Lecture, Discussion, Group Work, Repeat: Using Aerial Photography and Machine Learning to Study the Use of Writing-Related Pedagogies in STEM Courses and Their Impact on Different Student Subgroups[1]

Julia Voss, Santa Clara University
Navid Shaghaghi, Santa Clara University
Andres Mauricio Calle, Santa Clara University
Kristin Lee, Santa Clara University
Liam Abbate, Teaching Assistant Program in France

Abstract: Although Writing Across the Curriculum (WAC) has long focused on incorporating writing and related literacy activities into STEM education, the extent to which these pedagogies are widely used in STEM teaching remains unclear, as does their impact on student course performance, especially for underrepresented and marginalized student groups. Using a sample of 18 STEM courses at a private liberal arts university, this study uses unique empirical methods to reconsider, for STEM disciplines, Russell's (1990) claim that WAC has failed to make a "permanent impact" on higher education by a) using photography to document classroom activities in real time and b) using machine learning to categorize these images to determine which learning activities are used in STEM instruction and in what proportions. We find that (a) lecture continues to dominate in STEM education and that (b) some active learning pedagogies (discussion and group work) have ambivalent relationships to course performance (which differ according to student subgroups defined by gender, race, national origin, and other factors) while WAC pedagogies like reading and writing, although rare, are associated with improved student course performance. In light of these findings, we suggest implications for STEM pedagogy, best practices, and future research to prioritize equitably designed pedagogy in STEM.

Introduction

Writing Across the Curriculum (WAC) is an inclusion-based movement, focused on making disciplines accessible to students—especially as the college student population has become increasingly diverse beginning in the second half of the twentieth century—by using writing and other associated literacy practices. However, as Russell (1991) argued, WAC's inclusive mission has often inspired resistance in the disciplines because of its explicit goal of combating gatekeeping and opening disciplines to new groups by making the discourses of those disciplines explicit. These skeptics see a false dichotomy between equity and excellence when it comes to recruiting new members into disciplines, and argue that pedagogical changes explicitly aimed at inviting new members threaten the prestige and potential of the field. Russell (1991)

Across the Disciplines
A Journal of Language, Learning and Academic Writing
https://doi.org/10.37514/ATD-J.2022.19.1-2.07

wac.colostate.edu/atd
ISSN 554-8244

offered post-WWII American engineering as an example, describing how, when faced with a critical labor shortage in emerging technical fields, engineering faculty remained resistant to WAC because they perceived that inclusion-focused pedagogy would reduce the rigor and excellence by which they defined their disciplines, even though this adherence to gatekeeping resulted in serious worker shortages that hampered the disciplines' ability to contribute to the United States' Cold War efforts.

As scholars like Genevieve García de Müeller (see Syracuse University "Antiracist WAC Toolkit," 2021) work to address WAC's historical exclusion and complicity in oppression, there is more research to be done concerning the best pedagogical practices for student learning. This paper introduces a research method that uses a neural network to classify classroom photos depicting different pedagogical activities (such as discussion, lecture, writing, group work, etc.) in order to describe the proportion of class time devoted to each activity. This classroom data was combined with institutional demographic and grade data to allow the TailorEd Team of Santa Clara University's (SCU) Ethical, Pragmatic, and Intelligent Computing (EPIC) research laboratory to examine the effects of pedagogical practices on individual student subgroups so that educators can design their classes to support student learning, particularly for students who have been marginalized and discriminated against in STEM disciplines.

Literature Review

Overview of Writing to Learn STEM Pedagogies

Although this project focuses on all pedagogies practiced in STEM classrooms (broadly categorized as lecture, discussion, group work, reading, and writing), it is particularly interested in writing and writing-related pedagogies like discussion, group work, and reading as alternatives to the lecture formats that have long dominated STEM pedagogical instruction. In order to situate this focus within the longstanding literature on WAC in STEM, we begin with an overview of the research on Writing to Learn (WtL) in STEM classrooms, which uses writing to support the learning/mastery of disciplinary concepts and conceptual understanding across the curriculum (Herrington, 1981; Kovac & Sherwood, 1999). Across the United States, three kinds of pedagogies seem to dominate uses of WtL: (a) taking notes on readings for a class or other informal writing; (b) using innovative assignment types; and (c) project-based learning, documented in the classroom photos used in this study. In this approach, writing has been the means, rather than the end, of learning in a course and can be an effective learning tool because rhetorically effective writing on a topic requires deep comprehension of that topic (Herrington, 1981). There has been a strong positive association between instructors' perceptions of quality of scientific content and students' success in a course (Kovac & Sherwood, 1999). Furthermore, WtL seems to be more effective when it has been fundamentally integrated into course content and students understand how and why writing is being used as a learning activity (Kovac & Sherwood, 1999).

While there has been an overall consensus that writing is fundamental to disseminating knowledge in STEM fields, whether and how it is taught in undergraduate classrooms depends on faculty's conceptions of writing as a learning tool (Moon et al., 2018). Much of the existing WtL research focuses on faculty attitudes toward writing as an indirect indicator of impact on student learning (Moon et al., 2018). A long-established goal of WAC has been to shift faculty's belief that teaching is primarily a matter of transmitting content knowledge (Knoblauch & Brannon, 1983), which typically has encouraged faculty to rely on lecture as a means to achieve maximum content coverage (Scheurer, 2015). As professionals, STEM faculty write a lot (lab reports, conference papers and journal publications, grant proposals, etc.), but many have not felt comfortable teaching or grading writing (Barr, 2012; Moon et al., 2018; Kovac & Sherwood, 1999). However, Barr (2012) suggested that, under certain circumstances, students may appreciate WAC courses because they otherwise have very limited time devoted to developing writing skills in their academic careers.

Another obstacle is that institutional priorities may not have encouraged faculty to invest in teaching writing (Barr, 2012). For WtL to be effective, writing assignments should be linked to a course's objectives (Herrington, 1981), which often entails change at the departmental or programmatic level. The grading labor entailed in WtL has often been a barrier in bringing it to the classroom (Barr, 2012), which can be a major disincentive depending on criteria for faculty hiring, evaluation, and promotion. For example, Kramer et al.'s (2019) study of increased writing integration in kinesiology courses focused on a large-scale intervention where faculty added or revised a major assignment and collaborated with writing studio instructors to deliver one-on-one writing instruction, feedback, and practice. This was a massive undertaking contingent on Kramer's involvement as writing center director and an existing program of "studio" companion courses which supported writing-intensive additions to kinesiology "content" courses. Another obstacle often raised regarding such writing-intensive teaching is lack of transparency to students about how writing is evaluated (Moon et al., 2018; Kovac & Sherwood, 1999). In a lower-effort intervention-based study using informal writing where students kept a chemistry reading journal of notes, observations, and answers to reading questions, there was some student resistance to the intervention because some perceived it was busy work not essential to the main content of the class (Kovac & Sherwood, 1999).

However, when WtL interventions are well-integrated into the curriculum, they can have a major impact. Winfield et al. (2019) designed informal writing via worksheets to reinforce concepts of organic chemistry and found that students who completed these WtL exercises performed better on exams than those who did not. Furthermore, speaking back to Kovac and Sherwood's (1999) findings on student buy-in, students reported to Winfield et al. (2019) that they found WtL activities "valuable or extremely valuable" and recommended that WtL activities be used more frequently to teach disciplinary content, suggesting the potential of fully integrated WtL STEM pedagogies. Research by Bunker and Schneider (2015) in human physiology courses and by Achen and Lumpkin (2015) in sports management courses on the use of more creative WtL assignments in STEM classes such as narratives, blogs, minute papers, group work, and independent projects found that giving students more opportunities and more freedom in how to approach assigned writing reduced student anxiety about difficult course content, helped students see the significance of what they were learning, and encouraged them to think originally/creatively about the material, which resulted in students seeing these activities as the most valuable parts of their courses.

Impact of WtL STEM Pedagogies for Specific Student Subgroups: Connecting the WAC and Active Learning Literatures

In STEM contexts, considering differential impacts on student groups is particularly necessary due to the "pipeline" problems identified by Miller et al. (2020). National data shows the extent to which Black-, Latino/a-, and female-identifying students are underrepresented in STEM majors and in the professions these programs feed into. These findings have motivated the NSF, IEEE, and other organizations to call for and fund efforts aimed at improving recruitment and retention of underrepresented students in STEM disciplines (Miller et al., 2020). As Russell (1991) has argued, WAC pedagogies' emphasis on introducing and inducting students into the professional discourses that constitute and gatekeep disciplines positions WAC as a valuable tool for promoting equity and inclusion, especially in STEM disciplines struggling to recruit and retain diverse student populations. However, as illustrated by the previous section, much of the WAC research on WtL pedagogies reports on impacts on *students in general*, rather than considering how writing assignments, writing activities, and interactive literacy activities affect different subgroups of students. That is, STEM WAC studies reporting on the impact of writing-intensive pedagogical interventions are often small projects focused on a single course or teacher, and therefore subdividing participants according to race or gender is difficult. As Montenegro and Jankowski (2020) argued with regard to educational research, this is especially problematic for students from underrepresented groups, who may appear in studies in extremely small numbers, making it difficult to make generalizable claims about marginalized students. WAC studies also often do not report on the racial, gender, and other identities of participating students, which Anson (2012) noted, suggests that race and gender have rarely been seen as salient factors that affect students' experience of or response to WAC pedagogies. This contrasts, for example, with Winfield et al.'s (2019) explicit situation of their research at Spelman College, focusing on the impact of WtL chemistry pedagogies on Spelman's all-female, predominantly Black student population.

The WAC research that does attend to race shows how important these factors are to consider. For example, Kells's (2007) account of revamping the WAC program at the University of New Mexico (a Minority Institution, Minority Serving Institution, and Hispanic Serving Institution) showed how focusing on the racial and ethnic backgrounds of UNM's student population helped the program resist the tendency of WAC programs to reify White supremacy by trying to uncritically assimilate students—especially historically excluded ones—into disciplinary discourses, rather than examining how those discourses may need to change to accommodate new members.

To build on the empirical WAC work like Poe (2013) and Poe and Craig (2011) that focuses on diversity and inclusion, it's useful to refer to related work about the scholarship of teaching and learning (SoTL) that focuses on recruiting, retaining, and supporting diverse students in STEM fields, especially through the use of active learning pedagogical techniques. Active learning research in STEM often focuses on measuring the impact of new techniques like group-based problem solving, flipped teaching, and other methods that often use writing to resist STEM's traditional reliance on lecture-based pedagogies (Michael, 2006). These initiatives are often run as experiments on entire programs, including many sections and thereby encompassing enough students to offer disaggregated analysis of outcomes. Furthermore, because these studies often double as program assessments (rather than teacher-research), they typically include standardized measures of student learning via course performance (grades) and experience outcomes (course evaluations), allowing for further (although not unproblematic) generalization. When Latulipe et al. (2018) examined the use of flipped classroom techniques that assigned readings and videos to cover content as homework and using class time to collaborate on application activities, they found that female students in flipped computer science classes had higher retention rates than did female students in classes using traditional lecture pedagogy. Similarly, in a study of the impact of writing-intensive flipped classroom approaches in entry-level chemistry courses at Spelman College, a women's HBCU, Winfield et al. (2019) showed that students scored higher on the American Chemical Society standard exam, got better grades in advanced chemistry classes, were more likely to persist in their STEM major, had a greater sense of agency, and reported more peer engagement than students enrolled in the same courses using traditional lecture pedagogies.

Addressing Limitations of Existing Research on Writing-Intensive STEM Pedagogies

There is a broad range of research performed on writing-intensive STEM pedagogies, but since this research is conducted across numerous disciplines, there is little consistency or structure to define active learning activities and measure their outcomes (Hartikainen et al., 2019). To address these consistency issues, we define activities in terms of their physical presentation as captured in classroom photography, and use grades as an outcome measure, allowing for comparability across courses, programs, and institutions. Another significant limitation of much of the research on classroom pedagogy, Gierdowski (2013) has noted, is the need to compromise between detailed, real-time, comprehensive data on classroom activities, on the one hand, and scalability, on the other. Typically, researchers seeking to capture detailed, real-time data on classroom teaching have relied on observational methods such as teacher research or ethnography that are extremely time-consuming (see Achen & Lumpkin, 2015), which gather valuable data but limit the scale of their research. On the other hand, many large-scale projects have used surveys to ask students and faculty to retrospectively self-report on classroom pedagogy (see Lee et al., 2014; Hibbard et al., 2015), or used course descriptions/schedules (see Latulipe, 2018) to document the pedagogical activities used in STEM courses. These methods are more scalable, but, as Gierdowski (2013) cautioned, provide less comprehensive and reliable data on what actually happened in class on a daily basis. Our photographic research method and neural network categorization address shortcomings of both detail, recall, and scale: the photos capture classroom activities in real time, and the neural network allows for scalability.

We rely on grades to measure student learning via course performance for similar reasons of scalability, which raises some issues, especially given our interest in just and equitable STEM education, since grades have been demonstrated to reflect and perpetuate educational inequality rather than objectively reflect student learning (see Warikoo et al, 2016), in addition to research demonstrating the risk of mismatch between grades and learning for all students (Blum, 2020). However, because grades are universally mandated and archived by this and other institutions, grades allow us to gather data on the performance of every student in the study without placing additional demands on research participants. Secondly, although grades are flawed measures of learning, they carry considerable weight within academia and the professions for recent college graduates, determining things like class rank; eligibility for scholarships, and awards; access to graduate school and other professional attainments; et cetera. That is, although ample research has questioned the validity of grades as measures of learning—especially for diverse student populations—grades continue to serve as significant gatekeepers on student opportunity, making them an appropriate measure of the degree to which a course contributes to narrowing, widening, or maintaining educational inequities.

Methods

Our methods involve 3 phases:

1. A human coding phase, where trained coders labeled the activity depicted in each image
2. A machine learning phase where team members created a neural network trained on these human-labeled images to automate the analysis of aerial classroom photos according to activity
3. A quantitative analysis phase where we examined correlations between pedagogical activities and student course performance, disaggregated according to student and course characteristics

Phase 1: Qualitative Coding of Classroom Photos

This study analyzes aerial classroom photos (see Figure 1), taken at 1-minute intervals, from STEM courses taught in 2014-2016, as part of a larger IRB-approved research project on the relationship between classroom design and student course performance.[2]

Figure 1: Sample aerial classroom photos

Based on a previous study of this photo data, the coders began with a basic list of activities (see Shaghaghi et al., 2019). Over three months, every week each coder categorized a shared set of 100 images, randomly selected from the photo corpus and met to discuss categories and category definitions, focusing on images that had been labeled differently by different coders. At the conclusion of this grounded-theory-based norming process, the coders achieved an overall agreement rate of 80% or above around the following classroom activity categories:[3]

- **Empty:** no/few students are present in classroom[4]
- **Lecture:** students are receiving information from a single presenter (including instructor lecture, student presentation, and/or media viewing)
- **Discussion:** students are engaged in common discussion activity, attending to a member of the class
- **Group work:** students are formed into small groups working on a shared task (including discussing, reading, and/or writing)
- **Writing (solo):** students are writing individually, either by hand or typing on a device
- **Reading (solo):** students are reading individually, either from print materials or digital devices

Following this norming process, for the next six months coders received 50 images to label each week, randomly selected from across the entire photo corpus. Each image was assigned to two randomly paired coders, with pairings changing from week to week. If both coders labeled the image the same way, its categorization was final. If the coders disagreed, the image was submitted to a third coder for a "tie-breaker" vote. If the three coders could not agree, the project PIs (Voss and Shaghaghi) used the category definitions to assign a final label to the image. These "problematic" images were tagged as ambiguous and not used to train the neural network (described below), although all photos were kept in the corpus. The 3,700 non-ambiguous images labeled by the human coders were used to train the neural network built by the project's machine learning team, which was then used to label all 18,000 photos taken in the STEM courses analyzed for this paper.

Phase 2: Neural Network Categorization of Classroom Photos

The AI phase of this project involved two stages: standardizing images to account for inconsistencies in photography and adapting a previous iteration of a neural network created in 2019 to use machine learning to identify the classroom configurations documented in these photos. To standardize images taken in the nine different classrooms, we implemented image processing techniques using MATLAB to regularize the classroom images, specifically to deal with differences in photos caused by differences in camera lenses, lighting, and focus (visible in the raw photos in Figure 1). After standardizing the photos, we used transfer learning—a machine learning method where a model developed for one task is used as the starting point to develop a model for a second task—to develop a new model for analyzing classroom pedagogical activities. We developed the new model for this task using the existing Classroom Configuration Identifier (CCID, described in Shaghaghi et al., 2019), which itself was constructed using transfer learning on AlexNet (Krizhevsky et al., 2012), a convolutional neural network used to analyze visual imagery by recognizing visual patterns in images. Although CCID was capable of identifying classroom configurations in photos as depicting forward-facing, circular/u-shaped, small group, or empty layouts with 97% accuracy, it did not classify images according to pedagogical activities.

In order to shift the focus of the neural network from classroom configuration (layout) to classroom pedagogy (activity), a new Classroom Activity Identifier (CAID) was developed for this paper which categorizes the images into lecture, discussion, group work, reading, and writing activity categories. Unlike CCID, which used transfer learning on AlexNet, CAID uses the more modern ResNet CNN to obtain more accurate results (He et al., 2016). As with CCID, we applied hyperparameter tuning to find the optimal learning rate, batch size, dropout rates, and number of epochs to maximize the model's accuracy.

CAID accurately predicts the pedagogical activity occurring in the class 86.17% of the time. It should be noted, however, that this accuracy result doesn't tell the whole story, as illustrated in Table 1. Because the model was trained on a representative sample of pictures from classrooms, the number of images depicting each activity varied considerably because some activities were used much more frequently than others. The number of images depicting each activity affects the accuracy rate of each category, with activities with fewer examples having lower accuracy rates because CAID can't generalize the category parameters as effectively. For example, because by far the greatest number of images in the dataset show lecture activities, this category has a very high accuracy rate. For activities that are rarely depicted in photos (like reading), the accuracy is lower.

Table 1: Activity Confusion Matrix for CAID. The diagonal of blue squares counts the test images correctly labeled by CAID, where "true class" matches "predicted class." The pink squares tally the images mislabeled by CAID, where predicted class (the label assigned by CAID) does not match true class (the label assigned by human coders). A high ratio of the blue square count to the pink square count indicates a high degree of accuracy for the individual category

		Predicted Class					
		Discussion	Empty	Group Work	Lecture	Reading	Writing
True Class	Discussion	7	1			11	
	Empty		83				
	Group Work	1		12	5		
	Lecture	6		4	145		2
	Reading	1			6	1	4
	Writing	1			7		35

Phase 3: Quantitative Analysis

Using CAID, we analyzed all 18,000 STEM classroom photos to determine the percentage of class time devoted to different activities by determining the percentage of each class session devoted to discussion, group work, lecture, reading, and writing. We then merged these per-course percentage breakdowns with a database of student information from these courses containing students' demographic information (e.g., sex, race, major, first-generation status, et cetera) and course grades obtained from Santa Clara University's

Office of Institutional Research. The merged dataset allowed us to draw comparisons across courses on the basis of student demographics and discipline using group-specific bivariate correlations (Pearson correlations run in SPSS) between student grades and classroom activity. Pearson correlations, as the name implies, show correlation, not causation. That is, our findings show the grade performance for student groups that are associated with rates of use of different classroom activities. They do not tell us that these classroom activities *cause* these student course performance outcomes, because we did not control for the myriad of other factors that affect student grades, such as work completed outside of class or individual student prior experiences with the subject not represented in the demographic characteristics students provide to the university.

Study Context

The data for this study was gathered from 2014-2016 at Santa Clara University, a mid-sized, nationally-ranked private liberal arts university in California's Bay Area with a strong STEM focus. The 18 STEM courses analyzed for this article were part of a larger study of classroom pedagogy in seven "active learning" rooms with mobile furniture, writable walls, and multiple projectors, and two "control" classrooms with stationary desks and a single board/projector at the front of the classroom. The study included the following courses, listed in Table 2:

Table 2: STEM courses included in study (n=18)

Discipline	Course Title	Level	Number of sections
Biology/Environmental Science	Biostatistics	Upper division undergraduate	1
Biology	Bioinformatics	Upper division undergraduate	1
Chemistry	Bioanalytical Chemistry	Upper division undergraduate	1
Civil Engineering	Geology	Lower division undergraduate	1
Civil Engineering	Mechanics III: Strength of Materials	Lower division undergraduate	3
Civil Engineering	Civil Engineering Materials	Upper division undergraduate	2
Civil Engineering	Geotechnical Engineering	Upper division undergraduate	1
Civil Engineering	Advanced Concrete Design	Upper division undergraduate/ graduate	1
Civil Engineering	Construction Operations and Equipment	Upper division undergraduate/ graduate	1
Computer Engineering	Applied Programming in MATLAB	Lower division undergraduate	1
Computer Engineering	Computer Forensics	Upper division undergraduate/ graduate	1
Computer Engineering	Artificial Intelligence	Upper division undergraduate/ graduate	1
Mechanical Engineering	Intro to Math Methods in Mechanical Engineering	Upper division undergraduate	1
Mechanical Engineering	Thermodynamics I	Upper division undergraduate	1

This study's use of aerial photography focuses on the use of writing and associated classroom literacy activities like reading, group work, discussion, and lecture. However, as much of the WAC literature attests, there are many other ways writing and literacy can be integrated into curricula, many of which fall outside the scope of this study. To partially address this limitation, we analyzed course syllabi to determine the role of writing and writing-related activities based on the course description, schedule, and grade breakdown to place the in-class activities on which this study focuses within the context of the courses' full engagement with writing, reported in Figure 2.

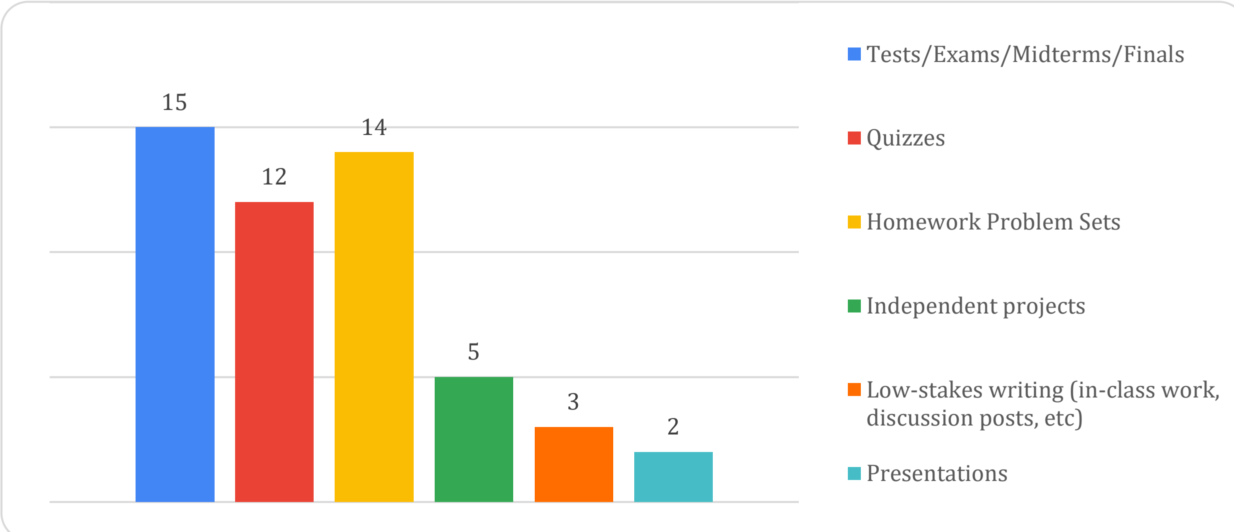

Figure 2: Writing and writing-related activities included in courses (n=18).

Our syllabus analysis suggests that the majority of in-class student writing activities visible in the classroom photos are likely tests or quizzes. Eighty-three percent of the courses assign homework problem sets, a writing activity not typically performed during class meetings. Twenty-eight percent of the courses assign independent projects that require larger-scope problem solving and writing up solutions, and about half of those courses also require one or more in-class presentations. Presentations were visible in classroom photos, but unless they involved discussion, group work, reading, or writing activities, they were classified as "lecture" since they involve passive learning for the majority of the class as audience members. For independent projects, class time spent working on the projects was categorized as group work (for group projects) or solo writing/reading (for individual projects), but the majority of independent project work likely occurred outside of class and is thus outside the scope of this study.

One other important contextual factor is the nature of WAC at Santa Clara, which does not have a dedicated WAC program, but rather uses an "advanced writing" general education requirement to complement the university's required first year writing course. Although all engineering undergraduates take an engineering-specific version of advanced writing (ENGL 181: Engineering Communications, no sections of which are included in this study), these courses are taught by English Department faculty and are not substantially integrated into the rest of the engineering curriculum. SCU's natural science and math departments do not offer their own advanced writing courses, so students in these disciplines take advanced writing courses offered by other departments. There are also university-level faculty professional development efforts that support WAC-type curriculum design in the form of sponsored participation in the Association of College and University Educators' [Effective Practice Framework] training program and the locally-designed and facilitated Success in Writing, Information, and Research Literacy Initiative (see Serviss & Voss, 2019). However, participants self-select into these programs and STEM faculty have not been a specific focus. Therefore, unlike institutions with established WAC programs and histories of WAC curriculum and faculty development, Santa Clara's STEM classes reflect an institutional context where WAC is subsumed within the institution's overall teaching mission, but has not (yet) become an institutional priority.

Results

We report the results of this study in two ways: first by summarizing the overall distribution of in-class activities for each course and then by reporting correlations between the use of different classroom activities and students' grades, disaggregated by student subgroup.

Course Activity Profiles

Figure 2 visualizes the distribution of classroom activities in each course included in this study. As the literature on typical pedagogies in STEM education suggests, the majority of class time in most courses was devoted to lecture, with the exception of the upper division undergraduate Bioanalytical Chemistry course that relies heavily on discussion, and the upper division undergraduate/graduate course Construction Operations and Equipment that incorporates a significant amount of writing. Group work occurs regularly in one class (Biostatistics [upper division undergraduate]), and occasionally in two others (Mechanics III: Strength of Materials [lower division undergraduate], Civil Engineering Materials [upper division undergraduate], and only in some sections of these courses). Most courses devote 5-15% of their class time to writing.

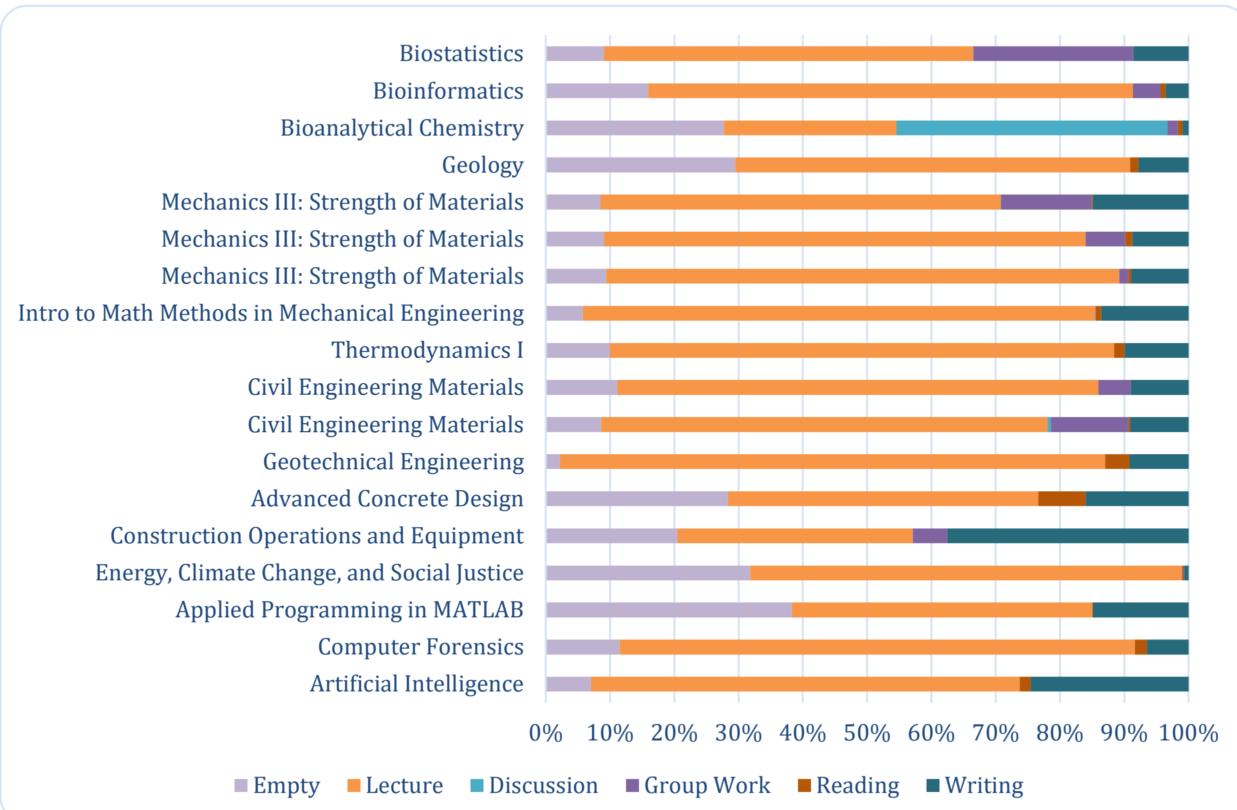

Figure 3: STEM course activity profiles, showing the percentage of total class time devoted to lecture, discussion, group work, reading, and writing activities, as well as the percentage of time during class meeting times the classroom was empty. Courses are grouped by discipline: first natural science courses, then mechanical engineering, civil engineering, and computer science/engineering courses.

A surprisingly large number of the photos taken during scheduled class meeting times showed few/no people and were categorized as “empty.” By reviewing the syllabi and corresponding with instructors, we learned that classrooms were empty during class meeting times for a variety of reasons including field trips, lab days, class cancellations due to school holidays or closures, and class cancellations due to instructor absence.

Comparing Classroom Activities and Student Performance

To establish relationships between student course performance learning (as measured by grades) and classroom activities, we compared students' course grades and their grades relative to the class average to the proportion of class time devoted to each pedagogical activity. The results of this analysis are reported in Table 3 and Table 4, and discussed in the following subsections focused on each type of classroom activity.

Table 3: Correlations between classroom activities and student grades. Results are reported for correlations whose two-tailed significance test meets the <.05 threshold.[5]

		Total n=376	Discussion	Group work	Lecture	Reading	Writing
Sex	**Male**	266	.168	-.158	-.203	.219	
	Female	110				.23	
Race/Ethnicity6	**White**	186	.191		-.265		
	Hispanic of any race	62		-.385		.378	
International/ Domestic	**International**	35	-.361	-.348			.438
	Domestic	341	.159		-.200	.203	
Pell Grant	**Ineligible**	252	.186		-.252	.191	
Parent educational background	**1st generation**	46				0.266	
	Continuing generation7	330	.136		-.207	.194	
Course meta discipline	**Engineering**	313	-.187	-.205	-.144	.264	
	Natural science	63	.286	-.392		.381	-.394
Major meta discipline	**Engineering**	304	-.196	-.204	-.144	.258	
	Natural science	58	.295	-.452		.402	

Table 4: Classroom activities correlated with students' grades relative to peers, calculated by subtracting average course grade from student's grade. Results are reported for correlations whose two-tailed significance test meets the <.05 threshold.

		Total n=376	Discussion	Group work	Lecture	Reading	Writing
International/ Domestic	**International**	35					.402
Pell Grant	**Ineligible**	252	-.271				

Several BIPOC (Black, Indigenous, and People of Color) and low-income student subgroups are missing from Table 3, Table 4, or both, reflecting (a) the failures of institutions like Santa Clara to recruit and retain Black, Indigenous, Multiracial, Pacific Islander, and Pell Grant-eligible students both generally and particularly in STEM disciplines and (b) our study's failure to create a large enough sample to generate statistically significant results for these groups. As we discuss in the Discussion section, this outcome has methodological implications for studying marginalized students at predominantly white institutions like Santa Clara.

While the small number of Black, Indigenous, and Pacific Islander students included in our sample did not allow us to identify subgroup-specific trends for these students, we were surprised not to identify any group-specific trends for Asian students, who account for 20% of our sample, suggesting that there were significant differences among Asian students which created significant variation between them. Because of the comparatively large number of Asian student participants, we were able to subdivide Asian students into different subgroups based on other demographic factors, identifying distinct trends for Asian male/female students, domestic/international students, and engineering/natural science students, reported in Table 5.

Table 5: Classroom activities correlated with course grades (upper section) and difference from average course grade (lower section) for subgroups of Asian students. Results are reported for correlations whose two-tailed significance test meets the <.05 threshold.

		Total n=76	Discussion	Group work	Reading	Writing
Correlation with Course Grade	Male	51		-.339	.301	
	International	20	-.500	-.590		
	Domestic	56				
	Natural science course	12		.753	-.724	.760
	Engineering course	64		-.265		
	Engineering major	64			.263	
Correlation with Difference from Average Course Grade	Female	25	.568	.482		
	International	20		-.507		
	Natural science course	12		.753	-.724	.760
	Natural science course	13		.628		

Lecture Activities

Where lecture has a statistically significant relationship to students' grades—for male students, White students, engineering majors and students enrolled in engineering classes, domestic students, students not eligible for Pell Grants, and continuing generation students—that relationship is negative, and small in effect (*r* values ranging from -0.265 to -0.144). That is, in courses using more lecture, students in these subgroups had, on average, slightly lower course grades compared to similar students in classes that used less lecture. It is important to note that the student groups showing this relationship between lecture and grades reflect many of the largest and most privileged subgroups in each demographic category, raising questions about whether similar trends in the relationship between students' grades and the use of lecture activities would be found if the total number of participants in the smaller comparison groups (female students, BIPOC students, Pell Grant eligible students, first generation college students) were larger (see Discussion section below). When we compared students' grades to the average course grade to compare students' performance within an individual class, we found an additional negative correlation between lecture and Asian international students (r = -.252), meaning that they had disproportionately lower grades than their classmates in classes that used more lecture-based activities.

Discussion Activities

The relationship between the use of discussion activities and grades is stronger, but more variable across student groups. For some student groups, the use of discussion activities is negatively related to student grades. For students enrolled in engineering classes or majoring in engineering, the inverse relationship was modest (r = -0.187 and r = -0.196, respectively). However, the inverse relationship between discussion activities and course performance was stronger for international students (r = -0.361), which seems driven especially by Asian international students (r = -0.500), for whom discussion activities had a strong inverse correlation with grades.

However, for other students, increased use of discussion was associated with higher student grades. Grade increases were modest for some student subgroups—continuing generation students (r = 0.136), domestic students (r = 0.159), male students (r = 0.168), students not eligible for Pell Grants (r = 0.186), and White students (r = 0.191)—many of the same privileged subgroups whose grades were negatively associated with lecture activities. The strongest positive associations between discussion activities and student grades were seen for students taking natural science courses (r = 0.286) and majoring in natural sciences (r = 0.295). Compared to their peers, higher course grades for Asian female students were most strongly associated with increased use of discussion (r = 0.568).

Group Work Activities

For most student subgroups, there was an inverse relationship between group work and grades, showing that, on the whole, increased use of group work was associated with lower grades. This inverse relationship ranged from small (male students [r = -0.158], students in engineering courses [r = -0.205], and engineering majors [r = -0.204]) to medium (Asian male students [r = -0.339], international students [r = -0.348], Hispanic students [r = 0-.385], students taking natural science courses [r = -0.392], and natural science majors [-0.452]) to large (Asian international students [r = -0.590]). The grades of Asian international students (r = -0.507) also showed a larger negative relationship to group work when compared to the grades of their peers.

Group work correlated with higher grades and better performance relative to peers for some student subgroups, however. There was a very strong positive association between the grades of Asian students taking natural science courses and the use of group work (r = 0.753), which was also reflected in Asian natural science students' superior performance relative to their classmates (r = 0.753) and in the higher grades of Asian natural science majors compared to their classmates (r = 0.628) in courses using more group work. Asian female students also performed significantly better, on average, than their peers in courses that used more group work (r = 0.482).

Reading Activities

While group work was inversely correlated with grades, reading—although rarely used (see Figure 3)—was positively correlated with grades: most student subgroups' grades improved when more reading was used during class. The effect was small for most groups: students not eligible for Pell Grants (r = 0.191), continuing generation students (r = 0.194), domestic students (r = 0.203), male students (r = 0.219), female students (r = 0.239), engineering majors (r = 0.258), students taking engineering courses (r = 0.264), and first-generation students (r = 0.266). The effect size of the correlation between grades and reading was greater for Hispanic students (r = 0.378), students enrolled in natural science courses (r = 0.381), and natural science majors (r = 0.402). However, just as Asian students enrolled in science courses were an outlier for the effects of group work, these students also showed a strong inverse relationship between the use of reading and course grades (r = -0.724), which was also reflected in their performance relative to their peers (r = -0.724).

Writing Activities

Writing activities showed significant relationships with course grades for only a few student subgroups, but the associations were comparatively strong. There was a medium correlation between the use of writing activities and the grades of international students (r = 0.438), and Asian students taking natural science courses (r = 0.760). This robust association was also present when considering these students' grades relative to their peers: there was a strong correlation between reading activities and Asian students taking natural science classes outperforming their peers (r = 0.760). The strong correlation between Asian natural science students' grades and writing activities is particularly striking because the only case of negative correlation between grades and writing activities described a medium relationship between grades and writing activities for all students enrolled in natural science courses (r = -0.394), suggesting that the experiences of Asian students in natural science courses are particularly distinct from those of their classmates.

Discussion

Limitations

While some of our findings parallel existing research about pedagogical best practices (such as the negative association between many subgroups' grades and the use of lecture), other findings challenge claims in published literature. One factor that mitigates our findings is the fact that this is a pilot study involving only 18 courses enrolling 376 students at a single institution, specifically an institution without a robust WAC program. As Montenegro and Jankowski (2020) caution, the limitations of our sample size are shown most problematically in our limited or nonexistent findings about many of the student subgroups (including female students, Black students, first generation college students, and Pell Grant eligible students) that equity-focused research in STEM education is especially concerned with. This study also focuses on in-class activities and does not consider writing or other literacy-related activities completed outside class (although our syllabi analysis suggests that few of these classes included this kind of work, suggesting their rarity and pointing to the need to integrate WAC methods into STEM courses on a large scale, especially at institutions like this one without a strong WAC tradition). Compared to studies using detailed ethnographic methods or self-reporting to account for both in- and out-of-class uses of WtL pedagogies, this study focuses exclusively on classroom pedagogy, which provides only a partial account of a course. With these limitations in mind, we discuss the implications of our findings as well as strategies for addressing these issues in the Future Research Steps section below.

Pedagogical Implications of Findings

Our findings are suggestive and offer some important questions to consider in light of the lengthy tradition of WAC scholarship and the growing attention in STEM disciplines to the impact of pedagogical techniques on student success and retention.

The simplest and most straightforward results of this study comment on the impact WAC pedagogies have had on classroom instruction in STEM. Most STEM courses in this study devote most of their time to lecture and do not make substantial use of writing or writing-related literacy activities during class (or outside of class, based on syllabus information). The majority of in-class writing activities are still test-taking, which limits our ability to comment on the impact of the integration of WAC/Writing to Learn activities into STEM classes more broadly but suggests the prevalence of pedagogical techniques in use in STEM courses during the focus period (2014-2016). This study deliberately focuses on STEM teaching "in the wild," that is, classes that are not associated with a faculty development program geared toward training faculty or promoting pedagogical interventions. This differs from much of the published research on STEM pedagogy, which is often associated with an initiative, faculty development program, or other intervention, which faculty often self-select into (see, for example, Wilson, 1994; Breslow, 2010; Foote, et al, 2016). Some of the faculty included in this study have participated in (or even led) such programs, but because faculty were recruited based on classroom assignment, the teaching methods depicted in the photos used for this study offer a broader sampling of STEM pedagogy.

Similarly, just as this study does not control for faculty training, it also leaves open the question of students' preparation for and receptivity to different types of classroom instruction. As a wealth of pedagogical research has shown, opportunities to apply concepts rather than listening to lectures about them typically results in deeper learning and better transfer. However, these pedagogical strategies also typically emphasize metacognitive scaffolding for such activities, since students often enter classes—especially STEM classes—expecting lecture and as a result are unprepared to engage with course content through student-centered application and inquiry activities (for example, using discussion, group work, and individual writing to solve problems). The impact of the disciplining students have undergone in 12+ years of formal education may also be a factor, especially in the "general population" context of our sample, composed of teachers who may or may not explicitly discuss with students the purpose of and expectations for classroom activities like discussion, group work, reading, and writing.

Implications for Specific Pedagogical Activities

Some of the most suggestive implications in our research have to do with our ambivalent findings about the impact of many of the pedagogies recommended by both active learning and WAC scholars, like group work, discussion, and in-class writing. We interpret these findings as raising questions about (a) how best to design and deliver these learning activities and (b) how to balance the learning needs of different student subgroups when those needs do not align. That is, this preliminary study adds nuance to existing findings, which is especially necessary as active learning and WAC scholars and practitioners work to implement these pedagogies more widely.

Discussion

While discussion was associated with higher grades for more privileged student groups like male students, White students, domestic students, affluent students, and continuing-generation students, discussion was negatively associated with the performance of international students, especially for Asian international students. These findings parallel Nielsen's (2014) research on Generation 1.5 ELL students at a majority-White institution, in which interviewees reported that while their teachers valued their multilingual and multicultural perspectives, their White, English-speaking classmates often did not, mocking them for their accents/vocabulary inside and outside class and engaging in more and less overt forms of shunning by self-selecting into racially homogenous peer working groups. These experiences align with STEM education research on how model minority stereotypes of Asian and Asian American students as categorically hardworking, compliant, intelligent, and motivated by academic/economic attainment result in bias and microaggressions directed at Asian and Asian American STEM students by both White faculty and students (Trytten, Lowe, & Walden, 2012; McGee, Thakore, & LaBlance, 2016). Such a classroom climate is not conducive to student-centered discussion for students who are actively discriminated against, harassed, and/or excluded by their majority-group peers (and even their teachers), pointing to the importance of curriculum design and scaffolding of discussion (and other WAC-related learning activities) that seek to improve student course performance and faculty development around inclusive teaching methods. Additionally, regardless of their country of origin, many international students report being unfamiliar with and unprepared for discussion-focused pedagogies found in American college classrooms, due to the emphasis on lecture-based learning in the educational systems in their countries of origin (Nathan, 2005; Mina & Cimasko, 2020; Overstreet, 2021), suggesting another area for improved scaffolding and curriculum development for US STEM faculty seeking to equitably integrate discussion pedagogies into their teaching.

However, the strongly positive association between discussion and the grades of Asian female students (including Asian female international students) illustrates the complexity here. This finding aligns with TESOL research showing that female Chinese students learning English outperformed their male counterparts in English language learning (Goh & Foong, 1997; Gu, 2002), supporting female students' greater confidence in spoken language performance in the classroom. This research on gender differences in English language proficiency among Asian students raises interesting questions about how students' specific backgrounds and characteristics intersect with the racist and xenophobic classroom environment international students, especially Asian international students, have reported encountering in STEM classrooms, supporting calls like Heng's (2019) for research specifically investigating how their intersectional identities lead to different experiences of US learning environments.

Group work

The results for group work are the most surprising, since group work is one of the most widely recommended active learning strategies in both in WAC and SoTL literature for increasing student engagement and comprehension, especially in classes with diverse student populations (see Bean, 2011; Addy et al, 2021). Grades (ubiquitous, though problematic, measurements of learning) for most student subgroups were negatively correlated with group work, including for Hispanic students, countering research on equity-focused STEM education (see Burke et al., 2020). As with discussion activities, international and especially Asian international students were among the subgroups who did not benefit from increased group work. As with discussion, these findings parallel WAC research focusing on international ELL students, which have demonstrated how international students must work to avoid being excluded from groups and meaningful contributions to group projects based on their perceived lack of intelligence/potential due to their status as non-native English speakers (Phillips, 2014; McKee, 2018). These issues are exacerbated when faculty design curriculum (including prompts/topics for group assignments) around U.S.-centric assumptions about what is familiar, engaging, and accessible to students (Pelaez-Morales, 2018), setting international students up for failure rather than leveraging the diverse knowledge and perspectives they bring to U.S. classrooms.

Additional differences emerged here, however, for subgroups who reversed this trend: Asian female students again outperformed their peers in classes that used more group work, as did Asian students taking natural science courses. These findings may reflect the importance of *how* group work is implemented in writing-intensive classes. Bean's influential WAC handbook *Engaging Ideas* (2011) devoted a chapter to group activities, identifying issues that can compromise their effectiveness. In addition to poor group task design (see also Wolfe, 2010), Bean noted that "differences in learning style, gender, or ethnicity can explain some of the ways that various people behave in groups" (pg. 197). As Voss (2018) argued, making these cultural expectations for group work explicit is one important—and often lacking—form of scaffolding for setting up equitable and effective group work. Framing the issue as one of "differences," however, suggests that students whose learning style, gender, ethnicity, or other identity characteristics differ from the majority (i.e., marginalized students) should acculturate to the values of the majority, rather than scaffolding and facilitating group work so that it both recognizes and draws on the diversity of a class in order to promote all students' learning (see McKee, 2018).

The echoes between the negative correlations for international students—especially Asian international students—between course grades and both group work and discussion suggest that similar forms of stereotyping and xenophobic prejudice may be at work in both small and large group student-centered activities. Our findings suggest that a) implementing group work as an active learning strategy does not in and of itself necessarily correlate to more equitable learning outcomes along lines of race, ethnicity, and gender, and b) closer examination of the structuring of group work is needed, especially because for certain subgroups (e.g., Asian female students and Asian students in natural science classes), these broader inequitable trends were reversed.

Reading and writing

The outcomes we found for WAC-centric pedagogical activities were much more positive than for group work and discussion. Reading as a potential alternative to lecture, or as a lead-in to writing, discussion, or group work activities was correlated with higher grades for almost all student subgroups, especially in courses where reading was used more frequently. The results for writing–although a rarer classroom activity–were particularly encouraging for international students, suggesting that in-class writing activities may offer a way to offset some of the challenges discussion and group work activities pose for this student subgroup. The positive impact of reading and especially writing on student course performance, especially for international students, parallels existing WAC research connecting WtL activities in writing intensive disciplinary courses to higher grades and improved mastery of disciplinary discourses (see Hirsch, 2014).

As noted above in the analysis of course syllabi, much of the writing we observed was likely test- or quiz-taking, although it's worth noting that in the advanced, discipline-focused STEM courses included in this study exams typically call for complex problem-solving involving reading, writing, and critical thinking, rather than lower order multiple-choice testing seen in lower-level and more general college courses. The positive correlations between most student groups—and international students in particular—and (exam-focused) writing suggests that the testing practices in these STEM courses may align with the well-scaffolded, inclusive kinds of assessment design emphasized by Bailey & Durán (2020) that both scaffold the kinds of learning performances elicited by exams and provide space in the exam's design for students from a variety of backgrounds to perform well. Conversely, the nature of STEM exams—which often focus on problem solving, computation, programming, and other skills rather than extended composing of alphabetic texts—may avoid some of the linguistic bias Lindsey & Crusan (2011) found in studying faculty assessment of international students' writing. This is an area for expansion beyond this study, by directly studying the kinds of exams used in STEM courses and the types of pedagogies used to scaffold them (via both in- and out-of-class activities), to learn more about the specific types of writing characteristic of STEM courses and their relationships to the learning outcomes of different student subgroups.

Summary of Pedagogical Implications

The most important takeaway suggested by our findings about specific pedagogical activities is that it is unlikely that any single pedagogical technique will benefit all students, or benefit all students equally, echoing Asao B. Inoue's (2012) observations about the different effects of contract grading for students of different races. Although there is undoubtedly room for improvement in the design and implementation of the pedagogical activities documented in this study (as we did not focus on master teachers or best practices, but rather sought an "average" sample), the decisions faculty make about how to teach course content will likely require them to weigh the needs of some students above others. From an equity-based approach, this calculation should consider which students have been most disadvantaged in the discipline or program and prioritize their needs, recognizing that these choices may not cater to those students who have historically been best-served. This kind of approach aligns with the disparate impact analysis methods advocated by Poe and Cogan (2016), using evidence of historical disadvantage as a basis for making equitable choices about curriculum design. It is such comparative data that this study seeks to present as a beginning step in this direction: what are the relationships between different classroom learning activities for different student subgroups, and how can these relationships guide teachers to design more equitable, inclusive, and effective STEM pedagogies?

Future Steps

This study confirms some existing research about the negative impacts of relying on lecture in STEM education, documents the relatively low rate of adoption of WAC pedagogies in STEM classroom teaching (especially at institutions without robust WAC programs), and raises new questions about the relationship between student course performance and WAC/active learning teaching techniques in STEM classes that the data provided by this sample of classes are unable to answer. Especially because we did not focus exclusively on WtL pedagogies, and because many well-established WtL pedagogies occurring outside class (such as independent projects) were not included in the scope of this study, we do not see this study as questioning existing research that establishes the benefits of these pedagogies. Rather, we focus on the degree to which writing and writing-related pedagogies are used (or not) in STEM classes, and the associations these instructional practices have with course performance for different groups of students, seeking to complement existing WAC research on WtL pedagogies.

Future research—including ongoing, expanded research conducted by the authors—should increase its scope to include more classes from different STEM disciplines with the specific goal of including more underrepresented students in order to allow for quantitative analysis of the relationship between marginalized student subgroups and the teaching techniques documented here. The EPIC Lab has undertaken such expansion, scaling up the TailorEd Project between spring 2021 (when Santa Clara resumed in-person instruction after the onset of COVID-19) and spring 2022 to collect data from another 222 STEM courses, with plans to continue collecting data in subsequent academic years. Expanding the scope of the study will help us include larger numbers of underrepresented STEM students in order to investigate specific relationships between different pedagogical methods and student subgroups and thus extend our investigation of WAC pedagogies and equitable educational outcomes as well as to address some of the ambiguities and questions raised here. Furthermore, as we have worked with increasing numbers of faculty on this project, some have expressed interest in class-specific statistics that will help them track the allocation of time in their classes to different learning activities, heat maps visualizing the distribution of pedagogical activities across individual class periods and across the term, and other reports based on our data that they can use to get a bird's eye view of their teaching and compare it to their assessments of the class, student evaluations, and other outcome measures. These demonstrate the need for the kind of curriculum development tool the TailorEd Project has been working on creating: an interactive, data-driven visualization tool that uses data from previous classes to suggest the grade correlations for different student subgroups when different pedagogies are used, specified to different disciplines.

Additionally, while our research has prioritized scalability and a big-data approach that relies on available institutional data and restricts its focus to classroom activities that can be captured in still images, additional research that uses different measures of learning (such as holistic measures of learning and student self-assessment), detailed analyses of curricula (see Melzer, 2014; Voss, Sweeney, & Serviss, 2021), and other qualitative approaches (such as ethnographic classroom accounts, teacher research, and student interviews) that account more fully for the development, delivery, and reception of different teaching techniques is necessary to revise and further develop the preliminary findings we report here. As demonstrated here, our goal is to complement such studies' ability to address in depth the effects of contextual, interpersonal, intercultural, and institutional characteristics embedded in specific teaching and learning contexts that our methods do not fully address by helping to place their more specific and nuanced findings in a wider context.

References

Achen, Rebecca M., Lumpkin, Angela. (2015). Evaluating classroom time through systematic analysis and student feedback. *International Journal for the Scholarship of Teaching and Learning*, 9(2) https://digitalcommons.georgiasouthern.edu/ij-sotl/vol9/iss2/4/

Addy, Tracie Marcella, Dube, Derek, Mitchell, Khadijah A., & SoRelle, Mallory E. (2021). *What inclusive instructors do: Principles and practices for excellence in college teaching*. Stylus Publishing.

Anson, Chris. M. (2012). Black holes: Writing across the curriculum, assessment, and the gravitational invisibility of race. In Asao B. Inoue, & Mya Poe (Eds.), *Race and writing assessment* (pp. 15-28). Peter Lang.

Bailey, Alison L., & Durán, Richard. (2020). Language in practice: A mediator of valid interpretations of information generated by classroom assessments among linguistically and culturally diverse students. In Susan M. Brookhart, & James H. McMillan (Eds.), (pp. 46-62). Routledge/Taylor & Francis Group. https://doi.org/10.4324/9780429507533-4

Barr, Nancy. (2012). Revisiting Writing Across the Curriculum: Overcoming Faculty Reluctance to Integrating Writing into the Engineering Curriculum. Paper presented at the *Eleventh International Writing Across the Curriculum Conference,* 1-11. https://wac.colostate.edu/docs/proceedings/iwac2012/presentations/3HBarrPaper.pdf

Bean, John C. (2011). *Engaging ideas* (Second ed.). Jossey-Bass.

Blum, Susan. D., & Kohn, Alfie. (2020). *Ungrading: Why rating students undermines learning (and what to do instead)* (First ed.). West Virginia University Press.

Breslow, Lorl. (2010). Wrestling with pedagogical change: The TEAL initiative at MIT. *Null, 42*(5), 23-29. https://doi.org/10.1080/00091383.2010.503173

Bunker, Aaron, & Schneider, Jeremy. (2015). Writing Across the Curriculum in a Human Physiology Class to Build Upon and Expand Content Knowledge. *Athens Journal of Education, 2*(4), 331-344. https://doi.org/10.30958/aje.2-4-3

Burke, Christian, Luu, Rochard, Lai, Andy, Hsiao, Valerie, Cheung, Elwin, Tamashiro, Dustin, & Ashcroft, Jared. (2020). Making STEM equitable: An active learning approach to closing the achievement gap. *International Journal of Active Learning, 5*(2), 71-85. https://www.learntechlib.org/p/218451

Foong, Kwah Poh, & Goh, C. C. M. (1997). Chinese ESL students' learning strategies: A look at frequency, proficiency, and gender. *Hong Kong Journal of Applied Linguistics, 2*(1), 39.

Foote, Kathleen, Knaub, Alexis, Henderson, Charles, Dancy, Melissa, & Beichner, Robert J. (2016). Enabling and challenging factors in institutional reform: The case of SCALE-UP. *Physical Review Physics Education Research, 12*(1), 10103.

Gierdowski, Dana. (2013). Studying learning spaces: A review of selected empirical studies. In Russell G. Carpenter (Ed.), *Cases on higher education spaces: innovation, collaboration, and technology* (pp. 14-39). IGI Global. https://doi.org/10.4018/978-1-4666-2673-7.ch002

Gu, Yonqi. (2002). Gender, academic major, and vocabulary learning strategies of chinese EFL learners. *RELC Journal, 33*(1), 35-54. https://doi.org/10.1177/003368820203300102

Hartikainen, Susana, Rintala, Heta, Pylväs, Laura, & Nokelainen, Petri. (2019). The concept of active learning and the measurement of learning outcomes: A review of research in engineering higher education. *Education Sciences, 9*(4) https://doi.org/10.3390/educsci9040276

He, Kaiming, Zhang, Xiangyu, Ren, Shaoqing, & Sun, Jian. (2016). Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 770-778. https://doi.org/10.1109/CVPR.2016.90

Heng, Tang T. (2019). Understanding the heterogeneity of international students' experiences: A case study of Chinese international students in U.S. universities. *Journal of Studies in International Education, 23*(5), 607-623. https://doi.org/10.1177/1028315319829880

Herrington, Anne J. (1981). Writing to learn: Writing across the disciplines. *College English, 43*(4), 379-387. https://doi.org/10.2307/377126

Hibbard, Lisa, Shannon Sung, & Wells, Breché. (2016). Examining the effectiveness of a semi-self-paced flipped learning format in a college general chemistry sequence. *Journal of Chemical Education, 93*(1), 24-30. https://doi.org/10.1021/acs.jchemed.5b00592

Hirsch, Linda. (2014). Writing intensively: An examination of the performance of L2 writers across the curriculum at an urban community college. In Terry M. Zawacki, Michelle Cox & Jonathan Hall (Eds.), *WAC and second-language writers: Research towards linguistically and culturally inclusive programs and practices* (pp. 151-180). University Press of Colorado; WAC Clearinghouse.

Inoue, Asao. B. (2012). Grading contracts: Assessing their effectiveness on different racial formations. In Asao B. Inoue, & Mya Poe (Eds.), *Race and writing assessment* (pp. 79-94). Peter Lang.

Kells, Michelle Hall. (2007). Writing across communities: Deliberation and the discursive possibilities of WAC. *Reflections: A Journal of Writing, Service-Learning, and Community Literacy, 1*(1), 87-108.

Knoblauch, C. H., & Brannon, Lil. (1983). Writing as learning through the curriculum. *College English, 45*(5), 465-474. https://doi.org/10.2307/376844

Kovac, Jeffrey, & Sherwood, Donna W. (1999). Writing in chemistry: An effective learning tool. *Journal of Chemical Education, 76*(10), 1399-1403. https://doi.org/10.1021/ed076p1399

Kramer, Tereza Joy, Zeccardi, Joe, Concepcion, Rebecca, Emhoff, Chi-An, Miller, Steve, & Posell, Krista Varela. (2019). WID course enhancements in STEM: The impact of adding "writing circles" and writing process pedagogy. *Across the Disciplines, 16*(4), 26-37.

Krizhevsky, Alex, Sutskever, Ilya, & Hinton, Goeffrey E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM, 60*(6), 84-90. https://doi.org/10.1145/3065386

Latulipe, Celine, Rorrer, Audrey, & Long, Bruce. (2018). Longitudinal data on flipped class effects on performance in cs1 and retention after CS1. Paper presented at the *ACM Technical Symposium on Computer Science Education,* 411-416.

Lee, Young, Boatman, Eric, Jowett, Steven, & Guenther, Brendan. (2014). REAL: The technology-enabled, engaged, and active learning classroom. *International Journal of Designs for Learning, 5*(1) https://www.learntechlib.org/p/209652

Lindsey, Peggy, & Crusan, Deborah. (2011). How faculty attitudes and expectations toward student nationality affect writing assessment. *Across the Disciplines, 8*(4), 1-40. https://wac.colostate.edu/docs/atd/ell/lindsey-crusan.pdf

McGee, Ebony O., Thakore, Bhoomi K., & LaBlance, Sandra S. (2017). The burden of being "model": Racialized experiences of Asian STEM college students. *Journal of Diversity in Higher Education, 10*(3), 253-270. https://doi.org/10.1037/dhe0000022

McKee, Heidi. A. (2018). Intercultural communication and teamwork: revising business writing for global networks. In Shirley K. Rose & Irwin Weiser (Eds.), *The internationalization of US writing programs* (pp. 185-200). University Press of Colorado. https://doi.org/10.2307/j.ctt22h6qmq.13

Melzer, Dan. (2014). *Assignments across the curriculum; A national study of college writing.* University Press of Colorado.

Michael, Joel. (2006). Where's the evidence that active learning works? *Advances in Physiology Education, 30*(4), 159-167. https://doi.org/10.1152/advan.00053.2006

Miller, Emily R., & King, Tara. (2019). *Promoting transformation of undergraduate STEM education: Workshop summary report.* (American Association of Universities). https://www.aau.edu/sites/default/files/AAU-Files/STEM-Education-Initiative/Promoting-Transformation-Report.pdf

Mina, Lillian. W., & Cimasko, Tony. (2020). Expectations, mismatches, and translingual dispositions in teaching multilingual students. In Alanna Frost, Julia Kiernan & Suzanne Blum-Malley (Eds.), *Translingual dispositions: Globalized approaches to the teaching of writing* (pp. 63-80). The WAC Clearinghouse; University Press of Colorado.

Montenegro, Erick, & Jankowski, Natasha. A. (2020). *A new decade for assessment: Embedding equity into assessment praxis (Occasional Paper No. 42).* University of Illinois and Indiana University: National Institute for Learning Outcomes Assessment (NILOA). https://www.learningoutcomesassessment.org/wp-content/uploads/2020/01/A-New-Decade-for-Assessment.pdf

Moon, Alena, Gere, Anne R., & Shultz, Ginger V. (2018). Writing in the STEM classroom: Faculty conceptions of writing and its role in the undergraduate classroom. *Science Education, 102*(5), 1007-1028. https://doi.org/10.1002/sce.21454

Nathan, R. (2005). *My freshman year: What a professor learned by becoming a student.* Cornell University Press.

Nielsen, Kathryn. (2014). On class, race, and dynamics of privilege: Supporting generation 1.5 writers across the curriculum. In In Terry M. Zawacki, Michelle Cox & Jonathan Hall (Eds.), *WAC and second-language writers: Research towards linguistically and culturally inclusive programs and practices* (pp. 129-150). University Press of Colorado; WAC Clearinghouse.

Overstreet, Matthew. (2021). Critical thinking, writing, and language learning: A report from northwest China. In Bruce Morrison, Julia Chen, Linda Lin & Alan Urmston (Eds.), *English across the curriculum: Voices from around the world* (pp. 223-238). WAC Clearinghouse.

Pelaez-Morales, Carolina. (2018). Internationalization from the bottom up: Writing faculty's response to the presence of multilingual writers. In Shirley K. Rose & Irwin Weiser (Eds.), *The internationalization of US writing programs* (pp. 234-254). University Press of Colorado. https://doi.org/10.2307/j.ctt22h6qmq.16

Phillips, Tallin. (2014). Developing resources for Success: A case study of a multilingual graduate writer. *WAC and second-language writers: Research towards linguistically and culturally inclusive programs and practices (pp. 69-91). University Press of Colorado; WAC Clearinghouse.*

Poe, Mya. (2013). Reframing race in teaching writing across the curriculum. In Asao. B. Inoue, & Mya Poe (Eds.), *Race and writing assessment* (pp. 87-105). Peter Lang.

Poe, Mya, & Cogan, J. A. (2016). Civil Rights and Writing Assessment: Using the Disparate Impact Approach as a Fairness Methodology to Evaluate Social Impact. *The Journal of Writing Assessment, 9*(1) http://journalofwritingassessment.org/article.php?article=97

Poe, Mya, & Craig, John Aloysius. (2011). Modeling multivocality in a U. S.-Mexican collaboration in writing across the curriculum. *Journal of Academic Writing*, 1(1), 29-36. https://doi.org/10.18552/joaw.v1i1.23

Russell, David R. (1991). *Writing in the academic disciplines, 1870–1990: A curricular history* (1st ed.). Southern Illinois University Press.

Russell, David R. (1990). Writing across the curriculum in historical perspective: Toward a social interpretation. *College English, 52*(1), 52-73.

Scheurer, Erika. (2015). What do WAC directors need to know about "coverage"? *WAC Journal, 26*, 7-21.

Serviss, Tricia, & Voss, Julia. (2019). Researching writing program administration expertise in action: A case study of collaborative problem solving as transdisciplinary practice. *College Composition and Communication, 70*(3), 446-475.

Shaghaghi, Navid, Khadadeh, Mohammed, McGinnis, Meghan, Liang, Liying, & Calle, Andres. (2019). Classroom configuration identifier (CCID). *2019 IEEE 11th International Conference on Engineering Education,* 204-209. https://doi.org/10.1109/ICEED47294.2019.8994827

Syracuse University Department of Writing Studies, Rhetoric, and Composition. (2022). *Antiracist WAC toolkit.* https://thecollege.syr.edu/writing-studies-rhetoric-and-composition/writing-across-curriculum/antiracist-wac-toolkit/

Trytten, Deborah A., Lowe, Anna Wong, & Walden, Susan E. (2012). "Asians are good at math. What an awful stereotype": The model minority stereotype's impact on Asian American engineering students. *Journal of Engineering Education, 101*(3), 439-468. https://doi.org/https://doi.org/10.1002/j.2168-9830.2012.tb00057.x

Voss, Julia. (2018). Who learns from collaborative digital projects? Cultivating critical consciousness and metacognition to democratize digital literacy learning. *Composition Studies, 46*(1), 57-80.

Voss, Julia, Sweeney, Meghan A., & Serviss, Tricia. (2021). A heuristic to promote inclusive and equitable teaching in writing programs. *WPA: Writing Program Administration, 44*(2), 13-39.

Warikoo, Natasha, Sinclair, Stacey, Fei, Jessica, & Jacoby-Senghor, Drew. (2016). Examining racial bias in education: A new approach. *Educational Researcher, 45*(9), 508-514. https://doi.org/10.3102/0013189X16683408

Wilson, Jack M. (1994). The CUPLE physics studio. *The Physics Teacher, 32*(9), 518-523. https://doi.org/10.1119/1.2344100

Winfield, Leyte L., Hibbard, Lisa B., Jackson, Kimberly M., & Johnson, Shanina Sanders. (2019). Cultivating agency through the chemistry and biochemistry curriculum at Spelman College. *Broadening participation in STEM* (pp. 153-181). Emerald Publishing Limited. https://doi.org/10.1108/S1479-364420190000022007

Wolfe, Joanna. (2010). *Team writing: A guide to working in groups*. Bedford/St. Martin's.

Notes

[1] This research was made possible by the financial support of Nancy Cutler, Santa Clara University's Deputy CIO for Academic Technology; the technical support of Joel Bennet, Media Systems Group Manager, and Sean Kennedy, Senior Media Systems Specialist, both of Santa Clara University's Academic Technology Office; the institutional data support of Conor Roycroft, Assistant Director for Business Intelligence and Data Analytics in Santa Clara University's Office of Institutional Research; and the data collection support of Denise Krane, Director of the HUB Writing Center and Lecturer in the Department of English at Santa Clara University. Past TailorEd machine learning team student members Stefan D'Costa, Mohammed Khadadeh, Liying Liang, and Meghan McGinnis contributed to the development of the CCID model, which was used to create the CAID model used in this paper. 2020-2021 TailorEd undergraduate student researchers Liam Abbate, Roland Afaga, Laurynn Diby, Katya Keklikian, Kristin Lee, Justin Ling, Ann Nguyen, Julia Perry, Kishann Rai, Ashika Rajesh, Gabe Reed, Malika Singh, Fiona Sundy, Sneha Vinod, Morgan Yazdi, and Kristina Yin analyzed and labeled thousands of classroom photos, enabling the development of CAID. Additionally, Ashika Rajesh processed and organized thousands of photos for use in multiple iterations of manual photo labeling and CAID development, and Kristin Lee cleaned, organized, scaled, and transformed institutional research data to enable the quantitative analysis done in this paper.

[2] This study is part of a larger project "What's in a Classroom? Examining Embodied Teaching and Learning Experience in Classrooms Designed for Active Learning" (Santa Clara University IRB Protocol #14-01-463), which includes both STEM and non-STEM courses. The photos used in this study were obtained by requesting permission from instructors to visit their classes to invite their students to participate in the project. Photos were taken only in classes in which all students and the instructor gave written informed consent to participate, following a short presentation describing the study's goals and procedures and displaying sample classroom photos (like the ones shown in Figure 1).

[3] There were two additional categories used in the human coding phase, but they occurred so rarely in the dataset that they were dropped from the analysis for this paper:

- **Unstructured:** when students are standing up or moving around the classroom, when there is no majority activity while students are engaged in a wide variety of activities, especially unstructured ones such as texting, chatting, or asking the professor a 1-on-1 question (as often happens at the start or end of class or during breaks)
- **Head turned:** when students confer briefly by only turning their heads together, not their bodies

[4] If there were fewer than 5 people in a classroom, the room was counted as empty, because photos like this often showed the class arriving or departing, a few students staying to talk with an instructor, or students using an empty classroom to study in. Because these photos did not show structured learning activities that included full classes of students, they were marked as "empty."

[5] A note on data interpretation: The correlation coefficients (r values) describing the relationship between classroom activities and grades (disaggregated according to student and course characteristics) reported in Tables 2, 3, and 4 can range from -1 to +1. The higher the absolute value of the coefficient, the stronger the relationship (either positive or negative). Positive r values mean that as one variable increases, so does the other, showing a direct relationship. Negative r values mean that as one variable increases, the other decreases, indicating an inverse relationship. This study uses the conventional threshold of statistical significance of a significance test (p value) less than or equal to .05, which suggests that there is a 5% or less chance that the relationships represented by the r value coefficient is due to chance: all coefficients reported in Tables 3, 4, and 5 have a p value of .05 or less (coefficients with a p value higher than .05 were not reported).

[6] Because we relied on data from SCU's Office of Institutional Research, we use the following Integrated Postsecondary Education Data System (IPEDS) categories for race and ethnicity that students use to self-identify when they apply to the university: American Indian or Alaska Native, Asian, Black or African American,

Hispanic of any race, Native Hawaiian or other Pacific Islander, Two or more races, White, and Race and ethnicity unknown (used for any applicant who does not answer the optional race/ethnicity question).

[7] "Continuing generation" refers to students whose parents attended college.

Contact

Julia Voss
Associate Professor of Rhetoric and Composition
Department of English
Santa Clara University
Email: jvoss@scu.edu

Navid Shaghaghi
Lecturer
Departments of Math & Computer Science, Computer Science and Engineering, Bioengineering, and Information Systems and Analytics
Director, Ethical, Pragmatic, and Intelligent Computing (EPIC) Research Laboratory
Faculty Fellow, Frugal Innovation Hub & BioInnovation and Design Lab
Santa Clara University
Email: nshaghaghi@scu.edu

Andres Mauricio Calle
Ph.D. Student
Department of Computer Science and Engineering
Santa Clara University
Email: acalle@scu.edu

Kristin Lee
B.S. in Economics and Commerce, Accounting, & Information Systems
Santa Clara University
Email: krstnn.lee@gmail.com

Liam Abbate
English Teaching Assistant
Cusset, Auvergne-Rhône-Alpes, France
Teaching Assistant Program in France
Email: lp2abbate@gmail.com

Complete APA Citation

Voss, Julia, Shaghaghi, Navid, Calle, Andres Mauricio, Lee, Kristin, & Abbate, Liam. (2022, November 4). Lecture, discussion, group work, repeat: Using aerial photography and machine learning to study the use of writing-related pedagogies in STEM courses and their impact on different student subgroups. [Special issue on *STEM and WAC/WID: Co-Navigating Our Shifting Currents*] *Across the Disciplines*, 19(1/2), 99-121. https://doi.org/10.37514/ATD-J.2022.19.1-2.07