A Biology-based CS1: Results and Reflections, Ten Years In

Zachary Dodds  
Harvey Mudd College  
Computer Science  
Claremont, CA, USA  
dodds@hmc.edu

Malia Morgan  
Harvey Mudd College  
Computer Science  
Claremont, CA, USA  
morganm@hmc.edu

Lindsay Popowski  
Harvey Mudd College  
Computer Science  
Claremont, CA, USA  
lpopowski@hmc.edu

Henry Coxe  
Harvey Mudd College  
Computer Science  
Claremont, CA, USA  
hmcoxe@gmail.com

Caroline Coxe  
Harvey Mudd College  
Computer Science  
Claremont, CA, USA  
carolinecoxe@gmail.com

Kewei Zhou  
Harvey Mudd College  
Computer Science  
Claremont, CA, USA  
zzhou@hmc.edu

Eliot Bush  
Harvey Mudd College  
Biology  
Claremont, CA, USA  
bush@hmc.edu

Ran Libeskind-Hadas  
Harvey Mudd College  
Computer Science  
Claremont, CA, USA  
hadas@cs.hmc.edu

ABSTRACT
For a decade, our institution has offered both a biology-based CS1 (CS1-B) and a traditional, breadth-based CS1. This project follows the paths of students in both courses – tracking their subsequent interests (what courses do the two groups choose afterwards?) and their grades in those courses. Within the biology-based cohort, we also contrast the futures of the students who chose a biology-themed introduction with the group who expressed no preference or requested the breadth-based approach. Even when student preference was not accommodated, equitable downstream performance results hold. We discuss the implications of these results, including the possibility that, like introductory writing, introductory computing is a professional literacy in which many disciplines have a stake.

CCS CONCEPTS
- Social and professional topics → CS1.

KEYWORDS
CS1, Computing-as-Literacy, CS for Insight, Computing in Biology

1 MOTIVE: COMPUTING-AS-A-LITERACY
This paper shares the results of a decade-long partnership in which CS and Biology faculties have taught computing through an interdisciplinary collaboration. Since 2009, every first-semester student has been required to take either a traditional CS1 or a biology-based CS1 (CS1-B), with subsequent performance and subsequent pathchoice (of courses and major) tracked. The outcomes downstream from CS1-B are heartening, both in computing skills gained (no difference in subsequent performance relative to a peer control-group) and in spurring overlapping interest (CS1-B students matching CS1 students in CS interest and exceeding them in biology interest). Even when students’ choice of an introductory computing course could not be accommodated, these equitable outcomes hold.

Our motivation for this experiment has been the hypothesis that computing is valuable beyond CS: CS does not own computing. Yet CS has a unique role in ensuring all students, regardless of academic identity, have access to computing’s skillsets. The “CS for All” initiative intended to equip K-12 students with the experiences needed to be “creators in the digital economy.” [15] This philosophy, we believe, should extend into college curricula. As CS1-B’s results suggest, it is possible for a venerable discipline to preserve its academic identity while still leveraging the tools computing has to offer.

For us, the collaboration between the CS and Biology departments has had lasting benefits beyond the shared curriculum. Specifically, computing need not have a CS identity to be valuable. We have found that, in many cases, it is more valuable without that burden. CS for All’s “CS,” in fact, may be misleading: the effort’s value lies not in proselytizing CS majors, but in framing computing as a literacy shared by all.

2 CS1: BIOLOGY VS. BREADTH?
With computing’s rise as a professional literacy [9], CS1 has become an increasingly common part of the college experience. Many feel that, just as writing is exercised through lenses across a university
curriculum, so too is computing a valuable means of inquiry, useful for developing insights in diverse disciplines. This increase in interest prompts an as-yet unresolved question, “How should college students develop computing’s mindsets and skillsets?” Does CS1 need to be factored out, or might disciplinary-developed approaches work as well? This ten-year experiment suggests the latter.

2.1 Related work
Contextualized computing abounds, as it should, e.g., [3, 11, 12, 14], CS1-B differs from such efforts in two ways. First, ownership: the class is a collaboration, with equal authorship from Biology and CS professors. Second, in curricular intersection: CS1-B intends to serve Biologists and CS’ers equally, acknowledging both the intersection and the union of those identities. Prior assessments of such collaborations have been less common than both contextualized computing and CS+X efforts; examples such as [10] emphasize the importance of reaching across not only disciplinary, but also identity, boundaries. Our experiment has maintained shared ownership of CS1-B, featuring five different instructors (including only-Biology, only-CS, and hybrid efforts) over the past decade. That is, CS1-B has become part of the institution, and it does not depend on any individual for its existence. Figure 1 tracks our cohort sizes in CS1 and CS1-B since 2009.

![Figure 1: CS1 and CS1-B cohort sizes, per year](image)

This represents all 1170 CS1 and 268 CS1-B students from 2009 to 2019. Every first-semester student must take introductory cs. These are students without prior computing experience; other paths serve students with substantial background. CS1-B was not able to be offered in 2013.

Both CS1-B and CS1 cover (almost) the same CS1 topics: (a) scripting with Python; (b) problem-solving using iteration and recursion; (c) data-structuring using lists, strings, dictionaries, and self-defined objects; and (d) a contextual arc that seeks to unify these skillsets and mindsets into a coherent narrative. This last point is the crucial difference: for CS1, that narrative is “Computing’s Applications,” writ large. Course examples draw from physics (simulations), language (prose analysis and generation), mathematics (fractals), CS itself (automata), and beyond (sound, images). For CS1-B, the story is “Computing’s Applications in Biology,” with a slate more focused and equally rich: molecular transformations, gene-finding, phylogenetic tree-building, RNA folding, approximate-string matching, and E. Coli pathogenicity modeling via genetic analyses, among many others. Additional details and assessments of each course’s curricular content appear elsewhere [4–6]. In both courses, all of the examples and assignments – and about half the classes’ meeting time – are dedicated to their respective narrative arcs.

2.2 Post-CS1 paths and performance
Having run these two introductory-computing experiences for a decade, we have tracked the post-computing experiences of the 1170 students in CS1 and 268 in CS1-B. For each student, we gathered data on the the CS and biology classes they took in subsequent semesters, including grades, as well as their choice of major.

We first examined whether taking CS1-B took a toll on those who continued studying CS. We found that students continued to take CS courses and declare CS majors at similar rates, regardless of which CS1 flavor they took, as shown in Table 1. That CS1-B students, who often have a particular prior interest in biology, took more biology courses and majored in biology at a higher rate than the “control” CS1 cohort is not a surprise. Indeed, this is heartening, because it suggests that contextualized, early-career computing can complement student interest and identity in a non-CS field. Computing, after all, is as much biology as it is CS.

<table>
<thead>
<tr>
<th>CS Flavor</th>
<th># Courses Taken</th>
<th>% Majoring in</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CS Bio</td>
<td>CS Bio</td>
</tr>
<tr>
<td>CS1</td>
<td>5.4 ±0.2</td>
<td>2.0 ±0.1</td>
</tr>
<tr>
<td>CS1-B</td>
<td>5.3 ±0.4</td>
<td>5.6 ±0.4</td>
</tr>
</tbody>
</table>

Table 1: Downstream choices in Biology and CS

This is an average, and includes the Intro CS course taken. One standard deviation (s.d.) σ is shown.

Relative to CS1, over its first decade, more of CS1-B’s students have majored in CS; many more have majored in Biology. CS course-choice is comparable (enrollment pressures are also a factor at play in the left-hand column).

As Table 2 attests, students who took CS1-B opted to take CS2 and CS3 at approximately the same rate, even slightly more often, as those in CS1. Gradewise, they performed as well as their peers who took CS1 in those subsequent CS classes. Here, we do not seek to adjudicate a “better” path (so no p-values). In fact, it has been essential to the long-term success of this departmental partnership that the effort sustains both academic worldviews! Our goal is to well-serve all students, regardless of academic identity.

We do not anticipate that either intro-cs path should replace the other. Rather, we seek to highlight that the evidence – over the first decade of this experiment – indicates that alternative paths can successfully convey CS1’s skillset, while simultaneously serving as an identity-sustaining introduction to computing. CS1-B has prepared its students for later CS courses at least as well as CS1.

We believe this result could be seen as a freeing one for some CS departments. With a collaboratively-constructed vision in place, 1CS2 and CS3 reflect their sequencing, not their course numbers.
sibling departments can fashion student-paths into computing, whether per se, or towards a CS major, or as a resource in service to their own or another discipline. The following section, on students’ preferences, reinforces this conclusion – even when inevitable compromises and constraints arise.

2.3 Comparison across student choice

Before arriving, students have an opportunity to express a preference for CS1-B or CS1. Often, the schedule is able to accommodate their preference, but other choices and other constraints occasionally get in the way. As a result, not every student who takes CS1-B has chosen to do so. Among the students who took CS1-B, n = 100 expressed a desire to do so, n = 71 opted for the default path of CS1. (Not all students provide information as to their preferences.)

Among downstream courses with at least ten students from each of the prior-interest and no-prior-interest cohorts, Table 3 compares the downstream grades earned. The two-tailed p-value shows that the differences do not constitute statistical significance (at a 95% confidence level).

<table>
<thead>
<tr>
<th>CS1 Flavor</th>
<th>Percent Taking</th>
<th>Avg. Grade†</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS2</td>
<td>73.3</td>
<td>3.4 ± 1</td>
</tr>
<tr>
<td>CS3</td>
<td>45.6</td>
<td>3.1 ± 1</td>
</tr>
<tr>
<td>CS1-B</td>
<td>73.1</td>
<td>3.4 ± 1</td>
</tr>
<tr>
<td>CS2</td>
<td>53.4</td>
<td>3.2 ± 1</td>
</tr>
</tbody>
</table>

† Grades given on a four-point scale, e.g., 2.0: C, 3.0: B, 4.0: A

We do not see the lack of statistical significance here as a failure. Instead, we see a reassuring reminder – that who students think they will be when arriving is not necessarily who they will be, when their undergraduate stories are being lived. We find it both humbling and heartening that CS1-B prepared students equally well for their future courses, regardless of self-perceived prior inclinations. The CS content of the course was communicated effectively, even for those without prior interest in a biological “flavor” of CS1.

3 Looking Back

3.1 Student Reactions

Because CS1-B has been fully institutionalized, we have digested – and responded to – many years of student feedback. Overall,
students have appreciated the integration of biological insight-development and problem-solving via computing. As one student put it, “it was interesting to see the applications of what we were learning, especially when given real data.” In fact, we were surprised by how CS1-B’s focused context produced many comments about computing’s general applicability – not only in biology, e.g., “the course taught me that CS is an extremely useful tool in research, regardless of what field that research is in.”

Reactions like these reflect that computing can be contextualized - without being compartmentalized. Many found the disciplinary bridge-building reassuring, bolstering their computing confidence through a specialty in which they already feel comfortable, e.g., “[this course] made someone who was scared about CS excited to take another CS course!” and “I liked how applications were emphasized. This helped me understand and stay interested in the material.” This formal student feedback – and we have sensed this through informal channels, as well – suggests a growing awareness of the breadth of computing’s reach and a growing acknowledgment of computing’s deep role in many of our era’s academic and professional paths.

3.2 Demographic comparisons

Academic identity and demographic identity intersect in intricate ways: if “Computing for All” is to make good on its quantifier, it must do so across all identity differences. In particular, it is one thing to create distinct curricular paths for developing computing skills; it is a very different thing to ensure that all of those paths serve all of their students inclusively and equitably.

3.2.1 Gender across CS1 and CS1-B. In many computational-identity fields women are underrepresented. Women are not underrepresented in many fields with bioscience-identities. Across CS1-B these broad trends have interwoven into the percentages shown in Figure 2.

3.2.2 Race and ethnicity across CS1 and CS1-B. Relative to the gender-identity data, the race and ethnicity trends across CS1 and CS1-B more closely follow the overall student cohort, as Figure 3 depicts.

3.2.3 First-generation status across CS1 and CS1-B. Optionally, too, students may identify as a first-generation college student, i.e., neither parent has a college degree; partial, with one college-graduate parent; or not-first-gen, with two college-graduate parents. Figure 4 shows the percentages of the full student-cohort, the CS1 cohort, and the CS1-B cohort across these backgrounds.

3.2.4 STEM-Major identity across CS1 and CS1-B. Figure 5 shows the percentages of students who pursue each of the six largest STEM majors, according to their introductory CS experience. The lefthand...
CS1-B seems to interest first-generation college students somewhat less than CS1. This is a possible trend that deserves additional follow-up.

gray bars show the choices of the whole student body, e.g., 7% are Biology majors. The second (amber) column in each group is the percentage of CS1-taking students who major in that discipline, e.g., 4% of CS1 students major in Biology. The third (green) column in each group is the percentage of CS1-B-taking students who major in that discipline: over 30% of CS1-B students major in Biology. The totals add to (well) over 100% because of the opportunities for double- and combined-majoring.

The left bar in each group is the percentage among all students majoring in one of the six largest STEM-disciplines. The central bar is the percentage among CS1-takers; the right bar (in each group) is the percentage of the CS1-takers who choose each major. Because of combined and double majors, totals add to over 100%.

Here, the two-way impact of CS1-B is perhaps most striking. The number of Biology (and Chemistry) majors whose foundational computing skillsets come from CS1-B is quite large (over 30% each; here joint-majors play a role, because biochemistry and compbio are popular). Program details aside, CS1-B is not detracting from, but contributing to the shared-cohort experience of those academic identities. Put another way, this experiment has shown a path by which early-incorporation of computing can serve not as a threat, but as an advantage to both biosciences majors and CS majors in both skillset-building and academic-identity building. For both groups, the door remains open for pursuing more computing, as desired.

3.3 Verdict

These results suggest that, accompanying the outward expansion of computing’s audience may be an expansion of its practitioners. Scientists of all sorts – and, increasingly, analysts, professionals, those pursuing humanities, the arts, and business – are owners of computing as much as traditional disciplinary practitioners, e.g., computer scientists and software engineers. The compatibility of downstream student experiences after CS1 and CS1-B suggests that, increasingly, the undergraduate literacy of “Comp 1” might refer to Composition 1 or Computing 1 (or both!)

To date, the data show that students taking CS1-B retain interest and identity in Biology, but are not at a disadvantage in – nor discouraged from – later CS courses. That it is possible to present the fundamental computing skills we ask of introductory CS through the lens of another academic discipline begs the question, “With which other disciplines might we collaborate?” Discipline-specific offerings, hosted in a variety of “home” departments may successfully blend computing into the practices and paths of their students. Indeed, this process is well underway.

3.4 Looking Outward

A distinctive contribution that we in the CS community can make is to foster a shared ownership of computing. When it comes to undergraduate curricula, crossover between CS and other disciplines grows naturally within Biology and many other STEM fields.

We are equally energized by institutions where arts and humanities disciplines are building bridges with and to computing. Bates College’s Digital and Computational Studies (DCS) program, for example, does not aim to train software engineers, nor scientists (though it precludes neither). The program embraces computing as a more general and powerful means of inquiry, with the goal of “transform[ing] a selection of courses featuring computation into a curriculum that integrates computing throughout the liberal arts.” [1] As in DCS, RISD’s Computation, Technology, and Culture concentration asks its artists and designers to own and leverage computing in their own professional practice: “students gain an understanding of the ideas and techniques of writing in programming languages, while engaging with critical analysis, history, and theory concerning software systems, computational platforms, and associated technologies shaping society.” [2]

Any two proxies necessarily miss entire dimensions in computing’s landscape. Yet these, like CS1-B, highlight the premise that computing, like other professional literacies, is too large to be captured by a single path, discipline, or curriculum.
Looking Forward

Our experiences with CS1-B have been that students – even those who may have found themselves in the class without having chosen it – are as eager and able to succeed in subsequent computing paths as their contemporaries who take more traditional CS1. As computing’s audience expands, we feel this is an important result. As summarized above, our first decade of CS1-B has yielded objectively positive outcomes for participants, without detracting from other students’ paths.

More subjectively, both we as an institution and we as individuals feel our CS1-B experiment has been enormously worthwhile. The benefits have not only been for CS1-B’s students, who have excelled and enjoyed doing so, but also for the faculty and departments involved. The shared curricular ownership between Biology and CS, we believe, has been essential to this experiment’s success thus far. That shared ownership diffuses tensions among students who want to leverage computing’s mindset and skillsets – but who want to do so without adopting a CS identity. Even as we look forward to an era in which the phrase “CS identity” connotes a wholly inclusive worldview, today it is a phrase that is lived, via too many pre-college experiences, as exclusionary. Neither we nor other post-secondary institutions benefit from waiting for others to change that landscape.

We hope results such as CS1-B’s will free institutions from an “only one way” or a “CS owns computing” mindset. CS doesn’t own computing, and CS1-B suggests that it need not. We foresee advantages for institutions looking to make “Comp 1” a part of the professional path of all of their students.

Those paths may prompt adjustments to academic identity in two ways. For non-CS disciplines, they may mean embracing the era’s computing resources and skillsets, to the extent they bear fruit. For CSers, even as we relish our role as stewards of computation, “CS for All,” or at least “Computing for All,” may mean less ownership and more fellowship with respect to computing.

The call is to the many opportunities for collaboration – and to those partnerships’ widely-branching curricular paths – paths to which all disciplines contribute.

It’s a great era for computing. We look forward to the journey!

Acknowledgments

We acknowledge and thank the NSF for making this work possible, specifically through the support of projects #1612451 and #1659805.

References


---

[4] this can also be true in college and post-college, unfortunately