

Educational Measurement & Evaluation

Practice Test 7

60 original multiple-choice questions for teacher test preparation (FPSC / FGEI / PPSC style). Choose one answer for each question. The full answer key is on the last page — worked explanations for every question are free on the test page at QuaidianPoint.com.

Name: _____ Class: _____ Score: _____ / 60

Norm-Referenced and Criterion-Referenced Testing

Q1. A school reports Ali is at the 90th percentile. The best interpretation is that Ali scored

- (A) 90 percent of the test items correctly
- (B) 90 marks in total
- (C) Better than 90 students in every school
- (D) Higher than about 90 percent of the norm group

Q2. A teacher says, 'Sana has mastered all five objectives in Unit 2.' This statement is

- (A) Norm-referenced
- (B) Aptitude-based
- (C) Interest-based
- (D) Criterion-referenced

Q3. Which report is norm-referenced?

- (A) 'Bilal can add three-digit numbers correctly.'
- (B) 'Bilal ranks third in a class of forty.'
- (C) 'Bilal has met the passing standard for spelling.'
- (D) 'Bilal completed all ten grammar objectives.'

Q4. A pass mark of 40 percent fixed before the test is an example of

- (A) A norm group
- (B) A percentile
- (C) A criterion
- (D) A quartile

Q5. Norm-referenced tests need a norm group that is

- (A) Made up only of top students
- (B) Representative of the students being compared
- (C) Selected after results are known
- (D) Limited to one classroom in all cases

Q6. If a test is very easy for a whole class, a norm-referenced comparison will

- (A) Prove that all students mastered every objective
- (B) Become a criterion-referenced test automatically
- (C) Lose all scoring value
- (D) Still rank students, but differences may be small and unstable

Q7. A criterion-referenced test is fairer for a mastery decision because it

- (A) Shows learning against the same standard for everyone
- (B) Ranks students even when all have learned well
- (C) Changes the standard after seeing results
- (D) Rewards only the fastest student

Q8. A selection test for ten seats from 500 applicants is best served by

- (A) Norm-referenced comparison
- (B) A diagnostic interview only
- (C) A daily homework check
- (D) An ungraded class activity

Q9. Which pair contrasts the two approaches correctly?

- (A) Norm: mastery; Criterion: rank
- (B) Norm: diagnosis; Criterion: prediction only
- (C) Norm: rank; Criterion: mastery of objectives
- (D) Norm: feedback; Criterion: no feedback

Grading Systems and GPA

Q10. A student with grades A, B, C and B in four equal-credit courses has a GPA of

- (A) 3.0
- (B) 3.25
- (C) 2.75
- (D) 3.5

Q11. If a 3-credit course gives grade B (3 points) and a 1-credit course gives A (4 points), the GPA is

- (A) 3.5
- (B) 3.0
- (C) 3.25
- (D) 4.0

Q12. A CGPA of 3.6 after six semesters drops after a weak seventh semester because CGPA

- (A) Forgets earlier semesters
- (B) Counts only failed courses
- (C) Includes the new semester in the cumulative average
- (D) Uses only the best semester

Q13. Letter grades help communication mainly because they

- (A) Show every answer a student gave
- (B) Summarise performance in a familiar, compact form
- (C) Replace all written feedback
- (D) Measure intelligence directly

Q14. Absolute grading can be seen as fair when

- (A) Grades depend on who else is in the class
- (B) The standard is clear and known to students in advance
- (C) The teacher changes cut-offs after marking
- (D) Only a fixed number of A grades are allowed

Q15. Relative grading can feel unfair to students because

- (A) It always uses the same cut-off for every class
- (B) It ignores all test scores
- (C) A good performance may still earn a low grade if peers do better
- (D) It gives every student the same grade

Q16. A school that wants to encourage cooperation rather than competition would prefer

- (A) Strict ranking on a curve
- (B) Absolute or criterion-based grading
- (C) Grading only the top student highly
- (D) Publishing class ranks weekly

Q17. Grade inflation means that grades have

- (A) Fallen because tests became harder
- (B) Become impossible to calculate
- (C) Risen over time without a matching rise in learning
- (D) Stopped being used by schools

Q18. When a transcript shows GPA without course titles, a reader cannot tell

- (A) Whether the student attended school
- (B) The student's name
- (C) Which subjects were taken or how demanding they were
- (D) The year of study

Standardized Tests and Their Characteristics

Q19. A test manual that gives age norms helps a user to

- (A) Interpret a child's score against children of the same age
- (B) Set the pass mark for the whole school
- (C) Rank children across different age groups directly
- (D) Diagnose a specific reading difficulty on its own

Q20. Why do standardized tests use a large tryout sample?

- (A) To make the test longer for every school
- (B) To build reliable norms and check items broadly
- (C) To avoid any scoring key
- (D) To let teachers set their own time limits

Q21. A standardized reading test given in a noisy hall with interruptions has lost

- (A) Its printed answer key
- (B) Its syllabus coverage
- (C) Its grade scale
- (D) Standardization of administration

Q22. A school uses a standardized test to check progress over three years. This works best if

- (A) A different local test is used each year
- (B) Only the top class is tested each year
- (C) Results are kept secret from teachers
- (D) The same standardized procedures and comparable scores are used each year

Q23. One valid criticism of heavy reliance on standardized tests is that they

- (A) Always improve teaching in every subject
- (B) Measure every school goal equally well
- (C) May narrow teaching toward tested content only
- (D) Remove the need for teacher judgement entirely

Q24. Standardized test scores are most useful when combined with

- (A) Classroom evidence such as classwork and teacher observation
- (B) No other information at all
- (C) A single student's opinion of the test
- (D) The school's building plan

Q25. A norm table is used to

- (A) Convert a raw score into a comparative score such as a percentile
- (B) Set the school timetable
- (C) Write new syllabus topics
- (D) Choose textbooks for a class

Q26. Which is a characteristic of a well-standardized test?

- (A) Documented reliability and validity evidence
- (B) No instructions for examiners
- (C) A different scoring key in each centre
- (D) Items that change during the test

Formative, Summative and Diagnostic Evaluation

Q27. A teacher checks understanding with questioning during a lesson and reteaches a point at once. This is

- (A) Summative evaluation
- (B) Certification testing
- (C) Formative evaluation
- (D) Admission testing

Q28. An end-of-unit test used to decide promotion to the next unit is mainly

- (A) Diagnostic only
- (B) Summative
- (C) Formative only
- (D) Informal observation only

Q29. A reading teacher gives a short screening, then a detailed check of decoding skills for struggling readers. The detailed check is

- (A) Diagnostic
- (B) Summative
- (C) Norm-referenced only
- (D) A final examination

Q30. Formative assessment improves learning most when students

- (A) Receive only a total mark at the end
- (B) Are ranked publicly each week
- (C) Are tested without any comment
- (D) Use feedback to correct errors before the final judgement

Q31. Which example shows diagnostic evaluation in mathematics?

- (A) Finding that a student confuses place value when subtracting
- (B) Reporting the class average for the term
- (C) Ranking students for a prize
- (D) Giving a final grade for the year

Q32. Summative results are often reported to parents as

- (A) Daily hints for the next lesson
- (B) A summary of achievement for a term or course
- (C) A list of every classroom question asked
- (D) A diagnosis of one error type

Q33. A teacher who uses exit tickets at the end of each lesson is practising

- (A) Final certification
- (B) Placement testing for admission
- (C) School inspection
- (D) Formative assessment

Q34. Diagnostic testing differs from a general achievement test mainly in that diagnostic testing:

- (A) Gives only a total score for ranking
- (B) Is used only at the end of the year
- (C) Covers the whole syllabus in one sitting
- (D) Pinpoints the specific nature of a pupil's difficulty rather than giving an overall score

Q35. A portfolio reviewed at the end of term to judge overall growth is mainly

- (A) Diagnostic
- (B) Formative questioning
- (C) A placement interview
- (D) Summative

Test Construction Steps

Q36. A blueprint that gives 40 percent of items to one unit and 10 percent to another reflects

- (A) The relative weight of the units in teaching and objectives
- (B) The alphabetical order of units
- (C) The number of students in each class
- (D) The teacher's favourite topics only

Q37. When writing a multiple-choice item, the stem should

- (A) Contain the answer in its wording
- (B) Use double negatives wherever possible
- (C) Be as long as the options combined
- (D) Present a clear problem without clues to the answer

Q38. A good distractor in a multiple-choice item is

- (A) Obviously wrong to every student
- (B) Correct under a different interpretation
- (C) Unrelated to the topic of the item
- (D) Plausible to a student who has not fully learned the point

Q39. During tryout, item difficulty is judged by

- (A) The proportion of tryout students who answer correctly
- (B) The length of the item
- (C) The number of words in the stem
- (D) The teacher's guess about difficulty

Q40. Item discrimination tells test makers whether an item

- (A) Separates high achievers from low achievers on the total test
- (B) Is liked by all students
- (C) Uses difficult vocabulary
- (D) Appears early in the paper

Q41. A test with clear instructions, adequate time and a scoring key is more likely to be

- (A) Valid for every possible purpose
- (B) Reliable in scoring
- (C) Free from all syllabus limits
- (D) Accepted without any tryout

Q42. After tryout, items with poor discrimination are usually

- (A) Kept because they are unusual
- (B) Revised or removed
- (C) Given extra marks
- (D) Moved to the end of the paper only

Q43. Standardizing a newly built test includes

- (A) Trying it on a representative sample and preparing norms
- (B) Asking one class to set the pass mark
- (C) Printing it in colour
- (D) Letting each school choose its own time limit

Q44. A test that samples only two of ten taught units lacks

- (A) A scoring key
- (B) A title page
- (C) Content validity in coverage
- (D) Enough time for students

Basic Statistics for Teachers

Q45. The scores are 5, 7, 9, 11, 13. The mean is

- (A) 7
- (B) 9
- (C) 11
- (D) 8

Q46. The scores are 2, 4, 6, 8. The median is

- (A) 4
- (B) 5
- (C) 6
- (D) 3

Q47. In the data 10, 12, 12, 14, 15, the mode is

- (A) 12
- (B) 10
- (C) 14
- (D) 15

Q48. Scores range from 40 to 95. The range is

- (A) 40
- (B) 95
- (C) 55
- (D) 135

Q49. If the mean is 50 and the standard deviation is 5, a score of 60 has a z-score of

- (A) 1
- (B) 10
- (C) 0
- (D) 2

Q50. A z-score of -1 means the score lies

- (A) One standard deviation below the mean
- (B) One point below the mean
- (C) At the lowest possible value
- (D) One percentile below the median

Q51. In a normal curve, about what share of scores falls within one standard deviation of the mean?

- (A) About 50 percent
- (B) About 68 percent
- (C) About 95 percent
- (D) About 99 percent

Q52. The median is preferred over the mean when data contain

- (A) Only whole numbers
- (B) An equal number of boys and girls
- (C) Scores that are all identical
- (D) Extreme scores that would distort the average

Q53. The first quartile (Q1) is the value below which about

- (A) 50 percent of scores fall
- (B) 75 percent of scores fall
- (C) 25 percent of scores fall
- (D) 100 percent of scores fall

Q54. A teacher reports that a score is at the third quartile. This means about

- (A) 25 percent of scores are below it
- (B) 75 percent of scores are below it
- (C) The score is the highest in class
- (D) The score equals the mean

Evaluation Models

Q55. Tyler's model is sometimes criticised because it

- (A) Always requires expensive equipment
- (B) Cannot be used in schools
- (C) May ignore unintended outcomes that were not stated as objectives
- (D) Never uses test results

Q56. In CIPP, input evaluation mainly looks at

- (A) Final examination results only
- (B) Student behaviour after graduation
- (C) Daily attendance records only
- (D) Available resources, plans and strategies before action

Q57. Process evaluation in CIPP is concerned with

- (A) Needs in the community before planning
- (B) Final outcomes after the programme ends
- (C) The budget approved last year
- (D) How the programme is actually implemented

Q58. A CIPP evaluator who checks whether the programme met community needs is doing

- (A) Product evaluation only
- (B) Item analysis
- (C) Context evaluation
- (D) Grading of students

Q59. Tyler's model and CIPP differ mainly because CIPP

- (A) Ignores objectives completely
- (B) Uses no data at all
- (C) Covers decisions at several stages, not only final objectives
- (D) Applies only to one lesson

Q60. A school reviews its new reading programme at the end of the year to judge results. In CIPP terms this is

- (A) Context evaluation
- (B) Product evaluation
- (C) Input evaluation
- (D) Process monitoring only

Answer Key

Q1: D	Q11: C	Q21: D	Q31: A	Q41: B	Q51: B
Q2: D	Q12: C	Q22: D	Q32: B	Q42: B	Q52: D
Q3: B	Q13: B	Q23: C	Q33: D	Q43: A	Q53: C
Q4: C	Q14: B	Q24: A	Q34: D	Q44: C	Q54: B
Q5: B	Q15: C	Q25: A	Q35: D	Q45: B	Q55: C
Q6: D	Q16: B	Q26: A	Q36: A	Q46: B	Q56: D
Q7: A	Q17: C	Q27: C	Q37: D	Q47: A	Q57: D

Q8: A	Q18: C	Q28: B	Q38: D	Q48: C	Q58: C
Q9: C	Q19: A	Q29: A	Q39: A	Q49: D	Q59: C
Q10: A	Q20: B	Q30: D	Q40: A	Q50: A	Q60: B

Worked explanations for all 60 questions:

quaidianpoint.com/pedagogy-educational-measurement-evaluation-practice-test-7/