

Scale Assessment: A Practical Example of a Rasch Testlet (Subtest) Approach

**Mike Horton¹ Alan Tennant² Alison Hammond³, Yeliz Prior³,
Sarah Tyson⁴, Ulla Nordenskiöld⁵**

1 Psychometric Laboratory for Health Sciences, University of Leeds, UK

2 Swiss Paraplegic Research, Nottwil, Switzerland

3 Centre for Health Sciences Research, University of Salford, UK

4 Nursing, Midwifery and Social Work, University of Manchester, UK

5 Sahlgrenska Academy, Göteborg University, Sweden

Background: the EDAQ

Patient Reported Outcome Measure (PROM)

Clinical

Assist:

- People to identify own abilities/difficulties in daily activities
- Occupational Therapists to help them find solutions

Research

- Evaluate occupational therapy, rehabilitation and pharmacological therapies in clinical trials (*eg Nordenskiold et al 1998*)
- Epidemiological studies (*eg TIRA: Thyberg et al 2004, 2005*)

Audit

- Identify service needs
- Evaluate services

Background: the EDAQ

- Originally developed in Sweden within a Rheumatoid Arthritis population (*Nordenskiold et al, 1996*)
- Adapted and validated for use in the UK (*Hammond et al, 2010*)

Scoring:

- 0-3 scale:
 - 0: Without Difficulty
 - 1: Some Difficulty
 - 2: Much Difficulty
 - 3: Unable
- Section A: ability WITHOUT altered methods, assistive devices or help
- Section B: ability WITH altered methods and/or assistive devices

EDAQ example page: Eating/Drinking (12 items)

EATING / DRINKING										
		A. How do you do it without an aid, alternate method or help?				B. How else do you do it with an aid or alternate method?				
		Without Difficulty (0)	Some Difficulty (1)	Much Difficulty (2)	Unable (3)	Describe which other aids or methods you use	Without Difficulty (0)	Some Difficulty (1)	Much Difficulty (2)	Unable (3)
1	Lift a glass	✓ (0)				n/a	(0)			
2	Lift a cup/mug		✓ (1)			2 hands	✓ (0)			
3	Use a knife and fork		✓ (1)			fat handled cutlery	✓ (0)			
4	Slice food (eg bread)			✓ (2)		Angled knife	✓ (0)			
5	Open jar			✓ (2)		Jar opener	✓ (0)			

Score A = 6


Score B = 0

“n/a” = not applicable; does not need to use any alternate method/aid

Background: the EDAQ

Scoring:

- 0-3 scale:
 - 0: Without Difficulty
 - 1: Some Difficulty
 - 2: Much Difficulty
 - 3: Unable
- **Scoring 'C'**: B scores (WITH altered methods/assistive devices) used where available; A score used where no B score is available



EDAQ (UK): 14 domains (138 activity items in total)

1. Eating & drinking (10 activities/items)
2. Going to the bathroom/personal care (12)
3. Getting dressed/undressed (11)
4. Bathing/showering (11)
5. Cooking (14)
6. Moving around Indoors (12)
7. Cleaning the house (9)
8. Laundry/clothes care (9)
9. Moving and transfers (6)
10. Communication (6)
11. Moving around outdoors/ shopping (13)
12. Gardening/ Household Maintenance (7)
13. Caring (9)
14. Hobbies and Leisure (9)

EDAQ in rheumatoid arthritis (RA)


- **Tested in RA in UK:**
- **Traditional Psychometric Testing:**
 - Internal consistency, test-retest reliability, concurrent and discriminative validity
- **Modern Psychometric Testing (Rasch Analysis using RUMM2030 software):**
 - Fit to the Rasch model, unidimensionality, response category ordering, response dependency, differential item functioning (DIF) (item bias), reliability, targeting

EDAQ in RA - Rasch analysis summary for EDAQ

(A scoring - without assistive devices or other methods)


Analysis #	Domain	Item Residual		Person Residual		Chi-Square Interaction			PSI	Unidimensionality	
		Mean	SD	Mean	SD	Value	(DF)	p		% t-test	(CI)
1	Eating	0.200	0.538	-0.491	1.023	17.41	15	0.30	0.89	7.1	4.8-9.5
2	Personal Care	-0.702	1.805	-0.604	0.921	14.56	15	0.48	0.77	1.6	-0.1-4.4
3	Dressing	-0.097	0.641	-0.494	1.090	15.83	28	0.73	0.84	3.1	0.1-5.5
4	Bathing	-0.338	1.260	-0.395	1.026	38.82	25	0.09	0.85	4.8	1.9-6.6
5	Cooking	-0.007	1.270	-0.405	0.934	22.39	20	0.32	0.90	3.4	0.9-5.8
6	Moving indoors	-0.635	2.684	-0.427	1.006	36.57	20	0.01	0.87	3.2	0.8-5.6
7	Cleaning	-0.617	1.252	-0.394	0.880	62.14	45	0.05	0.88	5.9	3.6-8.2
8	Laundry	0.053	0.753	-0.469	1.056	22.81	20	0.30	0.83	4.3	1.6-6.9
9	Transfers	0.028	1.578	-0.385	1.041	31.99	25	0.16	0.80	2.8	0.4-5.2
10	Communicate	0.011	1.205	-0.324	0.834	31.18	30	0.41	0.75	2.7	0.0-5.3
11	Moving outdoors	0.423	1.345	-0.402	1.038	36.29	25	0.07	0.84	4.4	2.1-6.7
12	House & Garden	-0.194	0.945	-0.466	0.670	12.06	10	0.28	0.91	1.9	-1.0-4.9
13	Caring	-0.411	1.358	-0.948	2.050	43.34	45	0.54	0.84	5.2	1.2-9.2
14	Hobbies	-0.073	1.264	-0.704	1.291	20.55	20	0.42	0.31	3.8	0.4-7.2
15	14 domains	-0.199	2.205	-0.355	1.116	75.44	70	0.31	0.94	10.0	7.8-12.3
16	Personal Care Component	-0.517	1.961	-0.434	1.045	28.72	35	0.76	0.91	4.5	2.1-6.8
17	Mobility Component	-0.357	1.099	-0.588	1.062	23.45	45	0.93	0.87	5.3	2.8-7.7
Fit Criteria		0.0	<1.4*	0.0	<1.4	> 0.05 [†]			>0.85	Lower CI < 5%	

- *Where testlets are used, this may be inflated;
- [†] Bonferroni adjusted (average is about 0.005)



EDAQ in rheumatoid arthritis (RA) – Rasch Analysis Summary

- Most domains function well, and have good psychometric properties
 - Almost all items within all domains have ordered response category thresholds (items with apparent disorder are marginal)
 - Overall fit statistics reflect good fit for almost all domains
 - Very few items displaying DIF by age, gender, or employment status
 - All domains displayed strict unidimensionality
 - Reliability (Person Separation Index) good – very good in almost all domains (very low in Hobbies domain)



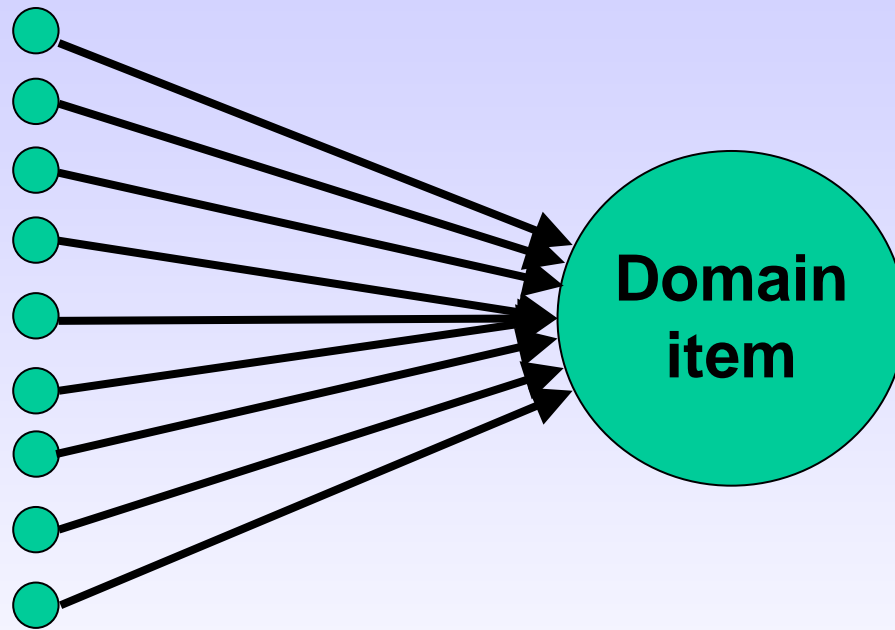
EDAQ in rheumatoid arthritis (RA) – Rasch Analysis Summary

- **However...**
- Issues with item response dependency within each domain (and especially when considered at the 138-item level)
- Individual problems with the ‘Caring’ and ‘Hobbies’ domains

A higher-order factor?

- Each domain treated as a testlet / subtest / 'super-item' to investigate whether the domains contribute to a single underlying higher-order factor
- All items within each domain are summed to create domain-level items
 - Takes account of the dependency within a set of items

A higher-order factor?




**Individual
items**

EDAQ (UK): 12 domains in 2 components:

Self-Care & Mobility

1. Eating & drinking (10 activities)
2. Going to the bathroom/personal care (12)
3. Getting dressed/undressed (11)
4. Bathing/showering (11)
5. Cooking (14)
6. Moving around Indoors (12)
7. Cleaning the house (9)
8. Laundry/clothes care (9)
9. Moving and transfers (6)
10. Communication (6)
11. Moving around outdoors/ shopping (13)
12. Gardening/ Household Maintenance (7)
13. Caring (9)
14. Hobbies and Leisure (9)



EDAQ in Musculoskeletal Conditions

Aim:

Test validity and reliability of EDAQ in musculoskeletal conditions (MSCs) in UK

- Base validation on 12-Domain, 2-Component structure












Musculoskeletal Conditions:

- Ankylosing Spondylitis (AS);
- Osteoarthritis (OA);
- Systemic Lupus (SLE);
- Systemic Sclerosis (SS);
- Chronic Pain (CP),
- Chronic Upper Limb conditions (CUL);
- Primary Sjogren's Syndrome (PSS).

Participants:

- 18 years +
- Able to read and write English
- Recruited from 20 Rheumatology departments; 10 organisations.

Sample Demographic data

Condition	n	male	female	M/F split	age mean (SD)
AS	165	118	45		54.7 (14.1)
Chronic Pain/FMS	194	24	170		53.3 (13)
Chronic hand upper limb conditions	157	48	109		55.5 (13.6)
Osteoarthritis	184	41	142		64.8 (9.8)
Systemic Lupus	164	7	157		54.3 (13.5)
Systemic Sclerosis	170	12	158		65.9 (10.1)
Primary Sjogrens	171	10	160		63.2 (11.1)
RA	383	97	286		60.4 (11.2)
Total	1588	357	1227		59.2 (12.8)

Self-Care Component

- Within each condition, overall fit is good
- **However...**
- Potential problem within CP sample
- A few isolated issues regarding individual item fit, response dependency and DIF (age & gender).

Self-Care Component

Analysis	n	Item Residual		Person Residual		Chi-Square Interaction			PSI	Unidimensionality	
		Mean	SD	Mean	SD	Value	(DF)	p		% t-test	(CI)
AS - Self Care	136	-0.34	1.09	-0.4	0.8	20.82	14	0.106	0.74	3.45%	*
CH - Self Care	130	-0.45	1.6	-0.44	0.88	16.06	14	0.31	0.77	4.20%	*
CP - Self Care	188	0.02	1	-0.45	1.17	14.06	14	0.45	0.89	9.88%	6.50%
OA - Self Care	170	-0.16	1.55	-0.43	1.03	20.1	14	0.127	0.82	5.77%	2.30%
RA - Self Care	365	-1.37	1.65	-0.56	1.04	37.45	35	0.357	0.88	2.46%	*
SJ - Self Care	142	-0.54	1.02	-0.46	0.85	20.65	14	0.111	0.75	1.69%	*
SLE - Self Care	147	-0.3	0.84	-0.39	0.87	13.83	14	0.462	0.85	6.11%	2.40%
SSc - Self Care	162	-0.29	0.84	-0.41	0.93	10.23	14	0.745	0.85	3.55%	*
All merged - Self Care	1440	-0.88	3.34	-0.47	1.06	132	63	1E-06	0.83	4.13%	*
Fit Criteria		0.0	<1.4*	0.0	<1.4			> 0.05†	>0.85		Lower CI < 5%

Mobility Component

- Within each condition, overall fit is good
- However...
- Again, a few isolated issues regarding individual item fit, response dependency and DIF (age & gender).

Mobility Component

Analysis	n	Item Residual		Person Residual		Chi-Square Interaction			PSI	Unidimensionality	
		Mean	SD	Mean	SD	Value	(DF)	p		% t-test	(CI)
AS - Mobility	155	-0.16	0.52	-0.53	0.95	6.89	10	0.736	0.86	4.07%	*
CH - Mobility	123	-0.2	1.02	-0.35	0.75	12.93	10	0.227	0.71	4.62%	*
CP - Mobility	189	-0.69	1.29	-0.71	1.03	4.36	10	0.93	0.85	8.05%	3.50%
OA - Mobility	168	-0.25	1.24	-0.49	0.9	5.89	10	0.824	0.83	5.08%	0.00%
RA - Mobility	358	-0.57	1.39	-0.5	0.94	23.13	25	0.57	0.85	6.58%	3.70%
SJ - Mobility	143	-0.29	0.82	-0.53	0.95	7.54	10	0.674	0.72	6.25%	0.00%
SLE - Mobility	146	-0.46	0.88	-0.43	0.75	8.48	10	0.582	0.82	1.49%	*
SSc - Mobility	148	-0.19	0.91	-0.44	0.87	13.47	10	0.199	0.79	5.08	0%
All merged - Mobility	1430	-0.29	3.59	-0.57	1.03	77.31	45	0.002	0.81	4.54%	*
Fit Criteria		0.0	<1.4*	0.0	<1.4			>0.05 [†]	>0.85	Lower CI < 5%	

Stability across MSCs (DIF-by-Condition)

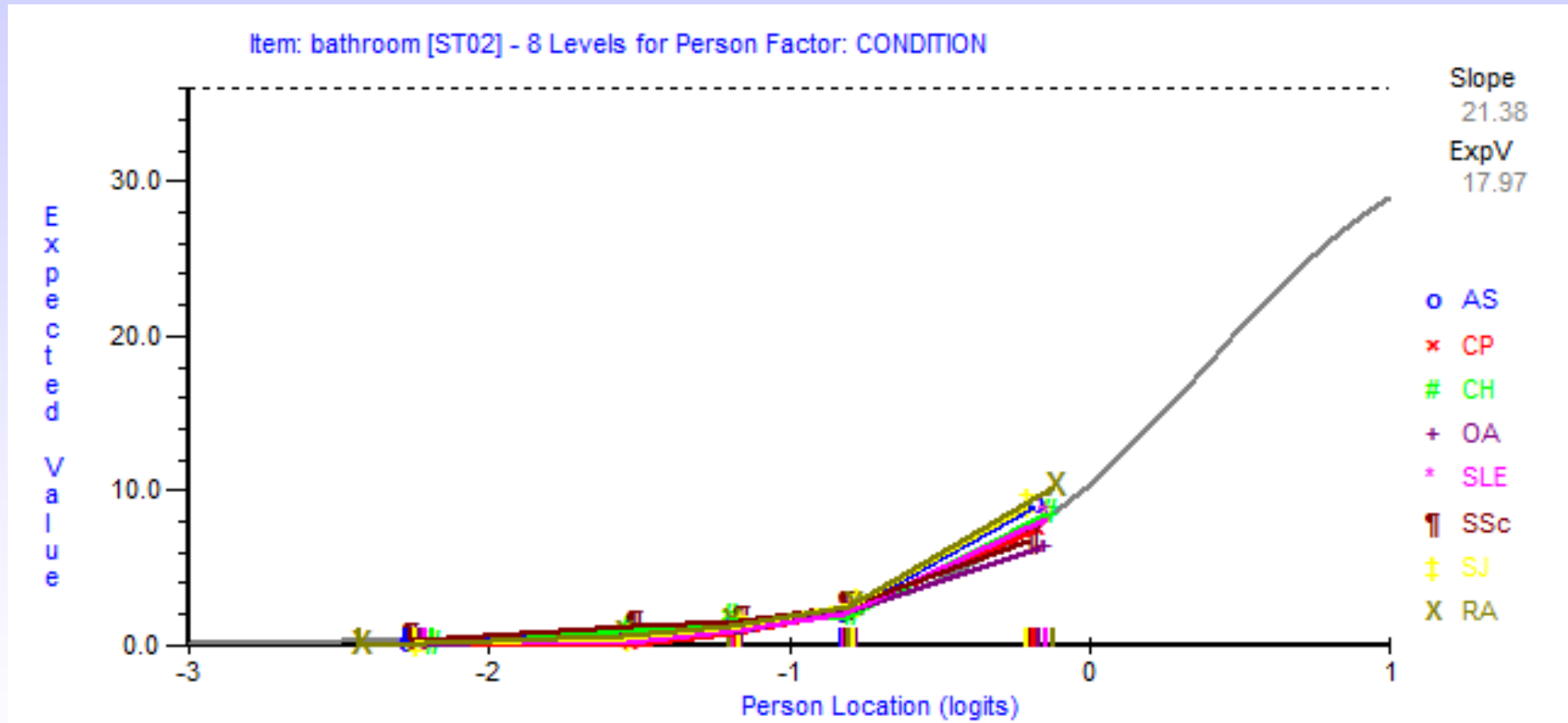
- Does the EDAQ work in the same way across conditions?
- If so, then EDAQ Component scores can be directly compared across conditions

Stability across MSCs (DIF-by-Condition)

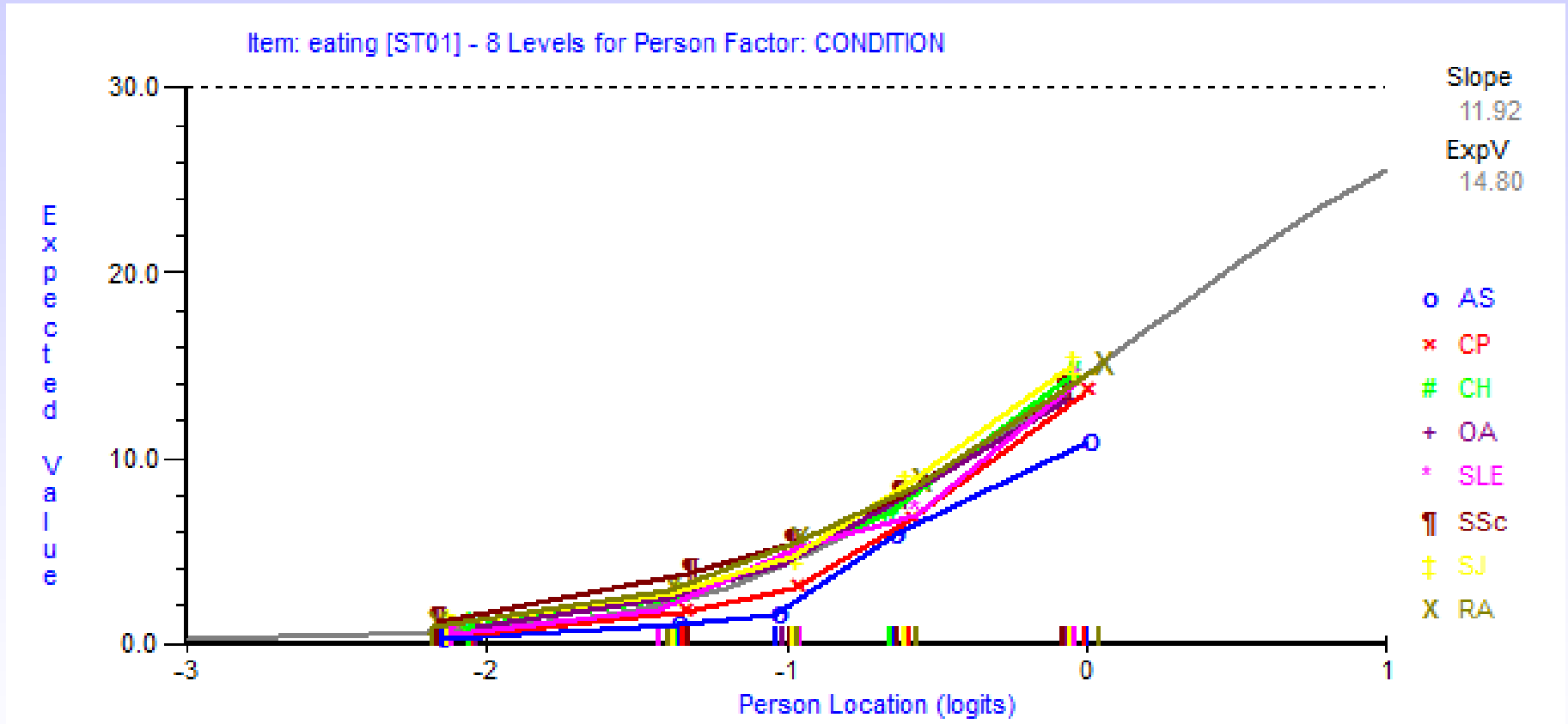
- Does the EDAQ work in the same way across conditions?

- In short, no!

Example of Item with no DIF



Example of Item displaying DIF



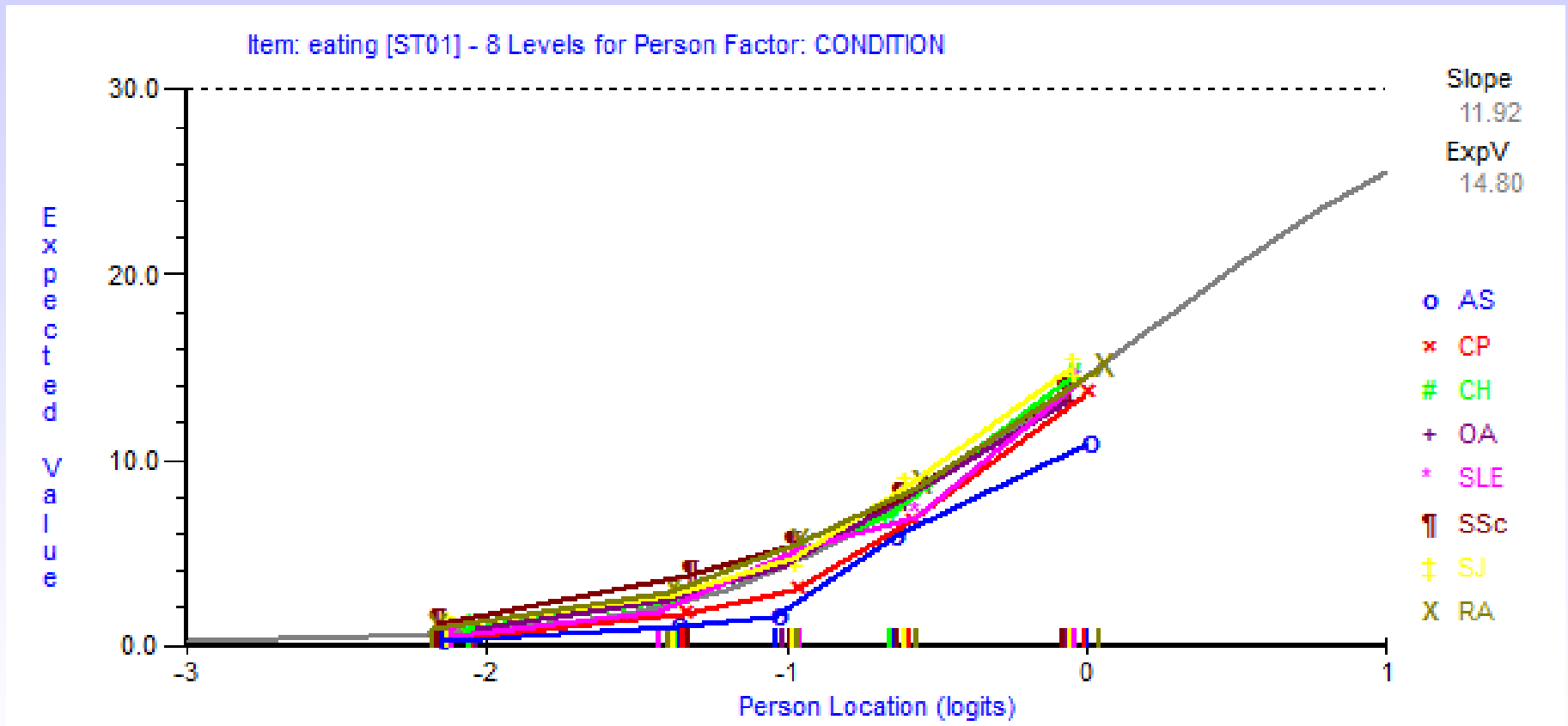
A logical explanation?

- Yes
- AS involves very little upper limb involvement, so eating activities would not necessarily be impaired for AS group (lower score = less impairment)

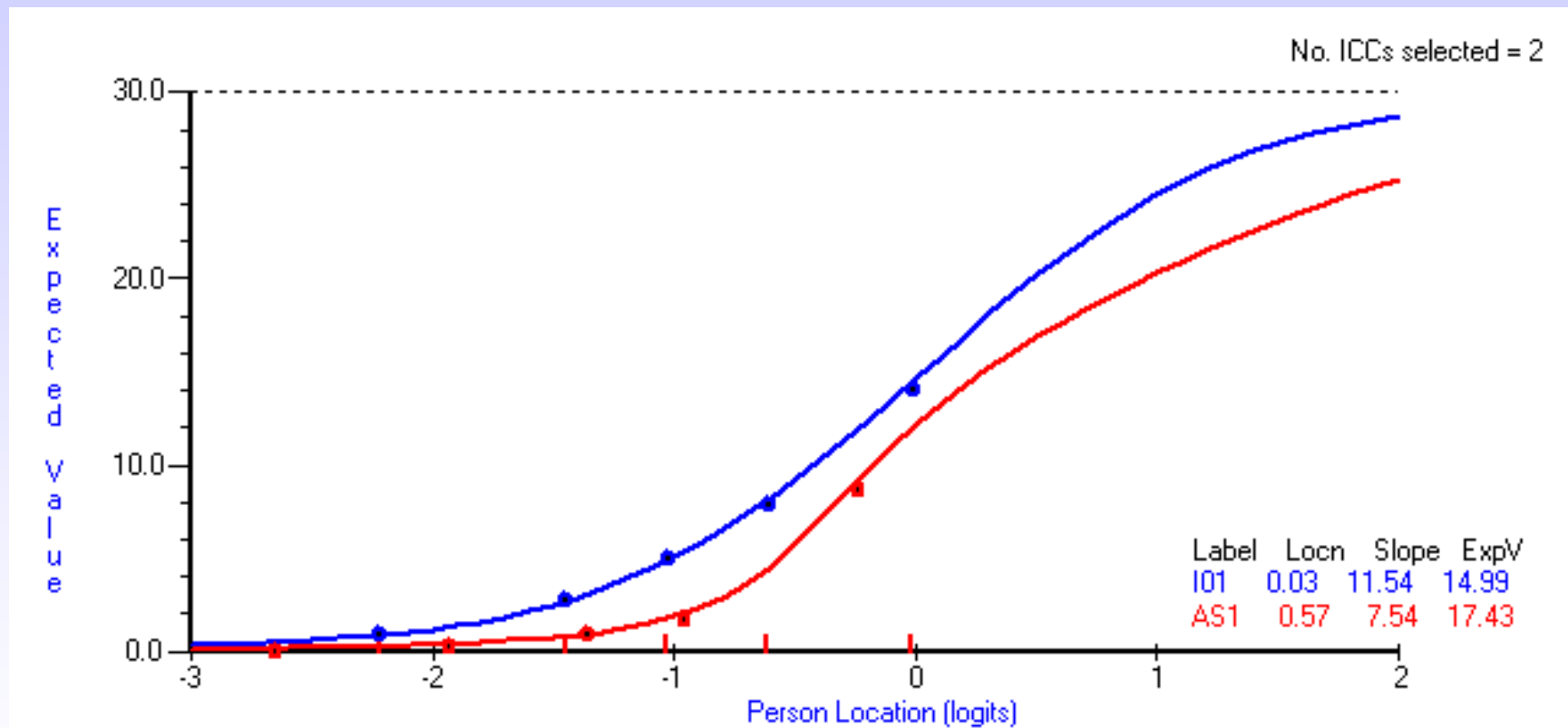
Stability across MSCs (DIF-by-Condition)

- Does the EDAQ work in the same way across conditions?
- In short, no!
- However, adjustments can be made to account for the differences through a DIF-splitting procedure...

Eating item displaying DIF



Item-split resolved DIF for eating item(s)



Self-Care Component - Equivalence

Original Domain	Domain Name	AS	CP	CH	OA	SLE	SSc	SJ	RA
1	eating								
2	bathroom								
3	dressed								
5	cooking								
7	clean								
8	laundry								
10	communication								

 =calibration

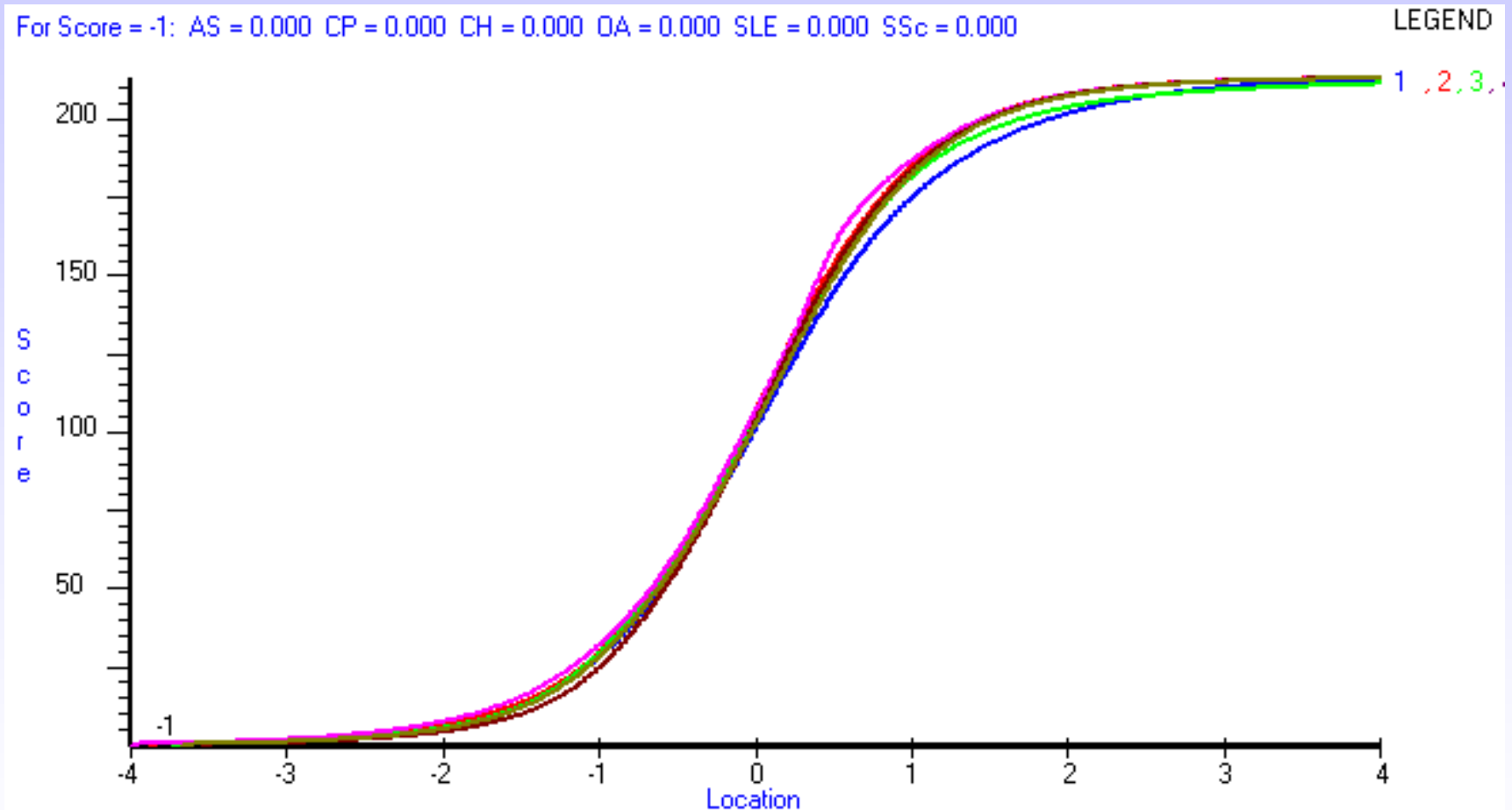
Self-Care Component - Equivalence

Original Domain	Domain Name	Split	AS	CP	CH	OA	SLE	SSc	SJ	RA
1	eating	AS1								
		CP1								
		I01								
2	bathroom	ST02								
3	dressed	AS3								
		SLE3								
		I003								
5	cooking	SSc5								
		CPSL5								
		I005								
7	clean	SSc7								
		CPSL7								
		I007								
8	laundry	SLE8								
		I008								
10	communication	CH10								
		I010								

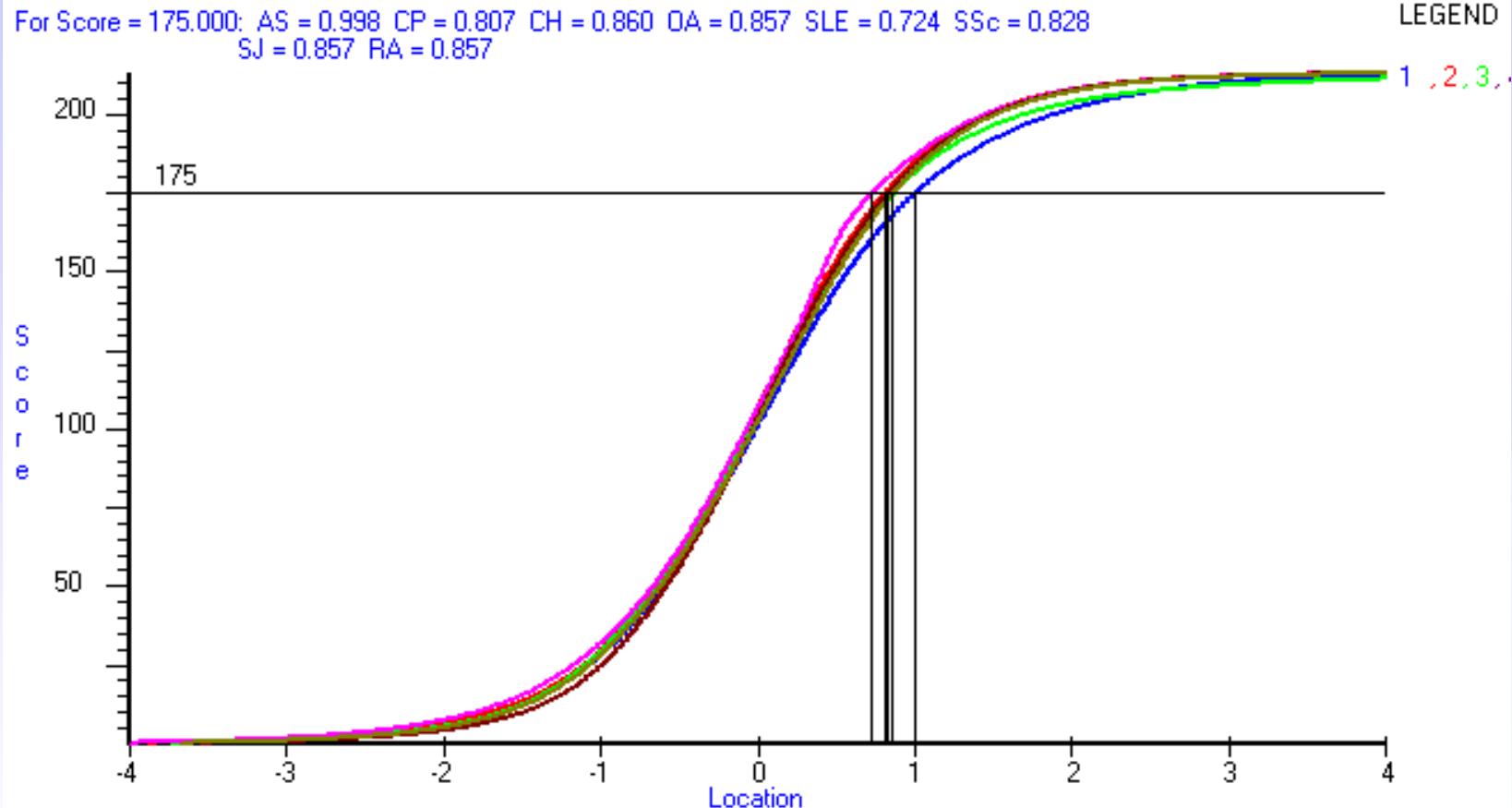
=calibration



Self-Care Component - Equivalence

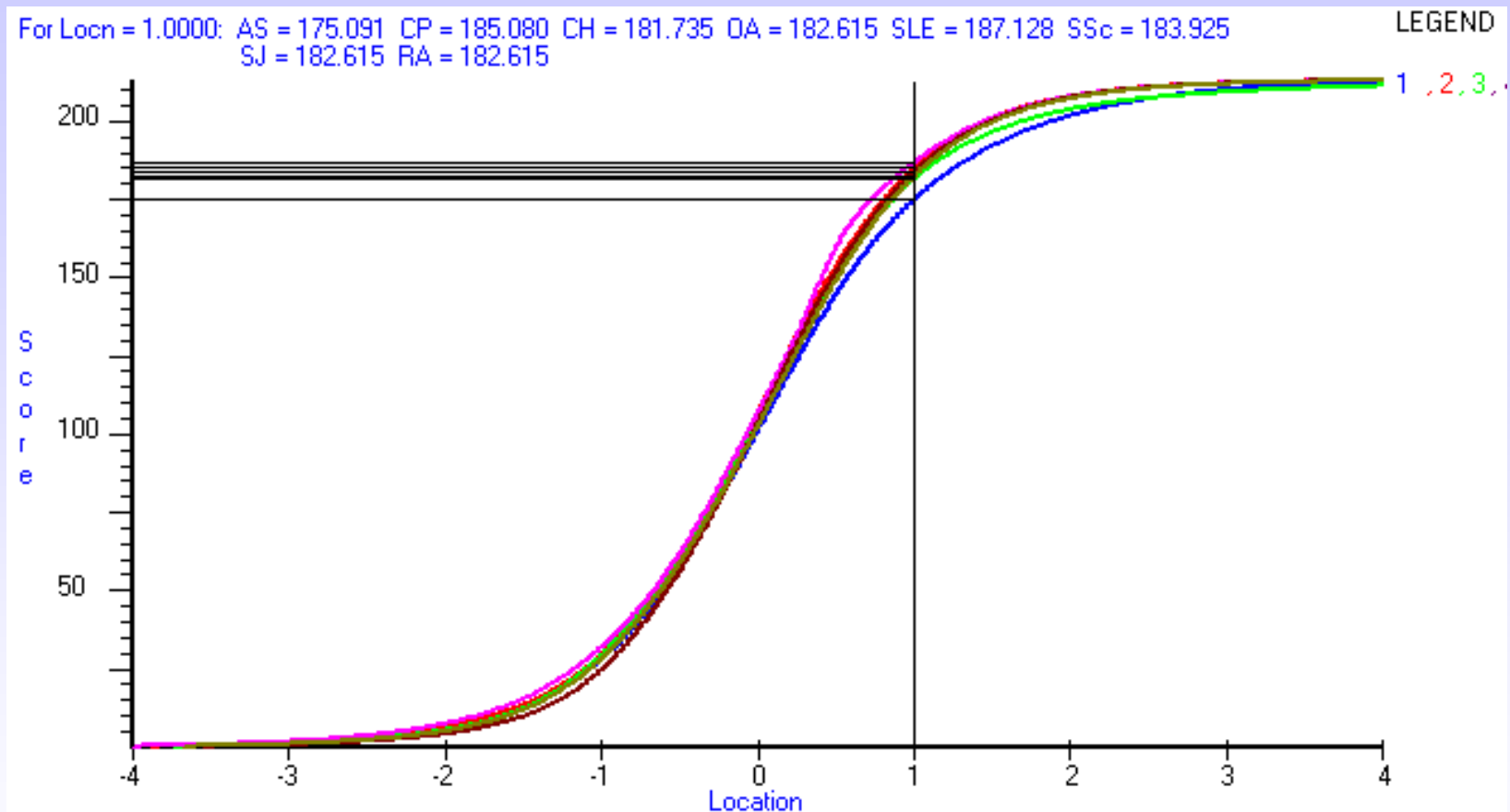


Self-Care Component - Equivalence



Max difference = 0.274 logits (between AS & SLE)


Self-Care Component - Equivalence



Max difference = 12 raw score points (between AS & SLE)

Mobility Component - Equivalence

Original Domain	Domain Name	AS	CP	CH	OA	SLE	SSc	SJ	RA
4	bathing								
6	move indoors								
9	transfers								
11	move outdoors								
12	house								

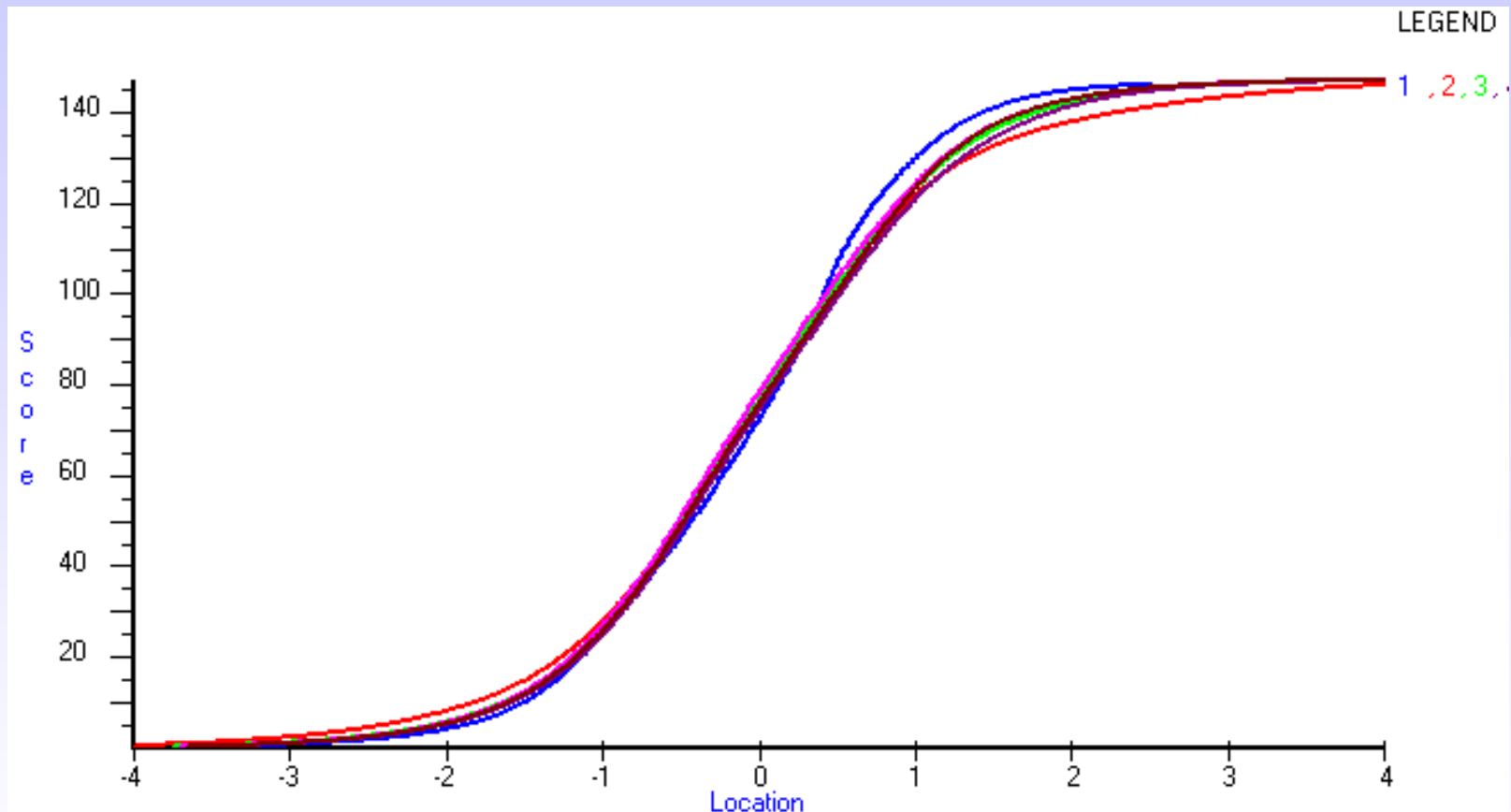
 =calibration

Mobility Component - Equivalence

Original Domain	Domain Name	Split	AS	CP	CH	OA	SLE	SSc	SJ	RA
4	bathing	CHRA4								
		i004								
6	move indoors	CH6								
		AS6								
		i06								
9	transfers	i09								
		CP9								
		AS9								
		SSc9								
11	move outdoors	ST11								
12	house	CH12								
		i012								

 =calibration

Mobility Component - Equivalence



Conclusions

- 12-Domain, 2-Component structure is fairly stable WITHIN each MSC
- Total scores are NOT directly equivalent across MSCs (due to DIF)
- Adjustments can be made to provide a scoring matrix of equivalent scores across conditions (although the trait measured is different from the original scale)

Acknowledgments

Available from:

- EDAQ User Manual: <http://usir.salford.ac.uk/30752/>
- EDAQ Parts 1 and 2: <http://usir.salford.ac.uk/30755/>
- EDAQ Parts 1 to 3: <http://usir.salford.ac.uk/30754/>
- United Kingdom Occupational Therapy Research Fund
- Contact: m.c.horton@leeds.ac.uk

