

Measurement Challenges in On-going Testing Environment: Potential Solutions

Süleyman Olgar, Ph.D., Florida Department of Education
Lauren White, Ed.S., Florida Department of Education
Alvaro J. Arce, Ph.D., Pearson
Leah Kaira, Ed.D., Pearson

National Council on Measurement in Education (NCME)
April 16, 2018

www.FLDOE.org

1




Coordinated Session Outline

- Speaker 1: Leah Kaira, Ed.D.
 - *Topic 1: Reliability Estimates and Continuous Testing*
- Speaker 2: Alvaro J. Arce-Ferrer, Ph.D.
 - *Topic 2: Small Sample Size DIF Challenges in on-going Testing*
- Speaker 3: Lauren White
 - *Topic 3: Test Speededness: Potential Solutions*
- Speaker 4: Süleyman Olgar, Ph.D.
 - *Topic 4: Setting/Validating CBT Time Limits for Continuous Test Administrations: A Modified Standard Setting Approach*
- Discussant: Ahmet Turhan, Ph.D.

www.FLDOE.org

2




Session Overview

- Present measurement challenges as well as potential solutions in the areas of item try-out, testing time, and reliability; and introduce a modified standard setting methodology to establish or validate computer-based testing times.
- Research for these areas will be presented with recommendations, drawing from operational testing experiences to overcome these measurement challenges in a testing program operating continuously.

www.FLDOE.org

3



Session Overview Continued

- What are potential solutions to address challenges regarding test form reliability estimates for professional certification testing programs?
- What are potential solutions to improve the performance of differential item functioning detection with data collection designs prevalent in professional certification testing programs?
- What implications do psychometric qualities of standardized assessments have for test speededness?
- How can a modified standard setting methodology be utilized to establish or validate computer-based testing time limits in a continuous testing environment?

www.FLDOE.org

4



Test Speededness: Potential Solutions

Lauren White, Ed.S.
Florida Department of Education (FLDOE)

www.FLDOE.org

5



History of FTCE/FELE

- The Florida Teacher Certification Examinations (FTCE) is part of the educator certification process:
 - Subject area examinations
 - Professional Education
 - General Knowledge
- The first subject area examinations were administered in October 1988.
- Currently there are 39 subject area examinations.
- The Florida Educational Leadership Examination (FELE) is part of the administrator certification process.

www.FLDOE.org

6



FTCE/FELE Changes

A “game-changing” shift in educator certification testing in Florida from the previous standard of “minimally competent” to a new standard of “beginning effective”

www.FLDOE.org

7



Threshold Description I: Beginning Effective Teacher Candidate

A beginning effective teacher candidate (BETC) demonstrates sufficient knowledge of subject matter, curriculum, research-based instructional strategies, professional responsibilities and standards, and use of student data to inform practice so that the candidate has a **high likelihood** of **positively impacting** the learning, as well as the behavior and dispositions, of diverse students.

www.FLDOE.org

8



Threshold Description II: Effective Teacher Education Applicant

Effective applicants for teacher education will be able to demonstrate **college-level** language and writing skills appropriate to college-level prompts, to read and analyze complex college-level texts from multiple disciplines, and to demonstrate the ability to gather, analyze and interpret data and make data driven decisions.



FTCE/FELE Changes

More rigorous content expectations, test formats, and cut scores

Notable impacts:

- **General Knowledge (GK) Test**
- Elementary Education K-6
- Prekindergarten/Primary PK-3



General Knowledge (GK) Test

4-Subtest Basic Skills Test:

Subtest 1: Essay

Subtest 2: English Language Skills (ELS)

Subtest 3: Reading


Subtest 4: Mathematics

GK Reading Competency 3: The integration of information and ideas based on text selections.



Determining Testing Time


- Is test completion speed an integral test construct?
 - No – Testing time should be adequate for examinees to have equal opportunity to demonstrate skills and knowledge.
 - Yes – Testing time should be determined deliberately and empirically to support the validity of the outcomes and associated interpretations.


GK READING TESTING TIME COMPARISON

Basic Skills Reading Test/Subtest	Number of MC Items	Average Length of Reading Passages (word count)	Seat Time (Minutes)	Average Minutes per MC Item	Average Percent of Non-Responses to Last 5 Items
GK Reading Subtest New Generation (827)	40	200-300	40	1.00	15-20%
GK Reading Subtest Old Generation (823)	40	200-300	40	1.00	5-10%
NYSTCE ALST	40	900-1000	220*	*	0-5%
ILTS TAP Reading Subtest	60	200-300	300 (of 4 subtests)* 150 (single subtest)	*	0-5%
FRANIS Core Academic Skills for Educator Test	56	NA	85	1.52	NA
National Evaluation Series (NES) Essential Academic Skills Test	45	NA	60	1.33	NA
Washington Educator Skills Test	60	NA	120	2.00	NA
California Basic Educational Skill Test (CBEST)	50	75-115	82	1.64	5-1%
AVERAGE	49	NA	71	1.42	NA

* Testing time is provided for multiple subtests so no subtest level average minute per item is available.


13
www.FLDOE.org



Testing Time: Speededness Methodology

- < 90% examinees complete test**
 - Wilhelm & Schulze (2002)
- < 80% examinees complete test OR;**
< 75% of examination completed by all examinees
 - Secolosky (1989)
- ≤ 75% examinees complete test**
 - Bugbee & Bent (1992)

14
www.FLDOE.org



Florida Teacher Certification Examinations
Descriptive Statistics and Speededness Analysis Summary

Select Fields

Subject Area	Prior					Post				
	Avg. Form Difficulty	Pass Rate**	% Pass	Y/N	Speededness Level	Avg. Form Difficulty	Pass Rate**	% Pass	Y/N	Speededness Level
GK: Reading*	62.59	8,089	50%	Y	70%	65.39	6,382	63%	Y	70%
GK: ELA	70.34	8,027	68%	Y	75%	68.69	6,362	65%	Y	75%
GK: Mathematics	65.92	8,245	58%	Y	70%	65.67	6,475	56%	Y	80%
Biology	74.42	357	69%	N	N/A	73.19	269	70%	N	N/A
Mathematics	64.60	558	55%	Y	70%	64.60	446	55%	Y	70%

*Current generation examination content, rigor, and cut score with previous timing condition. Current timing condition began July 1, 2015.
**Pass rates based on first-time testing events for examinees
*** Number of examinees corresponds with pass rate data ONLY


15
www.FLDOE.org


Timing Analysis – Last 5 Questions Data

General Knowledge Reading Subtest Examinee Response Data

GK Reading Subtest Timing Analysis	40 Minutes (N=1,767)		55 Minutes (N=2,161)	
	Frequency	%	Frequency	%
Saw all questions	1271	71.9	2012	93.1
Did not see the last question	91	5.2	49	2.3
Did not see the last 2 questions	71	4.0	28	1.3
Did not see the last 3 questions	65	3.7	20	0.9
Did not see the last 4 questions	49	2.8	21	1.0
Did not see the last 5 questions	220	12.5	31	1.4
Did not see at least last 3 questions	334	18.9	72	3.3


16
www.FLDOE.org



Findings

- **% of examinees seeing all test items increased**
 - From 71.9% of examinees to 93.1%
- Pass rates increased
- GK Reading remains speeded
- Average form difficulty decreased
- Continued monitoring


www.FLDOE.org 17



Setting/Validating CBT Time Limits for Continuous Test Administrations: A Modified Standard Setting Approach

Süleyman Olgar, Ph.D.
Florida Department of Education (FLDOE)

www.FLDOE.org 18




Setting Testing Time

Setting realistic testing times in computer-based testing – issues and considerations:

- How do we set adequate time for a test?
- What to do if empirical evidence continuously shows large number of examinees failing to complete an examination within the given testing time?
- Should speededness be considered as a natural effect of certification testing?

www.FLDOE.org 19



General Knowledge (GK) Test

General Knowledge (GK) Test:

4-Subtest Basic Skills Test:

- Subtest 1: Essay (1 Writing Prompt – 50 min)
- Subtest 2: English Language Skills (40 MC items – 40 min)
- Subtest 3: Reading (40 MC items, 4-5 passages – 40 min)
- Subtest 4: Mathematics (45 MC items – 100 min)

www.FLDOE.org 20



Background – Speededness Study Findings

- The number of items range from 5 to 15 per passage.
- New generation GK Reading subtest average difficulty increased by 10%.
- The average minutes per MC item for the comparable basic skills Reading tests is 1.42.
- Regardless of background, 90% of the examinees used all the allotted testing time to complete the exam.
- 90% of the examinees are taking an average of about 15 minutes to read the 4 passages and respond to the first item associated with each passage.
- The findings suggest that 90% of the examinees would take about 51 minutes or less to complete the whole test.



Methodology

- Committee of Florida educators from K-12 and higher education institutions
- Educators' judgments to validate testing time with data collected from a quasi-experimental design manipulating testing time
- A method to set or validate testing time limits with a framework of reference based on standard-setting methodology (Cizek & Bunch, 2007)




Meeting Purpose

- Solicit judgments regarding the length of testing time needed for individuals, whom we would consider to demonstrate sufficient core knowledge of subject matter to meet effective teacher proficiency expectations for beginning teachers, to successfully complete a General Knowledge Reading subtest form.



Overview of the Meeting

- Provide two types of completion time judgments for an *Effective Teacher Education Applicant* who have a high likelihood of becoming a *Beginning Effective Teacher Candidate*:
 - a series of passage-level completion time judgments
 - a test-level completion time judgment




Meeting Activities

4 Steps:

- Test Taking and Time Recording
- Round 1 Passage-Level Completion Time Judgments
- Round 2 Passage-Level Completion Time Judgments
- Test-Level Completion Time Recommendations

www.FLDOE.org 25




Test Taking and Time Recording


Directions for SMEs:

- Read the passages then answer questions in order.
 - Circle your responses in the booklet.
- Record the **cumulative time** as seen on the screen following each passage, group of questions, or single type question.
- Make notes in the item booklet if necessary.
- After completing the test and time recording sheet, bubble in your responses on the answer documents.

www.FLDOE.org 26



Sample Time Recording Sheet




SME ID: _____

Florida Teacher Certification Examinations (FTCE)
General Knowledge Reading Timing Analysis
SME Time Recommendation Sheet for **PRACTICE ACTIVITY**

Please record the time on the following recording sheet by documenting the exact time displayed on the timer after finishing each of the specified jobs explained in the "Notes" column.

Questions	Notes	Time
Reading Passage (Training Manual, p.18)	Please record the time in the next column upon finishing the reading passage.	
Questions # 1-5	Questions 1 through 5 are related to the first reading passage. Please record the time in the next column after answering all of the questions about the passage.	

www.FLDOE.org 27



Training Activity

- Reading passage from Test Information Guide
- 5 questions representing blueprint and complexity/difficulty of items
- Independent activity
- Discussion/Calibration
- Clarification

www.FLDOE.org 28



Passage-Level Completion Time Judgment Procedures for Rounds 1 and 2

- Review the Competencies & Skills.
- Think about *Effective Teacher Education Applicants* and *Beginning Effective Teacher Candidates*.
- Consider the knowledge and skills that the effective teacher education applicants and beginning effective teacher candidates should be able to demonstrate and the **amount of time it would take them to read a passage, respond to group of questions or single type question.**
- Fill in your time estimation on the Completion Time Recommendation form corresponding to your passage-level completion time judgment.



Considerations While Making Judgments

“Imagine a hypothetical group of effective teacher education applicants who are just at the level of knowledge and skills required to demonstrate:

- English language skills appropriate to college-level and
- the ability to read and analyze complex college-level texts from multiple disciplines;
- and has a high likelihood of becoming a Beginning Effective Teacher Candidate (BETC).

What is the approximate amount of time it would take for this population to read the passage, answer the group of questions or a single type of question?”



Sample Passage-Level Recommendations

Florida Teacher Certification Examinations (FTCE)
General Knowledge Reading Subtest Form X
Timing Analysis (Round -2 Judgements)

Item #	SME ID #						Statistics		
	82701	82704	82707	82710	82713	82716	Median	Min	Max
Reading Passage (p.1)	4.0	3.5	3.0	3.5	2.0	4.0	3.50	2.00	4.00
Questions # 1-13	20.0	12.0	15.0	12.0	14.0	14.0	14.00	12.00	20.00
Question # 14	3.0	1.0	2.0	1.0	1.0	1.5	1.25	1.00	3.00
Reading Passage (p.6)	5.0	4.5	5.0	4.0	3.0	4.0	4.25	3.00	5.00
Questions # 15-29	24.0	14.5	17.0	15.0	17.5	17.0	17.00	14.50	24.00
Selections 1-2 (p.11)	3.0	2.0	3.0	2.0	1.5	2.5	2.25	1.50	3.00
Questions # 30-36	10.0	6.5	7.5	7.0	8.1	7.0	7.25	6.50	10.00
Passages 1-2 (p.14)	6.0	3.0	3.5	3.0	1.3	3.0	3.00	1.30	6.00
Questions # 37-40	8.0	4.0	5.0	4.0	3.5	5.0	4.50	3.50	8.00
SME Total Time	83.0	51.0	61.0	51.5	51.9	58.0	54.95	51.00	83.00



Test-Level Completion Time Recommendations

FLORIDA DEPARTMENT OF EDUCATION

Florida Teacher Certification Examinations
827 GK Reading Subtest
Test Completion Time Recommendation Summary (Round 2)
Friday, September 11, 2015

Form X	Form Y	Form Z																																								
Number of Scorable Items: 40 Number of SMEs: 6 Median Time: 55	Number of Scorable Items: 38 Number of SMEs: 6 Median Time: 52	Number of Scorable Items: 40 Number of SMEs: 5 Median Time: 51																																								
Distribution of Individual SME Results	Distribution of Individual SME Results	Distribution of Individual SME Results																																								
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr><th>SME ID</th><th>TIME</th></tr> </thead> <tbody> <tr><td>82701</td><td>83</td></tr> <tr><td>82704</td><td>51</td></tr> <tr><td>82707</td><td>61</td></tr> <tr><td>82710</td><td>52</td></tr> <tr><td>82713</td><td>52</td></tr> <tr><td>82716</td><td>58</td></tr> </tbody> </table>	SME ID	TIME	82701	83	82704	51	82707	61	82710	52	82713	52	82716	58	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr><th>SME ID</th><th>TIME</th></tr> </thead> <tbody> <tr><td>82702</td><td>63</td></tr> <tr><td>82705</td><td>57</td></tr> <tr><td>82708</td><td>44</td></tr> <tr><td>82711</td><td>48</td></tr> <tr><td>82714</td><td>53</td></tr> <tr><td>82717</td><td>51</td></tr> </tbody> </table>	SME ID	TIME	82702	63	82705	57	82708	44	82711	48	82714	53	82717	51	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr><th>SME ID</th><th>TIME</th></tr> </thead> <tbody> <tr><td>82703</td><td>51</td></tr> <tr><td>82706</td><td>52</td></tr> <tr><td>82709</td><td>66</td></tr> <tr><td>82712</td><td>36</td></tr> <tr><td>82715</td><td>49</td></tr> </tbody> </table>	SME ID	TIME	82703	51	82706	52	82709	66	82712	36	82715	49
SME ID	TIME																																									
82701	83																																									
82704	51																																									
82707	61																																									
82710	52																																									
82713	52																																									
82716	58																																									
SME ID	TIME																																									
82702	63																																									
82705	57																																									
82708	44																																									
82711	48																																									
82714	53																																									
82717	51																																									
SME ID	TIME																																									
82703	51																																									
82706	52																																									
82709	66																																									
82712	36																																									
82715	49																																									

www.FLDOE.org

FLORIDA DEPARTMENT OF EDUCATION

Sample Test-level Completion Time Recommendation Form

Name: _____ SME ID# _____

Signature: _____ Date: _____

Imagine a hypothetical beginning effective teacher candidate (BETC) who you would consider to demonstrate sufficient core knowledge of subject matter to meet effective teacher proficiency expectations for beginning teachers.

How much time (minutes) would a qualified effective teacher education applicant (ETEA), who have a high likelihood of becoming a beginning effective teacher candidate (BETC), likely need to complete a General Knowledge Reading subtest?

A)40 B)45 C)50 D)55 E)Other: _____

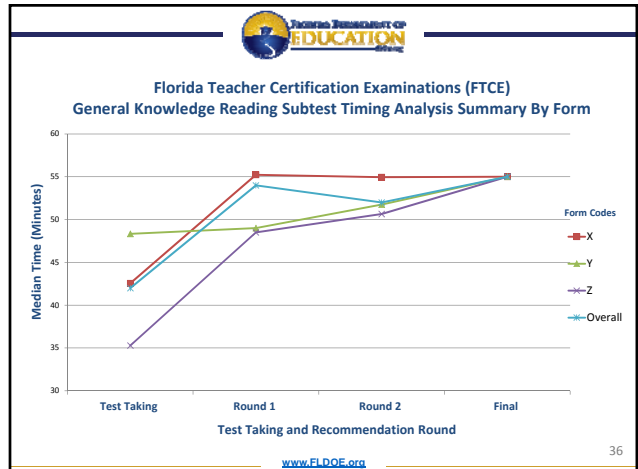
www.FLDOE.org


FLORIDA DEPARTMENT OF EDUCATION

Florida Teacher Certification Examinations (FTCE)
General Knowledge (GK) Reading Subtest Timing Analysis Summary

SME ID #	Form Code	Actual Test taking (Minutes)	Round 1 Recommendation (Minutes)	Round 2 Recommendation (Minutes)	Final Recommendation (Minutes)
82701	X	62	73	83	60
82702	Y	62	56	63	60
82703	Z	26	47	51	60
82704	X	40	91	51	55
82705	Y	53	69	57	55
82706	Z	35	65	52	55
82707	X	39	54	61	55
82708	Y	32	36	44	55
82709	Z	32	46	66	55
82710	X	39	30	52	55
82711	Y	58	61	48	55
82712	Z	43	49	36	55
82713	X	52	57	52	55
82714	Y	44	42	53	55
82715	Z	42	56	49	55
82716	X	45	49	58	55
82717	Y	35	41	51	55
Mean		43	54	54	56
Median		42	54	52	55

www.FLDOE.org





**FLORIDA DEPARTMENT OF
EDUCATION**
fldoe.org

Questions

Suleyman Olgar
Suleyman.Olgar@fldoe.org
Lauren White
Lauren.White@fldoe.org

www.FLDOE.org

37



www.FLDOE.org



www.FLDOE.org

38