Exploring Science Education Faculty Attitudes Toward Standardized Testing

Tyler L. St.Clair, Oregon State University
Jennifer L. Maeng, University of Virginia
Randy L. Bell, Oregon State University
Lindsay B. Wheeler, University of Virginia

Abstract

This investigation explored science education faculty perceptions about the issue of standardized testing prior to and following a professional development for science educators. All 16 participants taught preservice teachers preparing to enter STEM classrooms; 9 participants came from education departments and 7 were from STEM departments at their respective institutions. Prior to the PD, each participant wrote a narrative describing an issue they encounter related to standardized testing and proposed a solution to that issue. During the PD, participants engaged in small-group discussions about their narratives and collaborated to better understand issues surrounding standardized testing and explore possible solutions.

Data consisted of participants' narrative texts, follow-up phone interviews, field notes and artifacts from observations of the PD. Analytic induction was employed to find emergent patterns in the data. Results indicated participants were knowledgeable about the problems surrounding standardized testing and had a variety of ideas about possible solutions to those problems. Faculty members from STEM departments were more likely to report problems related to student motivation toward the sciences. Faculty from education departments more often discussed the loss of instructional time due to testing and proposed ventures that involved equipping teachers with effective strategies. Following the PD, participants reported a high degree of consensus around the issues, but there was a marked change in participants' proposed solutions to include a greater perception of agency in effecting change at the level of state and national policy, both in terms of individual impact and leveraging the voice of the larger science education community to effect that change.

Introduction

Standardized testing has a long history in the United States. The rationale for the implementation of standardized assessments in the early 20th century came from educators concerned that grading practices in schools were too subjective (Giordano, 2005). In 1958, congress passed into law the National Defense Education Act. The purpose of this Act was to emphasize the importance of STEM subjects so the U.S. could compete internationally. This Act also provided support for standardized testing on a large scale (Heubert & Hauser, 1999). Since that time, the No Child Left Behind Act and Race to the Top have expanded the use of standardized testing and raised the stakes of those tests by linking test score results to student advancement, teacher evaluation, and school viability (McGuinn, 2011). In addition to domestic testing, U.S. performance on standardized international tests such as the Program for International Student Assessment and the Trends in Mathematics and Science Study have received a large amount of media attention and influenced U.S. educational policy decisions (Riley & Torrance, 2003). The development of the Next Generation Science Standards (NGSS) represents the most recent shift in the direction of school science assessment. The NGSS framework emphasizes coherence of curriculum, instruction, and assessment across all grade levels and also stresses the need to move away from traditional multiple-choice formats to other styles of assessment that assess a broader range of science skills such as the formulation of scientific explanations and engaging in scientific argumentation (NRC, 2012). The NGSS assessment practices however have not yet been finalized.

Many teachers perceive standardized tests as a barrier to good teaching and feel frustrated and powerless in relation to high-level educational policy decisions that impact their classrooms (e.g. Barksdale-Ladd & Thomas, 2000). Teachers report the pressure to perform well on the tests comes from many directions, including their school administrators, peers, and the media (Barksdale-Ladd & Thomas, 2000). This pressure leads teachers to use instructional and assessment strategies that mirror the format of state assessments, even though they recognize these strategies do not reflect high quality science instruction (Abrams, Pedulla & Madaus, 2003). In other words, standardized tests strongly influence curriculum and instruction.

High-stakes standardized testing is a charged issue for school administrators, students, and parents, all of whom are stakeholders affected by the current educational climate. Administrators are under pressure for their schools to show rising test scores and are often forced to cut back or eliminate school programs in the arts, recess, or other elective courses to achieve their goal (Kohn, 2001). Test scores are often used to determine a student's ability to advance through the grades; however, these tests may be biased against English Language Learners (Menken, 2008) and minority students (Walpole, et. al., 2005). Finally, parents have reported that they see little value in current standardized tests and that they cause undo mental stress for children (Barksdale-Ladd & Thomas, 2000).

While many studies focus on the perceptions of the groups most directly impacted by standardized testing, as indicated above, science education faculty perceptions of the various issues surrounding standardized testing have not yet been systematically studied. In this investigation, the term *science education faculty* refers to any professor in higher education that prepares preservice teachers to teach science at the K-12 level. These individuals are most often found in colleges of education or STEM disciplines within their universities and may teach a variety of courses in both the fields of science and education. It is important to better understand science education faculty perceptions of standardized testing as they are responsible for preparing future teachers to work in school environments where standardized testing is currently

the norm. In addition, science education faculty members are both well-informed and well-positioned to leverage their opinions to shape the future of standardized testing policy at both the state and national level.

A professional development (PD) experience specifically designed to meet the needs of science education faculty served as the context for the present investigation. This PD experience allowed faculty members a chance to explore new research and share effective teaching strategies in science education (McDonnough, Sterling, Matkins, & Frazier, 2012). Prior to the PD, participants submitted a "Vexation and Venture" text related to the topic of standardized testing (Johnston & Settlage, 2008). Before writing this text, participants were asked to outline an issue or problem related to standardized testing and to proposed a solution or way forward to address that issue or problem. These texts also served as a basis for structured discussions during the PD. During these discussions, participants made a short presentation of their vexation and venture to their peers, followed by a round of short clarification questions from the group. Next, the group discussed the issue without the original author's input, and finally the author rejoined the discussion to respond to issues the group raised. These structured discussions were designed to help science education faculty gain new insights from their peers through creating a professional learning community environment in which new knowledge is created through group discussions.

Social Constructivism

Social constructivism is the theory of learning that guided this study. Social constructivism views learning as a phenomenon that is first socially co-constructed and then internalized as individual learning (Duit & Treagust, 1998; Hodson & Hodson, 1998; Vygotsky, 1980). This lens acknowledges and highlights that learning takes place within a social context rich with dominant metaphors, shared understandings, and implicit meanings (Mitchell & Sackney, 2011), and it is an ideal choice for a PD experience that uses a professional learning community format. Professional learning communities are made up of individuals who are engaged in collaborative reflective inquiry, and who possess a shared vision and set of values (Stoll, et. al., 2006). Social constructivism also acknowledges that though learning first occurs in the social realm, learning ultimately takes place within the individual via internalization (Vygotsky, 1980). Thus, research that uses social constructivism as a theoretical framework is guided by research questions, data sources, and data analysis designed to explore individual cognition.

Purpose

The purpose of this investigation was to better understand science education faculty members' perceptions about standardized testing. A second purpose was to gauge the influence of a PD designed specifically for science education faculty to share and refine their ideas on standardized testing with other members of the higher education community. The research questions that guided this study were:

1. What patterns exist among the vexations and ventures of science education faculty related to standardized testing? Specifically, who or what was perceived as the cause of the vexation and who or what was affected? Who or what is perceived as being responsible for the solution?
2. What differences exist, if any, in the vexations and ventures of faculty from education departments vs. faculty from STEM departments?
3. In what ways did the PD influence participants' thinking about their vexations and ventures related to standardized testing?

Methods

Participants

Participants included 3 males and 13 females from 10 different colleges/universities in a mid-Atlantic state. Of the 16 participants, 9 were members of education colleges or departments and 7 were from science colleges or departments at their institutions. All but one of participants were first time attendees of the PD. Numerical codes are used in places names to protect the identities of participants. These codes indicate the participants' gender and position in their university.

Table 1. *Participant Demographic Data* (n=16)

Gender		Position		Ethnicity[1]		
Male	Female	Education Department	STEM Department	Caucasian	African American	Asian American
3	13	9	7	13	1	3

Note: [1] Participants may self identify with more than one ethnicity and percentages may add to >100%.

Context

The PD was held at a mid-Atlantic university in late spring. The five-day (37 contact hours) PD was implemented by a team of six facilitators. The purpose of the PD was to provide opportunities for participants to learn about new research relating to effective science teacher development and science teaching, to provide a venue for the sharing of syllabi and effective teaching strategies, and to help participants collaborate and build professional networks across institutional lines. Additionally, the PD was designed to foster collaboration among science education faculty who often find themselves as the sole science education faculty at their institution (Johnston & Settlage, 2008). Thus, participants in the PD explored an issue of importance to the science education community through a *vexation and venture* (V&V) text (Johnston & Settlage, 2008). Participants were asked to submit their vexation and venture texts of approximately 1,000 words related to standardized testing prior to attending the PD.

On the first day of the PD, participants were given a booklet containing all participants' texts and then split into two groups. A member of the PD implementation team acted as a facilitator for each group. Each group engaged in discussions surrounding their texts according to the Vexation and Venture model originally developed by Johnston and Settlege (2008) for use at the *Science Education at a Crossroads* conference. The framework for these discussions was as follows:

1. A participant presents his or her vexation and venture to the group (10 minutes)
2. The participant responds to brief clarification questions (5 minutes)
3. The participant remains silent while all other members of the group may speak about the issue (15 minutes)
4. The participant rejoins the group discussion and may respond, ask questions, or offer further information (5 minutes)

All participants shared their vexations and ventures during the first day of the PD according to this format. There were short breaks between each presentation. The issue of standardized testing was revisited later in the week when participants were asked to create a poster summarizing the main points raised in their V&V discussions. The focus on standardized testing in this PD provided an excellent opportunity to explore science education faculty thinking

around issues of standardized testing and the ideas generated within this professional learning community.

Data Collection

In addition to the vexation and venture texts, other data sources consisted of qualitative observations of the PD, follow-up semi-structured interviews with 12 participants (75%), and other PD artifacts.

Observations. Qualitative observations included both type written observations and personal inferences of sessions over all 5 days of the PD in order to capture any and all relevant discussions relating to standardized testing.

Interviews. The purpose of the approximately 25 minute semi-structured interviews was to clarify and elaborate upon participants' Vexation and Ventures as well as to better understand the role of the PD in influencing their thinking about their vexations and ventures. Interviews were conducted over the phone, recorded and transcribed. Questions about participants' vexation and ventures were added to an already existing protocol designed to assess other aspects of the PD. The questions relating specifically to the vexation and ventures were:

1. How would you describe the importance of the issue you raised in your vexation and venture, and how long has this been an issue?

2. How did you develop your thinking about your proposed solution of your venture before attending the PD?

3. In what ways did your experience at the PD shape your thinking about your vexation and venture, and was there some specific session or sessions that did that the most?

4. How did interactions with other participants during the PD influence your thinking about your specific V and V and your solution?

(Probe: Did any participants offer specific solutions to your vexation that you hadn't thought of before attending?)

Artifacts. Collected artifacts included initial vexation and venture text instructions to participants from the implementation team and handouts and PowerPoint slides from all PD sessions to better understand the context of their writings and discussions. Also, pictures were taken of posters that were made by the whole group during lunch on the 3rd day of the PD. These posters were created by the group to summarize important themes that emerged during the vexation and venture session, as well as possible actions that science education faculty could take. All artifacts served to help triangulate data gathered from the vexation and venture texts, session observations, and interviews.

Data Analysis

Analytic induction, as described by Bogdan and Biklen (2007), was used to analyze the texts, interview transcripts, and observation field notes. First, these data sources were examined for patterns, and from these patterns preliminary categories were developed and refined through comparison with the original data set. Reported themes consisted of patterns common to at least two participants. For example, preliminary categories for participant venture texts included the two codes "help prepare teachers for teaching best practices" and "help prepare teachers with better assessment strategies." On a second comparison with the data, these were collapsed into the single code "teachers need to be better prepared in a variety of ways." Frequencies of the codes were calculated to elucidate how often particular patterns were present in the data set. Following this initial analysis, themes between science education faculty from STEM and education departments were compared to uncover any similarities and differences between these groups. Themes present in data from qualitative notes of the PD and interviews following the

PD were compared to pre-PD vexation and venture texts to understand how the PD affected participant thinking around their issue. To enhance reliability, two researchers first independently analyzed the data and a consensus was reached through discussion about any disagreements before reporting the final themes.

Results

Themes emerging from the vexation and venture texts will be presented first followed by a description of any changes in participant thinking following the PD along with a discussion of these findings and their significance.

Vexations

Overall, participants were quite knowledgeable, expressing detailed opinions about the variety of issues and research related to standardized testing. There were a number of prominent themes in participants' vexations. Most participants mentioned more than one issue in the vexation portion of their texts.

Standardized testing has some positive aspects. All participants raised concern about standardized testing in one way or another, but many participants acknowledged that there are some positive aspects to standardized testing. They reported that standardized tests are an important piece of information about student achievement. "On the surface, our current model of giving multiple-choice standardized tests allows us to quickly measure at least some aspect of student growth without incurring huge resource costs." (STEM-F5, V&V) Another participant alluded to the fact that standardized testing may have some role in ensuring some level of accountability for teachers. "The intention of standardized testing at its onset was to ensure quality instruction for all students." (STEM-M2, V&V)

Questions are basic and don't value science practice/process. Another theme focused on the style of questions that most commonly appear on standardized tests. Participants were concerned that multiple choice questions are almost always focused on breadth over depth and focused on the easiest skills to assess like remembering facts or understanding basic concepts. "Discrete isolated factual knowledge and memorization of vocabulary are often the focus, while conceptual understanding, critical thinking, problem solving, creativity, and argumentation fail to be addressed." (ED-M1, V&V) This basic questioning style may not reveal the kind of deep learning that may be occurring for students.

> One of the problems I have with standardized tests is that they are looking for a specific answer to each question. Those students who think outside the box are ultimately penalized and may even be labeled as being unintelligent; however they may be quite brilliant. (STEM-M1, V&V)

Student skills such as analyzing, evaluating and arguing from evidence may simply be too difficult to test using a multiple-choice format. Other participants specifically mentioned that standardized test questions do not emphasize science process skills (or scientific practices). "The multiple choice exams don't allow for students to be creative in designing experiments, communicate scientific ideas, or demonstrate how they think like a scientist." (STEM-F5, V&V) Participants perceived that the tests are not able to show deep student learning, and that many of the valuable skills students learn and practice in science classes are not measured through current standardized assessments.

Testing pressure negatively affects students, teachers, and administrators. Participants described standardized testing in schools as having a negative effect on student motivation and buy-in to learning science. One participant compared the current state of schools with his own experiences as a student prior to the implementation of widespread testing.

> The majority of my current students are not curious about science, which I think stems directly from the "teach to the test" mentality that many teachers must adopt to enable their students to score well on [state] tests… I was able to dig deep into a topic, wrestle with it, and glean additional information about the topic not presented to me during class. The students I see now either cannot do this or do not want to do this. (STEM-M1, V&V)

Participants made many such statements about standardized testing being the cause of a decline of enjoyment for students in their desire to learn and do science. Participants perceived attaching high stakes, such as teacher evaluation, to the tests caused more testing to take place in schools. The following quote from one participant's V and V described an exchange between two students overheard by the participant's sister, who is a first grade teacher

> Student 1: *What special do we have today after lunch? Do we have music?*
> Student 2: *Hmm. Oh, I know! Assessment!* (STEM-F2, V&V)

This exchange concisely portrays the concern from participants that too much class time is being spent preparing for and practicing testing.

Participants perceived that it was unfair for schools to use standardized assessments to evaluate teacher quality or strongly influence teacher pay as this demotivated teachers. For example, one participant wrote, "School systems have become so dependent on the standardized test system that they now hold merit pay and tenure over the teachers' heads as a carrot, and if not enough students meet minimum qualifications, teachers are punished." (ED-F7, V&V) Participants also wrote that teachers are pressured to teach in less effective ways, and that school curriculum and classroom teaching has changed to mirror the assessments.

> In an effort to meet state standardization requirements, teachers find themselves restricted to repetitive-type teaching, i.e., "teaching to the test," which primarily involves drilling students to recall rehearsed information rather than engaging in reflective instruction that involves science explorations and investigations (ED-F3, V&V)

Using class time to prepare students for the format of assessment by having them practice state assessments did not go unnoticed by participants. One participant noted, "Instead of teachers building in classroom assessment to plan how to differentiate their instruction, countless hours are spent testing." (ED-F5, V&V) In many states, standardized testing extended down into the elementary grades and emphasized mathematics and reading disproportionately to other subjects. This has led to increased time reviewing mathematics and reading, and a marginalization of time spent on learning science and social studies. "I also hear about how little time my elementary teachers have to teach science, as teachers are busy getting students ready for mathematics and reading standardized tests." (ED-F4, V&V) Together, these quotes about teacher pressure indicate that high-stakes assessments drive instruction in ways that leave teachers feeling helpless, fearful, and with no good options about how to structure class time.

Student performance on standardized tests also affects school administrators in important ways. One participant wrote:

> For administrators, standardized testing can be a source of pride if the students in their schools are doing well on the tests, or a source of fear of intervention from various agencies if test scores do not improve to meet the demands of adequate yearly progress. (STEM-F2, V&V)

One participant gave a vivid picture of the various conclusions the public might rush to make about the quality of a school that produces low test scores:

> The perspective apparently is that the highly publicized scores tell us about the effectiveness of a school, how much the students know, how effective the teachers are in

> doing their jobs, and the quality of the leadership in the school. When the scores are published in our local newspapers, we can see whether or not a school is doing a good job of teaching its students. If the scores are low (unacceptable), we assume that students are not being taught well, that differentiation is not occurring, that the curriculum is not appropriate or challenging, or that appropriate resources have not been channeled to that school to ensure that students are getting the quality of education to which they are entitled. (ED-F1, V&V)

As long as test scores are equated with teaching quality and administrative effectiveness, participants perceived that the quality of instruction in schools is likely to be devoted largely to test preparation, and this high stakes culture of fear is likely to continue.

Tests are biased against certain students. Participants pointed out that standardized tests unfairly disadvantage certain subgroups of students. These included students of low SES backgrounds, students from minority ethnicities, English Language Leaners, students who move from one state to another, and international students moving to the U.S. In talking about some of these subgroups, one participant mentioned:

> Furthermore, research has shown that minority and low-income students are more affected by standardized tests. In states with higher percentages of African Americans, Hispanics, and poorer students, there is a greater focus on test preparation and required high school graduation exams are more common (STEM-F5V&V)

This quote implies that the participants perceived the more pressure there is on schools to improve scores, the more time will be devoted in these schools to test preparation. Another participant elaborated on the ways in which she perceived these tests are unfair to low SES students; the pressure on teachers to show good scores dissuades teachers from taking jobs in high-needs districts.

> If we judge teachers solely on the performance of their students on an end of the year standardized test, then those teachers in affluent, high-achieving schools (or school systems) will routinely have higher evaluations and therefore judged to be "better" teachers. This contributes to an exodus of good teachers from poorer, rural, or lower achieving schools, the very places where they stand to make the most impact, to more affluent, higher achieving schools. (STEM-F3, V&V)

Tests are an incomplete picture of student achievement. Another perception among participants was that standardized tests do not show the full range of student achievement. "It is an unfair tool of measurement and an incomplete evaluation of student knowledge" (ED-F2V&V). Additionally, participants noted that many factors complicated summative assessments measuring student performance at only one point in time. For example, another participant commented "Since standardized tests results are a single data point, they represent a snapshot of the student's performance. There are so many uncontrolled variables contributing to a student's performance at the assessment time" (STEM-F3, V&V).

Students are not being properly equipped for their futures. Finally, participants mentioned some different ways in which they felt the overemphasis on standardized testing does not adequately prepare students for their futures. Some participants expressed this in a general way, while others gave some specific examples. One example was that students are not being equipped with skills they need to enter colleges and universities. "In my opinion the students I see in my classes now are less prepared, less motivated, and less inquisitive than the students I had at the start of my career. I attribute this in part to SOL testing." (STEM-F1, V&V) Other participants emphasized that students are not being prepared well for their future professions

with the skills they will need. "If education is relegated to picking and choosing what is taught in order for students to pass the state test, what are we doing to foster creativity, inventiveness, inquiry, and investigation, which ultimately support the development of our future leaders, scientists, writers, artists, and musicians?" (ED-F1, V&V) Participants were concerned that the skills most easily assessed on multiple choice tests are not aligned with the skills students will need when they go onto pursue higher education or enter the workforce.

Ventures

While many clear patterns emerged from participants' vexation texts, their ventures were far more varied and unique to each individual, or sometimes to a particular school or classroom context. In general, relative frequencies were lower for each of the venture themes compared to participant vexations for these reasons. Below, emergent themes that arose across participants are described.

Current tests should be modified or supplemented with other assessment types. One suggestion from participants was that the standardized tests should be modified from the current multiple-choice format. Though this was a common theme, the specific approaches suggested varied between participants. For example, one participant suggested trying to modify the question format to test a wider range of skills:

> Instead of a multiple-choice test based purely on lower-level learning objectives, why not add a free response portion that asks the students to design an experiment to investigate a phenomenon or apply their content understanding to complete some task? An even more progressive idea would be to give students actual opportunities for these application/discovery experiments where the students would be asked to perform a practicum as part of the exam… Asking students to explain their experiment on an exam would also encourage teachers to focus on developing their students' communication skills. Students would have the flexibility of using a variety of modes to describe their experimental set up (pictures, text, graphs, etc.) but they would need to learn how to explain science in every-day language. (STEM-F5, V&V)

This participant proposed including questions designed to assess students at a more complex level of understanding than basic facts and concepts, and incorporated some science practices. Another participant was hopeful that the newly released Next Generation Science Standards (NGSS) were written in a way that could help these changes be realized:

> Science education has an opportunity that it has not had since the mid 1990s to make significant changes to the ways in which we conduct assessment, and the NGSS has the language to help facilitate this move. The Performance Expectations sections have both the initial appearance and the philosophical underpinnings to support authentic and performance-based state-administered assessments. (ED-M1, V&V)

Similarly, other participants focused on supplementing traditional multiple-choice tests with other assessment types versus changing the style of questions on the assessments. Providing some examples, one participant said:

> Standardized tests can be one part of a comprehensive assessment system but should not be the sole or major assessment tool utilized to measure student achievement or teacher effectiveness… These solutions for the use of more authentic assessment should be incorporated in the summative evaluation of students. Other examples would include a collection of student work in portfolios, e-portfolios, and written explanations such as essays and projects. (ED-F3, V&V)

All of these suggestions indicate participants perceived assessing skills related to scientific practices and problem solving around authentic experiences to be important in a balanced science education.

Teachers need to be better prepared in a variety of ways. Participants reported that another possible venture would be to better prepare teachers to use best practices in their classrooms. One participant expressed this with regard to culturally responsive best practices:

> If education is to be viewed as a pathway out of the pedagogy of poverty and as promoting student achievement, then there is a significant need to implement pedagogical methodologies that are culturally responsive and connect to student lives. (ED-F3, V&V)

Others focused specifically on the need to prepare teachers with strategies that could allow them to better assess their students. "I also believe that we need to give teachers the independence to create their own assessments beyond traditional multiple choice tests." (ED-F8) While many participants mentioned that teachers should be prepared to teach or assess in different ways, no comments were made specifically about how to help teachers effectively implement these changes within the high-stakes testing environment of their schools.

The role of standardized tests should be re-evaluated. Some participants emphasized that ventures must be based on a change of how the many stakeholders in the education system value and think about education in the U.S. Some participants mentioned specific changes in how test scores should be used. "Schools and states need to stop tying tenure and promotion to the test scores and stop tying the amount of aid a school will receive based on their participation and scores on the test" (ED-F7, V&V). Other participants spoke of more general changes in attitudes as exemplified by the following description:

> Time and time again I have heard colleagues quote studies saying that employers are not "that concerned with a student's major" but rather they are looking for students that can "think critically, communicate effectively, and solve problems creatively." If this is what we truly value in education, then there has to be to be a shift in the way we measure students' abilities in these areas. (STEM-F1, V&V)

These participants' sentiments indicate that a systematic change in thinking is necessary prior to implementing specific changes in schools or classroom instruction.

The problem of limited resources must be addressed in a solution. Some participants cited resource limitations in any widespread testing system and sought to provide some solutions. For example, one participant wrote, "If we offered fewer exams over a student's academic career, it is feasible that we could find the resources to grade these open-ended assessments" (STEM-F5, V&V). Saving money by reducing the number of assessment times, or streamlining state standards to a uniform national set of standards was one way participants thought to save money. Others had more elaborate suggestions:

> The role of authentic assessment as a valid tool for measuring meaningful workforce skills and understandings may lead to cost saving opportunities to engage students in public service efforts. For example, if students were able to demonstrate their knowledge through a public works effort there may exist not only a potential to cut testing costs, but also an opportunity to save tax payer money regarding labor associated with that public work effort. (ED-M1, V&V)

Regardless of specific differences in approaches to save money, participants indicated that solutions should be cost effective and economically feasible for schools.

The power to change the situation resides with the policy makers. A final theme emerging from the venture texts was that some participants felt that the onus of reforming the

current system must come from the top down. "Policy makers at the national, state, and district level must begin discussing alternatives to standardized testing" (ED-F4, V&V). Others discussed the need for more transparency from those in power. "The stakeholders who reinvent these tests and standards need to become a more transparent entity and more available to those they are serving." (ED-F7, V&V) What did not emerge as a theme in participant ventures was any perceived ability to influence those in positions of power. Participants did not express solutions related to helping create standards, modify standards, or lend their voice to influence top down change.

Comparing STEM and Education Faculty V&Vs

When comparing emergent themes from faculty in STEM departments (7 participants) to those in education departments (9 participants), more consensus than divergence in their vexations and ventures existed. With regard to vexations, two large differences between STEM and education faculty emerged:

(1) STEM faculty members were more likely to mention vexations related to decreased student motivation toward science disciplines (71% of STEM faculty vs. 11% of education faculty).
(2) Education faculty members were more likely to mention vexations related to negative effects on classroom instruction (56% of education faculty vs. 14% of STEM faculty).

The biggest difference between education and STEM faculty ventures was that education faculty were more likely to mention solutions related to teacher preparation, such as preparing teachers to use best practices and effective/alternative assessments in the classroom (55% of education faculty vs. 29% of STEM faculty).

Participant Thinking Immediately Following the V&V Sessions. Following the vexation and venture session, the participants had a group discussion to summarize the major themes emerging from the V&V sessions and specific actions that science education faculty could take to address these issues. The consensus issues/problems about standardized testing that participants decided upon closely aligned with themes found in the V&V texts. The document that the group created during this discussion included the following:

1. Scores are used to exclude/marginalize
2. Fear (teachers/students)
3. Causes lack of resources for teaching and poor instruction
4. Poorly conceived and designed assessments
5. Loss of instructional time
6. Lack of understanding of legislators

Enhancing preservice teacher training in best practices was a theme common to participant ventures and was also mentioned during the collaborative discussion following the session. The other three possible actions agreed upon in this collaborative discussion were not patterns present in the V&V texts. In addition to these ventures, the specific actions that science education faculty collaboratively agreed they might take with regard to these issues were:

1. Visible involvement with policy makers at all levels
2. Involvement in the standards review process
3. Equip elementary teachers to not fear science content.

Since these actions were mentioned in the summary large group discussion after the V&V activity during the PD, it is possible that the V&V small group discussions influenced participant thinking. The first two new ventures involved using leverage as a science education professor to influence policy, both through actively helping to shape future standards as well as

through direct interaction with policy makers. The third new venture of equipping elementary teachers with skills so they are not afraid to teach science was not mentioned by any participant in the V&V texts arose from the V&V session discussions.

Participants' Thinking Following the PD. Further evidence about changes in participant thinking came from phone surveys conducted within two weeks following the PD. Of the 16 participants attending the PD, 12 responded to interview requests (75%). The following themes arose from interview transcripts.

There was much consensus around the issues, and this was empowering. Participants mentioned hearing opinions similar to their own during the V&V sessions, especially with regard to the vexations surrounding standardized testing was a positive experience. One participant said, "I just realized there's a lot of other really talented people out there that I can tap and ask for questions and network with and, again, I'm not out there on my own." (STEM-F4, Interview) Participants were happy to realize they were part of a community that shared similar concerns.

Thinking was mixed as to whether the V&V session changed participants' perceptions. Participants were specifically asked in the interview as to whether or not their thinking changed following the V&V session and also whether or not participants offered helpful opinions or feedback. Two opposing themes emerged. Some participants said the V&V session did not change their thinking (42% of interview respondents). When asked if the V&V had changed her thinking, one participant said, "I would say not at all, to be fair… Nobody wrote a paper about how awesome standardized testing is." (ED-F4, Interview) On the other hand, many respondents indicated they heard interesting ideas during the session or that others gave them helpful feedback (50% of interview respondents). One participant talked about others' comments on her presentation, "In fact, I had a few during my presentation, who made a comment that that was a very good idea." (ED-F2, Interview) Participant self-perceptions as to changes in their own thinking clearly varied.

Participants expressed greater agency to influence policy. Prior to attending the PD, no pattern existed in V&V texts that participants felt a connection to policy makers or an ability to affect policy. As in the immediate post-V&V group discussion, participants expressed more opinions about policy and expressed both a desire and ability to influence policy makers in interviews:

> Well it allowed me to look at some other areas of where I could help with just not knowing the problem or seeing the problem, but also taking the problem outside of the college classroom to the legislative areas and looking at some of the other stakeholders' interest in, I guess, implementing standardization… Whenever we can, voice our opinions to legislators, to administrators and kind of move up the ladder and take it outside of the college to help get our concerns out there. (ED-F3, Interview)

This participant recognized that she had the ability to share her knowledge and opinions to those in positions of power.

Another participant voiced the desire to involve people at all levels in the science education community in policy decisions:

> Science educators could be more heavily involved in policy in order to have practitioners' voices heard… rather than other voices that may not be as informed about education in general and, yeah, I think it's important… A lot of the comments were really thought provoking and helped me refine some of my thinking about what might work and what might not work in terms of us trying to help pre-service teachers become more engaged in policy in their careers. (ED-M1, Interview)

It seems that the PD experience helped participants realize they were part of a larger community with shared opinions and concerns about standardized testing, and that it could be possible to work together in realistic ways to make a difference at all levels of the policy decision making process.

Discussion

Analysis of Vexation and Venture texts and observations of PD discussions revealed science education faculty thinking about issues surrounding standardized testing. The high degree of consensus on these issues and deep knowledge in the group is understandable considering that this charged topic relates directly to their careers and their work with preservice teachers. Participants described a culture of fear for many of the stakeholders in standardized testing and that the tests are too limited in their scope and often biased. One of the main goals of the standardized testing movement in the U.S. was to identify and better serve under-performing schools and high-needs students. In fact, participants mention these tests are having just the opposite effect.

The proposed ventures may have been more varied simply because they have yet to be tried. Also, while vexations are easy to state in a relatively broad way, many ventures were rooted in specific classroom or school contexts. Some minor differences among science education faculty coming from STEM vs. education departments existed. Professors from STEM departments focused their vexations more on student motivation to learn science, while faculty from education departments were concerned about loss of instructional time and focused their ventures around how to equip teachers with effective skills. These differences may be influenced by many factors such as the types of classes they teach, the values of colleagues in their respective departments, their background in learning science or the student populations they teach.

Based on the lack of discussion about policy-based solutions to standardized testing issues in the pre-PD texts and the feeling that it was the sole responsibility of policy makers to be involved in any policy changes, it seems that science education faculty may feel as disconnected from policy decisions as do practicing teachers (Barksdale-Ladd & Thomas, 2000). A noticeable changed occurred following the PD however. Participation in the professional learning community appeared to help foster a sense of shared vision around these issues with a greater perception of agency to affect changes in policy.

Standardized testing is a charged topic in the U.S. with many stakeholders weighing in on issues surrounding the tests. The science education faculty members in this study were clearly well informed about these issues and uniquely positioned to have impact on a large number of teachers entering the teaching profession. They could also serve to help educate in-service teachers, administrators and the public about the complex issues around standardized testing, and to inform and influence future policy decisions at the state and national level. This study helps to better understand science education faculty perceptions about standardized testing and how PD initiatives could help provide them with greater agency to collaborate, to refine their understandings, and develop potential solutions to these issues.

It is important to mention that as a qualitative case study of a subgroup of science education faculty, results are not generalizable to a larger population. This study does suggest that future research is needed about specific strategies to help science education faculty be more involved in the policy arena.

This research was supported by funding from the U.S. Department of Education Investing in Innovation (I3) grant program. However, the results presented here do not necessarily represent the policy of the U.S. Department of Education, and you should not assume endorsement by the Federal government.

References

Abrams, L. M., Pedulla, J. J., & Madaus, G. F. (2003). Views from the classroom: Teachers' opinions of statewide testing programs. *Theory into Practice*, *42*(1), 18-29.

Barksdale-Ladd, M. A., & Thomas, K. F. (2000). What's at stake in high-stakes testing teachers and parents speak out. *Journal of Teacher Education*, *51*(5), 384-397.

Bogdan, R. C., & Biklen, S. K. (2007). *Qualitative research for education: An introduction to theories and methods (5th Ed).* Boston, Massachusetts: Pearson Education, Inc.

Duit, R., & Treagust, D. F. (1998). Learning in science- from behaviourism towards social constructivism and beyond. *International Handbook of Research on Science Education*, 3-25.

Giordano, G. (2005). *How testing came to dominate American schools: The history of educational assessment*. Peter Lang.

Heubert, J., & Hauser, R. (Eds.). (1999). *High stakes testing for tracking, promotion, and graduation*. Washington, D.C.: National Academy Press, 13-15.

Hodson, D., & Hodson, J. (1998). From constructivism to social constructivism: a Vygotskian perspective on teaching and learning science. *School Science Review, 79*(289), 33-41.

Jablon, P. C. (2002). The Status of Science Education Doctoral Programs in the United States: The Need for Core Knowledge and Skills. *Electronic Journal of Science Education*, *7*(1).

Johnston, A., & Settlage, J. (2008). Framing the professional development of members of the science teacher education community. *Journal of Science Teacher Education*, 19(6), 513-521.

Kohn, A. (2001). Fighting the Tests: Turning Frustration into Action. *Young Children*, *56*(2), 19-24.

McDonnough, J. T., Sterling, D. R., Matkins, J. J., & Frazier, W. M. (2012). *Virginia science education at the crossroads: Connecting science education faculty to a professional community.* A paper for the Annual meeting of the National Association of Research in Science Teaching, Indianapolis, IN.

McGuinn, P. (2011). Stimulating reform: Race to the top, competitive grants and the Obama education agenda. *Educational Policy*, 0895904811425911.

Menken, K. (2008). *English learners left behind: Standardized testing as language policy* (Vol. 65). Multilingual matters.

Mitchell, C., & Sackney, L. (2011). *Profound improvement: Building capacity for a learning community*. Taylor & Francis.

National Research Council. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: The National Academies Press.

Riley, K., & Torrance, H. (2003). Big change question as national policy-makers seek to find solutions to national education issues, do international comparisons such as TIMMS and PISA create a wider understanding, or do they serve to promote the orthodoxies of international agencies? *Journal of Educational Change*, *4*(4), 419-425.

Stoll, L., Bolam, R., McMahon, A., Wallace, M., & Thomas, S. (2006). Professional learning communities: A review of the literature. *Journal of Educational Change, 7*(4), 221-258.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Walpole, M., McDonough, P. M., Bauer, C. J., Gibson, C., Kanyi, K., & Toliver, R. (2005). This Test is Unfair Urban African American and Latino High School Students' Perceptions of Standardized College Admission Tests. *Urban Education, 40*(3), 321-349.