2011-2012 AAC report on the HMC core curriculum

version 2.3 8/22/2012; due to the FEC 9/3/2012

This report is a part of the college’s long-term assessment of the HMC common core curriculum. The ‘11-‘12 Assessment and Accreditation Committee ("AAC" or "assessment committee") coordinated inputs from many members of the HMC community in order to provide this snapshot of our successes and failures in meeting the new core's goals. In addition, this reports points out what still remains to be assessed along with a plan for doing so.

**Background** In January, 2007 the college's Strategic Vision Curriculum Committee (SVCC) was charged with examining "student choice and flexibility in the first-year curriculum"[[1]](#footnote-1) among other facets of HMC's common core. The SVCC created *A Proposal for a Revision of the Common Core[[2]](#footnote-2)* in the fall of 2008. A year later, the Strategic Vision Curriculum Implementation Committee (SVCIC) provided a detailed plan for a new HMC common core; initial changes occurred in the core courses taken by HMC's 2009 cohort of first-year students. 2010's incoming cohort was the first to experience the entirety of the new core. This document represents an assessment update on the new core. We also include one possible timetable and vision for future assessments in section 9 (pp. ?-?).

The structure of this document closely follows the assessment plan in section 9 of the SVCIC’s Feb. 26, 2009 report to the faculty (pp. 10-12)[[3]](#footnote-3). In that section the SVCIC identified nine “Institutional Outcomes” that make up the goals for the new core. We restate those nine outcomes here as a table of contents for this report:

**Institutional outcomes**

(1) **retention** “We will retain and graduate a greater percentage of the students that we enroll” (pp. ?-? in this document)

(2) **diversity** “We will attract, enroll, retain, and graduate a greater percentage of students who contribute to the diversity of the college, as measured by gender, ethnicity, and economic background” (pp. ?-? in this document)

**Student outcomes**

(3) **language study** “The numbers of students participating in language study during their first year will increase” (pp. ?-? in this document)

(4a) **academic choice** “Students will be more satisfied with their ability to choose courses that satisfy their interests” (pp. ?-? in this document)

(4b) **academic choice** “Students will be more satisfied with their ability to shape their own academic programs” (pp. ?-? in this document)

(5) **breathing space** “Students will be able to create breathing space within their first two years to accommodate academic, social, or emotional needs” (pp. ?-? in this document)

(6) **interdisciplinary thinking** “Students will be more able to employ interdisciplinary thinking” (pp. ?-? in this document)

(7) **writing** “Students will be more proficient writers” (pp. ?-? in this document)

For each of these seven outcomes, this report presents (a) an assessment-committee-defined standard for success, (b) the data collected to date for that outcome, (c) conclusions we might draw thus far, and (d) a suggested plan for continuing the long-term core assessment in 2012-2013 and beyond. That plan is designed to culminate with the WASC mid-accreditation report that will come due in four years.

**Disciplinary outcomes**

(8) **success in major** The SVCIC hoped that “students will be as able to achieve success in their majors as they were prior to the core reform (each department will choose one outcome they would like to use to assess success-in-major).” This committee began what it imagined was a multi-year effort to follow this SVCIC charge by asking each department to help examine a *downstream impact* of the new core of their own choosing. We fully acknowledge that this is (possibly) very different than defining "success-in-major" as the SVCIC called for, but we believe that, if 100% of HMC's departments materially contributed to this initial effort, it would put the committee in good stead to fulfill the charge as the next WASC reporting deadline approached. We're delighted to report that 100% of departments did materially contribute; the results of their efforts appear in the "downstream impact" portion (pp. ?-?) of this document.

**Plans**

(9) **perspective** Section 9 concludes with a bit of reflection on this process, a longer-term timetable, and an assessment of this year’s assessment committee's efforts.

Outcome 1: *retention*

(1) “We will retain and graduate a greater percentage of the students that we enroll”

(a) *Standard for success*

Defining success for this outcome is relatively straightforward: the college and new core will be successful if the retention and graduation percentages for the classes entering since 2010 are greater than those for classes entering before 2010.

(b) *Data collected*

This outcome is one we can track using enrollment data in the college’s databases. Through Herculean data-wrangling efforts by Eric Ditwiler, we have access to the semester-by-semester dataset of students who joined HMC in cohort years from 2005-2011. Figure 1.1, below, presents the percentage of the incoming cohort remaining after each semester, along with the percentage to graduate. It is important to acknowledge that, within a particular semester, these data will change along with students’ enrollment status. Figure 1.1's retention data are a snapshot as of Aug. 20, 2012:

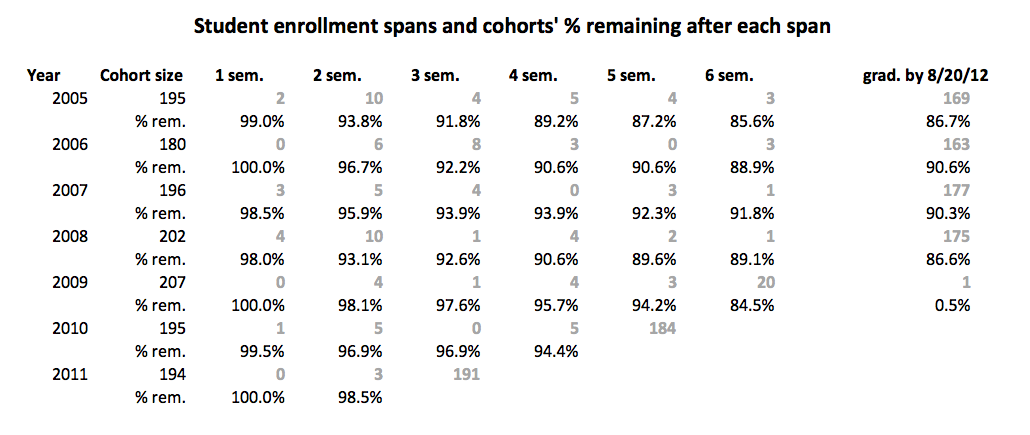


Figure 1.1: The percentages of each year’s incoming cohort who remain after N semesters at Harvey Mudd College. Above, in gray, are the numbers of students from each cohort who enrolled for precisely N semesters. For example, 9 students among the 2005 cohort stayed enrolled at HMC for precisely 2 semesters; 94.4% of that cohort stayed enrolled at HMC for more than 2 semesters. Not all students are continuously enrolled; some ITR or leave for non-academic reasons and return.

Relative to Figure 1.1's data, we are early in the roll-out of the new core: Mudders who entered in 2010 and 2011 experienced its full curriculum. Thus, graduation rates are not yet available for comparison. However, the retention data through the early semesters are available; thus far, the trends look promising, with six students departing in or after the first three terms among the 2010 cohort and two (so far) in the 2011 cohort – neither for academic performance reasons.

In addition, Dean Jacobsen and Registrar Mark Ashley contributed a slightly different set of numbers related to student retention, including (i) the number of students on warning, (ii) the number of students on probation, and (iii) the number of students who ITR’ed, i.e., were ineligible to re-register for their second semester. These are summarized in Figure 1.2, summing across all three of these categories:

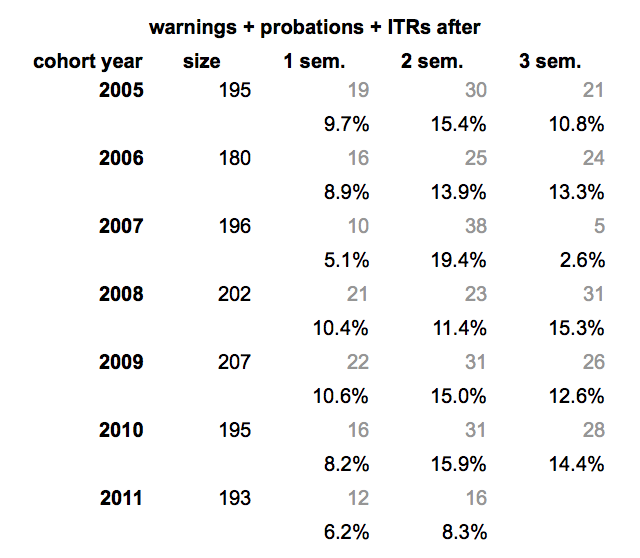


Figure 1.2: Data on the academic status of each incoming class’s students after their first two or three semesters for the cohorts entering 2005-2011, courtesy of Jon Jacobsen and the Office of the Dean of Academic Affairs.

Note that students opting to leave HMC for reasons other than academic difficulty would not show up in Figure 1.2’s data but would show up in Figure 1.1’s.

The above charts show that retention seems to be on a positive trend since the new core. In fact, a breakdown of the Dean of Academic Affairs data by situation (warning vs. probation vs. ITR) show that, in addition to absolute improvement in the numbers, the severity of the situations that do arise has lessened:

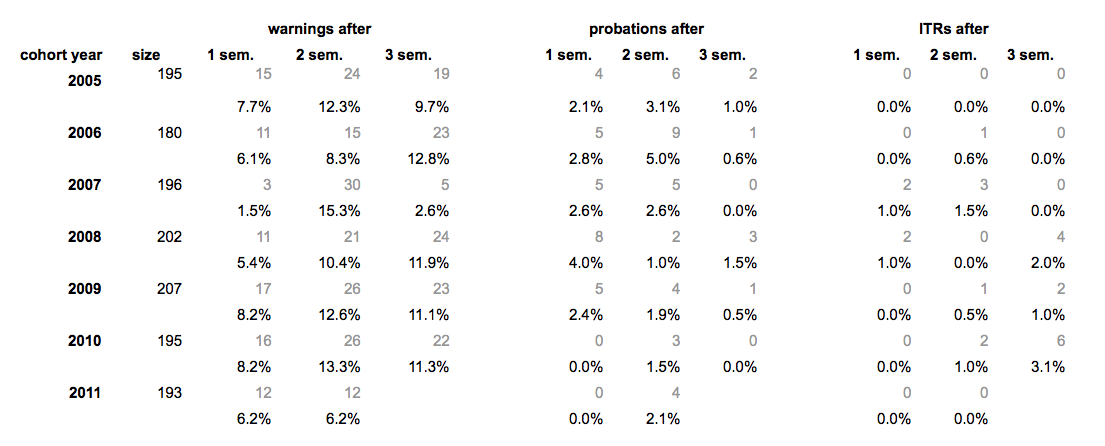


Figure 1.3: This is a breakdown of Figure 1.2's data with each of the three academic situations reported separately: warning, probation, and ITR.

(c) *Conclusions thus far*

Certainly, it is too early to provide a definitive judgment on this outcome as of this writing. However, the trends in all of the data are positive, so “the outlook is good.”[[4]](#footnote-4)

(d) *Future tracking and assessment*

This year’s assessment committee would encourage the yearly tracking of these two data sources:

* Semester-over-semester retention data from CX (Figures 1.1 and 1.2)
* Semester-over-semester academic standing data (Figure 1.3)

Low-effort improvements to this assessment could come through

* Tracking more than the first semester of academic standing data

We look forward to seeing these trends emerge in the coming years.

Outcome 2: *diversity*

(2) “We will attract, enroll, retain, and graduate a greater percentage of students who contribute to the diversity of the college, as measured by gender, ethnicity, and economic background”

(a) *Standard for success*

Here, success is defined as in outcome 1: the college and new core will be successful if the enrollment, retention, and graduation percentages of STEM-underrepresented groups within the classes entering at or after2010 are greater than those entering before 2010.

(b) *Data collected*

For this report we use the same data as shown in Figure 1.1, now separated out by gender and Figure 2.1’s definition of STEM-underrepresented racial/ethnic categories. We do acknowledge that these data were collected one semester earlier than those of Figure 1.1 and, with permission of the FEC, are reported as such.

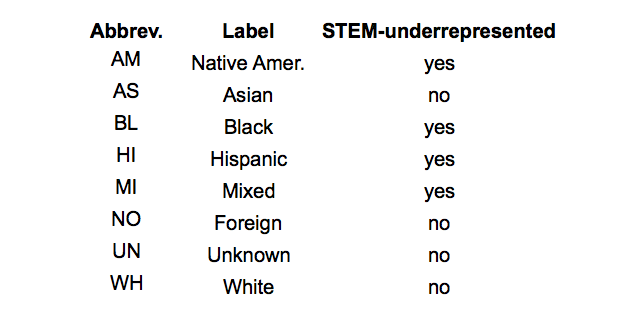


Figure 2.1: The racial/ethic categories maintained by HMC’s data-collection resources, along with whether each one is considered part of the STEM-underrepresented aggregate shown in Figure 2.3’s data.

Figure 2.1’s categories are far from comprehensive, but they are those handled by HMC’s current data-collection resources. In addition, we note that students choose the category into which they are placed. Those choices can and do change; we have not attempted to track those changes over time; rather, the set of figures below are snapshots of the categorizations that students gave themselves as of February 16, 2012.

Figure 2.3 shows the retention statistics for women in HMC’s most recent seven cohorts; Figure 2.2 shows the enrollment statistics across these cohorts among HMC’s most recent seven entering classes. Figures 2.4 and 2.5 show the enrollment and retention statistics for STEM-underrepresented students in that same span.

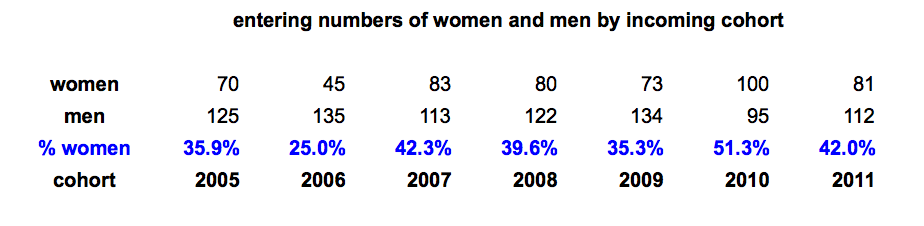


Figure 2.2: The numbers and percentages of women and men who enrolled in the most recent seven incoming HMC classes.

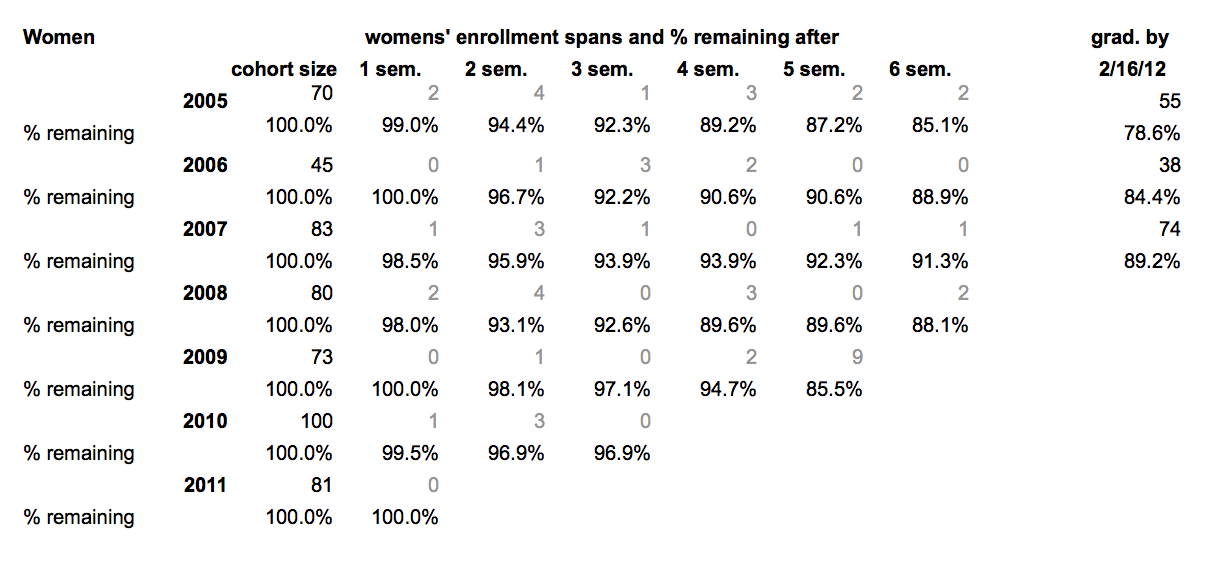


Figure 2.3: The retention statistics for women across seven years.

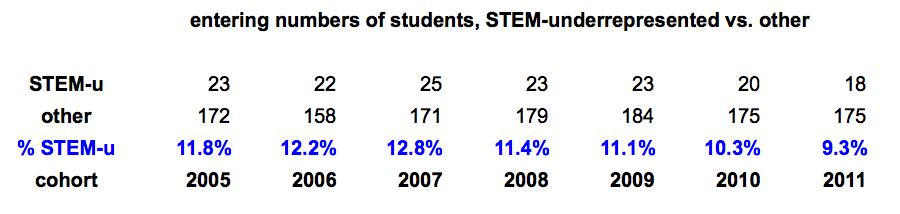


Figure 2.4: The numbers and percentages of STEM-underrepresented and other students who enrolled in the most recent seven incoming HMC classes.

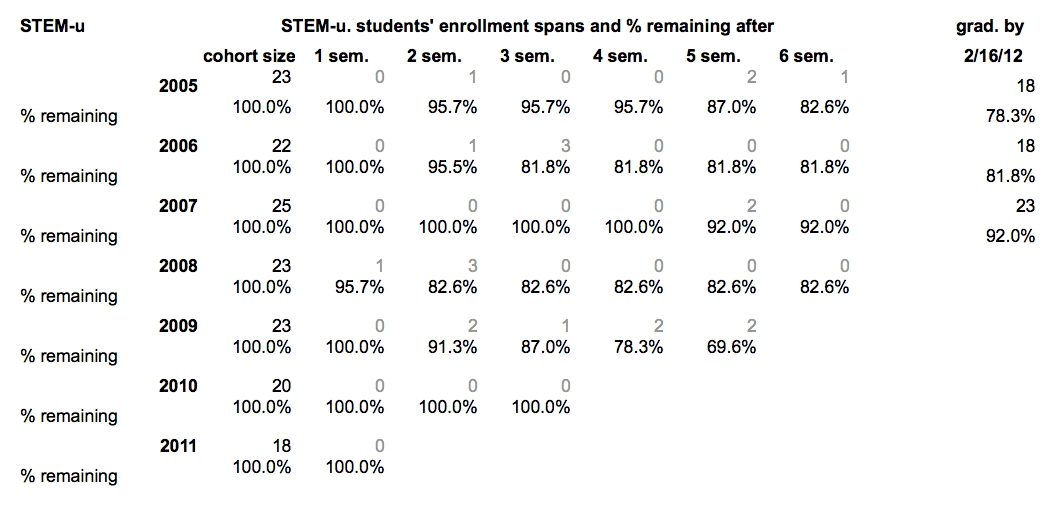


Figure 2.5: The retention statistics for STEM-underrepresented (*STEM-u*) students across seven years.

In addition, Eric Ditwiler and Gilma Lopez were able to obtain a rich dataset on the incomes of students’ families. Because this data arrived relatively late in the writing of this report – and because of its sensitivity -- the assessment committee opted to break down the enrollment and retention in three categories:

* Students at or under the 2012 U.S. poverty threshold for a family of 4.5, which is $25,000 of family income[[5]](#footnote-5)
* Students at or under four times the previous threshold, which is $100,000 of family income
* Students over that amount of family income.

Figure 2.6 captures the differences among these three groups:

***This data may be tricky to get, but I think that – at least in practice – it will be possible. Is this a reasonable summary or should we do something else?***

Figure 2.6: The enrollment and early-semester retention statistics for students in the three family-income categories described above.

(c) *Conclusions thus far*

The data show promising signs in some of the categories examines, e.g., SOMETHING and SOMETHING. However, they also demonstrate that for other STEM-underrepresented groups participation in the HMC community is either not improving or worsening, e.g., SOMETHING and SOMETHING. Thus, this outcome is not only necessarily incomplete because no new-cores students have yet graduated, but even the current trends are a mixture of failures and successes.

(d) *Future tracking and assessment*

This year’s assessment committee would encourage the yearly tracking of the same data as listed under outcome #1, as it supports the data reported here in outcome #2. This is one outcome for which the institution needs to deliberately balance assessment feasibility against curiosity. For example, we recognize that this year’s assessment committee has failed to run a finer-grained analysis of the economic data – and that such an analysis might well be of interest to the community, as long as it could be done in a way that does not compromise students’ privacy. In addition, we acknowledge that there are many other ways in which HMC students define themselves in relation to those around them. Such additional identity criteria could certainly be part of future investigations. We encourage future assessment committees to follow the FEC's, DOF's, and the institution's lead in these pursuits.

Outcome 3: *language study*

(3) “The numbers of students participating in language study during their first year will increase”

(a) *Standard for success*

Like the retention and enrollment outcomes, success for this outcome is relative and thus straightforward: the college and new core will be successful if the number of first-year students participating in language study increases from old- to new-core cohorts.

(b) *Data collected*

We have a provisional set of data that describe the numbers and percentages of first-year students in foreign-language classes. We call them “provisional,” because we are still seeking the registrar’s approval that the queries run were the queries intended. However, to the best of the committee’s knowledge, Figure 3.1 fully and accurately describe language participation trends:

Figure 3.1: Language-participation trends among HMC first-year students

In this case, curiosity got the better of us. We also asked *what* languages HMC first-years are studying at these increased rates. Those are summarized in Figure 3.2:

Figure 3.2: The languages chosen in SOMETHING

(c) *Conclusions thus far*

This is the clearest-cut outcome: the new core has succeeded here.

(d) *Future tracking and assessment*

We will continue to track these numbers. Again, for curiosity more than by mandate from the core’s assessment plan, we will also examine how many students pursue foreign language study throughout their time at HMC. Anecdotal evidence – as well as the change in the first and second semester seen in the data above – suggests that some HMC students who study foreign language early do choose to pursue other fields later on in their undergraduate years.

Outcome 4: *academic choice*

(4a) “Students will be more satisfied with their ability to choose courses that satisfy their interests”

(4b) “Students will be more satisfied with their ability to shape their own academic programs”

The assessment committee felt that these two objectives, though distinct, share a homologous desire: that the new core increase students’ academic flexibility. As a result, we do not distinguish between these objectives for the purposes of this report.

(a) *Standard for success*

Because this year’s assessment committee is unaware of data on prior years’ students satisfaction with their ability to choose courses and shape their academic programs in the same stage of their core-curriculum careers, we tweak these objectives by using an absolute, rather than a relative, standard for success. We asked all of 2011’s first-year students and second-year students to answer the question, “SURVEY QUESTION HERE.” This was one of ten questions asked during the annual *Winter Break* survey shared by the Dean for Academic Affairs, the Core Division Director, and the Assessment Committee.

In addition to providing written responses, SOME# out of SOME# students answered that question on the following seven-point Likert scale:

1. SOMETHING

4. SOMETHING

7. SOMETHING

We will use an average of 5.535 as the threshold for “success.” Although this is certainly an arbitrary cutoff, we note that it was the average student score to the question “STIMULATED INTEREST.” Certainly we hope that students’ perceptions of their academic flexibility are commensurate with their perceptions of the interest in the courses they choose!

(b) *Data collected*

A histogram of the results from that survey question appear below.

We have these charts – got to get them…

Figure 3.1: summary of results on this question

In the spirit of the representative summaries of students’ free responses that Dean Jacobsen provided, we include these excerpts:

Although you don’t get harmony when everyone sings the same note, it does seem that many HMC students do feel they have the ability to choose courses that satisfy their interests and to shape their own academic programs, at least from this year’s survey.

(c) *Conclusions thus far*

Using the criterion stated above, the committee would declare this outcome not-quite successful. Yet several factors mitigate that verdict:

Mitigating factors here…

(d) *Future tracking and assessment*

Future tracking of students’ satisfaction of academic flexibilityAdmittedly, the committee has failed to examine evidence of students’ relative satisfaction before and after the new core curriculum.

We did examine the FSSE, NSSE, and YFCY surveys, from which HMC has data for earlier cohorts. We did find questions that might be indirectly influenced by students’ satisfaction with the academic flexibility available, e.g., YFCY’s inquiry “Since entering this college, how often have you … been bored in class” might correlate with respondents’ feelings of curricular self-determination. Yet that connection – and those among the other indirectly related questions – seemed too indirect to pursue here.

Another possibility to restore this outcome’s intent to compare perceived flexibility under the old vs. new core could involve asking this outcome’s survey question in departmental exit interviews. The committee will invite (or, perhaps, has invited) department chairs to add this question (and those for outcomes #6 and #7) to those valedictory traditions.

Outcome 5: *breathing space*

(5) “Students will be able to create breathing space within their first two years to accommodate academic, social, or emotional needs”

(a) *Standard for success*

We consider this outcome successful to the extent that students report that they feel they can create “breathing space” within their HMC experience. As many have pointed out, this is different than *actually* creating that breathing space: many HMC students thrive amidst a combination of externally and internally concocted pressures. Others do not thrive, but feel them just the same.

(b) *Data collected*

Like outcome 4, the annual *Winter Break* survey asks exactly this question after students’ first and third semesters. In contrast to outcome 4, there are NSSE and FSSE data that can corroborate those survey responses. (are there? – list here)

(c) *Conclusions thus far*

More cowbell!

(d) Future tracking and assessment

something along the lines of this Epic Rap battle pitting Dr. Seuss vs. Shakespeare: http://www.youtube.com/watch?v=l3w2MTXBebg

Outcome 6: *interdisciplinary thinking*

(6) “Students will be more able to employ interdisciplinary thinking”

(a) *Standard for success*

We take the same tack here as with outcomes 4 and 7: in the short term, we ask the extent to which students consider the new core curriculum supportive of interdisciplinary thinking. The Likert response values provide a sliding-scale definition of success; to create a single standard we binarize via the consistent cutoff of 5.535.

The committee acknowledges that there are two near-disjoint interpretations of outcome 6 hinging on the ambiguity in the word “able.” One interpretation makes a statement about the curricular resources available to students, i.e., “students entering in 2010 or later will have more opportunities to employ interdisciplinary thinking.” The second makes a claim about the skills students have internalized, i.e., “students entering in 2010 or later will leave HMC will a more effective and better-exercised repertoire of interdisciplinary-thinking skills.” At least for the purposes of this report, the committee will consider evidence informing both of these two interpretations.

(b) *Data collected*

Profs. Sullivan and Daub, the outgoing and incoming Core Curriculum Directors, drafted an assessment that Choice Lab instructors could give to their students as extensions to the usual *Student Evaluations of Teaching*. Although Choice Lab’s charge has broadened beyond an early interdisciplinary emphasis, the program still provides a group of courses whose hosting department(s) sometimes differ from participants’ anticipated major(s).

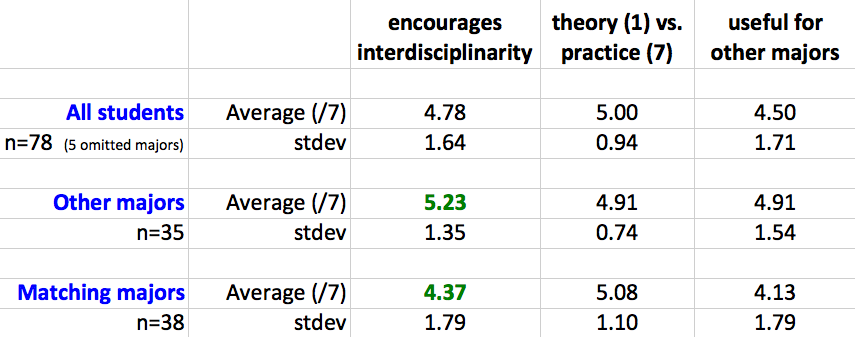
These are the three questions asked of Choice Lab participants in Spring 2012:

1. To what extent do you feel this choice lab encouraged your growth as an interdisciplinary thinker?
2. Where do you feel this choice lab balanced its content between theory and practice?
3. To what extent do you feel this course is useful for students of majors other than computer science?

Among the 78 students respondents, 5 opted not to fill in an anticipated major; of the others, 38 students indicated an anticipated major that overlapped with at least one of their Choice Lab’s hosting departments. Thirty-five indicated an anticipated major that did not overlap with the hosting department(s). It may be worth noting that this is far greater major-coincidence than random placement, i.e., students are unsurprisingly tending – though by no means exclusively – to opt for a Choice Lab overlapping a field of longer-term interest to them.

Figure 1 shows that, on the whole, students felt that their Choice Lab experience did encourage growth as an interdisciplinary thinker, though not overwhelmingly so. Perhaps it should not be a surprise that students who chose a Lab distinct from their anticipated major reported feeling interdisciplinary encouragement to a significantly greater extent (p < .05) than those who “stayed closer to home.”

The differences in the other two questions did not rise to that level of statistical significance, but it is of interest to the committee that students of other majors more strongly expressed that Choice Lab was useful to students of other majors (them!) than the major-matching cohort expressed. Also, both groups agreed that their Choice Lab balanced theory and practice well. The slight edge given to practice follows naturally from what is, after all, a lab experience.

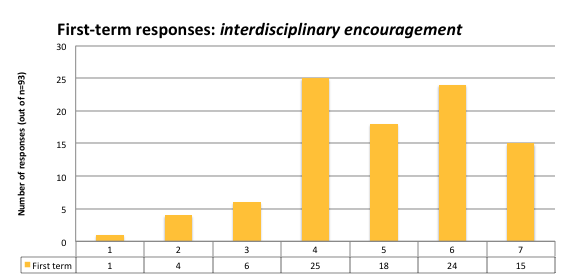


**Figure 1.** Seven-point Likert-scale averages and standard deviations for the Choice-Lab questions listed above. The results comprise three cohorts: all responding students (n=78, 5 of whom opted not to indicate an anticipated major), students anticipating a major *different* from the department(s) collaborating to offer the Choice Lab, and students at least one of whose anticipated major(s) did match at least one of the hosting departments. Differences unlikely to be explained by natural variation (p < .05) are highlighted in green.

In addition to these Choice Lab data, the Jan 2012 *Winter Break* survey asked the following question after students’ first term: “To what extent do you feel your first semester encouraged your growth as an interdisciplinary thinker?” The analogous question was asked after third-term students: “To what extent do you feel your first three semesters at HMC encouraged your growth as an interdisciplinary thinker?”

Figure 2’s results are histogrammed into Likert-scale bins ranging from 1 (I did not feel at all encouraged to grow as an interdisciplinary thinker) through 4 (I felt somewhat encouraged to grow as an interdisciplinary thinker) to 7 (I felt very encouraged to grow as an interdisciplinary thinker). The averages are heartening, at 5.0/7 (=1.4) for the first-term students and 5.7/7 (=1.3) for the third-term students. Using our consistent 5.535 as the cutoff at which to define “success,” HMC could call itself successful in students’ perception of its support for interdisciplinary thinking -- at least by the *end* of the core, if not earlier in the core experience.

It’s possible that the increased sense of core interdisciplinary encouragement after three terms reflects students’ additional certainty about their major paths. The full core, in contrast to the focus ahead, would seem even broader than it did after one term.



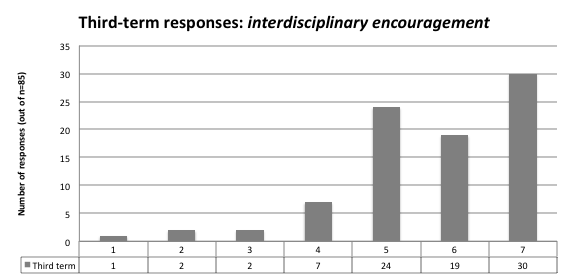


Figure 2. Histograms of student agreement (1-7) with *the core encouraged interdisciplinary thinking* after the first HMC term (gold) and third HMC term (gray).

The comments that (optionally) accompany these quantitative scores vary widely, though several of them hint at such an interpretation. For instance, in the first term, many students saw attempts at connection:

*-- My courses referred each other all over the place; that's definitely something to keep.*

Students gave "shout-outs" to different courses successful at building bridges across various fields:

*-- Chemistry labs consistently tied into concepts of the two chemistry core classes and tried to tie in ideas from various sciences.*

*-- Classes, especially in the math department, discussed material pertaining to many other disciplines*.

Some first-term students noted the substantial technical focus of HMC's curriculum -- and its challenges:

*-- I definitely felt I expanded my perspective as a scientist, especially getting a taste of all the different fields through the required courses, but didn't feel I had any courses that really made me understand the cultural significance of doing any work in society. Most assignments were mundane in that they offered little relevance to the political, cultural, and economic world that surrounds us. Incorporating more relevance with what's going on around us would be much appreciated.*

*-- Because of this [difficulty encountered in attempts to double-major in an HSA discipline], when asked if I think HMC has encouraged me as an interdisciplinary thinker, I would have to emphatically say no.*

Or, more succinctly,

*-- Still, pretty much all courses are science and math.*

After the third term, several students saw the core from a wider perspective:

*-- I noticed first semester that my professors from different classes seemed to be repeating each other on various points. It might be because I took E&M and E59 at the same time, but I can definitely see the connections between different fields much better now than I did in the past.*

*-- The multitude of subjects, though they do not each independently teach us interdisciplinary thinking, definitely do so together. That happens indirectly. In my case however, I feel like the mode of thinking that Mudd tries to encourage has mostly been transmitted to me by the people I've talked to and the people I lived with this year and last year. My conversion definitely happened the first year, but I've only become confident with my thoughts this semester.*

*-- Especially towards the end of the third semester, all these classes started coming together. Most notably Stems, E&M, & What Makes Things Tick. I tried as often as possible to make connections between different disciplines.*

Trying too hard might be read as a compliment:

*-- A clear effort was made to link concepts from various courses and subjects together, but sometimes, it seemed a bit too forced*

In the end, the plurality of opinions expressed an outright positive opinion of this facet of HMC's curriculum:

*-- Being an interdisciplinary thinker is perhaps the most useful skill I've gained so far.*

*-- Harvey Mudd does an astounding job here.*

**NSSE and FSSE data**

The committee, with particular help from Janel Hastings, examined the national NSSE and FSSE data that helped, at least in part, corroborate these HMC-local survey responses. The most relevant question on the National Survey of Student Engagement has been the following, “In your experience at your institution during the current school year, about how often have you… put together ideas or concepts from different courses when completing assignments or during class discussions?” (2011’s NSSE, question 1i). Figure 3 shows the percentages of first-year and fourth-year students who answered “often” or “very often” to that question.

In addition, the leftmost cluster of data in Figure 3 show the percentages of faculty respondants to the FSSE, the Faculty Survey of Student Engagement, who believe that their “students put together ideas or concepts from different courses when completing assignments or during class discussions” either often or very often.

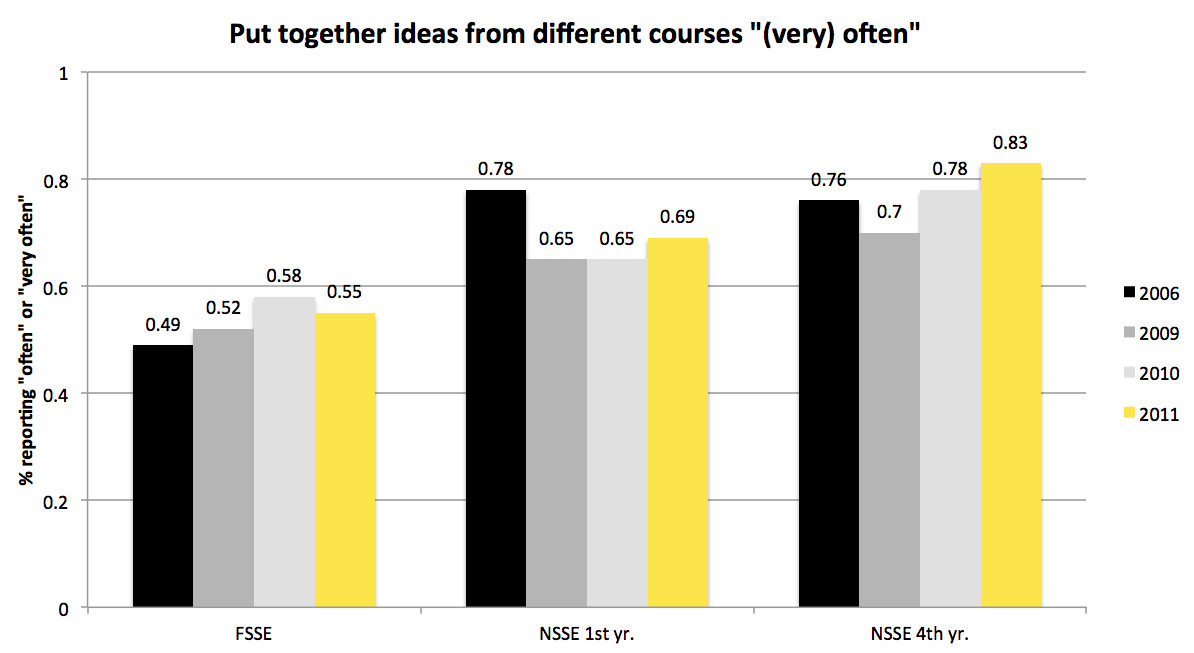


Figure 3 presents the percentage of faculty, first-year students, and seniors who answered “often” or “very often” to the question “how often have you (your students) put together ideas or concepts from different courses when completing assignments or during class discussions?” Several years of FSSE and NSSE data are shown. The cohort sizes are shown in Figure ?.

The trends show a slight student increase in recent years after a drop between the 2006-to-2009 groups. Faculty responses have been relatively flat, perhaps more notable is that the faculty responses are consistently lower than the students’. This may well be a side-effect from the *type* of FSSE survey HMC uses. We use the course-based option, in which faculty choose a *particular* course within which to gauge student engagement. In contrast, the student survey (NSSE) asks students to consider *all* of their courses, instead of a specific course, which may or may not have a large interdisciplinary component. Because of this difference in the surveys’ focus, the committee does not feel it appropriate to read too much into the disparity between faculty and student perceptions of interdisciplinary engagement.

(c) *Conclusions thus far*

These sources make it clear that at least some students and perhaps even more faculty feel there may be more room in the curriculum for interdisciplinary thinking. Very few, if any, voices feel that HMC *over*emphasizes such an integrative treatment of material; indeed, the criticisms that do arise express disappointment that cross-disciplinary work is less thoroughly, less deliberately, or less smoothly deployed than both students and faculty might want.

That all parties so deeply value interdisciplinary thinking provides a fantastic foundation on which to establish both clear expectations and curricular innovations in the future.

Outcome 7: *writing*

(7) “Students will be more proficient writers”

(a) *Standard for success*

We take the same tack here as with outcomes 4 and 6: in the short term, we ask the extent to which students consider the new core curriculum supportive of developing writing skills. Again, we use 5.535 as our cutoff for success.

Here, however, we have the opportunity – and obligation – to compare the work of pre- vs. post-new-core students’ writing. This has already begun at the departmental level, in both engineering and HSA, summarized in section 8, below. In addition, in just over two years we will be able to compare capstone writing – clinic reports and theses – between those groups. Section 9 proposes one possible timetable for that effort.

(b) *Data collected*

The annual *Winter Break* survey asks exactly this question after

(c) *Conclusions thus far*

Speaking should count as writing.

(d) *Future tracking and assessment.*

Outcome 8: *departmental studies of the new core’s downstream impacts*

(8) “Students will be as able to achieve success in their majors as they were prior to the core reform (each department will choose one outcome they would like to use to assess success-in-major)”

This year’s assessment committee, with the knowledge and approval of the Core Curriculum Director, opted to set aside outcome #8. In section 9 we have suggested a timetable for undertaking a thorough comparative assessment of students’ success in each of the college’s majors and in their work with the HSA department.

Instead of success-in-major, we invited each of the college’s departments to choose a possible *downstream impact* of the new core curriculum. The committee offered to help create and/or examine data for such an impact; other departments designed and created a downstream-impact study wholly independently.

For each department, we report

* the **question** asked about a possible downstream impact of the new-core
* the **background** and **context** of the question
* the **departmental goals** and **student learning outcomes** addressed
* the **data used** to inform the question asked
* conclusions drawn from the results and/or possible future directions

***HSA Department***

**QUESTION** The HSA department is interested in how the recent restructuring of first-year writing courses affect the students downstream in their ability to write effectively. More specifically, we’d like to investigate whether the students produce writings of higher quality (in both style and content) in HSA10, compared with the portfolio they produced in HUM1.

**BACKGROUND** Starting from 2011, freshmen at HMC were required to take a college-wide half-semester writing course (Writ1) in the fall and a full-semester writing course from the HSA department (HSA10) in the spring. HSA10 bears resemblance to Humanities 1 (HUM 1), a first-year writing course offered by the HSA department prior to the curriculum change, yet it has been restructured substantially in order to carry on and advance what students learned in Writ1, to introduce students to the wide variety of perspectives and methods in the field of humanities, social sciences and the arts, and to guide students towards independent, exploratory thinking and research.

The HSA department is already assessing how much progress, if any, students have made in writing in their first year by comparing papers producing in Writ 1 and HSA10. For this report, we will compare the papers in HUM1 with those in HSA10. For both assessments, we employ external readers to evaluate the thesis, support, coherence, prose and overall quality of individual papers. The papers are randomly selected, although we also strive to keep a gender balance during the selecting process.

**DATA USED** For this report, we will gather:

* The writing assessment of student midterm and final papers in HSA10 (2011 and 2012)
* The writing assessment of student midterm and final papers in HUM1 (2009 and 2010)

We already have the data for the HUM1 papers and the HSA10 papers from 2011; the evaluation of HSA10, 2012 will be carried out May 14-16. The data will be analyzed in summer 2012.

**GOALS AND SLOs ADDRESSED** This report directly informs the HSA department’s Goal 1, Part 4: *to provide a curriculum that fosters excellence in critical reading and thinking, and in writing.* Since both HUM1 and HSA are taught by faculty members representing a wide range of disciplines (11 for HSA10), the report also informs the department’s Goal 1, Part 1, *to provide a curriculum that allows students to develop an understanding of the questions, methods and content of our disciplines.*

***Physics***

**QUESTION** The Physics department is interested in how the new flexibility in HMC students’ core physics education affects their downstream interaction with physics labs, the physics major, and subsequent physics experiences in general. In particular, we investigate (1) whether there are noticeable differences among students’ enjoyment of -- and performance in -- Physics 54 *Modern Physics Lab* that may correlate with whether those students took the Physics Choice Lab or another of the Choice Lab options. More broadly, we will gather information on (2) other possible differences in the experience of the HMC physics major between these two Choice-lab cohorts (those taking the Physics choice lab and those not doing so.)

**BACKGROUND**

One of the experiences no longer present in the HMC core curriculum is Physics 53 *Electricity and Optics Laboratory*, a lab that traditionally accompanied the Physics 51 *Electromagnetic Theory and Optics* course that remains part of the core. The skills learned and practiced in Phys53 are important to the current-day practice of laboratory physics, as embodied in the physics major requirement, Phys54.

In 2011, sophomores at HMC began fulfilling the core curriculum’s new *Choice Lab* requirement. Four sections of students participated in the Physics department’s choice lab in fall 2011; this group included both students who continued into Modern Lab (Physics 54) and others who did not.

Thus, there are two cohorts of students who took Modern Lab in the spring of 2012 and thereafter: those (n=11) who have taken no physics laboratory beyond Phys22 and those who did take the Physics choice lab in addition to Physics 22 (n=19). This report summarizes the differences in experiences of those two cohorts in Physics 54.

**GOALS and SLOs ADDRESSED** This comparison informs at least two of the Physics department’s stated goals (based on the assessment office’s goals document as of Summer 2011):

* (from Part I, Goal 1, Student Learning Outcome b) ***Students will demonstrate the ability to apply physical principles through active engagement in an extensive laboratory program.***
* (from Part II, Goal 3, Student Learning Outcome b) ***Students will demonstrate skill with the operation of the basic scientific instruments necessary for physics investigation, and will show mastery of basic “bench skills” laboratory research practices common to the academic discipline of physics.***

**DATA USED** For this report, we have obtained

* The grades earned by students in Physics 54, at the course level, a granularity that we believe does not compromise the anonymity of students or instructors involved.
* Grades of Physics 54 students from the spring 2010 and spring 2011 cohorts, for comparison.
* Feedback on the students' coursework and overall lab demeanor, confidence, and experience, obtained in the spring of 2012 by Profs Lynn, Saeta, and Eckert.

**Results**

Discussions with Professors Lynn, Eckert, and Saeta identified five specific student objectives for Physics 54 and, more generally, for the lab skills that the department hopes all of its students develop and exhibit:

**1. Lab demeanor** This objective tries to capture students' organization skills, their lab-notebook skills. Ideally, students would observe and record everything that is needed to analyze a set of physical interactions and would include very little extraneous material in their records.

**2. Data analysis** Ideally, a student's mathematical reasoning is correct, as concise as possible, and includes/accounts for sources of uncertainty.

**3. Report quality** Lab reports should be detail-oriented and carefully composed, stating the results and conclusions from the experiment(s) without overreaching.

**4. "Fiddling ability"** This objective describes a student's problem-solving skills in a lab context – especially when facing difficulties that can arise on-the-fly with an experiment and apparatus. Students who excel at this facet of lab practice would not feel incapacitated or overwhelmed when unexpected events or problems occur. Rather, they would explore possibilities, try different work-arounds, seek out help, and have the equanimity, intuition, and persistence to work through or work around the problems.

**5. Physical intuition** This objective tries to capture how observant students are within the context of a set of physical interactions, especially in knowing which details are relevant and which are negligible, having an intuitive sense of the relative importance of different facets of an experiment, and finding/accounting for systematic biases and problems.

We recognize that these categories are informal (and imperfect!) reflections of different sets of skills that are in evidence in every laboratory interaction.

For each of these five categories, the lab instructors provided an assessment of the relative strengths and in-course improvement of each Physics 54 student close to the end of the term. These were not a numeric assignment but an overall sense of the students' engagement and effort with each facet above. In turn, the AAC then coded those conversations into a scale from 1-5, with 5 representing an enthusiastic and complete expression of each trait and 1 representing an absence the trait. The results, presented here, have been broken out into two groups: those students who took the Physics Choice Lab (PCL) and those students who did not. One might reasonably call *either* of these two cohorts the "control" group.

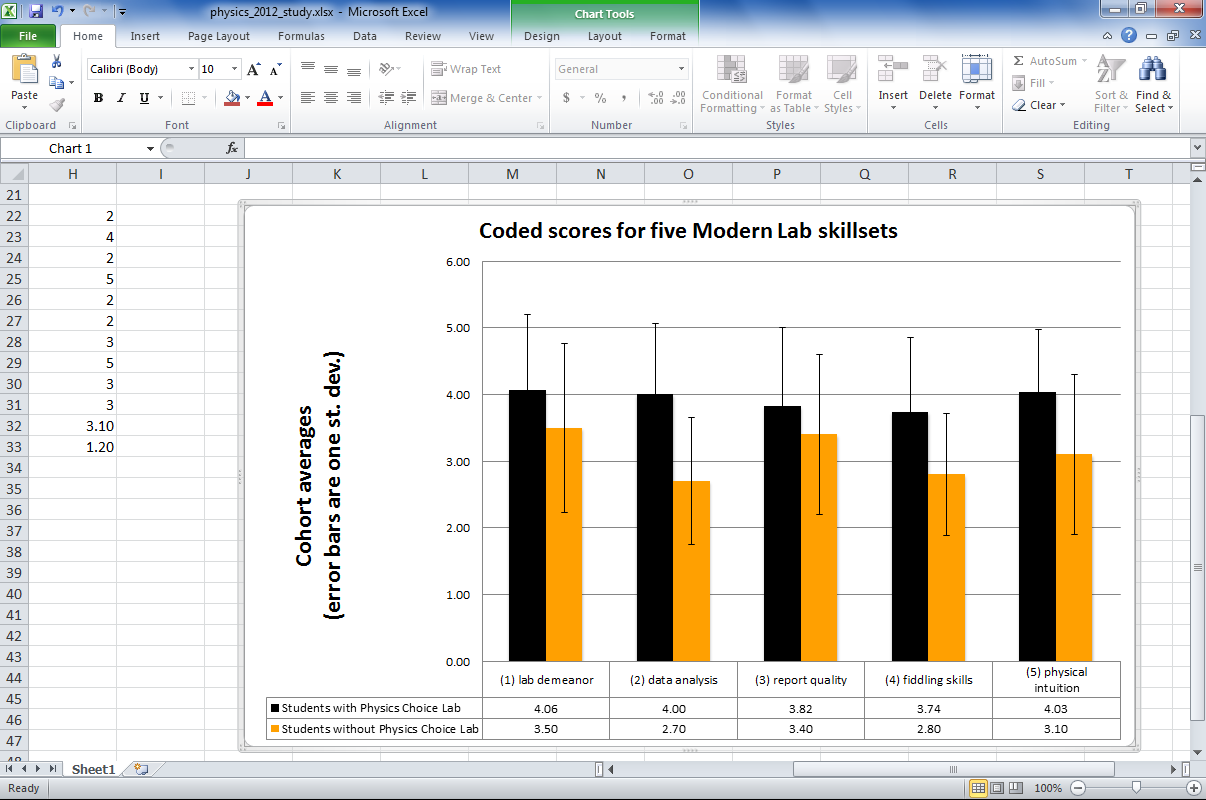


Figure 1. Side-by-side averages with errors bars of one standard deviation of the coded scores from each of the five student laboratory outcomes. The black bars show the coded averages among the Modern Lab students who took the Physics Choice Lab (PCL); the gold bars represent the averages among those not taking PCL.

Were there significant differences between the PCL and non-PCL cohorts? An F-test shows the variances are more likely to be equal than unequal across those cohorts for each of traits (1) through (4); the opposite was true for trait (5). Using the appropriate version of the one-tail t-test for each trait, the differences in averages were significant at p<0.05 for skillsets (2), (4), and (5) with p = 0.002, 0.002, and 0.03 respectively.

In addition, a comparison between the PCL and non-PCL cohorts show a significant difference in final grades: the PCL cohort earned a 3.4 average on a four-point scale as the non-PCL cohort earned a 2.7. An F-test shows 98% probability of unequal variances; a t-test shows a low probability of p=.034, i.e., p<0.05, that the data would arise from distributions with identical averages. Figure 2 shows the anonymized and sorted raw data from those two groups:

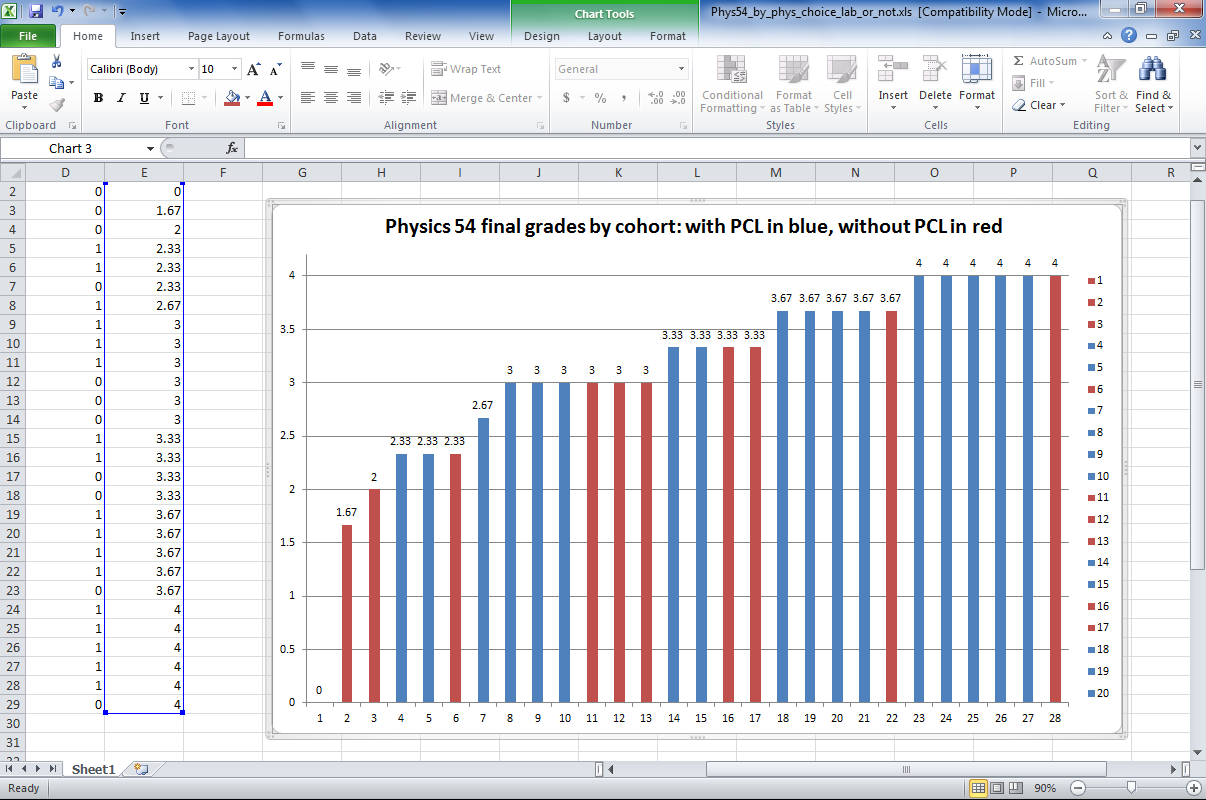
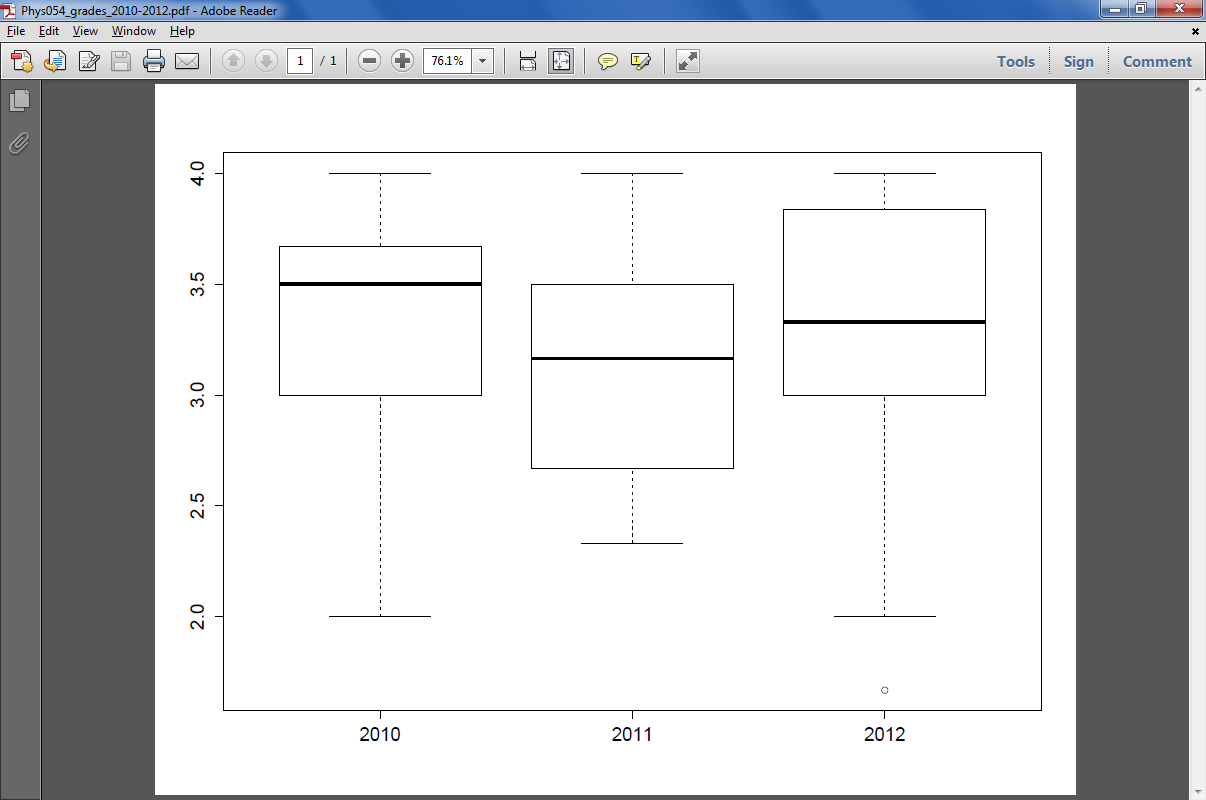


Figure 2. Final grades in Modern Lab for HMC students in Spring 2012: red bars indicate students who did not take the Fall 2011 physics choice lab; blue bars show students who did take that physics choice lab. These differences are significant in the sense described above.

**Conclusions**

Why the differences? Certainly an additional semester (PCL) of lab experience in which expectations are internalized and habits are strengthened seems likely to help students with the more demanding experiments that constitute Physics 54. It is also possible that students opting for PCL are a group especially enthusiastic about honing their skills. Although we do not have data analogous to Figure 1 for previous years in Modern Lab, we do have the overall grades from the spring 2010, 2011 and 2012 cohorts, presented here in a box-and-whisker plot including quartiles and data ranges:

**Figure 3** – quartiles and ranges for final

grades in Physics 54, 2010-2012

An analysis of variance examination of these three years does not suggest a significant difference (p=0.93) in the final grades of Modern Lab students through these terms.

It seems clear that an additional semester – the Physics Choice Lab, in particular – of physics lab work correlates with improved skills and results in Physics 54, Modern Lab.

**Possible future investigations**

In the future, we may have the opportunity to compare how physics majors’ experiences through a larger swath of the Physics department's curriculum correlate with the different core-curriculum paths they pursued.

In addition, it would be possible to repeat this 2011-2012 study in the coming academic year, adding – as suggested by Peter Saeta – a comparison of students' grades in Physics 22 and Physics 54 in order to get a better sense of how much of the difference seen might come from the experience of Physics Choice Lab.

The assessment committee will follow the lead and the priorities of the FEC and HMC administration in order to determine whether this is desired.

***Mathematics***

**QUESTIONS** The Mathematics department is interested in how students taking the new core curriculum fare in Mathematics 157, *Intermediate Probability*. In addition, the department maintains a longitudinal assessment of its core curriculum: this report shares that internal assessment, in case others find them useful or would like to coordinate a shared assessment with mathematics.

**BACKGROUND** Math 157 is required of mathematics majors and it is often taken immediately after the mathematics core. The department identified Professor Pippenger's section of the course as a curriculum that has remained relatively consistent across the changes in the core. As a result, it offers a basis for comparison across cohorts who experienced several different core mathematics curricula during the evolution of the new core.

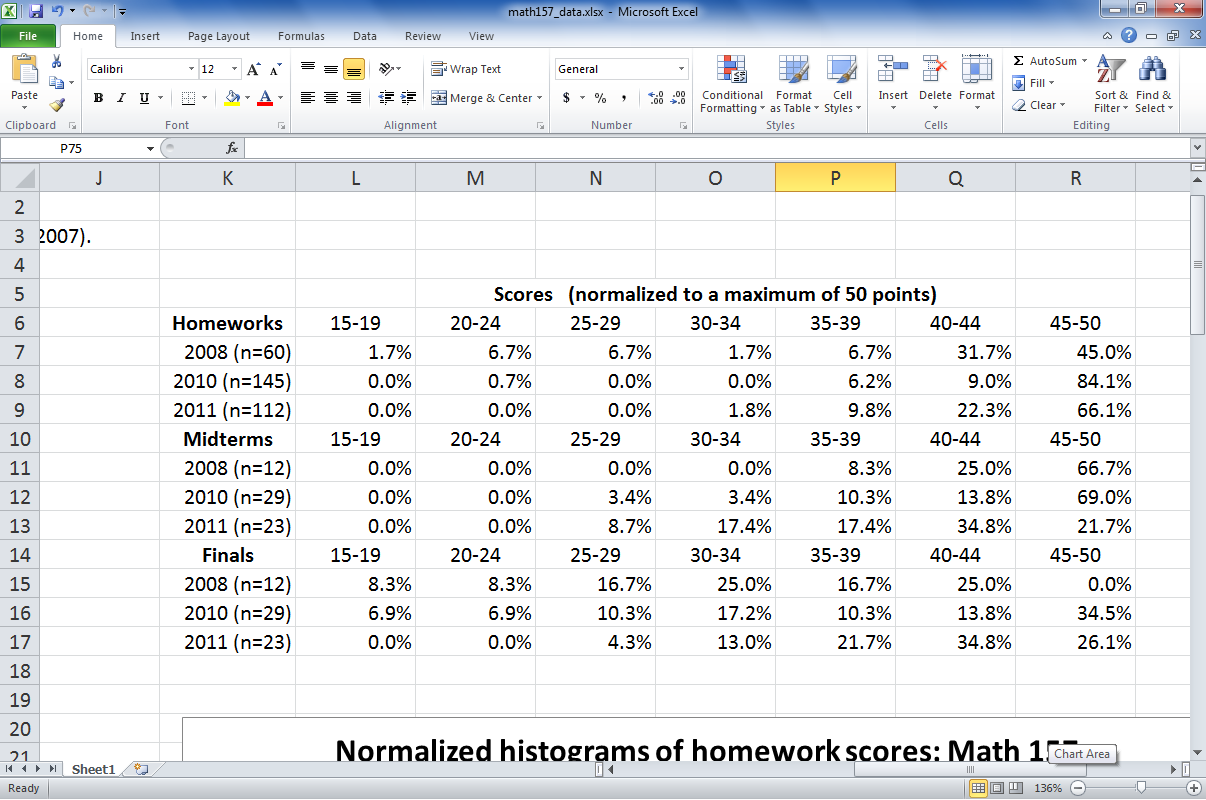
It is worth conveying Prof. Pippenger's observation that the content of Math 157 was not especially affected by the changes to the mathematics core. Yet because the treatment of probability and statistics did not undergo major core-content changes, Mathematics 157 may serve, at least in part, as a general barometer of the sophistication of students' mathematical reasoning and expression, rather than as a referendum on specific skills.

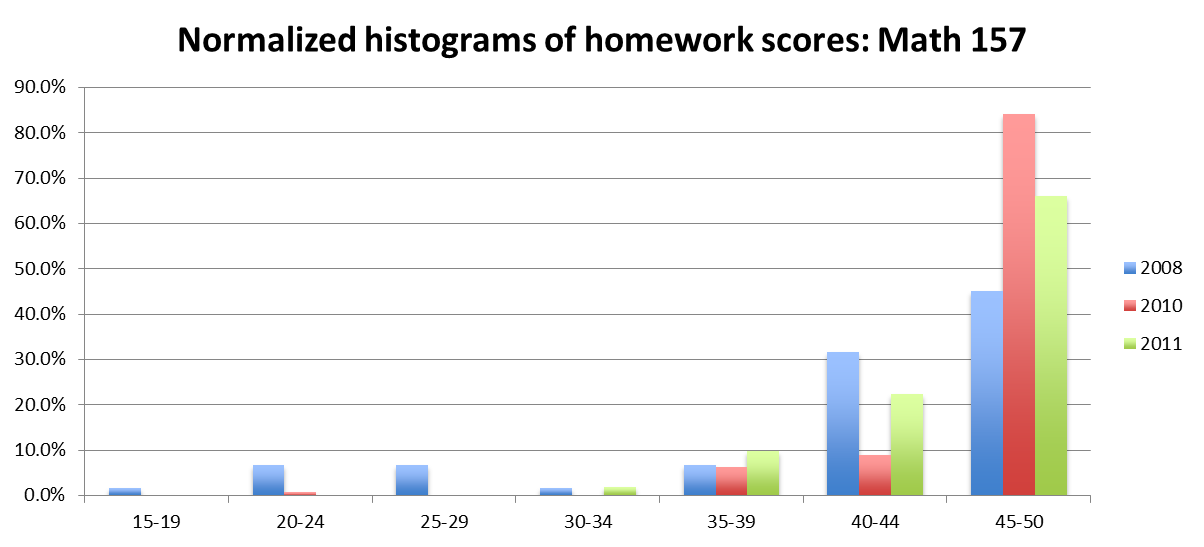
**DATA USED** For this report, we have obtained

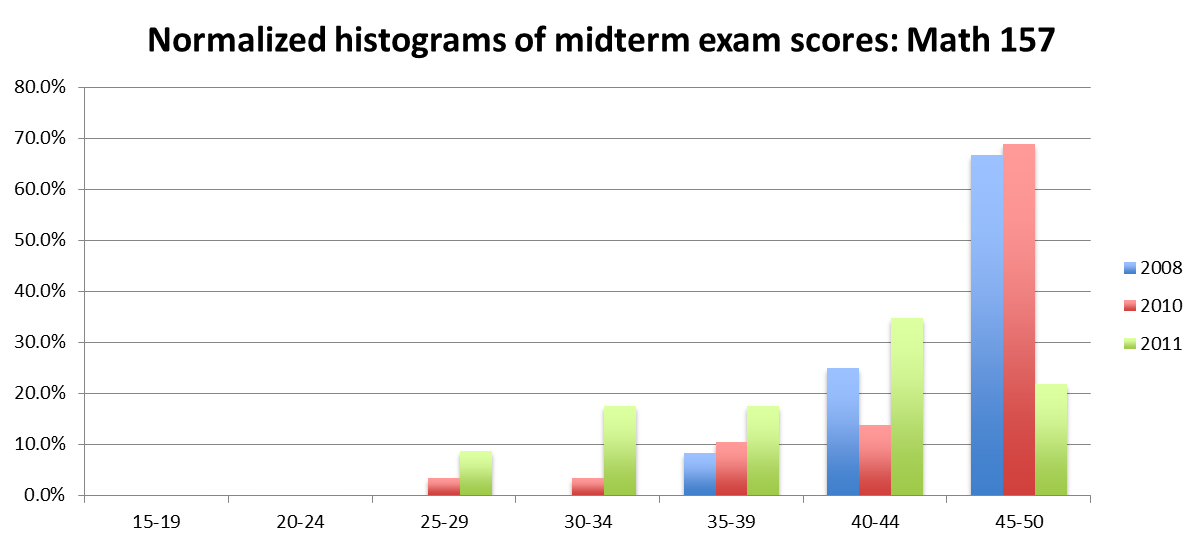
* the overall final grades in the Spring 2008, Spring 2010, and Spring 2011 offerings of Math 157. Because the mathematics core had undergone a series of changes even before the official advent of the new core, each of those three groups did see somewhat different mathematics curricula in the core. Here are Prof. Pippenger's summary of those differences:
  + 2008: The students in this class had taken Prob/Stat in the first semester of their *Sophomore* year (mostly in the fall of 2007).
  + 2010: The students in this class took Prob/Stat in the *first* semester of their Freshman year (mostly in the fall of 2008).
  + 2011: The students in this class took Prob/Stat in the *second* semester of their Freshman year (mostly in the spring of 2009).
* we thank Prof. Pippenger for sharing each of those sections' histograms of homework, midterm, and final-exam grades.
* the surveys of the mathematics core currently used by the department's assessment subcommittee

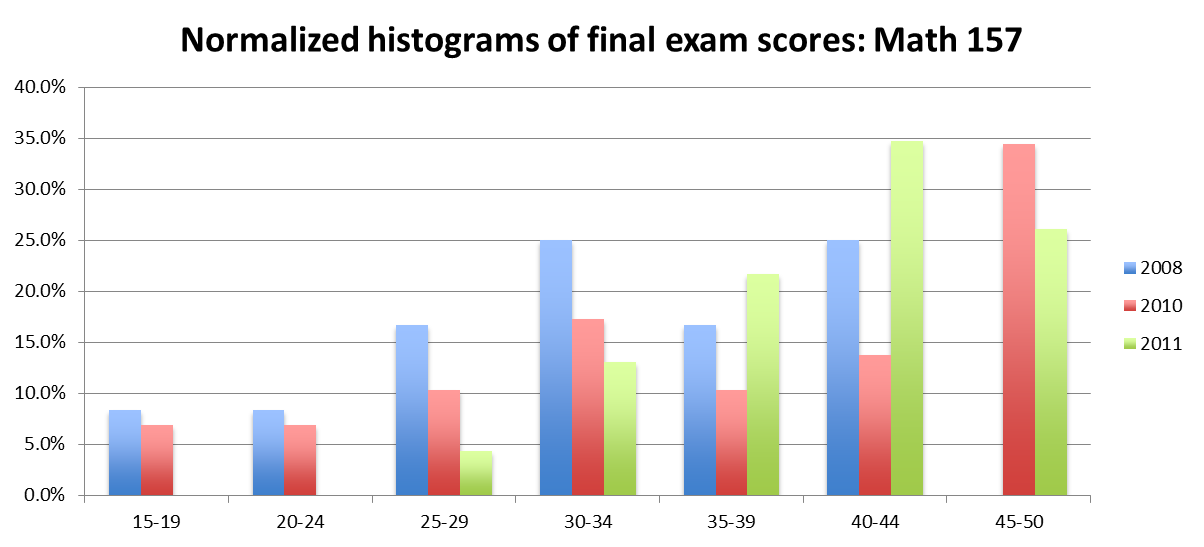
**Results**

As the histograms are already a statistical summary, computing a single statistic from them might be less illuminating and less indicative than the more qualitative trends provided by visualizing them. Figure 1 shows both the bar charts and the raw data from the three course components in each of the three cohorts:





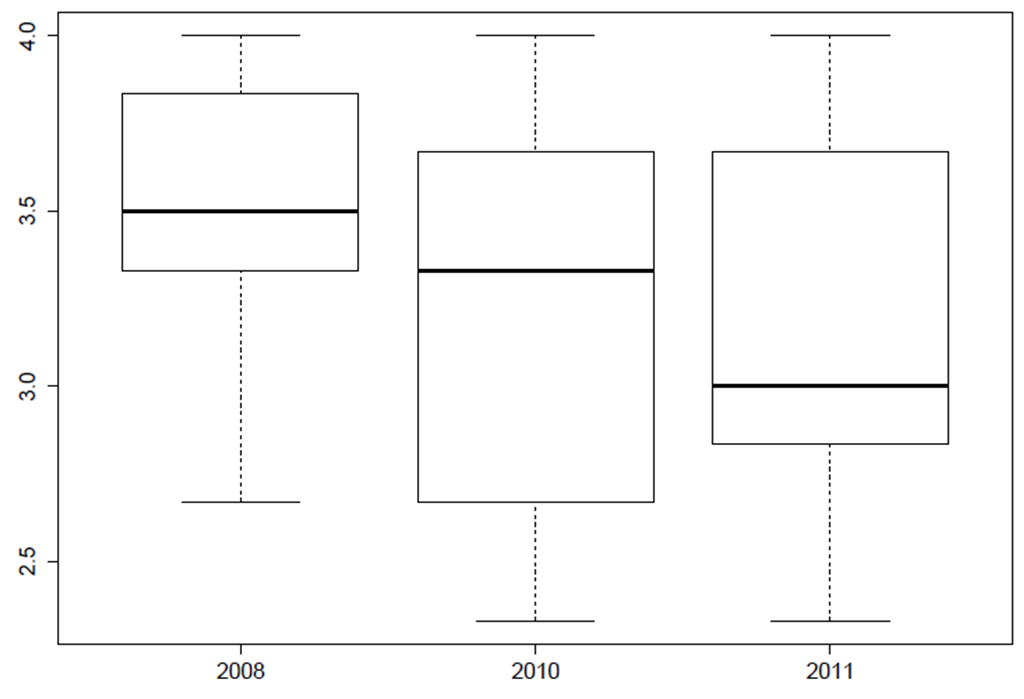




**Figure 1** (*top*) the histogrammed data of scores (out of 50 points in each case) for homework assignments, midterm exams, and final exams in Math 157 in each of three offerings (*below*) the same data in bar-chart form. Note that to remain commensurate the vertical axis is in units of *percentage of students* (of each cohort); in addition, note that the scale on the vertical axis changes in order to accommodate the largest value in each chart.

Prof. Pippenger elaborates on these data: *I used the same book and mostly the same homework assignments in all three years. The midterm and final exams were different from year to year---in 2010 the midterm was too easy and the final too hard; in 2011 I think I got things right. Another thing that changed was when students had taken Prob/Stat (the old Math 62 and current Math 35) as the core was revised. All scores were out of 50, and I've bucketed them into intervals of length 5 (except for 45-50).*

Final grades do provide a single statistic on which we have run an analysis of variance. Figure 2 provides a box-and-whisker chart of the quartile values on a four-point scale for each of the three cohorts:

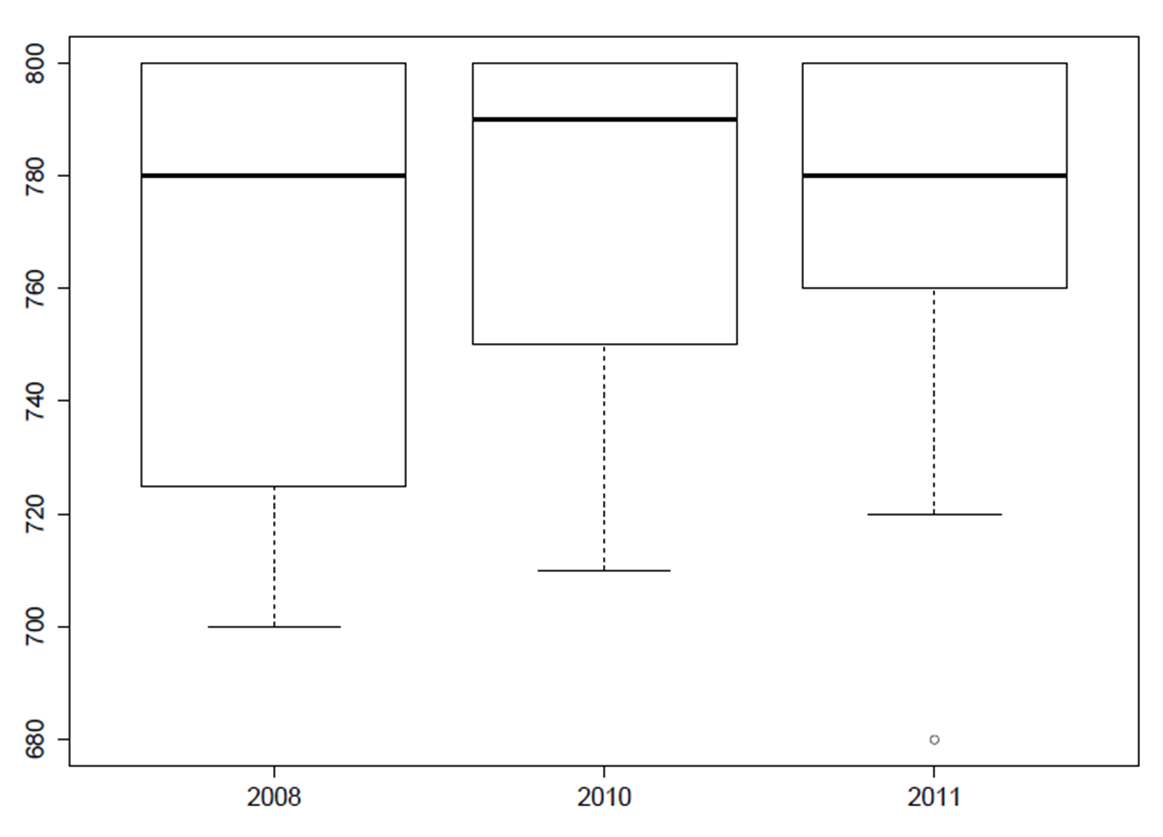
****

**Figure 2** – quartiles and ranges for final

grades in Math 157 in 2008, 2010, 2011

The ANOVA does not show a significant difference at the level we have been using in this document (p < .05). Paraphrasing Eric Ditwiler, *A political scientist might call the difference between 2011 and 2008 significant. A psychologist might not. The phrase I see in the literature for such cases is "marginally significant (p < .09)."*

In order to provide additional context for the relative mathematical backgrounds of these three cohorts, the mathematics department suggested examining the SAT scores of the three groups. Figure 3 shows the quartiles and ranges of those scores for each of the three groups above:

**Figure 3** – quartiles and ranges of math SATs scores for the cohorts in Figures 1 and 2

Here, no significant differences in the mathematics SAT scores emerge; we do note that there is a small correlation (explaining 8% of variance) between mathematics SAT score and grade, though this is not a surprise (except perhaps that the correlation is not stronger!)

Prof. Pippenger rightfully reinforces the importance of the individual students' stories when considering small-group statistics: *Associated with low grades there is usually a story that offers some explanation, but is impossible to study in a statistical way.  In a small class (one year there were 30 students in Math 157, but usually the enrollment is in the teens), these special cases can appear statistically significant.*

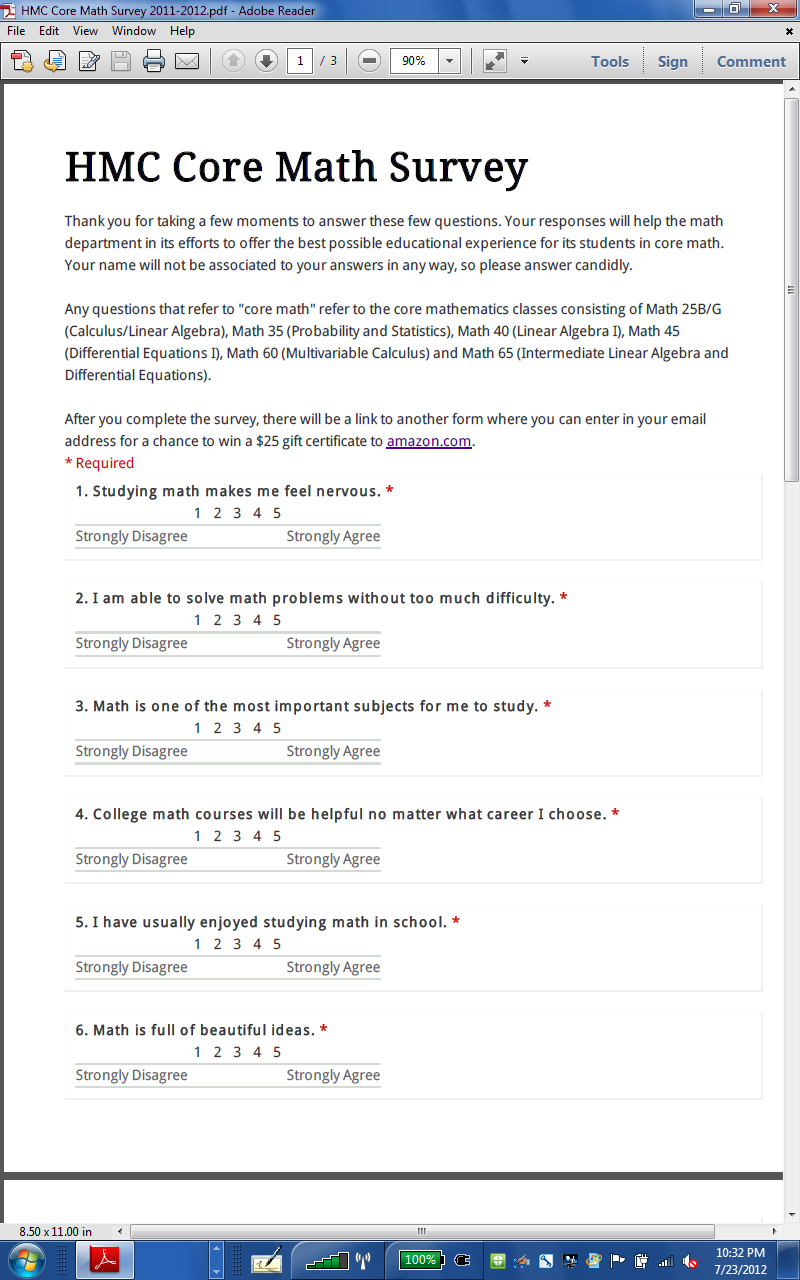
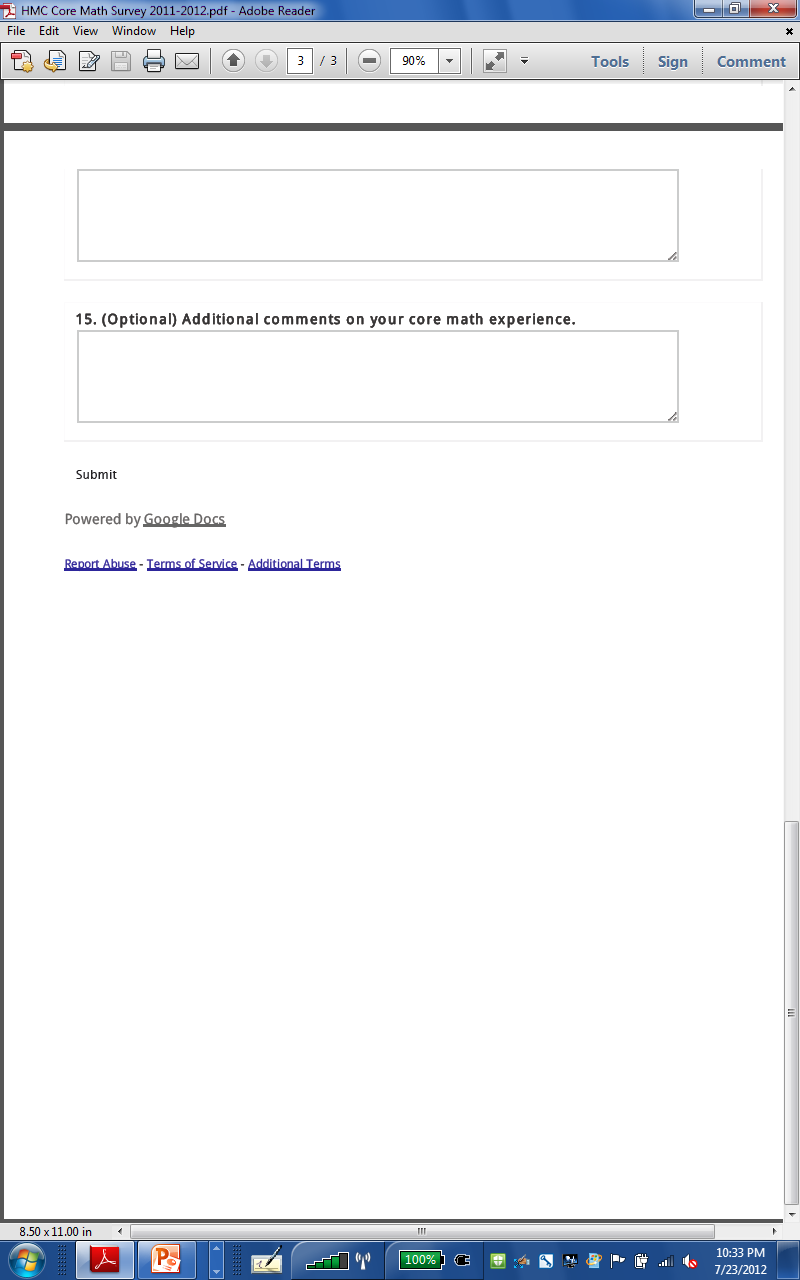
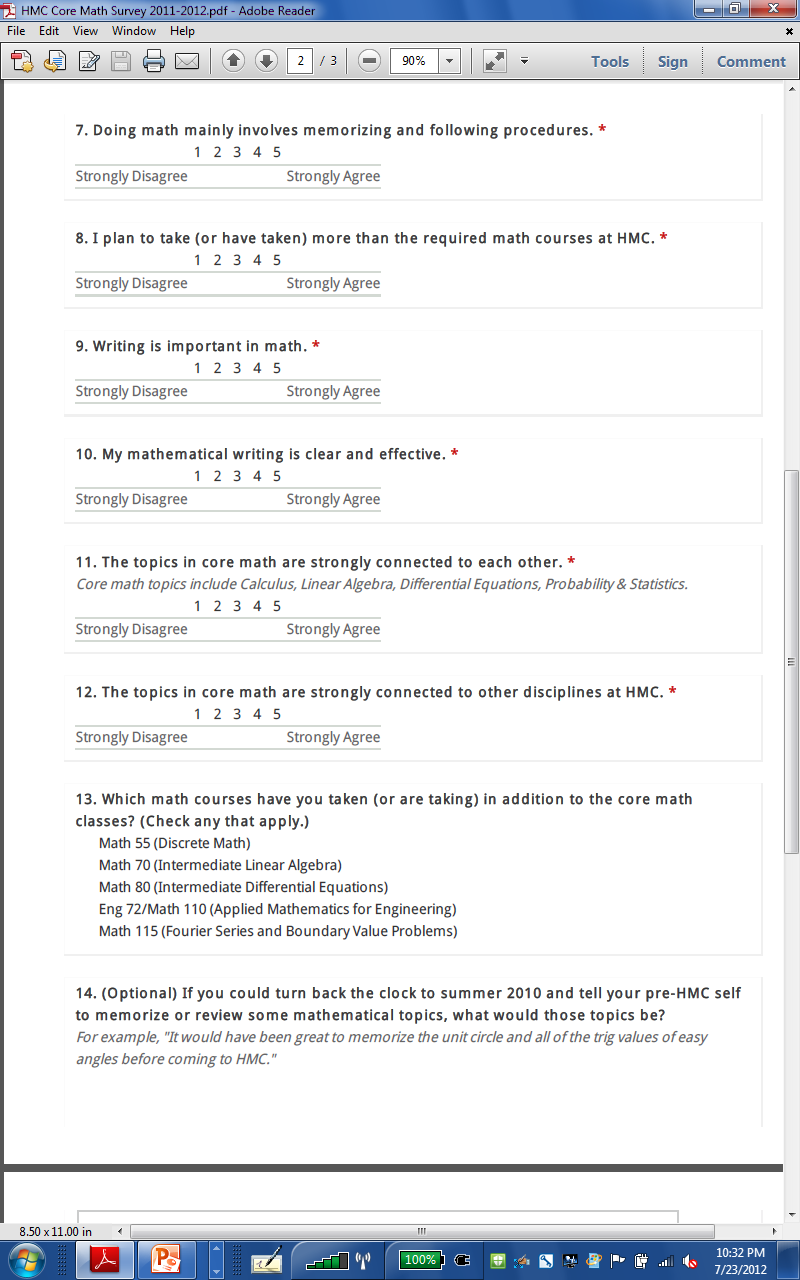
**Conclusions**

Certainly this is a decidedly partial measure of mathematics majors' work, as it draws from a single course, albeit one that has changed less than the curriculum preceding it. The results do not strongly suggest a decline in student performance, and whether they even weakly suggest one is not cut-and-dry.

**Current and possible future assessments**

Extending these data – either by including later offerings of this Math 157 or by adding other major requirements into consideration – would provide a chance to draw a more resolute conclusion than we can now. The committee will follow the mathematics department's lead as to whether to pursue such an extension.

In addition, the department maintains an ongoing set of internal assessments of the mathematics core, supported by the efforts of a mathematics-assessment subcommittee. The instrument they have created is intended to measure attitudes only, and it's all anonymous and self-reported data. The survey includes 12 Likert-scale and 3 additional questions, shown below in Figure 3. It is requested from all students before the first semester and after the third semester (after all core math is completed), and it is requested again from mathematics majors after the sixth semester and just before graduation as part of the exit process.



**Figure 3.** The mathematics department's longitudinal survey of attitudes towards the mathematics core curriculum.

**GOALS and SLOs ADDRESSED** As head of the mathematics subcommittee on assessment, Darryl Yong shared the departmental goals this survey addresses:

*The mathematics department hopes that its majors will*

*1. be competent in a broad range of mathematical skills and topics*

*2. be able to recognize and apply mathematics in a variety of settings*

*3. be able to articulate what mathematics is about and what mathematicians do*

*4. know how to learn and have the competence and confidence to build on their knowledge base independently*

*5. have strong professional and communication skills*

*6. be good citizens who are aware of the impact of their work on society*

*7. be able to function well as a part of a team, and have honed their leadership skills*

*8. be able to work and communicate with diverse groups of people with varying abilities and who come from a variety of cultures.*

In addition, the core-impact investigation of Math 157 informs the first of these goals, competence in a broad range of mathematical skills and topics, in its coverage of some of the probabilistic concepts and skills that all HMC mathematics majors should know.

This report does not report the results of this departmental longitudinal survey – those are still ongoing and not the purview of the AAC in any case. Even so, the committee is enthusiastic about the mathematics-specific feedback loops this effort makes possible!

***Biology***

**QUESTIONS** The Biology department is interested in how students taking the new core curriculum fare in a pair of courses that are part of the foundation of the biology major: *Evolutionary Biology* (Bio109) and *Biology Lab* (Bio 54). In addition, the department has articulated a longer-term desire to assess and further biology majors' sophistication and comfort in applying statistical techniques for data-analysis. Skill with technical writing, too, is fundamental to the department's goals. We look ahead to a comparative assessment of pre- and post-new-core biology majors' theses along these two criteria.

**BACKGROUND** As all majors at HMC, biologists rely on skills reinforced through the core, especially laboratory work, notebook-keeping, and basic mathematical and statistical skills.

Biology 109 is a requirement for all biology majors. The course uses algebra, probability, and statistics, e.g., determining independence and using a χ2 test for goodness-of-fit. Past experience with off-campus students in Bio 109 has shown, at least anecdotally, that non-Mudd students (who have not taken the HMC math core) use disproportionately more instructor office hour time. For this reason, we believe that any deficiencies in math background resulting from the changes to the core might appear in changes in office hour pressure, as perceived by the instructors for the course.

In changing from the old core to the new core, the chemistry core lab was reduced from two semesters to one semester. Previously, the chemistry core lab experience included a 3-week restriction mapping experiment led by biology faculty. As a result of the core changes, this lab was eliminated, and the remaining labs were trimmed down or replaced with shorter labs to fit within a single semester.

Biology 54 is the only biology laboratory course required of all biology majors, typically, though not always, taken in the sophomore year. This course depends on skills learned and practiced in the chemistry core lab, including following an experimental procedure, recording data and observations in a laboratory notebook, analyzing and interpreting data, and presenting results in written form. Although notebooks are not collected or graded in Bio54, changes in student preparation may be observed by instructors who have taught this course repeatedly both pre- and post-new core. Additionally, the laboratory exercises in Biology 54 have been carefully calibrated over several years to fit within the scheduled lab period, making it possible to observe any changes in student ability to finish lab exercises in the scheduled time.

**DATA USED** For this report, we have gathered:

* The grades earned by students in Bio 109 and Bio 54 across a number of years in which the instructor of those courses has remained constant
* Individual scores on assignments, tests, and other components within those two courses
* Reflections from Professors Ahn and McFadden as to changes in the running of those two courses

**GOALS and SLOs ADDRESSED** This report directly informs at least three of the Biology department’s stated goals (based on the assessment office’s goals document as of Summer 2011):

* (from Part II, Goal 6, Student Learning Outcome a) ***Biology majors at HMC will demonstrate skill with the operation of the basic scientific instruments necessary for biological investigation, including microscopes, centrifuges, spectrophotometers, electrophoresis equipment, and pH meters***
* (from Part II, Goal 6, Student Learning Outcome b) ***Biology majors at HMC will be able to demonstrate basic laboratory “bench skills” common to the discipline…follow and use experimental protocols including recording and maintaining accurate data records…***
* (from Part II, Goal 7, Student Learning Outcome a) ***Biology majors at HMC will be able to successfully complete assignments that require them to analyze and design experiments***

**Results**

**Bio 54**

Figure 1 summarizes the overall averages of student in Bio 54 during the years of 2007 through 2011. Professor Ahn taught the course through that span, and during that time its content did not change radically, though the "Muscle lab" in 2009 comprised a genetics paper and, in 2010, an EMG paper.

The error bars represent one standard deviation above and below the mean stated in the data table. The lighter gray overlay highlights a cohort who experienced *some* of the new-core curriculum, e.g., some took CS5 "green"; the darker overlay shows the group who experienced *most or all* of the new-core curriculum.

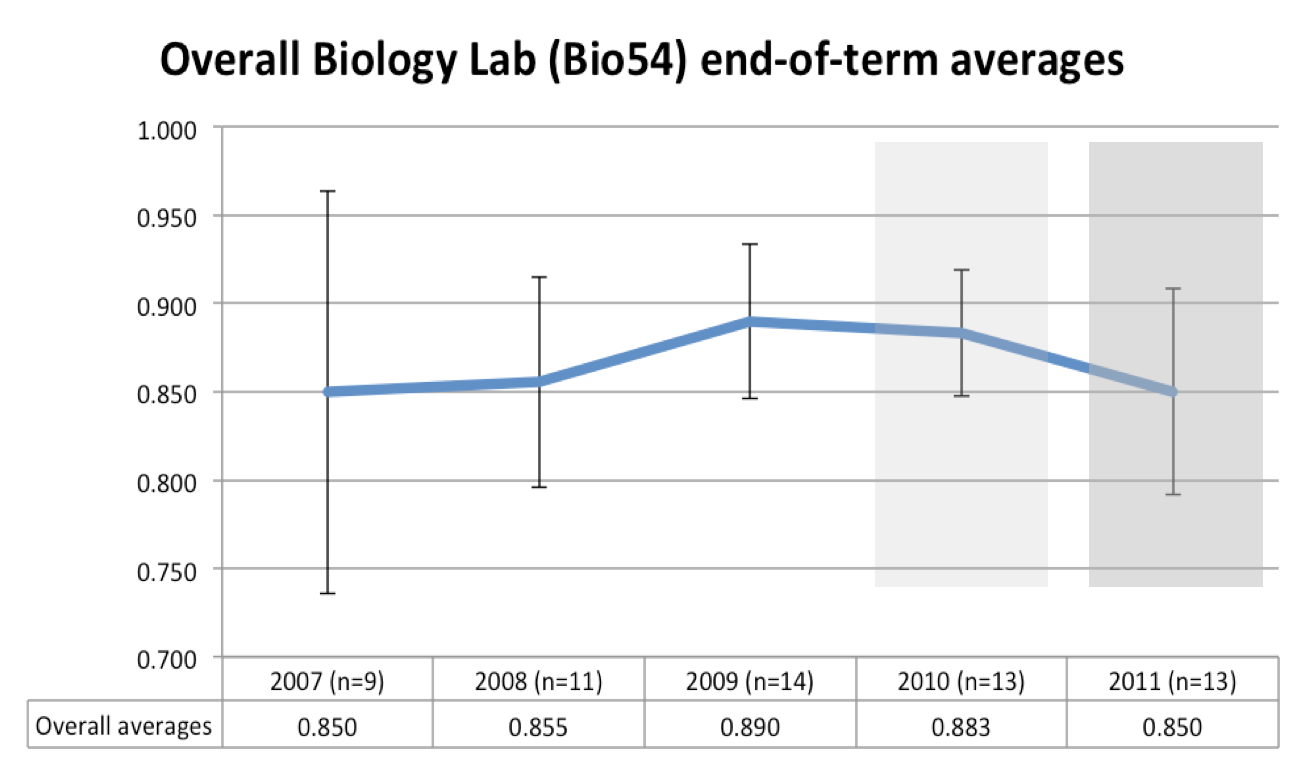


Figure 1. Bio 54 averages across the past 5 years with one-standard-deviation error bars. The gray overlays correspond to cohorts experiencing some and all of the new core curriculum, respectively.

The averages drop four points across the roll-in of the new core; because this is within one standard deviation and within longer-term experience, it is possible that that is not a trend at all. Alternatively, it may combine both an unusual cohort in 2009 and the possible need for Bio 54 to develop students' lab skills from a less experienced starting point in 2010 and beyond.

As suggested by Steve Adolph, we ran an analysis-of-variance test to determine if the change in mean was significant within the above data. The resulting F-statistic was 1.00, yielding a p-value of about .41, which suggests that the changes in the mean are well within the span expected, given the variances involved.

We examined the *T* statistic from the Cramer-von Mises two-sample test as an additional check for significant changes. This statistical technique allows the comparison of two ordered samples in order to determine their likelihood of having been drawn from the same distribution. Figure 3 summarizes those statistics in comparing each of the five cohorts against the others. The zero diagonal indicates, naturally, that there is no evidence that identical samples were drawn from distinct distributions; the diagonal symmetry arises equally naturally. The *T* values shown are not, in an of themselves, probabilities. For the population sizes considered, those *T* statistics are "significant"[[6]](#footnote-6) if greater than 0.32: the highlighted boxes show those.

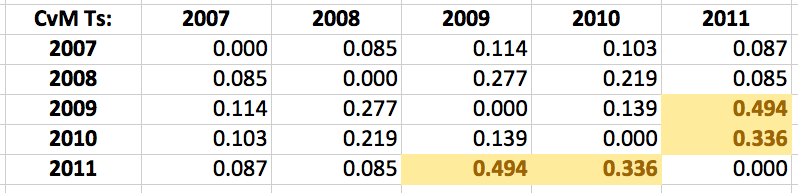


Figure 2: the Cramer-von Mises *T* statistics between Figure 1's cohort pairs.[[7]](#footnote-7)

As Steve points out, these pairwise comparisons suffer from the phenomenon that by making enough comparisons it's unlikely *not* to find one or more differences that would trigger such a definition of significance. That is, we would expect some "false discoveries." In order to legitimize the significance of Figure 2's differences, 2009 would have to be considered a baseline (without the context of previous years' data).

The individual components of these Bio 54 averages appear in Figure 3:

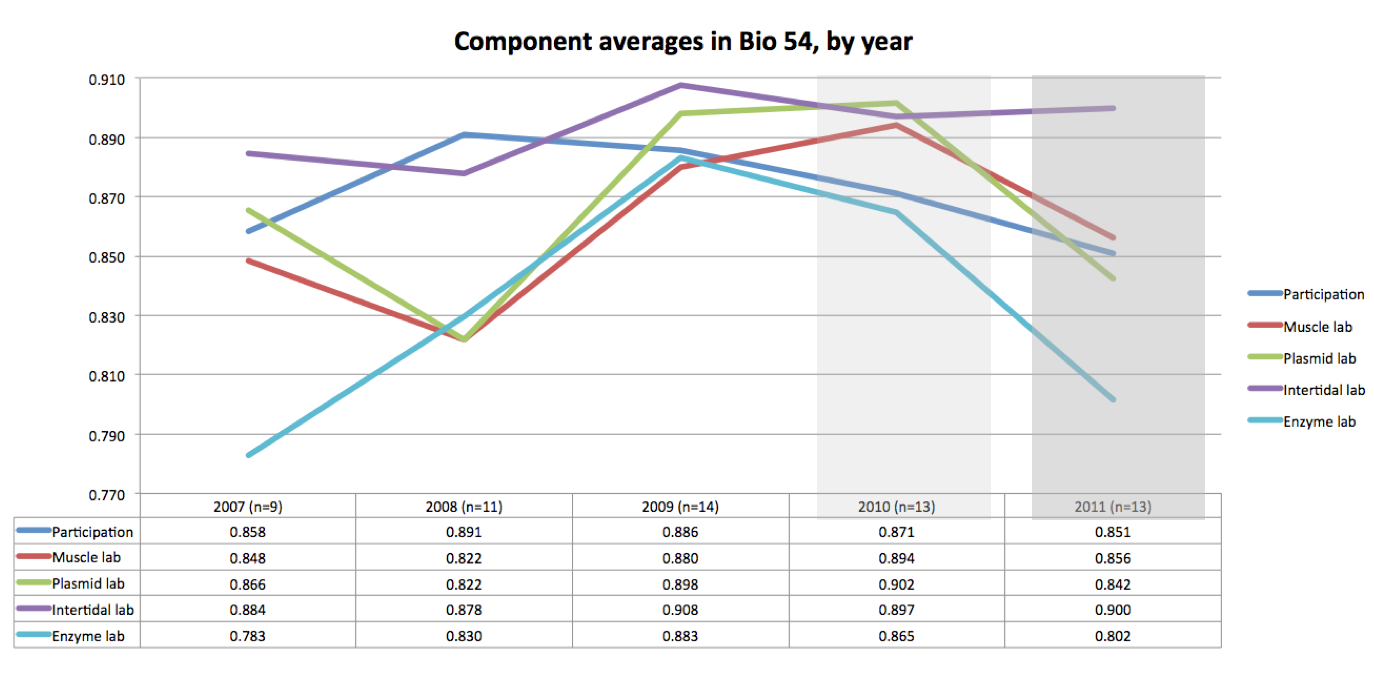


Figure 3. The component averages that make up Figure 1's data. Included are overall lab participation, the muscle, plasmid, intertidal, and enzyme labs.

Figure 3 demonstrates that the Enzyme and Plasmid labs contribute the lion's share of the changes from 2009 to 2011. This coincides with the subject matter now absent from the remaining core chemistry lab, so it's possible that the change has had an impact.

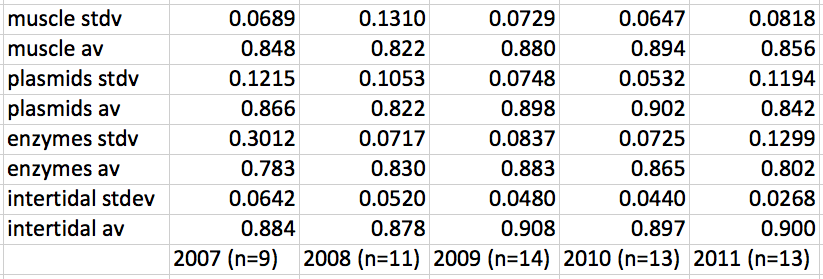


Figure 4. The within-year standard deviations for each lab in Figure 3.

Because placing error bars in Figure 3 turned out to be unwieldy, Figure 4 reports the standard deviations (within each year's cohort) of the four Bio 54 labs. That the enzyme and plasmid labs have four of the five the highest within-year variances may suggest that those are the "riskiest" labs, for which bench skills – and perhaps even luck – could play a greater role than for the other labs. This observation is mitigated by the fact that the large variance in 2007 was due to a missed lab by one student.

**Bio 109**

Evolutionary Biology, Bio 109, is a core requirement within the biology major most often taken by sophomores. Because it does not serve as a prerequisite for many other courses, students may choose to take it as a junior or even a senior. It is typically offered in the fall of each year.

Professor Cathy McFadden shared her data on both the components and overall grades of students in Bio 109 from 2000 to 2010. Those data are summarized in Figures 5 and 6:

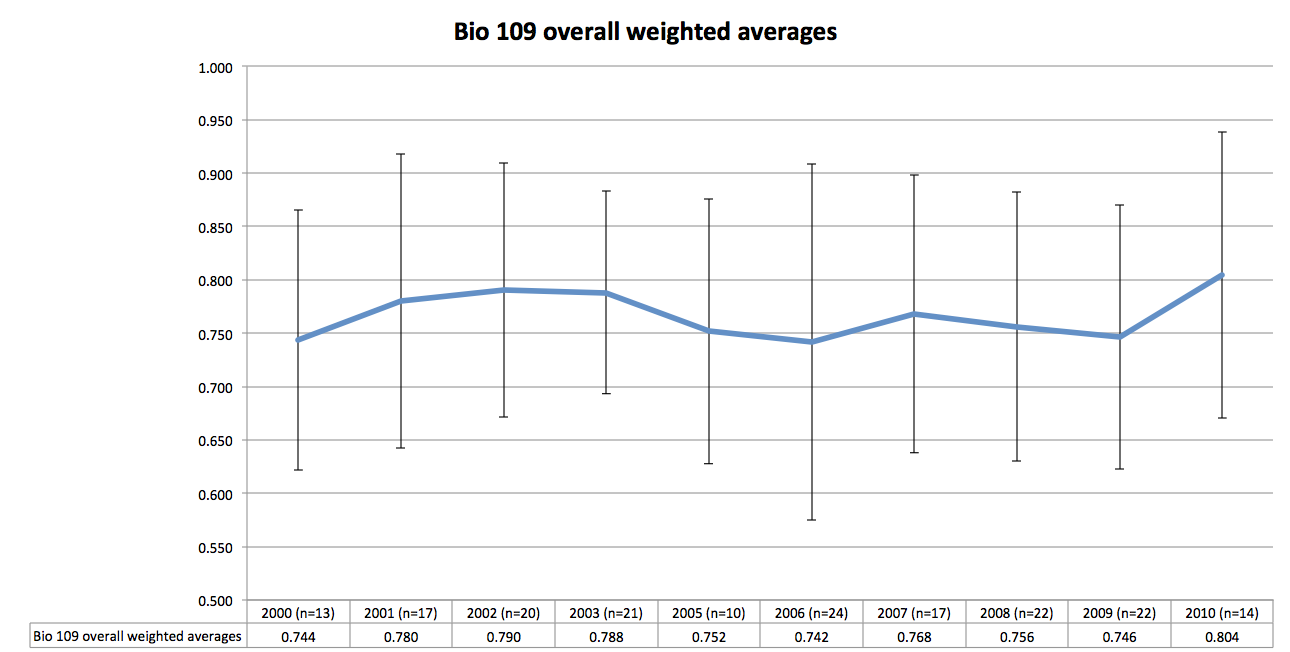


Figure 5. The overall averages from Bio 109's classes spanning 2000 to 2010 (Professor McFadden was on sabbatical in 2004.) Some of 2010's students experienced part of the new core. Error bars represent one standard deviation within each year's cohort.

Figure 5 shows that student scores in Bio 109 have remained relatively consistent over the past decade: the error bars, representing one standard deviation within each year's cohort, easily subsume the small year-to-year changes to the overall averages. The small rise from 2009 to 2010 reverses the change seen in Figure 1, and four of the five highest scores in 2010 were earned by (then) sophomores, who experienced some of the new core curriculum. Because 2011's course was taught by a different instructor (Professor Adolph), those data would not serve as a head-to-head comparison. In fall 2012 Professor McFadden will offer Bio 109 again; those data will be valuable in determining the impact, if any, of students' new-core experiences on Bio 109.

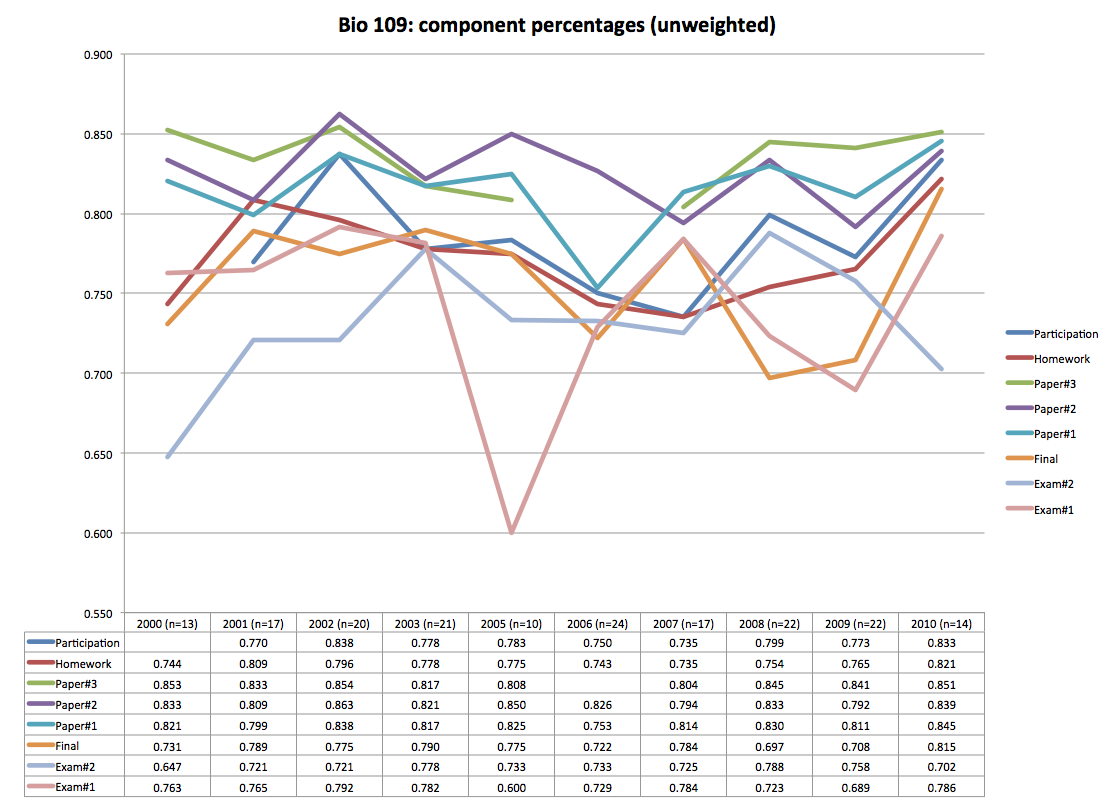


Figure 6. The unweighted component percentages from Figure 4's overall scores. The two empty data points result from small curricular changes during those terms.

Figure 6's summary of the components of Bio 109 scores shows an outlying data point in 2005's first exam, perhaps a result of the relatively small number of students in that class (only 10, in contrast to a steady state closer to or above 20). More generally, these data bear out the -- not uncommon -- tendency that exams result in lower scores than papers, but it does not seem that there have been dramatic changes in student performance across this timespan.

As data are added in 2012 and beyond, Figure 5 will serve as a wonderfully deep reservoir for informing the department and school of possible impacts of the new core.

**Capstone thesis assessment**

In 2013 the first cohort of biologists who experienced *some* of the new core will hand in their senior theses. In 2014 the first group experiencing the *whole* new core will be graduating. The biology department is interested in how the new core's emphasis on developing writing skills, e.g., in the early, cross-departmental Writ 1 course, may impact the writing seen in these senior theses. In addition, the department has expressed an interest in assessing – and improving – students' sophistication in applying statistical analyses in those capstone deliverables.

Currently, the department maintains a library of theses. We propose that in the 2014-2015 academic year the AAC hire a group of readers to assess a subset of theses from both the new-core and old-core cohorts for both writing quality and sophistication in the use of statistical and mathematical tools. We are fortunate that the HSA department has experience in running such studies and, absent a change in direction by those setting the AAC's agenda, we look forward to taking on that assessment of senior theses.

***Chemistry***

The assessment committee want to express its thanks to all of the chemistry department and, in particular, Katherine Maloney, Kerry Karukstis, and Lelia Hawkins for this report.

**QUESTIONS**

The Chemistry department is interested in how the changes in the chemistry core may affect downstream major courses. In particular, we investigate (1) student performance in facets of Physical Chemistry (Chem 51) over the past several years; and (2) differences in student performance in Analytical Chemistry (Chem 103) between Juniors (who took the pilot of the new chemistry core), and Seniors (who took the old chemistry core).

**BACKGROUND**

In 2010 the chemistry core was reduced from 2 full semesters of General Chemistry to three half-semesters divided into the topics of Structure, Energetics and Dynamics. A pilot version of the new core was offered in Fall 2009.

Chemistry 51 is the first course in the chemistry major, taken by all chemistry and joint chemistry & biology majors (typically in the fall of the sophomore year), along with a few students from other majors who elect to take it. The course relies on core math (especially calculus), physics, and chemistry knowledge.

Chemistry 103 is a required course for all chemistry and joint chemistry & biology majors, typically taken in the fall of the junior or senior year. The course relies on core math (especially statistics), along with topics from the chemistry core (*e.g.* acid-base equilibria).

Additionally, as part of the changes to the core, students can now proceed through the chemistry core in a variety of ways. Students can take the chemistry core lab (Chem 24) in either the fall or spring of the freshman year. Although virtually all students take a full semester of core chemistry in the fall of the freshman year, roughly half of these students take Structure before Energetics, while the other half take Energetics before Structure. The final half-semester core course, Dynamics, can be taken in the first- or second-half of the spring semester freshman year, *or* during the fall of the sophomore year.

**DATA USED**

For this report, we gathered:

* Exam data from several offerings of Chem 51, organized by course objective and offering a comparison of pre- and post-new-core results.
* Exam data from Fall 2011’s Chemistry 103, broken out for juniors (who took the pilot of the new core) versus seniors (who took the old core).
* Final grades in Chem 51, broken out by path (23E first vs. 23S first; 23D in first half spring vs. second half spring vs. fall)

**GOALS and SLOs ADDRESSED**

This report directly informs several of the Chemistry department’s stated goals (based on the assessment office’s goals document as of Summer 2011), including

* (from Part I, Goal 1, Student Learning Outcome a) ***Students will demonstrate a basic knowledge of chemistry in the area of structure, dynamics, and energetics***
* (from Part I, Goal 2, Student Learning Outcome a) ***Students will demonstrate sufficient background and depth in chemistry to be able to successfully complete any of the majors offered at Harvey Mudd College***
* (from Part II, Goal 4, Student Learning Outcome a) ***Senior chemistry majors will be able to demonstrate a mastery of factual knowledge comprehensively across the five principal areas of chemistry (organic, inorganic, physical, analytical, and biochemical),…***

**RESULTS**

**Chem 103**

The tables in Figure 1 show the distributions of exam 1 and exam 2 scores for each of the pre-new-core and pilot-new-core cohorts from Fall 2011’s offering of Chemistry 103. On the two exams, the pilot cohort’s averages were 64% and 82% (n=9); the pre-new-core’s corresponding averages were 66% and 83% (n=9). A t-test does not strongly support a significant difference between these two cohorts, with p > .3 for each.

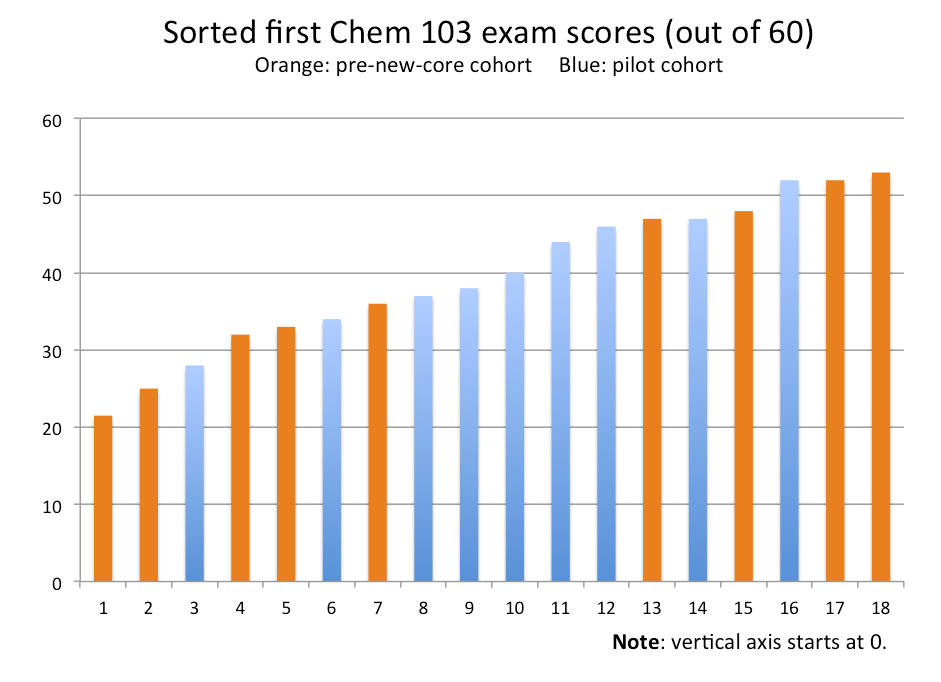
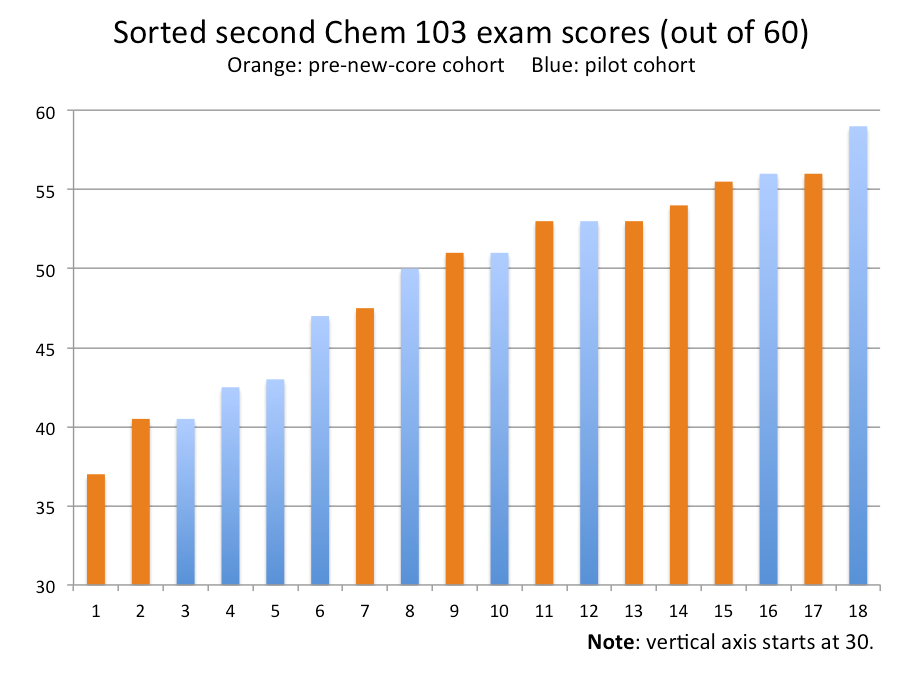
 

Figure 1. The distribution of the first and second exams in Fall 2011’s Chem 103. The datapoints corresponding to students from the class of 2012 (pre-new-core) are in orange; those in blue correspond to students from the class of 2013 (pilot new-core cohort). Averages do not suggest a significant difference between the two groups.

In addition to these exam scores and averages, Professor Hawkins maintained the scores earned on each (sub)question on each exam: these offer an opportunity to compare confidence and preparation at a finer resolution than the overall scores. Although none of the (sub)questions revealed cohort differences significant at p < 0.05, we list here the three questions that did generate the largest differences between the two groups:

Questions on which the pre-new-core cohort earned higher scores than the pilot cohort by the largest margins (the average difference, as a percent of the question’s value, is noted in parentheses):

1. [[Consider the solubility of calcium oxalate, CaC2O4. Please recall that oxalic acid is diprotic.]] Write any applicable acid/base expressions in terms of equilibrium concentrations. (39%)
2. [[Vitamin C (ascorbic acid, C6H8O6) can be measured in vitamin tablets or fruit drinks by a series of redox titrations. As a hint, they are listed in the order they would be performed. In this experiment, potassium bromate is a primary standard. Study the following reactions to determine exactly how this experiment would be conducted. (four reactions are then listed)]] Circle the reagent(s) whose *amount* (moles) in the following reactions must be known at the outset of the experiment (controlled by you but not in excess). (24%)
3. [[same setup as above in #2]] Cloud the reagent(s) whose *concentration* (molarity) must be known at the outset of the experiment. Box the reagent(s) whose amount depends critically on the amount of ascorbic acid in the unknown sample. (35%)

Questions on which the pre-new-core cohort earned lower scores than the pilot cohort by the largest margins:

1. [[same setup as above in #2 and #3]] Just as you finish running your titrations (planning to leave lab early to get in the Mongolian beef line at the Hoch), Prof. Hawkins comes by and says, “You know, we really need another set of titrations to measure the concentration of ascorbic acid accurately.” Explain what titration(s) you might do, what it would tell you, and why Prof. Hawkins is being so picky (creativity is appreciated but try to find a chemical reason, not a personal one). (29%)
2. [[Consider the solubility of calcium oxalate, CaC2O4. Please recall that oxalic acid is diprotic.]] Write a single mass balance equation for this system. (22%)
3. To one significant figure, what is the pH of a solution made by dissolving 0.1 mol of sulfurous acid (H2SO3) in 1.09 L of 0.10 M sodium hydroxide? (22%)

Beyond core changes there are a pair of potentially confounding factors that underlie these results. For one, the pilot cohort consist entirely of chemistry majors; the pre-new-core cohort consists entirely of chemistry-biology joint majors. The difference in curriculum or emphasis may (or may not) be relevant. In addition, the pre-new-core students are (necessarily) an additional year removed from their core chemistry experience: that additional time may (or may not) have had an impact. It is possible to use these data as a starting point and to expand them in the Fall of 2012. The assessment committee will be happy to take that path, if it is part of the coming year’s charge.

The committee is also excited about two in-class “survey assessments” taken by students in Chem 103 last year. One reviewed important ideas about buffers; the other reviewed electrochemistry. Though taken anonymously, students indicated their class year and major. The scores from the two review exercises did not differ significantly between the two cohorts in question, with p >> .3.

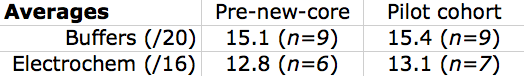


Figure 2. Average scores from two survey assessments asking students to review prerequisite topics for Chem 103. In addition, students indicated their experience with learning the material.

These reviews added an interesting innovation. Each asked students to indicate – overall – their recollection about how they learned and remembered the material in question. Below is a snapshot of one of the questions, showing the options provided (as well as an honest and reflective example answer!)

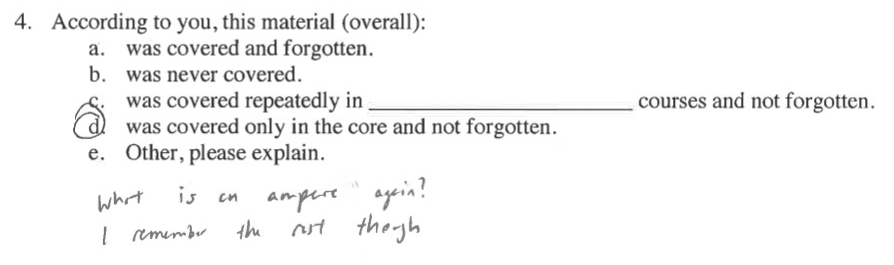


Figure 3. The question asking students to reflect on their past experiences with the review material.

The responses indicated that both cohorts felt the core did prepare them with the material (on buffers and electrochemistry) that is needed in Chem 103, even as they revealed that the reviews were important and welcome!

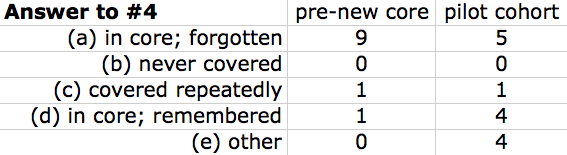
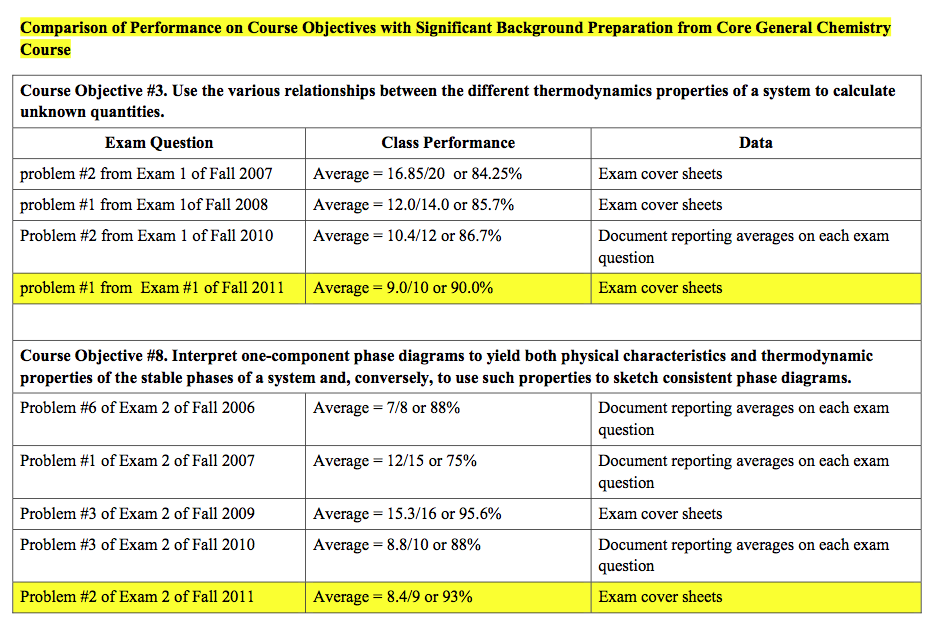


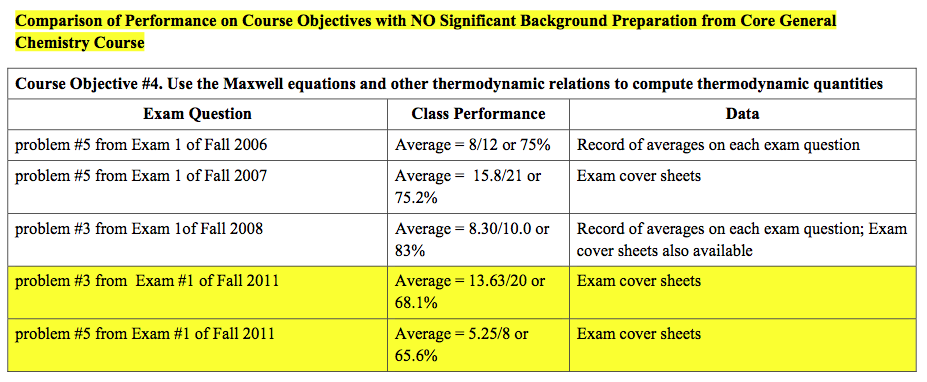
Figure 4. The distribution of the answers to the reveiws’ question #4.

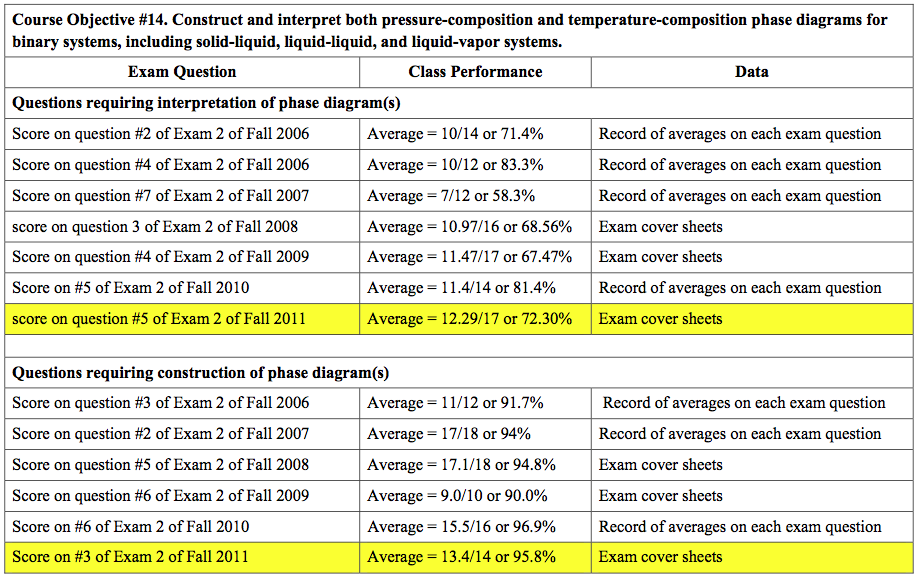
Although the number of responses is not huge, it seems plausible that the younger students remembered more of their core chemistry; the older cohort seemed to remember that they *had* learned the key concepts in the core, even if more of them appreciated the need for review. Notable, too, is that no student of either cohort reported *never* learning the material!

**Chem 51**

In the tables that follow, data are presented to examine student performance on exam questions that assess five objectives in the Chemistry 51 Physical Chemistry course. This course is generally the first chemistry course that students take after the core sequence of Chemistry 23 S, E, and D. Two of the course objectives are ones for which the students have a significant background from Chem 23; three of the objectives represent significantly new material. The Fall 2011 class of Chemistry 51 (results highlighted in yellow) represents the first class that took the new Chemistry 23 core in three half-courses.







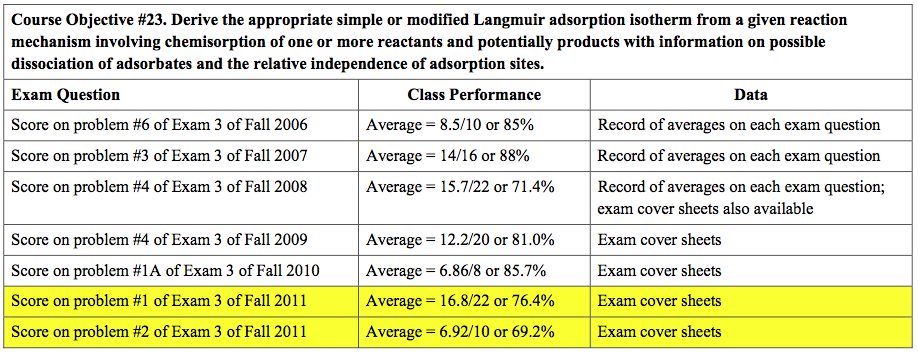


Figure 5 (two pages). Data across several years showing the average performance on two Chem exam questions that *do* rely on preparation from core chemistry and four questions that do not.

**Observations:**

• Students perform better on those course objectives for which they have significant background preparation from Core General Chemistry Course(s) - e.g., typical scores on questions representing 85-95% of the total point value – no immediately variation by year

• For those course objectives for which students have no significant background preparation from Core General Chemistry Course(s), the overall performance is lower than for those questions where they’ve had preparation – e.g. typical scores on questions representing 70-85% of the total point value

* One exception seems to be for questions requiring sketching a binary phase diagram
* For the Fall 2011 class – the class taking the Chem 23 S, E, D sequence for the first time in the reduced three half- course framework – there is some indication that their performance on questions where they have little background preparation is slightly worse than previous years

In particular, the Fall 2011’s scores were the lowest among all of the years for which data is available in both (1) the use of Maxwell equations (and other thermodynamic relations) to compute thermodynamic quantities (objective #4) and (2) deriving an appropriate Langmuir adsorption isotherm (objective #23). The assessment committee thanks Prof. Karukstis not only for all of this Chem 51 data, but also for the write-up and accompanying observations.

**Tracking different paths through the chemistry core**

Beyond the new- vs. old-core differences captured by the above individual exam-question data, students can experience the new core’s three half-courses in several different arrangements.

In the fall of the new core’s first year, the chemistry component of the core curriculum most often consists of the two half-courses Chemistry 23 E (Energetics) and Chemistry 23 S (Structure). Students may take 23E and 23S in either order. As the box-and-whisker plots of Figure 6 illustrate, the students taking 23E before 23S earned a slightly higher median grade than those taking 23S first. In addition, both of these groups had a median slightly, though not statistically significantly, higher than the cohort who had taken an alternative (or earlier) path through the core (the left-hand plot). The numbers in each of these groups, from left to right in Figure 6 (left) are n=5, n=8 and n=11.

To date, the differences are not large enough to indicate a significant statistical difference (p > .3), but the numbers thus far are relatively small.

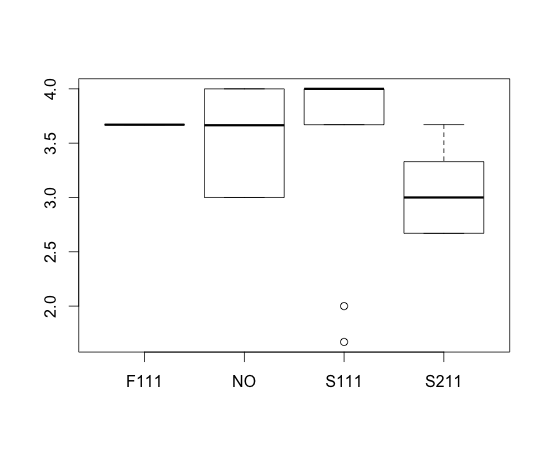
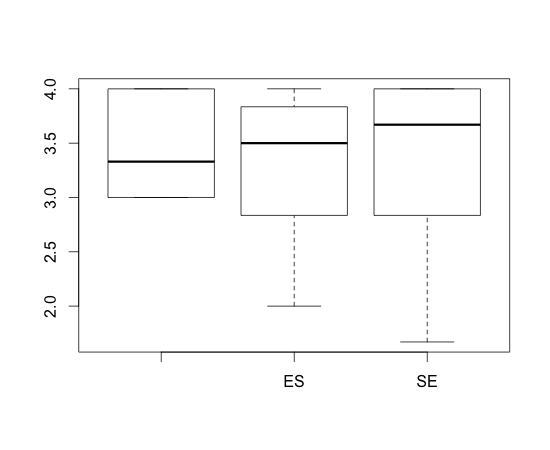


Figure 6. (left) Final grades earned in Chem 51 in Fall 2011 are slightly higher among students who took 23S before 23E; the differences corresponding to 23D’s timing appear at right.

Similarly, the differences in Chem 51 grades, broken out by the time at which 23D is taken, do not show statistically significant differences. That said, this is only the first opportunity for which this data could be collected; the cohort numbers in each case are still small: from left to right in Figure 6 (right), they are n=1, n=6, n=9, and n=8. It is certainly possible that significant differences will emerge in the future.

**Possible future directions for assessment**

The chemistry department and administration will decide whether and how these data-collection efforts will continue, e.g., in order to determine the extent to which these new-core differences persist and/or how they differ from class year to class year. In addition, this report has focused on only a small – albeit important – slice of the HMC chemistry curriculum in Chem 51 and Chem 103. In keeping with the SVCIC’s charge to define “success in major,” future work might broaden the courses, knowledge, skills, and outcomes assessed in order to inform how the new core has impacted majors according to that definition of “major success.”

***Engineering***

**QUESTIONS**

The Engineering department is interested in how the changes in the common core may affect downstream major courses. In particular, we hope to examine how students’ performance in E82, *Chemical and Thermal Processes,* has changed over time. Longer term (not in 2011-2012), the department seeks to investigate how the quality of written work in E4 *Introduction to Engineering Design and Manufacturing* may have changed in recent years.

**BACKGROUND**

Two changes in the new core have a clear potential to affect E82 and E4. The first is the replacement of a year-long Humanities, Social Sciences, and the Arts course sequence (Hum1 and, previously, Hum2) by the first-year sequence of Writ 1 and HSA 10. The second is the reduction of chemistry requirements from two semesters of lecture and two semesters of labs to 1.5 semesters of lecture and 1 semester of lab. This work’s E82 assessment aims to assess chemical engineering mastery, as affected by basic chemistry understanding, across years spanning the Core change.

The E4 assessment, a facet of a broad interest in students’ writing skills, is mentioned here because it is a longer-term goal of the engineering department. It will certainly be possible for the AAC to support that effort – perhaps analogous to the HSA assessment in this report – in the future.

**Data used**

For this report, we have gathered

* E82 final grades over several semesters
  + In addition to an overall comparison across several years, Professor Cardenas’s sections were compared: Spring 2001, Fall 2005, and Fall 2011/Spring 2012
* We also have scanned and archived a number of E82 exam contributed by Professor Cardenas; those are available as a baseline for future evaluations of student work in chemical engineering.

**Goals addressed**

This report directly informs the department – and the college – of Engineering’s success in its Program Educational Objective #1: “*to produce graduates who are exceptionally competent engineers whose work is notable for its breadth and its technical excellence.”*

**RESULTS**

Figure 1 shows the box-and-whisker plots that summarize students’ final grade distributions in E82 over the past decade.



An analysis of variance across these distributions shows that the differences are not likely to be caused by the sampling process alone (p < .01). Looks like there is a very significant difference between F2005 and F2011 but we can't attribute causality to the Core change.  
  
Half a grade better in 2011 and ten times the significance cut off:  
  
Welch Two Sample t-test  
  
data:  E82\_Cardenas$grade by E82\_Cardenas$year  
t = -2.8737, df = 95.203, p-value = 0.005004  
alternative hypothesis: true difference in means is not equal to 0  
95 percent confidence interval:  
 -0.9092486 -0.1662616  
sample estimates:  
mean in group 2005 mean in group 2011  
          2.904694           3.442449

Alas, S2001 is BJT (Before Jenzabar Time) and not something I have access to.

Is there any reason to go back that far?

Wouldn't it make the most sense to compare the last few classes before the Core change to the first few after?

Just in case something else is changing like the quality of high school preparation?

We have dependable data from F2002 on.

Figure 1 shows box-and-whisker plots

Looks like there is a very significant difference between F2005 and F2011 but we can't attribute causality to the Core change.  
  
Half a grade better in 2011 and ten times the significance cut off:  
  
Welch Two Sample t-test  
  
data:  E82\_Cardenas$grade by E82\_Cardenas$year  
t = -2.8737, df = 95.203, p-value = 0.005004  
alternative hypothesis: true difference in means is not equal to 0  
95 percent confidence interval:  
 -0.9092486 -0.1662616  
sample estimates:  
mean in group 2005 mean in group 2011  
          2.904694           3.442449

***Computer Science***

**QUESTION** The CS department is interested in how the different possibilities for core-CS experiences affect students downstream in the CS curriculum. In particular, for this report we consider (1) whether there are differences among the different introductory CS cohorts in choosing CS as a major and (2) how student participation in summer projects before sophomore year correlates with choosing CS as a major. In particular, we want to know whether the above data suggest that students from any of the four introductory paths puts their students is at a *disadvantage* in considering CS as a possible course of study.

**BACKGROUND** In 2008 and 2009, coincident with the development and early trials of HMC’s new core curriculum, the CS department added two introductory pathways into its departmental offerings. The first, CS 42, began in 2008 in response to the rapidly growing group of students with a great deal of computational background. The second, CS5 green*,* started in 2009. Though many factors influenced its genesis, CS5 green is one of the ways in which the computer science department hewed to the goals elaborated in the original “Charge to the Strategic Vision Curriculum Committee,” in particular the one that reads, “the committee should examine the core, including integration of the life sciences.”[[8]](#footnote-8)

CS5 green and CS 42 joined CS5 gold and CS5 black, both formed in 2006, to form four possible paths by which first-semester HMC students can satisfy the CS core requirement. Although the department allows students to choose which of the four sections they wish to join, we provide the following guidelines to help them choose:

* Students with a strong background equivalent to a semester or more of college computer science, e.g., a 5 on the AP exam, college coursework, or a year or more of high-school course work typically are steered towards CS42.
* Students with some background in programming, perhaps a semester at the high-school level, are advised to consider CS5 black.
* Students with modest or no background who are particularly interested in the interplay of computation and biology are encouraged to consider CS5 green.
* Students with modest or no background who have no preference (or would prefer a broader mix of computational applications) typically take CS5 gold.

In the spirit of investigating how core changes affect each department “downstream,” this report summarizes the extent to which these four first-semester cohorts choose to pursue CS as a major. Within that designation we include the combined majors named Computational and Mathematical Biology and the Joint Major in Mathematics and Computer Science. In addition, we report on how pre-sophomore summer experiences correlate with major choice.

Before gathering these data, we hoped that students would feel equally welcome to pursue computer science from any of the four tracks. We recognize that these data, which reflect student choices that depend on many factors, can not directly inform the department about how welcoming it is along each of the four introductory paths. We do hope, however, it can point out whether and where further investigation might be warranted.

**Data Used** For this report, we have gathered

* Major choices of students entering since 2008
* The introductory sections of students since 2008
* The paths of students who participated in pre-sophomore CS projects during the summers from 2005 through 2011. We extend a special thanks to Joyce Greene for maintaining these records over the years!!

All of these data are now in the AAC’s repository.

**GOALS and SLOs ADDRESSED** This report directly informs the department – and the college – of CS’s success in its departmental Goal 3, SLO (student learning objective) #1 *“students will opt to extend their computational education.”* That said, we recognize that there are several additional sources of data we might use to inform that goal, e.g., CS 60 numbers and surveys about self-directed computational projects. We summarize these and other possible future directions in the final section of this report.

**Results** We exclude data from the class of 2015, as almost none of those students have declared a major as of this writing. We note, however, that we will be able to leverage the work done for this report in order to extend this analysis to the class of ’15 quickly.

As a first-level breakdown of the data, Figure 1 presents a chart of the current sophomore, junior, and senior classes (’14, ’13, and ’12, respectively) comprising 574 students as of Feb. 2012. 154 of that total are majoring in CS -- again, with combined majors included in that number. The sub-chart indicates the number of students from each of the four introductory pathways through the CS core.

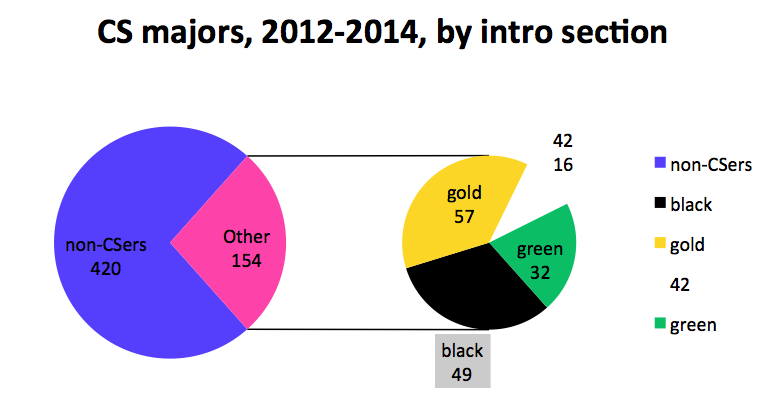


Figure 1. Breakdown of introductory CS experiences for the 154 majors, ’12-’14

These overview numbers neglect two important considerations: the very different sizes of the gold, black, green, and 42 cohorts and the possibility of year-to-year variation in those groups. Figure 2 presents the same absolute numbers that appear in Figure 1, now broken into the raw numbers of each of the past three years’ majors and non-majors from each section.

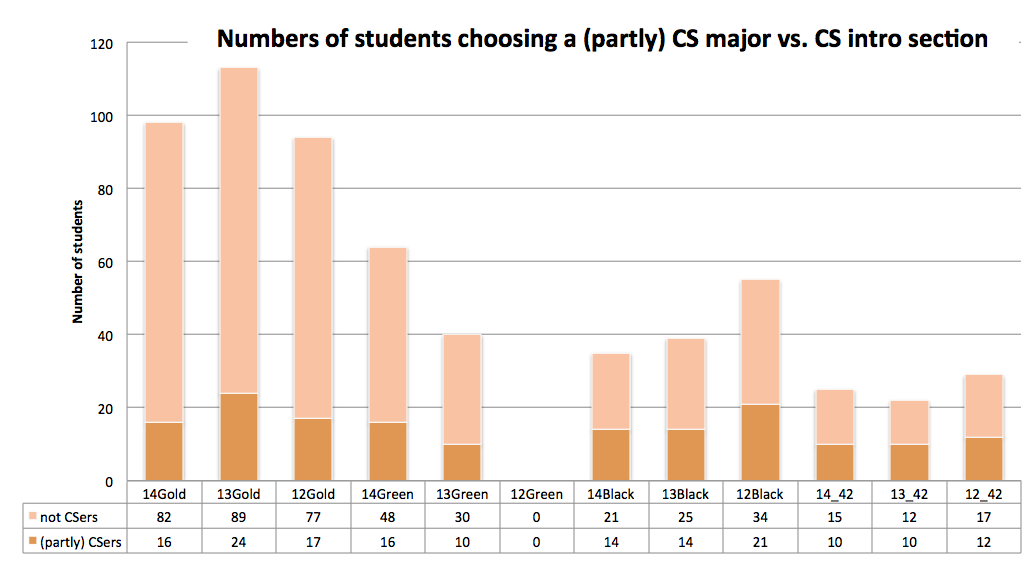
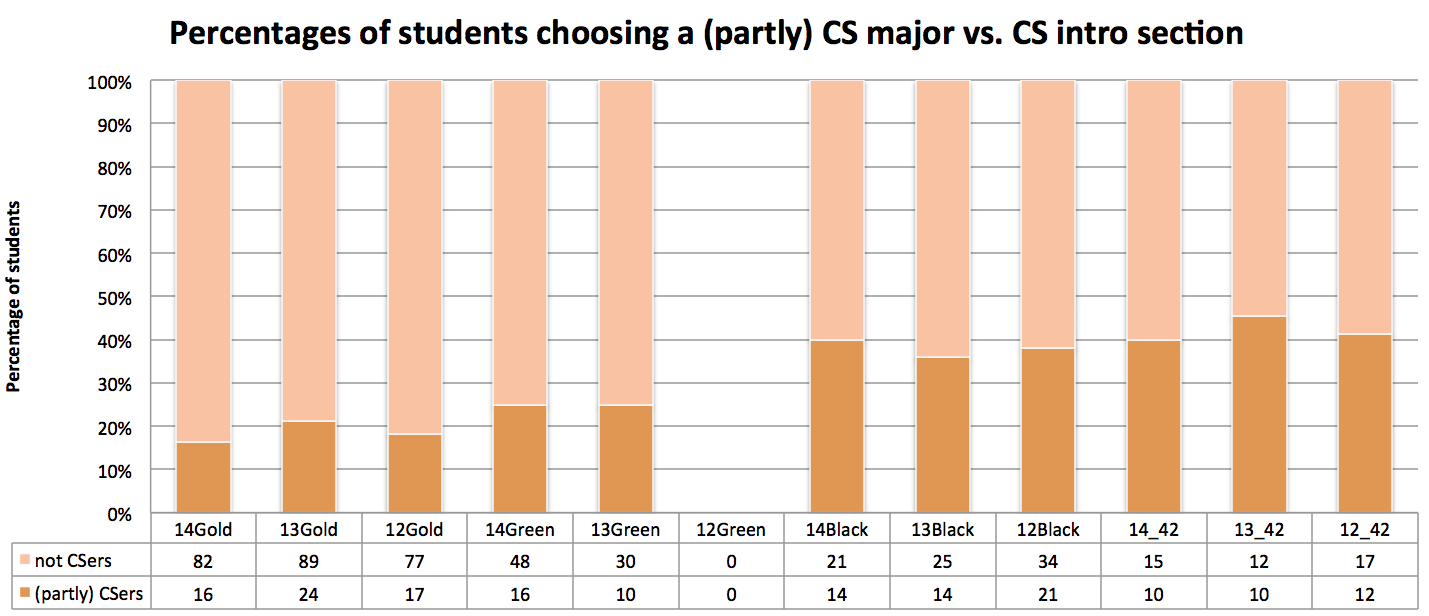


Figure 2. Detail of major and non-major numbers by year and introductory section. The (partly) is as a reminder that joint majors and math/cs/bio majors are included.

Figure 2’s data make it clear that, although more majors have joined CS via that CS5 gold course than any of the other paths, the rate of students continuing to the major from CS5 gold is lower than any of the other paths. Figure 3 recasts Figure 2’s data in terms of percentages, instead of raw numbers of students.



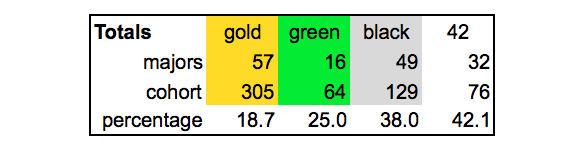


Figure 3. (top) Figure 2’s data shown as percentages along the vertical axis. The data table underneath summarizes these values across the introductory cohorts.

Here, it’s clear that the data support what might seem a natural presumption: a greater percentage of the students with experience in computation tend to go on to major in the field. One possible surprise is that CS5 green yields a greater percentage of CS majors than CS5 gold, its natural comparison group. Although with the current data, this difference is not statistically significant (p>0.2 for both one- and two-tailed Fisher exact probability tests using the online resources at faculty.vassar.edu/lowry/odds2x2.html), this is a trend that the department will keep tracking into the future.

In contrast, it is the case that the differences between gold and black and those between gold and 42 *are* significant at the p<0.01 level. These data do not shed light, however, on the factors involved in the failure of gold to match other sections’ interest. It seems likely that both classroom effects and students’ background differences contribute, though in unknown relative proportions. What these statistics do provide, however, is a baseline against which to compare future departmental data.

One experience that does turn out to correlate strongly with increased interest in the CS major is pre-sophomore summer work with the department. Figure 4 shows the number of students participating in a CS summer project since 2005, along with the number who opted a major overlapping with CS. Here, the percentage is about 60%. Certainly only students interested in CS would consider summer work with the department, but it’s also worth pointing out that the majority of these rising-sophomore participants did take the gold introductory section. At the very least, these data suggest that the pre-sophomore summer experiences benefit both the participating students and the CS department.

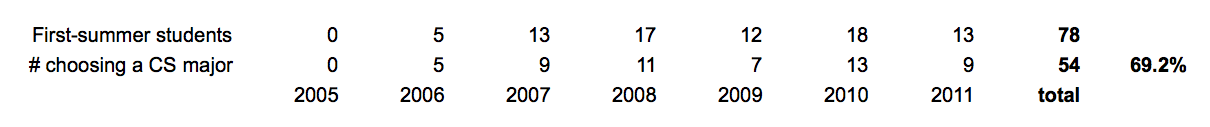


Figure 4. The number of students who participated in research or other open-ended CS projects in the summer (horizontal axis) before their sophomore year, along with the numbers choosing a major involving CS.

We don't have these summer-project numbers broken down by introductory CS section, but at least half of these students were in gold or green. Note that the percentage choosing to pursue additional CS is higher even than for students in CS 42 – this is to be expected, given that only the most interested students would apply for summer positions. Even so, it is clear that early summer participation is providing initial opportunities from which many students are building longer-term plans.

**Perspective**

In sum, the department is heartened that students from all four of the college’s introductory CS backgrounds have chosen to pursue CS as a major – on a consistent basis for the past three years. The pre-sophomore summer opportunities also seem to help generate interest in the field. The department will continue to monitor its least encouraging pathway, CS5 gold, in order to best serve the interests and needs of those students in the future.

**Possible future directions for assessment**

These data offer a number of opportunities for more deeply investigating how well the CS curriculum serves HMC students. First, the department should continue to monitor whether all of the pathways into the discipline are serving students well. Second, the department might consider emulating the example of mathematics, who take a carefully constructed student survey on students’ experience of their core curriculum at regular intervals. Such surveys would help understand why gold less effectively inspires further CS study than other introductory pathways.

Another direction in which we could expand these inquiries would ask how students from different core courses fare in subsequent experiences with CS, e.g., how well they do in CS 60 and beyond. In addition, many students opt to take CS 60 without choosing a CS major: those data could further inform the department’s – and this report’s -- goal that HMC students opt to extend their computational pursuits.

***Humanities, Social Sciences, and the Arts***

**QUESTION** The HSA department is interested in how the recent restructuring of first-year writing courses affect the students “downstream” in their ability to write effectively. More specifically, we’d like to investigate whether the students made greater and steadier progress during the two composition courses they were required to take in the first year.

**BACKGROUND** Prior to 2007, freshmen at HMC were required to take Humanities 1 and 2 (HUM1 and HUM2) in their first year, both of which were writing intensive and taught by the faculty members in HSA department. From 2007 to 2010, only HUM 1 was offered. Since the new curriculum was implemented in 2011, first-year students took a college-wide half-semester writing course (Writ1) in the fall and a full-semester writing course from the HSA department (HSA10) in the spring. The designing of HSA10 bears resemblance to HUM1, yet it has been restructured substantially to carry on and advance what students learned in Writ1, to introduce students to the wide variety of perspectives and methods in the field of humanities, social sciences and the arts, and to guide students towards independent, exploratory thinking and research.

In 2011 and 2012, the HSA department organized two writing assessments that compared student papers produced in Writ 1 and HSA10. For this report, we will compare the papers in early arrangement (HUM1 & HUM2) with those of the present (Writ1 & HSA10). For both assessments, we employ external readers to evaluate the individual papers according to provided rubrics. The papers are randomly selected, although we also managed to distribute the samples evenly across all sections and to have approximately the same number of male and female students.

**DATA USED** For this report, we use the following sets of data:

* The writing assessment of student final papers in Writ 1 as well as their midterm and final papers in HSA10 (2011 and 2012). 40 sets of papers were randomly selected in 2011, and 60 in 2012.
* The writing assessment of student midterm and final papers in HUM1 as well as their final papers in HUM2 (2004). 33 sets of papers were randomly selected.

We are aware that the recent samples greatly outnumbered the earlier ones. The 2004 assessment also examined different sets of papers across a varying time span. However the criterion and method we used to evaluate those papers remain consistent, and we believe that a comparative study will yield meaningful, though inconclusive, results. In the following analysis, we will focus more on paper 2 and paper 3 in the 2004 assessment, and paper 1 and paper 3 in the 2011 and 2012 ones, since they reflect the changes students went through from the end of the 1st semester to the end of the 2nd.

**GOALS AND SLOs ADDRESSED** This report directly informs the HSA department’s Goal 1, Part 4: *to provide a curriculum that fosters excellence in critical reading and thinking, and in writing.* Since both HUM1 and HSA are taught by faculty members representing a wide range of disciplines (11 for HSA10), the report also informs the department’s Goal 1, Part 1, *to provide a curriculum that allows students to develop an understanding of the questions, methods and content of our disciplines.*

**Results**

For all the three assessments, a stratified random sample of students was selected from the class rosters. For each student selected, three assigned papers were collected – two from Humanities 1 and one from Humanities 2 in 2004, and one from Writ 1 and two from HSA10 in 2011 and 2012, respectively.

To assess writing skills, a four-dimension rubric was developed. The four dimensions were related to the quality of the thesis and arguments, the logic and structure, the writing style, and an overall evaluation of the quality of the paper. For each dimension, there were five levels of performance, with concrete descriptors for each level (see **Appendix**). A higher score indicated a higher level of demonstrated skill/quality.

A workshop was conducted to train raters to apply the rubric. The workshop included an explanation of the rubric and procedures for applying it, question and answer periods, and several practice sessions in which all the participants read through the same papers, rated them according to the rubric, and exchanged their thoughts.

All identifying information was removed from the papers, and each paper was scored independently by two readers. Inter-rater reliability was calculated in both the 2004 and the 2012 assessments, and the results show satisfactory, though by no means perfect, consistency between the 1st and the 2nd readers. Mean scores were calculated for each rubric dimension on each paper.

1) Progress from the 1st semester to the 2nd semester, measure by mean scores.

As the charts below demonstrate, in the 2004 writing courses, the students showed visible though insignificant progress in all but the “logic & structure” category from the 1st to the 2nd paper in HUM1, yet the scores dipped slightly again at the final paper for HUM 2.

In 2011, the results show a small decline from the last Writ1 paper to the HAS 10 midterm paper, followed by a rebound on the HSA10 research paper, which raised the scores to a similar level as those in Writ1.

In 2012, however, the numbers are consistently going upward in all the four categories. If we compare the final papers in Writ1 (Paper 1) and those in HSA10 (Paper 3), the differences are significant and unmistakable in all but (perhaps) writing style, and the “overall grade” improved the most.

We will continue to investigate what has led to the impressive changes in 2012, but it is reasonable to deduce that our restructuring of HSA10 in 2012, which required students to revise their midterm papers and gave them more hands-on assistance during their research for the final papers, has produced desirable results.

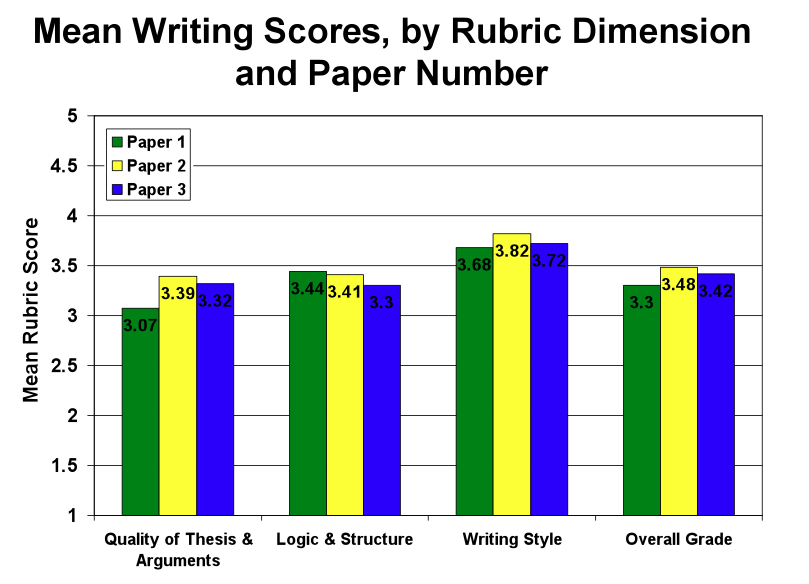


Figure 1: 2004 Assessment (n=33)

2) Progress made from the 1st semester to the 2nd, by groups with different “starting points”.

In order to measure the progress of students with different “starting points”, students’ overall average scores on the three assessed papers are disaggregated by the range into which their average first paper fell: 2.5 or lower (Low Score), 2.6-3.9 (Mid-range Score) and 4.0 or higher (High Score). In 2004, only “Low Score” and “High Score” were calculated.

We observe that, in all three years, the students with the lowest score in their first paper made the greatest progress by the end of their 2nd semester. The results in 2011 were the most drastic, in which those started “low” eventually earned the highest average scores. Apparently, those students worked really hard in HSA10 to catch up, and they usually received the greatest amount of attention from both the instructors and the tutors, who were notified of their performances in Writ1 at the very beginning of HSA10.

On the other hand, those started out with the highest scores tended to do worse in their following papers. They rebounded slightly by the end of the 2nd semester in both 2011 and 2012, yet still not to their initial level. We wonder whether they felt that they already could write well enough, therefore didn’t make much effort in HSA10.

Another notable difference is that, in 2012, students started with the “mid-range score” didn’t “dip” in their 2nd paper; instead, they made steady if varying progress throughout the 2nd semester. Since they are the largest group in the assessment, we deduce that their performances contributed to the changes we saw in the earlier analysis.

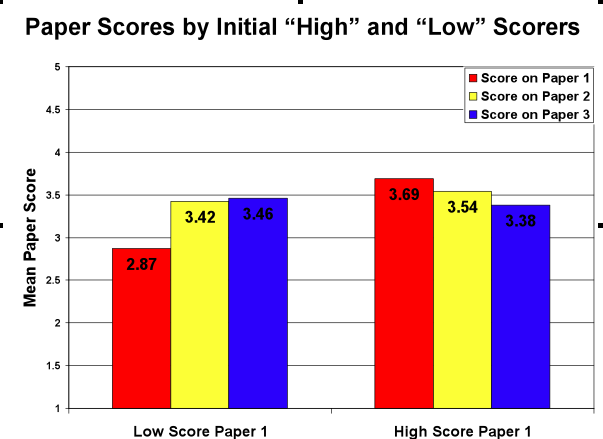
 Figure 2: 2004 Assessment (n=33)

Figure 3: 2011 Assessment Results (n=40)

Figure 4: 2012 Assessment Results (n=59)

3) Progress made from the 1st semester to the 2nd semester, measured by students in different improvement categories.

Finally, we looked at the percentage of students whose performances in the 1st year exhibit certain patterns of change. We defined those patterns in five categories:

*Steady increase* Each successive paper scores higher than the previous, with cumulative increase of 1.0 points or more

*Steady decrease* Each successive paper scores lower than the previous, with cumulative decrease of 1.0 points or more

*Tanked* One paper scored 2.5 or lower, and other two scored 3.5 or higher; or two papers scored 2.5 or lower

*Static* All three papers score within a .5 point range, decreasing or increasing

*Varied* Scoring increased and decreased between papers

In 2004, the pie chart illustrates the overall inconsistency in student writing performances. Few displayed steady increase (1) or decrease (2), while the majority fell into the “varied” category.

In 2011 and 2012, the patterns became more meaningful. Almost half of the students were either “static” or “tanked,” meaning that they maintained a certain level of performance though did poorly sometimes. While those whose scores changed steadily were still a minority, we are pleased to see that those who displayed “steady increase” greatly outnumbered those who displayed “steady decrease” in 2012. These results also corroborate well with the changes we noticed previously.

**Conclusion and Future Assessment**

The writing courses in the new core seem to have helped students to make more consistent and perceptible progress in their freshmen year. In the first year of its implementation, the results were still somewhat mixed, yet the restructuring of HSA10 in 2012 has produced unambiguous improvements. As before, those who started with low scores improved the most in their first year, yet those in the mid-range also saw their writings growing stronger over time, and fewer and fewer students experienced declining scores. However, more motivations need to be provided for the students who performed well in Writ1, who seem to have benefited little from an additional semester of writing.

Given the “almost too good to be true” results in 2012, we recommend one more assessment to be carried out in 2013, perhaps on the same scale (40 sample sets) as that of 2011. Since we don’t plan to make major structural changes in HSA10 in Spring 2013, the additional assessment will help us to evaluate the effectiveness of our current arrangement with more confidence.

Appendix: HSA 10 Assessment Rubric

*Quality of Thesis and Arguments*

**Does the paper present a significant and insightful thesis, supported with appropriate evidence/argument (including textual analysis where needed)?**

**5** The paper does an excellent job presenting and supporting a thesis.

**4** The paper does a good job presenting and supporting a thesis.

**3**  The paper does a fair job presenting and supporting a thesis.

**2**  The paper does a poor job presenting and supporting a thesis.

**1** The paper lacks a thesis or fails to offer any support for its thesis.

*Logic & Structure*

**Is the paper clearly structured, with a readable and logical flow in which all points relate to and are necessary for the writer's conclusion?**

**5** Yes. The paper's points are cohesive, logically connected, and easy to follow, with no extraneous material.

**4** The structure is generally cohesive and logical, though it may need improvement in some areas of the paper or there may be extraneous material that should have been omitted.

**3** Although one can follow the general contours of the argument, there are significant problems of structure or significant, unwarranted digressions from the paper's main line of argument.

**2** The paper's structure often obscures its main line of argument; the ideas in the paper often seem disconnected from one another.

**1** The paper lacks any sense of structure or cohesiveness.

*Writing Style*

**Is the writing clear, understandable, free of grammatical and mechanical errors?**

**5** Yes. The writing is engaging and free of errors.

**4** Some sentence structures could have been improved, but errors are minor. The writing is basically clear.

**3** Some sentences are confusing or contain serious grammatical/mechanical errors.

**2** There are substantial problems with sentence clarity and serious grammatical/mechanical issues.

**1** The paper is pervasively unclear and ungrammatical.

*Overall Grade*

**What overall grade would you assign this paper? (Feel free to take into account criteria other than those listed above.)**

**5** Excellent

**4** Good

**3** Fair

**2** Poor

**1** Failing

Section 9: *perspective and assessment*

It is unlikely that any educational experiment will yield entirely positive or entirely negative results. The new core curriculum is no exception. In encouraging students to take language courses and in increasing retention (at least in early semesters), the new core seems poised for success. However, early diversity data and some of the student voices undergirding this report show that other institutional goals remain works in progress.

In addition, there are effects of the new core that the assessment committee has not yet begun to investigate. We hope that this report augments the insights from 2010’s update by the SVCIC[[9]](#footnote-9), and that both represent downbeats in a rhythmic bass sustaining the college’s core assessment.

Fully recognizing that these melodic variations may change, this year’s committee proposes the following straw-man schedule in order to advance the discussion of how to proceed from here:

* [2012-2013: ***success in major I; “sidecar” courses***] The committee might revisit all goals in this report, especially diversity goals (WASC request (i), below), replacing the per-department section on “downstream impacts” of the new core with “success in major” criteria, defined by each department. This would leave 1 year to collect data on “old core” students, including data from students who have already graduated, in support of whatever criteria departments choose for defining “success” in their majors. Becausde pre-vs.post-core comparisons won’t be possible for senior year work, the committee could imagine an assessment of how “sidecar” courses have helped students succeed in each of HMC’s departmental offerings, regardless of major.
* [2013-2014: ***success in major II; choice labs***] Here, it will be possible to start gathering comparative data for students’ success in each major, before and after the new core. These data could become the per-department focus of the ’13-’14 report. In addition, the report might summarize the student and faculty responses to the new core’s *Choice Lab* requirement, particularly as they inform the interdisciplinary goals of the new core. This focus on new-core vs. old-core success in major would provide an opportunity to address WASC’s request that HMC pay earnest attention to its educational effectiveness (ii, below).
* [2014-2015: ***student writing***] For many students, senior theses and clinic reports are capstone pieces of writing that integrate their technical abilities and their communications and teamwork skills. During this academic year, the assessment committee will have a large number of new-core students’ capstone writing products. As noted in section 8, above, old-core students’ reports are currently being archived; those samples will be available, too. As a result, this may be an opportune year to focus on comparative assessment of students’ writing, using external readers in the way that the HSA department has done for their recent assessment of HSA10 and as the engineering department plans to do in order to assess E4 reports. This year’s focus on writing would serve as the per-department contributions to the core assessment and could also facilitate an examination of students’ recognition of the social impact in capstone experiences, as WASC would to see (iii, below).
* [March 1, 2016: **WASC interim report**] This is the date on which WASC has requested an interim report on (i) diversity, (ii) educational effectiveness, and (iii) capstone courses, as noted in the *Commission Action Letter* of July 5, 2011.

Although seismic shifts may occur within the college and its priorities between now and these future years, at least the above timeline provides one possible vision for a fuller assessment of the core than the 2011-2012 committee has composed.

**Assessing the assessment committee**

The faculty handbook lists a number of “functions” of the assessment committee:

*(1) To coordinate the assessment of institutional goals and outcomes;*

*(2) To assist the faculty and the college generally in developing and implementing assessment practices;*

*(3) To disseminate the results of assessments to key college constituent groups;*

*(4) To participate in accreditation reviews of the institution; and*

*(5) To coordinate its efforts with the activities of those of the Office of Institutional Research*

*(6) To assess its own efforts*

The 2011-2012 committee has sought to fulfill these functions, with this report contributing to (1), (3), (4), and (5).

In order to fulfill (6) and (2), we define the following metric for assessing the assessment committee:

*The assessment committee considers itself successful in proportion to the number of HMC faculty who substantively contribute to the committee’s assessment report, without requiring unusual or undue time commitment in those contributions.*

Although lofty, the appropriate cutoff for outright “success” on this metric should be *full* participation by (non-sabbatical) members of the HMC faculty. By involving everyone this criterion buttresses a culture of institutional self-reflection. In doing so, however, the assessment committee must take pains to make such efforts sustainable. That is, participation should require very little effort beyond that which is *already* part of everyday faculty activities. One hour per year as a general target might be a reasonable goal for investment, that is, beyond the day-to-day practices that make assessment one of our most widely shared bonds!

We note that there are genuine differences of opinion – both within the committee and outside of it – as to what number 2’s “practices” might mean. At least some voices have suggested that, in this context, praxis is practice. Thus, the need for number 2 could be satisfied by faculty members via the ongoing creation, dissemination, and reflection on students’ growth and learning, on our institution’s stated goals, and on the collimating influence of accreditation reviews – to which we hope this report will contribute.

In 2011-2012, the number of faculty who substantively contributed to these new-core assessment efforts was determined to be *at least* 30 out of 83, for a participation rate of about 36%. (To respect privacy, the related raw data are kept out of the repository.) Although this is short of the full participation needed for unqualified success, this failure is not abject. Rather, this result provides a starting point upon which future assessment committee efforts will improve. Alternatively, future assessment committees are equally welcome to redefine the criteria by which they measure their group’s success.

The 2011-2012 committee thanks all of the year’s many contributors for their time, their work, and – most of all – their lightheartedness and kindness throughout the process of putting together this report!

Respectfully submitted,

The 2011-2012 assessment committee

1. http://www.math.hmc.edu/~su/svcc/SVCC-charge.pdf, courtesy of Francis Su [↑](#footnote-ref-1)
2. http://www.math.hmc.edu/~su/svcc/SVCC-proposal-revised.pdf, courtesy of Francis Su [↑](#footnote-ref-2)
3. http://www.hmc.edu/files/institutionalresearch/standardsandcfrs2/svcic\_report-feb2009.pdf [↑](#footnote-ref-3)
4. quoting *The Magic Eight Ball*, a popular oracle [↑](#footnote-ref-4)
5. These thresholds were obtained from http://aspe.hhs.gov/poverty/12poverty.shtml [↑](#footnote-ref-5)
6. "significant" in the sense that the chance of so large a *T* value occurring under the null hypothesis of identical distributions is less than 1%. The critical value comes from http://www.weibull.com/RelGrowthWeb/Appendix\_B\_Critical\_Values\_for\_Cramer-von\_Mises\_Test.htm . [↑](#footnote-ref-6)
7. Personally, I had never heard of CvM. It's well-explained by Wikipedia and originally from [Anderson, TW (1962). "On the Distribution of the Two-Sample Cramer–von Mises Criterion" (PDF). The Annals of Mathematical Statistics (Institute of Mathematical Statistics) 33 (3): 1148–1159.] [↑](#footnote-ref-7)
8. http://www.math.hmc.edu/~su/svcc/SVCC-charge.pdf, courtesy of Francis Su, SVCC chair [↑](#footnote-ref-8)
9. http://www.hmc.edu/files/institutionalresearch/EER/Core/SVCIC\_2010\_report.pdf [↑](#footnote-ref-9)