

Scoring the National Center for Teacher Effectiveness Student Assessment

Judy Hickman, Educational Testing Service
Heather Hill, Harvard Graduate School of Education

June 4, 2013

This document is intended as a guide to scoring student tests. The general approach is to create a raw score for students based on the performance on their specific test, then convert the raw score to a scale score. The scale score takes into account differences between the forms of the test, and ensures that the score has the same meaning no matter what form a given student took.

There are three steps in this process, described below.

NOTE: All omits are counted as incorrect

Step 1: Data entry and verification

Step 2: Convert performance on items to points that count toward the student's "raw score"

There are three types of items on this assessment: traditional multiple-choice, nested sets, and gridded response. For each, we discuss how to convert students' answer to points that count toward their raw score.

Multiple choice items. The document "scoring key_final" contains correct answers for all multiple-choice items. Each multiple choice item correct corresponds to one raw score point.

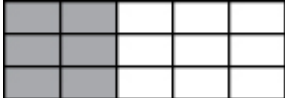
Nested sets. This instrument contains "nested sets" of items that would be scored using a rubric tied to the development model of the particular content. In other words, students would respond to a set of items, and the set would be scored as a single item (often a multiple-point item) rather than as discrete items. An example of this type of "nested set" is shown in Figure 1.


Figure 1. (Grade 4, Form G)

For numbers **21–24**, state whether or not each figure has $\frac{2}{5}$ of its whole shaded.

21.  (Y) Yes (N)

22.  (Y) Yes (N)

23.  (Y) Yes (N)

24.  (Y) Yes (N)

This item is more complex than a traditional multiple-choice item in that a student has to evaluate each part separately and decide whether the fraction $\frac{2}{5}$ can take different forms. Because two points are assigned to this problem, the item can also provide feedback at the item level as to the depth of understanding a student has about simple fractions. The total number of ways to respond to this item is sixteen. “Guessing” the correct combination of responses is much less likely than it would be for a traditional four-option multiple-choice item. The correct response for this item will receive 2 points, and the points will be earned based on the level of understanding the student has demonstrated. The scoring rubric used for this item follows in Figure 2.

Figure 2.

Scoring Rubric

Responses to this item will receive 0-2 points, based upon the following:

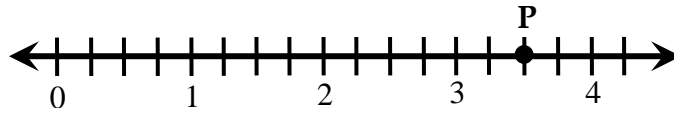
- 2 points: YNYN The student has a solid understanding of $\frac{2}{5}$ as well as an equivalent form of $\frac{2}{5}$.
- 1 point: YNNN, YYNN, YYYN The student has only a basic understanding of $\frac{2}{5}$. Either the student doesn't recognize an equivalent fraction for $\frac{2}{5}$ or doesn't understand that all 5 parts must be equal-sized in item #23.
- 0 points: YYYYY, YNNY, NNNN, NNYY, NYYN, NYNN, NYYY, YYNY, YNNY, NYNY, NNYN, NNNY The student demonstrates inconsistent understanding of $\frac{2}{5}$ or answers "Y" to item #25, clearly showing a misunderstanding of what $\frac{2}{5}$ means. Figure #25 is considered a "disqualifier," and an answer of "Y" to this part of the item would cancel out any other correct responses as "guesses" on the part of the student.

The document "Scoring Key_Final" provides information on how to convert patterns of responses to raw score points for each testlet item.

Gridded response. Finally, this test contains gridded-response items. The student essentially has to determine the solution and then grid-in a series of bubbles to show his/her answer. An example of this type of "gridded-response" item is shown in Figure 3.

Figure 3. (Grade 5, Form M)

27. Look at the following number line.



What decimal number is represented by point P ?

The student responds to this type of item by writing the answer (in this case, 3.5) in the boxes at the top of the grid and then filling in the corresponding circles below each number. The document “scoring key_final” contains correct answers for these items. Each item scored as correct corresponds to one additional raw score point.

0	0	0	0
1	1	1	1
2	2	2	2
3	3	3	3
4	4	4	4
5	5	5	5
6	6	6	6
7	7	7	7
8	8	8	8
9	9	9	9

Step 3: Converting Raw Scores to Scale Scores

Raw scores can then be converted to scaled scores using the following table.

Table 20.***Raw Score to Scale Score Conversion Table with Two Items Removed from Each Form***

Raw scores	Form F		Form G		Form H		Form J		Form K		Form L		Form M		Form N	
	Scale score	CSEM	Scale score	CSEM	Scale score	CSEM	Scale score	CSEM	Scale score	CSEM	Scale score	CSEM	Scale score	CSEM	Scale score	CSEM
0	40	24	40	22	40	21	40	22	45	23	45	21	45	22	45	21
1	61	13	57	14	47	18	46	19	62	14	62	13	49	20	59	14
2	71	8	67	9	61	12	62	12	73	9	72	9	66	12	70	9
3	76	7	73	7	70	9	70	9	79	7	78	7	74	8	76	7
4	80	5	77	6	75	7	76	7	83	6	82	6	79	7	80	6
5	83	5	80	5	79	6	80	6	86	5	85	5	83	6	84	6
6	85	4	83	5	83	5	83	5	89	5	87	5	86	5	87	5
7	87	4	85	4	85	5	85	5	91	5	90	4	89	5	89	5
8	89	4	87	4	88	4	87	4	93	4	92	4	91	5	91	5
9	91	4	89	4	90	4	89	4	95	4	94	4	94	5	93	4
10	92	4	91	4	91	4	91	4	97	4	96	4	96	5	95	4
11	94	3	92	4	93	4	93	4	99	4	97	4	98	4	97	4
12	95	3	94	4	95	4	95	4	101	4	99	4	100	4	99	4
13	97	3	96	4	97	4	96	4	102	4	101	4	102	4	101	4
14	98	3	97	4	98	4	98	4	104	4	102	4	104	4	102	4
15	100	3	99	4	100	4	100	4	106	4	104	4	106	4	104	4
16	101	4	100	4	101	4	101	4	107	4	105	4	107	4	105	4
17	103	4	102	4	103	4	103	4	109	4	107	4	109	4	107	4
18	104	4	103	4	105	4	105	4	110	4	108	4	111	5	108	4
19	106	4	105	4	106	4	107	4	112	4	110	4	113	5	110	4
20	108	4	106	4	108	4	109	4	114	4	111	4	115	5	111	4
21	110	4	108	4	110	4	111	4	116	4	113	4	118	5	113	4
22	112	4	110	4	112	4	113	5	118	5	114	4	120	5	115	4
23	114	4	112	4	113	4	116	5	120	5	116	4	123	5	116	4
24	116	5	114	4	115	4	119	5	122	5	118	4	125	6	118	4
25	118	5	116	5	117	4	122	6	124	5	120	5	128	6	120	5
26	121	5	118	5	120	5	126	7	127	6	122	5	132	7	122	5
27	125	6	121	5	122	5	132	9	130	6	124	5	137	8	125	5
28	130	8	124	6	125	6	145	17	134	7	127	6	142	9	128	6
29	139	13	128	7	129	7	160	26	138	8	131	7	150	12	132	7
30	160	24	134	9	134	9			144	10	135	8	164	18	136	8
31			145	15	146	16			155	14	141	10	170	21	143	10
32			160	22	160	25			170	20	153	15			155	15
33											170	23			170	22

Note. CSEM = conditional standard error of measurement.