

**NATIONAL AND LOCAL EDTPA RESULTS
BASED ON THE SPRING 2013 PILOT STUDY**

February 1, 2014

**School of Education
Contact Person: Dr. John Hoover, edTPA Coordinator
Osman Alawiye, Dean/ Kathy Dahlberg Assessment Director
Occasional Report No 14.1, February 1, 2014**

NATIONAL AND LOCAL EDTPA RESULTS BASED ON THE SPRING 2013 PILOT STUDY

Background Information

St. Cloud State University participated in the spring 2013 national pilot study, results of which are also available on the AACTE web site (EdTPA, 2013). Copies of this report are available electronically at the following location: <https://secure.aacte.org/apps/rl/resource.php?resid=313&ref=edtpa>. In addition, I can e-mail the report directly to members of the faculty who want it.

The Haertel (c.f., 2013; see also Cizek, 2012) system that considers both the reliability of the instrument, current performance on the tool, and primarily expert judgments of performance levels was employed by the SCALE partnership to set recommended performance standards (edTPA, 2013). This resulted in a professional performance standard recommendation of between 37 and 42 (39 is one SE_M below the raw cut score of 42, so I added that level):

During the edTPA standard-setting process (previously described), the practitioner panel and policy panel recommended a maximum score of no more than 42 as the professional performance standard that should be considered. The final panel (a subset of both the practitioner and policy panel) supported a similar cut score benchmark. Typically, in setting a cut score for a pass-fail decision, a standard error of measurement is applied to the recommended score so as to minimize erroneous decisions (e.g., false negatives). (EdTPA, 2013, p. 27)

Note that this recommended professional performance standard only includes instruments with scores running from 15 to 75, that is, assessments with 15 rubrics, thus excluding *Foreign Languages* from scoring. It is important to note that the Minnesota BOT has not set, nor have officials recommended a specific passing score; however, given the national pilot sample, the cut scores adopted by states is likely to fall between total scores of 37 and 42.

My calculation of the SE_M based on the national sample suggested a low point of 39. Thus, scores are calculated against three standards, the full AACTE/SCALE pass score of 42, the lower range suggested by SCALE (37) and my estimate of one SE_M below the mean, e.g., 39. [I can provide the calculation that produced a cut score of 39 upon request.]

Results

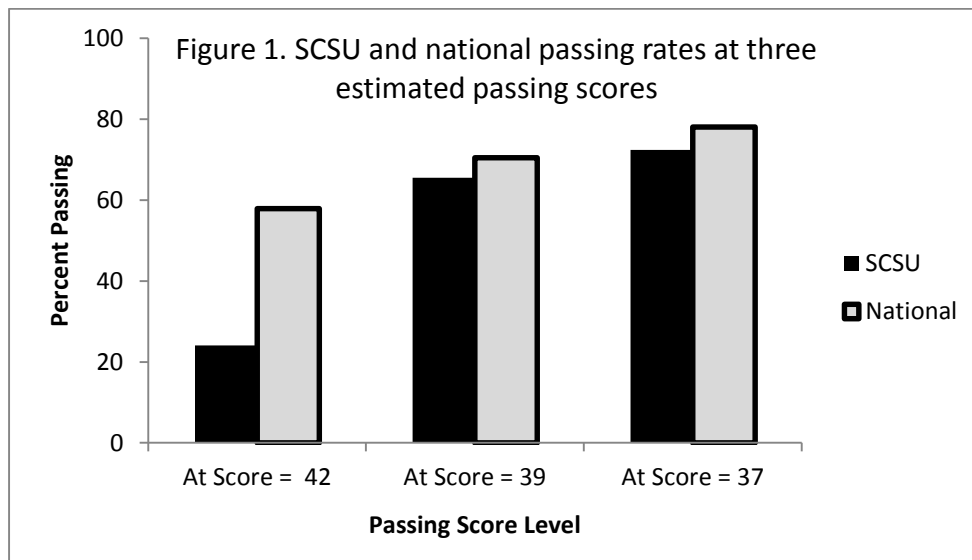
Passing rates by program. Passing rates for candidates in all fields except world languages are provided below, both for SCSU and for the national sample. The passing score for world languages can be calculated by first standardizing values (e.g., converting them to z scores) and then returning them to scalar values. Such manipulations must be interpreted with a great deal of caution and cannot be undertaken as yet because separate scores for WL were not provided in the initial report and we had too few cases with which to reasonably calculate them at SCSU. Table 1 shows SCSU passing rates by area. Passing rates for the institution are shown in Figure 1.

Table 1. Pass rates by area (minus world languages).

Field	N	M	N Pass (42)	% Pass (42)	N Pass (39)	% Pass (39)	N Pass (37)	% Pass (37)
Elementary Literacy	5	30.2	0	----	1	20.0	1	20.0
Elementary Mathematics	11	40.9	4	36.4	9	81.2	10	90.9
Elementary	16	25.0	4	25.0	10	62.5	11	68.7
Early Childhood (CFS)	3	40.7	1	33.3	2	66.7	2	66.7
Secondary English Language Arts	2	37.0	0	----	1	50.0	1	50.0
Physical Education	2	38.5	0	----	1	50.0	1	50.0
Performing Arts	1	37.0	0	----	0	----	1	50.0
Social Studies	2	35.0	0	----	1	50.0	1	50.0
Science	1	40.0	0	----	1	100.0	1	100.0
Visual Arts	2	50.5	1	50.0	2	100.0	2	100.0
Secondary-K-12	10	39.9	1	10.0	6	60.0	8	80.0
ALL SCSU	29	38.5	7	24.1	19	65.5	21	72.4
National Sample	----	42.8	----	57.9	----	70.5	----	78.0

As can be seen from Table 1 above, SCSU passing rates did not compare favorably with the national sample; we attained a low of 24.1 % at a passing score of 42 (national figure = 57.9%). About seven in 10 SCSU completers would have passed had a score of 37 been employed; the comparable national figure was 78%). If we employ the median figure of 39, then 66% of SCSU completers would have received teaching licenses (71% nationally).

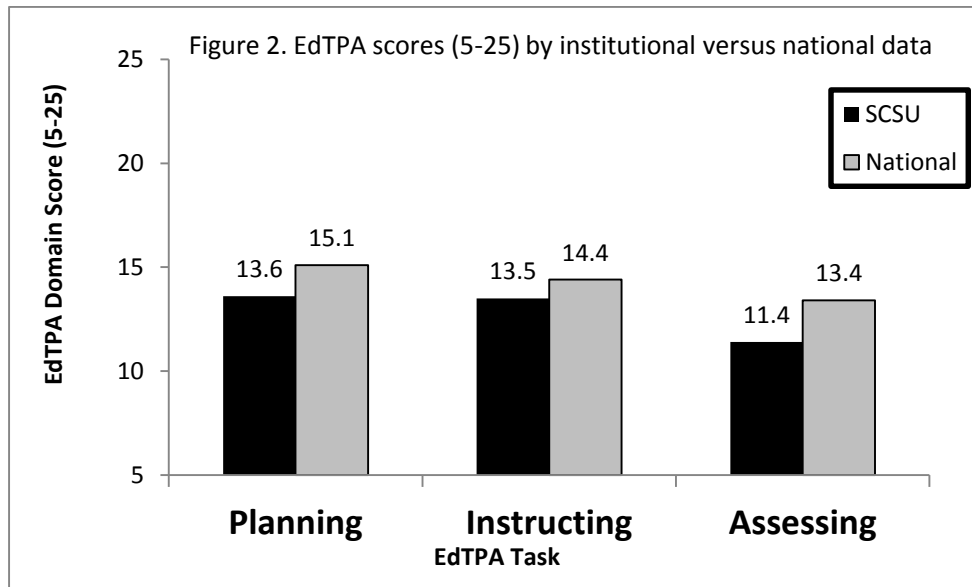
The same data are shown below in Figure 1. Despite the relatively small size of the sample, the difference (SCSU versus national performance) would likely prove significant at a passing rate of 42, but probably not at the two other passing levels.



Scores by domain. As can be seen in Table 2 and Figure 2, SCSU candidates scored qualitatively similar to members of the national sample. Specifically, SCSU candidates scored highest in Task 1 (Planning) and lowest in Task 3 rubrics (Assessing). As was true of the national sample, Tasks 1 and 2 probably did not differ, while Task 3 came in at a statistically significant lower value (e.g., Assessing vs Planning + Instructing, EdTPA, 2013).

Table 2. Scores by domain (all SCSU).

Area/Task	SCSU Sample Spring '13		National Sample Spring '13	
	Mean	SD	Mean	SD
Planning Instruction	13.57	3.29	15.06	3.15
Instructing	13.47	2.54	14.35	2.77
Assessing Instruction	11.37	3.21	13.36	3.39



Representatives of SCALE (2013) noted that candidates in the national sample scored significantly lower on *Assessment Task* rubrics than they did in the other two areas. St. Cloud State candidates scored lower generally than their national counterparts, while demonstrating a similar pattern. St. Cloud State candidates also scored lowest across the five *Assessment* rubrics.

The meaning of low scores for *Assessment* remains unclear. Several possibilities seem to exist. I have outlined these below:

- The rubrics for assessment show slightly lower reliability indices and slightly higher variance; thus, some of the difference may accrue to measurement error.
- Quite possibly, both SCSU students and the national sample evidence less or less effective preparation in assessment and thus the scores reflect real, that is to say validly-measured weaknesses suggesting the need for program improvements.

- Lower scores in assessment might reflect valid difficulties, not with assessment itself, but with assessment as measured via the EdTPA. Perhaps *Handbook* explanations need revision or the vernacular used to explain (and score) assessment practices does not match that employed in programs.
- Obviously, further analysis might reveal that an interaction between (low) reliability, poor performance, and EdTPA idiosyncrasies may produce the observed differences.

Fine-grained comparisons—SCSU. One last level of analysis is suggested by information in the *SCALE Pilot Study*. Evaluators noted that, not only did scores differ by task; they also differed by rubric. Table three shows all rubric scores for SCSU pilot subjects in reverse order by mean. Table 4 contains the same data from the manual.

To the extent that we consider the EdTPA values reliable and valid, data from Table 3 emphasize the above finding that SCSU candidates experienced the most difficulty with Task 3 *Assessment*. “Student use of feedback” and the analysis of language use proved most difficult at this level of analysis. On the other hand, content knowledge as reflected by the ability to plan for understanding appears as an emerging strength for candidates and for our program.

The rank order for candidates proved very similar to data from the national report. See Table 4.

Table 3. Score by individual rubric.

Rubric	Task	Explanation¹	Mean	SCSU Rank	National Rank
1	Planning	Planning: Planning for Subject-Specific Understandings	2.90	1	1
6	Instructing	Instruction: Learning Environment	2.90	2	2
7	Instructing	Instruction: Engaging Students in Learning	2.83	3	7
2	Planning	Planning: Planning to Support Varied Student Learning Needs	2.70	4	4
4	Planning	Academic Language: Identifying and Supporting Language Demands	2.70	5	6
5	Planning	Planning: Planning Assessments to Monitor and Support Student Learning	2.67	6	5
8	Instructing	Instruction: Deepening Student Learning	2.63	7	9
3	Planning	Analyzing Teaching: Using Knowledge of Students to Inform Teaching and Learning	2.60	8	3
9	Instructing	Instruction: Subject-Specific Pedagogy: Using Representations	2.57	9	11
10	Instructing	Analyzing Teaching: Analyzing Teaching Effectiveness	2.53	10	13
12	Assessing	Assessment: Providing Feedback to Guide Learning	2.53	11	8
15	Assessing	Analyzing Teaching: Using Assessment to Inform Instruction	2.40	12	12
11	Assessing	Assessment: Analysis of Student Learning	2.30	13	10
14	Assessing	Academic Language: Analyzing Students’ Language Use and Subject-Specific Learning	2.13	14	14
13	Assessing	Assessment: Student Use of Feedback	2.00	15	15

¹The rubrics differ slightly by discipline; please check the *Handbooks*.

Table 4. National scores by rubric.

Distribution of Scores across Rubrics and Tasks							
Task	Rubric	N	Average	SD	Median	Min	Max
Task 1: Planning	Rubric 1	3,669	3.15	0.75	3	1	5
	Rubric 2	3,669	2.98	0.87	3	1	5
	Rubric 3	3,669	3.02	0.75	3	1	5
	Rubric 4	3,669	2.95	0.71	3	1	5
	Rubric 5	3,669	2.96	0.84	3	1	5
Task 2: Instruction	Rubric 6	3,669	3.12	0.58	3	1	5
	Rubric 7	3,669	2.90	0.69	3	1	5
	Rubric 8	3,669	2.87	0.75	3	1	5
	Rubric 9	3,669	2.78	0.84	3	1	5
	Rubric 10	3,669	2.68	0.75	3	1	5
Task 3: Assessment	Rubric 11	3,669	2.83	0.86	3	1	5
	Rubric 12	3,669	2.90	0.88	3	1	5
	Rubric 13	3,669	2.38	0.82	2	1	5
	Rubric 14	3,669	2.52	0.74	3	1	5
	Rubric 15	3,669	2.73	0.91	3	1	5
Task 1: Planning		3,669	15.06	3.15	15	5	24
Task 2: Instruction		3,669	14.35	2.77	15	5	24
Task 3: Assessment		3,669	13.36	3.39	14	5	25
Total Score		3,669	42.76	8.17	43	19	70

References

Cizek, C. J. (Ed.) (2012). *Setting performance standards: Foundations, methods, and innovations*. New York: Routledge.

edTPA/ SCALE (2013). *2013 edTPA field test: Summary report*. Palo Alto, CA: Stanford Center for Assessment, Learning, and Equity (author).

Haertel, E. (2013). Getting the Help We Need. *Journal of Educational Measurement*, 50(1), 84-90. doi:10.1111/jedm.12002