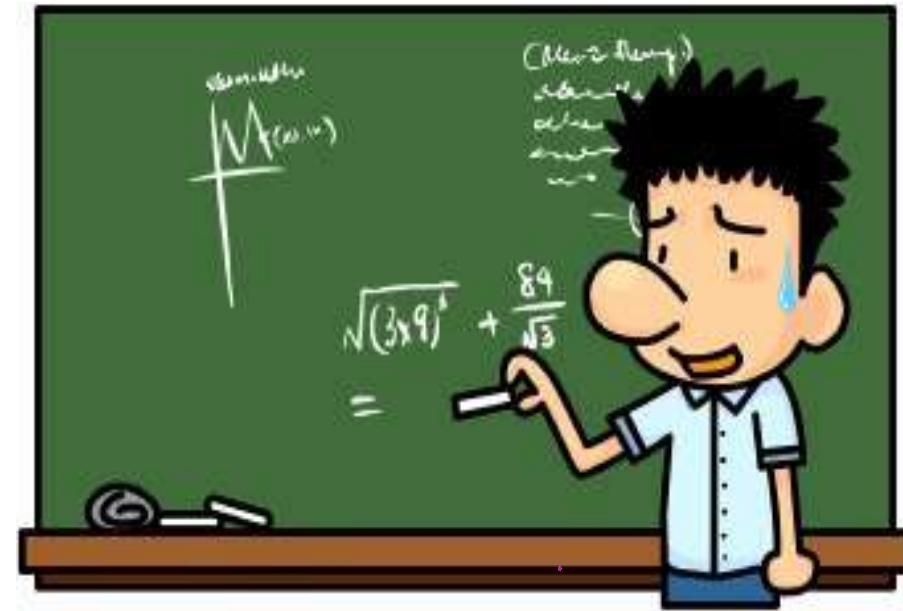# Psychometric evaluation of Measures

Ron D. Hays, Ph.D.

December 11, 2024 (2:30-4:00 pm)

Questionnaire Design and Testing Workshop

9th Floor Conference Room, 1100 Glendon Ave, Los Angeles, CA 90024

Impact Stratification Score (ISS)

- ISS Parts
  - PROMIS-29 physical function (4 items, 1-5)
  - PROMIS-29 pain interference (4 items, 1-5)
  - PROMIS-29 pain intensity (1 item, 0-10)

- ISS Score
  - A higher score is worse
  - Possible range: 8-50
    - Mild: 8-27
    - Moderate: 28-34
    - Severe: 35-50

Deyo, R. A., Dworkin, S. F., et al. (2014). Report of The NIH Task Force on research standards for chronic low back pain. Spine, 39(14), 1128-1143.

PROMIS® Impact Stratification Score

Please respond to each question or statement by marking one box per row.

Physical Function	Without any difficulty	With a little difficulty	With some difficulty	With much difficulty	Unable to do
PFA11 — Are you able to do chores such as vacuuming or yard work?	☐ 5 ①	☐ 4 ②	☐ 3 ③	☐ 2 ④	☒ 1 ⑤
PFA21 — Are you able to go up and down stairs at a normal pace?	☐ 5 ①	☐ 4 ②	☐ 3 ③	☐ 2 ④	☒ 1 ⑤
PFA23 — Are you able to go for a walk of at least 15 minutes?	☐ 5 ①	☐ 4 ②	☐ 3 ③	☐ 2 ④	☒ 1 ⑤
PFA53 — Are you able to run errands and shop?	☐ 5 ①	☐ 4 ②	☐ 3 ③	☐ 2 ④	☒ 1 ⑤

Pain Interference — In the past 7 days...	Not at all	A little bit	Somewhat	Quite a bit	Very much
PAININ9 — How much did pain interfere with your day to day activities?	☒ 1	☐ 2	☐ 3	☐ 4	☐ 5
PAININ22 — How much did pain interfere with work around the home?	☒ 1	☐ 2	☐ 3	☐ 4	☐ 5
PAININ31 — How much did pain interfere with your ability to participate in social activities?	☒ 1	☐ 2	☐ 3	☐ 4	☐ 5
PAININ34 — How much did pain interfere with your household chores?	☒ 1	☐ 2	☐ 3	☐ 4	☐ 5

Pain Intensity
In the past 7 days...

Global7 — How would you rate your pain on average?

| ☐ 0 No pain | ☐ 1 | ☐ 2 | ☐ 3 | ☐ 4 | ☐ 5 | ☐ 6 | ☐ 7 | ☐ 8 | ☐ 9 | ☒ 10 Worst pain imaginable |

$$(5+5+5+5) + (1+1+1+1) + 10 =$$
$$20 \quad + \quad 4 \quad + 10 = \boxed{34}$$

Psychometrics

- Item frequencies
- Item-scale correlations
- Factor analysis
- Reliability
- Scale means, medians, SD, % floor/ceiling
- Associations with other measures
- Item response theory

ACT: ClinicalTrials.gov (NCT01692275)

- **Objective**. This study examines Patient-Reported Outcome Measurement Information System (PROMIS)-29 v1.0 outcomes of chiropractic care in a multi-site, pragmatic clinical trial and compares the PROMIS measures to: 1) worst pain intensity from a numerical pain rating 0–10 scale, 2) 24-item Roland-Morris Disability Questionnaire (RMDQ); and 3) global improvement (modified visual analog scale). Design. A pragmatic, prospective, multisite, parallel-group comparative effectiveness clinical trial comparing usual medical care (UMC) with UMC plus chiropractic care (UMC+CC).

- **Setting**. Three military treatment facilities

- **Subjects**. 750 active-duty military personnel with low back pain.

- **Methods**. Linear mixed effects regression models estimated the treatment group differences. Coefficient of repeatability to estimate significant individual change. Results. We found statistically significant mean group differences favoring UMC+CC for all PROMIS VR-29 scales and the RMDQ score. Area under the curve estimates for global improvement for the PROMISVR -29 scales and the RMDQ, ranged from 0.79 to 0.83.

- **Conclusions**. Findings from this preplanned secondary analysis demonstrate that chiropractic care impacts health-related quality of life beyond pain and pain-related disability. Further, comparable findings were found between the 24-item RMDQ and the PROMIS-29 v1.0 briefer scales.

Hays, R. D., Shannon, Z. K., Long, C. R., Spritzer, K. L, Vining, R. D., Coulter, I. D., Pohlman, K. A., Walter, J.A., & Goertz, C.M. (2022). Health-related quality of life among United States service members with low back pain receiving usual care plus chiropractic care vs usual care alone: Secondary outcomes of a pragmatic clinical trial. Pain Med, 23(9), 1550-1559.

Baseline Item Frequencies in ACT Study (n= 749)

Response Option

ISS Item	1	2	3	4	5
Rating of pain on average*	1	23	51	25	<1
Pain interfered with your <u>day to day</u> activities**	7	31	26	28	8
Pain interfered with work around the house**	13	31	26	23	7
Pain interfered with household chores**	20	30	24	20	6
Pain interfered with ability to participate in social activities**	24	25	24	20	7
Able to do chores such as vacuuming or yard work***	26	34	29	10	1
Able to go up and down stairs at a normal pace***	32	33	23	9	2
Able to run errands and shop***	39	35	20	5	1
Able to go for a walk of at least 15 minutes***	45	31	17	5	1

* Recoded as 0 (no pain) -> 1, 1-3 -> 2, 4-6 ->3, 7-9 ->4, and 10 (Worst pain imaginable) -> 5.

** 1 = Not at all, 2 = A little bit; 3 = Somewhat, 4 = Quite a bit, 5 = Very much.

*** 1= Without any difficulty, 2 = With a little difficulty, 3 = With some difficulty, 4 = With much difficulty, 5= Unable to do.

ACT: 749 active-duty military personnel with low back pain. Mean age = 31; 76% were male, and 67% were non-Hispanic White.

ISS Item-Scale Correlations

Item	Physical function	Pain interference
Able to do chores such as vacuuming or yard work.	0.71*	0.63
Able to go up and down stairs at a normal pace.	0.77*	0.61
Able to go for a walk of at least 15 minutes	0.79*	0.58
Able to run errands and shop	0.83*	0.63
Pain interfered with your day to day activities	0.62	0.84*
Pain interfered with work around the house	0.64	0.89*
Pain interfered with ability to participate in social activities	0.62	0.81*
Pain interfered with household chores	0.68	0.86*
Pain intensity item	0.50	0.58

ACT

Categorical Confirmatory Factor Analysis of the ISS Items

CFI = 0.80, RMSEA = 0.27
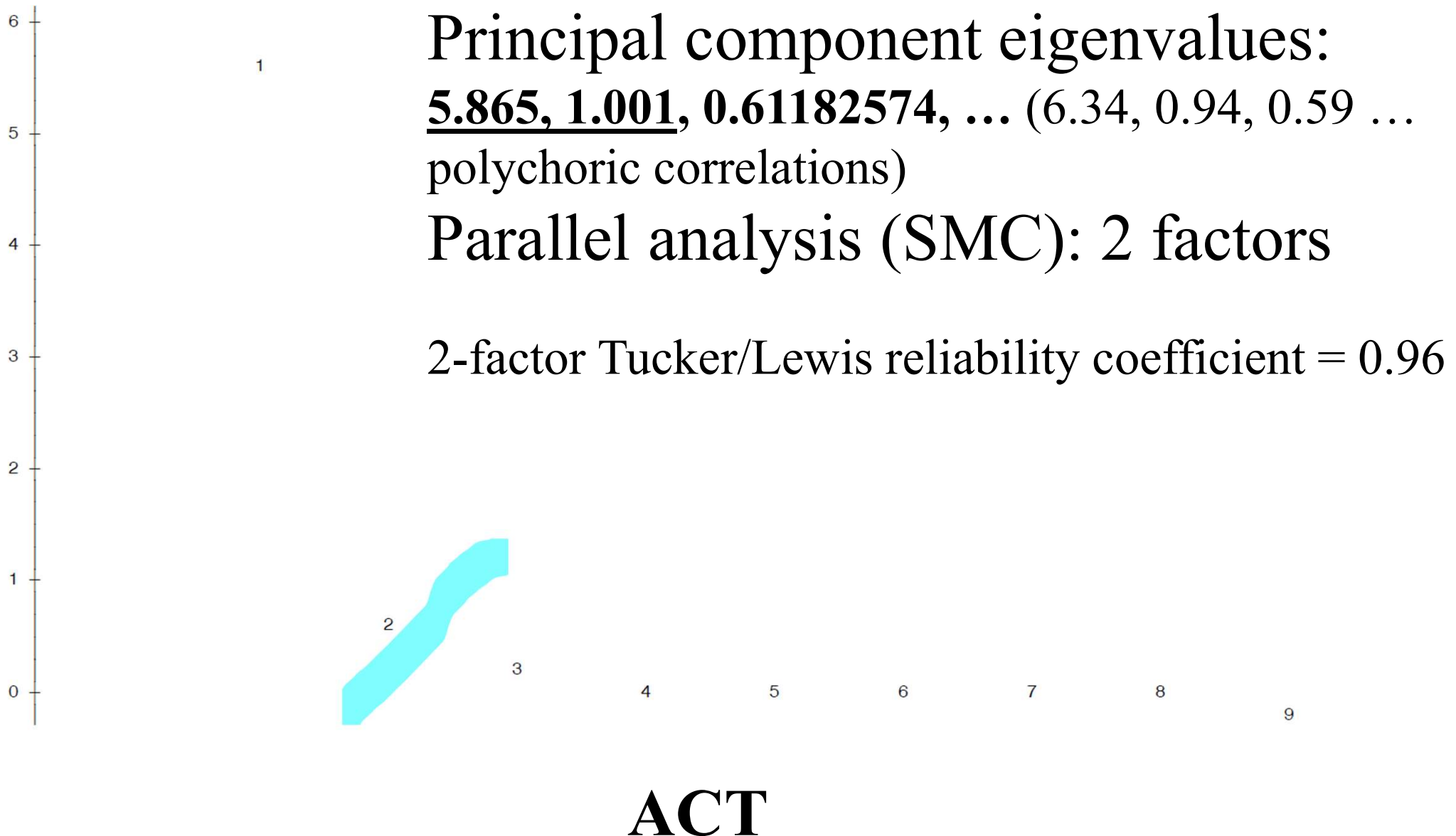
Rules of Thumb: CFI >=0.95, RMSEA <=0.06
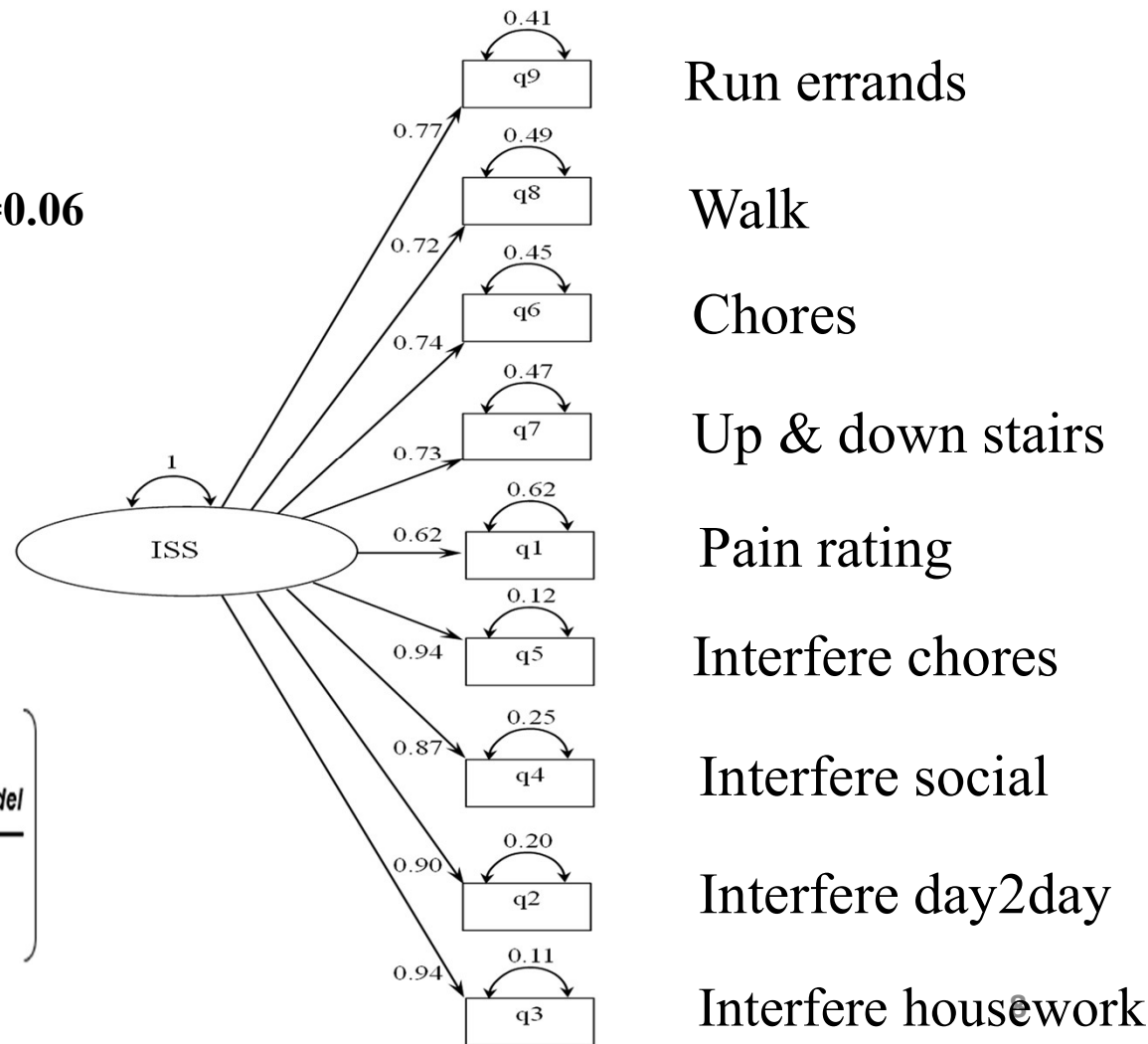
- Comparative fit index (CFI):

$$1 - \left(\frac{\chi^2_{model} - df_{model}}{\chi^2_{null} - df_{null}} \right)$$

$$RMSEA = SQRT\ (\lambda^2 - df)/SQRT\ (df\ (N-1))$$

Root mean square error of approximation

Exploratory Factor Analysis: Scree Plot

Principal component eigenvalues: **5.865, 1.001**, **0.61182574, …** (6.34, 0.94, 0.59 … polychoric correlations)

Parallel analysis (SMC): 2 factors

2-factor Tucker/Lewis reliability coefficient = 0.96

ACT

Promax Rotated Factor Loadings

Items	Physical function	Pain interference
Able to go for a walk of at least 15 minute	0.82	0.02
Able to run errands and shop	0.82	0.08
Able to go up and down stairs at a normal pace	0.74	0.12
Able to do chores such as vacuuming or yard work	0.62	0.21
Pain interfered with work around the house	0.02	0.92
Pain interfered with day to day activities	0.05	0.84
Pain interfered with household chores	0.12	0.82
Pain interfered with ability to participate social act.	0.12	0.75
Pain intensity	0.19	0.47

Factor correlation of physical function with pain interference = 0.66

Note: Correlation between these simple-summated scales = 0.70

ACT

Table 3 Exploratory factor analysis factor loadings for one factor, two correlated factors, and bifactor model

Item #	1 factor λ	2 factors PII-λ	PF-λ	Bifactor Gen—λ	PII—λ	PF -λ
1. Chores	0.70	-	0.52	0.64	-	0.30
2. Stairs	0.67	-	0.77	0.65	-	0.44
3. Walk15	0.67	-	0.89	0.67	-	0.50
4. Errands	0.75	-	0.80	0.72	-	0.45
5. Interfere daily	0.83	0.93	-	0.72	0.53	-
6. Interfere home	0.89	0.97	-	0.77	0.55	-
7. Interfere social	0.80	0.74	-	0.69	0.42	-
8. Interfere chores	0.87	0.82	-	0.75	0.46	-
9. Pain intensity	0.68	0.73	-	0.58	0.41	-

Factor loadings are denoted (λ)

Gen =General factor (i.e., impact), *PII* =Pain interference/intensity factor, *PF* =Physical function factor. EFA models estimated using minimum residual extraction with an oblimin rotation. Cross-loading values ($<.20$) are denoted (-) for ease of interpretation

Rodriguez et al. (2022)
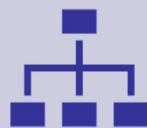
Types of Reliability

Inter-rater (rater)

Need 2 or more raters of the thing being measured

Test-retest (administrations)

Need 2 or more time points

Internal consistency (items)

Need 2 or more items

Test-retest Reliability of MMPI 317-362 (r = 0.75)

0.9 ≤ α	Excellent
0.8 ≤ α < 0.9	Good
0.7 ≤ α < 0.8	Acceptable
0.6 ≤ α < 0.7	Questionable
0.5 ≤ α < 0.6	Poor
α < 0.5	Unacceptable

MMPI 317

	True	False	
True	169	15 (14%)	184
False	21 (11%)	95	116
	190	110	

MMPI 362

I am more sensitive than most other people.

Hays, R. D., & Revetto, J. P. (1992). Old and new MMPI-derived scales and the Short-MAST as screening tools for alcohol disorder. Alcohol and Alcoholism, 27, 685-695.

ANOVA formulas to estimate reliability

Model	Reliability	Intraclass Correlation (single measure)
One-way	$$\frac{MS_{BMS} - MS_{WMS}}{MS_{BMS}}$$	$$\frac{MS_{BMS} - MS_{WMS}}{MS_{BMS} + (k-1)MS_{WMS}}$$
Two-way mixed	$$\frac{MS_{BMS} - MS_{EMS}}{MS_{BMS}}$$	$$\frac{MS_{BMS} - MS_{EMS}}{MS_{BMS} + (k-1)MS_{EMS}}$$
Two-way random	$$\frac{N(MS_{BMS} - MS_{EMS})}{NMS_{BMS} + MS_{JMS} - MS_{EMS}}$$	$$\frac{MS_{BMS} - MS_{EMS}}{MS_{BMS} + (k-1)MS_{EMS} + k(MS_{JMS} - MS_{EMS})/N}$$

Note: While coefficient alpha has been referred to as an intra-class correlation (consistency coefficient for average measures), the intraclass correlation column in Table 1 indicates estimates for a single measure (e.g., item, rater, time point).

BMS = Between Ratee Mean Square
WMS = Within Mean Square
JMS = Item, Rater, or Time Mean Square
EMS = Ratee x Item, Rater, or Time Mean Square
N = Number of ratees
k = Number of items, raters, or time

Coefficient alpha = 0.92
in ACT Study at baseline.

Test-Retest Reliability Formulas

Model	Reliability	Intraclass Correlation (single measure)
One-way	$$\dfrac{MS_{BMS} - MS_{WMS}}{MS_{BMS}}$$	$$\dfrac{MS_{BMS} - MS_{WMS}}{MS_{BMS} + (k-1)MS_{WMS}}$$
Two-way mixed	$$\dfrac{MS_{BMS} - MS_{EMS}}{MS_{BMS}}$$	$$\dfrac{MS_{BMS} - MS_{EMS}}{MS_{BMS} + (k-1)MS_{EMS}}$$
Two-way random	$$\dfrac{N(MS_{BMS} - MS_{EMS})}{NMS_{BMS} + MS_{JMS} - MS_{EMS}}$$	$$\dfrac{MS_{BMS} - MS_{EMS}}{MS_{BMS} + (k-1)MS_{EMS} + k(MS_{JMS} - MS_{EMS})/N}$$

Note: While coefficient alpha has been referred to as an intra-class correlation (consistency coefficient for average measures), the intraclass correlation column in Table 1 indicates estimates for a single measure (e.g., item, rater, time point).

BMS = Between Ratee Mean Square

WMS = Within Mean Square

JMS = Item, Rater, or Time Mean Square

EMS = Ratee x Item, Rater, or Time Mean Square

N = Number of ratees

k = Number of items, raters, or time

Six-WeekTest-Retest Reliability of ISS in ACT Study

Limited to those who reported the were the same at 6 weeks compared to baseline.

Source	Degrees of freedom	Mean Square	Label for mean square
Ratees (N-1)	177	112.94	BMS
Within	178	15.56	WMS
Time (K-1)	1	201.75	JMX
Time x Ratees	177	14.50	EMS
Total	**355**		

$$\frac{112.94-14.50}{112.94+14.50}$$

98.44/127.44 = 0.77

Cronbach's alpha	Internal consistency
$0.9 \leq \alpha$	Excellent
$0.8 \leq \alpha < 0.9$	Good
$0.7 \leq \alpha < 0.8$	Acceptable
$0.6 \leq \alpha < 0.7$	Questionable
$0.5 \leq \alpha < 0.6$	Poor
$\alpha < 0.5$	Unacceptable

Note: 0.76 for random effects.

ISS Score Distribution in ACT Study
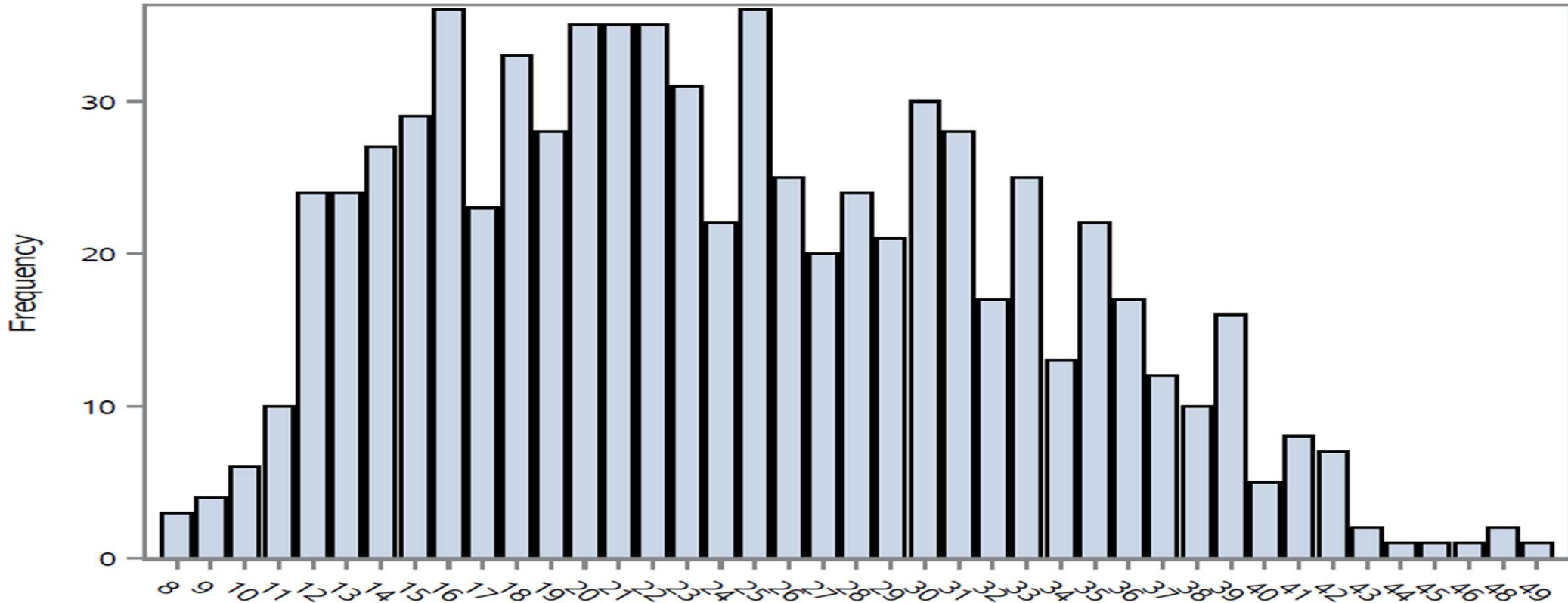
Analysis Variable : ISS0

N	Mean	Std Dev	Minimum	Maximum
749	24.2323097	8.3968719	8.0000000	49.0000000

8-50 is the possible range.

Median = <u>22.5,</u> Mode = <u>16</u> and <u>26</u>

Floor: <u>0.4%</u> and Ceiling: <u>0%</u>

Pain Intensity, Interference With Enjoyment of Life, Interference With General Activity (PEG) Scale

- What number best describes
  - your pain on average in the past week?
  - how, during the past week, pain has interfered with your enjoyment of life?
  - how, during the past week, pain has interfered with your general activity?

- 0-10 Response scale (10 = most severe pain)
- Mean scoring (0-10 possible range)

Standardized Factor Loadings for the ISS and 6 Other Pain Impact Measures Ranged from 0.78 (RMDQ) to 0.87 (ISS)

Impact Stratification Scale (ISS)	Graded Chronic Pain Scale (GCPS) disability score.	Oswestry Disability Index (ODI)
PEG (Pain intensity, interference with Enjoyment of life, interference with General activity)	Roland-Morris Disability Questionnaire (RMDQ)	Short form of the Orebro Musculoskeletal Pain Questionnaire (OMPQ)
	Subgroups for Targeted Treatment (STarT) Back Tool	

Hays, Herman et al. (2024)

Table 3 Product-moment correlations of the PEG with the PROMIS-29+2 v 2.1 scales and the Impact Stratification Score (n=1931)

PROMIS-29+2 Measures	PEG
Impact Stratification Score (ISS)	0.74
Pain intensity	0.70
Pain interference	0.68
Physical health summary score	—0.62
PROPr	—0.59
Mental health summary score	—0.58
Physical function	—0.57
Ability to participate in social roles and activities	—0.56
Fatigue	0.43
Anxiety	0.42
Depression	0.39
Sleep disturbance	0.28
Cognitive function	—0.27

NOTE. Higher scores mean better physical function, ability to participate in social roles and activities, cognitive function, physical health summary score, mental health summary score, and PROPr. Higher scores on the other measures indicate worse health.

Hays, Qureshi et al., 2023

Standardized Confirmatory Factor Loading Matrix for PROMIS-29+2, EQ-5D-5L, Personal Well-being, and Social Isolation Measures from the Bifactor Model (Estimates from Model Excluding EQ-5D-5L shown within parentheses)
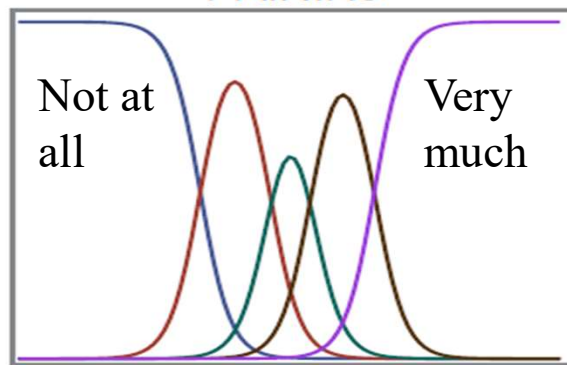
Scale	General Health	Physical Health	Mental Health
Fatigue	-.81 (-.81)		
Ability to participate in social roles and activities	0.78 (0.78)		
Depression	-.77 (-.77)		-.38 (-.35)
Personal well-being	0.75 (0.75)		0.30 (0.31)
Anxiety	-.72 (-.71)		-.28 (-.25)
Social isolation	-.69 (-.68)		-.44 (-.48)
Sleep disturbance	-.69 (-.69)		
EQ-5D-5L	0.68		
Cognitive function	0.63 (0.63)		
Pain interference	-.62 (-.62)	-.70 (-.70)	
Pain intensity	-.52 (0.52)	-.57 (-.57)	
Physical function	0.52 (0.52)	0.55 (0.56)	

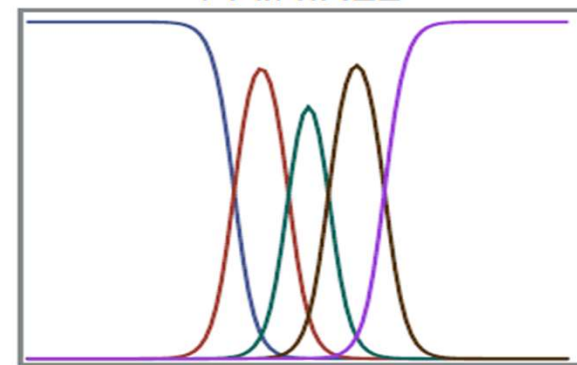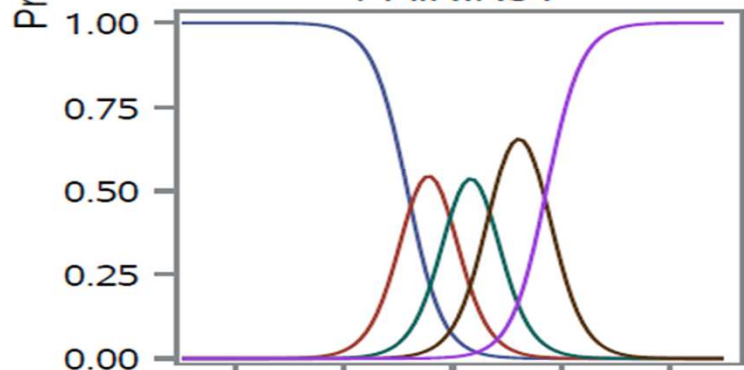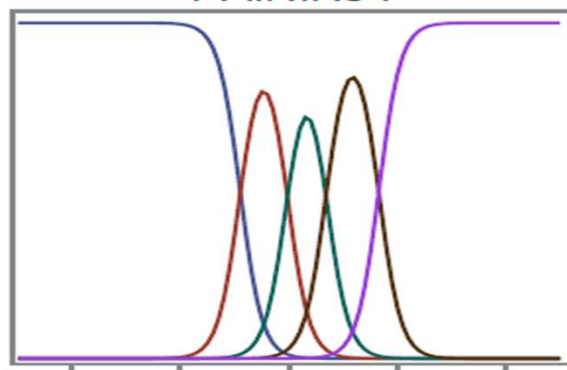Blank cells indicate that the loading was not estimated.
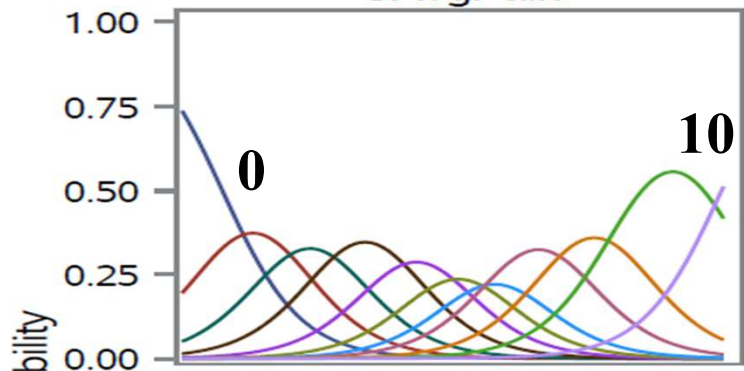
Hays, Rodriguez et al. (2024)

Test Information Curve

Reliability = 0.95

Reliability = 0.93

Reliability = 0.90

Reliability = 0.80

Rel = (INF-1)/INF

Information

ISS (ACT Study)

23
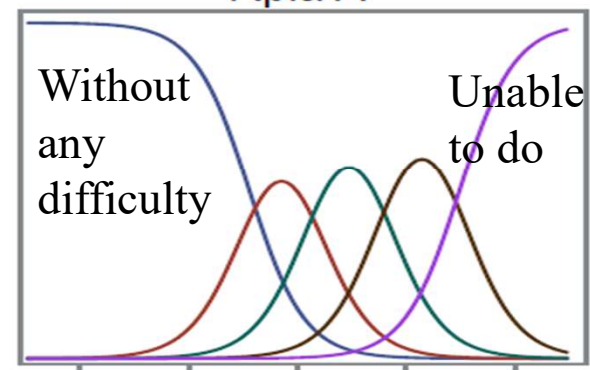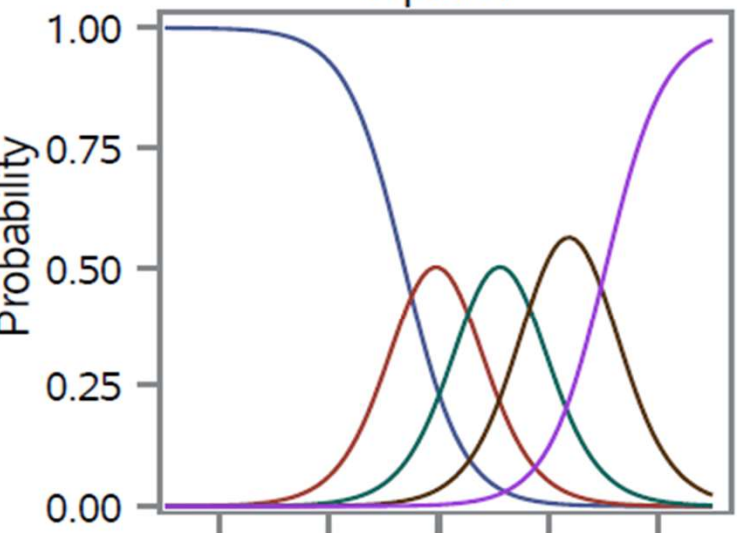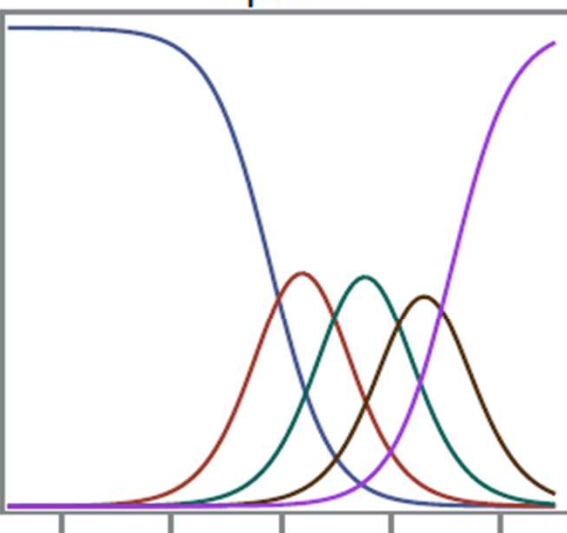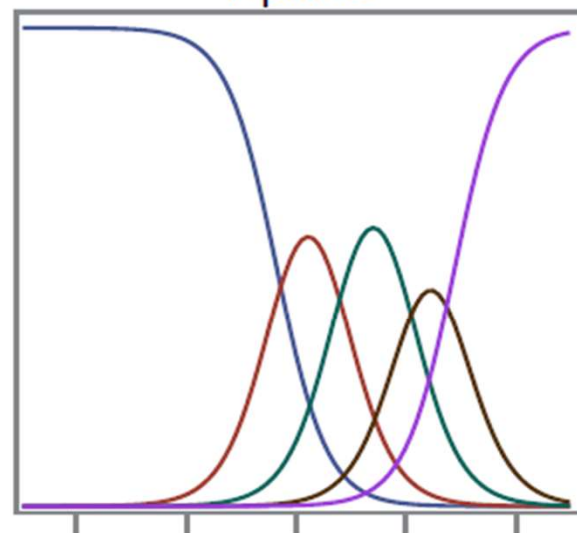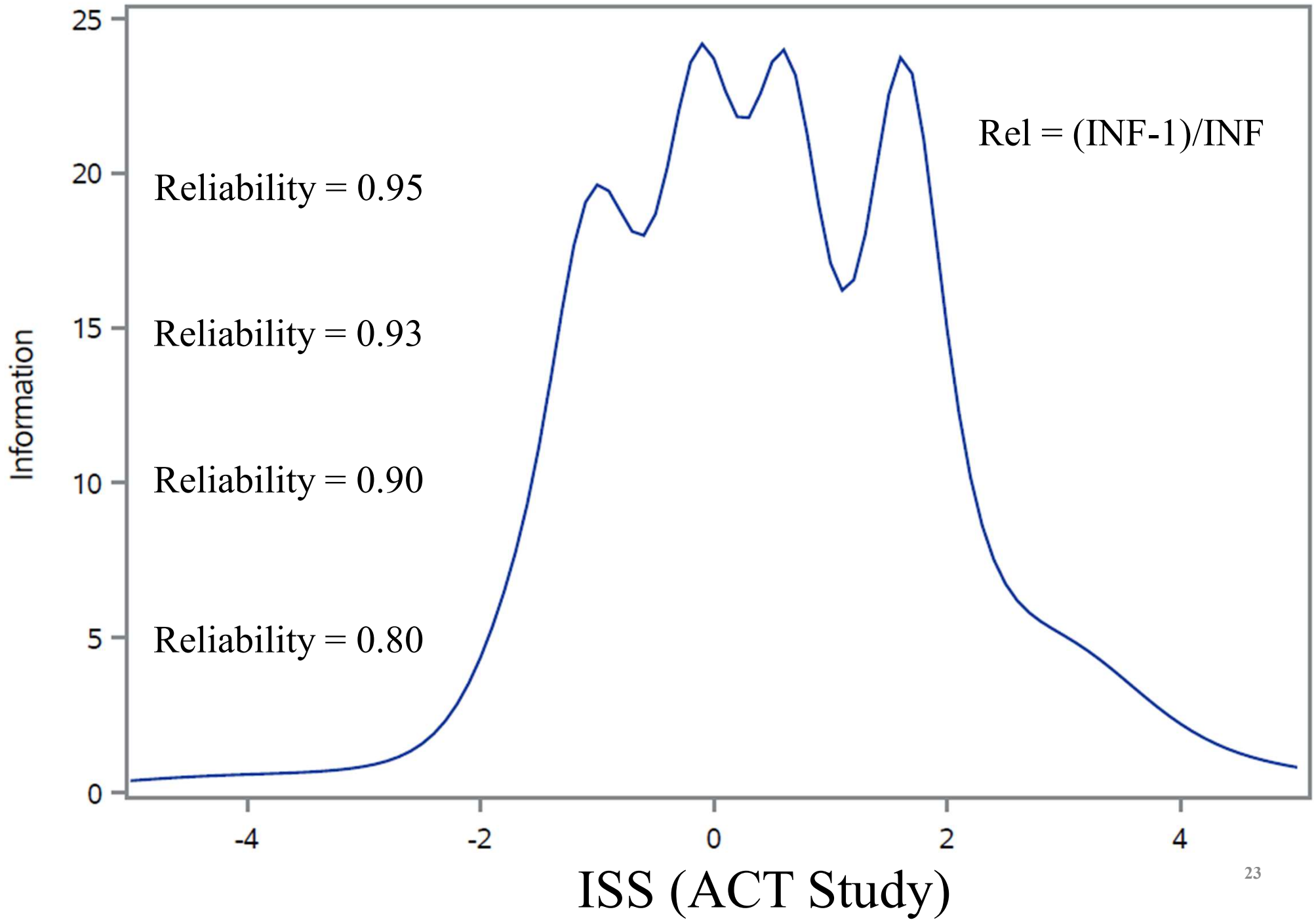
Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ISS0	749	24.23231	8.39687	18150	8.00000	49.00000
_Factor1	750	0.00835	0.95185	6.26600	-2.17936	3.11414

Pearson Correlation Coefficients
Prob > |r| under H0: Rho=0
Number of Observations

	ISS0	_Factor1
ISS0	1.00000	0.96983
		<.0001
	749	749
_Factor1	0.96983	1.00000
	<.0001	
	749	750

IRT graded response model estimates the relationship between a person's response Y_i to the question (i) and his or her level on the latent construct (θ):

b_{ik} estimates how difficult it is to score k or more on item (i). a_i estimates item discrimination.

$$\Pr(Y_i \geq k) = \frac{1}{1+\exp(-a_i\theta + b_{ik})}$$
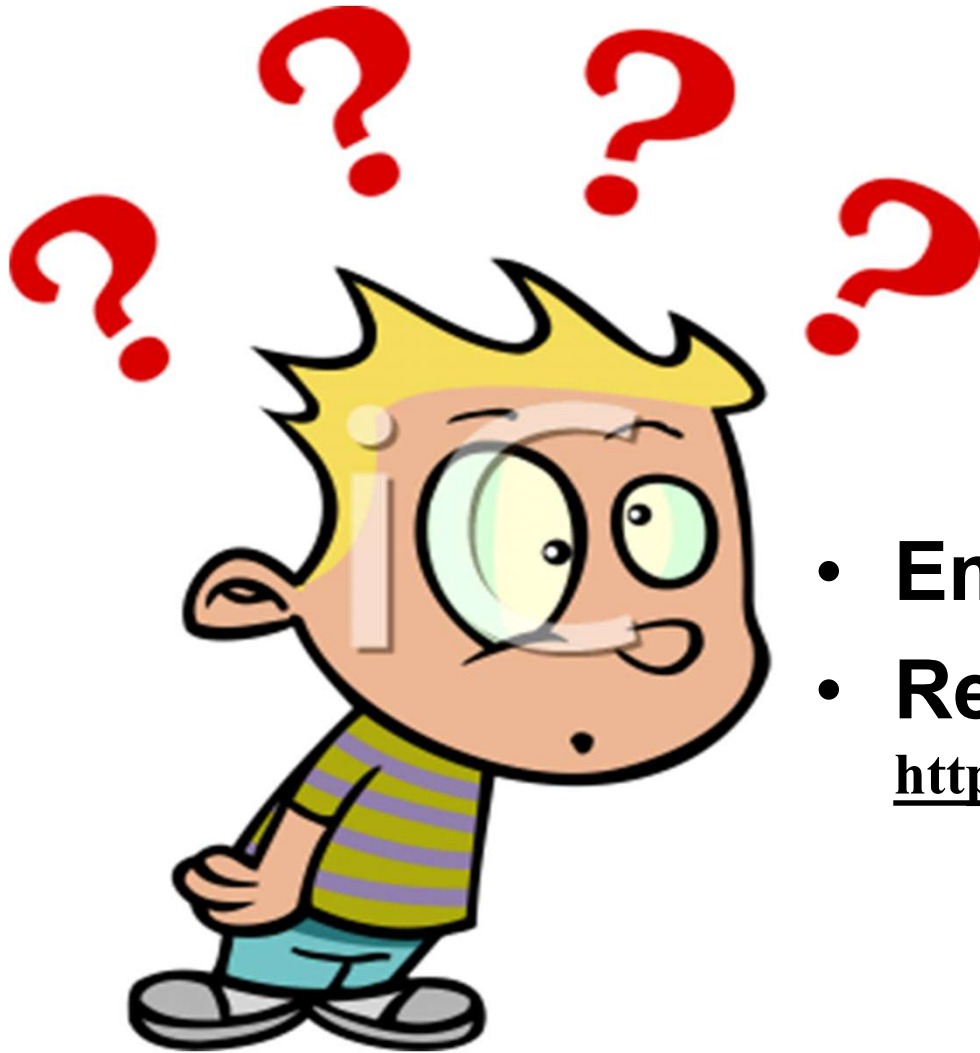
- **Email: drhays@ucla.edu**
- **Resources:**
  https://labs.dgsom.ucla.edu/hays/pages/

References

- Edelen, M. O., Rodriguez, A., Qureshi, N., Herman P. M., & **Hays, R. D**. (2023). Evaluation of the Impact Stratification Score in a sample of older adult patients with multiple chronic conditions. <u>J Geriatr Med Gerontol</u>, 9:141. doi.org/10.23937

- **Hays, R. D**., Edelen, M. O., Rodriguez, A., & Herman, P. (2021). Support for the reliability and validity of the National Institutes of Health Impact Stratification Score in a sample of active-duty U.S. military personnel with low back pain. <u>Pain Medicine</u>, <u>22</u>(10), 2185-2190. doi: 10.1093/pm/pnab175.

- **Hays, R. D**., Qureshi, N., Edelen, M., Rodriguez, M., Slaughter, M., & Herman, P. M. (2023). Crosswalking the National Institutes of Health Impact Stratification Score to the PEG. <u>Arch Phys Med Rehabil</u>, 104(3), 425-429. doi: 10.1016/j.apmr.2022.08.006

- **Hays, R. D**., Rodriguez, A., Qureshi, N., Zeng, C., & Edelen, M. O. (2024). Support for a single underlying dimension of self-reported health in a sample of adults with low back pain in the United States. <u>Applied Research in Quality of Life</u>, 19(5): 2213-2226. doi:10.1007/s11482-024-10327-8.

- Rodriguez, A., Edelen, M. O., Herman, P., & **Hays, R. D**. (2022). Unpacking the impact of chronic pain as measured by the Impact Stratification Score. <u>BMC Musculoskelet Disord</u>, 23: 884. https://doi.org/10.1186/s12891-022-05834-4

- Rodriguez, A., Herman, P. M., Slaughter, M.E., Edelen, M. O., & **Hays, R. D**. (2023). Classifying patients with non-specific low back pain using the Impact Stratification Score in an online convenience sample. <u>BMC Musculoskelet Disord</u>, 24(1):719. doi: 10.1186/s12891-023-06848-2.