

Educational Measurement & Evaluation

Practice Test 9

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 national test reports a student's stanine as 8. This is best understood as

- (A) A score of 8 marks out of 10
- (B) A norm-based band showing high performance relative to peers
- (C) Proof of mastery of 8 objectives
- (D) A grade in one school subject

Q2. A teacher sets a mastery level of 80 percent for a spelling list. A student scoring 82 percent has

- (A) Met the criterion
- (B) Beaten 82 percent of peers
- (C) Ranked first in class
- (D) Failed the unit

Q3. A district compares schools using the same test and norms. This comparison is

- (A) Norm-referenced
- (B) Criterion-free in all respects
- (C) Diagnostic for one student
- (D) Formative for daily lessons

Q4. A student scores at the 50th percentile. This student is at about the

- (A) Top of the norm group
- (B) Pass mark for the course
- (C) Median of the norm group
- (D) Lowest score in the group

Q5. A criterion-referenced test for a safety skill should use items that

- (A) Rank trainees against each other
- (B) Use only very easy questions
- (C) Avoid any performance check
- (D) Sample the skill directly and cover its key steps

Q6. Why can a norm-referenced test show a student as above average even when the student missed key objectives?

- (A) Because rank does not guarantee mastery of specific objectives
- (B) Because norms always equal objectives
- (C) Because percentile equals percent correct
- (D) Because ranking removes all errors

Q7. A test is used to decide which students need extra help in reading. Adding a criterion report would help because it would show

- (A) Only the class rank order
- (B) Which reading skills each student still lacks
- (C) Only the school average
- (D) Only the test date

Q8. A school wants to know if its new maths programme raised achievement versus last year. The most direct norm approach is to

- (A) Ask teachers for opinions only
- (B) Count the number of textbooks used
- (C) Compare this year's norm-based scores with last year's on a comparable test
- (D) Check attendance alone

Q9. A parent asks, 'Can my child do the work for the next grade?' The teacher should answer mainly with

- (A) Only the child's percentile rank
- (B) Only the child's age
- (C) Only the class size
- (D) Criterion evidence about mastered skills

Grading Systems and GPA

Q10. Grades for a course are A (4), A (4), B (3) and C (2), all equal credit. The GPA is

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

Q11. A 2-credit course with grade A and a 2-credit course with grade C give a GPA of

- (A) 4.0
- (B) 2.0
- (C) 3.0
- (D) 3.5

Q12. A student has a CGPA of 3.2 over 40 credits. After a 10-credit semester at 4.0, the new CGPA is about

- (A) 3.2
- (B) 3.36
- (C) 4.0
- (D) 3.6

Q13. A school uses grade points A=4, B=3, C=2, D=1, F=0. A student with B, C and D in equal courses has a GPA of

- (A) 3.0
- (B) 1.5
- (C) 2.5
- (D) 2.0

Q14. Absolute grading with a pass mark of 50 means a student with 49 has

- (A) Passed because the class average was low
- (B) Ranked above students with 60
- (C) Not met the preset standard, regardless of class rank
- (D) Met the standard by effort alone

Q15. Relative grading on a curve with 10 percent A grades means that in a class of 30, about how many earn A?

- (A) 10
- (B) 3
- (C) 30
- (D) 1

Q16. A teacher drops the lowest quiz score before grading. This practice mainly affects

- (A) How the final grade summarises performance over the term
- (B) The difficulty of the remaining quizzes
- (C) The weight given to homework in the term
- (D) The order in which the quizzes were taken

Q17. A grade boundary is moved after marking so that more students pass. This change makes grades

- (A) More valid automatically
- (B) Less comparable with the stated standard
- (C) More reliable automatically
- (D) Free from all bias

Q18. A transcript that lists both grades and course levels helps a reader to

- (A) Rank students across schools by GPA alone
- (B) Know the student's attendance record
- (C) Compare only total marks without course context
- (D) Judge achievement with course difficulty in mind

Standardized Tests and Their Characteristics

Q19. A standardized test reports a grade-equivalent score of 5.4 for a Grade 3 student. This means the student's score matches

- (A) A Grade 5 curriculum mastery
- (B) A score of 5.4 marks
- (C) The typical score of students in the fourth month of Grade 5
- (D) Placement in Grade 5

Q20. Grade-equivalent scores can mislead because parents may think a Grade 3 student scoring 5.4 should

- (A) Move to Grade 5 at once
- (B) Stay in Grade 3 with support
- (C) Receive harder work only
- (D) Be tested again later

Q21. A test developer pilots items, removes biased items, and norms the test. This process supports

- (A) Faster printing only
- (B) A shorter syllabus
- (C) Fair and comparable interpretation
- (D) Automatic validity for all subjects

Q22. A school gives a standardized test without following the manual's time limit. The main problem is that scores

- (A) Become automatically higher for all
- (B) Lose their printed format
- (C) Require a new syllabus
- (D) Cannot be fairly compared with the norms

Q23. A standardized test of science that uses very complex reading passages may unfairly affect

- (A) Only students who like science
- (B) Students with weak reading skills, even if their science is strong
- (C) Only students in large schools
- (D) No student at all

Q24. A test with high reliability but low validity for a purpose is

- (A) Useless in every way
- (B) Automatically valid for another purpose
- (C) Impossible to score
- (D) Consistent, but not measuring what the user needs

Q25. Standardized scores from two different tests should not be directly compared unless

- (A) They share a common scale or a sound linking method
- (B) They have the same number of items
- (C) They were given on the same day
- (D) They use the same colour paper

Q26. A district that reports only average standardized scores hides

- (A) The name of the test
- (B) The date of testing
- (C) The number of test centres
- (D) Variation among schools and students within the district

Formative, Summative and Diagnostic Evaluation

Q27. A teacher uses a short pre-test, finds gaps in prior knowledge, and adjusts the first lessons. The pre-test acted as

- (A) A final summative grade
- (B) Diagnostic and formative for planning
- (C) A certification decision
- (D) A norm ranking only

Q28. A student completes a draft essay, receives comments, and revises it before grading. The comments are

- (A) Summative judgement
- (B) A diagnostic label
- (C) Formative feedback
- (D) An admission score

Q29. A term examination is used both to award grades and to review which topics need reteaching next term. Its second use is

- (A) A diagnostic test for one student only
- (B) A formative quiz during a lesson
- (C) An evaluative review for improvement, beyond its summative role
- (D) A placement test for admission

Q30. A teacher notices that many students confuse area with perimeter. The best immediate response is

- (A) Waiting until the final examination
- (B) Recording a final grade at once
- (C) Moving to a new topic without comment
- (D) Formative reteaching with a quick recheck

Q31. A diagnostic profile shows a student is strong in calculation but weak in word problems. Remediation should focus on

- (A) Drilling only basic calculation
- (B) Repeating the whole course unchanged
- (C) Interpreting and setting up word problems
- (D) Giving a final grade

Q32. A weekly spelling quiz that counts toward the term grade but also guides practice is best described as

- (A) Neither formative nor summative
- (B) Both formative in use and summative in grading
- (C) Only diagnostic
- (D) Only a placement test

Q33. A teacher asks students to rate their own confidence after a topic. This self-check is formative when it

- (A) Leads students to review weak areas before the next lesson
- (B) Is filed without discussion
- (C) Is used only to rank students
- (D) Replaces all teacher feedback

Q34. A final practical examination in a science course is summative, but it lacks diagnostic value if it

- (A) Reports only a total mark with no skill breakdown
- (B) Is given at the end of term
- (C) Uses a scoring rubric
- (D) Takes a full day

Q35. Which sequence shows the three types in a sensible order for a struggling student?

- (A) Summative judgement, diagnostic check, no support
- (B) Formative support, summative judgement, diagnostic check only after failure is recorded
- (C) Diagnostic check, formative support, summative judgement
- (D) Admission test, final grade, no feedback

Test Construction Steps

Q36. A blueprint allocates 12 items to Unit A and 4 items to Unit B. A student who studies only Unit B can gain at most

- (A) One quarter of the items in this part of the plan
- (B) All items in the test
- (C) Half the items in the test
- (D) No marks at all

Q37. A constructor writes 80 draft items for a 40-item test. The extra items allow

- (A) Automatic doubling of the pass mark
- (B) Selection of the best items after tryout
- (C) Removal of the blueprint
- (D) Skipping the tryout stage

Q38. In tryout, 90 percent answer an item correctly. The item is

- (A) Very difficult for the tryout group
- (B) Highly discriminating for all groups
- (C) Invalid for every purpose
- (D) Very easy for the tryout group

Q39. In tryout, high scorers and low scorers answer an item equally often correctly. The item has

- (A) Strong discrimination
- (B) Perfect difficulty
- (C) No distractors
- (D) Weak discrimination

Q40. A constructor finds a distractor that no tryout student chooses. The best action is to

- (A) Keep it because it fills space
- (B) Remove the correct answer
- (C) Increase the marks for the item
- (D) Revise or replace that distractor

Q41. A test is assembled with easy items first and harder items later. This order helps because it

- (A) Builds confidence and reduces early anxiety
- (B) Makes the test invalid
- (C) Changes the scoring key
- (D) Removes the need for instructions

Q42. A scoring key is checked by two markers before printing. This step mainly protects

- (A) Syllabus coverage alone
- (B) Scoring accuracy and reliability
- (C) Test length alone
- (D) Student attendance

Q43. A final test review should check that instructions state

- (A) Only the school name
- (B) Only the date of the test
- (C) The teacher's qualifications
- (D) What students must do, how to answer, and the time allowed

Q44. A test with items that all measure recall cannot validly claim to measure

- (A) Basic factual knowledge
- (B) Spelling accuracy
- (C) Reading speed
- (D) Higher-order thinking such as analysis

Basic Statistics for Teachers

Q45. Scores: 10, 20, 30, 40, 50. The mean, median and range are

- (A) 30, 20 and 50
- (B) 25, 30 and 40
- (C) 30, 30 and 50
- (D) 30, 30 and 40

Q46. Scores: 4, 4, 9, 9, 9, 12. The mode and median are

- (A) 4 and 9
- (B) 9 and 9
- (C) 9 and 4
- (D) 12 and 9

Q47. A set has mean 60 and median 45. The distribution is likely

- (A) Skewed to the right by high scores
- (B) Perfectly normal
- (C) Skewed to the left
- (D) Without any spread

Q48. If every score increases by 10 marks, the mean will

- (A) Stay the same
- (B) Increase by 100
- (C) Increase by 10
- (D) Become zero

Q49. If every score increases by 10 marks, the standard deviation will

- (A) Increase by 10
- (B) Decrease by 10
- (C) Stay the same
- (D) Double

Q50. A score of 70 with mean 60 and standard deviation 5 has a z-score of

- (A) 1
- (B) 2
- (C) 10
- (D) 0.5

Q51. A z-score of -2 in a normal curve is at about the

- (A) 50th percentile
- (B) 98th percentile
- (C) 2nd percentile region
- (D) 75th percentile

Q52. In a normal curve, the area beyond two standard deviations from the mean on both sides is about

- (A) 32 percent
- (B) 68 percent
- (C) 5 percent
- (D) 50 percent

Q53. A class has $Q1 = 40$ and $Q3 = 70$. The interquartile range is

- (A) 110
- (B) 40
- (C) 70
- (D) 30

Q54. A teacher wants to compare scores from two tests with different means and spreads. The fairest tool is the

- (A) Z-score
- (B) Raw total alone
- (C) Range alone
- (D) Mode alone

Evaluation Models

Q55. A Tyler-style evaluation report that only lists whether objectives were met is incomplete if it ignores

- (A) Unintended effects of the programme
- (B) The school name
- (C) The number of classrooms
- (D) The test date

Q56. In CIPP, a needs survey of parents and employers before a new course is mainly

- (A) Product evaluation
- (B) Process evaluation
- (C) Context evaluation
- (D) Grading

Q57. A CIPP team reviews lesson observations and attendance during a programme. This is

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

Q58. A CIPP team compares planned resources with resources actually received. This is closest to

- (A) Product judgement only
- (B) Input and process checking
- (C) Student grading
- (D) Context survey only

Q59. Tyler's model judges success after teaching, while CIPP also supports judgements

- (A) Before and during teaching
- (B) Only after graduation
- (C) Only in admission
- (D) Only for one lesson

Q60. A programme meets its objectives but causes heavy student stress. A Tyler-only report might

- (A) Automatically report the stress fully
- (B) Miss the stress effect because it was not a stated objective
- (C) Reject the objectives as invalid
- (D) Measure stress without data

Answer Key

Q1: B	Q11: C	Q21: C	Q31: C	Q41: A	Q51: C
Q2: A	Q12: B	Q22: D	Q32: B	Q42: B	Q52: C
Q3: A	Q13: D	Q23: B	Q33: A	Q43: D	Q53: D
Q4: C	Q14: C	Q24: D	Q34: A	Q44: D	Q54: A
Q5: D	Q15: B	Q25: A	Q35: C	Q45: D	Q55: A
Q6: A	Q16: A	Q26: D	Q36: A	Q46: B	Q56: C
Q7: B	Q17: B	Q27: B	Q37: B	Q47: A	Q57: B
Q8: C	Q18: D	Q28: C	Q38: D	Q48: C	Q58: B
Q9: D	Q19: C	Q29: C	Q39: D	Q49: C	Q59: A
Q10: A	Q20: A	Q30: D	Q40: D	Q50: B	Q60: B

Worked explanations for all 60 questions:

quaidianpoint.com/pedagogy-educational-measurement-evaluation-practice-test-9/