

STATE OF FLORIDA
DIVISION OF ADMINISTRATIVE HEARINGS

JULIE MCCUE,

Petitioner,

vs.

Case No. 17-0423

PAM STEWART, AS COMMISSIONER OF
EDUCATION,

Respondent.
_____ /

RECOMMENDED ORDER

On June 13, 2017, Elizabeth W. McArthur, Administrative Law Judge, Division of Administrative Hearings (DOAH), conducted the hearing in this cause in Orlando, Florida.

APPEARANCES

For Petitioner: Robert F. McKee, Esquire
Robert F. McKee, P.A.
Suite 301
1718 East Seventh Avenue
Tampa, Florida 33605

For Respondent: Bonnie Ann Wilmot, Esquire
Darby G. Shaw, Esquire
Department of Education
Suite 1244
325 West Gaines Street
Tallahassee, Florida 32399

STATEMENT OF THE ISSUE

The issue for determination is whether Petitioner's challenge to the failing score she received on the essay section

of the Florida Educational Leadership Examination (FELE) should be sustained.

PRELIMINARY STATEMENT

Petitioner, Julie McCue (Petitioner), took the written performance assessment (essay) section of the FELE in September 2016. The score report she subsequently received showed that she did not earn a passing score, having received a score of six (on a scale of two to 12), when a score of seven was required to pass. Petitioner underwent the "score verification" process provided in statute and rule, and by letter dated November 22, 2016, the Department of Education (DOE) informed Petitioner of its determination that her essay had been scored correctly. Petitioner was informed of her right to an administrative hearing pursuant to sections 120.569 and 120.57, Florida Statutes (2017),^{1/} to dispute the decision.

Petitioner timely requested an administrative hearing, and the matter was transmitted to DOAH for assignment of an administrative law judge to conduct the requested hearing. In response to an Initial Order, the parties identified another DOAH case characterized as similar to this case, with overlapping witnesses for DOE. Consolidation was not requested, but DOE requested that if possible, the two hearings be coordinated and scheduled on back-to-back days. The parties also agreed to a "compromise" hearing location in Orlando and requested a live

hearing. The hearing was scheduled for March 13, 2017, in Orlando, by Administrative Law Judge J.D. Parrish, and the similar case was scheduled for March 14, 2017. Shortly thereafter, counsel for Petitioner filed his Notice of Appearance and DOE filed a Motion to Limit the Scope of Review in This Matter, followed by an amended motion. On February 22, 2017, counsel for Petitioner filed a Motion to Continue Hearing, asserting that, having newly appeared in the case, he needed time to conduct discovery and prepare. Counsel for Petitioner thereafter filed a response in opposition to DOE's pending motion to limit the hearing.

Judge Parrish granted Petitioner's motion for continuance and rescheduled the hearing for May 1, 2017, in Orlando. A continuance was also granted in what had become the companion case for scheduling purposes (DOAH Case No. 17-0424), and that hearing was reset for May 2, 2017.

On April 3, 2017, this case and the companion case were transferred to the undersigned. On April 5, 2017, the undersigned issued an Order denying DOE's motion to limit the scope of the hearing. Also on April 5, 2017, Petitioner filed a second motion for continuance, based on the uncertainty caused by DOE's motion while it had been pending, affecting such matters as the scope of permissible discovery and potential evidence to prepare for hearing. Since a similar DOE motion had been pending

in the companion case, also denied by Order issued on April 5, 2017, the undersigned scheduled a joint telephonic status conference, with counsel for parties in both cases participating. It was agreed that both hearings would be continued and rescheduled as soon as feasible, while allowing the parties sufficient time to complete discovery and hearing preparation. Based on the agreement of all parties regarding how much time was needed to prepare, the hearing in this case was reset for June 13, 2017, in Orlando (and the companion case was reset for June 14, 2017).

On June 5, 2017, a Joint Pre-hearing Stipulation was filed, in which the parties agreed to a few facts, incorporated below.

The parties entered into a Confidentiality Agreement on May 8, 2017 (Exhibit B to Petitioner's Motion for Clarification filed May 24, 2017, later withdrawn), to resolve confidentiality issues raised by DOE in its motion to limit the hearing. An agreed Motion for Protective Order was filed on June 6, 2017, to address the handling of confidential materials and testimony at the hearing. A Protective Order was issued on June 7, 2017.

On Friday afternoon, June 9, 2017, Petitioner's Attorney's Unopposed Motion to Withdraw as Counsel and Emergency Motion to Continue Hearing was filed. A telephonic motion hearing was held later that afternoon. At the conclusion of the hearing, counsel

for both parties and Petitioner, who also participated, were informed that both parts of the motion were denied.^{2/}

At the hearing, the parties offered Joint Exhibits 1 through 8, identified as confidential testing material subject to the Protective Order, which were admitted as such and are sealed. Petitioner testified on her own behalf. Petitioner's Exhibits 1 and 2, identified as additional confidential material subject to the Protective Order, were admitted as such and are sealed.

Respondent presented the testimony of the following witnesses: Christopher Small, a FELE chief reviewer; Michael Grogan, Pearson director of performance assessment scoring services; Phil Canto, DOE bureau chief of post-secondary assessment; Kelly Pelletier, a FELE chief reviewer; and Mary Jane Tappen, DOE vice chancellor for K-12 student achievement and student services. Respondent's Exhibits 1 through 10 and 12 (which are not confidential) were admitted in evidence.

As stated on the record, the undersigned took official recognition of the statutes and rules (including publications incorporated by reference) related to the FELE.

In addition to the confidential exhibits under seal, portions of the hearing were deemed confidential and the hearing room was cleared of persons not bound by the Protective Order. Those designated portions of the transcript are also under seal.

At the conclusion of the hearing, Petitioner requested an extended deadline of 30 days from the filing of the transcript to submit proposed recommended orders (PROs). DOE agreed and the request was granted. The two-volume Transcript of the hearing was filed on July 10, 2017. On August 2, 2017, Petitioner moved to extend the PRO deadline, and by amended motion, Petitioner clarified that the parties to the companion case agreed to the same extension. The amended motion was granted, and the PROs were timely filed by the extended deadline of August 30, 2017. The parties' PROs have been carefully considered in the preparation of this Recommended Order.

FINDINGS OF FACT

1. Petitioner is a teacher. She received her undergraduate degree in education with a major in social studies from Bowling Green State University in 1996. Since earning her bachelor's degree, she has taught history, psychology, and sociology over a 20-year span, at high schools in North Carolina, Ohio, and for the past three years, Florida.

2. Petitioner holds a Florida teacher certificate. She did not have to take an exam for that certificate. She likely was issued her Florida teacher certificate on the basis of the Ohio teacher certificate she held when she moved to Florida.

3. Petitioner aspires to add to her teacher certificate by attaining certification in educational leadership, which would require that she take and pass all subparts of the FELE.

4. Petitioner testified that in the district where she is employed as a teacher, she would qualify for a raise in her teacher's pay upon receiving a master's degree in educational leadership followed by DOE certification in educational leadership. Petitioner accomplished the first step by receiving a master's degree in educational leadership from Concordia University in Chicago, Illinois, in 2015.^{3/} She then initiated the process to take the FELE.

5. Educational leadership certification would also make Petitioner eligible for a leadership position, such as principal, vice principal, or a school district administrative leadership position, if she chooses to go that route. However, Petitioner's primary motivation in seeking this certification is for the additional compensation, and not because she wants an educational leadership position.^{4/}

6. Respondent, Pam Stewart, as Commissioner of Education, is the state's chief educational officer and executive director of DOE. §§ 20.15(2) and 1001.10(1), Fla. Stat.

7. One of DOE's responsibilities is to review applications for educator certification, and determine the qualifications of applicants according to eligibility standards and prerequisites

for the specific type of certification sought. See § 1012.56, Fla. Stat. One common prerequisite is taking and passing an examination relevant to the particular certification.

8. Respondent is authorized to contract for development, administration, and scoring of educator certification exams. § 1012.56(9)(a), Fla. Stat. Pursuant to this authority, following a competitive procurement in 2011, Pearson was awarded a contract to administer and score Florida's educator certification exams, including the FELE.

9. The State Board of Education (SBE) is the collegial agency head of DOE. § 20.15(1), Fla. Stat. As agency head, the SBE was required to approve the contract with Pearson. The SBE is also charged with promulgating certain rules that set forth policies related to educator certification, such as requirements to achieve a passing score on certification exams. DOE develops recommendations for the SBE regarding promulgating and amending these rules. In developing its recommendations, DOE obtains input and information from a diverse group of Florida experts and stakeholders, including active teachers and principals, district administrators, and academicians from colleges and universities.

FELE Essay Development and Scoring

10. DOE develops the FELE, as well as the other educator certification exams, in-house. The FELE is developed and periodically revised to align with SBE-promulgated standards for

educational leadership, as well as SBE-promulgated generic subject area competencies. In addition, as required by statute, certification exams, including the FELE, must be aligned to SBE-approved student standards.

11. Details about the FELE, such as the applicable generic competencies, the exam organization, and passing score requirements, are set forth in Florida Administrative Code Rule 6A-4.00821 (the FELE rule). The FELE rule has been amended periodically, but the current version includes a running history, setting forth FELE details that applied during past time periods, as well as those currently in effect.

12. The FELE consists of three subtests. Subtest one is a multiple choice test covering the area described as "Leadership for Student Learning." Subtest two, also a multiple choice test, covers "Organizational Development." Subtest three covers "Systems Leadership," and has two sections: a multiple choice section; and a written performance assessment, or essay, section.

13. The FELE has contained an essay component for many years (as far back as any witness could remember). Before January 2015, the essay score was included in a single composite score given for subtest three. The multiple choice part accounted for most of the weight of the composite score (70 percent); the essay portion accounted for 30 percent of the composite score.

14. Based on input from educators, academicians, and other subject matter experts, DOE recommended that the FELE subtest three be changed by establishing separate passing score requirements for each section, thereby requiring examinees to pass each section. The SBE adopted the recommendation, which is codified in the FELE rule, and has applied to FELE scoring since January 1, 2015. The effect of the change is that an examinee not as proficient in effective written communications can no longer compensate for a weak essay with a strong performance on the multiple choice section. To a lesser extent (given the prior 70:30 weight allocation), the reverse is also true.

15. The policy underlying this scoring change is to give more emphasis to testing writing skills, in recognition of the critical importance of those skills. By giving heightened scrutiny to writing skills, the FELE better aligns with increasingly rigorous SBE-approved student standards for written performance. This policy change is reasonable and within the purview of the SBE; in any event, it is not subject to debate in this case, because Petitioner did not challenge the FELE rule.

16. The generic competencies to be demonstrated by means of the FELE are set forth in the publication "Competencies and Skills Required for Certification in Education Leadership in Florida, Fourth Edition 2012," adopted by reference in the FELE rule and effective as of January 1, 2014.

17. The competency and skills generally tested by the FELE written performance assessment are:

Knowledge of effective communication practices that accomplish school and system-wide goals by building and maintaining collaborative relationships with stakeholders

1. Analyze data and communicate, in writing, appropriate information to stakeholders.
2. Analyze data and communicate, in writing, strategies for creating opportunities within a school that engage stakeholders.
3. Analyze data and communicate, in writing, strategies that increase motivation and improve morale while promoting collegial efforts.

18. This generic description provides a high-level view (aptly described as from the 30,000-foot level) of the competency and skills that an educational leader should possess, which are tested by the written performance assessment. DOE's job is to distill those qualities down to a test. As reasonably summarized by DOE's witnesses, the purpose of the FELE written performance assessment, as established by the SBE, is to test for effective written communication skills, and data analysis that drives appropriate strategies for improvement. These overall concepts are built into the general FELE rubric which serves as a guide to scoring, the individual essay prompts, and the supplemental rating criteria (essentially prompt-specific rubrics, making the general rubric specific to each essay prompt).

19. The FELE rule sets forth requirements for how the "test scoring agency" (Pearson) must conduct the scoring of the written performance assessment:

(a) Raters Judges. The test scoring agency shall appoint persons to score the written performance assessment who have prior experience as educational leaders, instructional leaders, or school building administrators.

(b) Chief Raters. The chief raters shall be raters who have prior experience as educational leaders, instructional leaders, or school building administrators and have demonstrated success as raters.

20. Pursuant to Pearson's agreement with DOE, DOE retains the right to approve raters who will be scoring the written performance assessments. Therefore, Pearson proposes raters who meet the specified qualifications, and then DOE approves or disapproves the proposed raters. Approved raters must undergo training before they are appointed by Pearson to conduct scoring.

21. There is currently one chief rater for the FELE written performance assessment. The chief rater was a rater before being trained for, and assuming, the chief rater position. The chief rater was trained by Florida DOE chief raters when Pearson became the contractor and the scoring was transitioned to Pearson's offices in Hadley, Massachusetts, during 2012 to 2013.

22. Pearson employs holistic scoring as the exclusive method for scoring essays, including FELE written performance

assessments (as specified in Pearson's contract with DOE). The holistic scoring method is used to score essay examinations by professionals across the testing service industry. Pearson has extensive experience in the testing service industry, currently providing test scoring services to more than 20 states.

Dr. Michael Grogan, Pearson's director of performance assessment scoring services and a former chief rater, has been leading sessions in holistic scoring or training others since 2003. He described the holistic scoring method as a process of evaluating the overall effect of a response, weighing its strengths and weaknesses, and assigning the response one score. Through training and use of tools, such as rubrics and exemplars, the evaluation process becomes less subjective and more standardized, with professional bias of individual raters minimized, and leading to consistent scoring among trained raters. Training is therefore an integral part of Pearson's testing services for which DOE contracted. In an intensive two-day training program conducted by the chief rater in Hadley, prospective raters are trained in the holistic scoring method used to score FELE essays.

23. Pearson's rater training program begins with a review of background about the holistic scoring method generally, including discussions about rater bias. From there, trainees are oriented to the FELE-specific training material. They thoroughly review and discuss the rubric, the score scale, the operational

prompt raters will be scoring, and exemplars (other responses to the prompt that have been pre-scored). The rater candidates then employ these tools to begin independently scoring exemplars. Raters-in-training conduct many rounds of independent scoring sessions, interspersed with group discussions regarding how the essays should have been scored. The trainees then move into the calibration test phase, in which they independently score essay exemplars, paired with an experienced rater who independently scores the same exemplars. The trainees score essay after essay, then compare scores with the experienced rater, with the goal to achieve consistency in scores, by equaling or coming within one point of the other rater's score. Ultimately, the raters must pass the calibration test by achieving scoring consistency to qualify for appointment as raters to score actual FELE essays.

24. Each FELE essay is scored independently by two DOE-approved raters who meet the qualifications in the FELE rule and who have successfully completed training. Pairs of raters receive scoring assignments, one prompt at a time. The assignments are received anonymously; one rater does not know who the other assigned rater is. And neither rater knows anything about the examinee, as the essay is identified solely by a blind number. FELE essay raters work in one room, at individual computer terminals, in Hadley. Security of all testing

information is vigilantly maintained, through confidentiality agreements and secure, limited, and protected computer access.

25. For each scoring assignment, raters adhere to a step-by-step process that reinforces their initial training. Raters must first score sample responses to a historic prompt that is different from the assigned prompt, as a training refresher to invoke the holistic scoring mindset. From there, raters review the assigned prompt and the scoring guides (general rubric and supplemental rating criteria). Raters then must score an anchor set of six sample responses, one exemplifying each score category; the historic scores are not revealed until the raters complete their scoring. Raters compare their scores with the anchor scores, and work through any discrepancies. Raters then go through a calibration process of scoring 10 more sample responses to the same prompt. After scoring all 10 essays, the raters learn the scores deemed appropriate for those responses, and must work through any discrepancies until consistency is achieved. Only after scoring many sample essays and achieving success in scoring consistency are the raters permitted to turn to the assigned FELE essay for review and scoring.

26. The chief rater supervises and monitors the raters while they are engaged in their scoring work. The chief rater is physically present in the same room with the raters, monitoring their work online in real time. As raters enter scores, those

scores are immediately known by the chief rater, so that any "red flag" issues in scoring results and trends can be addressed immediately. As another tool, "ghost papers," which are pre-scored essays, are randomly assigned to raters as if they are actual FELE essays. The chief rater monitors ghost paper scoring as another check on consistency with a predetermined measure.

27. The scores of the two raters assigned to score a FELE essay are added together for the total holistic score. Thus, the total score range for a FELE essay is between two points and 12 points: the lowest possible score of two points would be achieved if each rater assigns a score of one point; and the highest score of 12 points would be achieved if each rater assigns six points.

28. The sum of the two raters' scores will be the score that the FELE essay receives unless the raters' scores disagree by more than one point. If the two raters' scores differ by more than one point, then the chief rater steps in to resolve the discrepancy.

29. After FELE essays are scored, the examinee is informed of the final score of between two and 12 points, and the examinee is told whether the score is a passing or failing score. Seven points is a passing score, according to the FELE rule.

30. Raters do not develop written comments as part of their evaluation of FELE essays. Their holistic evaluation is expressed by the point value they assign to the essay.

31. Through the intensive training and the subsequent calibration and recalibration before each FELE essay scoring assignment, Pearson has achieved excellent consistency in rater scoring of the FELE written performance assessment. From September 12, 2016, through October 8, 2016, the four Pearson raters who were scoring FELE essays (including Petitioner's essay) achieved a coefficient alpha index of 98 percent, meaning that 98 percent of the time, the scores assigned to an essay by a pair of raters were either identical or adjacent (within one point), and when adjacent, were balanced (i.e., each rater was as often the higher scorer as he or she was the lower scorer). This exceeds industry standards. A comparable, high coefficient alpha index was achieved by FELE essay raters for each month in 2015 and 2016. The lowest coefficient alpha index, still exceeding industry standards, was 93 percent in a single month (February 2015). In two months (December 2015 and July 2016), the coefficient alpha index was 94 percent, with the remaining 21 months at between 95 percent and 98 percent.

Examinee Perspective: Preparation for the FELE Essay

32. DOE provides detailed information and aids on its website regarding the FELE, including the essay section, for potential examinees. This includes a 40-page test information guide for the FELE. The test information guide contains all of the SBE-adopted competencies and skills, including the competency

and skills tested by the written performance assessment. The guide also contains the general FELE essay scoring rubric, and a sample prompt that is representative of the essay prompts actually used. DOE also posts on its website three additional sample FELE essay prompts along with the supplemental rating criteria that correspond to those prompts.

33. Petitioner does not challenge the appropriateness of these materials generally, which she accessed and used to prepare for the FELE written performance assessment. However, Petitioner complained that DOE does not provide more study guide materials or endorse specific vendors of study guide materials so as to more thoroughly prepare potential examinees for their essay tests. Petitioner also complained that when an examinee fails an essay test, DOE does not provide substantive explanations to help the examinee understand the reasons for the failing score and how the examinee can perform better. DOE appropriately responded to this criticism by reference to standards for testing agencies adopted by three authoritative bodies: the American Educational Research Association, the American Psychological Association, and the National Council of Measurement Education. These standards dictate that as testing agency, DOE's responsibility is to develop tests that evaluate whether individuals are prepared with the necessary skills. It is not DOE's responsibility, and it would not be appropriate for DOE, as the testing agency, to

prepare individuals to pass its tests, or coach individuals on how to perform better on tests they do not pass.

34. The information DOE makes publicly available is appropriate and sufficient to explain the FELE essay exam and scoring process, and to allow an examinee to know what to expect in a prompt and what is expected of the examinee in a response. The DOE test information guide explains the FELE essay and scoring process, as follows:

Your response will be scored holistically by two raters. The personal views you express will not be an issue; however, the skill with which you express those views, the logic of your arguments, the quality of your data analysis and interpretation, and the appropriateness of your implementation plans will be very important in the scoring.

Your response will be scored on two constructs: communication skills, including ideas, focus, organization, and mechanics (capitalization, punctuation, spelling, and usage) and data analysis, interpretation, and evaluation, including data explanation, application, relevant implications, and analysis of trends.

The raters will use the criteria on the following page when evaluating your response. The score you receive for your written performance assessment will be the combined total of the two raters' scores. (R. Exh. 2 at 13 of 40).

35. On "the following page" of the test information guide, the general FELE essay rubric is set forth in its entirety. The rubric is also available on the DOE website as a separate, stand-

alone document. The rubric is simply a comparative description of the extent to which an essay demonstrates the generic competency and skills to be tested--effective written communication skills, with data analysis that drives appropriate strategies for improvement. For example, recognizing that part of effective written communication is use of proper grammar and syntax, the rubric describes that quality comparatively, differentiating between best, better, good, not-so-good, worse, and worst. Similarly, the rubric addresses whether proposed strategies are appropriate by comparing the extent to which the strategies are aligned with the data findings, relevant implications, and trends. But these are just parts--and not discrete parts--of the evaluation. As explained in the test information guide, holistic evaluation judges the overall effect of a response, considering all aspects of effective communication and data analysis, in a process of weighing and balancing strengths and weaknesses.

36. Of course, DOE does not make publicly available those essay prompts being used in FELE tests, or the supplemental rating criteria for those prompts; these are protected, confidential testing material. It would be unreasonable for examinees to expect more from a testing agency than what DOE makes available.

Score Verification

37. An examinee who fails the written performance assessment (or any other FELE subtest or section) may request score

verification, to verify that the failed exam was scored correctly. The score verification procedures are set forth in the FELE rule.

38. The score verification rule provides that DOE makes the determination as to whether an examinee's test was scored correctly. DOE is authorized to consult with field-specific subject matter experts in making this determination. In practice, though not required by the FELE rule, when a score verification request is directed to the score assigned to a FELE written performance assessment, DOE always consults with a field-specific subject matter expert known as a "chief reviewer."

39. Chief reviewers are another category of experts (in addition to raters and chief raters) proposed by Pearson pursuant to qualifications identified by DOE, subject to DOE approval. Once approved by DOE, prospective chief reviewers undergo the same rater training in the holistic scoring process as do all other raters, to gain experience in scoring essays and undergo calibration to achieve scoring consistency. In addition, chief reviewers are given training for the chief reviewer role of conducting review and scoring of essays when scores have been contested.^{5/} Unlike raters and chief raters, chief reviewers do not work at Pearson in Hadley, Massachusetts; they are Florida experts, actively working as principals of Florida schools.

40. Chief reviewers only become involved when an examinee who failed the FELE written performance assessment invokes the

score verification process. A chief reviewer is assigned to evaluate whether that essay was scored correctly. The chief reviewer conducts that evaluation by first going through the same step-by-step process as raters, following the same retraining and calibration steps that involve scoring many sample essays. Upon achieving success in the calibration test, the chief reviewer moves on to evaluate the assigned essay response independently, before reviewing the scores the raters gave to that essay. Upon reviewing the raters' scores, the chief reviewer offers his or her view as to whether the essay score should stand or be changed, and provides a summary rationale for that opinion. This information is conveyed to DOE, which determines the action to take--verify or change the score--and notifies the examinee of the action taken.

Petitioner's FELE Attempts

41. Petitioner took all parts of the FELE for the first time in the summer of 2015, in June and July. She passed subtest one, but failed subtest two and both sections (multiple choice and written performance assessment) of subtest three.

42. FELE examinees can retake failed subtests/sections, and need only retake the parts failed. There are no limits on the number of retakes. The requirements for retakes are that at least 30 days must have elapsed since the last exam attempt, and that examinees pay the registration fees specified in the FELE rule for each retake of a failed subtest and/or section.

43. On April 23, 2016, roughly nine months after her first attempt, Petitioner retook subtest two and both sections of subtest three. To prepare, Petitioner used the "very limited" resources on the DOE website, and purchased some "supplementals," which she described as materials "on the market that supposed FELE experts sell." (Tr. 33). She used the material to study and practice writing essays. Petitioner passed subpart two and the multiple choice portion of subpart three. However, she did not pass the written assessment section of subpart three.

44. Petitioner retook the written performance assessment 33 days later (May 26, 2016), but again, did not pass.

45. Petitioner did not invoke the score verification process to question the failing scores she received on her first three FELE essays. Those three failing scores stand as final, as she did not challenge them. Petitioner explained that she did not challenge them because she was embarrassed, because as a teacher, she believed that she would pass the test. However, while Petitioner has had many years of success as a teacher, the skills for teaching do not necessarily correlate to the skills required for educational leadership positions, as several DOE witnesses credibly attested.

46. Nonetheless, Petitioner tried again, in an effort to qualify for the pay raise her district would provide. She retook the FELE essay section for the fourth time on September 28, 2016.

Petitioner testified that, as she had done before, she reviewed the material on DOE's website, such as the test information guide with its general rubric, and she practiced writing essays using the sample essay prompts and supplemental rating criteria. In what was described as a "eureka moment," she also found what she described as "the rubric" on the website, which she proceeded to memorize. Rather than the rubric, however, what Petitioner memorized was the generic competency and skills tested by the written performance assessment. Petitioner made a point of incorporating words from the competency and skills document in her essay. Petitioner did not pass.

47. Each of the four times Petitioner took the FELE written performance assessment, including the most recent attempt at issue in this case, both raters assigned to score her essay gave the essay three points, for a total score of six points. Since in each of her four attempts, Petitioner's essay was scored the same by both raters, Petitioner's essays were never reviewed by a chief rater, because there was never a discrepancy in the raters' scores for the chief rater to resolve.

Petitioner's Challenge to Her Fourth Six-Point Essay Score

48. When Petitioner was notified that her fourth essay attempt resulted in the same score--six, on a scale ranging from two points to 12 points--this time Petitioner took the next step, by requesting a score verification session.

49. Following the procedures in the FELE rule for score verification, Petitioner registered, paid the required fee, and went to the designated Pearson site. There, she was able to review the essay prompt, as well as her written response.

50. Petitioner testified that she prepared a "statement of specific scoring errors" (so named in the FELE rule--more aptly, in her case, a statement explaining why she thinks her essay score was erroneous), which she submitted to Pearson at the end of her session. By rule, the statement is then filed with DOE.

51. The statement Petitioner prepared was not offered into evidence, apparently by choice, as Petitioner was looking for it at one point, stating that it was "part of the confidential stuff" (Tr. 78) that had been produced by DOE.

52. Petitioner attempted to describe the statement of scoring errors that she recalls completing. She described it as primarily demonstrating where in her essay she addressed what she characterized as the "rubric" that she had found on DOE's website and memorized. As noted previously, this was not the rubric, but rather, was the high-level description of the competency and skills tested by the FELE written performance assessment. As described, Petitioner's statement explaining that she "memorized" the competency/skills ingredients, and showing where she included competency/skills buzz-words in her essay (e.g., "morale"; she also said "celebration," but that word does not appear in the

competency/skills), would not seem to be the sort of statement that would be persuasive as to a claim of an erroneous score. It would be a mistake to memorize and repeat words from the generic competency/skills without regard to whether they are used in a way that makes sense in the responding to the specific instructions of the essay prompt.

53. DOE conducted its review, and the score was verified through a process consistent with DOE's practice of consulting a chief reviewer retained by Pearson with DOE approval, who was qualified as a subject matter expert in the field of Florida educational leadership. The assigned chief reviewer was also qualified by Pearson training in the holistic scoring method and in conducting score verification reviews.

54. The chief reviewer who undertook to verify Petitioner's essay score did not review Petitioner's statement explaining why she believed her essay score was erroneous. Instead, he independently evaluated Petitioner's essay, following the same holistic method, including the step-by-step retraining and calibration process, used by all raters to score a FELE essay. Then the chief reviewer reviewed the scores separately assigned by the two raters who scored Petitioner's essay. He concluded that the assigned scores of three were appropriate for Petitioner's essay, and that no change should be made. The chief reviewer provided a summary rationale for his determination.^{6/}

55. Petitioner complains that the chief reviewer should have been given her statement explaining why her score was erroneous, because that might have affected the chief reviewer's decision. However, pursuant to the FELE rule, the chief reviewer's role is consultative only; DOE makes the determination of whether Petitioner's essay was scored correctly, which is why the rule provides that the statement of asserted scoring errors is filed with DOE. Petitioner presented no evidence proving that DOE did not consider Petitioner's statement explaining why she believed her essay score was erroneous. No testimony was offered by a witness with personal knowledge of any review given to Petitioner's statement; that review would have been done by a member of DOE's "scoring and reporting team" (Tr. 260-261), none of whom testified. If Petitioner had proven that the statement was not considered by DOE, the failure to offer that statement into evidence would make it impossible to determine the import, if any, of such failure.

56. Petitioner was notified by DOE that the "essay score that you questioned has been reviewed by a Chief Reviewer. As a result of this review, the Department has determined that the written performance section that you questioned is indeed scored correctly." Petitioner was informed that if she was not satisfied with the outcome, she was entitled to dispute the decision pursuant to sections 120.569 and 120.57. Petitioner

availed herself of that opportunity,^{7/} and was given the chance in a de novo evidentiary hearing to present evidence to support her challenge to her exam score.

57. At the hearing, Petitioner offered only her own testimony as support for her challenge to the scoring of her essay. She isolated portions of the supplemental rating criteria and attempted to identify where her essay addressed the isolated portions, for which, in her view, she ought to have been awarded "a point" here or "a half-point" there. She also referred to isolated parts of the summary comments from the raters and chief reviewers, and attempted to identify the parts of her essay that did or did not do what the comment portions stated.

58. Petitioner was not shown to be, tendered as, or qualified as an expert in either educational leadership or holistic scoring of essays. Her attempt to tally points by comparing isolated parts of the prompt-specific rubric to isolated parts of her essay is contrary to the holistic scoring approach used to score the FELE written performance assessment. Petitioner offered no comprehensive, holistic evaluation of her essay as a whole, nor was she shown to be qualified to do so.

59. Besides being contrary to the holistic scoring method, Petitioner's critique of the scoring of her essay was wholly unpersuasive. Without undermining the confidentiality of the ingredients of Petitioner's testimony (the essay prompt, her

essay, the supplemental rating criteria, and the historic anchors), overall, the undersigned did not find Petitioner's critique credible or accurate. Although awkward to try to explain in code, some examples follow to illustrate the basis for this overall finding.

60. As one example, Petitioner referred to data points that the prompt-specific rubric indicated should be identified in response to the prompt. If a "data point" that should have been identified was that A was consistently lower than B, Petitioner called attention to a part of her essay identifying A as low. She acknowledged that her essay did not expressly compare A to B at all, much less over time, but Petitioner argued that those comparisons were implicit. She said that she should have gotten at least a half-point for partially identifying the data point. That argument is rejected. The point that needed to be made was a comparative assessment over a time span.

61. Where another data point called for identifying that two things were "substantially lower" than other things, Petitioner said that she sufficiently identified this point by saying that one of those two things was "lowest" (or "worst"). However, the point that needed to be made was not just that something was lowest or worst, but also, that another thing was also lower, and that the degree of separation between those two things and other things was substantial.

62. Overall as to the data points, Petitioner failed to identify several significant trends, and failed to offer sufficient comparative analysis as to the trends she did identify. She reported data or averages of data without identifying the relevant implications of the data, as would have come from making the appropriate comparisons and identifying the appropriate trends. In terms of the competency/skills language, she did not analyze the data and communicate, in writing, appropriate information to the stakeholders identified in the prompt as the target audience.

63. The data point failures were particularly problematic when taken to the next step of proposing specific strategies that would lead to improvement in the areas shown to be needed from the data points. For example, Petitioner's failure to identify the second data point in the supplemental rating criteria resulted in Petitioner proposing action that was at odds with what the second data point showed.^{8/}

64. Petitioner's attempted critique of her essay score was riddled with other inconsistencies. For example, Petitioner acknowledged that she often failed to summarize specific data for each of the three years, choosing instead to provide three-year averages. Petitioner's explanation was that she did not want to repeat data in the prompt because that would be condescending to her target audience. This is a weak rationale, one which is at

odds with the instructions given with the prompt. Petitioner also said it should have been a positive that instead of just citing yearly numbers, she went to the trouble of calculating three-year averages. Instead, it appeared more negative than positive, by masking information needed to respond to the prompt.

65. While Petitioner defended her omission of specific data because of the target audience she was instructed to address, Petitioner inconsistently sought to explain an odd statement using the word "celebrated" (Jt. Exh. 3 at 1, first sentence of second paragraph) as being directed more to certain other stakeholders than to the target audience. She did this because the "rubric" (i.e., the competency/skills), said to communicate to stakeholders, and also "talks about morale and celebration." (Tr. 59). This is an example of Petitioner's ineffective strategy of throwing out words from the competency/skills in ways that were contrary to specific instructions in the prompt. The target audience identified in an essay prompt may be certain stakeholders, instead of all stakeholders. For example, the sample prompt in the test information guide (R. Exh. 2 at 34), instructs the writer to prepare a memorandum for school advisory council members. The use of the word "stakeholders" in the competency/skills would not justify ignoring the essay prompt instructions by writing with a communication style more suited to a different audience of other stakeholders.

66. Petitioner disagreed with the suggestion in both chief reviewers' written comments that the essay's responses to the third and fourth bullet points in the prompt (Jt. Exh. 1) were generalized, lacking specifics and examples. Petitioner failed to persuasively establish that her essay provided sufficient detail in this regard to avoid being fairly characterized as responding to these bullet points with "generalizations." By failing to adequately analyze the data, relevant implications, and trends, Petitioner's responses to these bullet points were either too general (e.g., research to find strategies), or in the one instance where specific action was described, the action was at odds with data points she missed. Her responses lacked appropriate specific action driven by data analysis.

67. Petitioner admitted that her essay had a number of misspellings, grammatical errors, and punctuation errors. She acknowledged that this is an area that the raters are supposed to consider. It is a necessary part of effective written communication. In this regard, by the undersigned's count, 29 of the 37 sentences in Petitioner's essay suffer from one or more errors of grammar, syntax, punctuation, or misspellings. More than half of those sentences (at least 15 of 29) suffer from errors of grammar and syntax, such as pairing "neither" with "or" instead of "neither . . . nor," using non-parallel structure, using plural subjects with singular verbs or singular subjects

with plural verbs, and using conditional language (such as "would do" and "would be") without a corresponding condition (e.g., that action would be appropriate, if the trend continues). In addition, the last sentence of the second paragraph on page one is not a complete sentence, ending in mid-word. Petitioner admitted that she ran out of time to complete the thought.

68. As to this consideration, Petitioner's essay appears to the undersigned to fall somewhere between the general rubric's description for a "three" ("The writer demonstrates some errors in the use of proper grammar and syntax that do not detract from the overall effect."), and the general rubric's description for a "two" ("The writer demonstrates serious and frequent errors in proper grammar and syntax."). Petitioner's essay admittedly did not meet the general rubric's description for a score of "four" ("The writer demonstrates satisfactory use of proper grammar and syntax."). This does not automatically doom Petitioner's essay to a score of three or less than three. However, it demonstrates the fallacy of Petitioner's approach of seizing on isolated parts of the prompt-specific rubric (supplemental rating criteria) to compare to her essay, without approaching the scoring process holistically. Even if Petitioner had persuasively critiqued parts of the essay scoring, as Respondent aptly notes, it is not simply a matter of checking off boxes and adding up points.

69. Petitioner failed to prove that the holistic scoring of her essay was incorrect, arbitrary, capricious, or devoid of logic and reason. She offered no evidence that a proper holistic evaluation of her essay would result in a higher total score than six; indeed, she offered no holistic evaluation of her essay at all. Petitioner's critique of various parts in isolation did not credibly or effectively prove that her score of six was too low; if anything, a non-expert's review of various parts in isolation could suggest that a score of six would be generous. But that is not the scoring approach called for here.

70. Petitioner failed to prove that there was anything unfair, discriminatory, or fraudulent about the process by which the written performance assessment exam was developed, administered, and scored.^{9/}

71. Petitioner pointed to the passage rate on the FELE written performance exam following the adoption of a separate passing score requirement. In 2015 and 2016, the passage rates for first-time test takers were 54 percent and 50 percent, respectively. The data is collected and reported for first-time test takers only, because that is considered the most reliable. Historically, performance on essay examinations goes down, not up, with multiple retakes.

72. The passage rates reflect a mix of both examinees prepared in an academic educational leadership program geared to

Florida standards, and those whose educational background does not include a Florida-focused program. Historically, examinees from academic programs aligned to Florida standards have greater success passing the FELE essay than those from out-of-state academic programs that are not aligned to Florida standards. Petitioner may have been at a disadvantage in this regard, as it does not appear that her master's program at Concordia University was aligned to Florida's educational leadership standards.

73. The passage rates, standing alone, do not prove that the written performance assessment is unfair, arbitrary, or capricious. It may be that the SBE's decision to increase scrutiny of the writing skills of FELE examinees results in fewer examinees achieving a passing score. Perhaps that is a good thing. Perhaps too many examinees achieved passing scores on the FELE in the past, despite weak written communication skills. In any event, the overall written performance assessment passage rates, standing alone, provide no support for Petitioner's challenge to the score given to her essay.

74. Petitioner failed to prove that the scoring verification process was unfair, arbitrary, capricious, or contrary to the procedures codified in the FELE rule. Petitioner pointed to evidence that essay scores are changed only on occasion, and that no scores were changed in 2016. Those facts, standing alone, do not support an inference that the score

verification process is unfair, arbitrary, or capricious. An equally reasonable or more reasonable inference is that the scores to be verified were appropriate.

CONCLUSIONS OF LAW

75. The Division of Administrative Hearings has jurisdiction over the parties and subject matter, pursuant to sections 120.569 and 120.57(1), Florida Statutes.

76. Petitioner has the burden of proving by a preponderance of the evidence that she is entitled to the relief she seeks. See Dep't of Transp. v. J.W.C. Co., 396 So. 2d 778 (Fla. 1st DCA 1981); § 120.57(1)(j), Fla. Stat.

77. As the one who has failed the essay component of a certification exam, Petitioner shoulders a heavy burden to prove that the subjective evaluation of her exam by Pearson raters, who are experts in the field, is arbitrary and capricious, devoid of logic or reason. Harac v. Dep't of Prof'l Reg., 484 So. 2d 1333, 1338 (Fla. 3d DCA 1986); State ex rel. Glaser v. Pepper, 155 So. 2d 383, 384 (Fla. 1st DCA 1963); State ex rel. Topp v. Bd. of Elec. Examiners, 101 So. 2d 583, 586 (Fla. 1st DCA 1958).

78. Harac was a rare successful challenge, based on unique circumstances established in the administrative hearing, to a failing grade received on the design portion of the exam for an architect's license. In particular, it was shown in the hearing that one of three expert graders did not follow the holistic

scoring method described in the design test handbook, and instead, gave a score of one, which all parties agreed was invalid. As the court noted, a score of one would only have been proper if the design solution was incomplete, which everyone agreed was not the case. Therefore, the invalid grade had to be thrown out. Two expert witnesses testified in the administrative hearing regarding their evaluations of the design and the grades they would assign. One expert used the holistic method and followed the original grading procedures as closely as possible without reconvening the original graders; this expert assigned a passing grade. The other expert did not use the holistic method or approved procedures in evaluating the examinee's design, but offered his opinion that the design should earn a failing grade. The grade assigned by the expert who used the holistic method and followed the approved procedures was accepted as substituting for the admittedly invalid grade, and licensure was approved.

79. In marked contrast to Harac, there was no proof in this case that either of the two raters' scores of three points was invalid, contrary to Pearson's scoring procedures, or improper in any way. Without such a showing (e.g., proof that one rater assigned a "zero," which is not a valid option), arguably, it would be inappropriate to reach the second level of Harac where, under the unique circumstance of an admittedly invalid grade, expert testimony was accepted to regrade the design test by

following the holistic grading method and approved procedures, to substitute for the invalid grade. See, e.g., The Florida Bar re Williams, 718 So. 2d 773, 778-779 (Fla. 1998) (in a certification examinee's challenge to the scores given to two essay answers, the Court refused the invitation to regrade the essays and award a higher score, "absent clear and convincing allegations establishing fraud, imposition, discrimination, manifest unfairness, or arbitrary or capricious conduct."). In this de novo hearing, Petitioner was given the opportunity to try to prove fraud, imposition, discrimination, manifest unfairness, or arbitrary and capricious conduct. Petitioner failed to meet her burden of proof in this regard.

80. If it were appropriate to reach the second level of Harac, Petitioner's proof would fall well short of the necessary showing to sustain her score challenge. Unlike the examinee in Harac, Petitioner failed to offer expert testimony by an expert in holistic scoring or by an expert in educational leadership who could offer an expert opinion after replicating as closely as possible the holistic scoring method used by Pearson to score Petitioner's exam. Petitioner's non-expert, self-serving testimony was far off the mark. As found above, she did not undertake an overall evaluation using the holistic scoring method, and her non-holistic comparison of isolated essay parts with parts of the prompt-specific rubric was wholly unpersuasive.

81. To the extent Petitioner contends that her challenge should succeed solely because essays are scored by humans, which makes the process subjective, that contention is rejected. The fact that subjectivity plays some role in the scoring process is not, standing alone, a basis upon which to overturn the results. That is particularly true where, as here, the unrebutted evidence showed that the scoring process in place is not only designed to minimize subjectivity, but that it actually functions that way. Instead, as shown by Harac and cases cited therein, to prevail, Petitioner was required to also prove that those who subjectively evaluated her examination acted arbitrarily or without reason or logic in giving her a failing score. Petitioner failed to meet her burden of proof in this regard.

82. Petitioner's criticism of the SBE's policy decision to toughen its certification standards by requiring examinees to achieve passing scores on the FELE written performance assessment does not provide grounds to invalidate Petitioner's failing essay score. The policy choice, codified in the FELE rule, was the SBE's prerogative and is not a matter subject to debate in this proceeding. Moreover, the heightened focus on effective writing skills is appropriate to align the FELE certification exam with SBE-adopted student standards, which have increased the focus on, and raised the expectations for, student achievement in writing. See § 1012.56(9)(f), Fla. Stat. Petitioner chose to take the

FELE exam after the FELE rule's tougher requirements, including the passing score required for the essay section, were in place.

83. Similarly, evidence of overall passage rates on the FELE written performance assessment following the SBE's policy change is inadequate to prove grounds for invalidating Petitioner's essay score, as found above.

84. As to Petitioner's complaints about the score verification process, Petitioner failed to prove that DOE did not follow the requirements and procedures in statute and rule, as found above.

85. DOE is required by statute to provide procedures for a certification applicant who failed an exam to review the exam question(s) and incorrectly answered response(s). The examinee is required to "bear[] the actual cost for the department to provide an examination review pursuant to this subsection." § 1012.56(9) (d), Fla. Stat. Petitioner was allowed to review the essay prompt and her response that got a failing score, after she paid the required fee. That is all the statute requires.

86. The procedures to review failed exam questions and responses are set forth in the FELE rule providing for a score verification process. Fla. Admin. Code R. 6A-4.00821(10) (b). These procedures were followed.

87. Petitioner's complaint about the score verification process was that the statement explaining why her essay score was

erroneous, which she says she prepared and submitted at her score verification session, was not provided to the chief reviewer. However, the score verification rule provides that the "statement of specific scoring errors" is filed with DOE, and then DOE proceeds to "review test items, verify examination keys, and consult with field-specific subject matter experts as needed." Fla. Admin. Code R. 6A-4.00821(10)(b)4. and 5. The rule does not require that the examinee's statement be provided to the subject-matter expert with whom DOE consults. DOE complied with the rule: DOE consulted with a chief reviewer who was shown to be a qualified field-specific subject matter expert, certified and actively serving as a principal at a Florida school. The chief reviewer was both trained and experienced in conducting score verification reviews using the holistic scoring method.

88. The final step in the score verification procedures is for DOE to "notify the individual [examinee] "of the action on the statement of scoring errors[.]" Fla. Admin. Code R. 6A-4.00821(10)(b)6. DOE did so here: DOE informed Petitioner of its determination that her essay was scored correctly. In this instance, Petitioner's statement (which was not offered in evidence) could only claim a single scoring error: she received a failing score of six for her essay response, but contended that she should have received a passing score of seven. DOE properly notified Petitioner of the action on that claim.

89. Petitioner seems to argue that DOE was required to prepare a detailed substantive response, addressing the merits of any and all reasons that may have been given in Petitioner's statement for her claim that her score was erroneous. Without the actual statement in evidence, Petitioner cannot prove that DOE's response would be inadequate, even if a point-by-point rebuttal were required. But the FELE rule only requires notice to the examinee of "the action," which was provided here. Nothing in the rule requires DOE to prepare a detailed response to explain the reasons for its action, or to provide a point-by-point rebuttal of any arguments contained in the examinee's statement that claims an essay score is erroneous. Indeed, it would be inappropriate, according to the established standards, to require a testing agency to provide that sort of detailed substantive response.

90. Petitioner raised in her petition, and alluded to in the Joint Pre-hearing Stipulation, several claims that are subordinate to her main contention that she should be given a passing score. Those subordinate arguments, such as Petitioner's claim for "back pay" to recoup the district pay raise she would qualify for if she becomes certified in educational leadership, and Petitioner's demand for refunds of examination and score verification fees, are rejected because Petitioner did not prove the main contention that her failing score should be changed to a

passing score. Nonetheless, even if Petitioner had prevailed, those subordinate claims, unsupported by any legal authority, would have to be denied.

91. Petitioner's entitlement to a pay raise for her teaching job is a matter between Petitioner and her employer. DOE is not a party to, and has no control over, Petitioner's salary arrangements with her employer. DOE's responsibility in this proceeding is to address Petitioner's challenge to the failing score she received on the essay exam she took in September 2016, and that is the sole issue in this proceeding. DOE's statutory sphere is the FELE, as part of the requirements for certification in educational leadership. DOE has no authority to reward successful FELE examinees with additional compensation along with their certificates.

92. For the same reason, even if it were determined that Petitioner's challenge to her essay score should be sustained, no authority has been identified that would support a remedy that includes a refund of fees. Those fees are set by rule to cover the costs of the processes that have already occurred, and they are not refundable. Fla. Admin. Code R. 6A-4.00821(4) and (10).

RECOMMENDATION

Based on the foregoing Findings of Fact and Conclusions of Law, it is RECOMMENDED that a final order be entered rejecting Petitioner's challenge to the failing score she received on the written performance assessment section of the Florida Educational Leadership Exam taken in September 2016, and dismissing the petition in this proceeding.

DONE AND ENTERED this 13th day of October, 2017, in Tallahassee, Leon County, Florida.



ELIZABETH W. MCARTHUR
Administrative Law Judge
Division of Administrative Hearings
The DeSoto Building
1230 Apalachee Parkway
Tallahassee, Florida 32399-3060
(850) 488-9675
Fax Filing (850) 921-6847
www.doah.state.fl.us

Filed with the Clerk of the
Division of Administrative Hearings
this 13th day of October, 2017.

ENDNOTES

^{1/} References herein to Florida Statutes are to the 2017 codification unless otherwise provided. Any amendments to the applicable substantive and procedural statutes in effect at the time Petitioner took her exam and at the time the hearing was held appear inconsequential; the relevant law addressed in this proceeding was not changed.

^{2/} With regard to Petitioner's attorney's "unopposed" motion to withdraw, the filing represented that Respondent did not object,

but made no representation regarding whether Petitioner objected to the withdrawal of her counsel. The grounds alleged for the request for leave to withdraw were general and conclusory: "Significant and irreconcilable differences have arisen between the undersigned and the Petitioner which render the undersigned unable to continue representing the Petitioner in this matter." As to the emergency motion for continuance, the asserted basis was that if counsel were permitted to withdraw just before the scheduled hearing, Petitioner would need time to prepare to represent herself and/or retain other counsel. The motion represented that Respondent did not agree to a continuance.

Since the final hearing was scheduled to begin in Orlando on Tuesday morning, June 13, 2017, with parties, lawyers, witnesses, and the undersigned all traveling on Monday, June 12, 2017, the motion was, as a practical matter, filed on the eve of the hearing. Accordingly, the undersigned scheduled a telephonic hearing on the motion beginning at 5:00 p.m. on Friday, June 9, 2017. Counsel for both parties and Petitioner participated.

During the telephonic hearing, inquiry was made regarding the basis for counsel's request for leave to withdraw, so as to determine whether the grounds required counsel's withdrawal (see Fla. Bar Reg. R. 4-1.16(a)), or gave rise to permissive withdrawal (see Fla. Bar Reg. R. 4-1.16(b)). Counsel stated that the circumstances resembled the description in rule 4-1.16(b)(2) ("the client insists upon taking action that the lawyer considers repugnant, imprudent, or with which the lawyer has a fundamental disagreement"). Counsel explained that the issue related to information learned earlier in the week regarding the opinions that Petitioner's expert witness would be able to offer at hearing. The dispute apparently also had a dimension regarding funding of fees for the hearing, another basis for permissive withdrawal, in rule 4-1.16(b)(4).

Petitioner said she opposed the motion to withdraw, but agreed with counsel's description of the expert witness issue. She said she wanted her counsel to continue to represent her, but that a continuance would be helpful even if her counsel did not withdraw. She said more time was needed, either for her expert witness to explore other matters or to replace the expert.

The undersigned rejected the new asserted basis for an emergency continuance. As of June 9, 2017, this case was ready for hearing. The parties had filed their Joint Pre-hearing Stipulation on June 5, 2017, with final exhibit and witness lists (including Petitioner's expert). At no time before the eve of

hearing did Petitioner indicate that the case was not on track for hearing as scheduled. Pursuant to Florida Administrative Code Rule 28-106.210, motions for continuance filed less than five days before the hearing require a showing of an emergency. No emergency was demonstrated, so Petitioner's "emergency" motion for continuance was denied. In making this determination, the undersigned considered the prejudice to Respondent, who prepared and made arrangements for its witnesses to travel to Orlando from around the state and from out of state.

The lack of an emergency justifying a continuance was not changed by Petitioner's counsel's motion to withdraw. The dispute between counsel and Petitioner may have supported permissive withdrawal under the Florida Bar's rules and the granting of a motion to withdraw, if presented at an earlier time. However, on the eve of hearing, an alternative option under the Bar rules was to order counsel to continue representing Petitioner despite cause for terminating the representation. See Fla. Bar Reg. R. 4-1.16(c). Indeed, under similar circumstances, it has been held to be an abuse of discretion to grant a motion to withdraw. See Garden v. Garden, 834 So. 2d 190 (Fla. 2d DCA 2002). In Garden, the appellate court affirmed the denial of a husband's motion for continuance of a once-continued trial in a divorce case. The motion for continuance was filed by the husband's lawyer at the beginning of the week in which the trial had been rescheduled, asserting as grounds that the client's out-of-state business travel had interfered with trial preparation. At the outset of the trial, the husband's lawyer moved to withdraw due to an inability to communicate with his client (who did not appear). The trial court's order granting the motion to withdraw was reversed on appeal. The court's discussion of these awkward last-minute continuance-withdrawal pairings is instructive:

When these situations occur, as they frequently do, the trial court is often faced with what it regards as a Hobson's choice. It may permit counsel to withdraw and, due to a party's lack of representation, continue the case. This choice detrimentally affects the opposing party, who is prepared for trial and has incurred costs and attorney's fees in doing so. The court's other option is to grant the withdrawal motion and to deny the continuance request, thereby facing the potential that the ruling may cause the moving party to suffer a denial of due process.

There is, we believe, a third choice. Rule 4-1.16(b) of the Rules Regulating the Florida Bar, governs an attorney's ethical termination of representation. In those situations where withdrawal is optional, as here, the rule 4-1.16(b) provides that withdrawal should be "accomplished without material adverse effect on the interest of the client." Withdrawal at the moment a trial is to commence can seldom be accomplished without material adverse effect on the client. Thus, rule 4-1.16(c) authorizes the court to require continued representation of the client by counsel, even in those instances where good cause to withdraw exists. . . .

We conclude that the trial court abused its discretion by granting the motion to withdraw and failing to require counsel to continue representing Mr. Garden. We acknowledge this may be burdensome to the moving attorney, but the risk should be borne by counsel who is most familiar with the client and events and not by the court or by the opposing party[.]

Id. at 192-193. Likewise, in this case, the third choice described by the court in Garden was the most appropriate choice to make among the less-than-ideal options. Accordingly, the motion to withdraw was also denied.

Counsel for Petitioner appeared at the final hearing to present Petitioner's case. Petitioner rested without calling her expert witness named in the Joint Pre-hearing Stipulation. Since the telephonic motion hearing was not recorded, counsel for Petitioner was asked to explain why the expert witness was not being called to testify. He stated: "We engaged the services of an expert witness. We received a report from the expert witness that did not satisfy our needs for this hearing, and, therefore, have chosen not to call that expert witness." (Tr. 165-166).

^{3/} No evidence was offered regarding the Concordia University educational leadership program, such as whether the program was geared to Florida educational leadership standards. DOE Bureau Chief Phil Canto did not believe Concordia University provided such a program, although he could not say for certain and noted that he is aware of at least one out-of-state graduate school (not Concordia) that does incorporate Florida standards. Absent

evidence showing that this out-of-state graduate program was geared to Florida standards, the most reasonable inference is that it was not.

^{4/} Petitioner was asked why it was important to her to pass the FELE. She said, "Essentially, once I pass those tests, those-- my master's degree credentials are added to my teaching license. So, after those credentials are added to my teaching license, I am entitled to compensation." After a momentary pause, she added, "And, also, I believe, that having the credential added to your teaching license and passing the FELE helps you progress on to lead programs and administrative opportunities, if I choose to go that route." (Tr. 32-33).

^{5/} Petitioner's proposed finding 13 and conclusion 25 erroneously contend that the two chief reviewers who testified at hearing did not "possess the requisite experience and training to serve as" chief reviewers. As to experience, while both chief reviewers testified that they were not employed as raters before becoming chief reviewers, they both had experience scoring FELE essays, gained in the intensive rater training in Hadley. Petitioner failed to identify any requirement in statute or rule, or even in contract, for chief reviewers (as contrasted with chief raters, the rater supervisors and trainers in Hadley), to have been employed as raters before becoming chief reviewers. Instead, chief reviewers are qualified by the FELE rule to serve as consultants to DOE, as Florida-based subject matter experts in the field, in the score verification process. Petitioner apparently confused the chief rater requirements with chief reviewer requirements. As to training, Petitioner represented: "Both [chief reviewers] testified that they received no training beyond the basic rater training provided by Pearson, prior to become a Chief Reviewer." (Pet. PRO at 5). That is contrary to the actual testimony. Dr. Small, who was the chief reviewer assigned to conduct the score verification review (as shown by the date of his comments), plainly testified to having been given training to conduct reviews of contested scores, in addition to the rating training in Hadley. (Tr. 117). Dr. Pelletier, the chief reviewer enlisted for a second review after Petitioner requested an administrative hearing (as shown by the date of her comments), generally described the two-day rater training in Hadley (Tr. 283), but she also testified that she was trained to be a chief reviewer while in Hadley. (Tr. 296). She was not asked about what kind of chief reviewer training she received; she certainly never testified that she "received no training beyond the basic rater training," as Petitioner represented.

Of course, by the time both chief reviewers scored Petitioner's essay, they had conducted numerous score verification reviews, having each served as a chief reviewer for two years.

^{6/} In this case, after Petitioner contested DOE's determination that her essay was scored correctly, DOE asked the two raters to prepare written justifications for their scores. In addition, although DOE's practice in the score verification process is to have a chief reviewer prepare written comments to explain why the original score should stand or why it should be changed, in this case, after Petitioner requested an administrative hearing, DOE had a second chief reviewer conduct an additional review and prepare written comments. Both chief reviewers testified at hearing, but did not specifically address their written comments; neither original rater testified. All of the written comments are in evidence under seal. (Jt. Exhs. 5 and 6). The written comments were utilized at hearing only by Petitioner in her critique of her essay score, comparing isolated parts of the comments with isolated parts of her essay, but not doing so effectively or persuasively.

^{7/} Petitioner alleged in her administrative hearing request (considered a petition) that the scoring process was invalid because of the probability of human error. She also contended that DOE's generic response to her score verification request should have been personalized to explain why she got a failing score and respond to points raised in her prior appeal (it is unknown what prior appeal she was referring to, as there was no evidence of a prior appeal). Finally, she asserted that she was entitled to have her master's degree added to her teacher's certificate, and receive back pay for the 2015-2016 and 2016-2017 school years, as well as refunds of the last two essay retake registration fees, plus the score verification fee. As stated in the Joint Pre-hearing Stipulation, Petitioner's position was that the FELE essay is graded "through human assessment," based on rubrics provided only to the raters, and that the raters are required to grade under extremely short time schedules, which leads to human error. In addition, the re-evaluation of her score was done "through human grading systems which leave room for error." Petitioner asserted that her score was made in human error, causing Petitioner lost wages and emotional distress.

^{8/} The disconnect between Petitioner's proposed action and the data in the prompt was pointed out by one rater who prepared an after-the-fact justification for the score, in which it was noted that Petitioner's proposed action infers that the need was limited to two groups. (Jt. Exh. 5-B). At hearing, Petitioner

criticized this rater's point because in her view, it was inappropriate for the rater to be inferring anything. Inconsistently, Petitioner argued that inferences should be made when in her favor to cover something omitted from her essay. In fact, the rater did not infer anything. Petitioner's proposed action described in the second sentence of the first full paragraph on page two of her essay (Jt. Exh. 3 at 2) was expressly directed to two groups in two areas, which would only be reasonable if the data showed need in both areas for only those two groups, and not for other groups. However, the data did not show that. The point is that Petitioner's action plan was not supported by the data, which Petitioner did not sufficiently identify, summarize, or analyze.

^{9/} Petitioner took the position in the Joint Pre-hearing Stipulation that Pearson raters are "required to grade under extremely short time schedules, which leads to human error." Petitioner failed to present any evidence to support that position, and proposed no such finding of fact, acknowledging the absence of supporting evidence. In fact, the evidence was to the contrary. Raters are not given a time limit within which they are required to score essays, nor are they given any kind of quota. Rater scoring time is monitored, along with all other aspects of scoring work, by the chief rater. However, the purpose is to look for patterns of raters scoring either too quickly or too slowly, as either may indicate a problem needing to be addressed with a rater. Dr. Grogan testified credibly that he has not received any rater complaints regarding time pressure, and since he is in charge of the Hadley office where the FELE essay scoring is done, he would know of any such complaints.

COPIES FURNISHED:

Julie McCue
1474 Harbour Side Drive
Weston, Florida 33326

Bonnie Ann Wilmot, Esquire
Department of Education
Suite 1244
325 West Gaines Street
Tallahassee, Florida 32399
(eServed)

Robert F. McKee, Esquire
Robert F. McKee, P.A.
Suite 301
1718 East 7th Avenue
Tampa, Florida 33605
(eServed)

Darby G. Shaw, Esquire
Department of Education
Suite 1244
325 West Gaines Street
Tallahassee, Florida 32399
(eServed)

Matthew Mears, General Counsel
Department of Education
Turlington Building, Suite 1244
325 West Gaines Street
Tallahassee, Florida 32399-0400
(eServed)

Pam Stewart, Commissioner of Education
Department of Education
Turlington Building, Suite 1514
325 West Gaines Street
Tallahassee, Florida 32399-0400
(eServed)

Chris Emerson, Agency Clerk
Department of Education
Turlington Building, Suite 1520
325 West Gaines Street
Tallahassee, Florida 32399-0400
(eServed)

NOTICE OF RIGHT TO SUBMIT EXCEPTIONS

All parties have the right to submit written exceptions within 15 days from the date of this Recommended Order. Any exceptions to this Recommended Order should be filed with the agency that will issue the Final Order in this case.