

The PULSE Vision & Change Rubrics, Version 1.0: A Valid and Equitable Tool to Measure Transformation of Life Sciences Departments at All Institution Types

Loretta Brancaccio-Taras,¹ Pamela Pape-Lindstrom,² Marcy Peteroy-Kelly,^{3*} Karen Aguirre,⁴ Judy Awong-Taylor,⁵ Teri Balsler,⁶ Michael J. Cahill,⁷ Regina F. Frey,⁷ Thomas Jack,⁸ Michael Kelrick,⁹ Kate Marley,¹⁰ Kathryn G. Miller,¹¹ Marcy Osgood,¹² Sandra Romano,¹³ J. Akif Uzman,¹⁴ and Jiuqing Zhao⁷

¹Department of Biological Sciences, Kingsborough Community College–CUNY, Brooklyn, NY 11235; ²Life Sciences Department, Everett Community College, Everett, WA 98201; ³Department of Biology, Pace University, New York, NY 10038; ⁴Department of Biology, Coastal Carolina University, Conway, SC 29528; ⁵Department of Biology, Georgia Gwinnett College, Lawrenceville, GA 30043; ⁶Curtin University, Faculty of Science and Engineering, Perth, Western Australia 6845, Australia; ⁷Center for Integrative Research on Cognition, Learning, and Education (CIRCLE) and ¹¹Department of Biology, Washington University in St. Louis, St. Louis, MO 63130; ⁸Department of Biological Sciences, Dartmouth College, Hanover, NH 03755; ⁹Department of Biology, Truman State University, Kirksville, MO 63501; ¹⁰Office of the Vice President for Academic Affairs, Doane College, Crete, NE 68333; ¹²Department of Biochemistry and Molecular Biology, University of New Mexico, Albuquerque, NM 87131; ¹³College of Science and Mathematics, University of the Virgin Islands, St. Thomas, VI 00802; ¹⁴College of Sciences and Technology, University of Houston–Downtown, Houston, TX 77002

ABSTRACT

The PULSE Vision & Change Rubrics, version 1.0, assess life sciences departments' progress toward implementation of the principles of the *Vision and Change* report. This paper reports on the development of the rubrics, their validation, and their reliability in measuring departmental change aligned with the *Vision and Change* recommendations. The rubrics assess 66 different criteria across five areas: Curriculum Alignment, Assessment, Faculty Practice/Faculty Support, Infrastructure, and Climate for Change. The results from this work demonstrate the rubrics can be used to evaluate departmental transformation equitably across institution types and represent baseline data about the adoption of the *Vision and Change* recommendations by life sciences programs across the United States. While all institution types have made progress, liberal arts institutions are farther along in implementing these recommendations. Generally, institutions earned the highest scores on the Curriculum Alignment rubric and the lowest scores on the Assessment rubric. The results of this study clearly indicate that the *Vision & Change* Rubrics, version 1.0, are valid and equitable and can track long-term progress of the transformation of life sciences departments. In addition, four of the five rubrics have broad applicability and can be used to evaluate departmental transformation by other science, technology, engineering, and mathematics disciplines.

INTRODUCTION

The disciplines of science, technology, engineering, and mathematics (STEM) play a vital role in our nation's economy, contributing to at least half of the economic growth in the United States during the past 50 years, and consistently providing a source of stable, high-earning jobs for appropriately skilled individuals (U.S. Congress Joint Economic Committee, 2012). However, there is currently concern about a shortage of qualified STEM workers. Based on data from the U.S. Bureau of

Debra Tomanek, *Monitoring Editor*

Submitted December 22, 2015; Revised May 17, 2016; Accepted June 1, 2016

CBE Life Sci Educ December 1, 2016 15:ar60

DOI:10.1187/cbe.15-12-0260

*Address correspondence to: Marcy Peteroy-Kelly (mkelly2@pace.edu).

© 2016 L. Brancaccio-Taras *et al.* CBE—Life Sciences Education © 2016 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

“ASCB®” and “The American Society for Cell Biology®” are registered trademarks of The American Society for Cell Biology.

Labor Statistics, employment in STEM occupations is expected to grow to more than 9 million between 2012 and 2022, an increase of ~1 million jobs above the 2012 employment level (Vilorio, 2014). An inventory of federal expenditures on STEM education conducted by the National Science and Technology Council (2011) revealed \$3.4 billion was spent, with 28% devoted to STEM workforce development and 72% expended on broader STEM education projects. Even with this substantial monetary investment, progress toward creating educational experiences that engage current students and result in an increase in the STEM talent pool and STEM graduates has fallen short.

In 2012, the President's Council of Advisors on Science and Technology (PCAST) released a report suggesting that the first 2 years of undergraduate study are the most critical for recruiting and retaining STEM majors needed to fill the STEM employment gap (PCAST, 2012). Furthermore, the report states that 60% of the students entering college intending to major in a STEM discipline do not graduate with a STEM degree. Many of the students who leave STEM majors reported that their introductory courses were uninspiring and unwelcoming, and those experiences were enough to discourage them from majoring in STEM disciplines. The PCAST report suggests that colleges and universities attempt to increase the retention of STEM majors from 40 to 50% by providing the students with an educational environment that uses evidence-based, best practices in teaching and learning, while offering the academic and social support students need to persist to earn a STEM degree.

In response to the suggestions in the PCAST report and similar reports published over the past decade (National Research Council [NRC], 2003, 2011; American Institutes for Research, 2012; National Science and Technology Council, 2013), STEM faculty, funding agencies, and stakeholders have looked more intentionally at the reasons students do not select or persist in life sciences majors. College students and faculty members have long argued that the approach to undergraduate education in the life sciences should be modernized to reflect what is known about how students learn. They assert that the pedagogies (Freeman *et al.*, 2014; Wieman, 2014) and high-impact practices known to enhance student learning (Kuh, 2008) should be incorporated into life sciences programs nationwide. Twenty-first-century science demands that students develop modern scientific and technical skills, as well as the capacity to work beyond traditional academic boundaries. Undergraduate students, regardless of their majors, deserve and need a transformed life sciences curriculum that teaches them foundational biological concepts and allows them to become adept in scientific competencies. Informed decision making, whether around managing one's health, understanding how individual actions influence the environment, or understanding political policy discussions on scientific issues (e.g., stem cell research, climate change) requires an appreciation of key biological concepts and the nature and process of science.

As a result of a nationwide conversation about the future of life sciences education, *Vision and Change in Undergraduate Biology Education: A Call to Action* was published by the American Association for Advancement in Science (AAAS) in 2011. It included a set of recommendations for transforming life sciences education. One of the most significant recommendations

of this report is the recognition that a 21st-century undergraduate education requires systemic changes to how biology is taught, how curricular decisions are made, and how academic departments support faculty in developing and implementing modern student-learning methods. Many dedicated faculty members are changing their individual courses; however, for systemic change to be effective and sustainable, it must begin at the departmental level across the range of postsecondary educational institution types.

To explore how this systemic change can be realized across the country, the National Science Foundation (NSF), the National Institute of General Medical Science of the National Institutes of Health, and the Howard Hughes Medical Institute collaborated to form the Partnership for Undergraduate Life Sciences Education (PULSE) in 2012 (Dolan, 2012). PULSE began with the selection of 40 Vision and Change (V&C) Leadership Fellows; all were current or former life sciences department chairpersons or deans from a variety of institution types, including community colleges, liberal arts colleges, regional comprehensive universities, and research universities. Initially, the V&C Leadership Fellows were charged with developing strategies to enact the recommendations of the *Vision and Change* report over a 1-year period. These strategies were intended to promote changes in the way life sciences departments institutionalize best practices in evidence-based teaching and learning, develop curricula and infrastructure, create effective strategies for motivating systemic educational change, and assess their progress with an eye toward continuous improvement. During the first year of work, the V&C Leadership Fellows developed key projects and strategies to facilitate this national effort for systemic change (www.pulsecommunity.org; Woodin *et al.*, 2012). The V&C Fellows membership has been expanded so that the concerted effort to promote and adopt the recommendations in the *Vision and Change* report can continue nationally.

A PULSE pilot recognition program was one strategy developed by a subset of the V&C Leadership Fellows. The PULSE pilot recognition program was designed to provide undergraduate life sciences departments the opportunity for guided self-reflection and peer-review feedback about their programs' progress in implementing the *Vision and Change* recommendations. Based on existing models, a set of rubrics was developed that would serve life sciences departments in this self-reflection process and measure the extent of adoption of the principles of *Vision and Change*. In 2013, the PULSE Vision & Change Rubrics, version 1.0 (V&C Rubrics), were released (Aguirre *et al.*, 2013) and made available to the life sciences community on the PULSE community website (Supplemental Material).

This paper reports on the V&C Rubrics development process, their validation, and their reliability in measuring departmental change aligned with the *Vision and Change* recommendations at different institution types. In addition, we present an analysis of the findings based on the rubric data that were collected. We address three questions: 1) Are the V&C Rubrics an appropriate measurement tool across all institution types? 2) Can the rubrics be used to evaluate the adoption of the principles of *Vision and Change* by life sciences programs across all institutional types in the United States? 3) Is it possible to measure the implementation of *Vision and Change* nationwide?

METHODS AND RESULTS

Creation of the V&C Rubrics

The development of the rubrics for a recognition program began with extensive research on existing certification/accreditation models starting with the Accreditation Board of Engineering and Technology, which accredits college and university engineering programs (www.abet.org/accreditation) through a voluntary review process. Additionally, other models that were simultaneously under development were uncovered. For example, the American Society of Biochemistry and Molecular Biology had been working for several years on an accreditation program for biochemistry and molecular biology departments based on the principles of *Vision and Change* (www.asbmb.org/accreditation). The Association of American Colleges and Universities (2010) released a set of specific guidelines to its member institutions describing how STEM departments can move toward offering more student-centered environments that include active-learning experiences (www.aacu.org/value/rubrics). And the Royal Society of Biology in the United Kingdom recently instituted an accreditation program (www.rsb.org.uk/education/accreditation) that incorporates principles similar to those outlined in the *Vision and Change* report.

The PULSE recognition team created draft versions of the Vision & Change Rubrics in January 2013. Feedback and comments with regard to rubric content and wording were collected from all PULSE V&C Leadership Fellows and life sciences faculty via the PULSE community website. In spring 2013, the face validity of the draft rubrics was tested by presenting them at professional meetings such as the National Meeting of the Society for the Advancement of Biology Education Research (SABER) and the American Society for Microbiology's Conference for Undergraduate Educators (ASMCUE). For instance, at ASMCUE, ~300 faculty members were divided into groups based on institution type, and three of the five rubrics were distributed. Attendees were asked to comment whether the rubrics would be useful and indicated modifications that were needed. Feedback was collected, revisions were made, and the PULSE V&C Rubrics were released to the life sciences community via the PULSE community website (Aguirre *et al.*, 2013). These rubrics assess 66 different criteria across five broad rubric

areas: Curriculum Alignment (11 criteria), Assessment (12 criteria), Faculty Practice/Faculty Support (21 criteria), Infrastructure (10 criteria), and Climate for Change (12 criteria). A sample of the rubric structure can be found in Table 1. For each of the 66 criteria, life sciences departments select their level of progress in implementing the recommendations in *Vision and Change* from a range of 0–4 (with 4 being equivalent to exemplary progress toward implementing the recommendations and 0 being equivalent to baseline progress toward implementing the recommendations). The rubrics are accompanied by an instruction manual designed to provide guidance on rubric completion (see the Supplemental Material).

Pilot Recognition Process

In addition to the development of the V&C Rubrics and the collection of rubric data, an NSF-funded pilot recognition program was conducted to motivate life sciences departments to adopt the recommendations of the *Vision and Change* report. More than 70 schools applied and eight were selected. In this paper, the following terminology is used: doctorate-granting universities = R1, comprehensive universities and colleges = RC, liberal arts colleges = LA, and 2-year colleges = CC. These terms were selected because they have been commonly used when describing institutions of higher learning. Two were chosen from each of the four institution types based on initial evidence of transformed and innovative educational practices (Pape-Lindstrom *et al.*, 2015). The eight selected pilot institutions were asked to submit written justifications for their rubric scores and other supplemental documentation, including course syllabi, sample exams, and faculty CVs. Each school received a site visit by two recognition-team members. During the 2-day site visits, the recognition-team members met with administrators, faculty, and students; observed classes; and toured the institutions' facilities. These site visits were conducted to corroborate the information that the pilot schools submitted. The self-reported rubric scores submitted by the departments were typically in agreement with the team's evaluation of the progress made toward implementation of the principles of *Vision and Change*.

Based on evaluation of all documentation and additional information gathered at the site visits, each department was

TABLE 1. Sample structure of the V&C Rubrics

Rubric	Sections	Criteria
Curriculum alignment	A. Core Concepts	<ol style="list-style-type: none"> 1. Evolution core concept integrated into curriculum 2. Structure and function core concept integrated into curriculum 3. Information flow, exchange, and storage core concept integrated into curriculum 4. Pathways and transformations of energy and matter core concept integrated into curriculum 5. Systems core concept integrated into curriculum
	B. Integration of Core Competencies	<ol style="list-style-type: none"> 1. Integration of the process of science into the curriculum 2. Integration of quantitative reasoning into the curriculum 3. Integration of modeling and simulation into the curriculum 4. Integration of interdisciplinary nature of science into the curriculum 5. Communication and collaboration through a variety of formal and informal written, visual, and oral methods integrated into curriculum 6. An understanding of the relationship between science and society is embedded in curriculum

assigned a PULSE Progression Level. PULSE Progression is modeled after the U.S. Green Building Council's Leadership in Energy and Environmental Design (LEED) certification, in which buildings evaluated for specific design features are recognized with LEED Silver, Gold, or Platinum certification. Each level of achievement is associated with a specific point threshold. PULSE Progression Levels provide independent verification of a life sciences department's transformative features (Pape-Lindstrom *et al.*, 2015) and are designed to reflect the progress the department has achieved in implementing the recommendations of *Vision and Change*. Every level of PULSE Progression indicates a dedicated and concerted effort by a department to remodel its approach to life sciences education so that undergraduate teaching and learning in the life sciences are improved.

Rubric Data Collection

A Qualtrics rubric data-collection portal was created to gather institutional rubric data, and a request was sent to all PULSE community members to submit their departmental rubric data. Some institutions completed all five rubrics and submitted full data sets ($n = 26$). Eight of the 26 institutions that submitted full data sets were the participants in the PULSE pilot recognition program. Other institutions submitted partial data sets. For data to be included in the analysis reported here, an institution must have completed at least one full rubric. This collection method resulted in variation in the number of reports submitted for each rubric. For example, 57 data sets were analyzed for the Curriculum Alignment rubric and 35 for the Assessment rubric (Table 2).

Weighting Scheme

To evaluate and compare rubric data from different institution types, the recognition team created a weighting scheme, emphasizing criteria critical for implementation of *Vision and Change* (Table 3). Generally, the weighting scheme was informed by the team's extensive and collective experiences teaching at different institution types, the research conducted on accreditation models (Aguirre *et al.*, 2013), feedback from face validity, observations from the pilot-school site visits, and the team's vision of a fully transformed curriculum. The vision was heavily influenced by discussions with the complete PULSE Fellows membership, and with faculty from around the country at conferences and workshops. A fully transformed curriculum would include features that are highly likely to enhance the student experience and transform student learning. Aspects of the rubrics that are typically associated with practices that enhance the student experience were given higher weights, such as elements of the Assessment rubric (Momsen *et al.*, 2013; Brame and Biel, 2015; Couch *et al.*, 2015) and the Faculty Practice/Faculty Support rubric (D'Avanzo, 2013; Smith *et al.*, 2013; Wieman and Gilbert, 2014; Eddy *et al.*, 2015). Other components of the rubrics, such as elements of the Infrastructure rubric, although important, are not as critical to fully drive the enhancement of student experiences. These rubrics were therefore given lower weights.

There is a abundant literature supporting the notion that providing students with opportunities to engage in the process and practice of science enhances their learning experiences (Russell *et al.*, 2007; Freeman *et al.*, 2014; Wieman, 2014; Connell *et al.*, 2016). It is not only essential to provide engaging

TABLE 2. Entire rubric data set organized by institution type and number of reports for each rubric with unweighted and weighted mean scores and SEMs by institution type reported for each rubric

Rubric	Number of programs/departments reporting these data	Sample size	Unweighted mean (SEM)	Weighted mean (SEM)
Curriculum Alignment	57	R1: $n = 13$	2.78 (0.15)	2.67 (0.17)
		RC: $n = 16$	2.77 (0.17)	2.72 (0.17)
		LA: $n = 11$	3.02 (0.17)	2.97 (0.18)
		CC: $n = 17$	2.62 (0.12)	2.52 (0.13)
Assessment	35	R1: $n = 9$	1.34 (0.17)	1.35 (0.19)
		RC: $n = 10$	1.21 (0.14)	1.16 (0.16)
		LA: $n = 8$	1.67 (0.17)	1.68 (0.18)
		CC: $n = 8$	1.52 (0.26)	1.54 (0.30)
Faculty Practice/Faculty Support	49	R1: $n = 11$	2.10 (0.15)	2.07 (0.16)
		RC: $n = 14$	2.10 (0.12)	2.09 (0.12)
		LA: $n = 12$	2.42 (0.16)	2.51 (0.16)
		CC: $n = 12$	1.77 (0.11)	1.72 (0.11)
Infrastructure	28	R1: $n = 6$	2.47 (0.48)	2.43 (0.49)
		RC: $n = 8$	2.33 (0.22)	2.33 (0.23)
		LA: $n = 7$	2.57 (0.21)	2.63 (0.23)
		CC: $n = 7$	2.43 (0.30)	2.44 (0.27)
Climate for Change	32	R1: $n = 7$	1.75 (0.29)	1.75 (0.29)
		RC: $n = 11$	1.59 (0.17)	1.59 (0.17)
		LA: $n = 7$	1.87 (0.29)	1.87 (0.29)
		CC: $n = 7$	1.76 (0.29)	1.76 (0.29)

TABLE 3. Rubric weighting scheme

Rubric category/section	Weighting factor	Number of criteria	Possible points
Curriculum Alignment		11	68 (11%)
A. Core Concepts	× 1	5	20
B. Integration of Core Competencies	× 2	6	48
Assessment		12	136 (23%)
A. Course Level Assessment	× 2	7	56
B. Program Level Assessment	× 4	5	80
Faculty Practice/Faculty Support		21	296 (50%)
A. Student Higher Level Learning	× 6	5	120
B. Learning Activities beyond the Classroom	× 4	6	96
C. Faculty Development	× 2	10	80
Infrastructure		10	48 (8%)
A. Physical Infrastructure	× 1	5	20
B. Learning Spaces	× 2	2	16
C. Resources and Support	× 1	3	12
Climate for Change (all sections)	× 1	12	48 (8%)
Total		66	596 (100%)

opportunities for students, but also important to assess whether or not those opportunities are indeed enhancing student learning (Momsen *et al.*, 2013; Freeman *et al.*, 2014; Wieman, 2014; Brame and Biel, 2015; Couch *et al.*, 2015). Our weighting scheme was designed to acknowledge departments that embrace these practices and to reward more fully transformed departments with higher overall scores. Because there are often roadblocks to the implementation and measurement of these practices, the higher weights on these elements may also encourage departments to fully implement these recommendations.

Another driver for the adoption of the weighting scheme is the unequal distribution of criteria in each rubric section. In the absence of rubric data weighting, institutions that have made gains in enacting practices to enhance their students' experiences may earn lower, overall rubric scores. This may result from lower scores on the other sections of the rubrics that highlight aspects not as essential to departmental transformation toward enhancing the student experience. The site visits enabled the recognition team to align the observations they made at the institutions they visited with the unweighted and weighted rubric scores to confirm the weighting scheme model.

Examination of the Rubrics for Reliability

Statistical analyses conducted for this study were performed using Statistical Analysis System (SAS, version 9.3, for Windows, 2002–2010) and R; significance was determined at 0.05. Each rubric was initially divided into sections that, a priori, were designed to target a specific component of the rubric. Using all available data for each rubric, the internal consistency or reliability of the rubric sections was tested by computing Cronbach's α for each (Cronbach, 1951). Generally, $\alpha \geq 0.7$ is considered acceptable reliability. All sections of the Curriculum, Assessment, and Faculty Practice/Faculty Support rubrics exhibited adequate reliability. However, not all original sections of the Infrastructure and Climate for Change rubrics met this condition (Table 4).

An exploratory factor analysis (EFA; Hotelling, 1933, Fabrigar *et al.*, 1999; Suhr, 2005) was conducted to determine the most

coherent structure for each section of the Infrastructure and Climate for Change rubrics. EFA examines the underlying correlation structure of a set of items (Browne, 2001; Brown, 2009) and identifies coherent groupings within the larger set of items. Using all data for each rubric, all rubric items were included in a factor analysis, using principal components extraction with a varimax rotation (Kaiser, 1958). A factor analysis generates a number of factors equal to the number of items included in the analysis, but not all factors are retained. Each factor has an eigenvalue (indicating the proportion of variance in the data the factor accounts for), and each item has a loading for each factor, indicating how strongly the item associates with the given factor. For each analysis, the number of factors to retain based on the Kaiser criterion (all factors with eigenvalues < 1 are dropped) was applied, followed by the scree test, in which all remaining eigenvalues were plotted from left to right in descending order. Factors were removed if they occurred at or to the right of the location of the plot in which the eigenvalues "leveled off." Once the retained factors were determined, each item was placed into the retained factor on which it loaded most highly.

Based on the EFA, new structures were generated for the Infrastructure and Climate for Change rubrics and Cronbach's α values were then recalculated. Table 4 shows the original rubric structure, section labels, and Cronbach's α coefficients and the revised structure, labels, and coefficients. The reclustering resulted in adequate reliability for sections, with Cronbach's $\alpha \geq 0.7$. The new groupings were also examined for conceptual coherence, to identify a conceptual underpinning and to create meaningful labels for all new sections. The reliability analyses and the EFA resulted in major revisions to the Infrastructure and Climate for Change rubrics. As a result of these revisions, all rubrics are now reliable measures of progress on the implementation of the *Vision and Change* recommendations.

Analysis of Full Rubrics Data Sets

The rubrics were developed with the hypothesis that they could be used to evaluate departmental transformation equitably across institution types. To address this hypothesis, the data

TABLE 4. Original and reclustered Infrastructure and Climate for Change rubrics based upon EFA analyses

Rubric (original rubric Cronbach's α)	Reclustered rubric with improved Cronbach's α^a
Curriculum	
A. Core Concepts ($\alpha = 0.79$)	
B. Integration of Core Competencies ($\alpha = 0.78$)	
Assessment	
A. Course Level Assessment ($\alpha = 0.70$)	
B. Program Level Assessment ($\alpha = 0.74$)	
Faculty Practice/Faculty Support	
A. Student Higher Level Learning ($\alpha = 0.79$)	
B. Learning beyond the Classroom ($\alpha = 0.80$)	
C. Faculty Development ($\alpha = 0.80$)	
Infrastructure	
A. Physical Infrastructure ($\alpha = 0.84$)	A. Learning Spaces ($\alpha = 0.87$)
Classrooms and teaching laboratories can accommodate special needs	Classrooms and teaching laboratories can accommodate special needs
Teaching spaces to encourage student interaction	Teaching spaces to encourage student interaction
Classroom IT infrastructure	Classroom IT infrastructure
Intelligently designed laboratory	<i>Informal gathering spaces that encourage collaboration</i>
Equipment/supplies in teaching laboratories	<i>Learning center for students</i>
B. Learning Spaces ($\alpha = 0.64$)	B. Laboratory Spaces ($\alpha = 0.76$)
Informal gathering spaces that encourage collaboration	<i>Intelligently designed laboratory spaces</i>
Learning center for students	<i>Equipment/supplies in teaching laboratories</i>
C. Resources and Support ($\alpha = 0.71$)	C. Resources and Support ($\alpha = 0.79$)
IT support for innovative teaching	IT support for innovative teaching
Staff support for teaching	Staff support for teaching
Institutional support for electronic resources	Institutional support for electronic resources
Climate for Change	
A. Administrative And Institutional Vision ($\alpha = 0.72$)	A. Institutional Awareness and Communication of Vision ($\alpha = 0.89$)
Vision is clear and specific	Commitment to vision is demonstrated through administrative action
Vision aligns with V&C priorities	<i>There is awareness and buy-in of national initiatives in higher education</i>
Commitment to vision is demonstrated through administrative action	<i>There is a collaborative communication process in place, including disseminating new ideas</i>
	<i>There is faculty support for the administrative vision within the department</i>
B. Administrative and Institutional Attitude ($\alpha = 0.59$)	B. Strategies for Promoting Systemic Change in Teaching Culture ($\alpha = 0.78$)
Administration is supportive of the need for change	Administration is supportive of the need for change
There is awareness and buy-in of national initiatives in higher education	<i>Vision aligns with V&C priorities</i>
Institutional evaluation and assessment reflects the importance of teaching	<i>Strategies are in place to recruit and retain diverse teaching faculty</i>
	<i>Resources exist for faculty to improve their teaching methods</i>
C. Administrative and Institutional Action ($\alpha = 0.71$)	C. Concrete Implementations Promoting Change in Teaching Culture ($\alpha = 0.71$)
Strategies are in place to recruit and retain diverse teaching faculty	Faculty incentives exist for transformative approaches in teaching
Faculty incentives exist for transformative approaches in teaching	<i>Fund-raising and development efforts support departmental transformation in alignment with Vision & Change</i>
Resources exist for faculty to improve their teaching methods	<i>Institutional evaluation and assessment reflects the importance of teaching</i>
Fund-raising and development efforts support departmental transformation in alignment with V&C	<i>Vision is clear and specific</i>
D. Departmental Support ($\alpha = 0.88$)	
There is a collaborative communication process in place, including disseminating new ideas	
There is faculty support for the administrative vision within the department	

^aReclustered criteria are italicized.

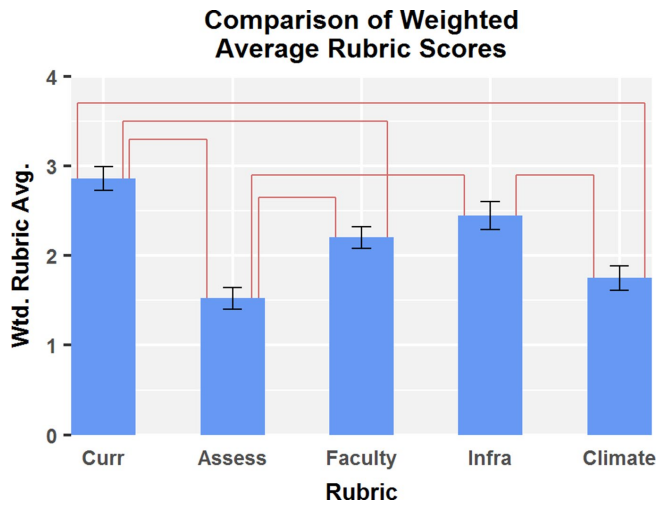


FIGURE 1. Weighted average rubric scores for 26 institutions with full data sets. Values represent scores, not ranks, with a possible range of 0–4. Error bars represent the SEM. Connecting lines represent statistically significant pairwise differences ($p < 0.05$), based on post hoc analysis (Tukey-Kramer least squared [LS] means). The rubric criteria can be found in the Supplemental Material. Curr = Curriculum, Assess = Assessment, Faculty = Faculty Practice/Faculty Support, Infra = Infrastructure, Climate = Climate for Change.

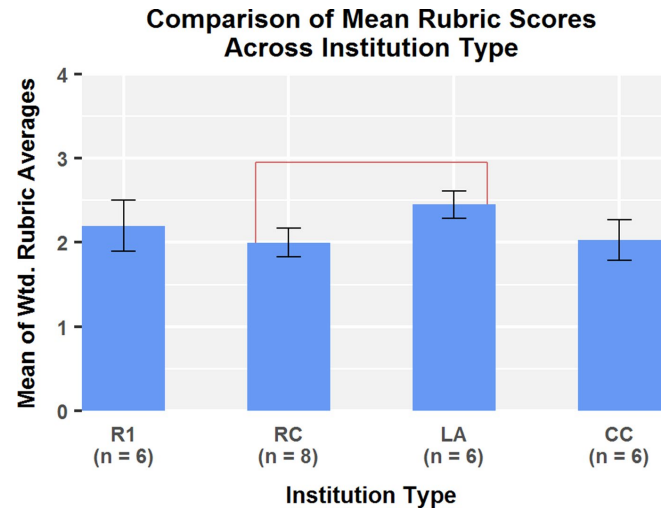


FIGURE 2. Weighted average scores, collapsed across the five rubrics and grouped by institution type, for the 26 institutions with full data sets. Values represent the scores, not ranks, with a possible range of 0–4. Error bars represent the SEM. Connecting lines represent statistically significant pairwise differences ($p < 0.05$), based on post hoc analysis (Tukey-Kramer LS means). The rubric criteria can be found in the Supplemental Material.

from the 26 institutions that completed all five of the rubrics were grouped by institution type: R1, RC, LA, or CC. Of the 26 complete data sets, six were submitted by R1 institutions, eight by RCs, six by LA colleges, and six by CCs. Each institution’s weighted mean score for each rubric was calculated, using the weighting scheme presented in Table 3. A two-way analysis of variance (ANOVA) of institution type versus rubric, with interaction term on ranked data (i.e., analogous to a Kruskal-Wallis test; Conover and Iman, 1981; Akritas, 1990) was performed, testing three effects: rubric main effect, institution-type main effect, and rubric \times institution type interaction. Significant effects were followed with post hoc pairwise comparisons. The design was unbalanced (i.e. differing in number of replicate institutions among type), so least-squared means were used for these post hoc tests, and the Tukey-Kramer method was used to adjust for multiple comparisons.

The rubric main effect directly tested whether implementation differed across the various rubrics, and the significant main effect ($F(4,110) = 15.46, p < 0.01$) indicates significant variation across rubrics. Notably, departments reported the highest degree of implementation on curriculum and the least implementation on assessment. Figure 1 and Table 5 display the

TABLE 5. p Values for post hoc analysis (Tukey-Kramer LS means) pairwise comparisons of weighted average rubric scores in Figure 1

	Curr	Assess	Faculty	Infra	Climate
Curr		<0.001	0.01	0.36	<0.001
Assess			<0.01	<0.001	0.75
Faculty				0.64	0.12
Infra					<0.01
Climate					

Curr = Curriculum, Assess = Assessment, Faculty = Faculty Practice/Faculty Support, Infra = Infrastructure, Climate = Climate for Change.

pattern of means and an indication of which rubric scores significantly differed from one another.

The rubrics were developed with the intent to evaluate departmental transformation equitably across all institution types. Figure 2 shows the mean scores, collapsed across rubrics and grouped by institution type. It is intended that the rubrics will be used to evaluate progress over time. The data presented here represent a baseline measurement. The question of whether the rubrics equitably measure progress across all institution types was addressed by the institution-type main effect and the interaction term of the aforementioned ANOVA. The institution-type main effect was significant ($F(3, 110) = 3.04, p < 0.04$), indicating significant differences across institution types, collapsed across all rubrics. Post hoc tests revealed that the LA institutions had significantly higher means than the RCs, and no other differences were significant (Figure 2). The interaction term was not significant ($F(12, 110) = 0.71, p > 0.7$), indicating that the relative standing of institution types does not significantly differ across the rubrics. Although LA and RC institutions significantly differ from each other, there is considerable overlap in the score distributions of these groups. The data show that even the institution type with the lowest mean score has representative institutions that score nearly as high as any other institution in the data set.

Overall, the analysis of full data sets reveals significant differences in progress across rubrics, with the most progress reported in the area of curriculum alignment and the least on assessment. However, examining the distribution of scores suggests no inherent bias exists that would prevent any particular institution from achieving high scores.

Analysis of Individual Rubrics

Many institutions did not complete all five rubrics (Table 2). Therefore, analyzing the data from each rubric separately

TABLE 6. ANOVA tables for analyses of rubric scores and ranked rubric scores

Measure ^a	SS _{Effect}	SS _{Error}	df _{Effect}	df _{Error}	MS _{Effect}	MS _{Error}	F	p Value
One-way (institution type) ANOVAs on weighted averages								
Curr	1.36	19.59	3	53	0.45	0.37	1.23	0.31
Assess	1.33	11.41	3	31	0.44	0.37	1.20	0.33
Faculty*	3.75	10.57	3	45	1.25	0.23	5.32	<0.001
Infra	0.34	15.37	3	24	0.11	0.64	0.18	0.91
Climate	0.36	13.65	3	28	0.12	0.49	0.24	0.86
Faculty-A*	4.86	18.08	3	45	1.62	0.40	4.03	0.01
Faculty-B*	8.81	17.94	3	45	2.94	0.40	7.37	<0.001
Faculty-C	0.62	17.28	3	45	0.21	0.38	0.54	0.66
One-way (institution type) ANOVAs on ranked weighted averages								
Curr	938.97	14439.03	3	53	312.99	272.43	1.15	0.34
Assess	374.54	3184.46	3	31	124.85	102.72	1.22	0.32
Faculty*	2338.99	7454.01	3	45	779.66	165.64	4.71	0.01
Infra	61.36	1755.14	3	24	20.45	73.13	0.28	0.84
Climate	60.59	2659.41	3	28	20.20	94.98	0.21	0.89
Faculty-A*	1946.22	7756.28	3	45	648.74	172.36	3.76	0.02
Faculty-B*	3816.51	5902.00	3	45	1272.17	131.16	9.70	<0.001
Faculty-C	320.64	9430.86	3	45	106.88	209.57	0.51	0.68

^aAn asterisk indicates that the main effect of institution type was significant for this measure ($p < 0.05$). The four categories of institution type are R1, RC, LA, and CC. SS = sum of squares, MS = mean sum of squares; A, B, and C refer to sections of the Faculty Practice/Faculty Support rubric; see Table 3.

allowed larger sample sizes for statistical analyses. In these analyses, a series of one-way ANOVAs was conducted with institution type as the independent variable and a given weighted rubric score as the dependent variable. These analyses were conducted with ranked data and weighted scores. Post hoc tests used least-squared means and the Tukey-Kramer method to correct for multiple comparisons. Results of ANOVAs on ranks (Kruskal-Wallis) and ANOVAs on scores yielded similar results (Table 6), with the only significant effect of institution type emerging on the Faculty Practice/Faculty Support rubric. Therefore, graphs of the data present ANOVAs on the scores themselves, not the ranked scores.

The data in Figure 3A and Table 6 indicate that the Faculty Practice/Faculty Support rubric shows significant differences by institution type. Figure 2 displays the mean weighted scores, grouped by institution type, and indicates statistically significant differences based on the post hoc comparisons. Overall, LA institutions scored the highest on Faculty Practice/Faculty Support. As shown in Figure 3A, the only significant pairwise comparison was between LA colleges and CCs. Further analysis examined the scores on the three sections of this rubric (A = student higher-level learning, B = learning activities beyond the classroom, and C = faculty development) to identify the sources of difference in scores for this rubric. A significant main effect of institution type was found for both sections A and B. Figure 3, B and C, shows the overall pattern of means for these sections and indicates which groups are significantly different from one another based on post hoc comparisons.

Analysis of Weighing Scheme Impact

Unweighted and weighted mean scores are shown in Table 2. For each rubric, a two-way ANOVA of institution type versus weighting scheme was conducted, with an interaction term. The interaction term, weighting versus institution type, was found to be significant for the Faculty Practice/Faculty Support

rubric ($F(3, 45) = 3.12, p = < 0.05$). For this rubric, the weighting scheme slightly increased the scores of the LA institutions and slightly decreased scores of the CC, RC, and R1 institutions (Table 2). This is likely due to LA schools reporting higher scores on sections of this rubric with higher weighting, student higher-level learning, and learning activities beyond the classroom (sections A and B, Table 3), while the other institution types score relatively well on Faculty Development (section C, Table 3). Indeed, we can think of LA institutions as models for the student experience and so it is not surprising these sections of the rubric showed a benefit to LA institutions.

Significance of Rubric Sections to Scores

An additional analysis was conducted to determine which sections were most important in terms of their association with overall rubric performance. First, principal components analysis (PCA) on the rubric section scores using the reclustered sections in the case of Infrastructure and Climate for Change was conducted. In PCA, linear combinations of the input variables, called principal components (PCs), are extracted from the data, such that PC 1 is the linear combination that extracts the maximum amount of variance from the data, and each successive PC extracts decreasing amounts of variance. In this way, much of the variance in the data can be retained with relatively few PCs. PC 1 can be considered a one-dimensional representation that best captures the overall variation in the 13-dimensional variable space. The results (Figure 4) indicate that an institution's performance on curriculum B, which measures progress on the six core competencies, indicates stronger performance on the rubrics overall and is most important in score discrimination between institutions. The A section of the Faculty Practice/Faculty Support rubric, which measures elements of student higher-level learning, is the second most important section in discriminating between schools.

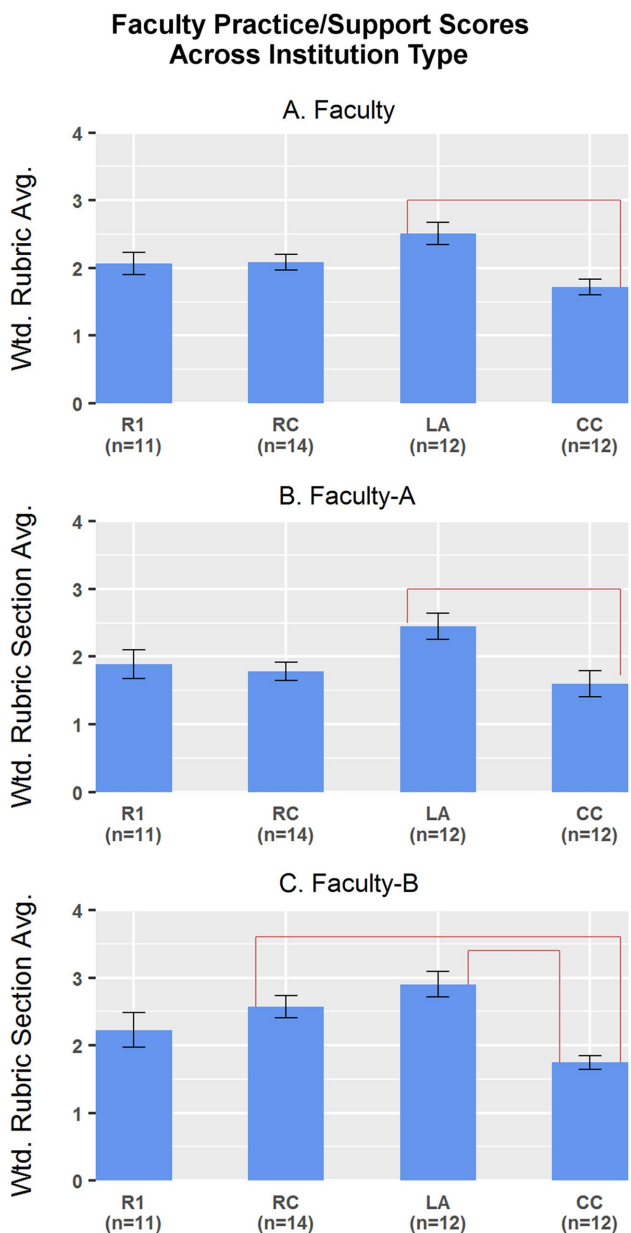


FIGURE 3. Weighted average scores, grouped by institution, for the Faculty Practice/Faculty Support rubric. (A) Overall rubric score, which is a weighted average of sections A, B, and C. (B) Score of section A, which contains five criteria that address inquiry, metacognition, and higher-order cognitive processes. (C) Score of section B, which contains six criteria that address learning activities beyond the classroom. Values represent the scores, not ranks, with a possible range of 0–4. Error bars represent the SEM. Connecting lines represent statistically significant pairwise differences ($p < 0.05$), based on post hoc analysis (Tukey-Kramer LS means). In addition to those marked as significant, the difference between LA and RC was marginally significant for section A ($p = 0.0504$), and the difference between R1 and LA was marginally significant for section B ($p = 0.06$). The specific rubric criteria can be found in the Supplemental Material.

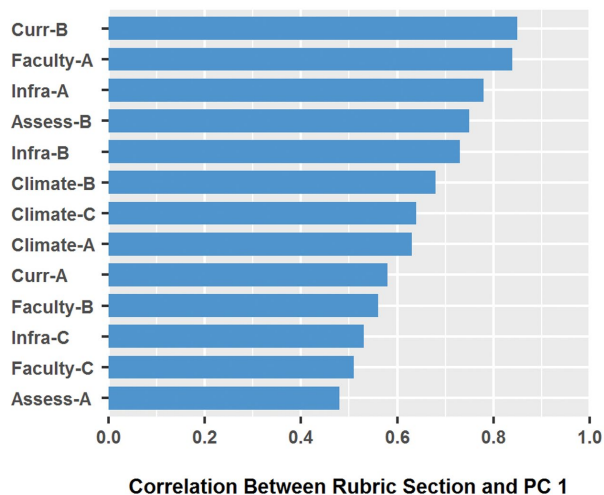


FIGURE 4. PCA including all 26 institutions with full data sets. PC 1 is the first PC extracted from a PCA including the full data sets from the 26 institutions. The inputs to the PCA were the weighted averages for the 13 rubric sections (listed along the y-axis), and PC 1 is the best linear combination of those rubric section scores, in terms of retaining the most variance from the original input variables. The horizontal bars represent the correlation between each individual rubric section, and PC 1, among the 26 full data sets. This correlation indicates how strongly each rubric section was associated with the overall pattern of variation in the data across all rubric sections.

National Progress with Regard to the Implementation of Vision and Change

Of the 26 complete data sets, six were submitted by R1 institutions, eight by RC institutions, six by LA colleges and six by CCs. For each institution, a total weighted score was computed to provide a single overall index of the progress made in adopting the *Vision and Change* recommendations. Out of a possible 596 points, total weighted scores ranged from 167 to 441 (Figure 5). The higher the total weighted score, the more progress the institution has made toward implementing the recommendations in *Vision and Change*.

Generally, institutions had the highest scores on the Curriculum Alignment rubric and the lowest scores on the Assessment rubric (Figure 1). The rubrics were capable of discriminating between institutions based upon their rubric scores, indicating the level of incorporation of *Vision and Change* report recommendations. Examination of the data submitted revealed that all institution types have made the most progress in terms of issues related to curriculum alignment; these scores were generally the highest across all institutions. Fifty-seven institutions submitted data for the Curriculum Alignment rubric with no significant differences found by institution type for these scores. The least degree of implementation appears to be in the area of course-level and program-level assessment. There were no statistical differences in the scores submitted among the 35 reporting institutions who reported data for the Assessment rubric. These data represent baseline scores. As institutions report their rubric scores in the future, comparison with baseline data will

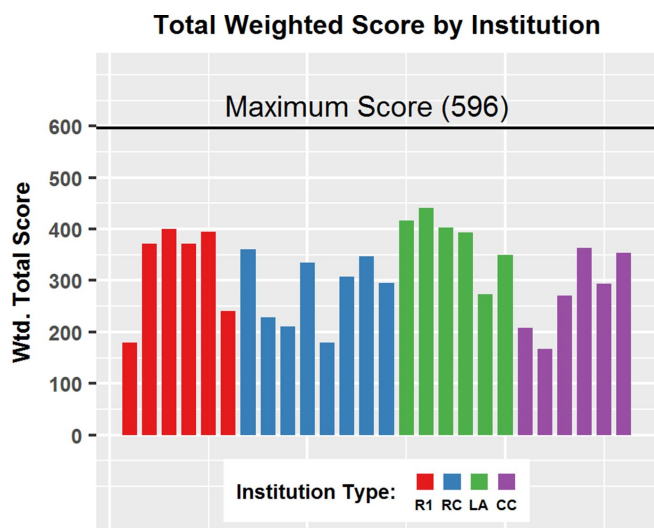


FIGURE 5. Values represent the total weighted scores of the 26 institutions that completed all five rubrics. Each bar represents the total score from a single institution. Bars are grouped by institution type for ease of comparison. The maximum possible score is 596. See Table 3 for the weighting scheme. All of the rubric criteria can be found in the Supplemental Material.

allow the determination of the transformational progress made in life sciences departments according to the recommendations of *Vision and Change*.

DISCUSSION

Rubrics are known to provide a reliable way to conduct assessment, foster self-analysis and self-reflection (Jonasson and Svingby, 2007), and serve as accountability structures required for successful change in higher education (Kezar, 2009). In this study, weighted rubric scores were analyzed as complete data sets for 26 institutions and further analyses with larger sample sizes (Table 2) were conducted on the five individual rubrics that comprise the V&C Rubrics. Based on the statistical findings, the V&C Rubrics are a valid measurement tool to assess the state of implementation of the recommendations of *Vision and Change*, regardless of institution type. Four of the five rubrics, Curriculum Alignment, Assessment, Infrastructure, and Climate for Change, show no statistical differences by institution type (Table 6). There are statistical differences by institution type for the Faculty Practice/Faculty Support rubric (Figure 3, A–C, and Table 6) and some benefit to LA institutions via the weighting scheme. However, overall, each institution type has the potential to receive any score (Figure 5), and thus, as a whole, the V&C Rubrics do not show institutional bias.

Curriculum Alignment

The Curriculum Alignment rubric addresses the degree to which the core concepts and competencies of *Vision and Change* are integrated throughout the life sciences curriculum. For the 26 institutions that reported complete rubric data sets, the majority had the highest scores on this rubric (Figure 1 and Tables 2 and 5). In addition, higher scores were achieved on the core concepts section and lower scores were reported for the core competencies. Of all of the *Vision and Change* recommen-

dations, the core concepts are probably the least controversial, because they focus on specific biological topics that are generally agreed upon. However, many schools report lower scores for the “systems” concept. One possible explanation is that an understanding of biological systems often requires a deep understanding of biological concepts and mathematical relationships and models, as well as higher-level cognitive skills. These skills are not easily acquired by undergraduate students and require repeated practice and feedback (Ambrose et al., 2010). Roadblocks to implementation of experiences to address systems biology may include a lack of faculty expertise in this area and/or a lack of emphasis on the development of higher-order cognitive abilities.

Alternatively, this may indicate a gap in curriculum development efforts. Henderson and Dancy (2011) report that most research-based instructional strategies have been developed at elite LA colleges or research universities; these curricula might not be directly transferable to other institution types. With the use of the V&C Rubrics, all institution types can evaluate their life sciences curricula in a systematic manner and identify their specific needs. In addition, curriculum review will inform all those engaged in its development as to which aspects are transferable and which require customization.

Assessment

The Assessment rubric evaluates a department’s emphasis on the development and assessment of student learning outcomes at the course and program level using common course assessment tools and pre- and postcourse assessment tools. Departments across all institution types generally reported lower scores on this rubric, indicating that work on assessment needs to be a priority (Figure 1 and Tables 2 and 5). Few STEM educators at the collegiate level have undergone formal training in the areas of effective teaching pedagogies and their evaluation. To remedy this situation, many disciplinary societies and professional organizations have offered faculty development experiences (Baldwin, 2009). Wieman (2007) contends there is a knowledge base for the development of authentic assessment tools to measure student learning. However, to carry out this work would require a substantial investment of institutional resources. Also, institutional culture has provided little motivation for departments to gather and analyze assessment data and implement pedagogical changes based on their findings. It is expected that scores on this rubric will increase in the future as more institutions are asked to become more reflective about what students are learning and educators begin to use assessment data gathered via validated instruments, such as concept inventories, to strategically examine their pedagogical practice, improve the classroom experience, and increase student learning (Anderson et al., 2002; D’Avanzo, 2008; Smith et al., 2008; Smith and Tanner, 2010; Nadelson and Southerland, 2010; Shi et al., 2010).

Faculty Practice/Faculty Support

The Faculty Practice/Faculty Support rubric evaluates the level of student-centered pedagogies, exposure to inquiry in course work, student access to authentic research experiences, and the extent and diversity of faculty development activities. Overall, LA colleges scored higher than R1 and RC institutions and CCs; the difference in scores between LA colleges and CCs was statistically significant (Figure 3A and Tables 2 and 5). When the

ANOVA was performed at the section level, LA colleges scored higher than both RC institutions and CCs on section A, “student higher-level learning” (Figure 3B). For section B, “learning activities beyond the classroom,” there were additional differences between institution types. LA colleges scored statistically higher than CCs and R1 institutions, and the scores of the RC institutions were also higher than those of CCs (Figure 3C). All of these findings fit with the typical mission of the different institution types. LA colleges are noted for their high teacher-to-student ratios and their emphasis on creative and critical thinking. Additionally, they enrich students’ experiences via faculty-mentored research projects and increased faculty–student interactions (Fortenbury, 2014).

Historically, providing extramural research opportunities for students has been considered outside the mission of CCs. However, as more faculty become informed that undergraduate research experiences are a documented high-impact practice (Kuh, 2008), CCs across the country are beginning to emphasize them and provide their student populations with authentic research programs (Wei and Woodin, 2011; Bangera and Brownell, 2014), such as the Community College Undergraduate Research Initiative (Berrett, 2012; Hensel and Cejda, 2014). This trend is particularly important, as CCs serve student populations more diverse than 4-year colleges (Labov, 2012). Participation in an authentic research experience has been shown to be an effective strategy to lessen the performance gap and increase the retention of students from backgrounds traditionally underrepresented in STEM (American Institutes for Research, 2012).

The main emphasis of R1 and RC institutions is research productivity. As such, support at these institutions for the practices measured by this rubric has traditionally been limited. Many of these institutions are beginning to recognize the importance of student-centered and inquiry-based learning and are now offering programs to help their faculty develop these teaching skills. Some of these institutions have realized that the transition to incorporate evidence-based teaching techniques known to foster student learning will be stimulated by hiring faculty with science education expertise (Bush *et al.*, 2006). It has been reported that departments that have created faculty positions to implement inquiry-based, high-impact practices and evidence-based research practices in their courses have been able to enact change (Wieman *et al.*, 2010).

Infrastructure

The Infrastructure rubric evaluates availability of flexible, reconfigurable teaching spaces, informal learning spaces, technological infrastructure, and well-designed laboratories. Although LA colleges scored highest on this rubric, the scores among the four institution types were fairly close (Table 2). Individual departments do not directly control infrastructure. The personnel, space, and equipment largely reflect the institution’s monetary resources and the commitment of the institution to national education reform efforts. It should be noted that Infrastructure rubric section A, Physical Infrastructure, was the third most important factor in determining differences between institutions (Figure 4).

Climate for Change

The Climate for Change rubric gauges the specificity and clarity of institutional and administrative vision, the effectiveness

of communication, and support for the development and modification of institutional policies and practice. The reported scores by all institutions are relatively low compared with scores on other rubrics (Figure 1 and Tables 2 and 5). Similar to infrastructure, individual departments do not directly control the entire institution’s climate. However, this rubric provides critical insights into whether departments are capable of implementing the recommendations of *Vision and Change*, particularly those that require institutional resources for faculty development and incentives to improve the students’ educational experiences. The culture of an institution needs to be considered for change to be effective (Henderson *et al.*, 2011). Although individual faculty can change their teaching approaches and implement assessment procedures to improve student learning outcomes, change will not be sustainable unless an institution values these efforts and reflects them in their reward systems.

Analysis of Full Rubrics Data Sets

Although many life sciences educators are familiar with the use of rubrics as instruments for assessing student work (e.g., AAC&U Valid Assessment of Learning in Undergraduate Education [VALUE] Rubrics), there are few rubrics available that evaluate departmental activities, and even fewer that measure institutional change. Recently, there has been some movement in this arena, as the National Center for Engineering Pathways to Innovation—Epicenter—has begun to address institutional change in engineering education (Nilsen *et al.*, 2015). Epicenter reports that the V&C Rubrics were influential in developing their tool. Similar to the V&C Rubrics, the Epicenter tool will enable the collection of an extensive data set from varied institution types that will inform large-scale improvement in undergraduate education.

The analysis of the 26 full data sets across various institution types has provided baseline knowledge and insights about the state of the adoption of the recommendations of the *Vision and Change* report. Some institutions have made more progress than others (Figure 5). Factors affecting the extent of progress may be the level of institutional commitment to change, the willingness of faculty to embrace new ideas about the student experience in life sciences education, and the support faculty receive to change their current practice.

Implications for STEM Transformation

Watkins and Mazur (2013) reported that the reasons students leave science majors at 4-year institutions include a lack of student–faculty interaction in the classroom and presentation of content in a manner that fails to engage the students. To retain students in STEM majors, Suchman (2014) recommends that institutions assign tenure-track faculty to teach introductory courses, as these faculty tend to be more invested in the institution. Active learning has been documented to increase student performance (Freeman *et al.*, 2014). The V&C Rubrics have taken this into account and reflect the importance of faculty use of validated tools to record the time students spend engaged in active-learning activities (Smith *et al.*, 2013; Wieman and Gilbert, 2014; Eddy *et al.*, 2015). Faculty will be able to assess the quantity and quality of the active learning taking place in their classes as they use these tools. As studies on active-learning techniques continue, this evidence will assist in determining

which specific active-learning techniques produce the greatest learning gains.

Providing incentives will help motivate faculty to learn more about evidence-based teaching practices and the cognitive science that supports such practices. Faculty evaluation metrics that take into account and reward use of best practices would also stimulate change in faculty teaching practice. These structural changes would motivate faculty to develop courses with active, collaborative, and inquiry-based learning. The V&C Rubrics can be used to document changes in the teaching practices of individual faculty members over time and to help motivate departments to initiate and sustain change through benchmarking progress and encouraging department-level reflection and discussion.

Research universities have been reported as having the most difficulty in changing their educational practices (Anderson *et al.*, 2011). The typical culture of these institutions places teaching and research in direct competition, with the status and progress of faculty members almost exclusively dependent on their ability to conduct research and acquire grant funding. However, teaching and research are equally valuable pursuits, as both are capable of generating new knowledge (Boyer, 1990). Research universities excel at postbaccalaureate education, conducting scientific research, and training new scientists, and historically have placed less emphasis on the development of their faculty as educators and on their work with undergraduates. Until chairpersons, deans, and college/university presidents increase the value placed on evidence-based, student-centered pedagogies, teaching will continue to be undervalued at these institutions. Although research universities seem to be viewed as having the greatest number of obstacles to transforming teaching and learning for undergraduates, the data suggest that all institutions are facing challenges. The V&C Rubrics provide an avenue for faculty to start conversations about the status of teaching and learning in their departments, reflect on accomplishments and opportunities for improvement, and determine their departments' future directions.

The magnitude and importance of the recommendations called for in the *Vision and Change* report have caused some authors to wonder whether the life sciences and larger STEM communities are up to the task of enacting the vision (D'Avanzo, 2013; Talanquer, 2014). D'Avanzo has specifically called out the lack of "evidence-based, realistic models for actually achieving the desired 'change' broadly." The PULSE V&C Rubrics can be used as a validated framework to evaluate the implementation of *Vision and Change* recommendations.

Overall, change in higher education is challenging. Many faculty are entrenched in the tradition of supplying content in a lecture format (Brownell and Tanner, 2012). College officials in leadership positions too often consider budgetary constraints rather than the current body of knowledge about how students effectively learn science. For improvements in teaching and learning to occur, science chairpersons need to enable faculty to become knowledgeable about effective teaching practices and to provide the time required to change one's teaching approach (Association of American Universities Undergraduate STEM Initiative, 2013). In addition, advocating and maintaining these departmental transformation efforts will require the development of leaders within the faculty ranks (Elrod and Kezar, 2014).

Few models exist that could provide possible schemes to successfully promote departmental and institutional change. Frechtling *et al.* (2015) developed the Innovation through Institutional Integration program (I³), which conducted six case studies on institutions with multiple science education grants. Participating schools submitted documents for review, and the I³ team conducted site visits and interviews. The schools most successful in the implementation and sustainability of their grant-developed programs were those in which high-level administrators were deeply involved. Change in life sciences education will need the support of administrators. The V&C Rubrics can support change by providing an institution's leadership with documentation on how well a particular department has implemented the practices called for in national reports such as *Vision and Change* (AAAS, 2011) and *Engage to Excel* (PCAST, 2012). For transformation to be effective and sustained, change agents must clearly articulate their strategies, collect evidence, and report the effectiveness of these strategies. The V&C Rubrics can supply feedback and assist in the monitoring of change as new directions in a department are sought. This tool is one of the few available measures of departmental transformation.

The V&C Rubrics are widely applicable to all STEM disciplines. Only the Curriculum Alignment rubric is specific for life sciences. For other STEM disciplines, such as chemistry and physics, resources are available from the American Chemical Society and the American Physical Society, respectively, that could be used to assist departments in these STEM disciplines in developing a rubric to measure discipline-specific curricula. All STEM disciplines can use the other four rubrics as a means of departmental and institutional self-reflection and evaluation of current practices. Although institutional effectiveness has been measured (e.g., accreditation by external agencies), these high-stakes evaluations have been slow to promote change. For desired and meaningful change to occur, institutions need to determine what is essential for their transformation using a collaborative and reflective approach. For example, the use of departmental collaborative management has been linked with faculty use of more student-centered instruction (Borrego and Henderson, 2014). When a collaborative approach is used to implement system-wide change, team members are typically more invested, leading to greater chances of success in institutionalizing the structural changes that will support the transformation of STEM curricula and lead to improved student learning outcomes.

A theory of change is a predictive assumption about the relationship between the anticipated changes and the actions that may create those changes (Kezar *et al.*, 2015). Kezar (2001, 2009) has reviewed the multidisciplinary-change research literature and recognized six major theories of change (evolutionary, teleological, life cycle, political, social cognitive, and cultural). Change in higher education is a complex and multifaceted process that requires elements of multiple theories of change to enable deep and complex changes (Kezar, 2009). Additionally, change in higher education needs to be contextualized to the specific institutional setting. Specific criteria of the V&C Rubrics give concrete examples of how to implement and institutionalize change, with several detailing specific structures that will enable change. Furthermore, the development of new structures is a significant element in both the evolutionary and

the teleological (planned change) theories of change. The social cognitive theory of change includes sense-making as an essential element. Sense-making is the process by which people give meaning to experience, and one of the levers for creating new sense is data (Kezar, 2009). Faculty are able to use the V&C Rubrics to gather data regarding the current status of their departments and discuss these with their colleagues. Various criteria of the five rubrics address many elements across these six theories of change, thus enacting features of multiple theories of change simultaneously. As groups of faculty collaborate to complete the rubrics, they will come to understand more completely the context or circumstances of their own institutions, which will better inform their change efforts.

Future Work

The recognition team has recently released a revised set of rubrics, Vision & Change Rubrics, version 2.0, available at www.pulsecommunity.org/page/recognition. Based on feedback from the life sciences community and the data described herein, the rubrics were revised so the criteria were more clearly delineated. Additionally, the instruction manual was revised to provide better guidance on how to complete the rubrics. The revised Vision & Change Rubrics will be used in an ongoing effort to gather additional data about the implementation of *Vision and Change* recommendations, creating a unique longitudinal data set that will track the progress of life sciences department in adopting the *Vision and Change* report recommendations.

As previously described, the V&C Rubrics are composed of 66 criteria. When departments use these rubrics, they are able to obtain a detailed view of their implementation of the recommendations in the *Vision and Change* report in the areas of Curriculum Alignment, Assessment, Faculty Practice/Faculty Support, Infrastructure, and Climate for Change. Departments may find it difficult to begin this self-reflective process. However, the authors of this paper are confident that the process is worth conducting, as information revealed to a department can be used to strategically guide future directions of the department and the institution. The V&C Rubrics were intentionally created to be highly detailed to enable STEM departments to gather information about their current status, successes, and opportunities for improvement.

Some departments might not be ready to conduct a complete analysis based on the full rubrics. With this in mind, the PULSE recognition team has also created the Vision & Change Snapshot Rubric (Supplemental Material). This abridged version evaluates 17 criteria and is accompanied by instructions to guide its completion. These criteria reflect elements of all five rubrics and provide an indication of the status of a department in areas significant to adoption of the *Vision and Change* recommendations. The Vision & Change Snapshot Rubric has been used at conferences and regional workshops to help faculty and administrators begin a collaborative, collegial review process that effectively reveals areas of strength and those that need greater attention.

Education research is conducted by a process similar to that of disciplinary research. In recent years, life sciences have focused on the collection and analysis of large data sets. Guided by these research principles, the recognition team is working to collect rubric data from departments throughout the country,

generating a national data set. This will represent one of the first comprehensive data sets in life sciences education and will allow long-term tracking of the progress of transforming life sciences departments nationwide. To create this data set requires the engagement of the science education community at large. Institutions will need to submit their baseline rubric data and then examine their progress by completing the rubrics after departmental change strategies to improve teaching and learning have been implemented. Once analyzed, these data will indicate the degree of national implementation of *Vision and Change*, drive the future directions of STEM education research, and further facilitate the transformation currently underway in classrooms, departments, and across higher education.

ACKNOWLEDGMENTS

This work was supported by NSF grants DBI-1323223 and EF-1350120. We also acknowledge Madison A. Lindstrom for her editorial assistance.

REFERENCES

- Aguirre KM, Balser TC, Jack T, Marley KE, Miller KG, Osgood MP, Pape-Lindstrom PA, Romano SL (2013). PULSE Vision & Change Rubrics. *CBE Life Sci Educ* 12, 579–581.
- Akritas MG (1990). The rank transform method in some two-factor designs. *J Am Statist Assoc* 85, 73–78.
- Ambrose S, Bridges M, DiPietro M, Lovett M, Norman M (2010). *How Learning Works: Seven Research-Based Principles for Smart Teaching*. San Francisco, CA: Wiley.
- American Association for the Advancement of Science (2011). *Vision and Change in Undergraduate Biology Education: A Call to Action*. <http://visionandchange.org/files/2011/03/Revised-Vision-and-Change-Final-Report.pdf> (accessed 24 September 2015).
- American Institutes for Research (2012). *Broadening Participation in STEM: A Call to Action*. www.air.org/sites/default/files/downloads/report/Broadening_Participation_in_STEM_Feb_14_2013_0.pdf (accessed 24 September 2015).
- Anderson DL, Fisher KM, Norman JG (2002). Development and validation of the Conceptual Inventory of Natural Selection. *J Res Sci Teach* 39, 952–978.
- Anderson WA, Banerjee U, Drennan CL, Elgin SCR, Epstein IR, Handelsman J, Hatfull GF, Losick R, O'Dowd DKO, Oivera BM, et al. (2011). Changing the culture of science education at research universities. *Science* 331, 152–153.
- Association of American Colleges and Universities (2010). *Valid Assessment of Learning in Undergraduate Education (VALUE) Rubrics*. www.aacu.org/value-rubrics (accessed 2 November 2016).
- Association of American Universities Undergraduate STEM Initiative (2013). *Framework for Systemic Change in Undergraduate STEM Teaching and Learning*. www.aau.edu/workarea/downloadasset.aspx?id=14357 (accessed 24 September 2015).
- Baldwin RG (2009). The climate for undergraduate teaching and learning in STEM fields. *New Dir Teach Learn* 117, 9–17.
- Bangera G, Brownell SE (2014). Course-based undergraduate research experiences can make scientific research more inclusive. *CBE Life Sci Educ* 13, 602–606.
- Berrett D (2012). With NSF support, research moves into science labs of 2-year colleges. *Chronicle of Higher Education*. <http://chronicle.com/article/With-NSF-Support-Research/130339> (accessed 10 December 2015).
- Borrego M, Henderson C (2014). Increasing the use of evidence-based teaching in STEM higher education: a comparison of eight change strategies. *J Eng Educ* 103, 220–252.
- Boyer E (1990). *Scholarship Reconsidered: Priorities of the Professoriate*. San Francisco, CA: Jossey-Bass.
- Brame CJ, Biel R (2015). Test-enhanced learning: the potential for testing to promote greater learning in undergraduate science courses. *CBE Life Sci Educ* 14, es4.

- Brown JD (2009). Principal components analysis and exploratory factor analysis—definitions, differences and choices. *Shiken JALT Testing & Evaluation SIG Newsl* 13, 26–30.
- Browne MW (2001). An overview of analytic rotation in exploratory factor analysis. *Multivar Behav Res* 36, 111–150.
- Brownell SE, Tanner KD (2012). Barriers to faculty pedagogical change: lack of training, time, incentives, and tensions with professional identity? *CBE Life Sci Educ* 11, 339–346.
- Bush SD, Pelaez NJ, Rudd JA, Stevens MT, Williams KS, Allen DE, Tanner KD (2006). On hiring science faculty with education specialties for your science (not education) department. *Cell Biol Educ* 5, 297–305.
- Connell GL, Donovan DA, Chambers TG (2016). Increasing the use of student-centered pedagogies from moderate to high improves student learning and attitudes about biology. *CBE Life Sci Educ* 15, ar3.
- Conover WJ, Iman RL (1981). Rank transformations as a bridge between parametric and nonparametric statistics. *Am Stat* 35, 124–129.
- Couch BA, Wood WB, Knight JK (2015). The molecular biology capstone assessment: a concept assessment for upper-division molecular biology students. *CBE Life Sci Educ* 14, ar10.
- Cronbach LJ (1951). Coefficient alpha and the internal structure of tests. *Psychometrika* 16, 297–334.
- D'Avanzo C (2008). Biology concept inventories: overview, status, and next steps. *BioScience* 58, 2–7.
- D'Avanzo C (2013). Post-*Vision and Change*: do we know how to change? *CBE Life Sci Educ* 11, 201–202.
- Dolan E (2012). Next steps for *Vision and Change*: moving from setting the vision to change. *CBE Life Sci Educ* 11, 201–202.
- Eddy SL, Converse M, Wenderoth MP (2015). PORTAAL: a classroom observation tool assessing evidence-based teaching practices for active learning in large science, technology, engineering and mathematics classes. *CBE Life Sci Educ* 14, ar23.
- Elrod S, Kezar A (2014). Developing leadership in STEM fields: the PKAL summer leadership institute. *J Leader Stud* 8, 33–39.
- Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychol Methods* 4, 272–299.
- Fortenbury J (2014). The perks of attending a liberal arts college. *USA Today*. <http://college.usatoday.com/2014/02/24/the-perks-of-attending-a-liberal-arts-college> (accessed 23 September 2015).
- Frechtling JA, Merlino FJ, Stephenson K (2015). The call to transform post-secondary STEM educational practices and institutional policies. *Am J Educ Stud* 7, 27–42.
- Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, Wenderoth MP (2014). Active learning increases student performance in science, engineering, and mathematics. *Proc Natl Acad Sci USA* 111, 8410–8415.
- Henderson C, Beach A, Finkelstein N (2011). Facilitating change in undergraduate STEM instructional practices: an analytic review of the literature. *J Res Sci Teach* 48, 952–984.
- Henderson C, Dancy MH (2011). Increasing the impact and diffusion of STEM education innovations. A white paper commissioned for the Characterizing the Impact and Diffusion of Engineering Education Innovations Forum, 7–8 February 2011. www.nae.edu/File.aspx?id=36304 (accessed 23 September 2015).
- Hensel NH, Cejda BD (eds.) (2014). *Tapping the Potential of All: Undergraduate Research at Community Colleges*, Washington, DC: Council on Undergraduate Research. www.cur.org/assets/1/7/tapping_potential_final_web.pdf (accessed 10 December 2015).
- Hotelling H (1933). Analysis of a complex of statistical variables into principal components. *J Educ Psychol* 24, 417–441/498–520.
- Jonasson A, Svingby G (2007). The use of scoring rubrics: reliability, validity and educational consequences. *Educ Res Rev* 2, 130–144.
- Kaiser HF (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika* 23, 187–200.
- Kezar A (2009). Synthesis of scholarship on change in higher education. Cent High Educ Policy Analysis. A white paper commissioned for Mobilizing STEM Education for a Sustainable Future, meeting series held in 2009. <http://mobilizingstem.wceruw.org/documents/synthesis%20of%20scholarship%20on%20change%20in%20he.pdf> (accessed 21 February 2016).
- Kezar A (2001). *Understanding and Facilitating Organizational Change in the 21st Century*, ASHE-ERIC Higher Education Report 28, San Francisco, CA: Jossey-Bass, 147.
- Kezar A, Gehrke S, Elrod S (2015). Implicit theories of change as a barrier to change on college campuses: an examination of STEM reform. *Rev High Educ* 38, 479–506.
- Kuh GD (2008). Excerpt from *High-Impact Educational Practices: What They Are, Who Has Access to Them, and Why They Matter*, Washington, DC: Association of American Colleges and Universities. www.aacu.org/leap/hips (accessed 24 September 2015).
- Labov JB (2012). Changing and evolving relationships between two- and four-year colleges and universities: They're not your parents' community colleges anymore. *CBE Life Sci Educ* 11, 121–128.
- Momsen J, Offerdahl E, Kryjevskaja M, Montplaisir L, Anderson E, Nate Grosz N (2013). Using assessments to investigate and compare the nature of learning in undergraduate science courses. *CBE Life Sci Educ* 12, 239–249.
- Nadelson LS, Southerland SA (2010). Development and preliminary evaluation of the measure and understanding of macroevolution: introducing the MUM. *J Exper Educ* 78, 151–190.
- National Research Council (NRC) (2003). *BIO2010, Transforming Undergraduate Education for Future Biologists*, Washington, DC: National Academies Press. www.nap.edu/catalog/10497 (accessed 24 November 2015).
- NRC (2011). *Expanding Underrepresented Minority Participation: American's Science and Technology Talent at the Crossroads*, Washington, DC: National Academies Press. www.nap.edu/read/12984 (accessed 24 November 2015).
- National Science and Technology Council (2011). *The Federal Science, Technology Engineering and Mathematics (STEM) Education Portfolio: A Report from the Federal Inventory of STEM Education Fast Track Action Committee on STEM Education*. www.whitehouse.gov/sites/default/files/microsites/ostp/costem_federal_stem_education_portfolio_report.pdf (accessed 24 November 2015).
- National Science and Technology Council (2013). *Federal Science Technology Engineering and Mathematics (STEM) Education 5-Year Strategic Plan*. www.whitehouse.gov/sites/default/files/microsites/ostp/stem_stratplan_2013.pdf (accessed 24 November 2015).
- Nilsen EA, Besterfield-Sacre M, Monroe-White T (2015). Landscape analysis as a tool in curricular change process. Paper presented at the Frontiers in Education Conference, 21–24 October 2015, in El Paso, TX. www.researchgate.net/publication/282158405_Landscape_Analysis_as_a_Tool_in_the_Curricular_Change_Process (accessed 1 December 2015).
- Pape-Lindstrom P, Jack T, Miller K, Aguirre K, Awong-Taylor J, Balsler T, Brancaccio-Taras L, Marley K, Osgood M, Peteroy-Kelly M, et al. (2015). PULSE pilot certification results. *J Microbiol Biol Educ* 16, 127–129.
- President's Council of Advisors on Science and Technology (2012). *Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics*. www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-engage-to-excel-final_feb.pdf (accessed 24 November 2015).
- Russell S, Hancock MP, McCullough J (2007). Benefits of undergraduate research experiences. *Science* 316, 548–549.
- Shi J, Wood WB, Martin JM, Guild NA, Vicens Q, Knight JK (2010). Diagnostic assessment for introductory molecular and cell biology. *CBE Life Sci Educ* 9, 453–461.
- Smith JI, Tanner KD (2010). The problem of revealing how students think: concept inventories and beyond. *CBE Life Sci Educ* 9, 1–5.
- Smith MK, Jones FHM, Gilbert SL, Wieman CE (2013). The classroom observation protocol for undergraduate STEM (COPUS): a new instrument to characterize university STEM classroom practices. *CBE Life Sci Educ* 12, 618–627.
- Smith MK, Wood WB, Knight JK (2008). The Genetics Concept Assessment: a new concept inventory for gauging student understanding in genetics. *CBE Life Sci Educ* 7, 422–430.
- Suchman EL (2014). Changing academic culture to improve undergraduate STEM education. *Trends Microbiol* 22, 657–659.

- Suhr D (2005). Principal component analysis vs. exploratory factor analysis. SUGI 30 Proceedings.
- Talanquer V (2014). DBER and STEM education reform: are we up to the challenge? *J Res Sci Teach* 51, 809–819.
- U.S. Congress Joint Economic Committee (2012). STEM Education: Preparing for the Jobs of the Future. www.jec.senate.gov/public/_cache/files/6aaa7e1f-9586-47be-82e7-326f47658320/stem-education-preparing-for-the-jobs-of-the-future-.pdf (accessed 24 September 2015).
- Vilorio D (2014). STEM 101: intro to tomorrow's jobs. *Occup Outlook Q.* www.bls.gov/careeroutlook/2014/spring/art01.pdf (accessed 10 December 2015).
- Watkins J, Mazur E (2013). Retaining students in science, technology, engineering, and mathematics majors. *J Coll Sci Teach* 42, 36–41.
- Wei CA, Woodin T (2011). Experiences in biology: alternatives to the apprenticeship model. *CBE Life Sci Educ* 10, 123–131.
- Wieman C (2007). Why not try a scientific approach to science education? *Change* 39, 9–15.
- Wieman C (2014). Large-scale comparison of science teaching methods sends clear message. *Proc Natl Acad Sci USA* 111, 8319–8320.
- Wieman C, Gilbert S (2014). The Teaching Practices Inventory: a new tool for characterizing college and university teaching in mathematics and science. *CBE Life Sci Educ* 13, 552–569.
- Wieman C, Perkins K, Gilbert S (2010). Transforming science education at large research universities: a case study in progress. *Change* 42, 6–14.
- Woodin T, Fesser J, Herrera J (2012). The vision is set, now help chronicle the change. *CBE Life Sci Educ* 11, 347–350.