

# VALIDATION OF A PHYSICAL ACTIVITY SCALE FOR OLDER ADULTS PARTICIPATING IN THE HEALTH AND RETIREMENT STUDY



Peter D. Hart, PhD  
Kinesmetrics Lab  
Health Promotion Research  
GSA 2025 Annual Scientific Meeting



## INTRODUCTION

The surveillance of physical activity (PA) at the population level generally involves administering a small number of questionnaire items. The combining of multiple related items to create a scale for scores has many psychometric benefits.

## PURPOSE

The aim of this research was to validate a new scale measuring PA using items contained in a large national survey of older adults.

## METHODS

Data from 15,335 adults 50+ years of age participating in the 2022 Health and Retirement Study (HRS) were used.

The assessment strategy involved six steps: 1) defining the PA scale (PAS) items and categories, 2) factor analysis, 3) internal consistency reliability, 4) item response theory (IRT) analysis, 5) convergent validity correlations, and 6) modeling PAS scores with a general health (GH) outcome.

Polychoric correlations between items were used for the classical analyses. A graded response model (GRM) for polytomous items was employed for the IRT analysis. Multinomial logistic regression was used to model GH categories with PAS scores.

## RESULTS

The PAS included 3 items of vigorous, moderate, and light PA, each with 3 categories of inactive, low/moderately active, and highly active. Factor analysis retained a single factor with 70% explained variance, whilst the reliability coefficient for items was 0.79. IRT calibration showed category thresholds ranging from -1.90 to 1.05 and item discrimination parameters between 1.39 and 5.00. IRT theta scores correlated with the PAS sum score ( $r=0.97$ ), age ( $r=-0.21$ ), GH ( $r=0.38$ ), and timed walk performance ( $r=0.35$ ). Modeling showed that for each point increase in PAS score, odds of poor ( $OR=0.31$ , 0.27-0.34), fair ( $OR=0.47$ , 0.43-0.52), good ( $OR=0.60$ , 0.56-0.65), and very good ( $OR=0.79$ , 0.73-0.85) GH, as compared to excellent (reference), continually decreased.

## RESULTS

Table 1. Physical activity (PA) rating scale development.

| Original item categories   | New scale | New ratings           |
|----------------------------|-----------|-----------------------|
| Hardly ever or never       | 0         | Inactive              |
| One to three times a month | 1         | Low/moderately active |
| Once a week                | 1         | Low/moderately active |
| More than once a week      | 2         | Highly active         |
| Every day                  | 2         | Highly active         |

Note. Each item's stem: How often do you take part in sports or activities that are [vigorous] or [moderate] or [mild] ....

Table 4. Graded response model item parameter estimates for the new PA scale (PAS).

| Item | Parameter | Estimate | SE    | p      |
|------|-----------|----------|-------|--------|
| VPA  | b1        | 0.270    | 0.015 | <.0001 |
|      | b2        | 1.048    | 0.023 | <.0001 |
|      | a         | 1.576    | 0.043 | <.0001 |
| MPA  | b1        | -0.784   | 0.015 | <.0001 |
|      | b2        | 0.033    | 0.011 | 0.0012 |
|      | a         | 5.003    | 0.477 | <.0001 |
| LPA  | b1        | -1.897   | 0.037 | <.0001 |
|      | b2        | -0.197   | 0.016 | <.0001 |
|      | a         | 1.393    | 0.035 | <.0001 |

Note. N = 15,335. Graded response model used a logit function and marginal maximum likelihood estimation [10].

Table 8. Multinomial logistic regression analyses for PAS sum score and GH outcome.

| Outcome Variable | PAS sum score |      |      |
|------------------|---------------|------|------|
|                  | OR            | LL   | UL   |
| GH - Unadjusted  |               |      |      |
| 1 vs. 5          | 0.31          | 0.27 | 0.34 |
| 2 vs. 5          | 0.47          | 0.43 | 0.52 |
| 3 vs. 5          | 0.60          | 0.56 | 0.65 |
| 4 vs. 5          | 0.79          | 0.73 | 0.85 |
| GH - Adjusted    |               |      |      |
| 1 vs. 5          | 0.34          | 0.30 | 0.39 |
| 2 vs. 5          | 0.50          | 0.46 | 0.55 |
| 3 vs. 5          | 0.62          | 0.57 | 0.67 |
| 4 vs. 5          | 0.79          | 0.73 | 0.85 |

Note. N = 12,136 (unadjusted). N = 11,820 (adjusted).

## CONCLUSIONS

The new PAS appears to represent a unidimensional construct with acceptable validity and reliability properties. These results support the use of a simple 3-item scale to measure PA in older adults.

## REFERENCES

Nunnally, J. C. (1978). Psychometric Theory (2nd ed.). New York: McGraw-Hill.

Health and Retirement Study. (2022). Data description and usage. 2020 core early version.

Matlock Cole K, Paek I. PROC IRT: A SAS Procedure for Item Response Theory. Appl Psychol Meas. 2017;41(4):311-320. doi:10.1177/0146621616685062

SAS Institute Inc. 2015. SAS/STAT® 14.1 User's Guide. The IRT Procedure. Cary, NC: SAS Institute Inc.

Hart, P. D. (2023). Perceived happiness and general health: An IRT investigation. Research in Psychology and Behavioral Sciences, 11(2), 49-55.

Boateng GO, Neilands TB, Frongillo EA, Melgar-Quinonez HR, Young SL. Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. Front Public Health. 2018;6:149. Published 2018 Jun 11.

## RESULTS

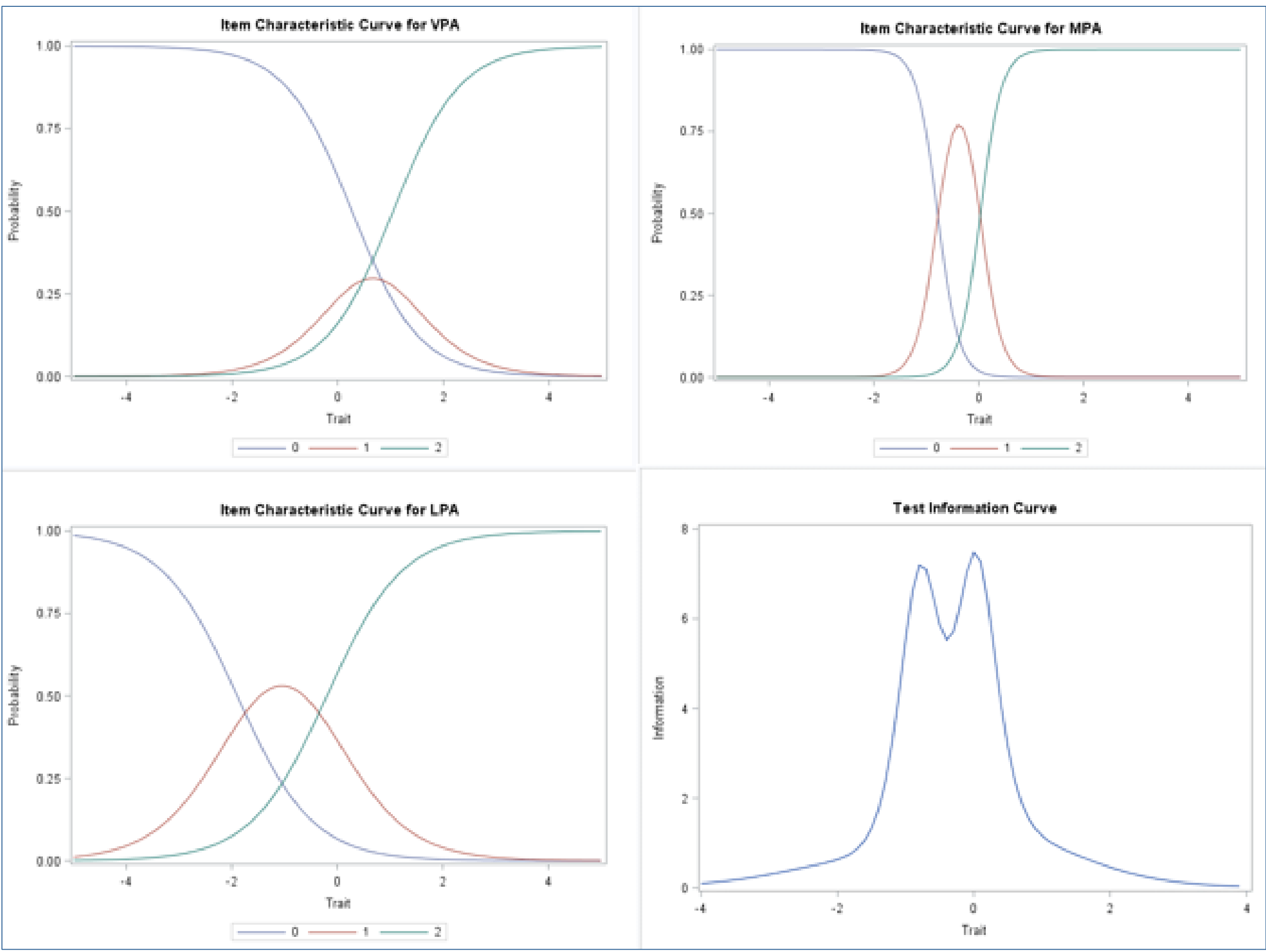


Table 6. Validity coefficients for the IRT-derived PA scale (PAS) score.

| Item                     | PAS    | Age    | Sex    | Educ   | GH     | BMI   | WC    | GS    | TW    | BT    |
|--------------------------|--------|--------|--------|--------|--------|-------|-------|-------|-------|-------|
| PAS IRT score - Pearson  | .962   | -.221  | .093   | .223   | .405   | -.228 | -.244 | .241  | .375  | .269  |
| PAS IRT score - Spearman | .967   | -.208  | .090   | .235   | .380   | -.150 | -.184 | .260  | .351  | .257  |
| N                        | 12,817 | 12,817 | 12,817 | 12,672 | 12,808 | 4,495 | 4,644 | 4,469 | 2,592 | 4,233 |

Note. Larger IRT scores represent greater amounts of PA. All coefficients are significant ( $p < .0001$ ). PAS is the PAS summed score ranging from 0 to 6 where larger values represent greater amounts of PA. Sex is coded 1 for males and 0 for females. Educ is years of education ranging from 0 to 17. GH is general health ranging from 1 to 5 where larger values represent better health. BMI is body mass index. WC is waist circumference. GS is grip strength. TW is timed walk test speed. BT is full tandem balance test time. Pearson coefficients are weighted. Spearman coefficients are not weighted.

Table 7. Multinomial logistic regression analyses for raw PAS sum scores and GH outcome.

| Outcome Variable | PAS sum score groups |        |        |                |       |       |                |      |       |
|------------------|----------------------|--------|--------|----------------|-------|-------|----------------|------|-------|
|                  | PAS: 0-1 vs. 6       |        |        | PAS: 2-3 vs. 6 |       |       | PAS: 4-5 vs. 6 |      |       |
|                  | OR                   | LL     | UL     | OR             | LL    | UL    | OR             | LL   | UL    |
| GH - Unadjusted  |                      |        |        |                |       |       |                |      |       |
| 1 vs. 5          | 300.23               | 120.30 | 749.27 | 36.00          | 16.85 | 76.94 | 7.14           | 3.00 | 16.99 |
| 2 vs. 5          | 42.12                | 24.80  | 71.52  | 13.71          | 9.80  | 19.16 | 4.50           | 3.25 | 6.24  |
| 3 vs. 5          | 10.82                | 6.94   | 16.87  | 5.59           | 4.25  | 7.34  | 2.97           | 2.33 | 3.79  |
| 4 vs. 5          | 2.71                 | 1.80   | 4.08   | 2.30           | 1.77  | 2.98  | 2.01           | 1.58 | 2.56  |
| GH - Adjusted    |                      |        |        |                |       |       |                |      |       |
| 1 vs. 5          | 158.24               | 58.78  | 425.98 | 27.50          | 12.09 | 62.59 | 6.15           | 2.49 | 15.18 |
| 2 vs. 5          | 27.67                | 15.68  | 48.83  | 11.27          | 7.75  | 16.38 | 4.26           | 2.99 | 6.05  |
| 3 vs. 5          | 9.01                 | 5.65   | 14.35  | 5.11           | 3.81  | 6.85  | 2.89           | 2.26 | 3.71  |
| 4 vs. 5          | 2.65                 | 1.70   | 4.14   | 2.28           | 1.72  | 3.02  | 2.00           | 1.57 | 2.55  |

Note. N = 12,136 (unadjusted). N = 11,820 (adjusted). Larger PAS scores represent greater amounts of PA. GH is general health where 1=poor, 2=fair, 3=good, 4=very good, and 5=excellent. OR is odds ratio. LL is lower limit of the 95% OR confidence interval. UL is upper limit of the 95% OR confidence interval. Adjusted model is adjusted for age, sex, race, employment, education, and marital status.