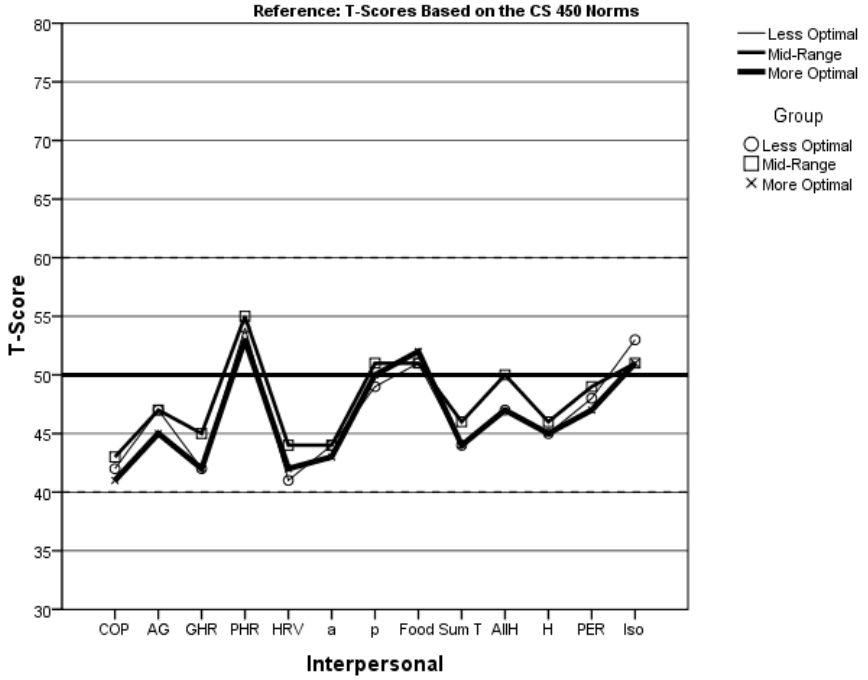
- Why change at all?
- A key question: How do the 3 quality-based samples look relative to the standard CS norms?
  - Created T-Scores using *M*s and *SD*s from:
    - New sample of 450 (on left in plots)
    - Older sample of 600 (on right in plots)

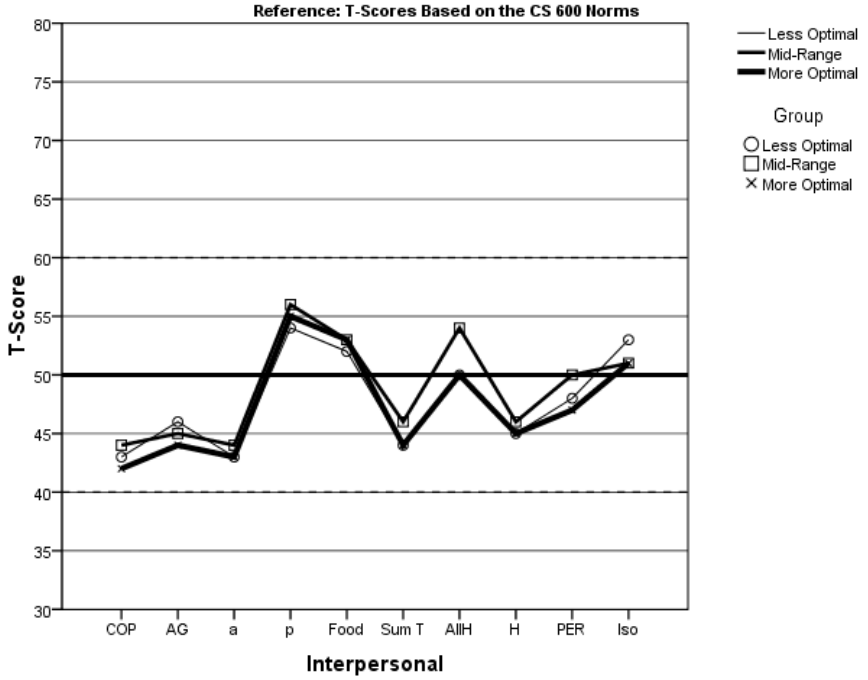

Reference: T-Scores Based on the CS 450 Norms
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal
T-Score
80
75
70
65
60
55
50
45
40
35
30
R
Lambda
PureF%
M
WSumC
EA
FM+m
SumSh
es
D Score
AdjD
FM
m
Sum C'
Sum V
Sum T
Sum Y
Resources and Controls

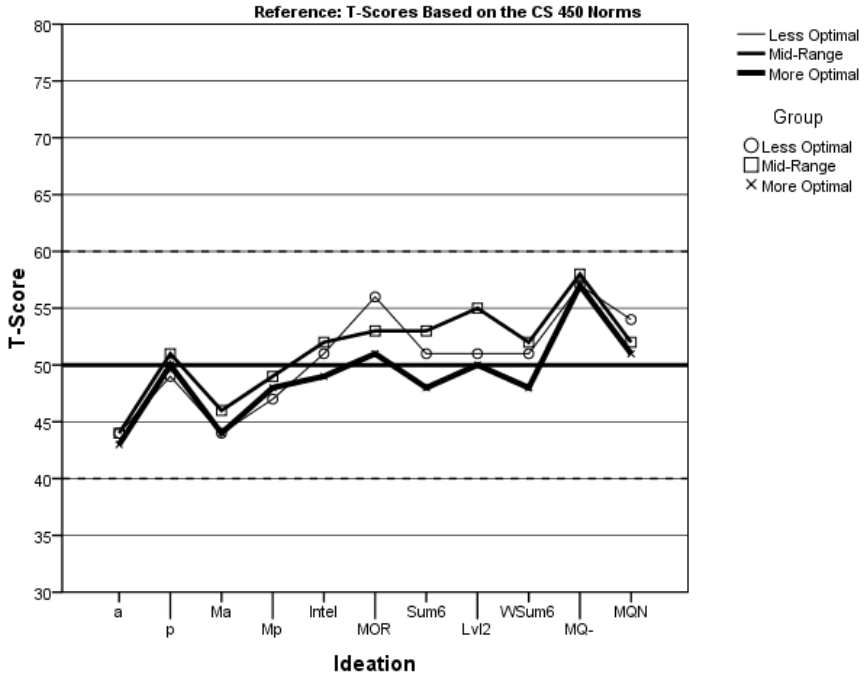

Reference: T-Scores Based on the CS 600 Norms
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal
T-Score
80
75
70
65
60
55
50
45
40
35
30
R
Lambda
M
WSumC
EA
FM+m
SumSh
es
D Score
AdjD
FM
m
Sum C'
Sum V
Sum T
Sum Y
Resources and Controls

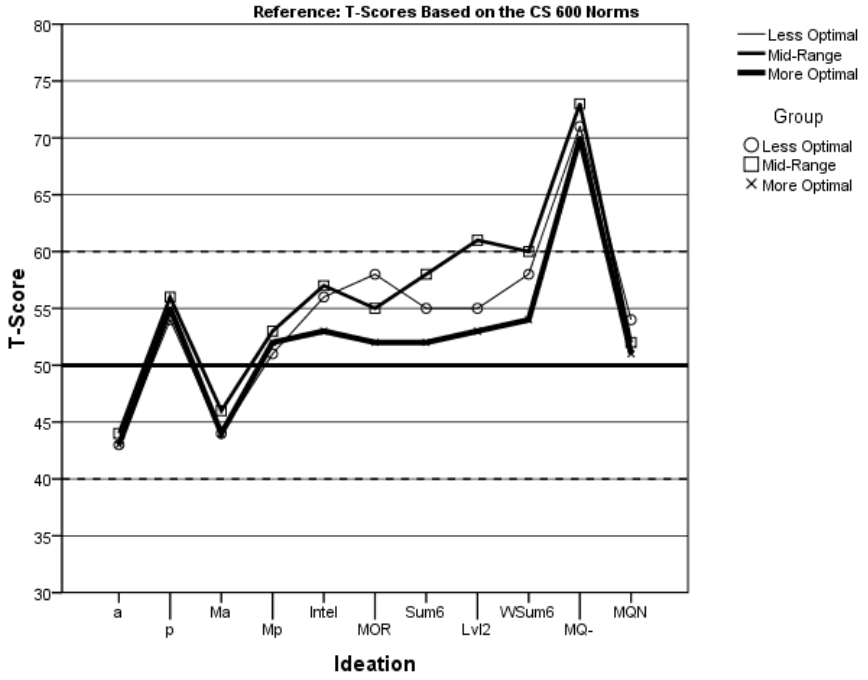

Reference: T-Scores Based on the CS 450 Norms
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal
T-Score
80
75
70
65
60
55
50
45
40
35
30
FC
CF
C
Sum C'
WSumC
Afr
S
Blends
Blend%
CP
Affect

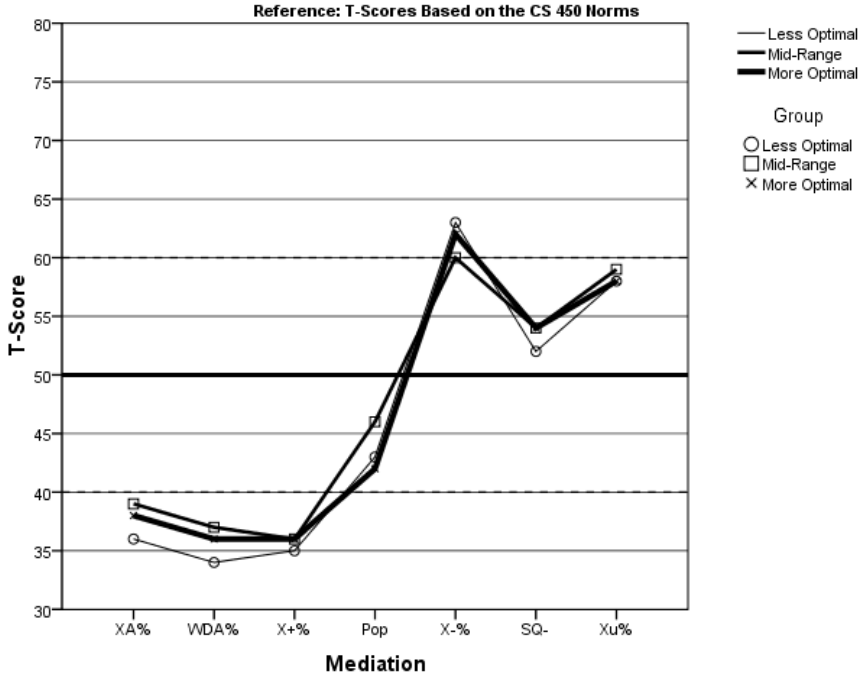

Reference: T-Scores Based on the CS 600 Norms
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal
T-Score
80
75
70
65
60
55
50
45
40
35
30
FC
CF
C
Sum C'
WSumC
Afr
S
Blends
CP
Affect

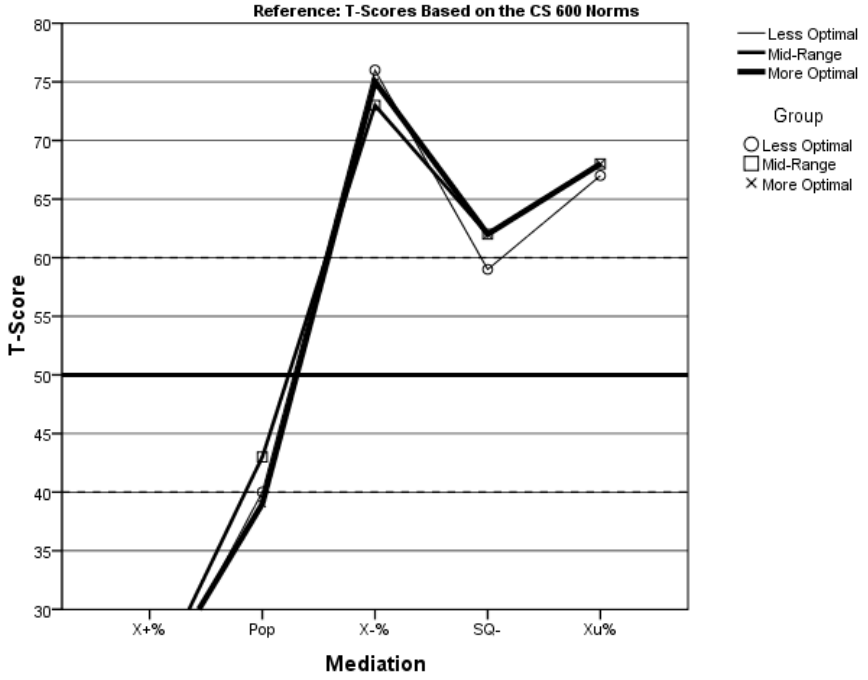

Reference: T-Scores Based on the CS 450 Norms
T-Score
80
75
70
65
60
55
50
45
40
35
30
COP AG GHR PHR HRV a p Food Sum T AllH H PER Iso
Interpersonal
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal

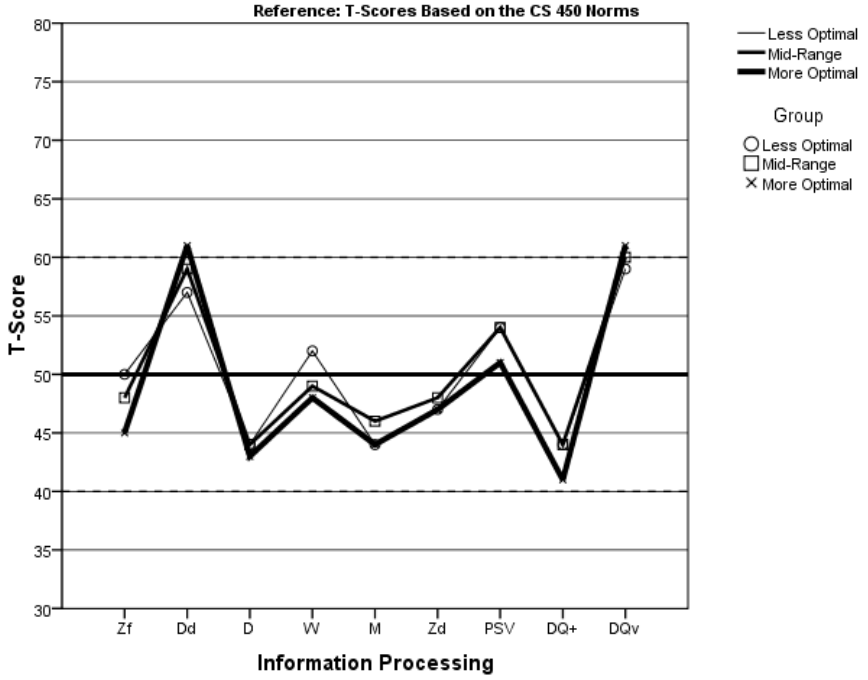

Reference: T-Scores Based on the CS 600 Norms
T-Score
80
75
70
65
60
55
50
45
40
35
30
COP AG a p Food Sum T AllH H PER Iso
Interpersonal
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal

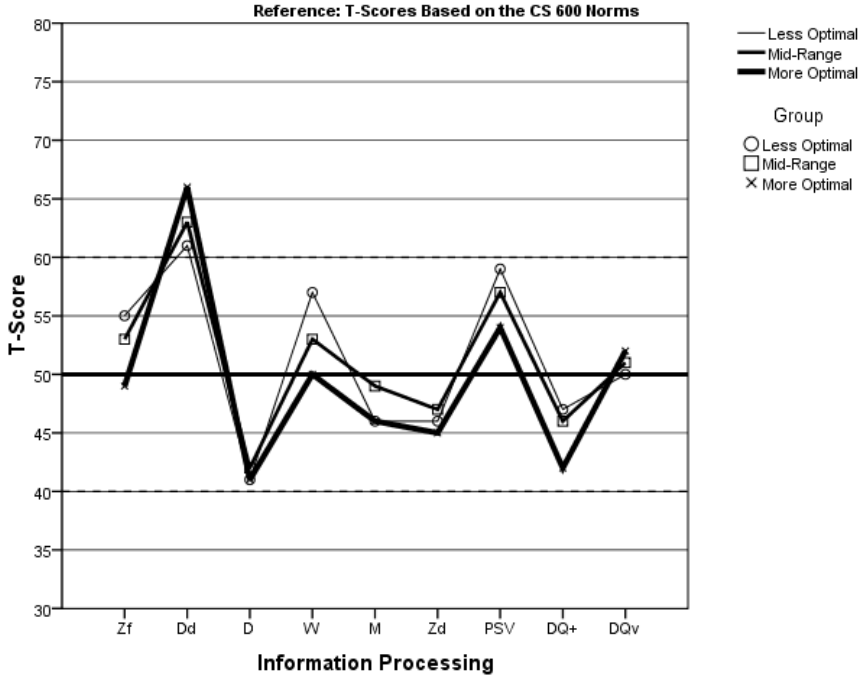

Reference: T-Scores Based on the CS 450 Norms
T-Score
80
75
70
65
60
55
50
45
40
35
30
a p Ma Mp Intel MOR Sum6 Lvl2 WSum6 MQ- MQN
Ideation
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal

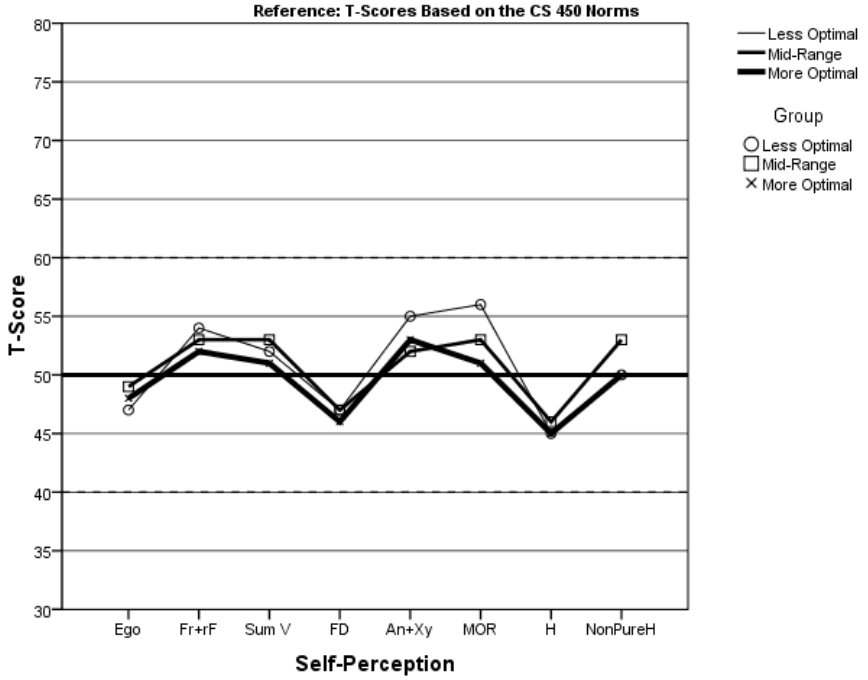

Reference: T-Scores Based on the CS 600 Norms
T-Score
80
75
70
65
60
55
50
45
40
35
30
a p Ma Mp Intel MOR Sum6 Lvl2 WSum6 MQ- MQN
Ideation
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal

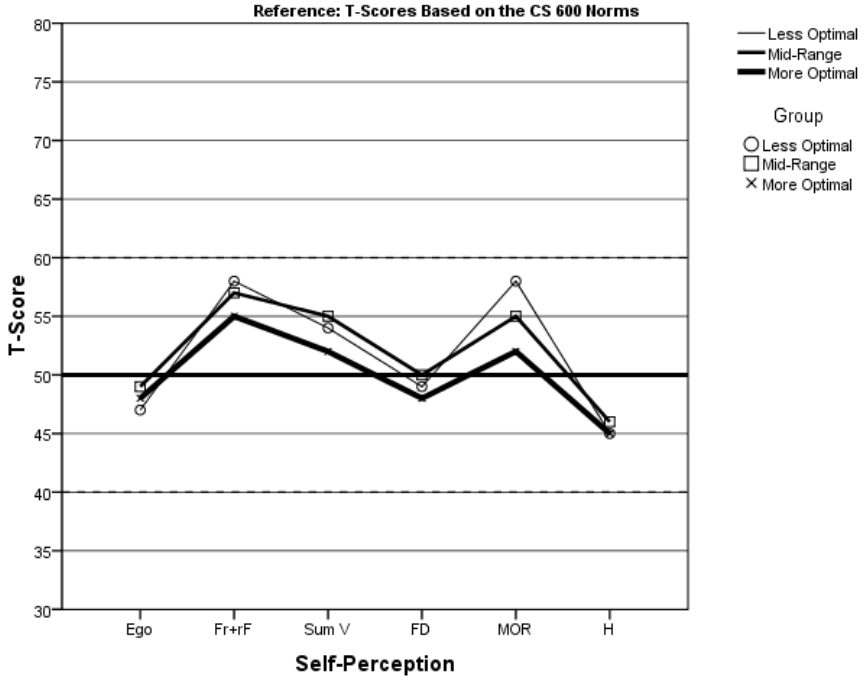

Reference: T-Scores Based on the CS 450 Norms
T-Score
80
75
70
65
60
55
50
45
40
35
30
XA%
WDA%
X+%
Pop
X-%
SQ-
Xu%
Mediation
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal

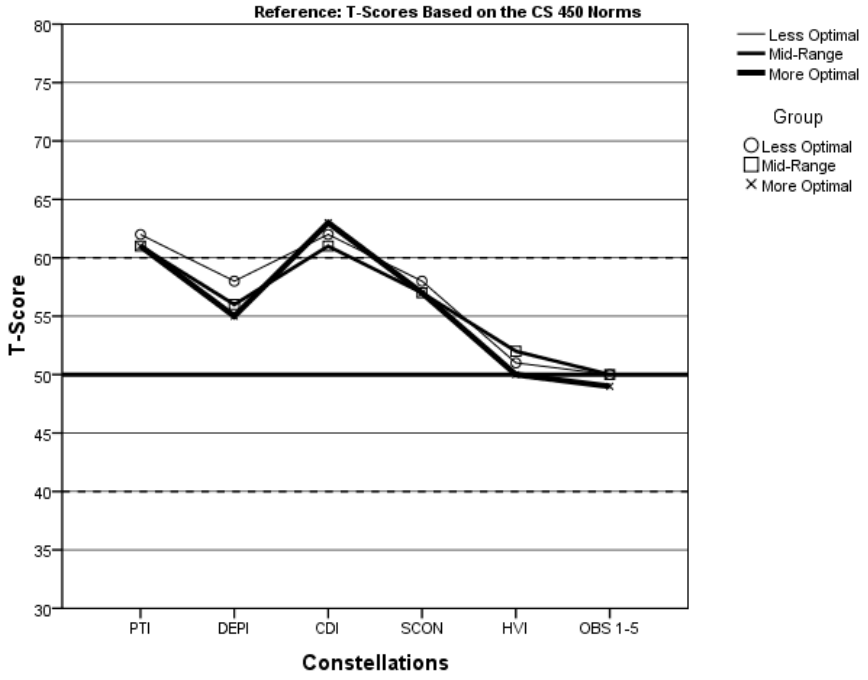

Reference: T-Scores Based on the CS 600 Norms
T-Score
80
75
70
65
60
55
50
45
40
35
30
X+%
Pop
X-%
SQ-
Xu%
Mediation
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal


Reference: T-Scores Based on the CS 450 Norms
T-Score
80
75
70
65
60
55
50
45
40
35
30
Zf
Dd
D
W
M
Zd
PSV
DQ+
DQv
Information Processing
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal


Reference: T-Scores Based on the CS 600 Norms
T-Score
80
75
70
65
60
55
50
45
40
35
30
Zf
Dd
D
W
M
Zd
PSV
DQ+
DQv
Information Processing
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal


Reference: T-Scores Based on the CS 450 Norms
T-Score
80
75
70
65
60
55
50
45
40
35
30
Ego
Fr+rF
Sum V
FD
An+Xy
MOR
H
NonPureH
Self-Perception
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal


Reference: T-Scores Based on the CS 600 Norms
T-Score
80
75
70
65
60
55
50
45
40
35
30
Ego
Fr+rF
Sum V
FD
MOR
H
Self-Perception
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal


Reference: T-Scores Based on the CS 450 Norms
T-Score
80
75
70
65
60
55
50
45
40
35
30
PTI
DEPI
CDI
SCON
HVI
OBS 1-5
Constellations
Less Optimal
Mid-Range
More Optimal
Group
Less Optimal
Mid-Range
More Optimal

Summary – Standard Norms

- Relative to the standard CS norms:
  - 3 quality-based samples show the same patterns
    - i.e., they still look very similar to each other
      - Maximum differences of ≈ 5 T-score points
  - But all 3 international samples look unhealthy on some variables
    - These would be considered "large" differences
    - With average T-Scores > 57 or < 43
      - Note these are differences in sample means, not individual scores

Summary – Standard Norms

- These data indicate that clinicians in the U.S. and in other countries using the standard CS norms will incorrectly infer that nonpatients:
  - Are prone to perceptual distortions
  - See the world in an atypical and idiosyncratic manner
  - Tend to be simplistic
  - Lack affective resources
  - Lack coping resources in general
  - Are prone to affective disturbances and dysregulation
  - Misunderstand others and misperceive relationships

Conclusion – International Norms

- The composite international reference samples may not seem ideal for any specific country but their consistency and generalizability across cultures, languages, and styles of administration and scoring appear to capture universal features of the human experience
- Patients would be better served if clinical inferences were drawn from them rather than the standard CS reference values

Establishing Normative Benchmarks for RPAS

- Goal: Use international reference data to
  - Closely estimate R-Optimized administration
  - Generate easily interpreted benchmarks
    - For raw scores and complexity adjusted scores
- The International Sample of 1496 Protocols
  - A composite of 16 adult samples
  - All but 2 in the 2007 *JPA* Supplement
  - Generally gave equal emphasis to each sample
    - Selected at most 100 records
- Many others offered data that could not be used
  - Not an adult sample
  - Response level scores not electronically available

Modeling R-Optimized Protocols

- From this pool we modeled an R-optimized administration
- Target database of 112 US protocols collected using the final R-optimized instructions
  - Prompt for a 2nd if only one is offered
  - Ask for the card back after the 4th
  - Tell examinees to give 2 responses, maybe 3
- Examiners:
  - Students being trained at AIU
  - All administered at least two protocols before contributing data

Modeling R-Optimized Protocols

- Sample of 112 R-Optimized protocols
  - No card rejections
  - Range of R: 16 to 35
  - M = 21.47, SD = 3.68
- Targeted this distribution
  - At Card level and Protocol level

Modeling R-Optimized Protocols

- Initially, precisely selected among the 2nd, 3rd, and 4th responses to each card
  - e.g., randomly selected 115 of the 312 3rd responses given to Card II so the % of people with a 3rd response in our sample equaled the % in the 112 records
  - Paradoxically, this resulted in a protocol level SD that was too small (2.65 rather than 3.7)
  - So took more liberal approach of approximating the M R per card
    - Worked well and greatly simplified the process

Modeling R-Optimized Protocols

- In the end, began with people who gave at least 2 responses to 6 of the last 9 cards
  - Card I not consider because already prompted
- Eliminated all responses after the 4th
- Randomly selected from the "excess" of 3rd and 4th responses to each card
  - Selected 60% of the 3rd responses
    - If a 3rd was eliminated in a protocol, so was any 4th
  - Selected 75% of the remaining 4th responses

Modeling R-Optimized Protocols

- Card Level:

Card	1	2	3	4	5	6	7	8	9	10
Target M	2.19	2.12	2.21	2.03	2.08	2.11	2.12	2.13	2.19	2.31
Our M	2.34	2.27	2.27	2.04	2.02	2.06	2.14	2.25	2.17	2.50

- Protocol level:

	Target	**Ours**
Mean	21.47	22.06
SD	3.68	3.90
Min	16	16
Max	35	37

- N = 849 from the initial pool of 1496
  - Example data on next slide

Variables	%Scored	%NotZero	Mean	SD	Sk	Ku	Min	Mdn	Max	Mode
Complexity	100.0%	100.0%	68.84	20.88	0.76	0.94	25.00	66.00	160.00	54.00
Loc/S/Obj Complexity	100.0%	100.0%	30.81	8.11	0.98	1.89	16.00	30.00	74.00	29.00
Determinant Complexity	100.0%	100.0%	17.60	7.49	0.69	0.74	2.00	17.00	48.00	15.00
Content Complexity	100.0%	100.0%	20.44	7.45	0.67	0.94	4.00	19.00	53.00	19.00
R	100.0%	100.0%	22.06	3.91	0.83	0.28	16.00	21.00	36.00	19.00
Form%	100.0%	99.6%	41.09	17.23	0.23	-0.40	0.00	41.00	88.00	32.00
Blend%	100.0%	91.3%	17.48	12.78	0.84	0.53	0.00	15.00	67.00	0.00
Synthesis% (Syn%)	100.0%	98.5%	29.20	15.53	0.35	-0.35	0.00	29.00	78.00	35.00
EA	100.0%	99.4%	6.79	3.53	0.78	0.91	0.00	6.50	22.00	7.00
EA - es	100.0%	94.7%	-1.89	4.44	-0.26	0.51	-20.00	-2.00	12.00	-2.00
M/(M+WSumC)	99.4%	94.2%	54.16	24.23	-0.29	-0.23	0.00	56.00	100.00	67.00
M	100.0%	94.2%	3.73	2.53	1.18	3.28	0.00	3.00	21.00	3.00
WSumC	100.0%	94.7%	3.06	2.08	0.83	0.63	0.00	3.00	12.00	1.50
FC/(FC+CF+C)	94.7%	78.0%	49.14	33.04	0.06	-1.02	0.00	50.00	100.00	0.00
FC	100.0%	78.0%	1.86	1.61	0.99	0.93	0.00	2.00	8.00	1.00
CF+C	100.0%	78.1%	1.97	1.72	1.05	1.41	0.00	2.00	11.00	1.00
EII-3	100.0%	100.0%	-0.12	1.06	0.67	0.63	-2.45	-0.21	4.13	-0.61
Thought & Percept. Comp. (TP-Comp)	100.0%	96.2%	0.58	0.99	0.70	0.62	-1.40	0.40	4.80	-0.20
WSum6	100.0%	82.8%	7.02	7.07	1.78	5.05	0.00	5.00	55.00	0.00
Level 2 Cognitive Scores	100.0%	18.3%	0.24	0.60	**3.84**	**24.45**	0.00	0.00	7.00	0.00
X-%	100.0%	97.5%	19.33	10.77	0.76	0.79	0.00	18.00	68.00	11.00
X+%	100.0%	100.0%	51.97	13.39	-0.05	0.19	5.00	52.00	94.00	50.00
WDA%	100.0%	100.0%	82.65	10.78	-0.84	1.51	24.00	84.00	100.00	88.00
Popular	100.0%	99.8%	5.43	1.92	0.24	0.23	0.00	5.00	14.00	6.00
m	100.0%	69.3%	1.49	1.51	1.25	1.68	0.00	1.00	8.00	0.00
Y	100.0%	64.8%	1.33	1.50	1.72	4.71	0.00	1.00	12.00	0.00
m+Y	100.0%	86.5%	2.82	2.32	1.28	2.66	0.00	2.00	16.00	1.00
MOR (S&D)	100.0%	61.5%	1.17	1.34	1.55	3.00	0.00	1.00	8.00	0.00
Suicide Composite (S-Comp)	100.0%	100.0%	4.50	1.27	0.58	0.21	1.90	4.30	9.20	4.20

Modeling Normative Data

- Goal: Create readily interpretable norms that apply a common benchmark across scores
  - Move away from needing to know raw score distributions for every variable
- Vexing problem is contending with non-normal count distributions
- We computed percentiles and converted them to their standard score equivalents
  - Every percentile has a SS equivalent

Modeling Normative Data

Percentile	Standard Score
16th	85
50th	100
84th	115

- For scores that encompass multiple percentiles the average value is assigned
  - T = 0 encompasses the lowest 57.7 of the distribution so has a percentile of 28.85
- Transformation from raw score to percentile to standard score allows us to put all variables on a common metric while not distorting the underlying distributions

Modeling Normative Data

- Next, adjusted for the 1st factor of complexity
  - Like WAIS strengths and weaknesses
- Typical linear regression or ANCOVA is not an option with skewed count variables
- Considered alternatives:
  - Poisson regression
  - Negative binomial regression
  - Zero-inflated negative binomial regression

Modeling Normative Data

- All alternatives could handle the task
  - Many Rorschach scores are distributed as ZINB with a large number of zero values and a long, skewed tail of counts to encompass the few individuals with very large scores
- But in the end used quantile regression, which can predict any percentile in a distribution
  - Developed in the 18th century but rarely used
  - Just recently available in SPSS with R

Modeling Normative Data

- Steps for each variable:
  - Predicted the 50th percentile from complexity
  - Computed the difference (residual) between each person's actual score and the predicted value
  - Determined the frequency of the difference
  - That frequency defines the percentile, which then is converted to the Complexity Adjusted Standard Score
- Residual distribution has same basic shape as original variable distribution

Organized Output

- Next 4 slides show computer program output
  - Coding sequence
  - Summary scores (Counts and Calculations)
  - Page 1 Profile; upper and lower halves
    - Page 1 = Variables with the best support
    - Page 2 = Variables with some support
- Program written by Fabiano Miguel, Ph.D.
  - From Ricardo Primi's lab at Universidade São Francisco – Brasil

Page 1 and Page 2 Profiles

- 4 Logically Organized Domains
  - Engagement and Cognitive Complexity
  - Perception and Thinking
  - Stress and Distress
  - Self- and Other-Representation
- 4 Types of Scores
  - Raw
  - Percentile
  - Standard Score
  - Complexity Adjusted Standard Score
- 8 Plotting Options
  - R-Opt Standard Scores
  - R-Opt Complexity Adjusted Standard Scores
  - "Legacy" (non-R-Opt) Standard Scores
  - "Legacy" (non-R-Opt) Complexity Adjusted Standard Scores
  - Hashmarks, that show raw score equivalent values on the plot, or not

Card	#	R-Opt	Or	Loc	Loc #	Space	Content	Sy	Vg	2	FQ	Pop	Det	Cognitive	Thematic	ROD (RP Only)
I	1			W		Si	(Ad)				o		F			
I	2			W			(H),Cg	+			o		F		GHR	
I	3		v	Dd	99		NC				u		F		AGC	
II	4			D	6,3		An	+		2	-		CF,FD			
II	5			D	6		A	+		2	o	P	FMp,FD			
II	6		v	W			A				-		Y			
III	7			D	1		(H),Sx			2	o		Mp	INC1	PHR	
III	8		v	Dd	99	Si	A				-		FC,C′		PER,AGC	
III	9		>	W			(A)	+			-		FMp,Y,r	DV1	AGC	
IV	10		v	W			(A)				u		F		AGC	
IV	11			D	7		(H)				o	P	FD		AGC,GHR	
V	12			W			A				o		FD			
V	13		v	W			A				o	P	F			
V	14		>	W			NC	+			-		F		MOR	ROD
VI	15			W			Art				u		F		PER,AGC	
VI	16			W			A				u		FD			ROD
VI	17		v	W			NC				u		F		AGC	
VII	18			W			H,Ay	+		2	o	P	Ma,mp		COP,MOAH,GHR	
VII	19		v	W			(A)				-		Mp		PER,PHR	
VII	20		v	D	7	Sr	NC				o		F		PER	
VIII	21		v	D	6		Ad,An	+			-		FC,FD		MOR	
VIII	22			W			A,NC	+		2	u		FMa			
IX	23			W			Art				o		FC			
IX	24		v	W			(H)				u		F	INC1	PHR	ROD
IX	25			W			(Ad)	+			o		Mp,FC,C′	DR1	PER,AGC,PHR	
X	26			D	1,12		(A)	+		2	u		FMa,C		AGM,MOAP,PHR	ROD
X	27			D	11,3,6,10		A	+			u		FMa			
X	28		v	Dd	22	Si	Ad				-		F	DV1		

RPAS Protocol Level Counts & Calculations

Domain	Counts	Counts	Calculations
Resp. & Adm.	R = 28	R8910 = 8	R8910% = 29%
	Prompts = 0	Pulls = 0	
	< or > = 2		
Location	W = 17	Dd = 3	W% = 61%
	D = 8	WD = 25	WD% = 89%
Space	Sr = 1	Sri = 0	SR = 1
	Si = 3	SumS = 4	SI = 3
Content	H = 1	Art = 2	AllH = 5
	(H) = 4	Ay = 1	NPH = 4
	Hd = 0	Bl = 0	PureH/AllH = 20%
	(Hd) = 0	Cg = 1	An+Xy = 2
	A = 8	Ex = 0	2ABS+Art+Ay = 3
	(A) = 4	Fi = 0	CritCont = 6
	Ad = 2	Sx = 1	DramCont = 4
	(Ad) = 2	Xy = 0	TCI = 6
	An = 2	NC = 5	TCI% = 21%
Syn	All + = 11		Syn% = 39%
Vague	All v = 0		Vague% = 0%
Pairs	Pair = 6		
Form Quality and Popular	FQo = 11	WDo = 11	X+% = 39%
	FQu = 9	a = 8	Xu% = 32%
	FQ- = 8	M- = 1	X-% = 29%
	FQn = 0		WDA% = 76%
	Popular = 4		

Domain	Counts	Counts	Calculations	
Det.	M = 4	FC = 4	WSumC	= 4.5
Blends	FM = 5	CF = 1	FC/(FC+CF+C)	= 67%
CF,FD	m = 1	C = 1	EA	= 8.5
FMp,FD	C′ = 2	Y = 2	M/(M+WSumC)	= 47%
FC,C′	T = 0	V = 0	SumShd	= 4
FMp,Y,r	r = 1	FD = 6	m+Y	= 3
Ma,mp		F = 11	Form%	= 39%
FC,FD			es	= 10
Mp,FC,C′			EA-es	= -1.5
FMa,C	a = 4	p = 6	a/(a+p)	= 40%
	Ma = 1	Mp = 3	Ma/(Ma+Mp)	= 25%
	Blends = 8	C-Sh Bl = 2	Blend%	= 29%
Cognitive Scores	DV1 (1) = 2	DV2 (2) = 0	WSum6 = 9	
	INC1 (2) = 2	INC2 (4) = 0	Level 2 = 0	
	DR1 (3) = 1	DR2 (6) = 0		
	FAB1 (4) = 0	FAB2 (7) = 0		
	ALOG (5) = 0	CON (7) = 0		
Thematic Scores	ABS = 0	PER = 5		
	MOR = 2	AGP = 0		
	AGC = 8	AGM = 1		
	COP = 1	ROD = 4	ROD%	= 14%
	MOAH = 1	MOAP = 1	MOAH/(MOAH+P)	= 50%
	GHR = 3	PHR = 5	GHR/(GHR+PHR)	= 38%
Other Calc.	EII-3 = 0.98	Complexity = 91		
	TP-Comp = 1.5	L/S/O = 45		
	V-Comp = 5.0	Det = 27		
	S-Comp = 6.1	Cont = 19		

Profile shows: Standard Scores Hashmarks: Hide

RPAS Summary Scores and Profiles – Page 1

Name: Case Present		ID: Workshop Case Presentation		Age:	Gender: NA	Education:
Domain/Variables	**Raw Score**	**%ile**	**Stand. Score**	**Compl. Adj. SS**	**Standard Score Profile**	
Engagement and Cognitive Processing					60 70 80 90 100 110 120 130 140	
Complexity	91	86	116			
Location/Space/Object Complexity	45	94	124			
Determinant Complexity	27	90	119			
Content Complexity	19	47	99			
Responses (R)	28	91	120	116		
Form% [Lambda=0.65]	39%	47	99	109		
Blend%	29%	82	114	106		
Syn%	39%	74	110	99		
EA	8.5	72	109	92		
EA-es	-1.5	52	101	102		
M:C [M/(M+WSumC)]	47%	36	94	94		
M	4	59	103	89		
WSumC	4.5	78	112	104		
FC:CF+C [FC/(FC+CF+C)]	67%	70	108	108		
FC	4	89	119	115		
CF+C	2	57	103	94		
Perception and Thinking					60 70 80 90 100 110 120 130 140	
EII-3	0.98	86	116	113		
Thought & Perception Composite (TP-Comp)	1.5	84	115	112		
WSum6	9	71	108	104		
Level 2 Cognitive Scores	0	41	97	97		
X-%	29%	83	114	113		
X+%	39%	15	84	86		
WDA%	76%	26	90	90		
Popular	4	24	90	86		

Stress and Distress					60	70	80	90	100	110	120	130	140
m	1	45	98	83					●				
Y	2	74	110	110						●			
m+Y	3	61	104	91					●				
MOR	2	78	112	112						●			
Suicide Composite (S-Comp)	6.1	89	119	115							■		
Self & Other Representation					**60**	**70**	**80**	**90**	**100**	**110**	**120**	**130**	**140**
ROD%	14%												
Space Reversal (SR)	1												
MOAH:MOAP [MOAH/(MOAH+MOAP)]	50%												
MOAH	1												
MOAP	1												
GHR:PHR [GHR/(GHR+PHR)]	38%	16	85	87				●					
GHR	3	37	95	84					●				
PHR	5	86	116	112							■		
AGC	8												
AGP	0												
MOR	2	78	112	112						●			
Vigilance Composite (V-Comp)	3.5	63	105	92						●			
COP	1	56	102	92					●				

Note: All ratios are computed as the numerator over the sum of numerator and denominator.

"Sounds Great, But Can I Take It To Court?"

Robert E. Erard, Ph.D.

SPA in San Jose

March 2010

Frye v. United States 293 F 1013 (1923)

- An early polygraph case
- Promulgated a “general acceptance” standard for federal expert testimony
- Did not require ‘universal acceptance’
- Mostly applied to novel techniques
- Was adopted by many of the states

Federal Rule of Evidence 702

"If scientific, technical, or other specialized knowledge will assist the trier of fact to understand evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto on the form of an opinion or otherwise., IF (1) THE TESTIMONY IS BASED UPON SUFFICIENT FACTS OR DATA, (2) THE TESTIMONY IS THE PRODUCT OF RELIABLE PRINCIPLES AND METHODS, AND (3) THE WITNESS HAS APPLIED THE PRINCIPLES AND METHODS RELIABLY TO THE FACTS OF THE CASE."

Daubert v. Merrell Dow Pharmaceuticals

509 US 579, 1135 S Ct. 2786, 125 L Ed 469 (1993)

The purpose of the relevancy and reliability requirements is, “to make certain that an expert, whether basing testimony upon professional studies or personal experience, employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field.” (119 S.Ct. at 1176).

Daubert's Progeny

- ***General Electric v. Joiner, 522 US 136, 118, S. Ct. 512, 139 L Ed 2d 508 (1997)***
  --Firmly mandated judicial gatekeeping
  --Established abuse of discretion standard
  --The "analytical gap"

- ***Kumho Tire v. Carmichael, 526 US, 199 S Ct 1167 (1999)***
  --Extended *Daubert* to non-scientific (experience-based or 'clinical' testimony)
  --It is up to the judge to decide what criteria will be used to establish reliability according to the needs of the particular case
  --The expert's testimony must have a factual basis in the admitted evidence of the case.

Justice Steven Breyer's Bottom Line

From the Federal Judicial Center's Reference Manual on Scientific Evidence

"The search is not a search for scientific precision. We cannot hope to investigate all the subtleties that characterize good scientific work. A judge is not a scientist, and a courtroom is not a scientific laboratory. But consider the remark made by the physicist Wolfgang Pauli. After a colleague asked whether a certain scientific paper was wrong, Pauli replied, 'That paper isn't even good enough to be wrong!' Our objective is to avoid legal decisions that reject that paper's so-called science. The law must seek decisions that fall within the boundaries of scientifically sound knowledge."

Daubert v Merrell Dow Pharmaceuticals

The seven *Daubert* guidelines are as follows:

1) Is the proposed theory (or technique), on which the testimony is to be based, testable?

2) Has the proposed theory (or technique) been tested using valid and reliable procedures and with positive results?

3) Has the theory (or technique) been subjected to peer review?

Daubert Criteria (cont.)

4) What is the known or potential error rate of the scientific theory or technique?
5) What standards, controlling the technique's operation, maximize its validity?
6) Has the theory (or technique) been generally accepted as valid by a relevant professional community?
7) [Added later] Do the expert's conclusions reasonably follow from applying the theory (or technique) to this case?

Rorschach Training Programs Commentary on RPAS (Ritzler & Sciara, 2010)

"Much effort will be required to establish a new Rorschach system. For a new system to be unique and have the potential for added value, it should include research from the 'ground up.' That means that administration, coding variables, and interpretative strategies must each be researched. It is impossible to develop a new system based on research from the CS. For example, while the international norms may be a cornerstone of a new system, it is inappropriate to use those norms for a new system as the protocols were collected using the administration, coding, and interpretation methods from the CS."

RTP Commentary (cont.)

"Any new system must undergo the rigors of scrutiny by the professional community. The international norms have yet to undergo this scrutiny. A single publication of these norms does not meet any legal standard for inclusion as a foundation for use in court. Likewise, new administration procedures, new variables, and new interpretative strategies must be scrutinized over time to satisfy the needs of the forensic community for inclusion in evaluations to be presented in court."

Is the proposed technique testable?

Yes, this is an evidence-based method.

- The personality descriptions and predictions can be tested individually and in combination.
- Techniques may include convergent and discriminant correlations with other measures of the relevant construct, confirmatory factor analysis, and behavioral predictions.

Has the proposed technique been tested using valid and reliable procedures and with positive results?

- Each variable, particularly on pg. 1, has been tested in research and in practice, with valid and reliable procedures and positive results.
- The international norms have shown remarkable convergence.
- Modeling simulations have been used to predict the impact of various administrative and scoring changes on CS-derived norms and additional new normative data are being gathered.

Has the technique been subjected to peer review?

- The Rorschach is one of the procedures most widely studied in peer-reviewed literature.
- The RPAS is built primarily upon peer-reviewed research, including actual clinician experience as studied by survey.
- The RPAS as a whole has not been subjected to peer review, but by design, it will be responsive to ongoing peer-reviewed research in its future refinements and revisions.

What is the known or potential error rate of the scientific technique?

- The concept of error rate does not apply very well to personality descriptions, where “goodness of fit” is a better criterion.
- Those RPAS variables that are susceptible to use in classifications and predictions have been or will be tested and error rates established.
- We expect a typical error rate for our Page One variables to approach the typical ceiling found in personality testing: ~33%.

What standards, controlling the technique's operation, maximize its validity?

- Carefully specified and standardized rules for administration and clarification
- R-Optimized administration
- Complexity-adjusted norms
- Elimination of many variables with low coding reliability

Has the technique generally been accepted as valid by the relevant professional community?

- Other Rorschach systems, most prominently the CS, has been recognized by nearly all courts as generally accepted by the relevant professional community
- RPAS builds upon the CS with a focus on strengthening the evidence base.
- RPAS had a very positive early reception by many leaders in the Rorschach community.
- Other personality tests with substantial changes from previous versions (e.g. MCMI-III, MMPI-2 and MMPI-2-RF) seems to be passing this hurdle with fairly little difficulty.

Do the expert's conclusions reasonably follow from applying the technique to the case?

RPAS

We think that RPAS will be easier to explain to judges, lawyers, and jurors than many other psychological tests and more persuasive when correctly used because:

a) Easier to read Standard Score presentation

b) Organization by strength of the evidence (Page 1/Page 2)

Foundations of Interpretation Applied to a Case

S. Phillip Erdberg
Eva Christiansen

Case 1 – Brief History

- Case 1 is a middle-age male who is a HS graduate with interests more in work and solitary activities than sports or social gatherings. He married, had children, and divorced several years ago after years of strain and conflict. He remains close with his ex and has frequent contact due to their children's activities. Case 1 was to have married his fiancée recently but called it off shortly before the date. Case 1 made a suicide attempt following a public and hostile exchange from his ex-fiancée toward he and his ex-wife. He drank a great deal of alcohol, emailed a family member and said he was going to shut himself in the garage with the car running, and did so. He was hospitalized for three days and before discharge he was referred for a more complete assessment. He was guarded on self-report measures, wanting to be perceived as a respected, psychologically healthy person who acted atypically in the suicide attempt. But on the WAIS-IV and Rorschach his cooperation was good.

Card	#	R-Opt	Or	Loc	Loc #	Space	Content	Sy	Vg	2	FQ	Pop	Det	Cognitive	Thematic	ROD (RP Only)
I	1			W		Si	A				o	P	F			
I	2	Pr		W			A				o		F			
I	3			W			H	+			o		Mp,r		GHR	
II	4			D	5,3	Sri	NC	+			o		mp			
II	5			W			H,Cg	+		2	o		Mp,FC		MOAH,GHR	
II	6			D	1		A			2	u		FMp		AGC	
II	7	Pu		W			NC		v		u		CF,C'			
III	8			D	9,7		H,NC	+		2	o	P	Mp		GHR	ROD
III	9			D	1		Xy				o		F			
III	10			W		Si	(Hd)				-		CF,C'		PHR	
IV	11			D	7		Cg			2	u		Y			
IV	12			D	7		H				o	P	FD		GHR	
IV	13			W			Ad				o		T			
IV	14	Pu		Dd	99		NC				-		F			
V	15			W			A				o	P	F			
V	16			W			A				o		F			
V	17			W			A				o	P	F			
V	18			W			Ad,Cg				o		F			
VI	19			W			Cg				o		F		MOR	
VI	20			W			NC	+			o		V			
VI	21			W			A				-		F			
VII	22			W		Si	NC	+	v		o		F			
VII	23			D	7	Sr	Ay				u		F		AGC	
VII	24			D	1		A	+			u		FMp			
VIII	25			W			A,NC	+		2	o	P	FMa			
VIII	26			W			Cg	+			-		F			
VIII	27			W			NC				o		FC			
IX	28			W			Art				o		CF		PER	
IX	29			D	3		(H),Cg	+		2	-	P	Ma		MOR,AGC,COP,PHR	
IX	30			W			NC	+	v		u		CF			
X	31			Dd	99		NC				-		F			
X	32			D	9		A				u		F			
X	33			Dd	99,8,7		A			2	u		F			
X	34	Pu		D	4,9		NC		v		-		CF			

Domain	Counts	Counts	Calculations
Resp. & Adm.	R = 34	R8910 = 10	R8910% = 29%
	Prompts = 1	Pulls = 3	
	< or > = 0		
Location	W = 20	Dd = 3	W% = 59%
	D = 11	WD = 31	WD% = 91%
Space	Sr = 1	Sri = 1	SR = 2
	Si = 3	SumS = 5	SI = 4
Content	H = 4	Art = 1	AllH = 6
	(H) = 1	Ay = 1	NPH = 2
	Hd = 0	Bl = 0	PureH/AllH = 67%
	(Hd) = 1	Cg = 6	An+Xy = 1
	A = 11	Ex = 0	2ABS+Art+Ay = 2
	(A) = 0	Fi = 0	CritCont = 3
	Ad = 2	Sx = 0	DramCont = 2
	(Ad) = 0	Xy = 1	TCI = 2
	An = 0	NC = 11	TCI% = 6%
Syn	All + = 11		Syn% = 32%
Vague	All v = 4		Vague% = 12%
Pairs	Pair = 7		
Form Quality and Popular	FQo = 19	WDo = 19	X+% = 56%
	FQu = 8	a = 7	Xu% = 24%
	FQ- = 7	M- = 1	X-% = 21%
	FQn = 0		WDA% = 84%
	Popular = 7		

Domain	Counts	Counts	Calculations	
Det.	M = 4	FC = 2	WSumC	= 6.0
Blends	FM = 3	CF = 5	FC/(FC+CF+C)	= 29%
Mp,r	m = 1	C = 0	EA	= 10.0
Mp,FC	C' = 2	Y = 1	M/(M+WSumC)	= 40%
CF,C'	T = 1	V = 1	SumShd	= 5
CF,C'	r = 1	FD = 1	m+Y	= 2
		F = 16	Form%	= 47%
			es	= 9
			EA-es	= 1.0
	a = 2	p = 6	a/(a+p)	= 25%
	Ma = 1	Mp = 3	Ma/(Ma+Mp)	= 25%
	Blends = 4	C-Sh Bl = 2	Blend%	= 12%
Cognitive Scores	DV1 (1) = 0	DV2 (2) = 0	WSum6 = 0	
	INC1 (2) = 0	INC2 (4) = 0	Level 2 = 0	
	DR1 (3) = 0	DR2 (6) = 0		
	FAB1 (4) = 0	FAB2 (7) = 0		
	ALOG (5) = 0	CON (7) = 0		
Thematic Scores	ABS = 0	PER = 1		
	MOR = 2	AGP = 0		
	AGC = 3	AGM = 0		
	COP = 1	ROD = 1	ROD%	= 3%
	MOAH = 1	MOAP = 0	MOAH/(MOAH+P)	= 100%
	GHR = 4	PHR = 2	GHR/(GHR+PHR)	= 67%
Other Calc.	EII-3 = -1.50	Complexity = 101		
	TP-Comp = 0.4	L/S/O = 51		
	V-Comp = 4.3	Det = 22		
	S-Comp = 6.0	Cont = 28		

RPAS

Name: Christiansen/Broberg Protocol		ID: Christiansen/Broberg Protocol			Age:		Gender: NA		Education:					
Domain/Variables	Raw Score	%ile	Stand. Score	Compl. Adj. SS	Standard Score Profile									
Engagement and Cognitive Processing					60	70	80	90	100	110	120	130	140	
Complexity	101	93	122											
Location/Space/Object Complexity	51	98	132											
Determinant Complexity	22	74	110											
Content Complexity	28	85	116											
Responses (R)	34	99	136	135										
Form% (Lambda=0.89)	47%	65	106	124										
Blend%	12%	42	97	69										
Sym%	32%	57	103	82										
EA	10.0	84	115	94										
EA-es	10	76	110	112										
N:C [N/(N+WSumC)]	40%	25	90	99										
N	4	59	103	82										
WSumC	6.0	90	119	111										
FC:CF+C [FC/(FC+CF+C)]	29%	29	92	92										
FC	2	59	104	91										
CF+C	5	94	123	116										
Perception and Thinking					60	70	80	90	100	110	120	130	140	
EII-3	-1.50	8	79	69										
Thought & Perception Composite (TP-Comp)	0.4	48	99	94										
WSum6	0	9	90	65										
Level 2 Cognitive Scores	0	41	97	97										
X-%	21%	61	104	103										
X+%	56%	62	105	107										
WDA%	94%	50	100	100										
Popular	7	79	112	109										
Stress and Distress					60	70	80	90	100	110	120	130	140	
m	1	45	99	77										
Y	1	51	100	100										
m+Y	2	43	97	76										
NOR	2	79	112	112										
Suicide Composite (S-Comp)	6.0	88	117	112										
Self & Other Representation					60	70	80	90	100	110	120	130	140	
ROD%	3%													
Space Reversal (SR)	2													
MOAH:MOAP [MOAH/(MOAH+MOAP)]	100%													
MOAH	1													
MOAP	0													
GHR:PHR [GHR/(GHR+PHR)]	67%	61	104	109										
GHR	4	57	103	88										
PHR	2	42	97	90										
AGC	3													
AGP	0													
NOR	2	79	112	112										
Vigilance Composite (V-Comp)	2.9	42	97	75										
COP	1	56	102	96										

Note: All ratios are computed as the numerator over the sum of numerator and denominator

Domain/Variables	Raw Score	%ile	Stand. Score	Compl. Adj. SS	Standard Score Profile
Engagement and Cognitive Processing					60 70 80 90 100 110 120 130 140
Location Array					
W	20	98	129	125	
D	11	64	105	109	
Dd	3	55	102	102	
W:DDd [W/(W+D+Dd)]	59%	79	112	105	
WD:Dd [(W+D)/(W+D+Dd)]	91%	68	107	104	
Source Integration (SI)	4				
SumS	5	87	117	113	
Vague%	12%	82	114	114	
Prompts	1				
Pulls	3				
Sideways (< or >)	0				
V	1	75	110	110	
FD	1	60	104	99	
R99.0%	29%	34	94	96	
WSumC	6.0	90	119	111	
Pure C	0	38	95	95	
Na: Np [Na/(Na+Np)]	25%	19	87	85	
Na	1	32	93	67	
Np	3	83	114	106	
ZA:BS+Art+Ay	2	56	102	97	
Perception and Thinking					60 70 80 90 100 110 120 130 140
Xu%	24%	41	96	95	
Stress and Distress					60 70 80 90 100 110 120 130 140
es	9	58	103	79	
V	1	75	110	110	
SumShd	5	67	106	93	
C-Sh Bl	2	90	120	116	
C'	2	65	106	97	
FM	3	55	102	92	
Critical Contents (CritCont)	3	38	95	81	
Dramatic Contents (DramCont)	2	47	99	84	
Traumatic Content Index Percent (TCI%)	6%	25	90	84	
Self & Other Representation					60 70 80 90 100 110 120 130 140
Mp-	1	74	110	110	
All Human Content	6	56	102	93	
r	1	83	114	114	
a:p [a/(a+p)]	25%	9	80	80	
a	2	19	87	62	
p	6	85	115	105	
AGM	0	32	93	93	
T	1	71	109	109	
Pure H:NPH [H/(H+(H)+Hd+(Hd))]	67%	84	115	113	
Pure H	4	82	113	100	
NPH [(H)+Hd+(Hd)]	2	28	91	79	
PER	1	73	109	109	
FD	1	60	104	99	
An+Xy	1	52	101	101	