Developing CALI: An Inventory to Capture Collaborative Active Learning and Inclusive Practices in Introductory CS Courses

Celine Latulipe*, Sri Yash Tadimalla†, Mary Lou Maher†, Tonya K. Frevert†, Marlon Mejias†, Jamie Payton‡, Audrey Rorrer†, John Fiore‡, Gene Kwatny‡, Andrew Rosen‡, Leilani Bell‡

*University of Manitoba
Winnipeg, Canada
celine.latulipe@umanitoba.ca

†University of North Carolina at Charlotte
Charlotte, U.S.A.
{stadimal, mmaher9, tfrevert, mmejias, audrey.rorrer}@uncc.edu

‡Temple University
Philadelphia, U.S.A.
{payton, john.fiore, gkwatny, andrew.rosen, leilani.bell}@temple.edu

***Abstract*—This full paper presents CALI 1.0: an inventory to capture the various components of collaborative active learning and inclusive practices in introductory CS courses, as a basis for studying their impact on a student's ability to engage with course material and be successful in the course. Collaborative active learning is defined as a pedagogy where learners work with one another on activities that involve actively engaging with the course material through discussions, problem-solving, case studies, role plays, and other methods. CALI consists of three separate indices: Structure, Sociality, and Inclusiveness. The Structure Index includes components related to course setup, organization, assessment, grading, and communications. The Sociality Index includes components related to opportunities for students to interact with each other. The Inclusiveness Index includes components related to how the instructor communicates a sense of belonging to the students through a growth mindset and inclusive policies and practices. In this paper, we present the inventory and results from a focus group of faculty to gauge responses to using this inventory to capture the teaching components of their courses. We present CALI as a tool that can be used to study how teaching practices, e.g. inclusive and collaborative active learning pedagogies, impact course experiences for students in different demographic groups as well as a reflective tool for use by faculty while designing their courses.**

***Index Terms*—Active Learning, Student Engagement, Diversity, Sense of Belonging, Faculty Engagement.**

I. Introduction

Introductory CS courses may inadvertently act as "gatekeepers" that prevent student success in the CS major, affecting retention and contributing to the "leaky pipeline", especially for women, non-binary, and racially marginalized students [1]–[3]. Student success in introductory CS courses has been generally linked with prior coding experience [4], [5], course design/teaching practices [6], as well as students' sense of belonging [7], [8]. To our knowledge, however, research to-date has not yet attempted to quantify specific course design and teaching practices in a way that allows for explicit examination of the relationship between variations in teaching practice and student experiences and success in these courses [9]. To address this gap, we present an inventory to capture the effects of various teaching practices such as collaborative and active learning practices and inclusive communications strategies on student experiences in introductory CS courses.

When students enter a CS course setting, they bring – along with their interest in learning computer science – their demographic and cultural identities. In CS education research, there has been some analysis of the role that students' race, gender, and social class plays in student outcomes across the K-20 CS education pipeline [10]. Gaps in equitable race and gender participation in CS have been widely acknowledged in CS education research, with much effort put forth to rectify them [11]. Emphasis has focused on the CS course setting, with many change efforts aimed at shifting CS teaching practices from passive learning to active learning [12], [13], and from individual learning to collaborative learning [14], based on evidence that these practices yield a more inclusive course setting, thus helping to close equity gaps [15], particularly in introductory CS courses [16]. There is increasing attention paid to the need for culturally-relevant computing that helps students see how computing can be used to address inequities in their communities [17], [18]. CS faculty are also paying more attention to their explicit communication strategies, how to better communicate and foster a growth mindset amongst

Supported by NSF Grant: 2111376

students, based on evidence that this can foster inclusive success [19]–[21].

While many best practices to promote inclusive success are being applied, the CS education field lacks a nuanced understanding of how such teaching practices, in combination, impact different student groups. For example, Quille and Bergen showed that a growth mindset intervention had positive impacts on programming performance, but did not effect all student groups equally [21]. Many faculty try to combine a variety of techniques, attempting to employ multiple best practices [16]. Variation in adoption makes it difficult to tease out the effectiveness of such practices. Some CS education researchers have attempted more controlled studies to isolate the impact of a particular technique, using it in one section of a course, with another section as a control group. However, experimental approaches have ethical implications in educational settings because they deny access to hypothesized best practices to one set of students (the students in the control group), and do not capture the effects of combining various best practices together. Thus, there is a need to study how teaching practices (collaborative, active and inclusive classroom techniques) used in combination, impact student experiences and success. There are exponential combinations of teaching techniques, and so attempting to enumerate and evaluate all potential combinations is not likely to be useful.

In this paper, we present the Collaborative Active Learning and Inclusiveness (CALI) inventory, which can be used to record teaching practices applied in a course along three dimensions: structure, sociality and inclusion. By characterizing the amount of structure in a course, the amount of opportunities for social learning, and the number of different techniques used to build an inclusive classroom environment, we expect to be able to consider the impact of these components on different groups of students more generally. In this paper we explain how the inventory was developed, introduce the inventory, and present results from both a pilot study of the inventory and a focus group of faculty who were asked to use the inventory to characterize one of their own courses.

II. Related Work

Lack of diversity in the CS field can lead to the recreation of existing system inequities from society into current and future technologies [22]. A lack of diversity is an increasingly alarming pattern as racially marginalized students in the United States currently make up 30% of the U.S. population, but students from these populations represent a very small portion of those who earn a science, technology, engineering, and mathematics (STEM) undergraduate degree [23]. There is mounting evidence that active learning is beneficial for all students, but leads to especially significant gains for students from marginalized groups [15]. However, there are many different ways to employ active learning [24], and mixed results from the literature suggest that not all active learning interventions 'work'. In the life sciences, there has been a shift from talking about specific active learning practices to a focus on 'course structure', which is often increased in classes to accommodate active learning activities and to ensure student participation [25]–[27]. For example, to accommodate in-class active learning activities, an instructor may flip a class, require students to watch content videos or complete guided readings before class, and then use homework quizzes to act as a forcing function to get students to come to class prepared. Eddy and Hogan first noted that it was possibly this increase in structure and forcing functions, as much as the nature of the active learning activities themselves, that led to increased student success, especially for first generation college students [28]. Gross et al. showed that the increased completion of prep-work in flipped, active learning classes led to better learning outcomes in a Chemistry course [29]. There is early evidence of similar impacts in computing [14], [30].

There have been previous efforts to measure active learning. Van Amburgh et al. developed an Active Learning index that is a low-level observational protocol for use in the classroom [31]. That tool is similar in nature to the COPUS protocol, which collects minute-by-minute observations in a classroom to characterize what is happening [32]. While these are useful tools, they are more-fine grained than we desire for our purposes. Our intention is to record the general practices in a course, rather than the percent of time spent on particular activities in a given class meeting. The Structural Index in the CALI inventory aims to capture the level of structure provided in a computing course overall.

Another pedagogical element that often, but not always, goes hand-in-hand with active learning is collaboration and social interaction [33]. This element may come in many forms, such as pair programming [34], team assignments or capstone projects [35], lightweight teams for in class activities [36], peer instruction quizzing [37], peer mentoring [38], and class discussion forums [39]. These elements can all contribute to helping students engage in their learning, see content from other people's perspectives, learn to work together, and build community within the classroom. As described by Deil-Amen [40], incorporating such elements into a course can provide socio-academic integrative moments that can help students feel more connected to their major. The Sociality Index in the CALI inventory attempts to capture the variety of ways that courses can provide connective moments.

There is increasing attention given to what constitutes equitable access in post-secondary computing education, including the ways that different curricula, tools, and guiding pedagogies impact equitable learning for different groups of students [41], [42]. Students who enroll in introductory CS courses often come with different motivation levels, expectations, computing experience, and computing perceptions based on their childhood upbringing, and K-12 experiences [20], [43]. Background variations can significantly impact how students experience pedagogical practices that they encounter in introductory CS courses. Equity in introductory CS courses must focus not only on who gets access to the courses, but the modes of pedagogy used within classrooms and curricular policies of those courses. And it is not sufficient to look at individual demographic groups when considering equity concerns - there

is growing consensus that an intersectional lens [44] is needed to understand the complexity of experiences for students at the intersection of underrepresented identities [45], e.g., female students of color [45], particularly in Computer Science [46]. The growth mindset of the instructor teaching the course has been shown to play a role in student success in STEM, and a lack of faculty growth mindset is particularly harmful to students from under-represented groups [47]. The CALI inventory Inclusiveness Index attempts to capture equity-related practices by looking at growth mindset messaging, inclusive pedagogical policies, and the classroom cultural environment.

III. The CALI Inventory

We developed the Collaborative Active Learning and Inclusiveness (CALI) inventory to enable investigation of the various components of classroom instruction and course design that impact a student's ability to effectively engage with introductory CS course materials. CALI 1.0, the initial version presented in this paper, is expected to be further refined and eventually serve to both study practices and effects on students, as well as to serve as a reflective guide for faculty seeking to create more collaborative, active learning course experiences for students. The inventory consists of three indices: Structure, Sociality, and Inclusiveness. Structure is defined as the way the course is set up and delivered, and is captured by items such as the way materials and activities are organized and delivered, how grading is implemented, and the mechanisms for communication and attendance, etc. Sociality is defined as the level of collaborative learning and peer-to-peer engagement in the course, captured by items about the quality and quantity of opportunities for students to interact and work together. Inclusiveness characterizes how the course policies, materials, grading, and feedback mechanisms communicate a culture of belonging and growth mindset to students.

A. Development History & Design Goals

The Connected Learner project at the University of North Carolina Charlotte aimed to transform faculty attitudes towards teaching by creating a community of practice shifting faculty attitudes and teaching practices away from knowledge transmission and didactic teaching toward a student-centered mindset. The project researchers examined the literature to understand best practices and experimented with a variety of innovative teaching methodologies. Through CS course observations, the research team noted that a higher level of structure in active learning courses was often coupled with higher levels of sociality, and that these instructors communicated inclusive messaging in their content delivery and course policy decisions. However, detangling the effects of each of these components was difficult, making it challenging to discern which components were most effective at improving student engagement and performance. The desire to untangle teaching practice effects led to the development of the CALI inventory.

Active learning pedagogies have been shown to promote success and narrow the achievement gap for students who have historically been underrepresented in the computing field [14], [15]. The CALI inventory, while focused on active learning (it is in the name!), does not have any individual items that explicitly record active learning. This omission is intentional because active learning takes so many different forms under many different names, making it difficult to include them all [24]. Thus, items on the three indices capture the course structure and social interactions that tend to occur as an instructor implements active learning approaches, regardless of the specific activity.

The inventory can be viewed as taking a systems ecology approach. The classroom is a complex environment, and there are many interacting teaching components that can work together to create a healthy learning environment. The CALI inventory enables analysis by grouping components into systems: the structural system, the social system, and the inclusiveness (policy) system. These systems together form the learning environment.

The inventory is designed to be used to record practices within a particular section of a course, since different sections are typically taught by different instructors and in-class practices vary, even if certain high-level elements of the course are kept the same across sections. Throughout the rest of this paper we use the term course and, for multi-section courses, we would envision the inventory being recorded separately for each section. To use these indices, which are quite wide-ranging in the information recorded, a researcher would gather information from the learning management system (LMS) course site, as well as through class observation and/or interviews with the instructor.

B. The Structural Index

The Structural Index focuses on how much structure there is in the course, and allows courses to be characterized along a spectrum from free-form to highly structured. This index measures the level of detail about content provided to students upfront, how grades are assigned, how material is delivered, and puts particular emphasis on the amount of different activities students are expected or required to participate in and complete. Figure 1 shows the Structural Index. For most of the dimensions, an observer or researcher checks the description that most closely resembles what is happening in the class. For grading and for feedback the observer checks all applicable elements.

To illustrate the use of the Structural Index, we describe two hypothetical courses that would be at either end of the structure spectrum:

- **Free-Form**: Course has a single final exam worth 100% of the grade, no required attendance, weekly outline, assignments or in-class activities. A textbook is listed on the syllabus. Students decide whether to attend lectures or not, and whether to read the textbook or not.
- **High-Structure**: Course has 4 term tests, weekly quizzes, 2 graded assignments, a final project, regular in-class activities, a comprehensive course page on the LMS that shows what will be covered each week and what readings

Dimension	Description of Application in Course	
Grading (check all that apply)	Midterm	☐
	Final	☐
	Multiple Tests (3+)	☐
	Assignments	☐
	Quizzes	☐
	Blog posts or writing reflections	☐
	Graded discussions	☐
	Graded classroom participation	☐
Student Learning Outcomes	Not stated	O
	Listed in syllabus	O
	Referred to in learning activities	O
Assigned work outside of class	Completely up to the student	O
	Reading lists, videos, sample exercises provided, but not required	O
	Homework or prep-work required but not graded	O
	Homework or prep-work required and graded or tested	O
Attendance	Not monitored	O
	Occasionally taken, but not part of grade	O
	Occasionally taken, part of grade	O
	Taken regularly, part of grade	O
Participation	Not monitored	O
	Occasionally monitored, but not part of grade	O
	Occasionally monitored, part of grade	O
	Regularly monitored, part of grade	O
Feedback (check all that apply)	Returned exams/assignments	☐
	Early tests, quizzes to help students get feedback	☐
	Regular opportunities for feedback, even if not graded	☐
Scheduling	Ad hoc schedule, things happen at instructor's discretion	O
	Some regularly scheduled activities/exams/assignments	O
	Full semester calendared: quiz, assignment, test dates in advance	O
	Full semester calendared, plus regular weekly rhythm of work	O
LMS Use	Course materials not provided online	O
	Course materials provided on static web site	O
	Course materials provided using LMS with integrated gradebook	O
	Participation on LMS is an integral part of the course completion	O

Fig. 1. The Structural Index. A researcher or observer selects the description within each dimension that most accurately represents what is happening in the course, based on examination of course materials as well as class observations and/or instructor interviews.

Dimension	Description of Application in Course	
Student Teamwork During Class (check all that apply)	Occasional ad hoc teams	☐
	Long term lightweight teams (ungraded teamwork)	☐
	Middleweight teams (teams for graded assignments)	☐
	Project teams (large, long-term graded assignment)	☐
Student Teamwork in Lab	No lab	O
	No teams in lab	O
	Occasional or optional pair activities	O
	Regular pair activities in labs (pair programming, etc.)	O
Student Teamwork Outside of Class	No teamwork outside of class	O
	Facilitated study groups	O
	Long term project teamwork	O
	Long term project teamwork and facilitated study groups	O
Online Discussion Forum	None provided	O
	Discussion forum with anonymous posting allowed	O
	Regular discussion forum (not graded)	O
	Discussion forum (graded)	O
	Discussion forum with points for active participation	O
Online Chat Room	None provided	O
	Chat room usage is encouraged during class time	O
	Chat room encouraged within and outside of class time	O
In Class/Lab Collaborative Active Learning	None	O
	Occasional class discussion	O
	Occasional group activities	O
	Occasional class discussion and occasional activities	O
	Occasional class discussion and regular activities	O
	Active learning activities with group/partner every class	O
In Class Testing / Assessment	No group testing	O
	Partial group quizzes (may answer individually first)	O
	Group-based quizzes/tests	O
	Group quizzes and 2-stage exams	O
In Class Warmups	None	O
	Occasional warmups (ice-breaker questions/activities)	O
	Regular (weekly or every meeting) warmups	O
Outside Help	None	O
	Individual office hours/tutoring online	O
	Group help sessions/tutoring sessions	O
Classroom Movement	Students stay seated throughout class period	O
	Occasional opportunities to move and talk to others	O
	Regular opportunities to move and talk to others	O

Fig. 2. The Sociality Index (in-person version).

or prep work students are expected to complete before class. Attendance is mandatory and recorded.

C. The Sociality Index

The Sociality Index is designed to characterize how much opportunity a student has to interact with other class peers. The interaction may be highly structured and/or part of a graded assessment (e.g. pair programming or group tests) or very informal and ungraded (e.g. think-pair-share or group discussions). The course may provide online forums for interaction. The Sociality Index has two versions: one for in-person or hybrid courses and one for courses that are completely online (not shown). The dimension that differs for online courses is Movement, which has descriptions related to breakout rooms instead of physical movement. Figure 2 shows the Sociality Index for in-person and hybrid courses.

To illustrate the use of the Sociality Index, we describe two hypothetical courses from each end of the sociality spectrum:

- **Individual**: Course is a lecture-based course, with all individual assignments, quizzes and exams. There are no group or pair activities in class or lab. Communication with the instructor is all handled through private email, there are no online discussion forums.
- **Collective**: Course is a flipped, active learning class with permanent lightweight teams. Students do all in-class activities in a group and are often encouraged to move and talk with other teams. The course has an online discussion forum and biweekly group tutoring sessions.

D. The Inclusiveness Index

The Inclusiveness Index, shown in Figure 3, aims to characterize ways that a course design may convey a more inclusive and welcoming infrastructure and environment to all students, particularly to students from marginalized groups.

To illustrate the use of the Inclusiveness Index, we describe two hypothetical courses that would be at either end of the inclusiveness spectrum:

- **Traditional**: Course has strict attendance and participation policies, with no latitude or accommodations for student groups with special needs. The materials do not include perspectives or examples from marginalized

Dimension	Description of Application in Course	
Course Syllabus (check all that apply)	Provides information about campus resources	☐
	Includes statements of accessibility and accommodations related to disability, sexuality, religion, veteran status or need	☐
	Includes SLOs promoting ethics and growth mindset	☐
	Incorporates diversity statement and students invited to discuss	☐
Diverse Interaction	No opportunities for students to work in diverse teams	○
	Formation of teams is left to students or automated	○
	Teams are created with diversity in mind	○
	Encourage and support opportunities for peer support	○
	Explicitly fosters collaboration and community, acceptance of diverse identities through discussion and reflection activities	○
Attendance	No attendance policy / attendance not required	○
	Attendance policy is fixed	○
	Attendance policy is fixed with clear processes for exceptions	○
	Attendance policy sets expectations & allows student feedback	○
Grading (check all that apply)	Explain criteria used for evaluation and give a rationale	☐
	Uses anonymous grading when appropriate to avoid bias	☐
	Include assessment of inclusive behavior during group work	☐
	Multiple ways to meet SLOs and demonstrate learning	☐
	Learning assessed early and often using variety of methods	☐
	Encouraged to demonstrate development (growth mindset)	☐
	Grading policies set expectations & encourage student feedback	☐
Content Context	Traditional material not reviewed for diverse perspectives	○
	Material reviewed for inaccuracy and absence of relevant perspectives. Students alerted to limitations.	○
	Some culturally responsive materials content included	○
	Culturally responsive materials written by people of different backgrounds or by members of the group discussed by the class	○
	Opportunities to critique materials, discuss experiences, correct inaccuracies or misrepresentations related to identity groups	○
Feedback	Predominantly provided through course assessments	○
	Feedback at checkpoints/before final	○
	Required check-ins with instructor or TAs	○
	Reflection based checkpoints (self-feedback)	○
	Timely feedback provided to students not meeting expectations	○
Participation	Participation based on attendance or not observed	○
	Participation is based on subjective instructor/TA observation	○
	Participation assessed through provided rubric	○
	Variety of inclusive approaches to encourage participation	○

Fig. 3. The Inclusiveness Index. For the two shaded dimensions, the researcher or observer selects all that apply.

groups. There is no opportunity for student reflection, interaction with diverse groups, or to provide feedback on course policies.

- **Inclusive**: Course has clear policies for attendance, feedback, and participation, and the instructor communicates flexibility and responsiveness to course policy feedback. Students work with peers within the class in diverse groups. Materials have been updated with readings from/about marginalized voices. The instructor and course policies communicate a growth mindset belief in students' ability to succeed.

E. Index-Based Course Classification

The completed inventory for multiple course sections provides a basis for classifying a single section as free-form vs highly structured, individual vs. collective in sociality, and traditional or inclusive in terms of course climate. Once an observer or researcher completes the three indices for multiple course sections, the items selected in each index form a vector that describes the section in terms of structure, sociality, and inclusiveness. The dimensions of the vectors are the options for the items in the index, and the value for each dimension is a binary indicating presence or absence. The vectors for each index can be clustered, using an algorithm such as k-means, to identify which course sections are similar to each other in each index. The courses in each cluster can then be inspected to determine whether those sections record the presence of relatively more structure, more sociality, and more inclusive practices than the sections in other clusters. Again we emphasize that these categorizations are not meant as value judgments for the sociality and structure indices: there may be very good reasons for a course to be free-form vs. highly structured or individual vs. collective. Though there are likely few good reasons to have a course that is traditional rather than inclusive. What this categorization allows is a grouping of courses such that multi-course analysis can be conducted at a higher level, as we detail in the following section on potential uses.

IV. Potential Uses of CALI

The main purpose of developing the CALI inventory is as a research tool, to enable investigation into how introductory CS course pedagogy practices and classroom environments impact the student experience and learning. Usage for research is explained in the first two subsections below. However, the inventory has other potential uses which are detailed in subsequent subsections C-F.

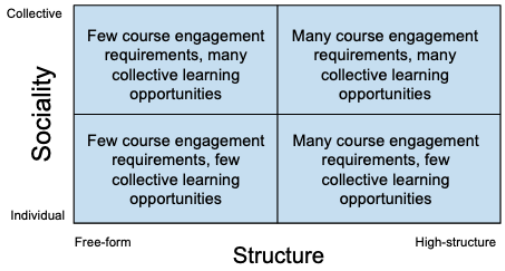

Fig. 4. A structure-sociality matrix of indexed courses could be used for analysis of sets of courses.

A. CS Education Research

Researchers in CS Education often work to determine how a particular pedagogical practice impacts student experience or learning outcomes [9]. There are so many possible interventions to study, and the intersection of different interventions, teaching approaches, mindsets, and inclusiveness practices may have differing effects. By grouping interventions and practices along the three dimensions of structure, sociality, and inclusiveness, we hope to be able to categorize large sets of courses for further analysis. The matrix in Figure 4 provides an example; once the indices are recorded for a set of courses, the courses could be clustered into categories and then researchers can look at correlations between course sections in each cluster with student success and student mindset. Similarly, researchers could look at DFW rates (% of students receiving D or F grades, or withdrawing) of the courses in each quadrant, or student engagement ratings. Of

course, other matrices could be drawn to reflect different index combinations such as sociality vs. inclusiveness, or structure vs. inclusiveness. Once data has been collected on a set of courses, various analysis trajectories can help to uncover effectiveness of structure, sociality and inclusiveness practices.

B. Intersectional Analysis

There are challenges associated with studying a set of interventions in a single course or between two sections. Because of the relatively small sample size, teasing out differential effects on demographic subgroups is difficult, and often can only be done at the level of binary gender distinctions (leaving out non-binary and transgender students) or the majority of racioethnic groups vs. 'everyone else' (leaving out groups with typically small sample sizes, such as indigenous students). Using the approach described above, data for a larger set of courses could allow a new approach to course categorization; broader analysis could then be done to understand companion data from those courses for specific sub-groups of students that typically lack a sufficient n size for intersectional analysis within a course. For example, a researcher could look at the final grades of a particular subgroup of students (e.g., Latina first-generation students) and map their performance in courses across the quadrants in Figure 4 to see what types of course environments they are most successful in, and whether that success differs from the course success of students as a whole.

C. Institutional/Departmental Analysis

A computer science unit (department, college, or school) could use the indices to generate a snapshot in time to understand how the courses in their program are being taught. This information could then be used by researchers and administrators, for example, to map student pass rates to understand how the student body in that particular institution performs in classes with higher or lower structure, or how students respond to courses that are low in inclusivity. This information could help to provide insight on what could be done to improve student experiences and performance in classes that have low pass rates. Such insights could lead to recommendations such as providing more structure or scaling back the social interaction, to meet the needs of that particular student body.

D. Undergraduate Advising

It is also possible that CALI information for a unit could be used by undergraduate advisors to guide students into sections of courses that are taught in a way that would meet their needs. For example, students with severe social anxiety may fare better in courses with low sociality, and students who are first generation college students may be directed towards sections of courses that are high in structure.

E. Course Intervention Tool

The CALI inventory could also be used as a process-oriented intervention tool. Where a class with issues (such as a high DFW rate) has been identified, the instructor could be provided with the inventory to analyze the course in a reflective manner. With the help of mentors, training, or through self-reflection, the instructor could restructure or revisit the way they teach/implement the course. After implementing changes, the inventory could be repeated, and along with student performance, feedback and instructor self-reflection, provide a way to capture course improvement progress.

F. Faculty Reflection Tool

The CALI Inventory can also be used by faculty as a reflection tool. The objective is to provide faculty with a new lens through which to consider their own teaching practice and its impact on student experiences and learning. The tool could be used in two ways: as a post-course evaluation lens to understand what worked and did not work in a completed course as well as a planning tool for a course that has yet to be taught. The Structural Index can be used to help faculty think explicitly about how much structure they currently provide or want to provide to students for a given class, and different ways to provide that structure. The Sociality Index can be used by faculty to think about how they might go about creating engagement and building community in class. The Inclusiveness Index can be used by faculty to consider the many ways that their course policies, communications and content can work together to create an inclusive and welcoming environment.

V. Pilot Test

We conducted a pilot test of a beta version of the CALI inventory in Fall 2020, using it to record information on the CS1 course at ⟨anonymized university & unit⟩. The beta version of the inventory had numeric scores associated with each dimension description, allowing calculation of a score out of 26 for each index. The CS1 course at (redacted university) is managed exclusively by the Integrated Critical Core (ICC), an academic sub-unit that includes all the instructors for core courses taught in the first two years of the Computer Science degree. This unit ensures consistency across the introductory courses in addition to promoting best practices and sharing knowledge and resources. For each course in the first two years, the ICC creates a base course in the LMS that synchronizes course specific modules including links to textbook content, activities, assessments, and descriptions of learning outcomes. That base course is then used as a template by instructors of each section of that course, with each section instructor modifying the template to reflect their specific policies and pedagogical approaches.

The beta Structure and Sociality Indices were recorded for 6 sections of the sample CS1 course in two phases. First, we recorded data by investigating the LMS setup for each section and conducting a close reading of the syllabus. The LMS setup provided much of the information needed for the Structural Index. Second, we observed classes and interviewed the section instructors to gather data that could not be gleaned from the LMS pages and syllabi. Due to the consistency of the base LMS course template, the courses scored similarly on the Structural Index.

We observed marginal differences in the Sociality Index scores, but significant differences in particular dimensions within that index, due to the way the instructors customized their individual CS1 sections. The most important takeaway from the pilot test was the conclusion that while the two Indices captured much of what we wanted, issues around equity, diversity and inclusion, including factors such as communication about growth mindset, were not being captured. Thus, this pilot test led to the development and addition of the Inclusiveness Index.

VI. Focus Group

In Fall 2021, we conducted a focus group among 5 faculty to gather feedback from faculty about the three Indices. The goal was to study how the items in the three Indices aligned with the ways that instructors talk about and think about pedagogy and inclusiveness. The faculty saw the beta version of the inventory that had points assigned to each element. We asked the faculty to indicate separately for each item on the inventory if they planned to use that item and/or if they actually did use that item. The focus group was conducted over Zoom at ⟨anonymized institution⟩ and lasted approximately 90 minutes. It was moderated by Author 2. The question prompts were:

- Were there any parts of CALI that you liked or struggled with when reflecting on your course?
- Which elements of CALI include aspects of your course design that you think are effective in creating a positive learning environment?
- Which course design elements do you think impact a positive learning environment?
- What prompts would you add that address course design?
- Can CALI help understand classroom climate or capture students' learning; why or why not?
- Which parts of CALI made you think about classroom climate or student learning?
- What are some barriers or equity gaps that exist in the way intro CS courses are designed/taught?
- How was your experience of using CALI, and what might have improved it?

A. Focus Group Results

The focus group was recorded, transcribed by Author 2, and then coded iteratively by three authors using grounded theory techniques [48]. The transcript was coded via open and selective coding [11] independently, with researchers applying qualitative verification strategies [49] to compare codes, discuss thematic labels, and derive a consensus of themes.

Themes about the instrument included confusion, use, and suggestions. Emergent themes regarding teaching practice included feeling overwhelmed, uncertainty about inclusion practices, and teaching challenges. All faculty expressed confusion about the inventory and were unclear about the meaning of some response options. For example, being asked about planned versus actual teaching strategies was a limiting choice, which prompted suggestions, which are described in the Discussion. The theme of use emerged as a concern among all the faculty, who perceived that the inventory would be used to assess their teaching performance.

> "I do believe the [inventory] is about the classroom climate.....it's about what me [sic] the instructor provides... about my teaching philosophy."

They believed that the inventory options about teaching practices were a ranked scale of ideal practices. The activity of discussing the inventory served as a reflection tool among the faculty about how they teach,with one faculty noting that it "increased my awareness of many things." The challenges they face, and support needs, also emerged. All faculty described feeling overwhelmed in meeting the expectations of teaching large courses amid a pandemic.

> "We have a bunch of things that [the students] have to do, and then we have a lot of this conversation about an inclusive environment.....and there's just no opportunity in the class."

The Inclusiveness component prompted discussion about faculty's trepidation regarding how to address diversity and equity within their courses. Each faculty member mentioned specific challenges they face in their teaching practice, such as the poor quality of the LMS and compatible tools, and the varying degrees of student attitudes about learning. Variation in the students' preparation for the courses was a major concern. One faculty member noted:

> "the background is so different, it's overwhelming."

Another faculty member made a similar comment:

> "a lot of [students] have exposure to computer science very early, and a lot of them don't."

VII. Discussion

Suggestions from the faculty focus group resulted in various clarifications to the inventory, which are reflected in the Indices presented here. Faculty were interested in the inventory functioning as a checklist and including options about class size, teaching assistant presence and type of delivery (e.g., in person, online, hybrid) to capture the context of a section. A choice option matrix was suggested where responses could distinguish between planned activity and implemented activity. To make the inventory feel less like a teacher assessment and more like a research and development tool, removal of the scoring structure was suggested. Additionally, providing suggestions for how to implement the teaching strategies was desired by the faculty, which was an unexpected discussion outcome since the purpose of the tool is not to provide pedagogical instruction.

Faculty repeatedly described being overwhelmed with the responsibilities of teaching and a need for support. They view the inventory as a potential resource.

> "There's just a lot going on. And I feel like [the inventory] it's high level, and it's sort of like how we feel like, yeah, I'm a cool teacher. Look at the cool stuff I do. Or maybe, maybe I should think about doing that. That's a good idea, but boy, I don't have time. So I'll put it in the back burner or something.

> ...it'd be cool to have this just as a list of things to consider in the beginning of every [LMS] template."

There are nuances to consider when applying the inventory, so that the intended use is clear to faculty and not viewed as a threat. As one faculty noted:

> "The [inventory] is like, are you doing quizzes? And are you doing discussions? And are you doing presentations? So how many, you know, because it would be easy for a, a really lousy teacher to waffle this and make themselves sound awesome."

In response to such concerns about the inventory being used for instructor assessment, we defined neutral terms (**individual** vs. **collective** and **free-form** vs **high-structure**) for the two ends of the structure and sociality spectra that could reduce the value-laden assessment concerns. By showing a visual such as the matrix in Figure 4 to instructors, we could show them that their course fits somewhere in the matrix, and we can describe good reasons to have courses in all four of the quadrants. For example, some economically disadvantaged students may fare better in more free-form courses due to family care-giving or job-related scheduling demands, while some introverted students or students with serious social anxiety issues may fare better in courses that are more individual than collective.

Challenges with stress around culturally relevant teaching strategies were a central concern among the majority of the faculty. This uncertainty focused both on a lack of knowledge about how to make programming content culturally relevant, as well as how best to promote equity and diversity in course policies. While there was recognition among the group that they desire understanding of what teaching practices work best for enhancing student learning outcomes, they also feel uncertain about how to address inequities in student preparation/backgrounds in computing [43]. As one faculty member commented about the Inclusiveness items:

> "I mean, I just, how do we create a diverse team without automatically, say, you know, segregating people according to race or gender or something like that... I, I really don't have an answer for that."

While we have applied more neutral terms to the ends of the sociality and structure spectra, the labels that we apply to the end of the inclusiveness spectrum are purposefully value-laden. Courses that are low in inclusiveness are labeled **traditional** and courses that are high are labeled **inclusive**. Our position is that promoting inclusive teaching practices is critical at this juncture in the history of computing education, given the numerous and serious equity issues the field is facing [22], [23], [41].

A. Limitations

We have presented v1.0 of CALI, which is a work in progress. We anticipate the inventory will evolve as we continue to study its application and efficacy as an analytical tool. In its current state, it has not been validated psychometrically. Face-validity of the items are a first step in this process, and we plan to continue capturing faculty design input. Incidentally, this iterative process of collaborating with faculty in the inventory development and use is expected to assuage their concerns about the assumed value assigned to teaching strategies that are more structured and collective.

VIII. Conclusion

We have presented the Collaborative Active Learning and Inclusiveness (CALI) inventory, version 1.0. This inventory consists of three separate indices that capture the level of structure, sociality and inclusiveness in a particular course offering. The inventory can be used for both analytical research purposes to study teaching practices that work with various student backgrounds, within institutional contexts by Computer Science unit administrators, and by faculty as a pedagogical reflection tool. We have described how the inventory evolved, and how initial feedback through piloting and focus groups led to the refinements that are present in this version. We invite researchers and colleagues at other institutions to use the inventory and share results, and anticipate that future iterations of the inventory will incorporate feedback from a broader set of constituents.

Acknowledgment

This paper is based upon work supported by the National Science Foundation under IUSE Grants 2111376 and 2111113 titled "Examining the Effects of Course Climate, Active Learning, and Intersectional Identities on Undergraduate Student Success in Computing". The study was supported by the Center for Education Innovation research at UNC Charlotte, we also acknowledge the support from The Integrated Critical Core (ICC) group at College of computing and Informatics at UNC Charlotte.

References

[1] J. Margolis, *Stuck in the Shallow End, updated edition: Education, Race, and Computing*. MIT press, 2017.

[2] C. Stephenson, A. D. Miller, C. Alvarado, L. Barker, V. Barr, T. Camp, C. Frieze, C. Lewis, E. C. Mindell, L. Limbird *et al.*, *Retention in computer science undergraduate programs in the US: Data challenges and promising interventions*. ACM, 2018.

[3] T. Camp, W. R. Adrion, B. Bizot, S. Davidson, M. Hall, S. Hambrusch, E. Walker, and S. Zweben, "Generation CS: the mixed news on diversity and the enrollment surge," *ACM Inroads*, vol. 8, no. 3, pp. 36–42, 2017.

[4] A. Tafliovich, J. Campbell, and A. Petersen, "A student perspective on prior experience in CS1," in *Proceeding of ACM SIGCSE 2013*, 2013, pp. 239–244.

[5] B. C. Wilson and S. Shrock, "Contributing to success in an introductory computer science course: a study of twelve factors," *ACM SIGCSE bulletin*, vol. 33, no. 1, pp. 184–188, 2001.

[6] J. P. Cohoon and L. A. Tychonievich, "Analysis of a CS1 approach for attracting diverse and inexperienced students to computing majors," in *Proceedings of ACM SIGCSE 2011*, 2011, pp. 165–170.

[7] S. Krause-Levy, W. G. Griswold, L. Porter, and C. Alvarado, "The relationship between sense of belonging and student outcomes in CS1 and beyond," in *Proceedings of ACM ICER 2021*, 2021, pp. 29–41.

[8] A. Salguero, W. G. Griswold, C. Alvarado, and L. Porter, "Understanding sources of student struggle in early computer science courses." New York, NY, USA: Association for Computing Machinery, 2021.

[9] A. Luxton-Reilly, I. Albluwi, B. A. Becker, M. Giannakos, A. N. Kumar, L. Ott, J. Paterson, M. J. Scott, J. Sheard, and C. Szabo, "Introductory programming: a systematic literature review," in *Companion fProceedings of ACM ITICSE 2018*, 2018, pp. 55–106.

[10] V. Pournaghshband and P. Medel, "Promoting diversity-inclusive computer science pedagogies: A multidimensional perspective," in *Proceedings of ACM ITICSE 2020*, 2020, pp. 219–224.

[11] P. Kinnunen and B. Simon, "Phenomenography and grounded theory as research methods in computing education research field," *Computer Science Education*, vol. 22, no. 2, pp. 199–218, 2012.

[12] G. Sprint and E. Fox, "Improving student study choices in CS1 with gamification and flipped classrooms," in *Proceedings of ACM SIGCSE 2020*, 2020, pp. 773–779.

[13] Y. B. Kafai and Q. Burke, "The social turn in k-12 programming: moving from computational thinking to computational participation," in *Proceeding of ACM SIGCSE 2013*, 2013, pp. 603–608.

[14] C. Latulipe, A. Rorrer, and B. Long, "Longitudinal data on flipped class effects on performance in CS1 and retention after CS1," in *Proceedings of ACM SIGCSE 2018*, 2018, pp. 411–416.

[15] E. J. Theobald, M. J. Hill, E. Tran *et al.*, "Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math," *Proceedings of the National Academy of Sciences*, vol. 117, no. 12, pp. 6476–6483, 2020.

[16] A. Salguero, J. McAuley, B. Simon, and L. Porter, "A longitudinal evaluation of a best practices CS1," in *Proceedings of ACM ICER 2020*, 2020, pp. 182–193.

[17] R. Eglash, J. E. Gilbert, and E. Foster, "Toward culturally responsive computing education," *Communications of the ACM*, vol. 56, no. 7, pp. 33–36, 2013.

[18] A. Yadav and M. K. Heath, "Breaking the code: Confronting racism in computer science through community, criticality, and citizenship," *TechTrends*, pp. 1–9, 2022.

[19] Q. Cutts, E. Cutts, S. Draper, P. O'Donnell, and P. Saffrey, "Manipulating mindset to positively influence introductory programming performance," in *Proceedings of ACM SIGCSE 2010*, 2010, pp. 431–435.

[20] D. F. Shell, L.-K. Soh, A. E. Flanigan, and M. S. Peteranetz, "Students' initial course motivation and their achievement and retention in college CS1 courses," in *Proceedings of ACM SIGCSE 2016*, 2016, pp. 639–644.

[21] K. Quille and S. Bergin, "Promoting a growth mindset in CS1: Does one size fit all? a pilot study," in *Proceedings of ACM ITICSE 2020*, 2020, pp. 12–18.

[22] A. Núñez, M. Mayhew, M. Shaheen, and L. Dahl, "Let's teach computer science majors to be good citizens - the whole world depends on it." *Edsurge*, 2021.

[23] M. C. Bottia, R. A. Mickelson, C. Jamil, K. Moniz, and L. Barry, "Factors associated with college stem participation of racially minoritized students: A synthesis of research," *Review of Educational Research*, vol. 91, no. 4, pp. 614–648, 2021.

[24] K. Sanders, J. Boustedt, A. Eckerdal, R. McCartney, and C. Zander, "Folk pedagogy: Nobody doesn't like active learning." New York, NY, USA: Association for Computing Machinery, 2017.

[25] J. F. Shaffer, "Student performance in and perceptions of a high structure undergraduate human anatomy course," *Anatomical sciences education*, vol. 9, no. 6, pp. 516–528, 2016.

[26] M. Wilton, E. Gonzalez-Niño, P. McPartlan, Z. Terner, R. E. Christoffersen, and J. H. Rothman, "Improving academic performance, belonging, and retention through increasing structure of an introductory biology course," *CBE—Life Sciences Education*, vol. 18, no. 4, p. ar53, 2019.

[27] S. Freeman, P. Pape-Lindstrom, A. Casper, S. Eddy, S. E. Alter, L. Gafney, K. M. Bonney, C. Waters, K. Cook, E. Fisher *et al.*, "Community college students rise to the challenge," *Journal of College Science Teaching*, vol. 49, no. 5, pp. 7–16, 2020.

[28] S. L. Eddy and K. A. Hogan, "Getting under the hood: How and for whom does increasing course structure work?" *CBE—Life Sciences Education*, vol. 13, no. 3, pp. 453–468, 2014.

[29] D. Gross, E. S. Pietri, G. Anderson, K. Moyano-Camihort, and M. J. Graham, "Increased preclass preparation underlies student outcome improvement in the flipped classroom," *CBE—Life Sciences Education*, vol. 14, no. 4, p. ar36, 2015.

[30] C. Latulipe and A. De Jaeger, "Comparing student experiences of collaborative learning in synchronous CS1 classes in gather. town vs. zoom," in *Proceedings of ACM SIGCSE 2022*, 2022, pp. 411–417.

[31] J. A. Van Amburgh, J. W. Devlin, J. L. Kirwin, and D. M. Qualters, "A tool for measuring active learning in the classroom," *American journal of pharmaceutical education*, vol. 71, no. 5, 2007.

[32] M. K. Smith, F. H. Jones, S. L. Gilbert, and C. E. Wieman, "The classroom observation protocol for undergraduate stem (copus): A new instrument to characterize university stem classroom practices," *CBE—Life Sciences Education*, vol. 12, no. 4, pp. 618–627, 2013.

[33] T. Boye, "The collaborative classroom in computing higher education." New York, NY, USA: Association for Computing Machinery, 2019.

[34] H. Yuan and Y. Cao, "Hybrid pair programming-a promising alternative to standard pair programming," in *Proceedings of ACM SIGCSE 2019*, 2019, pp. 1046–1052.

[35] M. Dorodchi, E. Al-Hossami, M. Nagahisarchoghaei, R. S. Diwadkar, and A. Benedict, "Teaching an undergraduate software engineering course using active learning and open source projects," in *2019 IEEE Frontiers in Education Conference (FIE)*. IEEE, 2019, pp. 1–5.

[36] C. Latulipe, N. B. Long, and C. E. Seminario, "Structuring flipped classes with lightweight teams and gamification," in *Proceedings of ACM SIGCSE 2015*, 2015, pp. 392–397.

[37] L. Porter, D. Bouvier, Q. Cutts, S. Grissom, C. Lee, R. McCartney, D. Zingaro, and B. Simon, "A multi-institutional study of peer instruction in introductory computing," in *Proceedings of ACM SIGCSE 2016*, 2016, pp. 358–363.

[38] B. W.-L. Packard, J. Solyst, A. Pai, and L. Yu, "Peer-designed active learning modules as a strategy to improve confidence and comprehension within introductory computer science," *Journal of College Science Teaching*, vol. 49, no. 5, pp. 76–83, 2020.

[39] M. Vellukunnel, P. Buffum, K. E. Boyer, J. Forbes, S. Heckman, and K. Mayer-Patel, "Deconstructing the discussion forum: Student questions and computer science learning," in *Proceedings of ACM SIGCSE 2017*, 2017, pp. 603–608.

[40] R. Deil-Amen, "Socio-academic integrative moments: Rethinking academic and social integration among two-year college students in career-related programs," *The Journal of Higher Education*, vol. 82, no. 1, pp. 54–91, 2011.

[41] R. Santo, L. A. DeLyser, J. Ahn, A. Pellicone, J. Aguiar, and S. Wortel-London, "Equity in the who, how and what of computer science education: K12 school district conceptualizations of equity in 'CS for all' initiatives," in *Proceedings of IEEE RESPECT 2019*, 2019, pp. 1–8.

[42] J. Weng and C. Murphy, "Bridging the diversity gap in computer science with a course on open source software," in *Proceedings of IEEE RESPECT 2018*. IEEE, 2018, pp. 1–4.

[43] S. Salehi, S. Cotner, and C. J. Ballen, "Variation in incoming academic preparation: Consequences for minority and first-generation students," in *IEEE Frontiers in Education 2020*. Frontiers, 2020, p. 170.

[44] K. Crenshaw, "Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory, and antiracist politics [1989]," in *Feminist legal theory*. Routledge, 2018, pp. 57–80.

[45] S. Lunn, L. Zahedi, M. Ross, and M. Ohland, "Exploration of intersectionality and computer science demographics: Understanding the historical context of shifts in participation," *ACM Trans. Comput. Educ.*, vol. 21, no. 2, mar 2021. [Online]. Available: https://doi.org/10.1145/3445985

[46] S. L. Rodriguez and K. Lehman, "Developing the next generation of diverse computer scientists: the need for enhanced, intersectional computing identity theory," *Computer Science Education*, vol. 27, no. 3-4, pp. 229–247, 2017.

[47] E. A. Canning, K. Muenks, D. J. Green, and M. C. Murphy, "STEM faculty who believe ability is fixed have larger racial achievement gaps and inspire less student motivation in their classes," *Science advances*, vol. 5, no. 2, p. eaau4734, 2019.

[48] A. Strauss and J. Corbin, *Basics of qualitative research*. Sage publications, 1990.

[49] J. M. Morse, M. Barrett, M. Mayan, K. Olson, and J. Spiers, "Verification strategies for establishing reliability and validity in qualitative research," *International Journal of Qualitative Methods*, vol. 1, no. 2, pp. 13–22, 2002.