Country-related differential item functioning in the EQ-5D-5L:

Insights from the EuroQol Data Assessment of Population Health Needs and Instrument Evaluation (EQ-DAPHNIE)

Jiabi Wen, Ademola J Itiola, Jeffrey A Johnson, Mathieu F (Bas) Janssen, Paula K Lorgelly, Tolulope Sajobi, Fatima Al Sayah

MSR64

Objective

Cross-country comparisons of EQ-5D-5L assume **measurement invariance.**

Measurement Invariance

Same Health Status → Same EQ-5D-5L Response

Differential Item Functioning

*Two people with the **same injury***

	EQ-5D-5L Response
Country A	Moderate problems
Country B	Severe problems

Actual health: **Identical** ✅
Reported sores: **Different** ❌

If **differential item functioning (DIF)** exists, comparisons may be invalid.

Research Question:
Does country-related DIF exist in EQ-5D-5L responses?

Methods

Data

- Standardized online data collection from adults aged 18+ years in UK, US, CA, AU, NZ (English-speaking); ES, MX, CL, AR (Spanish); NL, GE (Germanic); FR, CN, JP, and BR
- SDoH, health and wellbeing, health behaviours, health utilization
- Quota sampling (age, sex, income, residence)

EQ-5D-5L

- Five dimensions: mobility, self-care, usual activities, pain/discomfort, anxiety/depression.
- Five levels: 1 (no problems) – 5 (extreme problems/unable to)
- Level sum score: summing across all dimensions (range: 5-25)
- EQ-VAS: 0 (worst health) – 100 (best health) visual analog scale

DIF analysis

3 nested ordered logistic regression models: EQ-5D-5L dimension as the outcome variable (Note: level 3-5 grouped to one category)

M1: $log(dimension) = b_0 + b_1\ trait$

M2: $log(dimension) = b_0 + b_1\ trait + b_2\ group$

M3: $log(dimension) = b_0 + b_1\ trait + b_2\ group + b_3\ group{\times}trait$

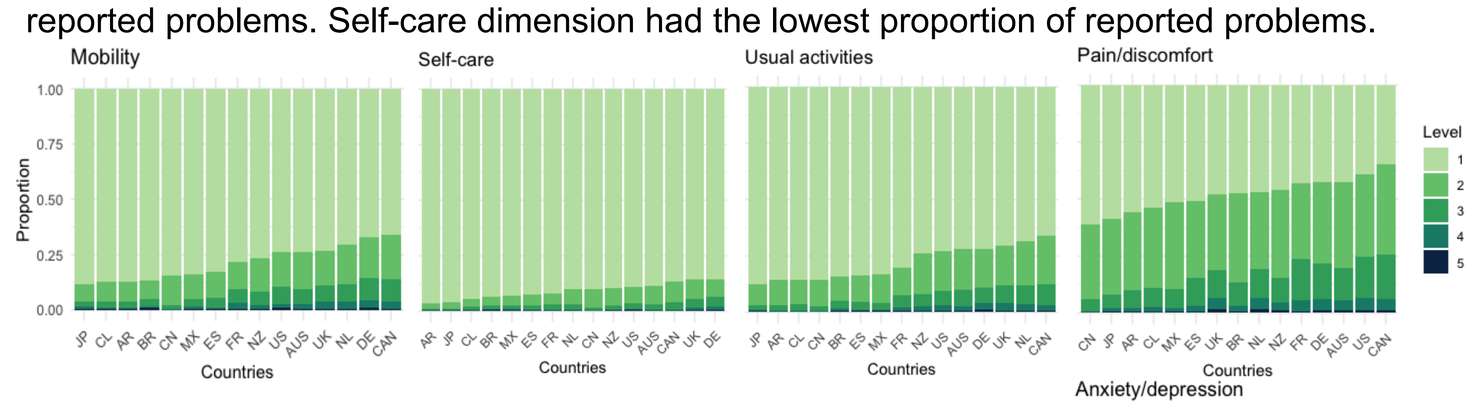
DIF exist: b_2 or (b_2 and b_3) statistically significant and reach the effect size threshold (change in model $R^2 > 3.5\%$)

Trait proxies

EQ-5D-5L level sum score; EQ-VAS; EQ-5D-5L latent score based on item response theory; PROMIS-10 physical/mental t-score; EQ-HWB-9 level sum score

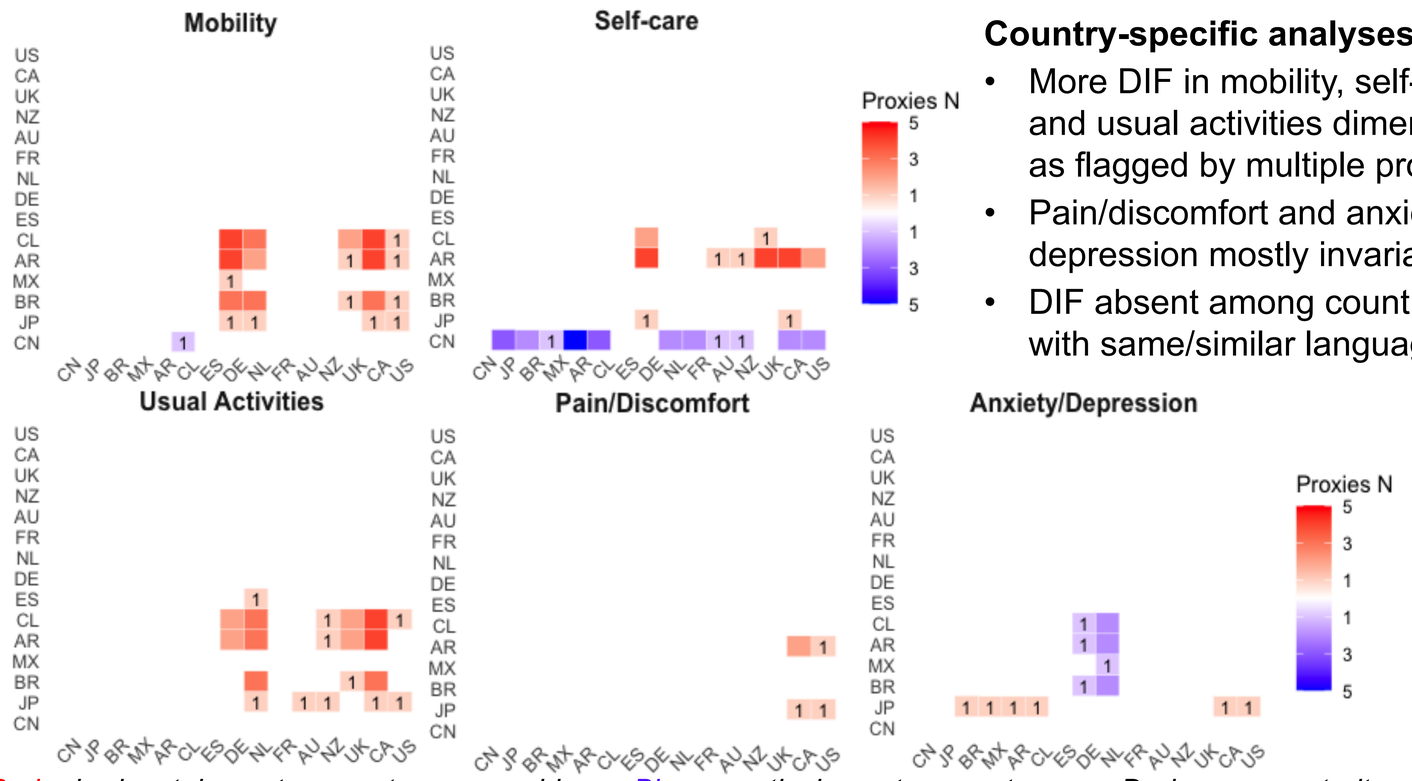
Results - Overall

- **Sample size:** 67,178 respondents across 15 countries with complete EQ-5D-5L responses (missing: 2%). Per country: 4389-4989.
- **Dimension-level patterns:** Pain/discomfort and anxiety/depression dimensions had the most reported problems. Self-care dimension had the lowest proportion of reported problems.



- **Country patterns:**
  - ➢ Japan & China: Best health across measures
  - ➢ Latin America (AR, CL, BR, MX): Fewer physical problems, more problems in the anxiety/depression dimension
  - ➢ Western countries: Poorer EQ-5D-5L level sum score & EQ-VAS, mixed PROMIS physical/mental t-scores and EQ-HWB-9 level sum score

Results - DIF



Red = horizontal country reports more problems; Blue = vertical country reports more; Darker = more trait proxies.

Country-specific analyses

- More DIF in mobility, self-care, and usual activities dimensions, as flagged by multiple proxies
- Pain/discomfort and anxiety/depression mostly invariant
- DIF absent among countries with same/similar language.

Language-specific analyses

- Findings above suggested language-specific patterns. Countries with the same or similar language were grouped and analyzed.
- When looking at language pairs that flagged by at least 2 proxies, mobility and self-care are the only dimensions.
- Germanic countries tended to be more likely to report more problems compared to Brazil
- China tended to report more problems than Japan, France, and Brazil

Dimension	Trait	Countries		Change in R^2
		Fewer problems	More problems	M2 vs M1
Mobility	PROMIS-10 physical	Brazil	Germanic	0.0399
Mobility	PROMIS-10 mental	Brazil	Germanic	0.0367
Self-care	EQ-5D-5L LSS	Japan	China	0.0466
Self-care	EQ-VAS	Japan	China	0.0438
Self-care	EQ-HWB-9	Japan	China	0.0349
Self-care	EQ-5D-5L LSS	France	China	0.0598
Self-care	EQ-5D-5L IRT	France	China	0.0396
Self-care	EQ-5D-5L LSS	Brazil	China	0.0501
Self-care	PROMIS-10 physical	Brazil	China	0.0320

All comparisons significant at p<0.001 level

Conclusion

- Overall invariance within language groups supports comparability.
- Potential bias exists for cross-country comparisons outside language groups.
- Implication: consider linguistic/cultural differences in cross-country EQ-5D-5L analyses.
- Next steps: explore whether observed DIF reflects translation vs. cultural interpretation