Working Paper (January 11, 2010)

State of Undergraduate STEM Education as Project Kaleidoscope Affiliates with AAC&U

Daniel F. Sullivan, President Emeritus, St. Lawrence University
Thomas F. Nelson Laird, Assistant Professor, Indiana University, Bloomington
Christine Zimmerman, Director of Institutional Research, St. Lawrence University

This January, 2010 annual meeting of AAC&U marks the official affiliation and integration into AAC&U of Project Kaleidoscope (PKAL), the twenty-year-old national collaboration and informal network of higher education campus leaders and faculty devoted to undergraduate STEM education improvement. We thought to mark this occasion with some reflections and new analyses of data on the state of undergraduate STEM education in America as a contribution to priority setting for the next stage of PKAL's life within AAC&U.

PKAL's Origins

Project Kaleidoscope (PKAL) began in the fall of 1989 in response to a request and charge from the National Science Foundation to Jeanne Narum, new head of the Independent Colleges Office in Washington, to bring together administrative and faculty leaders of highly science-active liberal arts colleges to outline and propose a plan of action for reform and improvement of undergraduate science and mathematics education in the liberal arts college setting. This work benefited not just from NSF support but received additional funding from the Exxon, Pew, Kellogg and Dreyfus Foundations. One of us (Sullivan) was among those college leaders and eventually became chair of the PKAL Executive Committee.

In February of 1991 PKAL held a National Colloquium at the National Academy attended by over 600 people from colleges and universities from across the country, federal and private funding agencies and educational associations to discuss its preliminary plan, to receive input, and to create new energy for the work of reform ahead. In June of that year PKAL published its report—What Works: Building Natural Science Communities, PKAL Volume One—data, analyses and conclusions drawn from experience of what leads to success in undergraduate STEM student learning. In our view that report remains today perhaps the single clearest and most inspiring case that we know of for the powerful effects of pedagogies of engagement on STEM student learning outcomes.

In a letter to Walter Massey, new Director of the NSF, letting him know that the report was on its way Sullivan said:

> We are convinced that science and mathematics education works wherever it takes place within an active community of learners, where students work collaboratively in groups of manageable size, and where faculty are deeply committed to teaching, devoted to student success, and convinced that all students can learn. It works where learning is active, hands-on, investigative, and experiential, and where the curriculum is rich in laboratory experiences, steeped in the methods of scientific research as it is practiced by professional scientists. This approach works for women, for minorities, for all students.[1]

In the aftermath of the report's publication the PKAL leaders committed themselves with Jeanne Narum at the helm to giving leadership nationally, not just in liberal arts colleges, to help faculty and institutional leaders come to appreciate the superior impact on student learning of these high-impact, engaged learning pedagogies and then to adopt them institution-wide. You

[1] Project Kaleidoscope, "What Works: Building Natural Science Communities—A Plan for Strengthening Undergraduate Science and Mathematics," Washington, DC, 1991, XI.

can see why the affiliation of PKAL with AAC&U, highlighted at this AAC&U annual meeting, is such a natural—PKAL is to undergraduate STEM education what AAC&U is to all of undergraduate liberal education.

20 Years Later

So where are we after 20 years as PKAL begins a new life within AAC&U? In June and October of 2008 The National Academies National Research Council Board of Science Education convened two workshops to bring together and assess the large base of evidence now available on what works in undergraduate STEM education and the progress of reform. A must read from those workshops for anyone interested in improving undergraduate STEM student learning is the summary paper by James Fairweather.[2] Here is just a small sample of what he says:

> *Do we need more evidence about the effectiveness of active and collaborative teaching strategies and related efforts to foster student engagement in their own learning?* Here the answer is a definitive *no*. [There is] clear research evidence that active and collaborative instructional pedagogies are more effective than traditional lecture and discussion across most if not all dimensions of student learning.[3]

---

[2] James Fairweather, "Linking Evidence and Promising Practices in Science, Technology, Engineering, and Mathematics (STEM) Undergraduate Education: A Status Report for The National Academies National Research Council Board of Science Education," Center for Higher and Adult Education, Michigan State University, 2009. See also Jay B. Labov, Susan R. Singer, Melvin D. George, Heidi A. Schweingruber, and Margaret L. Hilton, "Effective Practices in Undergraduate STEM Education Part 1: Examining the Evidence," CBE—Life Sciences Education, Vol. 8, Fall 2009, 157-161.

[3] Ibid, Fairweather, 11.

> [And] the pedagogical strategies most effective in enhancing student learning outcomes are not discipline dependent. Instead, active and collaborative instruction coupled with various means to encourage student engagement invariably lead to better student learning outcomes irrespective of academic discipline.[4]
>
> The key to improving STEM undergraduate education lies in getting the majority of STEM faculty members to use more effective pedagogical techniques than is now the norm in these disciplines. Additional research evidence will play only a small role in this process.[5]

Fairweather's review is a striking validation of the claims made in the 1991 PKAL report.

But a recent paper by Nelson Laird et al. brought some sobering news.[6] Using data from the National Survey of Student Engagement, or NSSE (80,124 randomly sampled seniors from 517 of all varieties of 4-year colleges and universities) and the Faculty Survey of Student Engagement, or FSSE (10,365 faculty members from 109 institutions representing the full range of 4-year colleges and universities), Nelson Laird et al. examined disciplinary differences in the extent to which students are exposed to educational environments that promote what they call "deep approaches to learning. Students who use deep approaches to learning tend to perform

---

[4] Ibid., 4.

[5] Ibid., 13.

[6] Thomas F. Nelson Nelson Laird, Rick Shoup, George D. Kuh, and Michael J. Schwartz, "The Effects of Discipline on Deep Approaches to Student Learning and College Outcomes," Research in Higher Education (2008) 49: 469-494.

better as well as retain, integrate, and transfer information at higher rates than students using surface approaches to student learning."[7] Let us give you just a few highlights:

- Despite the resounding affirmation in Fairweather of the positive impact of pedagogies of engagement on student STEM learning, Nelson Laird et al. found—using models with extensive statistical controls—that, nationally, STEM faculty use pedagogies that encourage higher-order, integrative, and reflective learning[8] significantly less than faculty in non-STEM fields and therefore STEM students experience "deep approaches to learning" less than students in non-STEM fields.[9] The differences were small for "higher-order learning," the scale that is concerned with analysis, synthesis, and judgment regarding evidence—relatively good news—but quite large for the integrative and reflective learning scales. Expecting "deep learning" of all three kinds from undergraduate study is, we believe, a high but appropriate standard for all disciplines. Indeed, these learning goals are central components of the Essential Learning Outcomes AAC&U strongly advocates for what the focus of liberal

---

[7] Ibid., 470.

[8] The **Higher-Order Learning** scale comes from these NSSE and FSSE questions—Course work emphasized: analyzing the basic elements of an idea, experience, or theory; emphasized synthesizing and organizing ideas, information or experiences into new, more complex interpretations and relationships; emphasized making judgments about the value of information, arguments, or methods. **Integrative Learning**: Worked on a paper or project that required integrating ideas or information from various sources; Included diverse perspectives (different races, religions, genders, political beliefs, etc.) in class discussions or writing assignments; Put together ideas or concepts from different courses when completing assignments or during class discussions. **Reflective Learning**: Examined the strengths and weaknesses of your own views on a topic or issue; Tried to better understand someone else's views by imagining how an issue looks from his or her perspective; Learned something that changed the way you understand an issue or concept.

[9] "The average student in a hard field [essentially STEM] used deep approaches to learning nearly a quarter of a standard deviation less than the average senior in a soft field. For faculty, the difference was nearly three quarters of a standard deviation." Ibid., 484.

education should be. These national results for STEM faculty and STEM student majors are disappointing.

- Consistent with Fairweather's summary, Nelson Laird et al. also found ". . . . . relative consistency across disciplines in the strength of the relationships between deep approaches to learning and . . . three student outcomes."[10] Nelson Laird et al. emphasize that even in fields with a high degree of consensus on paradigms and underlying knowledge base—fields where faculty often insist on high levels of memorization by students and where students, consequently, often use shallow short-term approaches to learning that result in low retention of information—the strength of the relationship between deep approaches to learning and student outcomes is the same as in lower-consensus disciplines.

For those of us who have committed significant portions of our time in the last two decades to undergraduate STEM reform, it is discouraging that, nationally, faculty in STEM fields have expectations that students exhibit integrative learning and reflective learning significantly less often than faculty in other disciplines. Let's pause for a moment to consider what these STEM majors are missing. The integrative learning scale is made from seniors' responses to three questions on the NSSE survey—worked on a paper or project that required integrating ideas or information from various sources; included diverse perspectives (different races, religions, genders, political beliefs, etc.) in class discussions or writing assignments;[11] and

[10] Ibid., 481.

[11] We are aware that some STEM faculty may find including this question in the integrative learning scale problematic. We do so for several reasons. It was included in the integrative learning scale in the Nelson Laird et al. paper, so we keep it for continuity. Second, while the argument is too complex to make in detail in this paper, we know that some premises that find their way into scientific arguments have non-science sources. Much of the best of the feminist critique of positivism is about this, and see the seminal works of Gerald Holton—*Thematic Origins of Scientific Thought* (Harvard University Press, 2nd ed., 1988) and *The Advancement of Science, and its Burdens*

put together ideas or concepts from different courses when completing assignments or during class discussions. Two of these questions speak directly to the central recommendations of the seminal National Academies' Board on Life Sciences 2003 report BIO 2010:

> Like research in the life sciences, undergraduate education must be transformed to prepare students effectively for the biology that lies ahead. Life sciences majors must acquire a much stronger foundation in the physical sciences (chemistry and physics) and mathematics than they now get. Connections between biology and the other scientific disciplines need to be developed and reinforced so that interdisciplinary thinking and work become second nature. Connections within biology are equally important and the relevance of fields such as population biology, plant biology, and cognitive science to biomedical research should not be ignored. Equally important, teaching and learning must be made more active to engage undergraduates, fully prepare them for graduate study, and give them an enduring sense of the power and beauty of creative inquiry. [12]

The Nelson Laird et al. data (and our additional analyses reported below) indicate clearly that STEM seniors—and especially engineering and physical sciences majors, as our analysis shows—lag behind seniors majoring in the other arts and sciences disciplines with respect to developing this critically important form of thinking.

(Harvard University Press, 1998). Discussing these premises enriches scientific inquiry. So we include the question on that ground as well. Finally, while the STEM/non-STEM differences are smaller with that question removed from the scale, they are still present. So we keep the question in on that ground as well.

[12] Committee on Undergraduate Biology Education to Prepare Research Scientists for the 21st Century, Board on Life Sciences, Division on Earth and Life Studies, National Research Council of the National Academies, *BIO 2010: Transforming Undergraduate Education for Future Research Biologists*, The National Academies Press, Washington, D.C., 2003, 1-2.

The reflective learning scale is a combination of seniors' responses to these three questions—examined the strengths and weaknesses of your own views on a topic or issue; tried to better understand someone else's views by examining how an issue looks from his or her perspective; learned something that changed the way you understand an issue or concept. The kinds of intellectual self-reflection skills these questions are about are surely as important in the STEM disciplines as they are in other disciplines, but we see that STEM majors experience far fewer opportunities to develop these skills than majors in the other disciplines. Indeed, one might argue that it is especially in the STEM disciplines that students should acquire these skills given the way in science empirical evidence tends to be seen as harder evidence than the evidence for arguments in other disciplines. Discovering a bad premise or assumption is just as important in STEM disciplines as in other disciplines.

The Nelson Laird et al. findings caused us to want to look more deeply at the learning experiences STEM and non-STEM majors have in college. Are the results different for students attending different kinds of colleges and universities? Are there types of institutions where the mean scores for each of the three deep learning scales do not differ by seniors' disciplinary major?

Looking Deeper

For our deeper analysis we used responses to the 2008 NSSE survey from 614 institutions and 73,606 seniors. The disciplinary[13] breakdown is:

Biological Sciences	11,315 (15.4%)
Engineering & Phys Sciences	16,113 (21.9%)
Arts & Humanities	22,871 (31.1%)
Social Sciences	23,307 (31.6%)

Let's look quickly in Table 1, before we examine differences by institutional type, at the STEM/non-STEM differences on which Nelson Laird et al. focused as they exist in the 2008

Table 1. Overall Senior Means by Field[14]

	STEM			
	Biological Sciences	**Engineering & Physical Sciences**	Arts & Humanities	Social Sciences
Deep Approaches to Learning				
Higher-Order Learning	**-0.01**	**-0.05**	-0.01	0.05
Integrative Learning	**-0.20**	**-0.43**	0.15	0.25
Reflective Learning	**-0.14**	**-0.30**	0.14	0.13
Academic Effort	**0.13**	**0.06**	-0.01	-0.09
Amount of Reading and Writing	**-0.17**	**-0.38**	0.16	0.19
Active and Collaborative Learning	**0.00**	**0.00**	-0.01	0.01
Student-Faculty Interaction	**0.03**	**-0.15**	0.08	0.02
Enriching Educational Experiences	**0.02**	**-0.21**	0.04	0.10
Supportive Campus Environment	**0.03**	**-0.07**	0.01	0.03
Academic and Personal Gains	**-0.01**	**-0.09**	-0.04	0.11
Satisfaction	**0.00**	**-0.01**	-0.02	0.03

Variables standardized across entire sample.

[13] For purposes of this analysis, we excluded business and education majors to focus on traditional arts and sciences disciplines and because some institutions in our sample do not offer these majors.

[14] For ease of understanding by those not familiar with statistical modeling involving multiple controls, we present our data through a simple comparison of means. We have run extensive models controlling for institutional control (public/private) and various student characteristics: gender, race/ethnicity, international status, parents' level of education, enrollment status and, if living on campus, Greek membership, and intercollegiate athletic participation. For models run with St. Lawrence data, we also controlled for SAT and high school GPA, and the differences we discuss in this table and subsequent tables remain.

NSSE senior survey. What you see in Table 1 are deviations from the means (expressed in the number of standard deviations) for the entire sample, where the total sample means are set to zero.[15] With the exception that differences by discipline in the higher-order learning scale means reported in Nelson Laird et al. are further reduced in our data, the differences they reported in the integrative and reflective learning scales that we highlighted above persist—STEM majors experience environments that encourage integrative and reflective learning significantly less than non-STEM majors do, with engineering and physical sciences majors having the lowest scores. Goods news in Table 1 is that for several variables STEM/non-STEM means are essentially the same: active and collaborative learning, supportive campus environment, academic and personal gains, and overall satisfaction with college. STEM seniors report somewhat higher levels of academic effort and, not surprisingly, a lot less reading and writing. Notable (though we won't try to provide an in depth analysis here) is that engineering and physical sciences majors report the lowest levels of student-faculty interaction and enriching educational experiences.

That is the overall picture that Nelson Laird et al. alerted us to. Are the results different for students attending different kinds of colleges and universities? In a word, yes, there are

[15] Active and Collaborative Learning, Student-Faculty Interaction, Enriching Educational Environment, and Supportive Campus Environment are standard NSSE benchmarks. The fifth standard benchmark—Level of Academic Challenge—has been pulled apart for this analysis since some of the questions for which responses are included in this benchmark are also part of Nelson Laird et al.'s "higher-order learning" scale. The questions remaining we have put into two groupings: "Academic Effort" and "Amount of Reading and Writing required." Then, to eliminate overlap with other scales, we have removed two items from Nelson Laird et al.'s "Integrative Learning" scale so that it includes only "worked on a paper or project that required integrating ideas or information from various sources," and "put together ideas or concepts from different courses when completing assignments or during class discussions." Our "Academic and Personal Gains" scale is the same as in Nelson Laird, and "Satisfaction" is the response to "How would you evaluate your entire educational experience at this institution?" and, "If you could start over again, would you go to the same institution you are now attending?"

significant differences by Carnegie Classification institutional type[16]—some of them large—in the extent to which seniors experienced pedagogies that we know lead to higher-order learning, integrative learning, and reflective learning, in how much effort students put into their academic work, in the academic and personal gains they report, and in their satisfaction with college. With

Table 2. Overall Senior Means by Institutional Type

	Research	Doctoral/ Research	Master's - Large	Master's - Medium	Master's - Small	Bac -Arts & Sci	Bac - Diverse
Deep Approaches to Learning							
Higher-Order Learning	-0.10	0.03	-0.04	-0.01	0.02	0.18	-0.02
Integrative Learning	-0.17	0.07	0.00	0.00	0.07	0.19	0.00
Reflective Learning	-0.07	0.05	-0.01	-0.02	0.02	0.10	-0.01
Academic Effort	-0.08	-0.12	-0.07	-0.03	-0.04	0.24	-0.01
Amount of Reading and Writing	-0.13	0.04	-0.08	-0.06	-0.04	0.30	0.01
Active and Collaborative Learning	-0.21	0.09	0.01	0.05	0.16	0.15	0.19
Student-Faculty Interaction	-0.26	0.04	-0.03	0.05	0.17	0.26	0.22
Enriching Educational Experiences	-0.10	0.07	-0.22	-0.13	0.01	0.46	0.03
Supportive Campus Environment	-0.20	-0.04	-0.03	0.02	0.15	0.23	0.22
Academic and Personal Gains	-0.11	0.02	-0.05	0.00	0.09	0.17	0.10
Satisfaction	-0.04	-0.08	-0.10	0.01	-0.03	0.20	0.01

Variables standardized across entire sample.

one exception—Enriching Educational Experiences—students least likely to experience high-impact practices attended research universities, while, again with one exception—Active and Collaborative Learning—those most likely to experience high-impact practices attended baccalaureate arts and sciences colleges. With regard to the three so-called "deep learning" scales—Higher-Order Learning, Integrative Learning, and Reflective Learning—seniors least likely to experience practices leading to "deep learning" attended research universities and those

[16] Unfortunately, we are unable to include community colleges—a very important group of institutions within AAC&U—as a category in this analysis because two-year programs, since they have no "seniors", are not surveyed with NSSE.

most likely to experience them attended baccalaureate arts and sciences colleges. With regard to Academic Effort the doctoral/research universities were lowest, and with regard to Amount of Reading and Writing and Academic and Personal Gains the research universities were the lowest while the baccalaureate arts and sciences colleges were the highest on all three. Finally, overall student satisfaction was the lowest for seniors who attended large master's universities and the highest for seniors who attended baccalaureate arts and sciences colleges.

The differences between the highs and lows were at least .5 standard deviations in two comparisons—Student-Faculty Interaction and Enriching Educational Experiences—at least .4 standard deviations in two comparisons—Amount of Reading and Writing and Supportive Campus Environment—and at least .3 standard deviations in three more—Integrative Learning, Active and Collaborative Learning, and Overall Student Satisfaction. These differences remain when we look across institutional types by disciplinary major. In each category of discipline, and with only a few exceptions, the baccalaureate colleges are the highest and the research universities the lowest (see, for example, Tables 2a and 2b for the biological sciences seniors and engineering and physical sciences majors across institutional types).

The biggest differences, however, are not on the deep learning scales: only for integrative learning is the difference larger than .3 standard deviations. While the variation across institutional types is very large for several kinds of student engagement, differences among types of institutions in expectations of students in the areas of deep learning we are analyzing are smaller. These are not new findings, of course—differences on these scales by institutional type have been evident from the very first administration of NSSE—but they have been under-

Table 2a. Biological Science Senior Means by Institutional Type

	Research	Doctoral/ Research	Master's - Large	Master's - Medium	Master's - Small	Bac -Arts & Sci	Bac - Diverse
Deep Approaches to Learning							
Higher-Order Learning	-0.18	-0.01	0.00	-0.03	0.09	0.17	0.02
Integrative Learning	-0.36	-0.24	-0.18	-0.26	-0.10	0.01	-0.13
Reflective Learning	-0.16	-0.20	-0.16	-0.14	-0.06	-0.08	-0.18
Academic Effort	0.02	0.06	0.08	0.12	0.13	0.34	0.10
Amount of Reading and Writing	-0.30	-0.15	-0.24	-0.23	-0.16	0.05	-0.10
Active and Collaborative Learning	-0.32	0.05	0.03	0.03	0.21	0.25	0.21
Student-Faculty Interaction	-0.27	-0.01	0.04	0.07	0.27	0.28	0.23
Enriching Educational Experiences	-0.09	-0.08	-0.20	-0.12	0.13	0.42	0.08
Supportive Campus Environment	-0.20	-0.09	0.00	0.02	0.20	0.30	0.24
Academic and Personal Gains	-0.13	-0.03	-0.08	0.00	0.09	0.15	0.04
Satisfaction	-0.07	-0.15	-0.10	0.01	-0.06	0.22	-0.03

Variables standardized across entire sample.

Table 2b. Engineering and Physical Science Senior Means by Institutional Type

	Research	Doctoral/ Research	Master's - Large	Master's - Medium	Master's - Small	Bac -Arts & Sci	Bac - Diverse
Deep Approaches to Learning							
Higher-Order Learning	-0.06	-0.10	-0.06	-0.07	-0.11	0.05	-0.08
Integrative Learning	-0.49	-0.36	-0.44	-0.43	-0.38	-0.29	-0.37
Reflective Learning	-0.34	-0.25	-0.29	-0.37	-0.25	-0.19	-0.30
Academic Effort	0.08	-0.07	0.01	0.00	-0.03	0.23	-0.02
Amount of Reading and Writing	-0.39	-0.30	-0.43	-0.42	-0.51	-0.24	-0.31
Active and Collaborative Learning	-0.15	0.15	0.01	0.04	0.11	0.24	0.28
Student-Faculty Interaction	-0.36	-0.08	-0.16	-0.08	0.12	0.24	0.16
Enriching Educational Experiences	-0.21	-0.22	-0.42	-0.40	-0.24	0.24	-0.11
Supportive Campus Environment	-0.21	-0.05	-0.08	-0.04	0.11	0.23	0.18
Academic and Personal Gains	-0.14	-0.10	-0.16	-0.14	-0.09	0.16	0.07
Satisfaction	-0.03	-0.16	-0.10	0.01	-0.14	0.21	0.08

Variables standardized across entire sample.

discussed in analyses of NSSE data, in our view. Perhaps they have been under-discussed because non-baccalaureate arts and sciences institutions and policy makers attribute the higher scores of baccalaureate arts and sciences colleges to differences in institutional wealth, on the assumption that high-impact practices are expensive and that all baccalaureate arts and sciences colleges are wealthier than institutions of other types. The latter is, of course, not true across the board though some baccalaureate arts and sciences colleges are indeed wealthy. Or perhaps the assumption is that student selectivity at the baccalaureate arts and sciences colleges assures students who are already deep learners, but there is great diversity in student body quality in our sample of baccalaureate arts and sciences—we do not have an elite sample—and in the other institutional types. As we will see below institutional selectivity **is** associated with higher scores on the deep learning scales, but seniors at some less selective institutions also have high scores on these scales. Or perhaps it is because we have just given up nationally on any hope that universities with major research missions will ever apply the same ingenuity and commitment to undergraduate education that they give to research.

George Kuh, Alex McCormick and other NSSE staff caution that, in general, the variance on most NSSE variables within institutions is greater than the variance across institutional types,[17] which means that all institutions have much room for improvement, but these differences are striking nonetheless. It is clear that there are institutional settings—baccalaureate arts and sciences colleges are one—where the usual practices lead to very high levels of student

[17] National Survey of Student Engagement (2008). *Promoting engagement for all students: The imperative to look within*. Bloomington, IN: Indiana University Center for Postsecondary Research.

engagement, which we know leads to high levels of student learning generally, and to greater opportunities for deep learning.

At the same time, below we will highlight examples of private and public institutions in all Carnegie Classifications that have very high scores on the indicators we are analyzing here to further emphasize that at least some of the differences by institutional type that we see here are more a matter of individual institutional emphasis and commitment and that throughout the full wonderful diversity of American higher education high levels of student engagement are possible. As with most things in teaching and learning, if faculty expect and demand higher-order, integrative, and reflective learning of students, they will come through.

Where are the Successes, Especially with Regard to Deep Learning?

Are there campuses of all institutional types where students experience high levels student engagement and expectations for higher-order learning overall and where STEM majors' scores on these measures are similar to the scores of non-STEM majors? To answer this question we switch to a kind of profile analysis (see Table 3 below), selecting the 100 institutions from our total sample of 614 that had the highest means for each of five measures of special interest here—higher-order learning, integrative learning, reflective learning, and active and collaborative learning and student-faculty interaction, two critical high-impact practices. Next we sorted these institutions by their STEM effect sizes to identify those institutions in the top 100 on each variable that had small (greater than -0.1 and less than 0.1) STEM/non-STEM differences—looking, in other words, for institutions with a single culture of engagement (see Table 4).[18]

What we see in Table 3 is that research universities are under-represented on all measures—indeed, no research universities make the top 100 for higher-order learning or student-faculty interaction. Baccalaureate arts and sciences institutions, on the other hand, are generally overrepresented in the top 100 on all measures. Baccalaureate Diverse institutions are underrepresented among the top 100 institutions for the deep approaches to learning measures and overrepresented among the top 100 for student-faculty interaction. Private institutions are greatly over-represented among the top 100 institutions on all measures. Very competitive and above institutions are overrepresented on the deep approaches to learning measures, but they are not overrepresented on active and collaborative learning and student-faculty interaction—two high-impact practices that we know are associated with improved student outcomes. This deserves more comment and analysis than we can do here, so we just note it.

[18] To do this we ran a hierarchical linear model for each variable of interest (higher-order learning, integrative learning, reflective learning, active and collaborative learning, and student-faculty interaction) to estimate an adjusted mean and an adjusted effect size for the STEM major variable for each institution. The model adjusted for the number of students at the institution (a data quality issue—we didn't want to select institutions where the number of NSSE-takers was small) and student characteristics (gender, race, first generation college student status, living on campus, transfer status, foreign citizenship, enrollment status [full-time vs. part-time], Greek affiliation, and STEM/non-STEM major).

Table 3. Distributions of Top 100 Scoring Institutions for Each of Five Engagement Measures by Institutional Characteristics

	All Institutions n=614	Higher Order Learning 100	Integrative Learning 100	Reflectiv e Learning 100	Active and Collaborative Learning 100	Student Faculty Interaction 100	Engaging 10
Carnegie Classification							
Research	11%	0%	4%	7%	1%	0%	0%
Doctoral/Research	4%	6%	6%	6%	4%	0%	0%
Master's L	27%	15%	15%	17%	25%	15%	10%
Master's M	13%	10%	9%	14%	15%	16%	40%
Master's S	7%	7%	5%	5%	12%	8%	10%
Bac/A&S	22%	57%	53%	40%	29%	41%	40%
Bac/Diverse	15%	5%	8%	11%	14%	20%	0%
Sector							
Public	42%	13%	15%	22%	24%	18%	0%
Private	58%	87%	85%	78%	76%	82%	100%
Barron's Selectivity							
Competitive plus or below	64%	33%	43%	46%	70%	69%	90%
Very competitive or above	36%	67%	57%	54%	30%	31%	10%

Table 4. Distribution of STEM/non-STEM Differences on Five Engagement Measures by Institutional Groups

	Higher Order Learning		**Integrative Learning**		**Reflective Learning**		**Active and Collaborative Learning**		**Student-Faculty Interaction**	
STEM/non-STEM Effect Sizes	All 614 Inst	Top 100	All 614 Inst	Top 100	All 614 Inst	Top 100	All 614 Inst	Top 100	All 614 Inst	Top 100
>= 0.3	0%	0%	0%	0%	0%	0%	3%	1%	0%	0%
>= .1 and < 0.3	1%	0%	0%	0%	0%	0%	20%	37%	3%	3%
>-0.1 and <0.1	93%	81%	0%	0%	0%	0%	69%	57%	73%	91%
>-0.3 and <=-0.1	6%	19%	1%	5%	2%	7%	9%	5%	24%	6%
<=-.3	0%	0%	99%	95%	98%	93%	0%	0%	0%	0%

Table 4 shows us that, with regard to integrative and reflective learning, STEM seniors scored more than .3 standard deviations lower than non-STEM seniors in 99% and 98% of all 614 institutions in our sample, respectively. Results are only slightly better for the top 100 institutions on both measures. There are essentially no institutions where there is one culture rather than two on these scales!

For higher-order learning and student-faculty interaction, most institutions in the total sample and the top 100 have small STEM differences. For active and collaborative learning slightly more than half have small STEM differences, but the variance is in favor of STEM disciplines: that is, on this variable STEM seniors score higher than non-STEM seniors. Interestingly, the top 100 institutions were more likely to have negative STEM differences for the higher-order learning variable and less likely to have negative STEM differences for the student-faculty interaction variable.

Only 10 institutions are in the top 100 for higher-order learning, active and collaborative learning, and student-faculty interactions with small STEM differences across all three measures. These "Engaging 10" are all private, **less selective**, and master's or baccalaureate arts and sciences institutions.

What we see in these data are several things to encourage undergraduate STEM education reformers. First of all, we are excited that large fractions of top 100 institutions on three measures—higher-order learning, active and collaborative learning, and student-faculty interaction—have minimal STEM/non-STEM effects, or one culture rather than two. At the same time, for integrative learning and reflective learning, even in the highest performing institutions, the STEM/non-STEM differences we discussed above are very large. Second, with the exception of the research universities, all institutional types, public and private, more and less wealthy are represented in the profiles of high-performing institutions. In addition, the institutions that are most engaging across multiple measures and have small STEM/non-STEM effects (one culture rather than two) are private, but not elite, supporting the notion that institutional commitment may, in the end, be more important than resources.

High Levels of Student Engagement Are Not Just Good for Students: They Are Good for Institutions as Well

We cannot totally put to bed the issue of cost here, but we make one additional argument. While we do not have the data, it may in fact be that the pedagogies of engagement we focus on here in this paper are more expensive to provide than traditional, passive-learning-oriented approaches to teaching and learning. Surely it seems so on surface.[19] However, where we believe they pay for themselves, from the point of view of the overall American higher-education system and students and their families and aside from the educational benefits of deeper learning, is in greater efficiency measured by higher student retention, shorter time to completion of degree, and less need for students to attend multiple institutions to complete a degree.

A very important message that leaders and faculty from institutions of all types should take away from this paper in these very difficult economic times is that the more students experience pedagogies of engagement, and the more their academic work involves expectations that they will engage in deep, as opposed to surface, learning (higher-order learning, integrative learning, and reflective learning), the more satisfied they will be with their college experience. The more satisfied with their college experience they are, the higher their levels of retention.

Using first-year student 2008 NSSE data from St. Lawrence we created a "deep learning" index identical to that used in Nelson Laird et al., combining questions that asked students to rate

---

[19] On the other hand, in its broad-based transformation of its undergraduate STEM courses—especially the introductory courses—in the 1990's, RPI demonstrated that it is possible to convert from large lecture-lab courses to small studio courses and keep total costs—operating and capital (for facilities renovation and equipment) the same while improving learning outcomes. The RPI experience deserves a close look today.

how much their classroom experiences involved expectations of higher-order learning, integrative learning, and reflective learning. We then looked at first- to second-year retention among students in the four quartiles of this scale. Fully 95% of first-year students in the top quartile of this scale returned for their sophomore year, while only 82% of students in the bottom quartile of this scale returned for their sophomore year. High levels of student engagement are clearly good for students and good for their institutions—especially for tuition-dependent institutions. They save families and students money because of faster college completion (reducing opportunity costs—the income foregone by not having a job sooner), and this reduces the annual number of "slots" we must maintain nationally to allow for an overall student enrollment that produces the same output, saving system-wide "production" costs.

Furthermore, in STEM undergraduate education particularly, we still in many (perhaps most) institutions have a "weeding" rather than "cultivating" mentality. In the science- and PKAL-active institutions with which we are familiar, very high percentages of students who take an introductory course in science or mathematics go on to take additional courses. We have these data for St. Lawrence seniors from 2008 and 2009, as one example: 86% of those St. Lawrence seniors who took introductory biology took at least one more biology course; 86% who took introductory chemistry took at least one more chemistry course; and 86% who took calculus went on to take another mathematics course; while fully 95% of students who took an introductory course in one of those disciplines went on to take at least one more science or mathematics course. We have seen data from at least one large institution where less than 10% of students who took an introductory chemistry course went on to take even one more science or mathematics course. Some species lay thousands of eggs to get just a few to survive, and others lay just a few (or have a small number of young) but expend great energy per egg to get them to survive, using perhaps similar amounts of total energy. It takes many more prospective science majors at the beginning of the pipeline to produce the same number of science majors at the end of the pipeline where a weeding culture operates.

Priorities for the Future

Clearly there remains a great deal of work to do if America's colleges and universities are going to get undergraduate STEM education right. It was critically important for PKAL to come into being 20 years ago; it is even more important for it to continue today in the welcoming and synergistic environment it will find within AAC&U. Improvement needs to happen not just relative to non-STEM fields in the areas of integrative and reflective learning, but in overall STEM learning of all kinds, regardless of what happens in non-STEM fields. This is still a critical national priority.

But systemic institutional change is hard, as anyone who has tried it can attest. We find ourselves in agreement with a number of Fairweather's conclusions, which apply to non-STEM undergraduate education as well:

> On balance the research suggests that the greatest gains in STEM education are likely to come from the development of strategies to encourage faculty and administrators to implement proven instructional strategies rather than to carry out additional research on these strategies. . . . . The largest gain in learning productivity in STEM will come from convincing the large majority of STEM faculty that currently teaches by lecturing to use any form of active or collaborative instruction. . . . . .[20]

---

[20] Op. cit., Fairweather, 26.

- External networks of like-minded colleagues outside of one's institution can be important forces in promoting instructional reform. [We (this paper's authors) note that PKAL is exactly such an external network.]

But how, exactly, can we encourage more faculty and institutional leaders to adopt or support the adoption of these proven strategies, and especially to expect more higher-order, integrative and reflective learning on the part of students? We know that good programs of faculty development supported with appropriate institutional resources can help willing faculty adopt proven strategies, and that externally-organized programs such as PKAL can leverage campus-based programs to produce more change faster. Continuing this PKAL work into the future, with external financial support from government and private foundations to PKAL and to campuses, is one important priority. Here are some other ideas:

- Institutions should have a clearly written document for all teaching faculty, not just STEM faculty—"What We Expect"— laying out expectations that faculty will adopt and use, to the extent reasonable in the local context, pedagogies of engagement and high-impact practices, and that they will adopt learning goals for students including higher-order, integrative and reflective learning. Institutions should provide robust mentoring programs for new faculty to support them as they develop their pedagogical tool-kit.
- A similar document should go to all candidates for faculty positions at the institution, who should be asked to describe how their current approach to teaching fits with the institution's expectations and how, in the new institutional context, they would seek to meet these expectations. For a number of years at St. Lawrence, as part of their application for the position, candidates for teaching positions in STEM disciplines have written a short essay

describing their teaching philosophy and explaining how students might collaborate in their research. Candidates giving no evidence of commitment to the kinds of pedagogies of engagement valued at St. Lawrence inevitably received low ratings. The idea is not to hire faculty uncommitted to using best practices, or unwilling to become committed. It is far easier to support a faculty member already pointed in the right direction than to get a faculty member to completely change direction. We did not ask candidates to address how they would help students achieve higher-order, integrative and reflective learning goals, and we should have.

- We agree with Fairweather that institutional faculty reward systems must reward faculty adoption of best practices/proven strategies. This is, of course, easier said than done, especially in institutions with specific, major research missions. We have no magic bullets to offer. We just note that, unless faculty teaching practices change in the kinds of institutions that award the largest numbers of undergraduate degrees, we must essentially concede that we will not transform undergraduate STEM education sufficiently to meet the societal expectations we are facing on this matter. Is that what we want?
- We are of the view that institutional change will happen more rapidly when college and university trustee Academic Affairs Committees, having learned how to be appropriately sensitive to the academic culture in their institutions—perhaps with some help from the Association of Governing Boards, which has high quality writings on best practices for board academic affairs committees[21]—regularly receive and discuss good, carefully selected

[21] See Richard Morrill, *Strategic Leadership in Academic Affairs: Clarifying the Board's Responsibilities* (AGB, 2002) and Peter Ewell, *Making the Grade: How Boards Can Ensure Academic Quality* (AGB, 2006).

educational outcomes assessment data benchmarked, if possible, against the same data from comparable institutions. The idea here is to encourage good conversations among faculty, institutional leaders and trustees on the current state of student learning at the institution and especially, since these data are readily available through instruments like NSSE, the extent to which students experience the pedagogies of engagement that we know lead to better learning outcomes and are expected by faculty to achieve deep learning goals. We know that increased levels of student engagement lead to better educational outcomes, and we know what faculty and institutional practices lead to higher levels of student engagement. When the Board's Academic Affairs Committee begins to monitor institutional progress with regard to student engagement and deep learning, benchmarked against comparable institutions, there will surely be heightened efforts by faculty and institutional leaders to improve student engagement and to have higher expectations regarding deep learning, including seeking the resources to do so.

This suggestion will make many institutional leaders and faculty nervous, we know—boards of trustees can be inappropriately intrusive. But it is the institution's board that has the ultimate fiduciary responsibility for the achievement of its academic mission. Properly educated as to how to exercise that fiduciary responsibility, we believe the institution's board can have a very positive impact on educational outcomes.

- We feel the same way about the potential impact the regional accrediting groups can have. One of us has chaired a number of Middle States visiting teams. His experience leads him to believe that institutions up for re-accreditation take the new expectations regarding assessment very seriously and would benefit greatly from more guidance from the accrediting agencies on exactly how to meet the highly laudable new (over the last decade) standards regarding assessment. Progress is being made in the development of rubrics for the direct assessment of liberal learning goals based on students' best work (see AAC&U's own VALUES Project), though any sort of benchmarking across institutions is always going to be difficult. Nonetheless, accrediting agencies could point institutions to this work and ask them to include this kind of evidence of student learning in their self-studies.

For benchmarking against like institutions with regard to the extent to which the institution is using best practices—those practices that lead to higher-order learning, integrative learning, and reflective learning as well as substantive learning in specific fields—we think institutional leaders, and visiting team members, will find NSSE and similar instruments extraordinarily useful in accreditation. When you are behind your competitors in use of pedagogies of engagement, you will also be behind in retention and in student satisfaction with their learning. This kind of comparative analysis tells institutional leaders, faculty, and trustees what they will need to do to get better, and it lets them see if or how fast things are getting better, and whether things are getting better faster than the competition is getting better. Accrediting agencies could be more specific, in other words, in letting institutions know what they believe is good evidence of educational outcomes. If they do that, institutions will move quickly, in our view, to provide that evidence and they will get better quicker.

Conclusion

Our review of the current state of undergraduate STEM education in America is both sobering and hopeful. There is no doubt that Project Kaleidoscope, a critically important movement 20 years ago, is at least as important today. Nationally, STEM disciplines lag behind the arts and humanities and social sciences in providing several key forms of student engagement and expectations for integrative and reflective learning. But we know even more clearly than we did 20 years ago what works, and we have examples of institutions of all types and resource levels that have aggressively engaged reform. Do we have the institutional and national will to pull it off? Only time will tell.