

STATE OF FLORIDA
DIVISION OF ADMINISTRATIVE HEARINGS

DARYL BRYANT,

Petitioner,

vs.

Case No. 17-0424

PAM STEWART, AS COMMISSIONER OF
EDUCATION,

Respondent.

_____ /

RECOMMENDED ORDER

On June 14, 2017, Elizabeth W. McArthur, Administrative Law Judge, Division of Administrative Hearings (DOAH), conducted the hearing in this cause in Orlando, Florida.

APPEARANCES

For Petitioner: Jennifer Diane Rose, Esquire
Post Office 924
Melbourne, Florida 32902

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 he received on the essay portion of the Florida Teacher Certification Examination's (FTCE) General Knowledge (GK) test should be sustained.

PRELIMINARY STATEMENT

Petitioner, Daryl Bryant (Petitioner), took the GK essay test on June 25, 2016. The score report he subsequently received showed that he did not earn a passing score, having received a score of seven (on a scale of two to 12), when a score of eight was required to pass. Petitioner utilized the score verification procedures in statute and rule, and by letter dated September 26, 2016, the Department of Education (DOE) informed Petitioner of its determination that his essay had been scored correctly. Petitioner was informed of his right to an administrative hearing pursuant to sections 120.569 and 120.57, Florida Statutes (2017),^{1/} to dispute the decision.

Petitioner initially requested an informal hearing not involving disputed facts, but after discussions with counsel for DOE, Petitioner notified counsel by email on December 15, 2016, that he wanted a formal (i.e., disputed-fact) hearing instead of an informal hearing. Treating the email as a petition, DOE issued an Order dismissing the petition without prejudice, and allowed Petitioner to amend his petition to conform to the requirements for petitions in Florida Administrative Code Rule 28-106.201(2). Petitioner timely filed an amended request (petition) for a disputed-fact hearing on January 4, 2017, and the matter was transmitted to DOAH for assignment of an administrative law judge to conduct the requested hearing.

After being unable to reach Petitioner, DOE filed a response to the Initial Order that sought information related to scheduling the hearing. DOE 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. DOE suggested some dates, and requested a live hearing in Orlando. The hearing was scheduled for March 14, 2017, in Orlando, by Administrative Law Judge J.D. Parrish, and the similar case was scheduled for March 13, 2017. Shortly thereafter, DOE filed a Motion to Limit the Scope of Review in This Matter, followed by an amended motion. DOE learned that counsel would be representing Petitioner, although no notice of appearance was filed, and DOE served its motion and amended motion on Petitioner's counsel.

After a motion for continuance was filed on February 28, 2017, in the related case by the petitioner, DOE filed a motion for continuance in this case, so as to be able to agree to the requested continuance in the related case, while trying to keep the hearing schedules coordinated. Petitioner did not file a response opposing the continuance. Judge Parrish granted DOE's motion for continuance and rescheduled the hearing for May 2, 2017, in Orlando. A continuance was also granted in what had

become the companion case for scheduling purposes (DOAH Case No. 17-0423), and that hearing was reset for May 1, 2017.

On March 20, 2017, Petitioner filed a response to DOE's pending motion to limit the scope of the hearing. On April 3, 2017, this case and the companion case were transferred to the undersigned. On April 5, 2017, an Order was issued denying DOE's motion to limit the scope of the hearing; that Order was amended (corrected) by Order issued on April 7, 2017. A similar DOE motion to limit the hearing scope in the companion case was also denied by Order issued on April 5, 2017.

Meanwhile, a second motion for continuance was filed in the companion case by the petitioner, 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 the same issues would apply to this case, a joint telephonic status conference was scheduled, 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 14, 2017, in Orlando (and the companion case was reset for June 13, 2017).

To resolve confidentiality issues raised by DOE in its motion to limit the hearing, the parties entered into a Confidentiality Agreement. Pursuant to that agreement, DOE filed an agreed Motion for Protective Order 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 June 7, 2017, a Joint Pre-hearing Stipulation was filed, in which the parties identified their witnesses and proposed exhibits, and agreed to a few facts. The agreed facts are incorporated below.

On June 9, 2017, Petitioner filed a motion to permit a witness who was in New York to testify by telephone. The motion was granted by Order issued on June 12, 2017.

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 his own behalf. Petitioner did not offer the testimony of the witness in New York for whom leave to testify by telephone had been granted. Petitioner did not offer any additional exhibits besides the Joint Exhibits.

Respondent presented the testimony of the following witnesses: Michael Grogan, Pearson director of performance assessment scoring services; Phil Canto, DOE bureau chief of post-secondary assessment; Betsy Griffey, a FTCE GK essay chief

reviewer; and Mary Jane Tappen, DOE vice chancellor for K-12 student achievement and student services. In addition, the parties stipulated to adopt by reference as testimony in this case the non-confidential testimony given by Dr. Grogan, Mr. Canto, and Ms. Tappen the previous day in DOAH Case No. 17-0423, since both Petitioner and counsel for Petitioner were in attendance for that testimony, and counsel for Petitioner was permitted to cross-examine the witnesses as to their previous day's testimony.^{2/} Respondent's Exhibits 1 through 7 (which are not confidential) were admitted.

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

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, it was agreed that the deadline to submit proposed recommended orders (PROs) would be an extended deadline of 30 days from the filing of the transcript. The one-volume Transcript of the hearing was filed on July 10, 2017. By subsequent motions to extend the PRO deadline by the petitioner in the companion case and by Petitioner in this case, the PRO filing deadline was extended until August 30, 2017.

Both parties timely filed their PROs by the extended deadline, and the PROs have been carefully considered in the preparation of this Recommended Order.

FINDINGS OF FACT

1. Petitioner has been employed as a teacher for the past three years. He had a temporary Florida teacher certificate, but at the time of the hearing, he said that he believes it was expired. Petitioner is seeking to qualify for a (non-temporary) Florida teacher certificate. Petitioner first must pass the essay part of the GK test to complete the GK requirements. He would then be qualified to proceed to address the remaining certification requirements. See § 1012.56(2)(g), (h), (i), and (7), Fla. Stat.

2. 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.

3. 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 examinations relevant to the particular certification.

4. 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 FTCE.

5. 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, school district personnel, and academicians from colleges and universities.

FTCE Development, Administration, and Scoring

6. DOE develops the FTCE, as well as the other educator certification exams, in-house. The FTCE is developed and periodically revised to align with SBE-promulgated standards for teachers. In addition, as required by statute, certification exams, including the FTCE, must be aligned to SBE-approved student standards.

7. Details about the FTCE, such as the competencies and skills to be tested, the exam organization, and passing score

requirements, are set forth in Florida Administrative Code Rule 6A-4.0021 (the FTCE rule). The FTCE rule has been amended periodically, but the current version includes a running history, setting forth FTCE details that applied during past time periods, as well as those currently in effect.

8. The FTCE is not actually a single examination. It consists of multiple separate examinations to meet the different requirements for teacher certification and the different options for specific subject areas. Descriptions of the areas to be tested by each FTCE component are set forth in a publication incorporated by reference in the FTCE rule. The version of this publication that was in effect when Petitioner took the exam at issue in this proceeding is identified in the FTCE rule as: "Competencies and Skills Required for Teacher Certification in Florida, Twenty-Second Edition."

9. As set forth in the FTCE rule, the GK exam consists of four subtests. Subtest one is the essay test; subtest two, three, and four are multiple choice tests covering English language skills, reading, and math, respectively.

10. Petitioner met the requirements for GK subtests two, three, and four, by virtue of having taken and passed the College Level Academic Skills Test (CLAST) in those areas prior to July 1, 2002.^{3/} Therefore, Petitioner only had to take and pass subtest one, the essay exam, to satisfy all GK requirements.

11. The competency and skills to be tested by the GK essay test, as promulgated by the SBE and codified by reference in the FTCE rule, are as follows:

Knowledge of formal college-level writing

- Determine the purpose of writing to task and audience.
- Provide a section that effectively introduces the topic.
- Formulate a relevant thesis or claim.
- Organize ideas and details effectively.
- Provide adequate, relevant support by citing ample textual evidence; response may also include anecdotal experience for added support.
- Use a variety of transitional devices effectively throughout and within a written text.
- Demonstrate proficient use of college-level, standard written English (e.g., varied word choice, syntax, language conventions, semantics).
- Provide a concluding statement or section that follows from, or supports, the argument or information presented.
- Use a variety of sentence patterns effectively.
- Maintain a consistent point of view.
- Apply the conventions of standard English (e.g., avoid inappropriate use of slang, jargon, clichés). (Competencies and Skills Required for Teacher Certification

in Florida, Twenty-Second Edition, page 2 of 247, incorporated by reference in the FTCE rule).

12. Prior to January 1, 2015, a score of at least six (using a scoring range from two points to 12 points) was required to pass the GK essay test.

13. Based on input from educators, academicians, and other subject matter experts, DOE recommended that the passing score for the GK essay test be raised from a score of six to a score of eight (using the same range of two points to 12 points). The SBE adopted the recommendation, which is codified in the FTCE rule: eight is the required passing score for GK essays as of January 1, 2015.

14. Without question, the higher passing score requirement makes it more difficult to pass the GK essay. The policy underlying this scoring change is to make the GK essay test more rigorous, in recognition of the critical importance of writing skills. By raising the standards for demonstrating mastery of the writing skills tested by the GK essay test, the GK essay test 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 FTCE rule.

15. Not surprisingly, since the passing score was raised for the GK essay, the overall passage rates have dropped. The passage rates were 96 percent in 2013 and 93 percent in 2014, when the passing score was lower. After the passing score was raised, the passage rates were 63 percent in 2015 and 69 percent in 2016. While Petitioner characterizes the 69 percent passage rate as "low" (Pet. PRO at 4, ¶ 13), that is an opinion that is unsupported by any testimony offered at hearing. Petitioner did not offer any expert witness to testify on his behalf. Instead, based on the testimony offered on this subject at the final hearing, the more reasonable inference to draw from the overall GK essay passage rates is that the passage rates were too high prior to 2015. The overall GK essay passage rate, standing alone, is not evidence that the GK essay is arbitrary, capricious, unfair, or invalid.

16. Pursuant to its contract with DOE as the test administration and test scoring agency, Pearson administers and scores GK essay exams. Pearson employs holistic scoring as the exclusive method for scoring essays, including GK essays (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.

17. 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, prospective raters are trained in the holistic scoring method used to score GK essays.

18. 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 GK essay-specific training material. They thoroughly review and discuss the rubric, the score scale (which is one point to six points), 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 GK essays.

19. Raters who conduct scoring of the GK essay must meet qualifications specified by DOE (including teacher certification and experience). Pearson proposes qualified individuals to DOE, and then DOE must approve proposed raters. Then the approved raters must undergo and successfully complete Pearson's training.

20. Each GK essay is scored independently by two qualified raters. 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. GK 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.

21. 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 rubric. 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 scoring consistency are the raters permitted to turn to the assigned GK essay for review and scoring.

22. Pearson also employs chief raters to supervise and monitor the raters while they are engaged in their scoring work. Chief raters must meet specified qualifications and be approved by DOE. Chief raters must be certified and experienced in the field of teaching, plus they must have prior experience working

as raters. Chief raters conduct the training sessions to train raters in the holistic scoring method in Hadley.

23. A chief rater supervises and monitors raters by being physically present in the same room with the raters while they are engaged in their scoring work. The chief rater monitors rater 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.

24. The scores of the two raters assigned to score a GK essay are added together for the total holistic score. Thus, the total score range for a GK 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.

25. The sum of the two raters' scores will be the score that the GK 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.

26. After GK 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. Eight points is a passing score, according to the FTCE rule.

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

28. Through the intensive training and the subsequent calibration and recalibration before each GK essay scoring assignment, Pearson has achieved consistency in rater scoring of GK essays that meets industry standards for holistic scoring. Consistency in this context means that the scores assigned to a GK essay by a pair of raters are either identical or adjacent (within one point), and when adjacent, are balanced (i.e., each rater is as often the higher scorer as he or she is the lower scorer). DOE makes sure that Pearson maintains rater scoring consistency in accordance with industry standards, by monitoring monthly performance reports provided by Pearson.

Examinee Perspective: Preparation for the GK Essay

29. DOE provides detailed information and aids on its website regarding all four subtests of the GK exam, including the GK essay, for potential examinees. This includes a 39-page test information guide for the FTCE GK test.

30. The test information guide sets forth the complete SBE-adopted competencies and skills to be tested by each of the four GK subtests, including those specific to the essay test quoted in Finding of Fact 11.

31. The test information guide explains the GK essay and scoring process, as follows:

For your essay, you will choose between two topics. The 50 minutes allotted for this section of the exam includes time to prepare, write, and edit your essay.

Your work 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, and the degree to which you support your position will be very important in the scoring.

Your essay will be scored on both the substance and the composition skills demonstrated, including the following elements: ideas, organization, style (diction and sentence structure), and mechanics (capitalization, punctuation, spelling, and usage).

The raters will use the categories on page 14 when evaluating your essay. The score you receive for your essay will be the combined total of the two raters' scores. (R. Exh. 2 at 12 of 39).

32. At the referenced page 14, the test information guide sets forth in full the scoring rubric used by raters to evaluate GK essays. The rubric is simply a comparative description of the extent to which an essay demonstrates the competency and skills to be tested, on a scoring scale of one to six points. The rubric descriptions differentiate between the various skills to be tested in a way that identifies, as to each skill or group of skills, which essay is best, better, good, not-so-good, worse,

and worst. But the evaluation of each skill is not separately scored; instead, the essay response is evaluated as a whole, with the various strengths and weaknesses weighed and balanced.

33. Finally, the test information guide provides a sample essay test, with representative essay prompts in the same format that the examinee will see on the exam: two topics are set forth, with instructions that the examinee is to select one of the two topics.

34. The information DOE makes publicly available is appropriate and sufficient to explain the GK 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.

Score Verification

35. An examinee who fails the GK essay test (or any other FTCE test or subtest) may request score verification to verify that the failed exam was scored correctly. The examinee has the right, by statute and rule, to review the test question(s) and response(s) that received a failing score. The score verification procedures, providing this review opportunity, are set forth in the FTCE rule.

36. 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 FTCE rule, when a score verification request is directed to the scoring of a GK essay, DOE always consults with a field-specific subject matter expert known as a "chief reviewer."

37. Chief reviewers are another category of experts (in addition to raters and chief raters) retained by Pearson, pursuant to qualifications identified by DOE, and 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. Unlike raters and chief raters, chief reviewers do not work at Pearson in Hadley; they are Florida experts in the field, with certification and experience teaching in Florida schools.

38. Chief reviewers only become involved with GK essays when an examinee who failed the GK essay invokes the score verification process. A chief reviewer is assigned to evaluate whether that essay was scored correctly. As with the initial scoring, a chief reviewer is not given any information about the raters or about the examinee; the essay is assigned a blind, anonymous number. The chief reviewer conducts the 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. After 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.

39. In the 14-month period from January 2016 through February 2017, two failing GK essay scores were changed by DOE to passing scores as a result of the score verification process. As with the overall passage rates, Petitioner characterizes this reversal rate as low, but no evidence is offered to prove that characterization. It is as reasonable or more reasonable to infer from the fact that GK essay scores are only rarely reversed through score verification that the scoring process works well.

Petitioner's GK Essay Attempts

40. Petitioner took the GK essay test for the first time in July 2015. He received a failing score of four, with two points assigned by each of the two raters. Petitioner admits that he did little to nothing to prepare for the GK essay the first time. When taking the essay test, he ran out of time and recalls that

he left the essay incomplete. The time pressure "had a huge deal with me not being able to provide enough specifics for it to make any sense at all where I was going with the essay." (Tr. 75). Petitioner thought the passing score was six at the time, but his recollection is incorrect. The higher passing score of eight has been in place since January 2015, and has been the passing score for each of Petitioner's GK essay attempts.

41. FTCE 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 FTCE rule for each retake of a failed subtest and/or section.

42. Petitioner retook the GK essay test in February 2016. In preparation for this second attempt, Petitioner did not seek tutoring or spend much time training. As he explained, "I'm under the impression that I can write an essay." (Tr. 21). Instead, he focused mostly on preparing for the timed aspect of the exam, making sure that he started when the clock started. Although his score improved from four to six, it was still a failing score.

43. Petitioner did not invoke the score verification process to question the failing scores he received on his first

two GK essays. Those two failing scores stand as final, as he did not challenge them.

44. Petitioner took the GK essay test for the third time on June 25, 2016. This time, he prepared to some extent. In the month before the exam, Petitioner sought help from someone he described as a writing coach. The writing coach did not evaluate Petitioner's writing so as to identify weaknesses; instead, she asked him what he thought his weaknesses were, and he responded that he did not know what his weakness is besides not being able to formulate his plan and map out his essay faster. As a result, she coached him on some mapping techniques, and on how to structurally organize an essay--with an introduction, followed by three points in paragraphs begun with transitional phrases, and a conclusion. Petitioner practiced a little with his writing coach, by email: she would send a prompt and he would write an essay, which he timed, and then send it back to her. They did this "a few times." (Tr. 24). There is no evidence of record regarding the writing coach, other than that her name is Ms. Martin. She may have been Petitioner's proposed witness who was allowed to appear from New York by telephone, but who was not called to testify.

45. One of the things Petitioner learned from Ms. Martin was that in his introduction, he should "speak vaguely about" what will be covered. When asked if Ms. Martin actually said to

be "vague" in the beginning, Petitioner said, "She may not have used the word vague, but that is the meaning that I got from what she said." (Tr. 70).

46. In preparation for his third attempt at the GK essay test, Petitioner also sought help from Jordan Gibbs, who was described as an educator who taught language arts for over 20 years. Petitioner testified that Mr. Gibbs is "our academy leader there[.]" (Tr. 24). However, Petitioner did not elaborate; it is unknown which academy is led by Mr. Gibbs, or where "there" is. Like Ms. Martin, Mr. Gibbs also addressed mapping techniques with Petitioner. Petitioner never sent any essay drafts to Mr. Gibbs for his review.

47. Petitioner also reviewed GK essay preparation material on the DOE website. He reviewed sample prompts, but did not practice writing complete essays. He just looked at the sample prompts for purposes of mapping and planning an essay. Petitioner said that he found the preparation material useful to an extent, but did not think the sample prompts reflected the type of GK essay prompts in use when he took the test. A comparison of the sample GK essay prompts in the test information guide (R. Exh. 2 at 17 of 39) with the actual GK essay prompt Petitioner chose for his essay topic (Jt. Exh. 1 at 3 of 4) suggests otherwise. Although DOE obviously does not make available as samples the actual essay prompts actively being used

in GK examinations, the sample prompts appear to be similar to Petitioner's actual prompt in style, substance, and tone. It would be unreasonable for examinees to expect more from a testing agency than what DOE makes available.

48. Petitioner's score improved slightly in his third attempt at the GK essay test, but it was still a failing score of seven. One rater assigned the essay a score of three, while another rater scored the essay a four.

49. Each of the three times Petitioner took the GK essay test, the two raters assigned scores that were consistent, in that they were either identical or adjacent (within one point of each other). Accordingly, a chief rater was never assigned for discrepancy resolution, as there were no discrepancies.

50. After receiving notification of his third failing score, this time Petitioner invoked the score verification process. Petitioner completed a statement explaining why he believes his score was erroneous, which is in evidence as part of the confidential testing material. (Jt. Exh. 1 at 2 of 4). The statement set forth why he believes the essay demonstrated good organization, used transitional phrases, and addressed the topic. He acknowledged one misspelling, and acknowledged that his conclusion ended in mid-sentence, as he ran out of time. He added three words to complete the last sentence, and suggested that the ending should have been inferred from what he did say.

51. DOE conducted its review, and the score was verified through a process consistent with DOE's practice of consulting a chief reviewer who was qualified as a subject matter expert in the field of teaching in Florida and approved by DOE.

52. The chief reviewer who undertook to verify Petitioner's essay score conducted an independent evaluation of Petitioner's essay following the same holistic method. Then the chief reviewer considered the scores separately assigned by the two raters who scored Petitioner's essay. She concluded that the assigned scores of three/four should stand. The chief reviewer provided a summary rationale for her determination, offering her view that the essay borders on a three/three due to weak development.^{4/}

53. The chief reviewer's summary was provided to DOE for consideration. By letter dated September 27, 2016, 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 notified of his right to an administrative hearing pursuant to sections 120.569 and 120.57 to dispute the decision. Petitioner availed himself of that opportunity, and was given the chance in a de novo evidentiary hearing to present evidence to support his challenge to his exam score.

54. At the hearing, Petitioner offered only his own testimony as support for his challenge to the scoring of his essay. Petitioner was not shown to be, tendered as, or qualified as an expert in either formal college-level English writing or scoring of essays. His attempt to compare isolated parts of the rubric to isolated parts of his essay is contrary to the holistic scoring approach used to score the GK essay. Petitioner offered no comprehensive, holistic evaluation of his essay as a whole, nor was he shown to be qualified to do so.

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

56. Petitioner began his critique by reading the first three sentences--the introductory paragraph--of his essay. He said that each sentence had one topic, and that each of the subsequent three paragraphs in the body addresses one of those three topics. The problem with Petitioner's explanation for the substantive organization of his essay is that the essay prompt identifies a single topic, not three topics. Petitioner failed

to respond to the prompt's single topic by introducing that topic as the essay's theme, and developing that single theme in the body of that essay. Similarly, the concluding paragraph offers scattered thoughts, somewhat related to the three topics discussed in the essay. The essay's weakness in development was a prominent point in the scoring rationale summaries written by the raters and chief reviewers.

57. Petitioner specifically addressed only one aspect of the rubric considerations, addressing the extent to which an essay has errors in sentence structure, usage, and mechanics. As to this consideration, Petitioner stated that there were three spelling errors in his essay (up from the one error he identified in his score verification statement). He was critical of one rater's comments for referring to grammatical errors, because Petitioner does not believe there were any grammatical errors in his essay. Petitioner's assessment of his essay reflects his bias, because it fails under any objective analysis.

58. In fact, Petitioner's essay (Jt. Exh. 3) has both spelling errors and grammatical errors. In addition, the essay uses poor sentence structure in several instances, as well as poor word choices that interfere with an understanding of what Petitioner means. An example of a sentence with a grammatical error is the fifth sentence in paragraph 4. At the very least, the word "having" is required after the comma. With that

addition, the sentence would only be awkward, instead of grammatically incorrect.

59. An example of a poorly written sentence is the second sentence of the second paragraph. This sentence combines a misspelling, a misused word, and syntax that is awkward, at best.

60. Petitioner must also acknowledge that the last sentence of his essay is another example of poor sentence structure, since it is an incomplete sentence without punctuation. It would be inappropriate for raters reviewing essays to fill in the gaps left by writers, whether those gaps were because of running out of time or otherwise. What Petitioner meant to write to complete the sentence is not something that can be added after-the-fact to cure the defect on the face of the essay.

61. By the undersigned's count, there are five misspellings in the essay, unless one counts "in to," which should be "into," as an error of grammar or syntax. The other misspellings were: easire (easier); savy (savvy); yeild (yield); and evironment (environment). In addition, Petitioner made several punctuation errors, failing to hyphenate two compound adjectives preceding nouns and presenting a single idea: cutting-edge technology; tech-savvy students. Petitioner also improperly omitted a hyphen in "self discipline."

62. Petitioner acknowledged some repetitive use of a particular word, but thought he only used that word twice. In

fact, he used the word in both sentences one and two of the second paragraph, and then again in paragraph four. Only the first usage is arguably correct (but in an awkwardly written sentence). While used once, the word is an interesting one, Petitioner's overuse and misuse of this word suggests a mechanical, as opposed to thoughtful, approach of injecting interesting words into the essay.

63. Petitioner's essay demonstrated good superficial structure, with an introductory paragraph, three paragraphs in the body that begin with good transitional words, and a concluding paragraph. The organizational structure may have earned Petitioner a score of four, as stated in that rater's comments, but that same rater also repeated the comments of others that where the essay is weakest is in development.

64. Petitioner offered his view that the only reason his essay received a failing score was because the raters considered it to be too short in length. While Petitioner is correct in noting that length is not a criterion, he mischaracterized the comments on this subject, by ignoring the criticisms of his essay that were made when the length of the essay was noted. The comments only mention the length of Petitioner's essay as it correlates to other considerations, such as the weakness in development, the lack of specifics or examples, or the impact of a "number of misspellings, . . . usage issues, . . . and

punctuation errors," which accumulated to a notable level "given the shortness of the response." (Jt. Exh. 5-A). Petitioner failed to prove his contention that an unauthorized criterion--essay length alone--was applied in scoring Petitioner's essay.

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

66. Petitioner presented no evidence that any aspect of the GK essay process overall, including development, administration, evaluation, and score review, was arbitrary, capricious, unfair, discriminatory, or contrary to requirements imposed by law.

CONCLUSIONS OF LAW

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

68. Petitioner has the burden of proving by a preponderance of the evidence that he is entitled to the relief he 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.

69. 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 his exam by Pearson raters, who are experts in the field, is arbitrary and capricious. 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 5832, 586 (Fla. 1st DCA 1958).

70. In Harac, an applicant seeking licensure as an architect successfully challenged the failing grade received on the design portion of the exam, because of unique circumstances established in the administrative hearing. In particular, it was shown in the hearing that one of the three 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. In the administrative hearing, two expert witnesses provided testimony as to their evaluations of the design and the grades they would assign. One of the experts 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 evaluate the examinee's design in accordance with the holistic method or approved procedures, 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.

71. In marked contrast to Harac, there was no proof in this case that either of the two raters' scores was invalid, contrary to Pearson's scoring procedures, or improper in any way. Without such a showing, 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 credit given on 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 had his opportunity to prove fraud, imposition, discrimination,

manifest unfairness, or arbitrary or capricious conduct.

Petitioner failed to meet his burden of proof in this regard.

72. If it were appropriate to reach the second level of Harac, Petitioner's proof would fall well short of the necessary showing to sustain his score challenge. Unlike the examinee in Harac, Petitioner failed to offer expert testimony by an expert in holistic scoring and/or an expert teacher to attempt to replicate as closely as possible the holistic scoring method used by Pearson to score Petitioner's exam. Petitioner's non-expert testimony was far off the mark; as found above, he did not undertake an overall evaluation in accordance with the holistic scoring method. His attempted non-holistic, self-serving critique of the score given to his essay was wholly unpersuasive.

73. To the extent Petitioner contends that his challenge should succeed because scoring essay examinations are, by their nature, 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 his examination acted arbitrarily or without reason or

logic in giving him a failing score. Petitioner failed to meet his burden of proof in this regard.

74. Petitioner hints at criticism directed to the SBE for making the policy decision to toughen the certification standards by raising the score required to pass the GK essay. As noted, that choice, codified in the FTCE rule, was the SBE's prerogative and is not a matter subject to debate in this proceeding. Moreover, raising the standards for writing skills required to pass the GK essay was appropriate to align the FTCE 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.

75. Petitioner failed to prove his contention that the passage rates on the GK essay have been "low" since the passing score was raised. The fact that fewer GK essay examinees are passing is the expected consequence of the SBE's policy choice, codified in the FTCE rule, to increase the passing score requirement. Thus, while the passage rate was shown to be lower, there was no proof that the recent scores are "low," as opposed to the prior scores having been too high. The passage rates, standing alone, do not provide grounds for invalidating Petitioner's essay score.

76. Similarly, Petitioner's challenge to his essay score is not aided by the fact that the score verification process rarely

results in a change from a failing score to a passing score. The score change rate, standing alone, does not establish that the score verification process is arbitrary, capricious, unfair, discriminatory, or otherwise improper.

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 he received on the General Knowledge essay test taken in June 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 his exam and at the time the hearing was

held appear inconsequential; the relevant law addressed in this proceeding was not changed.

^{2/} The transcript portions from the hearing in DOAH Case No. 17-0423 that were adopted by reference have not been duplicated, but the following portions should be considered part of the record of this case: Volume I, pages 167 through 204 (Michael Grogan); Volume II, pages 219 through 257 (Phil Canto); and Volume II, pages 296 through 316 (Mary Jane Tappen).

^{3/} CLAST scores used to be accepted to demonstrate mastery of general knowledge for purposes of teacher certification. Use of CLAST scores was eliminated as of July 1, 2002, and instead, the Legislature directed the SBE to develop a "basic skills examination," which was the precursor to the GK four-part exam, to substitute for CLAST scores as the means to demonstrate mastery of general knowledge. See § 1012.56(3)(a) and (b), Fla. Stat. (2002). CLAST scores earned prior to July 1, 2002, are still accepted in lieu of passing scores in the corresponding subtests of the GK exam.

^{4/} In this case, after Petitioner contested DOE's determination that his essay was scored correctly, DOE asked the original raters to prepare written justifications for their scores. In addition, although DOE's practice in the score verification process is to have one chief reviewer prepare written comments to explain why the original score should stand or why it should be changed, in this case, after Petitioner had contacted DOE regarding a possible request for an administrative hearing, DOE had a second chief reviewer conduct an additional review and prepare written comments. Only the second chief reviewer testified at hearing, but her testimony was more about the process followed by chief reviewers in conducting score verification reviews, and she did not specifically address her written comments. The original raters did not testify. 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 his critique of his essay score, comparing parts of the comments with parts of his essay, but not doing so effectively or persuasively.

COPIES FURNISHED:

Daryl Bryant
3607 Brophy Boulevard
Cocoa, Florida 32926

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

Jennifer Diane Rose, Esquire
Post Office 924
Melbourne, Florida 32902
(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.