

**An Analysis of Principal Perceptions of Required  
Evaluator Proficiency Exams Used in the  
Primary Teaching Evaluation System in Seven U.S.  
States**

*Richard Dodson*

*Murray State University*

*This research examines how public school principals in seven U.S. states perceive the proficiency exam they must take and pass in order to evaluate their teachers. The test is centered on the states' primary teaching evaluation system, which is based on Charlotte Danielson's Framework for Teaching. An online survey was developed and 832 out of over 7,000 working principals across seven states responded, yielding a response rate of nearly 12%. States were selected to represent a cross section of high, middle, and low scorers in the annual Education Week "Quality Counts" report (Education Week, 2016). Results showed that most principals were not satisfied with the proficiency test that they must take and pass in order to evaluate their staffs. Many principals called for the elimination or drastic overhaul of their proficiency exam. Suggested changes also showed that most principals wanted better quality videos they must watch to evaluate teaching lessons, and they also wanted the test to be less subjective. The survey showed that more principals than not believed the test was unfair and should not stay the same.*

### **Introduction**

Teachers want, need and deserve evaluation processes that accurately identify their strengths as well as areas in which they need to improve (Almy, 2011). Almy stated that currently the evaluation systems in too many schools forgo high-quality feedback and fail to provide paths to improvement; this is "unfair to both the teachers themselves and the students who need their help" (p. 1). Evaluation systems for teachers have been under severe scrutiny for the

past decade; as a consequence, a majority of U.S. states have overhauled their teacher evaluation instruments in the past six years (Ruffini, Makkonen, Tejwani, & Diaz, 2014). Why has there been such a drastic push to change teacher evaluation systems? According to Weisberg, Sexton, Mulhern, and Keeling (2009), too often teacher evaluations are “too lenient, fail to adequately differentiate between teachers at different levels,” and recent teacher evaluation changes “are the result of dissatisfaction with evaluation systems that have largely failed to distinguish between effective and ineffective teaching” (p. 1).

Of the more than 30 states that have recently changed their teacher evaluation system, over 20 either have adopted Charlotte Danielson’s *Framework for Teaching* (FfT) as their single teacher evaluation system, created a modified version of it, or use it as one of their approved evaluation systems (The Danielson Group, 2013). These include Arkansas, Delaware, Kentucky, Idaho, Wisconsin, and South Dakota. FfT is also the default teacher evaluation framework for school districts in Illinois if they do not have their own (Teachscape, 2011). In Rhode Island, it is the foundation for the teacher evaluation system used throughout the state and is called the Rhode Island Model. New Jersey, Florida, and Washington also have approved its use. Before implementing an FfT evaluation system, principals must take and pass a proficiency exam. A state-wide analysis of how school principals feel about their FfT proficiency exam (PE) was conducted as part of a recent study of Kentucky principals. The results of that study showed that Kentucky principals were very unsatisfied with their state’s FfT proficiency exam (Author, 2015). The current study expands on that research by examining the perceptions that principals in seven other states have regarding their FfT’s PE.

While such efforts as using a new teacher evaluation system apparently aim at improving educational quality,

experienced administrators often find them troubling. Administrators themselves are not subject to evaluations based on student test scores, but they do face two different types of testing barriers. The first is that when student test scores are low, administrators themselves appear ineffective and may be held accountable. For example, in 2012, in Jefferson Parish, Louisiana, fifteen public school principals were fired because of low student test scores (Waller, 2012). The second hurdle is that administrators in some states face an additional high stakes testing challenge that comes before they can even evaluate any of their teachers. They must pass an exam showing their proficiency on teacher evaluation using FFT. This obstacle is addressed in this study.

#### **Significance of this study**

Ensuring student learning depends on conducting thorough and accurate teacher evaluations. Stronge and Tucker (2003) note that teaching is the fundamental part of schools and, “without capable, high quality teachers in America’s classrooms, no educational reform effort can possibly succeed” (p. 3). “Without high quality evaluation systems,” they add, “we cannot know if we have high quality teachers” (p. 3). At present, there is very little support for this claim (Taylor & Tyler, 2012), even though it seems to make good sense. While states “race to design new systems,” Taylor and Tyler (2012) say, “very little is known about how the availability of new information, or the experience of being evaluated, might change teacher effort and effectiveness” (p.1). Tucker and Stronge (2005) contend that most educators are in agreement that they are responsible for student learning, but “the profession as a whole has avoided evaluations based on measures of student learning, sometimes with good reason, given the unfair approaches that have been proposed.” They argue the solution is “not to continue with traditional strategies simply because they are benign and

comfortable, but rather to develop fair and reasonable means of assessing teacher success with students.” This study provides a first step towards examining the relationship between teacher quality and teacher evaluation systems by exploring how principals use and view such systems.

To be effective, teacher evaluation systems must be well understood by teachers and should result in the identification of authentic differences in performance (Danielson and McGreal, 2000; Milanowski, Prince, and Koppich, 2007). In addition, those using evaluation instruments must be equipped to do so effectively, yet many principals “have not received the mandate, the training, and the tools that will enable them to promote teachers’ professional growth as a result of evaluation” (Goe, 2013). Implementing an effective evaluation system involves including individuals with significant, recent experience in the classroom as evaluators, and everyone involved in the evaluation process should be well-trained to use the assessment instruments, including classroom observations, portfolio reviews, or whatever other methods are employed (Mathers, Oliva, & Laine, 2008).

Are current training models effective? This study examines principal perceptions of the evaluator proficiency test they must take in order to evaluate their teaching staff using Danielson’s FfT. By doing so, it provides a critical gauge of evaluation system efficacy: if those conducting evaluations do not feel well-prepared, the whole system may be undermined. This research uses a cross section of seven U.S. states found in the higher, middle and lower ranges of the Education Week state rankings over the eight-year period from 2009 through 2016. Wisconsin and Arkansas are in the high range; Delaware, Rhode Island, and Illinois are in the middle range; and Idaho and South Dakota are in the lower range.

### State education quality rankings

For the past 20 years, *Education Week* has ranked all U.S. states and the District of Columbia in education using six categories: K–12 Achievement; Standards, Assessments and Accountability; Teaching Profession; School Finance; Transitions and Alignment; and Chance for Success (an index that combines information from 13 indicators covering residents' lives from “cradle to career”). U.S. states and the District of Columbia also receive overall scores and letter grades based on the average of scores over the six categories (Education Week, 2016).

Table 1 shows individual yearly rankings of seven states that implement FfT over an eight-year period starting with 2009 and ending with 2016, according to *Education Week's* “Quality Counts” reports. It also shows each state's mean ranking over the same period. Wisconsin has consistently ranked in the top 20 states during this time. Arkansas was in the top ten from 2009 through 2014 but has slipped precipitously in rank for the past two years. It still averages 14<sup>th</sup> during the past eight years, however. Rhode Island has gradually improved its rank. Illinois also has progressively risen in rank before dropping to 17<sup>th</sup> this year. Idaho continues to rank near the bottom of the states, as does South Dakota. However, South Dakota improved to 40<sup>th</sup> in 2015 and then climbed up another notch to 39<sup>th</sup> this year. (The rankings include the District of Columbia, which increases the total number to be ranked to 51. *Education Week* did not rank states in their 2014 finding, but they assigned each state and the District of Columbia scores in each of the six categories and then provided an overall average score. The reader was then able to rank states in order, which this researcher did.)

**Table 1. Education Week State Rankings**

<b>State</b>	<b>2009</b>	<b>2010</b>	<b>2011</b>	<b>2012</b>	<b>2013</b>	<b>2014</b>	<b>2015</b>	<b>2016</b>	<b>8-yr. average</b>
<b>Wisconsin</b>	15	16	18	18	13	11	11	11	14.1
<b>Arkansas</b>	10	10	6	5	5	6	36	41	14.9
<b>Delaware</b>	18	22	22	25	19	18	15	16	19.4
<b>Rhode Island</b>	23	33	31	20	17	16	13	13	20.8
<b>Illinois</b>	41	41	40	29	28	24	15	17	29.4
<b>Idaho</b>	49	44	44	47	49	42	46	47	46.0
<b>South Dakota</b>	46	48	49	51*	51	51	40	39	46.9

### **FfT Proficiency Exams**

Charlotte Danielson teamed up with Teachscape, a company that delivers web-based learning content, to develop the teaching evaluation methods and instruments in her evaluation system (Teachscape, 2011). Teachscape also provides the proficiency exam that evaluators must take, along with training for them. This is also known as the Framework for Teaching Proficiency System or FfTPS. All of the states studied here use the proficiency test based on FfTPS, but the specific implementation varies. Some of these differences are highlighted below.

Illinois evaluators get two attempts to demonstrate proficiency; if an evaluator does not successfully complete the test on the first round, the Illinois State Board of Education (ISBE) offers face-to-face remediation before the second round. This introduces further stress: "... if a principal doesn't pass, it can impact their pride and confidence. They have to tell their superintendent and school that they haven't passed and can't evaluate in their school. So it's both high stakes and very personal," one principal said (Illinois Education Association, 2012). The Illinois State Board of Education acknowledges concerns about the number of evaluators passing the 7.5-hour test, but it stands by the stringent process (Illinois Education Association, 2012). Acknowledging the high-stakes nature of the test, ISBE also asserts that "Assessments with low reliability may lead to severe legal consequences and an inability to defend the fairness and reliability of the practice of observation. A highly reliable assessment ensures consistency in testing across a large base of examinees and from one administration instance to the next adding any well-constructed item to a test increases its reliability" (Morris, 2012, p. 37).

All new administrators in Arkansas who are designated as the person responsible for evaluating teachers and who are employees of a school district or open

enrollment charter school must successfully complete the Framework for Teaching Proficiency System test or FOCUS (Arkansas Department of Education, 2016). **Evaluator training in Arkansas involves** 20 hours of video and training content; 9 modules on Framework for Teaching; detailed rater training on each component and performance level in FfT Components for Domains 2 and 3; and over 100+ master-scored videos (Arkansas Department of Education, 2015). Similar to Illinois, Arkansas evaluators have two chances to pass each stage of assessment. If they pass both Stage 1 and Stage 2 they are considered “proficient.” If they fail a stage twice, they are deemed not proficient and are eligible for “retest.” There are specified wait times between failing an assessment and retesting; these range from 24 hours to 21 days. Wisconsin also uses lockout periods for failed test attempts. Wisconsin’s modified version of Charlotte Danielson’s FfT is called the Wisconsin Educator Effectiveness (WI EE) System (Wisconsin Department of Public Instruction, 2014). All evaluators here must pass the certification assessment in Teachscape. Those who initially fail the proficiency exam during Stage 1 or 2 are locked out for 24 hours. Two failed attempts result in a 30-day lockout; the tester must then begin again at Stage 1.

Rhode Island adapted Danielson’s 2011 Framework for Teaching Edition II rubrics to assess professional practice (Rhode Island Department of Education, 2012). This state’s proficiency test has two sessions, each a minimum of 3.5 hours of testing. A potential evaluator has up to six hours for each stage before the system times out, and he or she gets two attempts to pass each stage. An incomplete test counts as one attempt and the system logs the test taker out (Rhode Island Department of Education, 2014). Stage 1 of the Rhode Island FfT PE “has a combination of multiple-choice questions and video lesson scoring. Stage 2 is primarily video lesson scoring. ...each test session times out after 6 hours.



Videos can be paused for short breaks but if the computer is idle for more than 60 minutes or the user logs out they will fail the attempt” (Rhode Island Board of Regents Elementary and Secondary Education, 2012, p. 2).

All districts in Idaho except one use Danielson’s FfT as their observation and evaluation model. The one district not using the model has done a “crosswalk to the Framework showing how those components are being evaluated” (T. Carter, personal communication, Dec. 22, 2015). Would-be evaluators in Idaho have until September 2018 to pass the state’s proficiency assessment (Idaho Department of Education, 2014, p. 2).

South Dakota public schools began using FfT during the 2014-15 school year (South Dakota Department of Education, 2015). Known as the South Dakota Framework for Teaching, it is the state’s recommended teacher evaluation system; starting in the 2015-16 school year, all public schools in South Dakota must at least meet the minimum requirements of this model. The state’s training and proficiency assessment are provided online through Teachscape and are based on its Framework for Teaching Proficiency System or FfTPS. It takes approximately 30-35 hours to complete the training and assessment. Principals are eligible to receive state-paid licenses to complete FfTPS training (South Dakota Department of Education, 2011).

The proficiency exam is completed online. To illustrate, one test section asks about teacher “bias,” which in FfT training is described as something which “occurs whenever there is variability in an observer's application of the rubric based on a deep-seated belief about or stereotype of the individuals in the classroom (Teachscape, 2016).” The test-taker views a series of teaching videos. These videos focus on FfT teaching domains, which include: Domain 1: Planning and Preparation; Domain 2: Classroom Environment; Domain 3: Instruction; and Domain 4:

Professional Responsibilities (Arkansas Department of Education, (2015).

Depending on the state, the exam may focus on particular domains; however, in all states studied here the evaluator watches videos that emphasize domains 2 and 3, namely classroom management and instructional practices. Each domain has several components. For Domain 2 -- Classroom Environment -- the components include: 2a Creating an Environment of Respect and Rapport; 2b Establishing a Culture of Learning; 2c Managing Classroom Procedures; and 2d Managing Student Behavior. For Domain 3 – Instruction -- the components include: 3a Communicating with Students; 3b Using Questioning/Prompts and Discussion Techniques; 3c Engaging Students in Learning; and 3d Using Assessment in Instruction (Teachscape, 2016).

While watching each video, the evaluator is instructed to provide narrative **“evidence”** that supports his or her judgment as to the level of competency displayed by the teacher being viewed. In Rhode Island, for example, teachers are labeled on a performance level descriptor of Highly Effective (the highest level), Effective (the next highest level), Developing and Ineffective (the lowest level) (p. 8).

In the PE training, evidence comprises “only what you see, hear, or read—not your interpretation of what you see or your opinion about (Teachscape, 2016, p. 6).” The training adds that evidence “may be what the teacher and students say or what they do, including body language (e.g., teacher walking around the classroom, students waving their hands in the air to be recognized, students slumping in their chairs, students putting their heads on their desks)” (p. 6). In this way, the evaluation procedure should use only empirical evidence.

### Research Questions

Recognizing that effective teacher evaluations are essential for schools' success, this study examined how principals in seven states perceive their evaluator proficiency tests, all of which are based on Danielson's FFT. This research addressed the following questions:

1. How do principals perceive the difficulty of their evaluator proficiency exam (PE)?
2. What changes, if any, can be made to improve evaluator proficiency exams?

### Methodology

Seven states were selected to represent a cross section of high, middle, and low scorers in the *Education Week* "Quality Counts" report over the past eight years. Two states from the high range – Wisconsin (average ranking of 14.1) and Arkansas (average ranking of 14.9) – and two from the low range – Idaho (46.0) and South Dakota (46.9) – were chosen. Three states with middle range scores were also selected: Delaware (19.4), Rhode Island (20.8) and Illinois (29.4). While states were selected randomly from within each ranking category, they also represent different U.S. geographic regions. Working school principals in the states received an email letter with an electronic link to a survey on Survey Monkey. The link was first directly e-mailed to each superintendent of school districts in the seven states, requesting that they forward it to their principals. The survey was then directly e-mailed to all public school principals whose email addresses could be obtained from databases. An introduction letter accompanied the survey link. The survey used a Likert-scale attitude measure, forced choice (yes/no), and open-ended questions. Principals answered questions that examined their perceptions of their teacher evaluation system and the proficiency test administrators must pass to evaluate teachers.

### **Results**

Response rates varied by state. The states with the highest response rates were Arkansas (242 out of 1,100, or 24%), Idaho (134, or 19%), Rhode Island (66, or 13%), and South Dakota (70, or 12%). Lower rates were obtained from Illinois (234, or 9%), Wisconsin (79, or 7%), and Delaware (7, or 5%). In sum, 7,050 principals in the seven states combined could have responded to the survey; 832 actually participated, yielding a response rate of nearly 12% (11.8%). This falls within the average external online response rate range of 10 to 15% (SurveyGizmo, 2010; PeoplePulse, 2013).

### **Participant Demographics**

Of the 832 respondents who completed the survey, more than half (54.9%) were male (see Table 2). The majority of respondents (nearly 60%) were between the ages of 41 and 55 years; over 21% of respondents were 46 to 50 years old. Nearly 65% of respondents had a Master's degree plus 15 hours, and more than 11% had earned a Doctorate. Most had relatively little experience as a principal, with nearly 30% having been a school principal for less than five years and over half for less than nine years. Nearly 97% led public schools; about 3% worked in a charter school; and less than 1% listed their schools as a magnet or "other." Nearly half (over 48%) of all respondents worked in a rural setting. Slightly less than half (49.1%) were elementary principals, nearly 30% were middle/junior high school principals, and over 31% were secondary principals (over 8% reported they were Pre-K/K-12 principals). In sum, the categories with the largest proportion of respondents for each demographic measure were male; ages 41 to 55 years; masters plus 15 hours of education; fairly inexperienced; elementary principal; public school; and rural location.

**Table 2. Participant Demographics**

<b>N = 832</b>	<b>Percentage of Respondents</b>
<b>Gender</b>	
Male	54.9
Female	45.1
<b>Age in years</b>	
36 - 40	13.5
41 - 45	19.1
46 - 50	21.7
51 - 55	18.6
56 - 60	12.9
60+	07.9
<b>School setting/location</b>	
Rural	48.1
Town	19.5
Suburban	21.3
Urban	11.5
<b>Instructional level</b>	
(Pre)K-12	08.3
Elementary	49.1
Middle/Junior High	29.7
Secondary	31.3
<b>Highest level of Ed.</b>	
Master's degree	24.1
Master's degree + 15	64.6
Doctorate degree	11.27
<b>Type of school</b>	

<b>N = 832</b>	<b>Percentage of Respondents</b>
<b>Public</b>	96.6
<b>Charter</b>	02.9
<b>Magnet</b>	00.1
<b>Other</b>	00.5
<b>Years as a principal</b>	
<b>0 – 4</b>	30.6
<b>5 – 8</b>	21.7
<b>9 – 12</b>	21.3
<b>13 – 16</b>	11.8
<b>17 – 20</b>	06.7
<b>21 – 25</b>	03.7
<b>26+</b>	04.2

#### **FfT Proficiency Exam Completion & Perception of Difficulty**

In most of the states studied here, the vast majority of respondents had taken their state's proficiency exam (PE) and passed it on their first try (see Table 3). Idaho and South Dakota principals had the lowest first-time passing rates (55.6% and 41.1%, respectively), and Wisconsin principals had the highest first-time success rate (88.5%). Idaho and South Dakota also had a substantial portion of principals who had not taken their state's PE (22.6% and 50%, respectively). In every other state studied here, over 97% of principals had not taken their state PE; in Delaware, Rhode Island, and Illinois, all had. Overwhelmingly, respondents report that their state's PE is somewhat or very difficult (80% or above for every state). While most principals found their state's PE challenging, about half in each state believe their state's test is fair and should remain the same.

**Table 3. Respondents' Satisfaction with FfT Proficiency Exam (PE)**

<b>N=832</b>	<b>Percentage of Respondents Who Passed State PE on First Try</b>	<b>Percentage of Respondents Reporting Various Difficulty Levels of State PE</b>	<b>Percentage of Respondents Who Believe the State PE is Fair &amp; Should Remain the Same</b>
<b>Wisconsin</b>	88.5%	Very difficult 17.7% Somewhat difficult 65.8% Somewhat easy 13.9% Very easy 3.8%	44.3%
<b>Arkansas</b>	74.7%	Very difficult 22.9% Somewhat difficult 70.0% Somewhat easy 5.8% Very easy 1.3%	58.6%
<b>Delaware</b>	71.3%	Very difficult 14.3% Somewhat difficult 71.4% Somewhat easy 0.0% Very easy 14.2%	42.8%

<b>Rhode Island</b>	83.3%	Very difficult	19.6%	44.2%
		Somewhat difficult	68.8%	
		Somewhat easy	11.4%	
		Very easy	0.0%	
<b>Illinois</b>	85.1%	Very difficult	19.7%	53.5%
		Somewhat difficult	73.7%	
		Somewhat easy	6.6%	
		Very easy	0.8%	
<b>Idaho</b>	55.6%	Very difficult	27.3%	49.6%
		Somewhat difficult	66.7%	
		Somewhat easy	6.8%	
		Very easy	0.0%	
<b>South Dakota</b>	41.1%	Very difficult	18.0%	47.1%
		Somewhat difficult	74.0%	
		Somewhat easy	6.0%	
		Very easy	2.0%	



### **Suggested Changes to FfT Proficiency Exams**

Table 4 shows the top two changes that principals in each state suggested to improve their state's proficiency exam. As one can see, principals in five of the seven states wanted better quality videos provided for evaluators to watch as one of their two top choices. Principals in four states wanted to either remove or revamp the FfT proficiency exam. Two states' principals wanted to make the test less subjective, while two states' principals called for better test quality in general. Principals in one state also wanted to lessen the amount of time it takes to complete the exam, while in another state principals also wanted their state department to improve or provide more training for would-be evaluators.

Respondents from Idaho sent in 134 responses concerning their FfT PE. The most common responses (28) centered on the poor quality of their training videos. Among the comments were: "The videos are horrible and difficult to follow;" and "[Need] better quality videos. Leave out the fishbowl lens." The next most common response (23) revolved around general improvement or having assistance available. Comments included: "The training sessions offered were vague;" and "More support. We had to do it on our own." Fourteen Idaho principals called for eliminating the exam. Comments included: "The entire exam needs to be deleted;" "It is punitive at present...pass it or no recertification;" and "Put the control back at the local level." Of the 134 comments, only 3 were positive, with two saying "none" for changes and the other one saying, "It is adequate."

The most common response (5) from Delaware principals about the test centered on the test's quality. Comments included: "Some questions were very vague;" "Just give ratings and provide evidence to support;" and "Make [it] less subjective." None of the comments made by Delaware principals were positive.

**Table 4. Respondents' Perceptions of FfT PE**

<b>State</b>	<b>Total # of Responses</b>	<b>Most Common Responses</b>	<b># of Responses</b>
<b>Arkansas</b>	214	1. Have better quality videos and audio provided for evaluators to watch.	41
		2. Either remove or revamp FfT proficiency exam.	25
<b>Illinois</b>	234	1. Have better quality videos and audio provided for evaluators to watch.	66
		2. Either remove or revamp FfT proficiency exam.	17
<b>Wisconsin</b>	95	1. Either remove or revamp FfT proficiency exam.	17
		2. Make the test less subjective.	13
<b>Rhode Island</b>	64	1. Have better quality videos and audio provided for evaluators to watch.	16
		2. Either remove or revamp FfT proficiency exam.	15
<b>Delaware</b>	7	1. Improve test quality.	5
		2. Make the test less subjective.	1
<b>Idaho</b>	134	1. Have better quality videos provided for evaluators to watch.	28
		2. Have state department improve or provide more training available for would-be evaluators.	23
<b>South Dakota</b>	56	1. Have better quality videos provided for evaluators to watch.	9
		2. Improve test quality.	9
		3. Lessen the amount of time it takes to complete exam.	9

South Dakota principals sent in 56 responses. The three most common response themes (9 each) were training video quality; time demands; and test quality and format. Typical comments included: “The videos are TERRIBLE and almost impossible to hear or see clearly what is happening in the classrooms;” “...it is time consuming. It would be more helpful to have an individual train the administrators and then go through the process with them using information from a teacher they currently work with;” “The content is overly difficult;” and “This is a scapegoat for poor administrators to have tough conversations and in the end they do not have them because they blame it on the system. Have courage and have the conversations that are realistic.” Three South Dakota comments called for eliminating the exam. One positive comment said, “It is rigorous and spot on. No changes.”

Illinois principals sent in 234 responses regarding changes they believed needed to be made to their PE. As in Idaho and South Dakota, many (nearly 25%) focused on the quality of training videos. Comments included: “The videos are horrific. Hard to see, hard to hear and too, too many of them;” and “The audio was bad. I could not determine if students were talking about the lesson or talking about something else.” The second most common response (17) revolved around either eliminating or overhauling the exam. Comments included: “Completely revamp the exam;” and “Danielson’s framework reeks of her dislikes for principals, someone must have criticized her teaching when she was young.”

There were 64 responses from Rhode Island principals regarding possible changes to their PE. As in Idaho, South Dakota, and Illinois, the most common responses (16) targeted training videos. As Illinois, the second most common concern was the test itself: 15 responses called for eliminating or drastically changing the

exam. Comments included: “I don’t think the test was useful;” and “Get rid of it...I’m a professional – I went to college. I know good/poor teaching.” Three respondents wanted the exam to remain unchanged.

Arkansas principals made 214 comments regarding their PE and, once again, training video quality topped the list (41 comments). Echoing principals from Illinois and Rhode Island, Arkansas principals’ second most common response theme (25) called for eliminating or overhauling the state PE. Comments included: “Don’t make it like your [sic] having to pass the medical boards;” “Some principals had others take it for them;” “and “Delete it entirely. The logic is that everyone qualified to be a principal or superintendent has already proven to be qualified for the job of observer and they should not have to prove it once again.” In contrast to the other states studied, Arkansas principals provided more positive comments about their state’s system. Twenty-three responses positively commented on the state exam. These included: “I think it is good!” and “I feel the program that I participated in for TESS was a good indicator of what we see in the classroom and I was well prepared to carry out the Observation process.” Seventeen respondents said they wanted the exam to stay the same.

There were 95 responses from Wisconsin principals regarding their PE. The most common response (17) centered on eliminating or significantly changing the exam. Comments included: “Drop it. Let my Superintendent and School Board decide if I am evaluating appropriately;” and “Replace the Danielson Model with a more current and progressive mode.”

The second most common response (13) revolved around perceived subjectivity of the test. Comments included: “[Make it] less subjective;” and “Way too subjective.” As in other states, Wisconsin principals criticized training video quality, although it was only the third most

common response theme. Only one Wisconsin comment wanted the exam to stay the same, saying, “None at this time” about making changes to the test.

### **Conclusions**

This research makes clear that there are many principals in the seven states who are not satisfied with the proficiency test that they have to take in order to perform teacher evaluations using Danielson’s FfT model. In only two of the seven states do principals tend to think the test is fair and should stay the same. Of the 802 total suggestions made by principals regarding changes they would make to their FfT proficiency exam, 91 responses, or more than 11%, called for eliminating or drastically overhauling the proficiency exam. Only slightly more than 3 percent of comments about PEs were positive, and little more than 4 percent suggested the exam remain unchanged.

Significantly, the study showed that respondents in five of the seven states called for having better quality training videos as their top-priority for change in their state’s FfT PE. For respondents in four states, doing away with or overhauling their PE was one of the top two priorities, while one state’s principals (Wisconsin) made it their first priority. The second most common theme for principals from two states (Wisconsin and Delaware) was the exam’s perceived subjectivity.

The results of this study mirrored the results found in the recent Kentucky research which showed the top priority for Kentucky principals was to have better quality training videos (102 respondents). Kentucky principals’ second priority was to lessen the amount of time it takes to complete the exam (63 respondents). In the current study, only for South Dakota’s principals did time surface as a top-tier concern as measured by the number of suggestions concerning it.

This study suggests some relationship between a state's ranking in the annual *Education Week* "Quality Counts" report and principals' perceptions of FfT proficiency exams. Principals in three of the four high ranked states (Wisconsin, Arkansas, and Rhode Island) shared the desire to have the test either completely eliminated or overhauled. In addition, principals from two high ranked states (Wisconsin and Delaware) wanted their FfT PE to be less subjective. However, the study also showed that better training videos are needed in states from every ranking range -- in high (Arkansas and Rhode Island), middle (Illinois) and low (Idaho and South Dakota).

In sum, it is evident from this research that principals in the seven states studied wanted to voice their concerns about their version of the FfT PE, especially regarding the quality of the videos they have to watch to score the lessons. Nearly half of the respondents were elementary principals and nearly half worked in rural school settings. Rural schools tend to have smaller enrollments and, therefore, many have only one administrator who must handle all of the teacher observations and evaluations. The stress of sole responsibility may account for some of the dissatisfaction being expressed. It also points to a limitation of the study: the views of principals in middle/junior high and high schools and those in urban or suburban schools were represented less than those in elementary and rural schools. It may be that principals who did not respond to this survey felt no need to do so because they were satisfied with the teacher evaluation systems they use. Lack of knowledge regarding non-respondents is a common limitation of survey research; extensions of the current study should incorporate additional methods in an attempt to address this issue. Future studies should also focus on proficiency tests other than those based on Danielson's FfT.

Research such as this is vital in order to measure the perceptions that principals who use a version of FfT have about their proficiency exam. The survey showed that most principals in the seven states found the test difficult, and only two states had over 50% of respondents indicating the state test was fair and should remain the same. Will state decision-makers take to heart principal perceptions and modify their FfT proficiency exams? Hopefully, this research will catch the attention of state department officials and, at the very least, prompt them to improve the proficiency test training. This research offers a vital step in ensuring that the voices of those using FfT PEs are heard with the ultimate goal of supporting higher-quality education in the seven states studied and nationwide.

### References

- Almy, S. (2011). *The education trust*. Stronger teacher evaluation systems benefit teachers and students. 1. Retrieved from [https://edtrust.org/press\\_release/stronger-teacher-evaluation-systems-benefit-teachers-and-students/](https://edtrust.org/press_release/stronger-teacher-evaluation-systems-benefit-teachers-and-students/)
- Arkansas Department of Education. (2016). *TESS for administrators: Law & process training*. Administrator training. Retrieved from <http://www.arkansased.gov/divisions/human-resources-educator-effectiveness-and-licensure/office-of-educator-effectiveness/teacher-evaluation-system/tess-training-materials-and-resources-for-administrators>
- Arkansas Department of Education. (2015). Teacher valuation system, p. 13. Retrieved from <http://www.arkansased.gov/divisions/human-resources-educator-effectiveness-and-licensure/office-of-educator-effectiveness/teacher-evaluation-system>
- Dodson, R. (2015). Kentucky principal perceptions of the state's new teacher evaluation system: A survey

- analysis. *Educational Research Quarterly* 39, (2), pps. 53 – 74.
- Danielson, C. & McGreal, T.L. (2000). *Teacher evaluation: to enhance professional practice*. Alexandria, VA: Association for Supervision and Curriculum Development. Retrieved from [http://www.centerii.org/handbook/Resources/6\\_B\\_Improving\\_Staff\\_Evaluation.pdf](http://www.centerii.org/handbook/Resources/6_B_Improving_Staff_Evaluation.pdf)
- Delaware Department of Education. (2015). *DPAS-II guide (revised) for teachers*, updated August 2016. Delaware performance appraisal system: Building greater skills and knowledge for educators, pp. 2-3. Retrieved from [http://www.doe.k12.de.us/cms/lib09/DE01922744/Centricity/Domain/375/2016%20DPAS%20II%20Guide%20for%20Teachers\\_RevisedJN.pdf](http://www.doe.k12.de.us/cms/lib09/DE01922744/Centricity/Domain/375/2016%20DPAS%20II%20Guide%20for%20Teachers_RevisedJN.pdf)
- Education Week*. (2016). Quality Counts 2016: Reports and rankings. *Education Week*. Retrieved January 11, 2016, from <http://www.edweek.org/ew/toc/2016/01/07/index.html>
- Goe, L. (2013). Can teacher evaluation improve teaching? *National Association of Secondary School Principals*. Sept. 2012--May 2013. Retrieved from <http://www.nassp.org/tabid/3788/default.aspx?topic=CanTeacherEvaluationImproveTeaching>
- Idaho Department of Education. (2014). *Teacher and Principal Evaluation Rules IDAPA 08.02.02.120 and 121*. Retrieved from <https://www.sde.idaho.gov/.../Teacher%20and%20Principal%20Evaluation%20Rules%20IDAPA>
- Illinois Education Association. (2012). *Teacher, principal evaluator training hard but rewarding work*. Retrieved from <http://www.ieanea.org/2012/08/18/teacher-principal-evaluator-training-hard-but-rewarding-work/>
- Mathers, C., Oliva, M., & Laine, S. (2008). Improving instruction through effective teacher evaluation:



- Options for states and districts. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved from <http://www.tqsources.org/publications/February2008Brief.pdf>
- Milanowski, A., Prince, C., & Koppich, J. (2007). Observations of teachers' classroom performance. Washington, DC: Center for Educator Compensation Reform. Retrieved from <http://www.cecr.ed.gov/guides/CECRTeacherObservationModel.pdf>
- Morris, M. (2012). Consortium for educational change. *Illinois principal and teacher evaluation training program*, p. 37. Retrieved from [http://www.isbe.state.il.us/PEAC/pdf/prin\\_tchr\\_eval\\_trng\\_pres022412.pdf](http://www.isbe.state.il.us/PEAC/pdf/prin_tchr_eval_trng_pres022412.pdf)
- PeoplePulse. (2013). Quintessential marketing consulting pty ltd. Retrieved from <http://www.peoplepulse.com.au/Survey-Response-Rates.htm>
- Rhode Island Department of Education. (2012). RI model FAQs. Retrieved from <http://www.ride.ri.gov/TeachersAdministrators/EducatorEvaluation/RIModelFAQs.aspx#1283657-student-learning-objectives-faqs>
- Rhode Island Department of Education. (2014). Rhode Island model evaluation & support system. Retrieved from [http://www.ride.ri.gov/Portals/0/Uploads/Documents/Teachers-and-Administrators-Excellent-Educators/EducatorEvaluation/GuidebooksForms/Teacher\\_Guidebook\\_2015-16.pdf](http://www.ride.ri.gov/Portals/0/Uploads/Documents/Teachers-and-Administrators-Excellent-Educators/EducatorEvaluation/GuidebooksForms/Teacher_Guidebook_2015-16.pdf)
- Rhode Island Board of Regents: Elementary and Secondary Education. (2012). Training and support. Rhode Island model teacher professional practice, 2, 8, 10. Retrieved from <http://www.ride.ri.gov/Portals/0/Uploads/Documents/Teachers-and-Administrators-Excellent-Educators/FFTPS-and-FFTES-Overview.pdf>

- Ruffini, S., Makkonen, R., Tejwani, J., & Diaz, M. (2014). Principal and teacher perceptions of implementation of multiple-measure teacher evaluation systems in Arizona. IES National Center for Education Evaluation and Regional Assistance: Institute of Education Sciences. Retrieved from: <http://files.eric.ed.gov/fulltext/ED548539.pdf>
- South Dakota Department of Education. (2015). SD commission on teaching and learning develops model teacher effectiveness system. Retrieved from [https://doe.sd.gov/pressroom/educationonline/2013/oct/art\\_HotTopic.aspx](https://doe.sd.gov/pressroom/educationonline/2013/oct/art_HotTopic.aspx)
- South Dakota Department of Education. (2011). State-provided support and training. Evaluations of professional practice (Danielson Model). Retrieved from <https://doe.sd.gov/secretary/documents/doepdopts1.pdf>
- Stronge, J.H., & Tucker, P. (2003). Teacher evaluation and school improvement: Improving the educational landscape. Retrieved from [http://www.corwin.com/upm-data/7808\\_Stronge01.pdf](http://www.corwin.com/upm-data/7808_Stronge01.pdf)
- SurveyGizmo.com. (2010) Best practices. Retrieved from <http://www.surveygizmo.com/survey-blog/survey-response-rates>
- Taylor, E. & Tyler, J. (2012). Can teacher evaluation improve teaching? Evidence of systematic growth in the effectiveness of midcareer teachers. *Education Next*, (12)4, 1. Retrieved from [educationnext.org/can-teacher-evaluation-improve-teaching/](http://educationnext.org/can-teacher-evaluation-improve-teaching/)
- Teachscape. (2011). Charlotte Danielson's framework for teaching Sees record growth as states and districts redefine teacher evaluation. Retrieved from: <http://www.prnewswire.com/newsreleases/charlotte-danielsons-framework-for-teaching-sees-record>

- growth – as – states – and – districts - redefine-teacher-evaluation-124646653.html
- Teachscape. (2016). Learning and focus. Minimizing bias. Retrieved from <http://www.prnewswire.com/news-releases/charlotte-danielsons-framework-for-teaching-sees-record-growth-as-states-and-districts-redefine-teacher-evaluation-124646653.html>
- The Danielson Group. (2013). Framework. Retrieved March from <http://danielsongroup.org/charlotte-danielson/>
- Tucker, P. & Stronge, J.H. (2005). Linking teacher Evaluation and student learning by assessing teacher effectiveness. Association for Supervision of Curriculum Development. Retrieved from [www.ascd.org/publications/books/104136/chapters/The-Power-of-an-Effective-Teacher-and-Why-We-Should-Assess-It.aspx](http://www.ascd.org/publications/books/104136/chapters/The-Power-of-an-Effective-Teacher-and-Why-We-Should-Assess-It.aspx)
- Waller, M. (2012). 15 Jefferson Parish public school principals ousted under academic performance policy. *The Times-Picayune*. June 5, 2012. Retrieved from [http://www.nola.com/education/index.ssf/2012/06/sixteen\\_jefferson\\_parish\\_publi.html](http://www.nola.com/education/index.ssf/2012/06/sixteen_jefferson_parish_publi.html)
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. Brooklyn, NY: The New Teacher Project. Retrieved from <http://widgeteffect.org/downloads/TheWidgetEffect.pdf>
- Wisconsin Department of Public Instruction. (2014). Wisconsin educator effectiveness system. *Teacher evaluation process manuale Wisconsin educator effectiveness system 2014-2015 Version 4*. 9, 14. Retrieved from [ee.dpi.wi.gov/sites/default/files/imce/ee/pdf/eeteac\\_herevaluationprocessmanual-version4.pdf](http://ee.dpi.wi.gov/sites/default/files/imce/ee/pdf/eeteac_herevaluationprocessmanual-version4.pdf)