PSYCHOMETRICS

What Can We Measure About a Mind?

A SAHILTAYAL.COM PRODUCTION | Research meets reality

Field guide | November 2026 | Educational reference, not clinical or employment decision advice

1. What psychometrics studies

Psychometrics studies the theory and practice of psychological measurement: developing instruments, estimating scores, testing measurement models and evaluating interpretations. We observe responses and behaviours; latent constructs are inferred rather than directly read from the mind. Every estimate is conditional on a model and context.

2. The measurement workflow

Define the construct: Specify intended meaning, boundaries and decision context.

Generate observations: Items, tasks, ratings, performance samples or repeated observations.

Pilot and refine: Check comprehension, response processes, missingness and accessibility.

Study internal structure: Investigate dimensionality, local dependence and plausible scoring models.

Estimate precision: Reliability coefficients and standard errors, conditional if appropriate.

Investigate validity: Evidence from content, response processes, internal structure, relationships and consequences.

Norm and adapt responsibly: Document reference groups, languages, adaptations and differences.

Monitor in use: Recheck drift, fairness, consequences and changing contexts.

3. Reliability

Reliability is the consistency or precision of measurements, not proof that a construct has been correctly identified. Internal consistency, stability over time, inter-rater agreement and model-based precision target different error sources. Cronbach alpha alone is not a universal measure of quality.

4. Validity

Modern validity arguments concern support for score interpretations and uses. Distinguish associations with other variables from causal claims. Predictive association does not imply fairness or justify every downstream decision. A scale may work differently by age, language, setting, and test version.

5. Norms, percentiles and error

A percentile compares a result to a specific reference distribution. It is not percentage correct or a biological rank. In classical test theory, an illustrative standard error of measurement is SD × sqrt(1 – reliability). Its interpretation requires assumptions and cannot be indiscriminately attached to every measure.

6. Models and approaches

Classical test theory, item response theory, Rasch models, generalizability theory, factor analysis, measurement invariance and adaptive testing are different tools. Choice follows research questions, data types and assumptions—not fashion.

7. Clinical, occupational and educational use

A screener is not a diagnosis. A hiring test requires job-related validity and fairness review; classroom assessments need alignment with learning opportunities; clinical interpretation requires appropriate competence and context. Avoid reducing a person to a single score.

8. Ethics and limitations

Respect consent, minimal data collection, confidentiality, accessible administration, cultural and linguistic fairness, and the right to challenge harmful decisions. Be cautious of proprietary licensing, credential restrictions, false precision, automated profiling, test preparation effects and misuse of cut scores.

9. Reading sources

Standards for Educational and Psychological Testing — https://www.testingstandards.net/

COSMIN methodology — https://www.cosmin.nl/

International Test Commission — https://www.itc-testcommission.org/

APA testing and assessment — https://www.apa.org/science/programs/testing

National Academies: Psychological Testing in the Service of Disability Determination — https://www.ncbi.nlm.nih.gov/books/NBK305233/

10. Curated catalogue index

The accompanying website contains 160 named instruments and approaches, arranged by category. Links open PubMed research searches, not necessarily current official test manuals. Evidence status, licensing and population fit must be individually checked before any use.

Cognition and intelligence

WAIS-IV • WISC-V • WPPSI-IV • Stanford–Binet Intelligence Scales, Fifth Edition • Raven’s Progressive Matrices

Cattell Culture Fair Intelligence Test • Woodcock–Johnson IV Tests of Cognitive Abilities • Kaufman Assessment Battery for Children, Second Edition • Differential Ability Scales, Second Edition • Universal Nonverbal Intelligence Test, Second Edition

Wechsler Memory Scale, Fourth Edition • Montreal Cognitive Assessment • Mini-Mental State Examination • Trail Making Test • Stroop Color and Word Test

Wisconsin Card Sorting Test • Digit Span • Corsi Block-Tapping Test • Rey Auditory Verbal Learning Test • Rey–Osterrieth Complex Figure Test

Cambridge Neuropsychological Test Automated Battery • NIH Toolbox Cognition Battery • Addenbrooke's Cognitive Examination III • Clock Drawing Test • Test of Nonverbal Intelligence, Fourth Edition

Personality and traits

NEO Personality Inventory-3 • NEO Five-Factor Inventory-3 • Big Five Inventory-2 • IPIP-NEO • HEXACO Personality Inventory-Revised

Minnesota Multiphasic Personality Inventory-3 • Personality Assessment Inventory • Millon Clinical Multiaxial Inventory-IV • Eysenck Personality Questionnaire-Revised • Temperament and Character Inventory

California Psychological Inventory • Hogan Personality Inventory • Hogan Development Survey • Hogan Motives, Values, Preferences Inventory • 16 Personality Factor Questionnaire

Revised NEO Personality Inventory • Ten Item Personality Inventory • International Personality Item Pool • Dark Triad Dirty Dozen • Short Dark Triad

Levenson Self-Report Psychopathy Scale • Barratt Impulsiveness Scale • UPPS-P Impulsive Behavior Scale • Adult Temperament Questionnaire • Cloninger Tridimensional Personality Questionnaire

Mental health and wellbeing

PHQ-9 • PHQ-2 • GAD-7 • GAD-2 • Beck Depression Inventory-II

Beck Anxiety Inventory • DASS-21 • Kessler Psychological Distress Scale K10 • Kessler Psychological Distress Scale K6 • WHO-5 Well-Being Index

Warwick–Edinburgh Mental Well-being Scale • Perceived Stress Scale • General Health Questionnaire-12 • Patient Health Questionnaire-15 • PCL-5

PTSD Checklist—Civilian Version • Impact of Event Scale-Revised • Columbia–Suicide Severity Rating Scale • Young Mania Rating Scale • Hamilton Depression Rating Scale

Hamilton Anxiety Rating Scale • Insomnia Severity Index • Pittsburgh Sleep Quality Index • Epworth Sleepiness Scale • UCLA Loneliness Scale

Rosenberg Self-Esteem Scale • Satisfaction With Life Scale • PANAS • Brief COPE • Connor–Davidson Resilience Scale

WHOQOL-BREF

Education and achievement

SAT • ACT • GRE General Test • Graduate Management Admission Test • Woodcock–Johnson IV Tests of Achievement

WIAT-4 • Kaufman Test of Educational Achievement-3 • Peabody Picture Vocabulary Test-5 • Expressive Vocabulary Test-3 • Test of Word Reading Efficiency-2

Gray Oral Reading Tests-5 • Comprehensive Test of Phonological Processing-2 • Nelson–Denny Reading Test • Wide Range Achievement Test-5 • Progressive Achievement Tests

National Assessment of Educational Progress • PISA • TIMSS • PIRLS • Dynamic Indicators of Basic Early Literacy Skills

Curriculum-Based Measurement

Work and organisations

General Aptitude Test Battery • Wonderlic Personnel Test • Watson–Glaser Critical Thinking Appraisal • SHL Verify Interactive • Raven's Advanced Progressive Matrices

Situational Judgment Tests • Assessment Center Exercises • Structured Behavioral Interview Ratings • Organizational Commitment Questionnaire • Utrecht Work Engagement Scale

Maslach Burnout Inventory • Job Satisfaction Survey • Minnesota Satisfaction Questionnaire • Psychological Safety Scale • Work Design Questionnaire

Copenhagen Psychosocial Questionnaire • Perceived Organizational Support Scale • Leader–Member Exchange-7 • Multifactor Leadership Questionnaire • Workplace Big Five Profile

Development and social functioning

Ages and Stages Questionnaires-3 • Bayley Scales of Infant and Toddler Development-4 • Vineland Adaptive Behavior Scales-3 • Adaptive Behavior Assessment System-3 • Child Behavior Checklist

Strengths and Difficulties Questionnaire • Conners 4 • Autism Diagnostic Observation Schedule-2 • Autism Diagnostic Interview-Revised • Social Responsiveness Scale-2

Behavior Assessment System for Children-3 • Developmental Profile 4 • Denver Developmental Screening Test II • BRIEF-2 • Empathy Quotient

Reading the Mind in the Eyes Test • Interpersonal Reactivity Index • Social Skills Improvement System

Methods and measurement approaches

Classical Test Theory • Item Response Theory • Rasch Measurement Model • Generalizability Theory • Confirmatory Factor Analysis

Exploratory Factor Analysis • Measurement Invariance Testing • Differential Item Functioning Analysis • Computerized Adaptive Testing • Ecological Momentary Assessment

Experience Sampling Method • Implicit Association Test • Q Methodology • Semantic Differential Scale • Thurstone Scaling

Guttman Scaling • Likert Summated Ratings • Many-Facet Rasch Measurement • Multitrait–Multimethod Matrix • Criterion-Referenced Assessment

11. Reflection

The promise of measurement is not omniscience. It is disciplined uncertainty: asking what an instrument really tells us, what it fails to tell us, and what decision its evidence can justifiably support. The most humane score is one that retains its uncertainty and its purpose.

12. Editorial evidence policy

This field guide is an educational orientation, not an individually verified technical review of each instrument. Inclusion does not assert validation, legal permission, cultural equivalence or access to copyrighted items. Readers should consult manuals and peer-reviewed studies for the exact version, population and application.