Do they see it coming? Using expectancy violation to gauge the success of pedagogical reforms

Jon D. H. Gaffney,[1] Amy L. Housley Gaffney,[2] and Robert J. Beichner[1]

[1]*Department of Physics, North Carolina State University, Campus Box 8202, Raleigh, North Carolina 27695, USA*
[2]*Department of Communication, North Carolina State University, Campus Box 8104, Raleigh, North Carolina 27695, USA*

We present a measure, which we have named the Pedagogical Expectancy Violation Assessment (PEVA), for instructors to gauge one aspect of the success of their implementation of pedagogical reform by assessing the expectations and experiences of the students in the classroom. We implemented the PEVA in four physics classes at three institutions that used the Student Centered Active Learning Environment for Undergraduate Programs (SCALE-UP) pedagogy in order to gain an understanding of students' initial expectations, how those expectations are shifted during early classes, and what students report experiencing at the end of the semester. The results indicate appropriate shifts in student expectations during orientation, but some gaps between student expectations and experiences persisted. Students rated the communication aspects of SCALE-UP as desirable and indicated an overall positive affect toward the pedagogy, indicating that violations of their initial expectations were largely positive. By studying the patterns of the shifts in students' expectations and gaps between those expectations and their experiences, we gain insight for improving both the orientation of the students and the implementation of the course.

PACS number(s): 01.40.Fk, 01.40.Di

I. INTRODUCTION

Imagine the stereotypical introductory science class at a large university. Chances are, you pictured an amphitheater or rows of desks, a projector aimed at the front of the room, and an instructor facing a sea of students while lecturing. While we have all experienced this situation as students and instructors, some classes break this norm. Pedagogical reforms often push aside the elements that are foundational to our conceptions of classes—such as lecture—in favor of creating an environment conducive to multiple interactions. While we know that such reform can have a positive impact on learning outcomes, we do not often consider the jarring reality for a college student entering a classroom that is not only physically different, but also promotes participation in ways that are perhaps unfamiliar or uncomfortable.

In this paper, we explore the phenomenon of how students react to reformed pedagogy—specifically the Student Centered Active Learning Environment for Undergraduate Programs (SCALE-UP) pedagogy.[1,2] By first grounding this research in previous literature about SCALE-UP and then explaining the expectancy violation framework,[3,4] we set the stage for understanding students' expectations and experiences in reformed classes. We then present a structure for investigating those expectations and experiences in the form of a customizable survey. The results from our implementation of that survey have practical implications for instructors of reformed courses, as well as insights for considering how our students view their learning environments.

II. BACKGROUND

Students can learn more when they interact with faculty, collaborate with classmates, and are actively involved.[5–7] For instance, students in SCALE-UP showed better improvement in conceptual understanding than did students in lecture classes, as demonstrated by higher normalized gains on force and motion concept tests.[2] Such results are not uncommon among reformed pedagogies.[5] Furthermore, the problem-solving abilities of students in SCALE-UP sections are as good or better than the abilities of students in traditional sections. SCALE-UP sections report higher rates of class attendance and a substantial reduction in failure rates, particularly for women and minorities.

Although reform pedagogies have often been demonstrated[2,5–7] to yield vastly improved student learning over traditional instructional methods, some failures have been reported (see e.g., Fig. 1 of Ref. [5], where a few "interactive engagement" courses fall close to the average normalized gain line representing the average for "traditional" courses). Furthermore, over half of the articles addressing efforts to promote change in the instructional practices of science, technology, engineering, and mathematics (STEM) faculty as reviewed by Henderson *et al.*[8] offered little or weak evidence of support for change strategies, and no articles offered strong support. Indeed, we may wonder why some promising pedagogical approaches do not succeed and why others are only successful in certain implementations.

In this vein, there have been a number of attempts to understand student attitudes as a dynamic of reform. For example, in SCALE-UP implementations, students felt that although the class was worthwhile, they noted that they could receive the same grade with less effort in a lecture section.[2] A few survey tools have also been created to specifically to look at expectations in physics. One of the first of these was the Maryland Physics Expectations Survey (MPEX).[9,10] Its main finding was that student expectations of what science is and what goes on in science courses deteriorate (away from those desired by experts) after completing introductory coursework. Halloun and Hestenes' Views About Sciences Survey (VASS) (Ref. [11]) found that student perspectives correlated significantly with physics achievement and postulated that these views may be major determinants of what is learned. The most recent attitude survey used for measuring physics expectations is the CLASS, Colorado

Learning Attitudes about Science Survey,[12] whose findings replicate the earlier work in that most teaching practices caused drops in student scores. In any event, existing instruments mostly probed student attitudes toward science by comparing them with attitudes of expert scientists. However, when probing student expectations of instruction itself, there are no such "expert" attitudes: we have no reason to believe that there is a preferred set of pedagogical expectations that one has upon entering a physics classroom. In fact, we expect that students' expectations regarding the course will vary somewhat depending on their individual experiences. Moreover, existing instruments do not address the instructor's efforts to calibrate the students' pedagogical expectations or the manner in which those expectations are met throughout the semester.

We propose that a piece of the reform puzzle that has been missing is this interaction between what students expect from a college class and what they experience in a reformed pedagogy class. We utilized the framework of expectancy violations to explore this interaction. Expectancy violation uses the assumption that people have expectations about the interactions they participate in and specifically about the behavior of other people in those situations, an assumption that is not very different from those taken in literature about framing;[13] this set of expectations affects what people notice and how they act in those situations. If those anticipated behaviors do not occur (or alternative behaviors occur instead), this is called a violation of expectations. When expectations are violated, that action may be evaluated either positively or negatively.[3,4] We have all experienced both extremes; for example, a stranger holding open a door may represent a positive violation, while a stranger standing too close on the bus may represent a negative violation.

The positive or negative value people attribute to expectancy violations varies in terms of strength; for example, a violation may be seen as slightly negative or strongly positive. When a violation occurs, the person whose expectations have been violated will make an assessment based partially on communicator characteristics (e.g., gender, personality) and relationship characteristics (e.g., is the communicator a close friend or a stranger?). Such characteristics affect both the interpretation and the evaluation of the violation. For example, if an instructor and student are speaking one-on-one and the instructor takes a step closer to the student, the student may interpret that action either as an aggressive act or as a sign of affection.

A person's expectancies include a range of acceptable (and unacceptable) behaviors; this range is influenced by considerations such as knowledge of the other person. Actions that fall outside of this range are identified as violations.[14] When the other person is unknown—as is typically the case when students start a new class—cultural norms and contextual cues are the primary sources of expectancies. Therefore, our focus is primarily on classroom and instructor factors that may or may not deviate from what students expect based on the previous classes they have taken.

While students and instructors both enter the classroom with expectations about how the class will proceed, those expectations may not be the same.[15] Furthermore, expectations held by both instructors and students can impact behaviors in and perceptions of events in the classroom.[16] When students' expectancies are negatively violated, the violation can have negative implications for how students view the class and the instructor. Indeed, the strength of students' initial expectations of nonverbal immediacy (behaviors intended to reduce the psychological distance between communicators, e.g., eye contact), coupled with the strength of the violations of those expectations, accounted for more of the variance in learning performance than did simply assessing instructor immediacy.[17] As instructors' violations of expectations (such as displaying anger in the classroom) increased, students displayed more dissatisfaction toward the instructor and the class content.[18]

Students' initial set of expectations, which are based on previous experiences, are influential in both their perceptions of the class as well as their learning. For example, one study showed that students' expectations of future success in chemistry classes (and therefore students' willingness to enroll in future chemistry classes) were affected by past negative or positive experiences.[19] Other research has demonstrated that when students' expectations are met or the experience is more positive than expected, they report more learning and satisfaction than when expectations are not met.[20,21] Furthermore, when students' expectations were violated (e.g., professors were less clear than students expected), their motivation to learn and perceptions of their own learning were reduced.[22,23]

In summary, students' previous experiences in classrooms inform their expectations. These expectations and the extent to which they are met affect students' satisfaction, motivation, and perhaps even their ability to learn. While little work has been done that explores the connection between learning gains on standardized measures and expectancy violations, existing literature does suggest that we should consider the implications of not meeting students' expectations. Indeed, in light of these results, we suspect that investigating the complex relationship between students' expectations and their perceived experiences in reformed physics classes is an important step toward understanding what makes for successful implementations of pedagogical reforms that have demonstrated the potential to produce strong learning gains.

Toward that end, we approached this research with several goals. First, we sought to uncover students' expectations before they attended the first class meeting of SCALE-UP. Second, we wanted to examine how those initial expectations shifted when students were oriented toward the reformed pedagogy. We also wanted to compare what students reported experiencing during the semester with what they had expected. Finally, we sought affective responses to the reformed pedagogy in order to gauge whether the students had an overall positive or negative perception of SCALE-UP after the course was over.

III. STUDY DESIGN

To meet these goals, we created a measure called the Pedagogical Expectation Violation Assessment (PEVA), which is a structure for investigating students' initial expectations about a course and assessing how they are shifted

during an orientation to that course and also the degree to which those expectations are met. The particular questions that we used in this implementation of the PEVA are specific to SCALE-UP, but the method of analysis and interpretation are general to any course of interest. Notably, one could redesign the questions to use the PEVA for traditional lecture sections, laboratories, or recitations, but the PEVA is perhaps most useful as a tool to learn about students' reactions to pedagogical reform that is new to them. In this section, we will discuss the SCALE-UP classes we investigated and the design of the PEVA and supporting components that comprised the surveys we used.

A. Environment

SCALE-UP is a pedagogical reform that was developed at North Carolina State University in the 1990s, initially for large introductory physics classes. This pedagogy emphasizes collaborative learning and is organized around maximizing both the variety and quality of interactions in the classroom. Given the myriad of unique features that comprise the SCALE-UP approach, an examination of this setting provides an interesting insight into the effects of such a reform. Features of a SCALE-UP classroom include round tables, each of which seats 9 students in groups of three, an instructor station in the center of the room, and name tags for each student. Depending on enrollment, which can be 100 or more, one or more teaching assistants are present in the classroom during regular class time, and they share some of the teaching responsibilities with the lead instructor. While implementations vary among institutions, ones that take full advantage of the classroom generally require students to work together on problems, make informal presentations to the class, and participate in an ongoing dialog with the instructors. The instructor and teaching assistants also directly intervene with individual students, asking them questions as they work and providing assistance when required. SCALE-UP thus provides a variety of modes of communication: between students, from students to the class, between instructor and class as a whole, and between individual students and instructors.[24] These additional avenues of communication constitute a significant dimension of SCALE-UP's success, and we were particularly interested in this increase in communication opportunities, given the role that informal communication plays in learning[25] and the difference in the types of communication between SCALE-UP and traditional courses.[24]

However, with all of the advantages of SCALE-UP comes a challenge to implementation that is not unique among pedagogical reforms: students may not expect the kind of work they will be required to do during class time. Our pilot study indicated that many students enrolling in a physics class expect to passively take notes from a lecture in an amphitheater-style classroom, and they expect to attend a laboratory section at another time and in another place.[26] An instructor teaching a pedagogically reformed class, especially one with such a visibly unorthodox classroom, is likely to experience a room full of hesitant faces on the first day of class. This hesitation is emphasized when students are not fully aware of the nature of the course in which they have enrolled. For example, after a single SCALE-UP class meeting, one of the authors (J.G.) received an email asking whether he was "ever going to lecture and actually teach anything" by a student who subsequently dropped the course in favor of a traditional section.

To minimize the degree of this hesitation and resistance to change later in the semester, instructors of pedagogically reformed classes must prepare the students by realigning the students' expectations toward the instructor's design for the course as soon as possible. Because the implementation of SCALE-UP varies between institutions (and even between instructors), there is no generic script that is followed to introduce the students to the reformed class. The SCALE-UP website and wiki[27] provide example notes for first class meetings. Some activities that have been used include collecting and discussing rumors about what students previously heard about SCALE-UP, forming post-hoc groups and running a "game of science" activity,[28] and defining a coordinate system and origin in the classroom and having each table of students determine the location of their table by using meter sticks. These orientation activities do more than simply review the syllabus, as they actually engage the students in sample tasks that demonstrate the instructor's awareness of and intent to fully utilize the SCALE-UP room.

B. Survey

Because of this orientation of the students, implementation of the PEVA requires three surveys: Survey No. 1 obtains students' initial expectations about the class, Survey No. 2 obtains students' expectations after being introduced to the reformed pedagogy, and Survey No. 3 obtains students' perceived experiences after most of the semester was over. It is important to note that the experiences the students report may not precisely reflect an objective observation of what occurs in the classroom, although one of the authors (A.G.) performed such an observation of SCALE-UP and found through interviews that student perceptions are not far from her objective observations.[24] Nonetheless, our interest was in what students felt they had experienced, because it is those perceptions which help determine the students' affect.[20]

While Survey No. 1 would ideally have been given when the students first enrolled in the course, we gave it at the same time as Survey No. 2. Those two surveys were given during the first week, either in class or online, depending on the instructor's ability to provide class time for the survey. Survey No. 3 was given during the final two weeks of the course.

In addition to the PEVA, we included an open-ended question as Part 1 on Survey No. 1, and three open-ended questions as Part 1 on Survey No. 3. These questions were about what kinds of communication they expected or experienced in SCALE-UP, and were used to help establish the validity of our chosen PEVA questions. We also included an established measure of communication violation attitude as Part 2 on Survey No. 2 and Part 3 on Survey No. 3. Finally, Part 4 of Survey No. 3 is a short list of affective questions about the pedagogy itself. All three surveys are included as an appendix to this paper.

1. Pedagogical Expectancy Violation Assessment (PEVA)

To create a set of questions for use in the PEVA, we first identified thirteen salient features of a generic SCALE-UP physics class. Two additional items: "computer programming or simulations" and "to memorize equations" were included because of their relevance specifically to the classes at North Carolina State University. These fifteen items, which are listed in the appendix, were identified through an examination of previous literature on SCALE-UP and reflections on experiences in the SCALE-UP classroom by all three authors (two as instructors, one as an observer), and refined after a pilot study.[26] We did not expect that the thirteen items would span the space of reforms in SCALE-UP (for example, a SCALE-UP instructor who validated the relevance of our list noted that we omitted a feature: "the use of technology to provide real time feedback to the instructor"). However, we chose these items because of our particular needs. For example, we wanted to know if our orientation to the class made students expect the level of group work that we required of them.

In order to check the internal consistency of our PEVA items, we used the coefficient of reliability. We checked the PEVA items for each survey at each school; Cronbach's alphas ranged from 0.62 to 0.83, which all were within the acceptable ranges.[30] We did not use a factor analysis to establish a few specific features of SCALE-UP, or pedagogical reform in general, because reform efforts vary greatly, even among SCALE-UP implementers.

On Survey No. 1, students were asked to indicate how often they expected that they would experience each of the fifteen activities when they first enrolled in the class. On Survey No. 2, they were asked how often they expected to experience those same activities in their particular SCALE-UP course after orientation to the SCALE-UP pedagogy. On Survey No. 3, the students were asked how often they experienced those activities. The student responses were on a 7-point scale, ranging from "very infrequently" to "very frequently." Again, we were interested in the students' perceptions rather than an objective account of the students' activities that might be obtained through observing the classes themselves. This decision follows in the steps of previous expectancy violation and experience research.[20]

Our intention for using the PEVA was to show how much students' expectations shifted after being introduced to SCALE-UP, and whether those expectations were met by their perceived experiences. Unfortunately, because we did not have access to the students before they first attended class, we were unable to ask for their expectations when they first enrolled. Therefore, when students took Survey No. 1, they were asked to recall previous expectations rather than reporting how they currently felt. Seeing the unorthodox classroom might therefore have introduced some bias in their results. We were unable to eliminate this potential bias; however, informal discussions and interviews with students lend credibility to our results.[24]

2. Communication violation attitude measure

To assess how students evaluated the communication they were told to expect, we used items based on previous work with classroom expectations[29] to gauge students' attitudes. Students were asked to respond on a 7-point bipolar scale to ten pairs of adjectives such as expected or unexpected and appropriate or inappropriate. Several items were reversed on the measure when given to students; these items were subsequently adjusted so that a higher number was reflective of students having a more positive affect for the communication they were told to expect (in Survey No. 2) or that they experienced (in Survey No. 3). All ten items were combined into a single score, as appropriate for this particular measure. We used this score on the two surveys to determine whether students perceived the nature of the communication in SCALE-UP positively and whether their attitudes regarding that communication shifted over the course of the semester.

3. Affective questions

In part 4 on Survey No. 3, we included six affective questions regarding SCALE-UP (see appendix). Participants were asked to indicate their agreement with these statements on a 7-point scale ranging from "strongly disagree" to "strongly agree." We used the results from these questions to speak to the overall affective success of the pedagogy, which was then compared to the results from the PEVA. Further affective questions could have been added to determine whether students perceived each individual activity on the PEVA as positive or negative; however, we deferred that assessment for future work.

C. Implementation

During the Fall 2008 semester, the surveys were administered to four different SCALE-UP classes at three different institutions: North Carolina State University (NC State), Old Dominion University (ODU), and the University of Pittsburgh (Pitt). We chose ODU and Pitt because they used relatively new implementations of SCALE-UP that had been directly adapted from NC State, with significant contributions from two of the authors (J.G. and R.B.). Including these institutions allowed us to compare implementations between various student populations. Because populations varied, it is reasonable to expect that students may have different expectations of an introductory physics course; using the PEVA would allow us to identify such differences and could provide feedback to instructors for adjusting their orientation to reflect their particular students.

The NC State course was a second-semester course, and some of the students had previously been enrolled in a SCALE-UP course; however, the structure of the course was not advertised well, so we had reason to believe that many of the students in the class would not be expecting reformed pedagogy. ODU, which was using SCALE-UP for the first time, was running two sections of SCALE-UP for first semester physics, taught by two different professors. Class H was an honors course while class G consisted of a general population. One of the authors (J.G.) assisted at ODU as a consultant during the first week (but not the first day), performing the role of a teaching assistant during class and advising the teaching assistants and instructors. Pitt ran one section of the class. It was the second year that they used

SCALE-UP, but the previous year had been honors students while this year consisted of a general population. The curricula differed between the schools as well; NC State and Pitt both used *Matter & Interactions 2nd Edition*,[31] while ODU used *Understanding Physics*.[32]

Students were told that Surveys No. 1 and No. 2 were about their expectations of the course, and that Survey No. 3 was about their experiences. At NC State and ODU, the surveys were administered during class time. At NC State, the surveys were online and most of the students in the class (81) completed Surveys No. 1 and No. 2, and 61 completed Survey No. 3. At ODU, the surveys were given on paper, and nearly all students completed Surveys No. 1 and No. 2 (25 in class H and 36 in class G), while slightly fewer completed Survey No. 3 (22 in class H and 28 in class G). At Pitt, the surveys were online and students were asked to complete them outside of class. There were considerably fewer responses at Pitt, with only 16 students, less than half, completing Surveys No. 1 and No. 2 and only 8 completing Survey No. 3. No students received compensation or course credit for completing the surveys, which were completely anonymous. In total, 158 out of approximately 200 students took the first survey and 119 took the second. At Pitt, the low response rate created a potential bias in the responses, as many students did not complete it. We do not know the reasons for this lack of participation, but we suggest that it's due to the online, outside-of-class implementation of the surveys.

Below we present the results of our surveys as a starting point to discuss not only the success of the SCALE-UP implementations, but also the students' expectation shifts and perceived experiences. Such results point to the potential utility of this information to allow instructors to meet their goals. Additionally, the two different classes at ODU will be compared to demonstrate how different implementations of the same pedagogy, with different student populations, can address student expectations and experiences.

IV. RESULTS

A. PEVA results

The mean score for each question on the PEVA is displayed on Figs. 1 and 2, grouped by school, for each of the three surveys: initial expectations, expectations after orientation to the SCALE-UP pedagogy, and their experiences at the end of the semester. The error bars indicate the standard error for each of those means. We used the Wilcoxon Mann-Whitney nonparametric test[33] to compare the means and distributions where applicable. We did not track individual students, so when we speak of a shift in expectations, we refer to the expectations of the class as a whole. To whit, even when the mean undergoes a significant shift, there may be individual students who do not experience the shift (or may even experience that shift in the opposite direction). The differences between the two classes at ODU will be explored in a later section; for the purposes of this analysis, we grouped those classes together unless otherwise noted.

1. Initial expectations

When students enroll in a physics course, they expect to participate in certain activities or environments and not in others. Of the fifteen items proposed in our PEVA, students from all three schools reported expecting, with a mean score of 5.0 (somewhat frequently) or more: lecture, doing required reading, an amphitheater classroom, a separate laboratory section, that missed classes would be harmful to their learning, and memorizing equations. At Pitt, students also expected a grading curve. These aspects are common in traditional introductory physics courses; additionally, experiences from previous college courses probably contributed to these expectations. Worth noting is that students expected that they would do the required reading and that they understood the value of attending class.

Students from all three schools reported not expecting to explain their work to the class, with a mean score under 3.0 (somewhat infrequently). Indeed, it is uncommon in an introductory college physics course for students to be asked to make even informal presentations of their work. The means on the remaining eight items ranged between 3.0 and 5.0 among the schools and varied somewhat between them due to their specific environments.

2. Shifts in expectations

Most items on our PEVA characterized a significant shift of the students' expectations from when they enrolled for the course to the completion of their class orientation. In all cases where a shift in expectation occurred, that shift was in the direction of what the instructor wanted students to expect (see Table I).

Two items carried no noticeable shift at more than one school. "Missed classes would be harmful to my learning" did not change significantly at NC State or Pitt, but it did change at ODU, with students expecting missed classes to be more frequently harmful to their learning. "Doing required reading" did not change at ODU or Pitt, but at NC State, students came to expect more reading. Note that for both of these items, the orientation moved the students to approximately the same level of expectation regardless of the institution or the initial expectations of the students.

A third item, "memorizing equations," did not shift at Pitt but did at both NC State and ODU, with students reducing their expectations for memorization. Two further items may have shifted at Pitt, but the significance was borderline ($p = .05$): expectations for both "lecture" and "grading curve" may have been reduced. Again, at both NC State and ODU, students report lowered expectations after orientation.

3. Comparing expectations and experiences

Ideally, an instructor will be able to shift students' expectations during the orientation to prepare the students adequately for the course. However, this is not always the case. We note four distinct patterns that represent how the students' expectations were shifted in comparison with their reported experiences for specific items. In Fig. 3, we display prototypical examples of the four patterns: success, undershooting, overshooting, and failure.

a. Success. The signature of a success for a particular item was a shift in expectations in the desired direction thanks to the orientation, followed by a reported experience
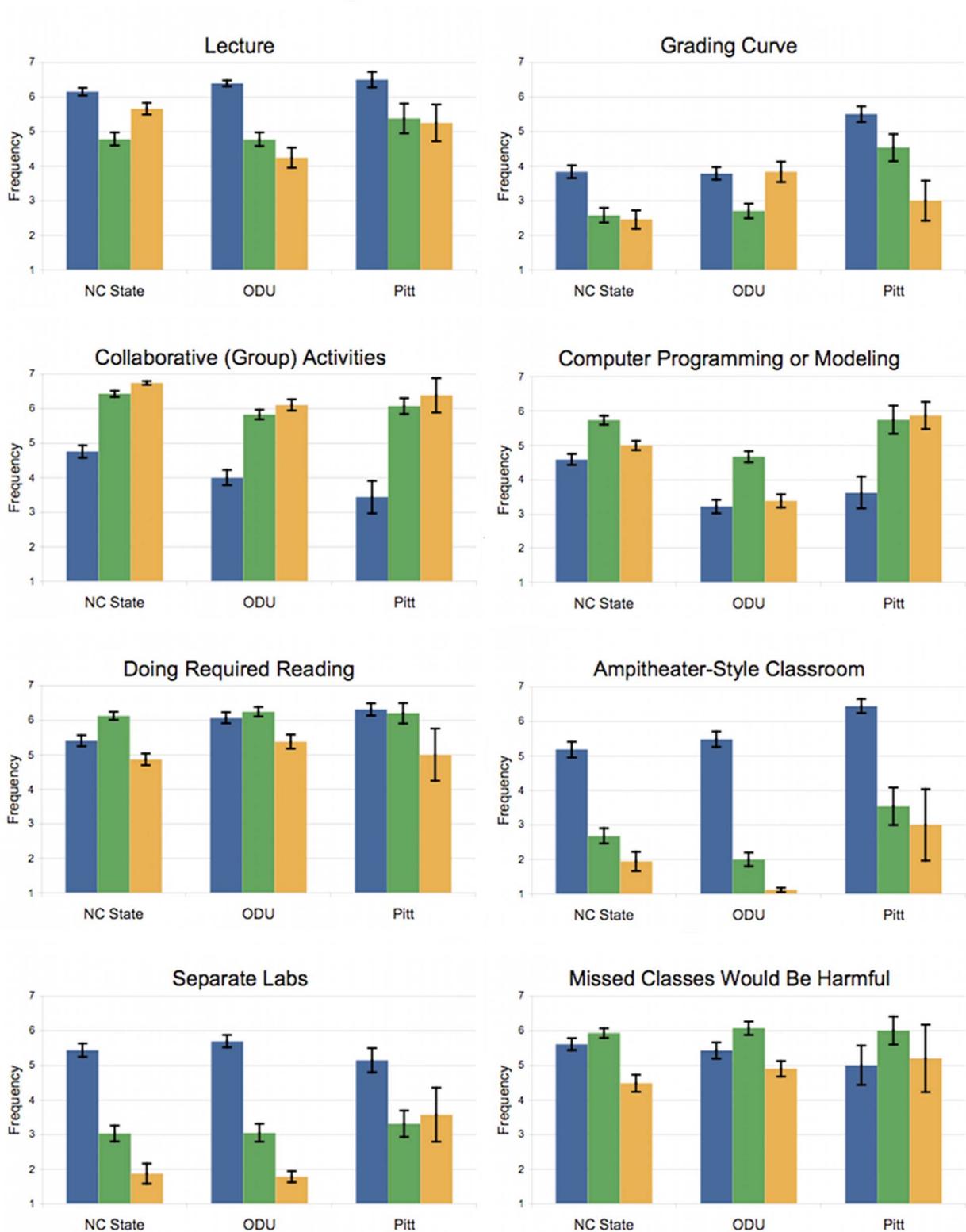
FIG. 1. (Color) The extent to which students at each school initially expected (blue, first column), expected after orientation (green, second column), and experienced (orange, third column) items 1–8 on the PEVA.

that was the same as those expectations. Ideally, all items on a PEVA would fit into this category, as a "success" pattern indicates that the implementation of the class matched what

the students expected after their orientation. Four items at NC State and six items at ODU can be considered successes. Two items were "to interact with my instructor during class
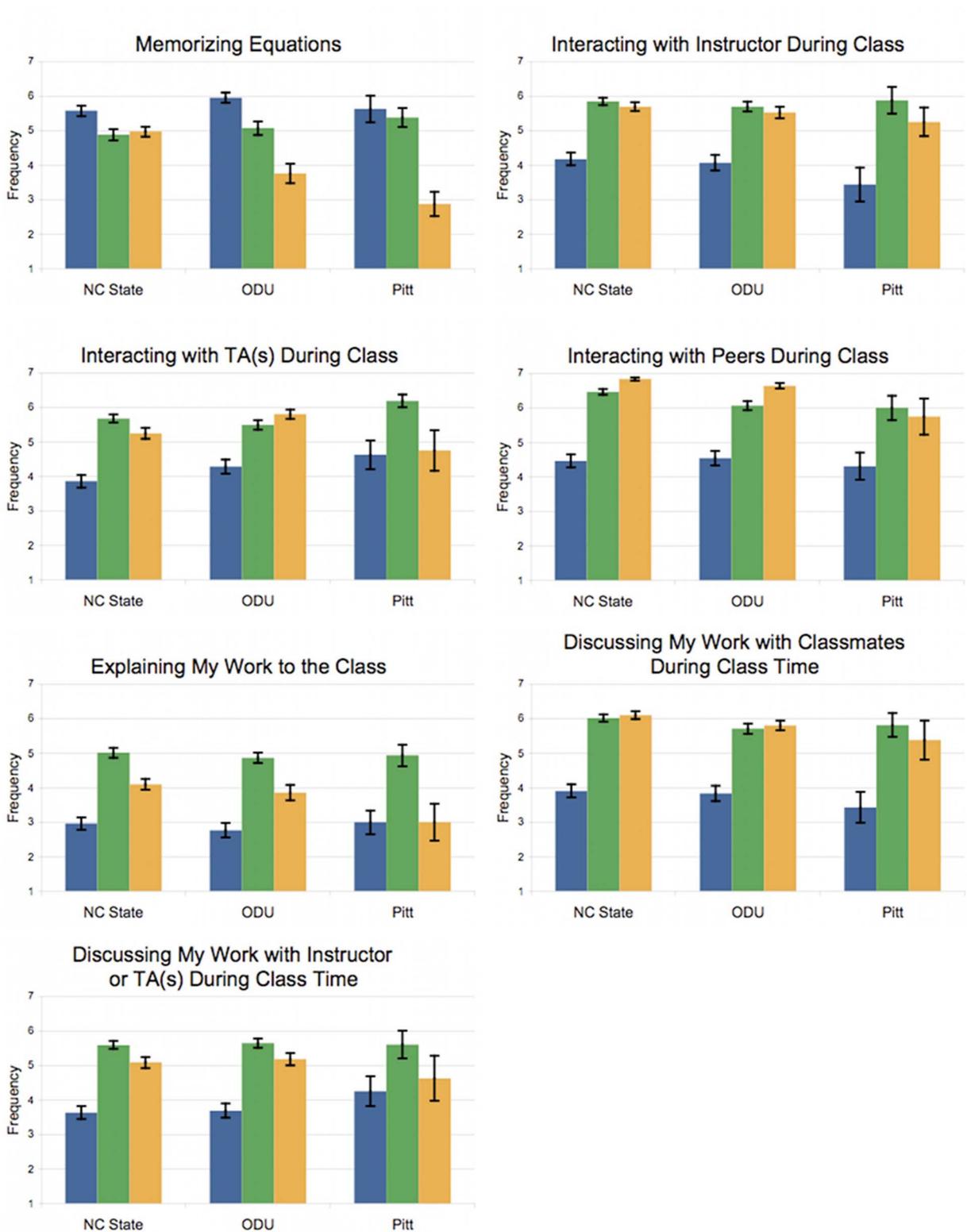
FIG. 2. (Color) The extent to which students at each school initially expected (blue, first column), expected after orientation (green, second column), and experienced (orange, third column) items 9–15 on the PEVA.

time" and "to discuss my work with my classmates during class time," both of which emphasize the increase in class-time communication for the reformed pedagogy. This result is particularly optimistic because we feel that the use of multiple communication methods is integral to the success of SCALE-UP.

TABLE I. Sample items demonstrating shifts in student expectations.

Item	School	Initial expectation	Shifted expectation	p-value
Missed classes would be harmful to my learning	NC State	5.6	5.9	
	ODU	5.1	6.1	$<.05$
	Pitt	5.0	6.0	
Doing required reading	NC State	5.4	6.1	$<.001$
	ODU	6.1	6.2	
	Pitt	6.3	6.2	
Memorizing equations	Pitt	5.6	5.4	
Lecture	Pitt	6.5	5.4	0.05
Grading curve	Pitt	5.5	4.5	0.05

b. Undershooting. In this category, the students' expectations shift in the appropriate direction, but not all the way to

meet their experiences in the class. In other words, the orientation to the class was not sufficient for the students to be prepared for these aspects of the class, which is indicative of the need for more training in these areas. Four items at NC State and four at ODU fell into this category, with an overlap on three items, which are shown in Table II. Particularly surprising is the students' resistance to shift their expectations regarding the classroom more fully after orientation, considering that in the case of NC State and ODU, they took the survey in that classroom.

One item at Pitt was possibly in this category, but the significance of a shift in expectations was borderline ($p = .05$). Students shifted their expectations after being told that there was no grading curve in the course, but they reported less frequent experiences of a grading curve than expected. This finding may be due to simple disbelief on the part of the students; if their previous science classes used a grading curve, then it is reasonable for them to expect this class to do the same even when the instructor informed them otherwise. Interestingly, neither ODU nor NC State reported the same pattern with the grading curve item, indicating a
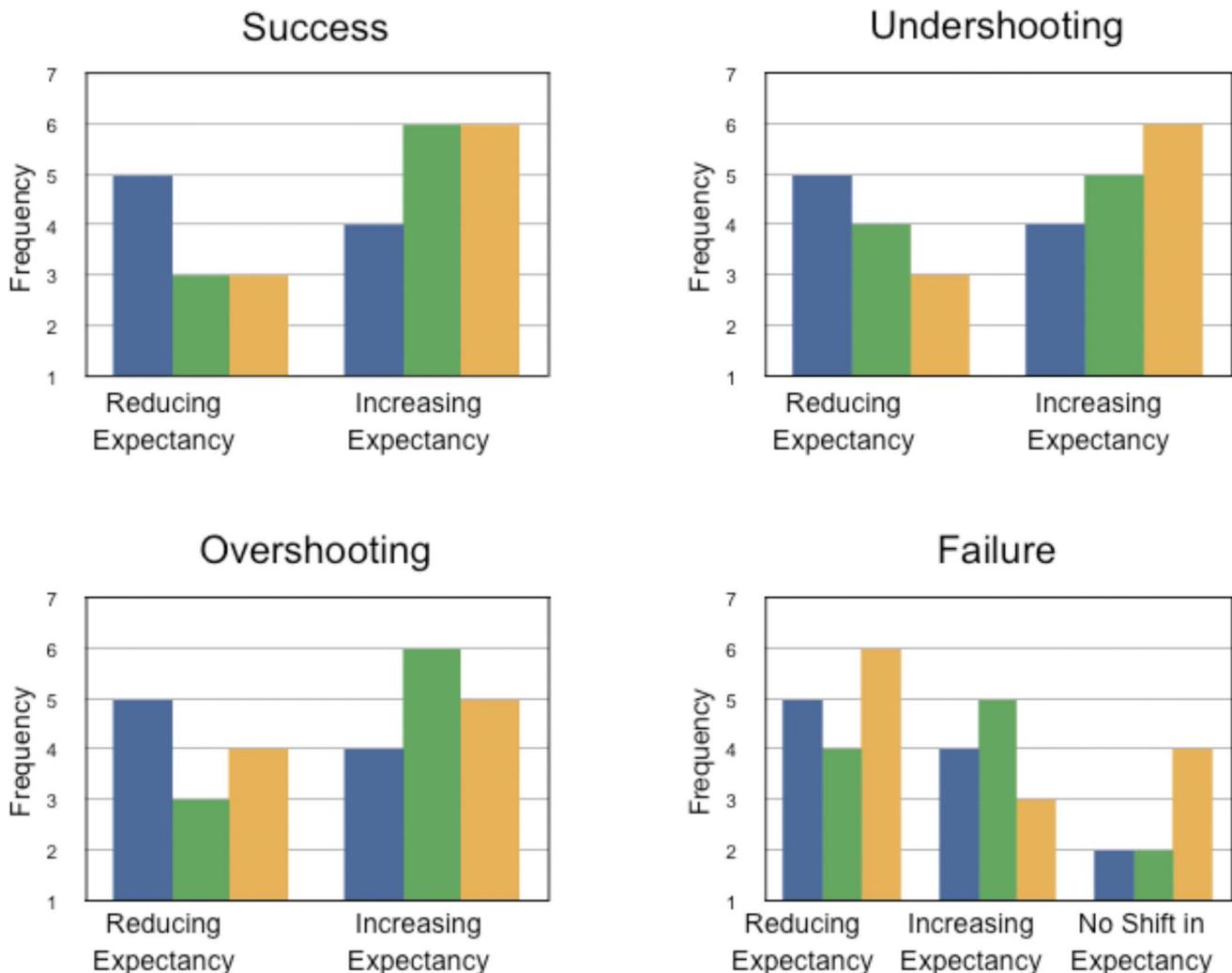

FIG. 3. (Color) Prototypical patterns of students' initial expectations (blue, first column), expectations after orientation (green, second column), and experiences (orange, third column).

TABLE II. Sample items fitting an "undershooting" profile.

Item	School	Initial expectation	Shifted expectation	p-value (initial to shifted)	Reported experience	p value (shifted to reported)
Laboratories separate from the rest of class	NC State	5.4	3.0	<.001	1.9	<.01
	ODU	5.7	3.0	<.001	1.8	<.01
Interacting with my peers during class time	NC State	4.5	6.5	<.001	6.8	<.01
	ODU	4.5	6.1	<.001	6.6	<.01
Amphitheater-style room	NC State	5.2	2.7	<.001	1.9	<.05
	ODU	5.5	2.0	<.001	1.1	<.001
Presence of grading curve	Pitt	5.5	4.5	0.05	3.0	<.05

possible difference in the student populations between the three schools.

c. Overshooting. On an overshoot, students' expectations changed in the desired direction after orientation, but their experiences did not meet either their original or their realigned expectations. This pattern is indicative of an instructor promoting an aspect of the pedagogy that is not ultimately provided as much as the students were expecting. Some examples are detailed in Table III.

Five items at NC State were overshoots. One of those items, "explaining my work to the class," was apparently true at ODU as well. However, this result is due to the differences between the classes at ODU. This item was not an overshoot in class H, because the students' experiences matched the expectations that they reportedly had when they enrolled, making it a failure. It was also not an overshoot in class G; in fact, for that class it was a success. This situation will be explored more in the comparison between classes at ODU. There were no instances of overshooting at ODU.

Also of note is a unique difference between NC State and ODU. NC State reported overshooting the lecture component of the course, while ODU reported lecture as a success. Both schools shifted the students' expectations to the same mean; however, ODU met those expectations while NC State produced more lecture than the students expected after orientation. While the students may vary in their interpretation of "lecture," it is clear they felt that their instructor did not

lecture as infrequently as he had prepared his class to anticipate.

d. Failure. A failure is similar to an overshoot in that the students' reported experiences did not match their initial expectations after they were oriented to the class. However, in a failure, the mismatch is more dramatic: not only do they not meet the students' shifted expectations, they are lower than or equal to their initial expectations, indicating that any shift in expectations was in opposition to the actual experience of the class. Examples of this are shown in Table IV.

For instance, students at all three schools expected that missing classes would be harmful to their learning. The training early in the semester either left those expectations alone or emphasized them. However, the students did not feel that the class meetings were as valuable as they had expected them to be: a message to instructors to carefully consider how class time is being used. Note that while the pattern at Pitt appears to be a failure, the lack of significance prevents us from making a clear assessment of the situation.

At ODU, students were prepared to have computer modeling and/or programming as an integral component of the course. However, the instructors found that their curriculum choice was not particularly conducive to maintaining this component, so they reduced the number of computer-based assignments from what they had intended. Therefore, their implementation of the course did not match what they had advertised, a difference that was detected by our PEVA. This

TABLE III. Sample items fitting an "overshooting" profile. Items marked with an asterisk are not overshoots but are provided for comparison purposes.

Item	School	Initial expectation	Shifted expectation	p-value (initial to shifted)	Reported experience	p value (shifted to reported)	Notes
Explaining my work to the class	NC State	3.0	5.0	<.001	4.1	<.001	
	ODU	2.8	4.9	<.001	3.9	<.001	
	ODU−H*	3.0	5.1	<.001	3.5	<.001	Failure
	ODU−G*	2.6	4.7	<.001	4.2		Success
Lecture	NC State	6.1	4.8	<.001	5.7	<.01	
	ODU*	6.4	4.8	<.001	4.2		Success

TABLE IV. Sample items fitting a "failure" profile.

Item	School	Initial expectation	Shifted expectation	*p*-value (initial to shifted)	Reported experience	*p*-value (shifted to reported)	*p*-value (initial to reported)
Missed classes would be harmful to my learning	NC State	5.6	5.9		4.5	<.001	<.001
	ODU	5.4	6.1	<.05	4.9	<.001	<.05
	Pitt	5.0	6.0		5.2		
Computer programming	NC State	4.6	5.7	<.001	5.0	<.001	
	ODU	3.2	4.7	<.001	3.4	<.001	

example lends credibility to the argument that our PEVA is capable of detecting instances of an implementation not matching the stated course goals.

B. Communication violation attitude measure results

Overall, students reacted positively to the communication they were told to expect in the class at the start of the semester and to the communication they experienced over the course of the semester (see Table V). None of the schools demonstrated a shift in students' perceptions of the communication from the start to the end of the semester. Because the communication violation attitude survey yields a single score between 1 (very negative) and 7 (very positive),[29] we do not report responses to individual items.

The results from the communication violation attitude measure are meaningful in the larger context of this study. The results indicate that even immediately after orientation—when expectations are first violated—students reported that the communication that they were going to engage in during the semester was overall positive. The shift in expectations seen in our PEVA further adds to this understanding. Students came into the class expecting mostly one-way communication: lecture and little collaboration or one-on-one interaction. However, these expectations were clearly violated, as seen in all of the PEVA items that dealt with communication (e.g., lecture, interactions). We interpret the size of this shift, along with the overall positive nature of the initial responses about the communication, to mean that although expectations were violated, students saw such communication as a positive aspect of the class. This result sheds hopeful light on attempts to institute a different, more richly communicative pedagogy than students may expect. Because students initially viewed the communication they were told to expect positively (despite its deviance from their expectations before the class), it is not surprising that they did not

view the communication any more positively at the end of the course. Indeed, the mere fact that no schools showed a negative slide in attitudes regarding communication is encouraging. Despite the initial violation of students' expectations, students maintained positive affect toward the communication they experienced throughout the semester.

C. Affective results

The combined results from the three schools indicate an overall positive affect toward SCALE-UP (see Table VI). The students at NC State were more positive than those at ODU ($p < .05$), but the difference between NC State and Pitt was not significant ($p = .078$).

A particularly encouraging note for instructors considering SCALE-UP or other reformed pedagogies is the overall positive affect these students had toward the communication in the class. Given the strength of the violation of students' initial expectations, such positive affect was by no means guaranteed; students could easily have become upset over the violation, particularly in light of research demonstrating a close connection between expectations and evaluations of the instructor or class.[18,34] The overall positive perception of what students were told to expect (as seen in the communication violation attitude measure) perhaps provides the best explanation for their affect. Students' expectations were violated, but they viewed this violation positively. That positive affect may have carried through to the end of the semester.

As a whole, students' experiences matched their expectations after their orientation to the class far better than before the class began, a result which provides another explanation for the positive affect. Previous research showed that when instructors' behaviors match expectations, students evaluate the instructor positively.[35] In this case, even though students experienced a different class than they expected when enrolling in the course, they were quickly oriented to the new pedagogy. Because all three schools matched those new ex-

TABLE V. Communication violation attitude measure results.

School	Communication you have been told to expect	Communication you have experienced	*p*-value
NC State	5.0	5.2	
ODU	4.8	4.9	
Pitt	4.7	4.9	

TABLE VI. Affective results. A score of 7 is "strongly agree."

Item	School	Score
The SCALE-UP environment is a useful style of teaching and learning.	Combined	5.9
I would have learned physics better in a more traditional setting than SCALE-UP.	Combined	3.2
The SCALE-UP environment is inappropriate for college classes.	Combined	2.0
The style of this course helped me learn physics.	Combined	5.5
Courses in other departments should use a SCALE-UP environment.	Combined	5.2
SCALE-UP is not for me.	Combined	2.6
Overall (7=strongly positive)	NC State	5.8
	ODU	5.1
	Pitt	5.1
	Combined	5.5

pectations fairly well (especially on the communication-based aspects of the course), the students may not have felt violated. Therefore, properly preparing students for what to expect in a class—even if it goes against their previous experiences—and then fulfilling those expectations helps to alleviate negative effects.

D. Between classes at ODU

As previously noted, ODU offers an interesting case study of two SCALE-UP classes offered at the same institution. The two classes' expectations when enrolling in the course differed only slightly. Class G, consisting of general-population students, expected lecture and an amphitheater classroom more than class H, which was an honors class. After both classes met, their difference concerning classroom type was quickly reconciled. However, their difference in expectations about lecture was accentuated; after one day,

TABLE VII. Differences in expectations at ODU.

Item	Class	Initial expectation	Shifted expectation
Amphitheater-style room	H	4.8	1.8
	G	6.0	2.1
	p-value	<.05	
Lecture	H	6.2	4.6
	G	6.5	5.4
	p-value	<.05	<.01
Memorize equations	H	5.9	4.6
	G	6.0	5.4
	p-value		<.05
Grading Curve	H	3.8	3.4
	G	3.8	2.2
	p-value		<.01

TABLE VIII. Differences in experiences at ODU.

Item	Class	Experiences
Computer modeling and/or programming	H	4.0
	G	2.9
	p-value	<.01
Memorize equations	H	2.8
	G	4.5
	p-value	<.01

class G expected more lecture than did class H even though both classes reported a lowered expectation of lecture (see Table VII).

After orientation, two differences emerged in the classes: class G expected to have a grading curve less, but to memorize equations marginally more, than class H. The grading curve might be explained by the expectations of the honors class being more resistant to change because of their previous experiences in honors classes. Similarly, students in the general section may have a similar resistance in terms of memorizing equations. These two differences are probably best explained by differences in the students rather than differences in the instructors.

At the end of the semester, the two classes reported very similar experiences with only two exceptions (see Table VIII). The honors class reported more computer modeling and/or programming and less memorization of equations than did the general population class. In both classes students were permitted to bring a "crib sheet" of equations to the examination, so the difference in their experiences may be due to differences in the class populations. The instructor of the honors class reported implementing more computer tasks than the other. It is worth noting that class G reported a failure pattern, whereas class H reported an overshoot. In both cases, the orientation shifted the students' expectations to more computer work, but the experiences during the semester failed to meet those shifted expectations in both classes (Class H: $M=4.7$ vs 4.0, $p<.05$; Class G: $M=4.6$ vs 2.9, $p<.001$).

Indeed, the two classes saw some differences in the patterns revealed by the PEVA. In four cases, one class observed a success pattern, while the other exhibited either an overshoot or undershoot. For three of those items: "labs separate from the rest of class," "to memorize equations," and "interact with peers," a combination of the two classes yielded an undershooting pattern. For example, class H showed a success for "interact with peers," while class G showed an undershoot. Both classes reported the same experience; therefore, class H was better prepared for the course than class G. In the fourth case, "discussing work with an instructor or TA during class," the pattern of the combined data was a success even though class H showed an overshoot. Some variation between the classes should be expected; while instructors may be able to learn from one another how to better prepare their students or meet their expectations, instances where one class succeeds and the other overshoots or undershoots are not a cause for concern.

In contrast to those four items, the item "explain work to the class" was a failure in class H but a success in class G.

The instructors should explore this difference between the classes; for example, one of them may have made the conscious decision to either emphasize student presentations or to de-emphasize them. However, there was not a significant difference in students' reported experiences between the classes ($M = 3.5$ vs 4.2, $p = .22$), and there are no significant differences between the classes' expectations (initial: $M = 3.0$ vs 2.6, $p = .41$; after orientation: $M = 5.1$ vs 4.7, $p = .14$). Yet, the classes exhibited different patterns in comparing the expectations with experiences. It is worth noting that both instructors at ODU were teaching SCALE-UP for the first time and may have inadvertently biased student expectations because the instructors did not know precisely how the class would shape up; in future implementations, they will be able to adjust the orientation to more accurately represent their course.

On the communication violation attitude measure, the two classes were equally positive about the communication they expected and experienced. Also, no differences were found in the affective measures, and both classes were positive about the SCALE-UP pedagogy.

The differences seen between these two classes offer several lessons regarding pedagogical reform. For example, the instructor of class H, who shifted students' expectations more during orientation, may have advice for the other instructor on how to cause such shifts more effectively. On the other hand, it may be the differences in the class type that are informing potential challenges in preparing different student populations for class. Specifically, it may be more difficult to convince honors students that there will not be a grading curve or to convince general-population students that they will not need to memorize equations. In either case, the PEVA provides feedback about specific adjustments each instructor can make to adapt to the particular classes they are teaching.

Instructors can also learn from comparing the experiences of their students. For example, if students claim to interact with the instructor more in one class than another, one instructor may find that observing the other class with that result in mind reveals some help for increasing those particular interactions.

At ODU, even though there were a few differences between the classes as reported on the PEVA, both classes were successful in terms of affect. Therefore, it is important to remember not to get hung up on small discrepancies between the students' expectations and perceived experiences. It is possible to be successful even if the instructor is unable to make all student experiences precisely match their expectations.

V. GENERAL DISCUSSION AND CONCLUSIONS

A. Discussion

Students entered the SCALE-UP classrooms with clear expectations about what they would encounter: lecture in an amphitheater classroom with some limited interactions with their instructors, TAs, and peers. However, thanks to their orientation to the pedagogy, these expectations were shifted to more fully reflect the SCALE-UP design. At the end of the semester, the students as a whole reported that some of these shifted expectations were met, although many were not. In each class under investigation, students reported positive affect toward the communication they experienced in the class as well as toward the class as a whole. These results are promising for instructors considering pedagogical reform by adding to the depth and breadth of our understanding of the effects of such reforms. Based on this research, we offer three main take-away lessons as well as some suggestions for extension of this work.

First, this investigation demonstrates the usefulness of examining students' expectations and experiences, particularly when implementing pedagogical changes. Using the methods outlined here, it is possible to gain a greater understanding of students' initial expectations for a class and then to shift those expectations to more closely align with your instructional goals. Subsequently, the course that you create for your students during the semester can be aimed at meeting those expectations. Based on our results, we believe that even if a large shift is necessary during orientation, meeting or exceeding those newly created expectations will help alleviate negative effects. However, further research needs to be done to establish the relationship between student affect and the shift in expectations and experiences the students undergo on particular features of SCALE-UP.

A second lesson is that such measures can be used to examine different classes that use the same pedagogy, as we did here with the data from ODU. Such an investigation provides insight into the strengths and weaknesses of implementations and may point out areas that, with a small amount of tweaking, can more closely meet instructional goals. The use of the PEVA to better understand a class can also provide a starting point for instructors to discuss the connection between expectations and perceived experiences of students, especially when differences arise.

Third, although these three schools have successfully integrated SCALE-UP into classes, there remains room for improvement. In some cases, there was a clear disconnect between students' expectations after orientation and their actual experiences. In other cases, instructors oversold students on the importance of class features, which students then experienced less than they were led to believe would be the case. For example, the NC State class reported experiencing more lecture than they were led to believe they should expect. Such differences serve as examples of where even successful implementations have room to grow and adapt. Regardless, students at all three institutions were positive about the class, and this positive perspective was also reflected in the end-of-semester evaluations.

It is also important to note that the specific items on a PEVA can easily be customized to meet an individual's (or department's) pedagogical goals. For example, although the items on our PEVA specifically highlight salient features of SCALE-UP in a physics context, a PEVA with different items was created for a communication theory course taught by one of the authors (A.G.). In that class, features such as memorizing equations were not relevant and were therefore replaced with instructional goals that were relevant for the intended instruction. If one wishes to create items for a PEVA from scratch, reflection on the unique pedagogical as-

pects of target class may indicate salient features of one's instruction. Then, items can be constructed based on these features. Even a short PEVA containing these items will provide valuable insight into potential disconnects between what students and instructors expect in a class.

B. Limitations and suggestions for future research

In considering the possibilities offered by the use of a PEVA, it is important to note that the results reported in this paper reflect only a few implementations of a specific pedagogy (SCALE-UP). The challenges and insights gained through our implementation should therefore not be interpreted as meaningful for all pedagogical reform; instead, the underlying approach for obtaining these results should be seen as both a recommendation of important issues in pedagogical reforms as well as a means of assessing how well a specific implementation is addressing those issues. As this type of assessment is utilized, each PEVA variation would need to be examined for issues of validity and reliability within its own context. With the limited research that is currently available on students' expectations in reformed classes, possibilities for application and exploration remain largely open to further refinement and adaptation.

One major suggestion for future research is to examine students' attitudes toward individual aspects of a specific course. For example, students may or may not perceive discussion with peers to be positive. Likewise, we recommend asking students to rate their perceived experiences. It is conceivable that a student may have had few, but very positive, interactions with the instructor, or may have had many negative interactions with the instructor. Our measure did not account for such differences, but given the positive outcome of the communication attitude measure (and the lack of a negative slide), along with the positive affect, we assume that such interactions were generally positive, while also not-

ing the increased depth of understanding we could gain from expanding the measure. We would also suggest an investigation into the relationship between violations of student expectations and performance on standardized assessments of physics learning such as the Force Concept Inventory (FCI) (Ref. [36]) or the Brief Electricity and Magnetism Assessment (BEMA).[37]

C. Conclusion

The framework of expectancy violation can help gauge the success of pedagogical reform in addition to providing insight for instructors using any teaching style. By examining students' initial expectations and how those expectations shift (or stay the same) after orientation, it is possible to gain a deeper understanding of how well our students understand what to expect from our classes. Furthermore, by examining students' actual experiences in relation to those expectations, we learn the extent to which we fulfill the goals that we set out for our classes. Such investigations add to the depth and breadth of our understanding of the effect of our instructional choices. Whether we teach students in rows of desks or seated in circles, by lecture or interaction, a further understanding of the success of our instructional choices will only serve to strengthen positive outcomes of instructional reform.

APPENDIX: SURVEY ITEMS

1. SCALE-UP Survey 1

Part 1. When you signed up for this course, what kinds of communication did you expect to witness and participate in during class time?

Part 2. Indicate how often you expected to experience the following when you signed up for an introductory physics course, using the following scale:

1	2	3	4	5	6	7
Very infrequently	Infrequently	Somewhat infrequently	Sometimes	Somewhat frequently	Frequently	Very frequently

1. Lecture.
2. A grading curve.
3. Collaborative (group) discussions.
4. Computer modeling and/or programming.
5. Doing required reading.
6. Amphitheater-style (lecture hall) classroom.
7. Laboratories separate from the rest of class.
8. Missed classes would be harmful to my learning.

9. To memorize equations.
10. To interact with my instructor during class time.
11. To interact with my TA during class time.
12. To interact with my peers during class time.
13. To explain my work to the class.
14. To discuss my work with classmates during class time.
15. To discuss my work with my instructor or TA during class time.

2. SCALE-UP Survey 2

Part 1. Indicate how often you expect to experience the following in this SCALE-UP physics class, using the following scale:

1	2	3	4	5	6	7
Very infrequently	Infrequently	Somewhat infrequently	Sometimes	Somewhat frequently	Frequently	Very frequently

1. Lecture.
2. A grading curve.
3. Collaborative (group) discussions.
4. Computer modeling and/or programming.
5. Doing required reading.
6. Amphitheater-style (lecture hall) classroom.
7. Laboratories separate from the rest of class.
8. Missed classes would be harmful to my learning.
9. To memorize equations.
10. To interact with my instructor during class time.
11. To interact with my TA during class time.
12. To interact with my peers during class time.
13. To explain my work to the class.
14. To discuss my work with classmates during class time.
15. To discuss my work with my instructor or TA during class time.

Part 2. The communication I've been told to anticipate in this SCALE-UP physics course during class time is

1	2	3	4	5	6	7
Strongly disagree	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree	Strongly agree

1. Expected.
2. Normal.
3. Unusual.
4. Typical.
5. Uncommon.
6. Natural.
7. Inappropriate.
8. Out of place.
9. Foreseeable.
10. Unacceptable.

3. SCALE-UP Survey 3

Part 1.
1. What sorts of communication did you have with your instructor in this SCALE-UP physics course during class time?
2. What sorts of communication did you have with your teaching assistant(s) in this SCALE-UP physics course during class time?
3. What sorts of communication did you have with your peers in this SCALE-UP physics course during class time?

Part 2. Indicate how often you experienced the following in this SCALE-UP physics class, using the following scale:

1	2	3	4	5	6	7
Very infrequently	Infrequently	Somewhat infrequently	Sometimes	Somewhat frequently	Frequently	Very frequently

1. Lecture.
2. A grading curve.
3. Collaborative (group) discussions.
4. Computer modeling and/or programming.
5. Doing required reading.
6. Amphitheater-style (lecture hall) classroom.
7. Laboratories separate from the rest of class.
8. Missed classes would be harmful to my learning.
9. To memorize equations.
10. To interact with my instructor during class time.
11. To interact with my TA during class time.
12. To interact with my peers during class time.
13. To explain my work to the class.
14. To discuss my work with classmates during class time.
15. To discuss my work with my instructor or TA during class time.

Part 3. The communication I am experiencing this semester in this SCALE-UP physics class during class time is:

1	2	3	4	5	6	7
Strongly disagree	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree	Strongly agree

1. Expected.
2. Normal.
3. Unusual.
4. Typical.
5. Uncommon.
6. Natural.
7. Inappropriate.
8. Out of place.
9. Foreseeable.
10. Unacceptable.

Part 4. Please indicate the extent to which you agree with the following statements, using the following scale:

1	2	3	4	5	6	7
Strongly disagree	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree	Strongly agree

1. The SCALE-UP environment is a useful style of teaching and learning.
2. I would have learned physics better in a more traditional setting than SCALE-UP.
3. The SCALE-UP environment is inappropriate for college classes.
4. The style of this course helped me learn physics.
5. Courses in other departments should use a SCALE-UP environment.
6. SCALE-UP is not for me.

[1] J. D. H. Gaffney, E. Richards, M. B. Kustusch, L. Ding, and R. J. Beichner, Scaling up education reform, J. Coll. Sci. Teach. **37**, 48 (2008).

[2] R. J. Beichner, J. M. Saul, D. S. Abbott, J. J. Morse, D. L. Deardorff, R. J. Allain, S. W. Bonham, M. H. Dancy, and J. S. Risley, The student-centered activities for large enrollment undergraduate programs (SCALE-UP) project, in *Research-based reform of university physics* (2007), Vol. 1, WWW Document, (http://www.compadre.org/Repository/document/ServeFile.cfm?ID=4517&DocID=183)

[3] J. K. Burgoon, A communication model of personal space violations: Explication and an initial test, Human Communication Research **4**, 129 (1978).

[4] J. K. Burgoon, in *Intercultural Communication Theory*, edited by R. L. Wiseman (Sage, Thousand Oaks, CA, 1995), pp. 194–214.

[5] R. Hake, Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses, Am. J. Phys. **66**, 64 (1998).

[6] P. Heller, R. Keith, and S. Anderson, Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving, Am. J. Phys.you. **60**, 627 (1992).

[7] A. Van Heuvelen, Learning to think like a physicist: A review of research based instructional strategies, Am. J. Phys. **59**, 898 (1991).

[8] C. Henderson, A. Beach, N. Finkelstein, and R. S. Larson, *Facilitating change in undergraduate STEM: Initial results from an interdisciplinary literature review*, PER Conference (unpublished), http://www.compadre.org/Repository/document/ServeFile.cfm?ID=8094&DocID=727

[9] J. M. Saul, Ph.D. Dissertation, University of Maryland, 1998.

[10] E. F. Redish, J. M. Saul, and R. N. Steinberg, Student expectations in introductory physics, Am. J. Phys. **66**, 212 (1998).

[11] I. Halloun and D. Hestenes, Interpreting VASS dimensions and profiles for physics students, Sci. Educ. **7**, 553 (1998).

[12] W. K. Adams, K. K. Perkins, N. S. Podolefsky, M. Dubson, N. D. Finkelstein, and C. E. Wieman, New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey, Phys. Rev. ST Phys. Educ. Res. **2**, 010101 (2006).

[13] See, for example, D. Hammer, A. Elby, R. Scherr, and E. Redish, in *Transfer of Learning from a Modern Multidisciplinary Perspective*, edited by J. P. Mestre (Information Age Publishing, Greenwich, CT, 2005), pp. 89–120.

[14] J. K. Burgoon and J. L. Hale, Nonverbal expectancy violations: Model elaboration and application to immediacy behaviors, Commun. Monogr. **55**, 58 (1988).

[15] A. B. Frymier and B. Weser, The role of student predispositions on student expectations for instructor communication behavior, Communication Education **50**, 314 (2001).

[16] e.g., J. B. Dusek, *Teacher Expectancies* (Erlbaum, Hillsdale, NJ, 1985).

[17] Seiffert & Wheeless, 1990 as reported in C. D. Koermer & J. L. Petelle, Expectancy violation and student rating of instruction, Commun. Q. **39**, 331 (1991).

[18] M. B. McPherson, P. Kearney, and T. Plax, The dark side of instruction: Teacher anger as classroom norm violations, J.

[19] Appl. Commun. Res. **31**, 76 (2003).

[19] C. Deboer, Predicting continued participation in college chemistry for men and women, J. Res. Sci. Teach. **24**, 527 (1987).

[20] C. D. Koermer and W. J. Seiler, *The Relationship between Student Communication Expectations and Evaluations of Teacher Performance*, Eastern Communication Association Conference (unpublished).

[21] R. J. Gigliotti, Are they getting what they expect? Teaching Sociology **15**, 365 (1987).

[22] M. L. Houser, We don't need the same things! Recognizing differential expectations of instructor communication behavior for nontraditional and traditional students, J. of Continuing Higher Education **52**, 11 (2004).

[23] M. L. Houser, Are we violating their expectations? Instructor communication expectations of traditional and nontraditional students, Commun. Q. **53**, 213 (2005).

[24] A. L. Housley Gaffney, *A Qualitative Investigation of Communication Purposes in the SCALE-UP Classroom*, Southern States Communication Association Conference (unpublished).

[25] A. Astin, *What Matters in College? Four Critical Years Revisited* (Jossey-Bass, Inc., San Francisco, 1993).

[26] J. D. H. Gaffney, A. L. Housley Gaffney, and R. J. Beichner, *Do They See It Coming? Expectancy Violations in Reformed Classrooms*, American Association of Physics Teacher Conference (unpublished).

[27] More information about SCALE-UP is located on the World Wide Web at http://scaleup.ncsu.edu

[28] D. P. Maloney and M. F. Masters, Learning the game of formulating and testing hypotheses and theories, Phys. Teach. **48**, 22 (2010).

[29] T. P. Mottet, J. Parker-Raley, S. A. Beebe, and C. Cunningham, Instructors who resist "college lite:" The neutralizing effect of instructor immediacy on students' course-workload violations and perceptions of instructor credibility and affective learning, Commun. Educ. **56**, 145 (2007).

[30] L. J. Cronbach, Coefficient alpha and the internal structure of tests, Psychometrika **16**, 297 (1951).

[31] R. W. Chabay and B. A. Sherwood, *Matter & Interactions*, 2nd ed. (John Wiley & Sons, Inc, Hoboken, NJ, 2007).

[32] K. Cummings, P. W. Laws, E. F. Redish, and P. J. Cooney, *Understanding Physics*, 1st ed. (John Wiley & Sons, Inc, Hoboken, NJ, 2004).

[33] M. Hollander and D. A. Wolfe, *Nonparametric Statistical Methods*, 2nd ed. (Wiley-Interscience, New York, 1999).

[34] R. Perry, P. Abrami, L. Leventhal, and J. Check, Instructor reputation: An expectancy relationship involving student ratings and achievement, J. Educ. Psychol. **71**, 776 (1979).

[35] M. B. McPherson and Y. Liang, Students' reactions to teachers' management of compulsive communicators, Commun. Educ. **56**, 18 (2007).

[36] D. Hestenes, M. Wells, and G. Swackhamer, Force concept inventory, Phys. Teach. **30**, 141 (1992).

[37] L. Ding, R. Chabay, B. Sherwood, and R. Biechner, Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment, Phys. Rev. ST Phys. Educ. Res. **2**, 010105 (2006).