Relationships Between Computing Students' Characteristics, Help-Seeking Approaches, and Help-Seeking Behavior in Introductory Courses and Beyond

Shao-Heng Ko
Duke University
Durham, NC, USA
shaoheng.ko@duke.edu

Matthew Zahn
North Carolina State University
Raleigh, USA
mszahn@ncsu.edu

Kristin Stephens-Martinez
Duke University
Durham, USA
ksm@cs.duke.edu

Yesenia Velasco
Duke University
Durham, USA
yvelasco@cs.duke.edu

Lina Battestilli
North Carolina State University
Raleigh, USA
lbattestilli@ncsu.edu

Sarah Heckman
North Carolina State University
Raleigh, USA
sarah_heckman@ncsu.edu

Abstract

Background. Academic help-seeking is a key metacognitive strategy that benefits post-secondary computing students' learning. Although recent works revealed rich relationships between students' characteristics and their help-seeking, most focus on a single type of help resource, a single computing context, and/or a single characteristic at a time.

Objectives. We seek to study the relationships between student characteristics and their *behavior* in course-affiliated (internal) help resources as well as their *approaches* in help resource selection/utilization. We then study whether any relationships are specific to a help resource or a course context and whether they persist when other related characteristics are controlled.

Method. We collected (1) students' help-seeking behavioral records in two course-affiliated help resources (office hours and discussion forums) from 40 offerings of eight courses across two institutions over 3.5 years and (2) students' self-reported help-seeking approaches in their preferred order of help resource usage from 18 of these offerings. We investigated relationships between these metrics and six student characteristics (gender, race, ethnicity, major, year/standing, and confidence coming into a course).

Findings. On students' behavior, we found students belonging to the gender minority and students who felt less confident coming into a course sought help from internal resources more than their peers. On students' approaches, we found students belonging to the gender minority preferred using course-affiliated resources over external resources and vice versa for students belonging to the gender majority. We found Latinx/e/a/o and non-CS major students relied on people outside of the course more than their peers, and first-year students prioritized course-affiliated resources less than their peers. We found the relationship between students' confidence and their help-seeking behavior is specific to office hours, and that both gender and confidence remain significantly related to students' usage of office hours when the other characteristic is controlled.

Implications. Our results deepen the understanding of students' use of common internal help resources and reveal relationships between student characteristics and their intent to use external resources. These insights inform computing educators' help landscape design, resource allocation, and teaching staff training.

CCS Concepts

• **Social and professional topics → Computing education**; *Gender*; *Race and ethnicity*; *Age.*

Keywords

Computing Education, Help-Seeking

ACM Reference Format:
Shao-Heng Ko, Matthew Zahn, Kristin Stephens-Martinez, Yesenia Velasco, Lina Battestilli, and Sarah Heckman. 2025. Relationships Between Computing Students' Characteristics, Help-Seeking Approaches, and Help-Seeking Behavior in Introductory Courses and Beyond. In *ACM Conference on International Computing Education Research V.1 (ICER 2025 Vol. 1), August 03–06, 2025, Charlottesville, VA, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3702652.3744214

1 Introduction

The landscape of help resources in computing education has been rapidly changing. Thanks to the widespread deployment of undergraduate TAs as support figures [28] and the ubiquitous usage of automated feedback systems [15, 31], students in modern post-secondary computing classrooms have a particularly rich landscape of available help resources to navigate [20, 21]. As not all kinds of help-seeking actions equally benefit students' long-term academic achievement [5], understanding computing students' help-seeking approaches and behavior remains a paramount task for educators.

Recent works in the computing education community revealed substantial individual differences in how students seek help from course-affiliated resources [16] (*internal* hereafter) and resources *external* to the course [42], as well as their preferred order of using available resources [17]. On the other hand, there are both qualitative and quantitative evidences suggesting that part of

ICER 2025 Vol. 1, Charlottesville, VA, USA

ACM ISBN 979-8-4007-1340-8/25/08
https://doi.org/10.1145/3702652.3744214

these differences may be related to student characteristics such as gender [4, 7, 22, 46, 47], race [38], prior experience [2, 39], year/standing [20, 42], and major [4]. While each revealing important insights, most of these works focus on a single type of help resource, a single computing instructional context, and/or a single student characteristic, leaving significant gaps in our collective understanding of help-seeking:

Gap in exploration across help resource interactions. Most works between student characteristics and help-seeking focus on one type of help resource, but do not consider students' usage, attitude, or intention to use other resources. Without taking other available resources into consideration, it is difficult to uncover whether a relationship between a student characteristic and a help-seeking behavior (e.g., "women seek help in office hours more frequently than men" [4, 7]) reflects a resource-agnostic effect (women seek help more frequently than men), a resource-specific effect (women seek help more frequently than men *only in office hours*), or a more nuanced effect (women prefer using *course-affiliated social help resources* over other resources more than men).

Lack of attention on non-introductory and non-programming contexts. The majority of relevant works are conducted in either an *introductory programming for majors* (CS1) context or in the context of a subsequent course within the introductory programming sequence. While these contexts are undoubtedly *important*, recent works have shown that computing students' help-seeking intention [42] and behavior [17] differs by the course context, and that first-year and upperclass students have different views towards the landscape of help resources [20]. Without studying other computing contexts, it is difficult to understand what parts (if any) of current knowledge is specific to an introductory context. Any insights specific to programming instruction (e.g., student seeking help on *debugging*) also do not immediately transfer to *non-programming-focused* course contexts such as discrete mathematics.

Lack of emphasis on the interplay of student characteristics. Most relevant works analyze the relationships between help-seeking and one or more characteristics separately without accounting for the interplay of student characteristics. However, it is widely known that certain characteristics (e.g., gender and prior programming experience) are interrelated in the computing context [39, 41]. For instance, Sibia *et al.* [39] found that significant relationships between prior programming experience and student behavior on discussion forums disappear when controlling for gender. Studies that emphasize one single student characteristic at a time without controlling for other confounding characteristics could misidentify relationships.

Motivated by the above gaps, we seek to expand and deepen our current understanding of the relationships between student characteristics and help-seeking, by studying computing students' help-seeking approaches towards the entire landscape of resources and behavior within internal resources across various computing course contexts, including service courses, non-programming-focused courses, and interdisplinary electives (see Section 3).

We tackle the following research questions:

RQ1. What are the relationships between computing students' (individual) characteristics and their help-seeking *behavior* in course-affiliated help resources?
RQ2. What are the relationships between computing students' (individual) characteristics and their self-reported help-seeking *approaches*, specifically in their order of preferred resource usage?
RQ3. Among all observed effects:
(RQ3a) Are any specific to a help resource?
(RQ3b) Are any dependent on the instructional context?
(RQ3c) Do these effects persist when controlling for other student characteristics?

We collected student characteristics and help-seeking records in office hours and discussion forums from 40 offerings of eight courses between 2021 and 2024 across two research-intensive institutions in the US (Table 1 in Section 3), which we used to study **RQ1** and **RQ3a-RQ3c**. In 18 of these offerings, we additionally surveyed students on their preferred order (approaches) of using all help resources available in their context, which is used to study **RQ2**.

Consistent with many existing works on gendered effects in student help-seeking behavior (**RQ1**), we found students of gender minority sought help from both internal help resources more frequently than their peers and used anonymity more often on class forums. We also observed students who felt more confident coming into a course sought less help from internal resources than their less confident peers in office hours.

On students' help-seeking approaches (**RQ2**), we discovered that students of gender minority placed more emphasis in using internal resources, while gender majority students preferred external resources more. We found Latinx/e/a/o students and non-CS majors relied on people outside of the course more than their peers. We also found that first-year students prioritized internal resources substantially less than their more senior peers, suggesting that seeking help from internal resources is a skill that needs to be explicitly acquired as students transition to post-secondary learning. Students who felt nervous coming into a course heavily preferred getting hands-on guidance from course staff in office hours, while more confident students mostly used external resources.

By cross-comparing inconsistent trends in our results from **RQ1**, we found that the negative relationship between students' confidence and their help-seeking frequency is specific to office hours that provide affective support (**RQ3a**). Additionally, the relationships between year/standing and race versus help-seeking behavior are highly dependent on course contexts (**RQ3b**). Lastly, we found that both gender and confidence remain significantly related to students' usage of office hours when the other characteristic is controlled (**RQ3c**).

2 Related Works

2.1 Help-seeking and help resource selection

Academic help-seeking is regarded as a metacognitive and self-regulated learning (SRL) strategy [13, 14, 24, 25] students utilize to aid their learning. The acts of "identifying a need to seek help" and "determining help resources" correspond to the *forethought* phase in Zimmerman's SRL model [53], and the process of "judging and reacting to the help received" match the *self-reflection* phase [45].

The seminal eight-stage model of help-seeking by Karabenick and Dembo [14] includes a "decide on whom to ask" stage, implicitly assuming all help comes from human. Makara and Karabenick [25] refines this stage with an expectancy-value model that ties a student's likelihood of seeking help from a specific help resource to (1) the accessibility of the resource and (2) the student's perceived usefulness of the resource.

Inspired by Makara and Karabenick's model, a long line of works [4, 8–10, 34, 35, 37, 45, 49] have identified many factors that influence students' help resource selection. Our recent work [20] synthesizes these factors and proposes the concept of the "help resource landscape" where all factors serve as *dimensions* with various levels of importance perceived by the students. The work also reports that men and first-year students perceive all dimensions of the help landscape as less important than their peers. Our other recent work [19] studies students' resource utilization approaches by asking students to group and rank all available help resources in their preferred order of usage (see Section 5.2 for details). Analysis on students' collective responses reveals a progression of resource *clusters* that can be interpreted using the help landscape, but that work does not take student characteristics into account.

2.2 Relationships between student characteristics and internal resources

Thanks to the availability of rich data collected by course-affiliated help platforms, there has been much effort on investigating the relationships between student characteristics and students' help-seeking behavior in internal resources. On the front of gender, multiple works have reported that women computing students seek more help compared to men both in one-on-one office hours [4, 7, 52] and in class discussion forums [12, 22, 46], and are willing to wait longer in line for getting help in office hours [6]. Although many also find women more likely to post anonymously in class forums [12, 22, 40, 46], Sharma *et al.* [38] do not find the same effect in their data and instead report race and intersectionality (between gender and race) as significant factors in the use of anonymity.

Sibia *et al.* [39] characterize the relationship between prior programming experience and the usage of anonymity in class forums through the lens of social status theory: students with less prior programming experience have lower perceived status and therefore are more likely to use anonymity to avoid being judged. They also report that this phenomenon disappears when students are subdivided into smaller groups of homoegenous prior experience levels and that the effect between prior experience and anonymity usage is moderated by gender. Our recent work [18] reports a negative relationship between prior relevant experience (including but not limited to programming experience) and students' usage of office hours. Zahn *et al.* [52] report that students with higher confidence in their CS skills utilized office hours less than their peers.

2.3 Relationships between student characteristics and external resources

Due to the difficulty in obtaining accurate help-seeking records, most prior works on computing students' utilization of external resources in post-secondary classroom contexts used self-reported frequencies [4, 9], small-scale interviews [34, 43, 51], or attitudinal surveys [42] (i.e., the *approaches* part in this study). Sibia *et al.* [40] find that women prefer seeking help within smaller groups of peers/friends over larger unofficial group chat, while men prefer the latter more. Doebling and Kazerouni [4] report that computing majors seek help from their peers more frequently than non-computing majors. Skripchuk *et al.* [42] find males to self-report as more intended to use web search and generative AI (GenAI) tools than females. Padiyath *et al.* [30] report that students with less self-efficacy initially use GenAI tools more often in an introductory Python programming course, while the effect disappears with time. Outside of classroom contexts, analyses on public forums like Stack Overflow [44] reveal that women have lower reputation and are less likely to give answers than men [27, 47].

2.4 Synthesis

Despite the prior efforts in understanding the relationships between student characteristics and help-seeking, most works in this direction focus on one type (internal or external) of help resources, one or two specific student characteristics, and one or two computing instructional contexts. Our study instead attempts to unify all these effects into the same agenda and explicitly compare between different types of resources, course contexts, and student characteristics.

3 Contexts and Student Characteristics

3.1 Courses

Table 1 summarizes our data collection across North Carolina State University (Inst1), a large-size, research-oriented public university, and Duke University (Inst2), a medium-size, research-oriented private university. Both institutions use 15-week semesters and are located in south-eastern US. Throughout this paper, we refer to the eight courses in Table 1 as *courses* and single offerings of a course as *classes* (40 in total). Our data spans from Fall 2021 (Fa21) to Fall 2024 (Fa24), although not every course was offered every semester, and not all datasets were collected in every class. Although some of the earlier semesters overlapped with the Covid-19 pandemic, all classes in the dataset were offered primarily in person.

The eight courses span the following different types of instructional contexts:

3.1.1 Inst1 introductory programming courses. Introduction to computing in MATLAB (MATLAB) is a required introductory course at Inst1 for non-CS-major engineering students delivered in the flipped learning model [1]. Object-oriented programming (OOP, taught in Java) is the second course in a three-course introductory programming sequence for Inst1 CS majors.[1] Note that the first course in the introductory sequence for computer science majors is different than MATLAB, and we did not collect data from that other introductory course.

3.1.2 Inst2 introductory programming courses. Introduction to programming in Python (PYTHON) is a service course at Inst2 for non-CS majors and also an optional entry course to the CS major. Inst2 CS majors with prior programming experience typically enter the major by directly taking Data Structures and Algorithms

[1]We refrain from calling any of these courses a "CS1" or "CS2" as our data scope spans two institutions with different CS curriculums. Furthermore, neither MATLAB nor PYTHON has a predominant CS major student body like a conventional "CS1" course.

Table 1: Overview of all classes and datasets.

Inst.	Type	Course	Student body	Has Small Sections?	Flipped?	# Consenting Students Fa21	Sp22	Fa22	Sp23	Fa23	Sp24	Fa24	Total
Inst1	Intro Programming	Matlab	Eng. majors	Yes	Yes	123	195	62	122	116	138	152	908
		OOP	CS majors	Yes	No	161	121	82	103	88	26	77	658
Inst2	Intro Programming	Python	Non-majors	Yes	Yes	**161**	**124**	**173**	159	175	**72**	56	920
		DSA	CS majors	Yes	No				238	272	174	129	813
	Non Prog.-focused	DM	CS majors	Yes	Sp24 only				**62**	94	**68**		224
		Algo	CS majors	Yes	No				202	119	220	**94**	635
	Electives	Data	Interdisciplinary	No	Yes	**142**	**140**	**149**	**159**	**66**	**115**	**62**	833
		DB	CS majors	No	No					183			183

The number in a class cell is the number of consenting students who provided some characteristic information (see Table 2). Bolded classes administered an intake survey that collected **Confidence**, while all other student characteristics (see Table 2) were collected in every class, except **Major** in Matlab and DM-Fa23. An empty class cell implies the class either was not offered or did not collect data. All help-seeking metrics for **RQ1** (see Section 4) were collected in every class, except **Requests** in Matlab-Fa21. Shaded classes surveyed students' order of preferred resource usage for **RQ2**.

Table 2: Summary of self-reported student characteristics, ordered by characteristic type.

Characteristic	Type	Groups	Analysis Scope
Gender	Binary	Man, Woman/Nonbinary	All classes
Ethnicity		Latinx/e/a/o, Non-Latinx/e/a/o	All classes
Major		CS, Non-CS	Python, DSA, and Data
Year	Ordinal	1 → 2 → 3 → 4+ → Other	All classes
Confidence		Really Nervous → ... → Really Confident (See Figure 2)	Bolded classes in Table 1
Race	Categorical	Asian, Black, White, Multiracial, Other	All classes

(DSA, taught in Java) as their first programming course, skipping Python. *Every* offering of Python therefore had more students self-reporting not intending/likely to major in CS than intending/likely to major (see Section 3.2.3).

3.1.3 Inst2 non-programming-focused courses. Discrete Math (DM) and Algorithm Design/Analysis (Algo) are two *non-programming-focused* required courses for Inst2 CS majors, forming a prerequisite chain with DSA (DSA→DM→Algo). No DM class in the scope of this study had any mandatory programming assignments. The Algo classes in the scope of this study had regular *hybrid* assignments, each typically consisting of three theory problems and one programming problem, with only the three highest-scoring problems contributing to the final grade. This design allows flexibility such that a student can achieve the highest grade in each assignment without writing any programs. Furthermore, Algo's course policy limits the help provided by course staff to conceptual discussions on algorithm design (i.e., the course does not provide official programming help such as implementation or debugging).

3.1.4 Inst2 elective courses. Data Science (Data) and Database Systems (DB) are two popular elective courses at Inst2. Unlike all other courses in our data scope, these two courses do not have any small-scale, TA-instructed sections as part of regular instruction (see Table 1). Data has an interdisciplinary student body (only 53% CS majors, see Table 3) and uses Python, with a heavy focus on the application of data analysis packages (such as `scipy` and `scikit-learn`) rather than programming. DB is taken almost exclusively by CS majors (97%) and emphasizes database systems design. Both courses have a semester-long open-ended project.

3.2 Student Characteristics

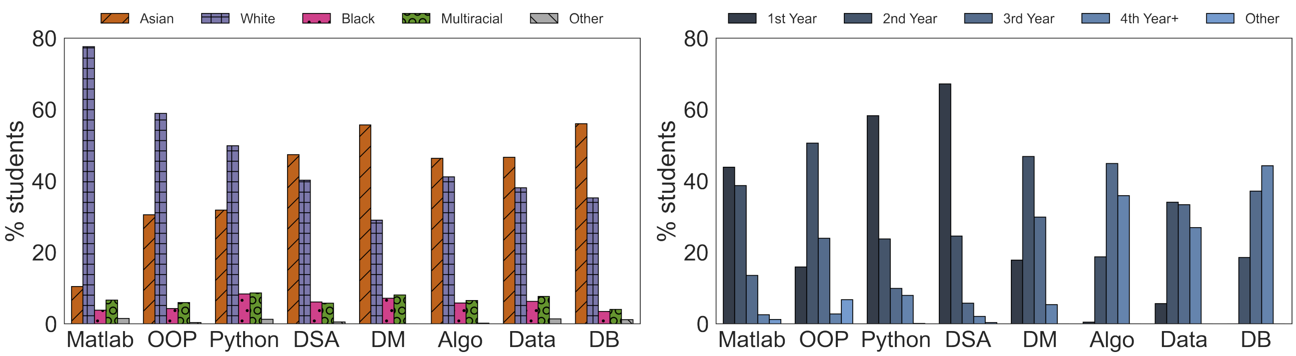
This study includes six self-reported student characteristics (see Table 2 for a summary) that we introduce below. We note that our scope of student characteristics do *not* include prior (programming) experience despite it being shown as an important factor in students' help-seeking [2, 18, 39]. This is because relevant types of prior experience depend on course context, e.g., prior programming experience may not be as relevant in non-programming-focused courses [18]. In contrast, all other characteristics hold the same meaning in any context. We revisit this choice in Section 7.

3.2.1 ***Gender****.* All but two classes (see Section 3.2.6) collected gender data via a single selection question. In addition to the *Woman/Man* or *Female/Male* options, each class included one or more additional choices, including *Nonbinary*, *Prefer not to specify*, and *Other:/I identify as:*, followed by an open text field. Since there were only a few students in each course that answered *Nonbinary*, *Other*, or gave a text response, we treated *Man/Male* as Man, all other valid responses as Woman/Nonbinary (W/NB), and *Prefer not to specify* as NA (no data) in this study (see Table 3). The decision to combine women and nonbinary was based on past works that found both groups to have lower self-efficacy compared to men in computing contexts [29].

Table 3: Distribution of Gender, Ethnicity, and Major per course.

Inst.	Course	Gender: Woman/Nonbinary	Gender: Man	Ethnicity: Latinx/e/a/o	Ethnicity: Non-Latinx/e/a/o	Major: Non-CS	Major: CS
Inst1	Matlab	212 (23.9%)	676 (76.1%)	54 (6.2%)	811 (93.8%)	(service course)	
	OOP	186 (29.2%)	451 (70.8%)	34 (5.4%)	601 (94.6%)	115 (19.8%)	466 (80.2%)
Inst2	Python	512 (62.1%)	312 (37.9%)	119 (14.4%)	705 (85.6%)	531 (71.9%)	208 (28.1%)
	DSA	355 (44.2%)	448 (55.8%)	86 (10.8%)	712 (89.2%)	461 (56.7%)	352 (80.2%)
	DM	108 (48.2%)	116 (51.8%)	27 (12.2%)	195 (87.8%)	18 (13.8%)	112 (86.2%)
	Algo	224 (36.4%)	391 (63.6%)	68 (10.9%)	558 (89.1%)	31 (4.9%)	604 (95.1%)
	Data	373 (45.0%)	455 (55.0%)	88 (10.6%)	740 (89.3%)	389 (46.7%)	444 (53.3%)
	DB	78 (42.9%)	104 (57.1%)	21 (11.6%)	160 (88.4%)	6 (3.3%)	177 (96.7%)

The Python distribution of **Major** shows students' *intent* to major in CS and excludes undecided students, and only shaded courses are analyzed for **Major** (see Section 3.2.3). The proportions are out of students who are not NAs for each characteristic, not out of all students in the data.



(a) Race **(b) Year**

Figure 1: Bar plots of Race and Year, aggregated by course.

3.2.2 ***Race** and **Ethnicity***. All but two classes (see Section 3.2.6) had a select-all-that-apply race question where options varied and sometimes included an "*I identify as:*" open text field. We categorized the options into five groups:[2] Asian, White, Black, Multiracial, and Other (see Figure 1a). Most Inst2 classes had a separate question asking *"Do you identify as Latinx?"* or *"Are you Spanish, Hispanic, or Latinx, or none of these?"*. We categorized each student as Latinx/e/a/o if they answered *Yes* to this question, Non-Latinx/e/a/o if they answered *No*/*None of these*, and NA if they answered *Prefer not to disclose/respond*. The other classes all had a *Latino/a/x*, *Hispanic*, or *Hispanic of any race* option in the race question; we coded students as Latinx/e/a/o if they selected these options and Non-Latinx/e/a/o otherwise.[3]

3.2.3 ***Major***. Most classes (see caption of Table 1) tracked majors, whereas Python instead tracked *Intended Majors*/*Likely Majors* as its student body consisted of many undeclared first-years. For Python, we coded students as CS if they reported that they intended/were likely to major in CS, Non-CS if they did not intend/were not likely to major in CS, and NA if they responded by "*Undecided*". For all other courses (except Matlab which is a service course for engineering majors), students were coded as CS if they indicated CS as one of their major(s), Non-CS if they had exclusively non-CS majors, and NA if they had not declared any majors or chose not to respond. Note that students who intended to major in CS but had not declared the major would be classified as Non-CS or NA in these courses.

As shown in Table 3, only Python, DSA, and Data have at least 25% of Non-CS students. We excluded other courses from all

[2]Asian includes the *Asian or Asian-American* option at Inst1 and the *Asian* or *Asian (including Southeast Asian, etc.)* options at Inst2; Black is worded as *Black or African American* at both institutions, with the "or" sometimes replaced by a slash; White is worded as *Non-Hispanic White* at Inst1, just *White* or *White/Caucasian* at Inst2. All students that either selected an existing *Other* option or gave a single text response are classified as Other. Students that (1) selected multiple options, (2) selected an option and added a text response, or (3) gave a text response that contains a comma, a semicolon, or the keywords "*and*", "*mixed*", "*multiracial*", or "*biracial*" are classified as Multiracial.

[3]Unless they answered *Prefer not to disclose/respond* in **Race**, in which case both their **Race** and **Ethnicity** are coded as NA. Furthermore, since the Latinx/e/a/o options were dropped from the **Race** parsing, a student that only selected *Latino/a/x* or *Hispanic* (and nothing else) is coded as NA in **Race**, while a student that selected both *Hispanic* and *White* is coded as White instead of Multiracial. Finally, we acknowledge our analysis does not differentiate between pan-ethnic labels such as *Hispanic*, *Latino*, and *Latinx*, mainly due to inconsistencies in the data collection.

subsequent analyses on the **Major** characteristic to avoid drawing conclusions on small Non-CS subpopulations.

3.2.4 ***Year***. Most classes asked either "*What is your current standing at [Inst1/Inst2]?*" (where options were *Freshman/First-Year*, *Sophomore*, *Junior*, *Senior*, *Graduate Student*, *Continuing Education*, etc.) or "*What year do you expect to graduate?*" (where options were calendar years). We consolidated the responses into 1 (first-year), 2 (second-year), 3 (third-year), 4+ (all other undergraduate students), and Other (all non-undergraduate) groups (see Figure 1b). Two classes (Matlab-Fa24 and OOP-Fa24) only collected students' age. For these classes, we treated all age-18, 19, 20, 21, and 22 (or older) students as 1st, 2nd, 3rd, 4th, and Other years, respectively; we revisit this choice in Section 7. Students under the age of 18 were excluded from this study.

3.2.5 ***Confidence***. Some Inst2 classes (bolded cells in Table 1) administered an entry survey during the first week of the semester. This survey had a question asking "*How do you feel about taking [course]?*" with Likert options ranging from *Really Nervous* to *Really Confident*. We plot the distributions of this characteristic in Figure 2.

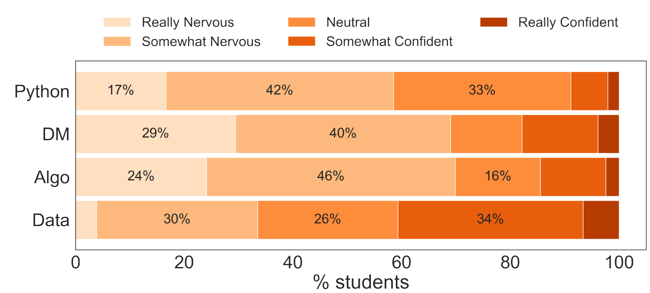


Figure 2: Stacked bar plots of Confidence.

3.2.6 Missing and duplicated data. Python-Fa22 did not collect **Gender**, **Race**, and **Ethnicity**. However, 86 of the 173 students in that offering (49.7%) self-reported these characteristics in other classes in our data. We opted to infer these three characteristics for these students while acknowledging the associated selection bias towards CS majors in doing so. DSA-Sp24 and Algo-Fa24 collected characteristics twice at different timepoints in the semester. For these courses, we merged the two datasets together and kept the earlier response whenever a mismatch happens.[4] Algo-Fa24's **Gender** and **Race** questions were an open text field, so we manually coded student responses into our two gender categories and five racial categories (see 3.2.1 and 3.2.2).

4 Help-seeking Data Collection and Methodology

Throughout the paper, we use the term *metric* to refer to help-seeking behavioral measures listed below in Section 4.1.

[4]Most mismatches happened in **Major**, which can be attributed to some students declaring (or switching) their major in between the surveys. There were only two other mismatches, both in **Year**. Since we inferred **Year** from students' self-reported *expected graduation year*, it is likely that these two students updated their expected graduation year in between the surveys due to some unknown reason.

4.1 Help-seeking behavioral data in internal resources (RQ1, RQ3a, and RQ3b)

4.1.1 Class forums. Every class utilized a discussion forum for students to seek asynchronous help from course staff (and sometimes from fellow students). Inst1 classes used the Piazza [33] platform up to Sp24, then switched to Ed Discussion [3] in Fa24; Inst2 classes used Ed Discussion throughout. Both platforms tracked the total number of posts/comments/responses each student read (**Reads**) [17, 39, 40],[5] and the number of questions/posts they initiated across the semester (**Posts**).[6] Both platforms allowed students to post a thread anonymous to their classmates, while most classes also let students post a thread *privately* (i.e., viewable to staff members but not classmates). For all Piazza datasets, we follow Sharma *et al.* [38] by inferring a thread as "private" if the number of unique users who viewed the thread at least once does not exceed the number of teaching staff members in that class. This attribute is available in Ed Discussion and does not need to be inferred.

Following Sharma *et al.* [38], we calculated a student's *anonymous rate* (**Anon%**) as the proportion of their *public* threads that are anonymous, excluding private threads, since these are not visible to peers regardless of anonymity settings.[7] Similarly, we define a student's *private rate* (**Priv%**) as the proportion of their threads that are private. We excluded DM-Sp24 data as private threads were typically not allowed in that class. Although these normalizations prevent overrepresenting frequent forum users, we note that **Priv%** excludes all students who never posted a thread (*non-writers*), while **Anon%** excludes all students who never posted publicly.

4.1.2 Consulting hours. Throughout the paper, we use *consulting hours* (CH) to refer to a designated time/place (which could be an online platform) for students to consult a teaching staff (an instructor, a graduate TA, or an undergraduate TA) on a *one-on-one* basis. All classes (except Matlab-Fa21, for which data was not collected) hosted regular CH several days per week via a queueing app. Following existing works [7, 16–18], for each student in every class, we define **Requests** as the number of *valid* interactions between the student and a teaching staff. An interaction is deemed valid if and only if (1) the wait time does not exceed four hours (the longest CH shift any class in our dataset offered) and (2) the interaction duration is less than 60 minutes. Although interactions with duration over an hour do happen occasionally, these records often reflect human errors, such as teaching staff forgetting to close the queue after each day's shift.

[5]The platforms also tracked the number of days each student logged into the forum during the semester. As a preliminary analysis, we ran a Spearman's ρ rank correlation analysis between that and **Reads** for each course. Results revealed strong rank correlation ($\rho = 0.65$) in OOP and extremely strong rank correlations ($\rho = 0.87$ - 0.92) in other courses. As such, we concluded that either metric encodes similar information under the nonparametric regime, thus only one needs to be studied. We thus analyze only **Reads**.

[6]Our earlier work [17] studied an alternative metric that includes the number of student self-comments on their own threads. We however found extremely strong rank correlations ($\rho \geq 0.93$ in all courses) between that definition and the one in the current study, so the choice of metric likely does not matter.

[7]Sibia *et al.* [39, 40] found most students either stick to the feature or never use them, suggesting that a binary metric could potentially be used. We are unable to follow that approach: in every Inst2 course, at least 92% of the students with at least one public thread have used anonymity at least once.

4.2 Statistical tests for RQ1, RQ3a, and RQ3b

In **RQ1**, we seek to understand how each characteristic is individually related to each metric, following many existing works [17, 18, 23, 38, 39, 46, 47, 52] by using exclusively non-parametric statistical methods throughout the paper.[8] This enables using the same techniques for numerical and ordinal metrics, as all methods only concern the relative order but not the absolute values of numerical statistics. As such, we used Mann-Whitney U tests to examine distributional differences in help-seeking metrics between two subpopulations. For **Race**, we used Kruskal-Wallis H tests to account for more than two population groups and the Dunn's posthoc test (with Holm-Bonferroni correction) for significant H test results. Finally, we used the Jonckheere-Terpstra test [11] for ordered differences for ordinal characteristics (**Year** and **Confidence**) such that the hypotheses preserve the ordered nature of population groups. For example, the alternative hypothesis between the **Year** and **Reads** would mean "the more senior a student is, the more (or alternatively, the fewer) posts they read on class forums".

To ensure a focused analysis of instructional contexts, no statistical test in this study draws from multiple courses; we revisit this choice in Section 7. For each (characteristic, metric) tuple, we conduct a test for each *course* (combining all class offerings of the same course), followed by a Holm-Bonferroni correction across all (up to 8) tests to control the overall type-I error rate at the 0.05 level. We chose to control the type-I error this way such that our results can be directly compared with past works that most focused on one specific characteristic and one specific help-seeking metric. After obtaining all individual results for **RQ1**, we compare the results across different help resources to answer **RQ3a** and across different instructional contexts to answer **RQ3b**. Finally, when two or more student characteristics are both significantly related to a help-seeking metric, we build a regression model suitable for that metric taking all characteristics as inputs to answer **RQ3c** (Section 5.4).

For all post-correction significant results, we report the common language effect size (CLES) as a unifying measure. For U tests and Dunn's post-hoc results, the CLES is the proportion of pairs (of students) in which the student in the first subgroup has a higher value of the metric. Similarly, for the Jonckheere-Terpstra test, the CLES is the proportion of pairs (of students with different levels of the characteristic) in which the student with the higher characteristic level has a higher metric (if the effect is positive) and vice versa. Note that in a U test with population sizes N_1 and N_2, the U statistic is defined as $U = \min\{U_1, U_2 = N_1 N_2 - U_1\}$, and CLES is defined as $U_1/(N_1 N_2)$. As such, given the population sizes N_1 and N_2 (which can be found in Table 3), CLES contains the same information encoded in the conventionally-reported U statistic while being less sensitive to differences in population sizes.

[8]Although these works have already documented the non-normality of help-seeking metrics, we still ran a Shapiro-Wilk normality test *per metric per course*. All tests rejected the normality hypothesis, with the largest p-value (for the **Priv%** distribution in DB) among these tests being 1.1×10^{-10}.

4.3 Help-seeking approach data via order of preferred help resource usage (RQ2)

In **RQ2**, we seek to identify the relationships between students' (individual) characteristics and their order of preferred help resource usage. In the latest few offerings of each course (18 classes in total; see Table 1), we administered the *order of preferred help resource usage* survey [19] (hereafter referred to as Order) either as part of the class's mid-semester survey or as part of a project reflection survey. This survey asks students to group and rank all available help resources in their preferred order of usage, where the descriptions of available help resources slightly varied and depended on each class' context (see caption of Figure 3 for a unified list). Students were instructed to describe their preferred help-seeking process by grouping available resources based on the order in which they would use them. They placed the resources they would seek first (assuming availability) in Group 1, then listed those they would turn to if the first group was unavailable or insufficient in Group 2, and continued this process for up to three groups. Resources they did not consider useful or never used could be omitted.

The Order survey treats every student equally, aiming at understanding students' self-reported help-seeking approaches (i.e., what each student would do should they feel they need help), regardless of their actual help-seeking behavior in either internal or external resources. Therefore, data from this survey is not inherently biased towards students who mainly utilize internal resources, nor does it exclude students who do not seek *any* help [19].

4.4 Analysis of the Order survey

Following our recent work that used the same survey (but does not analyze student characteristics) [19], we use the Borda Score as an aggregate metric to describe a group of students' collective preferences towards a help resource. Each response assigns a Borda Score of 3, 2, 1, or 0 to a resource based on whether the student placed it in their *used first*, *used second*, *used third*, or *omitted (would never use)* category, respectively. The Borda Score for a resource within a student group is then calculated as the mean Borda Score across all students in that group. This allows for easy *inter-resource* comparison within a student subgroup, e.g., comparing women's collective preferences towards reading class forum posts against their collective preferences towards posting new questions on class forums. Throughout this study, we do not perform individual-level analyses on Borda Scores, and instead aggregate the Borda Scores *per class* with respect to each offering's different help resource contexts (see Section 7 for more discussions).

5 Results

5.1 RQ1: relationships between individual characteristics and behavior in internal help resources

Table 4 summarizes the results of all statistical tests (see Section 4.2) between each characteristic and each behavioral metric across all 8 courses. In this section, we focus on describing trends across different course contexts at a high level; we defer all resource-specific (**RQ3a**) and context-specific (**RQ3b**) discussions to Section 5.3.

Table 4: Summary of RQ1 results at the (characteristic, metric) level.

Characteristic	# Courses Analyzed	Favored Group (+)	Detailed Results	**Reads** +	**Reads** −	**Posts** +	**Posts** −	**Anon%** +	**Anon%** −	**Priv%** +	**Priv%** −	**Requests** +	**Requests** −
Gender	8	Woman/Nonbinary	Table 5	6	.	7	.	3	.	.	.	8	.
Major	3	CS	Table 6	.	2	.	.	.	.	.	.	.	.
Year	8		Table 7	3	1	2	.	.	.	.	.	2	.
Confidence	4		Table 8	.	.	.	1	.	.	.	.	.	4
Race	8		Table 9	1*		.		.		.		3*	

Each cell summarizes the results of up to 8 statistic tests, one per course that collected the characteristic (see the # Courses Analyzed column). Holm-Bonferroni correction was applied separately per cell. The number under +/− indicates the number of courses with post-correction significant *positive/negative* effects where the favored group (for binary characteristics) or higher groups (for ordinal groups) had higher/lower metric values. All zeros are replaced by dots (.) for brevity. For **Race**, the numbers indicate the count of courses with post-correction significant H test results in which two different racial groups had significantly different metrics (see Table 9 for details). The **Ethnicity** row only contains zeros and thus is omitted.

Gendered effects on help-seeking frequencies are not specific to introductory programming contexts. Consistent with many existing works in computing education, we found Woman/Nonbinary students both read [22] and posted [12, 22, 46] more frequently on the class forum in most courses (6/8 and 7/8 of the courses, respectively; see Table 5 for details), as well as utilized CH more frequently in *all* courses [4, 7]. We found Woman/Nonbinary students used anonymity significantly more often in only 3 of the 8 courses; we discuss this phenomenon in Section 5.3.

CS majors sought help less frequently than non-majors from class forums. The only post-correction significant effects involving **Major** saw CS majors having fewer **Reads** on the forum in two of the three courses that analyzed **Major** (Python and DSA, see Table 6). Furthermore, *all* effects on **Reads**, **Posts**, and **Anon%** are *negative* (see Table 6), implying CS majors sought less help on class forums and used anonymous posts less than their Non-CS peers.

Year/standing and racial effects do exist but are highly context dependent. We found contradictory significant effects between **Year** and **Reads** (i.e., we observed significant effects *in both directions* in different courses; see Table 7 for details), implying these effects are highly dependent on course context and may be influenced by confounding factors such as potential correlations between year and prior background/experience. Similarly, we found contradictory significant effects between Asian and White students' **Requests** (in OOP and DSA; Table 9).

Students who felt more confident coming into a course sought less help from internal resources. We observed *negative* effects between **Confidence** and **Posts** in Python as well as between **Confidence** and **Requests** in *all four courses that collected **Confidence***. On the surface, this may be attributed to more confident students typically having a stronger prior background, reducing their need for help. However, the substantially stronger negative effects in CH **Requests** compared to any class-forum related metrics highlight the distinct roles these resources serve, which we further discuss in Section 5.3.

Non-effects. We observed no significant effects involving the **Priv%** metric and the **Ethnicity** characteristic. Furthermore, aside from the gendered effect that we revisit in Section 5.3, there were no other significant effects observed for **Anon%**.

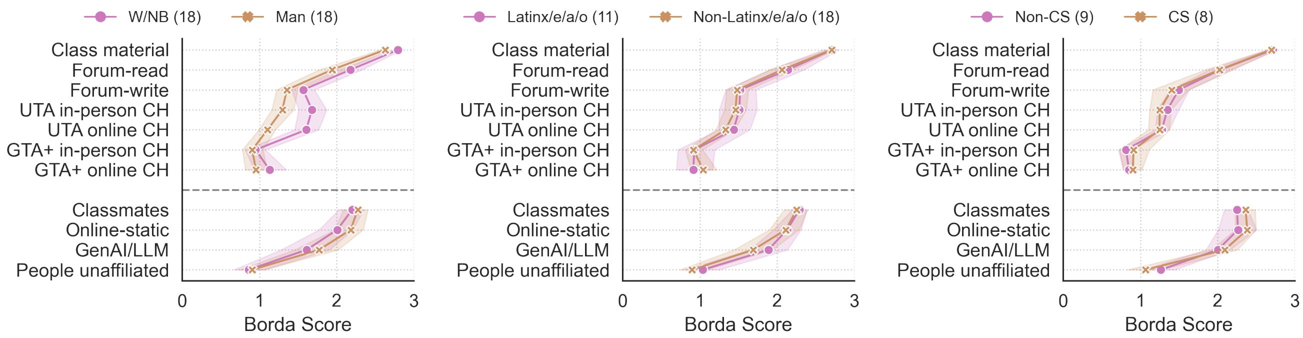
5.2 RQ2: relationships between individual characteristics and students' order of preferred resource usage

Figure 3 plots the Borda Score distributions against each characteristic, where each observation is the mean Borda Score of all students within a subpopulation in a given class. To avoid noisy aggregates from small sample sizes, we excluded classes with less than 10 students in a subpopulation. Therefore, we report no aggregate values for Other in Figures 3d and 3f as well as for *Really Confident* in Figure 3e, since no class had 10 students in those groups.

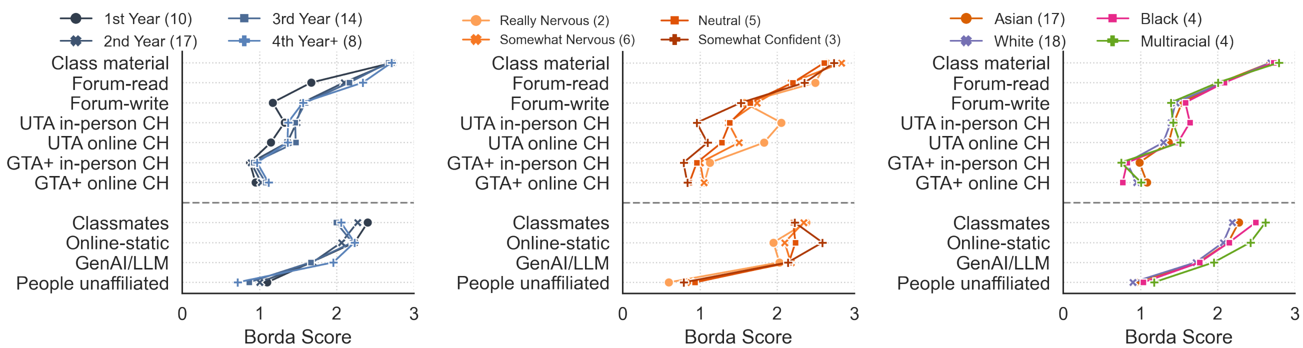
Gendered effects in resource usage order. Figure 3a reveals substantial differences between Man and Woman/Nonbinary's aggregate Borda Scores. Notably, Woman/Nonbinary had higher Borda Scores than Man in *all internal resources* (resources above the dashed line in Figure 3a), whereas *all external resources* (resources below the line) saw Man having the higher Borda Scores. At a high level, this phenomenon aligns with the gendered effects in internal resources found in **Reads**, **Posts**, and **Requests** in **RQ1**,[9] and also corroborates with the recent finding by Skripchuk *et al.* [42] that Man have higher intentions to use both Static Online Resources and GenAI/LLMs than Woman/Nonbinary. Furthermore, Woman/Nonbinary distinguish UTA CH and GTA+ CH to a substantially higher extent than Man do. Since Woman/Nonbinary are more likely to feel intimidated and judged by more authoritative figures in social help-seeking experiences [2], they may gravitate towards UTA CH which are more approachable as near-peers.

Minority groups rely on People Unaffiliated with class more than their peers. For both **Ethnicity** and **Major**, minority groups (Latinx/e/a/o and Non-CS) had a higher Borda Score for People Unaffiliated than their respective majority group. Non-CS students also had slightly lower Borda Score for Classmates than CS major, which is consistent with Doebling and Kazerouni's work [4] that reports that computing majors seek help from their peers more often than non-computing majors. They hypothesize that this phenomenon is due to computing majors knowing more peers than non-computing majors. A similar argument applies to our observed

[9] We note that the metrics in **RQ1** measure students' help-seeking *frequency* while the Borda Score measures students' preferences (i.e., relative ordering) of the resources, which are related but separate concepts. Therefore, our analyses in **RQ1** and **RQ2** are *not* to be directly compared with each other.



(a) Gender. **(b) Ethnicity.** **(c) Major.**



(d) Year. **(e) Confidence.** **(f) Race.**

Figure 3: Class-level Borda Score by characterstic. The number after each subpopulation is the number of observations, i.e., classes with at least 10 students in that subpopulation. Each subfigure plots internal resources at the top and external resources at the bottom. Within each type, resources are listed in overall decreasing order of median Borda Score, i.e., from resources used earlier to those used later. Forum-read and Forum-write stand for *reading* existing posts and *writing* new posts on the discussion forum, respectively. UTA CH stands for undergraduate TA consulting hours. GTA+ CH stands for graduate TAs/instructors consulting hours. Static Online Resources (Online-static) include all web-based external resources that are non-interactive (also called *web search* in related works [42, 43]), and GenAI/LLMs include all interactive external resources. The shaded regions in the Gender, Ethnicity, and Major subfigures represent confidence intervals obtained by bootstrapping. We omitted confidence intervals for characteristics with more subgroups for brevity.

phenomenon on People Unaffiliated: minority groups might have *less* access to Classmates at the same institution and as a result resort to seeking social help from People Unaffiliated more often.

First-year students have different help-seeking approaches. Figure 3d highlights that first-year students' Borda Score progression is drastically different than other year groups. Compared to other year groups, first-year students prioritized Classmates and People Unaffiliated more, while their Borda Scores in internal social resources all rank last among the four year groups. Consistent with a recent work [20] that finds first-year students to collectively have a different perception of the help resource landscape, our results show that first-year students do not prioritize seeking help from course-affiliated internal help resources as much as their more senior classmates, suggesting that help-seeking from internal resources (or mere awareness of the resources) is a skill that needs to be explicitly acquired.

Students' confidence is related to their help-seeking approaches. As shown in Figure 3e,[10] students who were really nervous coming into the course had the lowest Borda Score for most external resources and the highest Borda Score for most internal social resources. Considering that more nervous students often have lower self-efficacy, lower self-assessment, and less prior relevant experience, this phenomenon aligns with existing works that reported that less experienced students struggle with using external resources effectively [2, 43, 50] as well as effectively describing their own context on class forums [26, 48, 50]. These struggling students then heavily rely on getting hands-on guidance from course staff in CH. Conversely, students who are somewhat confident exhibit the opposite approach: they effectively use Static Online Resources and seldom need hands-on help from course staff.

[10]Although the number of observations is quite small, we note that each observation is itself the mean over at least 10 students.

5.3 RQ3a and RQ3b: resource- and context-specific effects

In this subsection, we report detailed, by-course results for all statistical analyses, with an emphasis on cross comparing the effects by resource (**RQ3a**) and by instructional context (**RQ3b**). We excluded **Ethnicity** from this subsection as there were no significant results involving **Ethnicity**.

Table 5: Summary of course-level Gender results.

Inst.	Type	Course	Reads	Posts	Anon%	Priv%	Requests
Inst1	Intro Prog.	MATLAB	+	+.54*	+	+	+.59***
		OOP	+.58**	+.59**	+.64***	+	+.62***
Inst2	Intro Prog.	PYTHON	+.60***	+.60***	+	+	+.60***
		DSA	+.61***	+.61***	+.55*	+	+.69***
	Non Prog.	DM	+.62**	+.62**	+	+	+.69***
		ALGO	+	+	+	−	+.65***
	Electives	DATA	+.56**	+.62***	+	−	+.60***
		DB	+.62*	+.64**	+.61*	+	+.66***

Plus (+) and minus (−) signs indicate the effect directions (regardless of significance), i.e., Woman/Nonbinary students (favored group) are in aggregate more likely to have a higher (+) or lower (−) value of that metric than Man students. For brevity, we report the CLES value only for results significant at the 0.05 level post Holm-Bonferroni correction. Results significant at the 0.05, 0.01, and 0.001 level post Holm-Bonferroni correction are followed by one (*), two (**), and three (***) asterisks, respectively. We follow these reporting conventions in the subsequent Tables 6, 7, 8, and 9.

Gendered effects in usage of anonymity. Table 5 reports all course-level **Gender** results. Although Woman/Nonbinary students only have a significantly higher **Anon%** in OOP, DSA, and DB, the effect direction is consistently positive across all eight courses. As the three courses with significant results range from introductory to advanced (see Table 1) and vary widely in gender composition (Table 3), we did not see a clear relation between instructional contexts and this effect. Nevertheless, our results (all effects pointing in the same direction, but some are not significant) align with what past works have collectively reported, as some past works found a positive effect [12, 22, 40, 46] but some did not [38]. As the high usage of anonymity by Woman/Nonbinary students is found in computing and STEM but not in other fields [36, 38], the gendered effect in the usage of anonymity is likely not solely a gender-driven effect and is instead interconnected with gaps in gender parity [22, 46, 47], prior experience [18, 39], and self-efficacy [29] in computing.

Table 6: Summary of course-level Major results.

Inst.	Type	Course	Reads	Posts	Anon%	Priv%	Requests
Inst2	Intro Prog.	PYTHON	−.43*	−	−	+	−
		DSA	−.44*	−	−	+	+
	Electives	DATA	−	−	−	−	+

Plus (+) and minus (−) signs indicate the effect directions, i.e., CS majors (favored group) are in aggregate more likely to have a higher (+) or lower (−) value of that metric than Non-CS students. See Table 5 for reporting conventions.

Major effects in forum behavior. Table 6 reports the course-level **Major** results. As mentioned in Section 5.1, all three courses we analyzed for **Major** (see Section 3.2.3) saw consistent *negative* effects between being a CS major versus **Reads** (significant for PYTHON and DSA, but not DATA), **Posts**, and **Anon%**. We note that DATA (the only course among the three in which the effect for **Reads** was insignificant) regularly used its class forum for live discussion during class meetings. This required all students (that attended class) to use the class forum and might have diluted the existing effect.

Table 7: Summary of course-level Year results.

Inst.	Type	Course	Reads	Posts	Anon%	Priv%	Requests
Inst1	Intro Prog.	MATLAB	−	−	−	−	−
		OOP	+.55**	+.55*	−	+	−
Inst2	Intro Prog.	PYTHON	−	+	+	−	−
		DSA	+.60***	+.57***	−	+	+.56**
	Non Prog.	DM	+.59**	+	+	−	+.64***
		ALGO	−.57***	−	−	−	−
	Electives	DATA	−	−	−	−	−
		DB	−	−	+	+	−

Plus (+) and minus (−) signs indicate the effect directions, i.e., higher-level students are in aggregate more likely to have a higher (+) or lower (−) value of that metric than lower-level students. See Table 5 for reporting conventions.

*Relationship between **Year** and help-seeking behavior is context dependent.* Table 7 reports all course-level **Year** results. We observe contradictory effect directions for *every* metric, which implies the effects **Year** has on help-seeking behavior is highly course-context dependent. Students' **Year** distribution in DSA is highly correlated with their prior programming experience: as mentioned in Section 3.1, Inst2 students could take DSA in their first semester at Inst2 *only if* they have prior programming experience and otherwise had to start in PYTHON. As such, students with higher **Year** in DSA are more likely to have lower prior programming experience and have more frequent help needs [18], resulting in *positive* effects between **Year** and all three frequency-based metrics (**Reads**, **Posts**, and **Requests**) in DSA.

We however also observed a *negative* significant effect between **Year** and **Reads** in ALGO, i.e., in the opposite direction of the positive effects in introductory courses. Upon further inspection, we found that in both ALGO and DB (the two highest-level courses in our data; see Figure 1b), the 4+ group (undergrad fourth years and above) had the *lowest* mean among the 2, 3, and 4+ groups *in all three frequency-based metrics*. This further indicates the **Year** effects are highly dependent on the specific course context and may not simply follow a unified direction.

Table 8: Summary of course-level Confidence results.

Inst.	Type	Course	Reads	Posts	Anon%	Priv%	Requests
Inst2	Intro Prog.	PYTHON	+	−.54*	−	−	−.58***
	Non Prog.	DM	−	−	−	−	−.64***
		ALGO	−	−	+	−	−.61**
	Electives	DATA	−	−	−	+	−.55***

Plus (+) and minus (−) signs indicate the effect directions, i.e., students who are more confident before taking the course are in aggregate more likely to have a higher (+) or lower (−) value of that metric than students who are more nervous. See Table 5 for reporting conventions.

*Effects of **Confidence** are resource-dependent.* Table 8 reports all course-level **Confidence** results in the courses that collected this characteristic. Consistent with our findings in **RQ2**, students who were less confident (more nervous) coming into the course actively sought more social help from course staff (in both **Posts** and **Requests**). However, this effect is highly significant in CH (all four effects for **Requests** being significant at the 0.01 or lower level even after correction) and mostly insignificant in class forums (only **Posts** in Python is significant). This again indicates that CH, being a more intimate, adaptive, and media-rich [19] resource than the class forum, provides more support to novices that struggle to phrase the need and context [50], exposes them less to judgment from more experienced peers [2], and thereby acts as a "safe place" that promotes students' sense of belonging [32].

Table 9: Summary of course-level Race post-hoc results.

Inst.	Course	Metric	Favored	Unfavored	CLES	Dunn's p (corrected)
Inst1	OOP	**Requests**	Asian	White	+.62	9.4×10^{-7}
Inst2	DSA	**Requests**	White	Asian	+.57	9.4×10^{-4}
			Black	Asian	+.61	3.8×10^{-3}
	DM	**Requests**	White	Asian	+.63	1.8×10^{-3}
	Data	**Reads**	Asian	White	+.58	9.8×10^{-4}

For brevity, we only report post-hoc results significant at the $\alpha = 0.05$ level post Holm-Bonferroni correction. Note that the Dunn's results are Holm-Bonferroni corrected across all pairs of racial groups within each H test independent to the Holm-Bonferroni correction performed in Table 4. When a Dunn's result is deemed significant as $p \leq \frac{\alpha}{k}$ during Holm-Bonferroni with multiplier k, we report $(p \cdot k)$ as the corrected p-value for that effect. We report CLES values for these results (obtained from a separate U test each) for ease of comparison with effects in other tables. The Favored group is the racial group with the higher metric for each result. Shaded groups indicate the largest racial subgroup in the course.

*Relationship between **Race** and help-seeking behavior is context dependent.* Table 9 summarizes the significant post-hoc results (after Holm-Bonferroni correction within all pairs of racial groups) for each significant **Race** results in Table 4.

We found Asian students using CH (in OOP) and read the class forum (in Data) significantly more frequently than their White peers. However, DSA and DM saw Asian students using CH significantly *less* often than White students (and also Black students in DSA). In all four significant effects on **Requests**, the largest racial subgroup in the course (shaded cells in Table 9; also see Figure 1a) is the unfavored group (i.e., utilizing CH less often), which might be attributed to these students having more same-race Classmates (and potentially more senior students of the same race) to seek help from in place of CH. However, this does not explain that in Data, the largest racial subgroup (Asian) reads the class forum the most often. More in-depth studies are necessary to understand these hidden racial effects on help-seeking.

5.4 RQ3c: controlling interdependent characteristics

Throughout our results, only **Gender** and **Confidence** exhibit consistent effects on students' help-seeking behavior across instructional contexts, where **Confidence**'s effect seems to be more specific to CH. We, therefore, focused on disentangling the effects associated with **Gender** and **Confidence** on **Requests**. Following our recent work on prior experiences [18], we built a hurdle model for each course for its suitability of modeling count data with many zeros and easier interpretation of parameter estimates [5]. Each hurdle model consists of a (binary logit) *zero part* modeling *whether a student is a CH-user* and a *count part* modeling *how many visits each user makes.* We used a truncated negative binomial model for the count part due to overdispersion, as the variance of the nonzero part of **Requests** is at least 4x its mean in every course.

Table 10: Hurdle model results for Requests with Gender and Confidence as input characteristics.

Course	N	Zero part: **Gender** Coef.	std.	p	**Confidence** Coef.	std.	p
Python	530	0.214	0.134	0.11	-0.287	0.076	$\mathbf{1.5 \times 10^{-4}}$
DM	129	0.818	0.249	**0.001**	-0.230	0.122	0.06
Algo	83	0.988	0.356	**0.006**	-0.258	0.207	0.21
Data	833	0.676	0.128	$\mathbf{1.3 \times 10^{-7}}$	-0.152	0.061	**0.013**

Course	N	Count part: **Gender** Coef.	std.	p	**Confidence** Coef.	std.	p
Python	530	0.147	0.181	0.42	-0.204	0.117	0.08
DM	129	-0.114	0.269	0.67	-0.338	0.156	**0.03**
Algo	83	0.784	0.482	0.10	-0.117	0.218	0.59
Data	833	0.260	0.201	0.20	-0.439	0.101	$\mathbf{1.5 \times 10^{-5}}$

Woman/Nonbinary is coded as 1 and Man is coded as 0. p-values for characteristics significant at the 0.05-level are bolded. Effect directions that match the results in **RQ1** are shaded. Note that the purpose of these models is to investigate whether the effects of **Gender** and **Confidence** on **Requests** remain substantial when the other is explicitly controlled, not to predict **Requests** accurately by using only these two characteristics. We therefore omit the intercepts, the dispersion coefficients, and the model performances for brevity.

***Gender** and **Confidence** are both relevant in students' CH usage, but in slightly different ways.* The results of the hurdle models (summarized in Table 10) collectively show that both students' **Gender** and **Confidence** are related to their zero part of **Requests** (i.e., whether they are a CH user) in some courses even when the other characteristic is controlled, while only **Confidence** seems to be relevant for the count part (their frequency of using CH if they are a user). Although the models for DM and Algo have relatively lower statistical power as fewer classes collected **Confidence** (see Table 1), many insignificant effects in these courses (e.g., **Gender** in the count part of Algo) are in fact *stronger* than that in the (more statistically powerful) Python and Data.

Together with the gendered effect in student approaches revealed in **RQ2**, this might indicate that **Gender** and **Confidence** influence different parts of students' help-seeking behavior in CH: Woman/Nonbinary students prefer CH more than their Man peers and thus are more likely to be a CH user (zero part), while among CH users, the usage frequency is driven by **Confidence** (which itself is related to students' frequency of *needing help*).

6 Discussions

Our results revealed many insights on gendered effects in help-seeking. Existing works on behavioral data in consulting hours [4, 7] and class forums [22, 46] found Woman/Nonbinary seek help more frequently than their Man peers in each respective resource. It is tempting to take these results out of context and conclude that Woman/Nonbinary in computing contexts simply seek help more frequently than their Man peers. However, our results indicate that we must be more precise with this claim: *Woman/Nonbinary seek more help* ***from internal resources*** *than Man*. Through triangulating behavioral data with approach data in **RQ2**, we see that Woman/Nonbinary collectively prefer internal resources over external resources; on the other hand, Man collectively prefer external resources, which are relatively understudied due to the difficulty in collecting accurate behavioral data. Furthermore, our models in Section 5.4 suggest that the gendered effect in CH usage frequency *among users* may be moderated by **Confidence**. These nuances reveal a need for *more research dedicated to attitudinal aspects of student help-seeking*: if we keep studying exclusively students' help-seeking behavior in internal resources just because the data is rich there, we risk not seeing the big picture.

Our results that reveal CS students using anonymity less than their Non-CS peers (Table 6), albeit not significant, mirror that of students with prior programming experience reported by recent works [18, 39]. This phenomenon can be explained by the social status theory via a similar argument posited by Sibia *et al.*'s work [39]: being a CS major is a *higher perceived status* in a computing course, similar to having prior programming experience. CS students thus may feel more belonged and may not feel as inclined to hide their identities as their Non-CS peers.

Finally, we note that the instructional context differences in our results do *not* seem to be well-explained by the types of courses we introduced in Section 3.1. For both **Year** and **Race**, the context differences we discussed in Section 5.3 seem more related to the distribution of the characteristic itself (e.g., different **Year** distributions in introductory vs. advanced courses) than the instructional styles in the courses (whether the course is flipped, or whether there is a semester-long project, etc.). As such, it is dangerous to draw conclusions on relationships between instructional styles and student help-seeking behavior without sufficiently comparable student populations.

7 Limitations and Future Work

7.1 Scope

Although we collected data from two institutions across a 3.5-year timespan, both institutions are research-intensive institutions located in the same geographical region in the US. We therefore cannot claim that our insights generalize to primarily-undergraduate institutions, minority-serving institutions, other geographical regions, or other institutions with smaller class sizes. More replication studies are needed to determine what (or whether any) findings apply to other contexts. As mentioned in Section 4, we chose not to include prior experience (either programming experience or other experience revelant to the course context) in the scope of this study because (1) the type of experience relevant to a course depends on the course topic, and (2) there are sometimes multiple types of relevant experience to a course, necessitating different analytical plans. This is *far from implying* prior experience does not have any impact on students' help-seeking approaches and behavior. Instead, we investigate the effects of prior experience in a separate study [18] and leave the interplay between prior experience and other student characteristics as future work.

While we briefly analyzed the interplay between other characteristics by studying the effects of **Gender** and **Confidence** while explicitly controlling for the other, we did *not* quantitatively address intersectionality broadly among student characteristics like that in some recent works [29, 38] in computing education. That is, none of our analyses performed in this study could identify any intersectional effect (e.g., how/whether being *Woman/Nonbinary and Latinx/e/a/o* simultaneously is related to a student's help-seeking approach and behavior). Due to sample size concerns (lack of statistical power) and IRB-regulations (unable to report aggregate data for any demographic subgroup of less than five students), we could not perform a comprehensive intersectional analysis. Nevertheless, we acknowledge that the quantitative and aggregate nature of our study biases all results towards majority groups and hides the unique and subtle experiences of small intersecting characteristics. More future work is needed to deeply understand individual students' unique experiences.

7.2 Data Collection

Our data is subject to fluctuating enrollment and consent rates of the 40 classes. Small details in the implementation of the informed consent process changed during the timespan, resulting in later classes having lower consent rates than earlier classes. In conducting course-level analyses, we assumed such biases were minor and tolerable. Our student characteristics data contains missing, duplicated, and inferred data, which we acknowledged in Section 3.2. We also mapped student ages in Matlab-Fa24 and OOP-Fa24 to their **Year** (Section 3.2.4). To understand the potential impact of this operation, we repeated all analyses on **Year** with these two classes excluded and found only marginal differences in the results: no result changed significance at the $\alpha = 0.05$ level, and the maximum difference in any CLES among all significant effects was 0.01. We therefore concluded that this mapping did not severely impact the results.

Our behaviorial data comes only from internal resources and not from external resources, and is subject to the fluctuating help ecosystem designs of each class (e.g., how much of consulting hours was typically available, how many interactions had invalid timestamps). We however do not have any reason to believe that such data inaccuracies would biase in favor of/against a particular student subpopulation. In some classes, teaching staff sometimes changed a public thread to private (or vice versa), and it is unknown whether they also changed the post to anonymous when changing it from private to public, which potentially biased both **Anon%** and **Priv%**. For Piazza datasets, we inferred whether a post was private from the number of users who viewed that post, which also could introduce biases in **Priv%**. However, we note that we did not find or claim any effects between any characteristic and **Priv%** throughout this study.

Our approach data comes from non-validated survey questions. Although the same survey instrument has been collected and studied in a few recent works [17, 19], it is unclear whether students fully understood and interpreted the Order survey question as intended. The survey was also administered at slightly different timepoints in each class. The later (in a semester) a student responded to the survey, the more their help-seeking approach was influenced by the current class's help context and experience; we cannot distinguish nor eliminate this confound. As such, our approach data has significantly lower fidelity than the behavioral data.

7.3 Methodology

Each student could be included in our analyses multiple times, once for each class they took in our data collection scope. This naturally biases the overall results towards students that took more classes. However, we note that no student appeared in our data in two different offerings of the same course,[11] and none of our statistical tests pools data from multiple courses together. Therefore, no single result in our analyses was impacted by repeated students.

We choose to perform course-level analyses (i.e., each test only draws data from one course) throughout the study over any kind of multilevel models. This is mainly to account for the possibility that the relationship between a certain characteristic and a certain metric goes in opposite directions in different instructional contexts (as we indeed revealed in Section 5.3 for **Year** and **Reads**), which would have been harder to detect in multilevel models. A side benefit of this choice is that single results between a characteristic and a metric can be easily compared with past results in a similar course, as part of our **RQ1** is a replication of existing studies.

We therefore designed our false-positive error controlling plan to ensure fair comparison between our results and existing results. The significance of some individual results could change had we implemented a more conservative (or less conservative) error controlling plan. To not hinge our insights on significance levels, we put much emphasis on cross comparing the directions of effects found in all courses, and reported overall trends when all effects point to the same direction (even if none is post-correction significant).

In the analysis of **RQ2**, we did not perform any inferential tests that involve comparing an individual student's ranking of a single resource with another student's ranking of the same resource. This is mainly due to the design of the Order survey; as each individual student only classified each resource into one of four groups, one of which (the *omitted* group) being rare, there are only four possible values for each individual observation (as opposed to metrics such as **Requests** taking values from 0 to 40-60 depending on the course), severely reducing the statistical power and robustness of the possible inferential tests. Individual-level inference analysis would have been more analytically suitable had each student provided a full ranking of all 10-12 resources, but we do not feel such granularity in help-seeking approaches is reasonable to ask students to provide. This again shows the relative lack of data granularity on external help resources.

8 Conclusions and Implications

We studied the relationships between computing students' characteristics and their help-seeking. Our results revealed gendered effects in both students' help-seeking approaches and behavior, i.e., Woman/Nonbinary students preferring course-affiliated help resources over external resources (and also using them) more than Man students. We observed students who felt more confident sought less help from course-affiliated resources and preferred using external resources more than their less confident peers. We found minority groups such as Latinx/e/a/o and Non-CS students rely on people outside of the course more than their peers, and first-year students prefer course-affiliated resources less than upperclass peers. We also identified **Year** and **Race** effects that are highly context dependent.

Our work expands the current understanding of the relationships between student characterstics and help-seeking from mostly in course-affiliated help resources to over the entire help resource landscape and from mostly in introductory programming contexts to other contexts. However, much further work is needed to understand the interplay of different student characteristics, as well as the connections between characteristics and psychoconstructs, such as self-efficacy and sense of belonging, in the context of student help-seeking. Nevertheless, our work provides insights that help inform computing educators to tailor their help ecosystem design to their targeted user populations while being subject to influences from their own course contexts. As our results show, the relationship between students' confidence and their help-seeking behavior is specific to one-on-one synchronous consulting hours, which calls for dedicated training for course staff on providing affective support in addition to technical help.

[11] In other words, no students repeated a course, consented multiple times in that course, and provided characteristics data multiple times in that course.

Acknowledgments

We thank all instructors, teaching staff, and students at both institutions for their contribution to the data collection, deidentification, and transferral. We thank the anonymous reviewers for their helpful and constructive feedback to the paper. This material is based upon work supported by the National Science Foundation under Award No. #1934965, #1934975, #2336804, and #2336805.

References

[1] Maria Ijaz Baig and Elaheh Yadegaridehkordi. 2023. Flipped classroom in higher education: a systematic literature review and research challenges. *International Journal of Educational Technology in Higher Education* 20, 1 (2023), 61. https://doi.org/10.1186/s41239-023-00430-5

[2] Anael Kuperwajs Cohen, Alannah Oleson, and Amy J. Ko. 2024. Factors Influencing the Social Help-seeking Behavior of Introductory Programming Students in a Competitive University Environment. *ACM Trans. Comput. Educ.* 24, 1 (2024), 11:1–11:27. https://doi.org/10.1145/3639059

[3] Ed Discussion. 2025. https://edstem.org/.

[4] Augie Doebling and Ayaan M. Kazerouni. 2021. Patterns of Academic Help-Seeking in Undergraduate Computing Students. In *ACM Koli Calling.* 13:1–13:10. https://doi.org/10.1145/3488042.3488052

[5] Cindy Xin Feng. 2021. A comparison of zero-inflated and hurdle models for modeling zero-inflated count data. *Journal of statistical distributions and applications* 8, 1 (2021), 8. http://doi.org/10.1186/s40488-021-00121-4

[6] Zhikai Gao, Adam M. Gaweda, Collin F. Lynch, Sarah Heckman, Damilola Babalola, and Gabriel Silva de Oliveira. 2024. Using Survival Analysis to Model Students' Patience in Online Office Hour Queues. In *ACM SIGCSE TS.* 1646–1647. https://doi.org/10.1145/3626253.3635517

[7] Zhikai Gao, Sarah Heckman, and Collin Lynch. 2022. Who Uses Office Hours?: A Comparison of In-Person and Virtual Office Hours Utilization. In *ACM SIGCSE TS.* 300–306. https://doi.org/10.1145/3478431.3499334

[8] Joan Giblin, Jill E. Stefaniak, Angela Eckhoff, and Tian Luo. 2021. An exploration of factors influencing the decision-making process and selection of academic help sources. *J. Comput. High. Educ.* 33, 1 (2021), 1–18. https://doi.org/10.1007/s12528-020-09252-0

[9] Irene Hou, Sophia Mettille, Owen Man, Zhuo Li, Cynthia Zastudil, and Stephen MacNeil. 2024. The Effects of Generative AI on Computing Students' Help-Seeking Preferences. In *ACM ACE.* 39–48. https://doi.org/10.1145/3636243.3636248

[10] Iris K. Howley and Carolyn P. Rosé. 2018. Empirical Evidence for Evaluation Anxiety and Expectancy-Value Theory for Help Sources. In *ICLS*. https://repository.isls.org//handle/1/891

[11] Aimable Robert Jonckheere. 1954. A distribution-free k-sample test against ordered alternatives. *Biometrika* 41, 1/2 (1954), 133–145.

[12] David A. Joyner, Lily Bernstein, Ian Bolger, Maria-Isabelle Dittamo, Stephanie Gorham, and Rachel Hudson. 2022. Anonymity: A Double-Edged Sword for Gender Equity in a CS1 Forum?. In *ACM SIGCSE TS*. 766–772. https://doi.org/10.1145/3478431.3499289

[13] Stuart A. Karabenick and Jean-Louis Berger. 2013. Help seeking as a self-regulated learning strategy. In *Applications of self-regulated learning across diverse disciplines: A tribute to Barry J. Zimmerman*, Héfer Bembenutty, Timothy J. Cleary, and Anastasia Kitsantas (Eds.). 237–261. https://www.infoagepub.com/products/Applications-of-Self-Regulated-Learning-across-Diverse-Disciplines

[14] Stuart A. Karabenick and Myron H. Dembo. 2011. Understanding and facilitating self-regulated help seeking. *New Directions for Teaching and Learning* 2011 (2011), 33–43. Issue 126. https://doi.org/10.1002/tl.442

[15] Hieke Keuning, Johan Jeuring, and Bastiaan Heeren. 2019. A Systematic Literature Review of Automated Feedback Generation for Programming Exercises. *ACM Trans. Comput. Educ.* 19, 1 (2019), 3:1–3:43. https://doi.org/10.1145/3231711

[16] Shao-Heng Ko and Kristin Stephens-Martinez. 2023. What drives students to office hours: individual differences and similarities. In *ACM SIGCSE TS*. 959–965. https://doi.org/10.1145/3545945.3569777

[17] Shao-Heng Ko and Kristin Stephens-Martinez. 2024. The Trees in the Forest: Characterizing Computing Students' Individual Help-Seeking Approaches. In *ACM ICER*. 343–358. https://doi.org/10.1145/3632620.3671099

[18] Shao-Heng Ko and Kristin Stephens-Martinez. 2025. Prior What Experience? The Relationship Between Prior Experience and Student Help-Seeking Beyond CS1. In *ACM ITiCSE*. 100–106. https://doi.org/10.1145/3724363.3729092

[19] Shao-Heng Ko and Kristin Stephens-Martinez. 2025. Rethinking computing students' help resource utilization through sequentiality. *ACM TOCE* 25, 1, Article 7 (2025), 34 pages. https://doi.org/10.1145/3716860

[20] Shao-Heng Ko, Kristin Stephens-Martinez, Matthew Zahn, Yesenia Velasco, Lina Battestilli, and Sarah Heckman. 2025. Student Perceptions of the Help Resource Landscape. In *ACM SIGCSE TS*. 596–602. https://doi.org/10.1145/3641554.3701851

[21] Sam Lau and Philip J. Guo. 2023. From "Ban It Till We Understand It" to "Resistance is Futile": How University Programming Instructors Plan to Adapt as More Students Use AI Code Generation and Explanation Tools such as ChatGPT and GitHub Copilot. In *ACM ICER*. 106–121. https://doi.org/10.1145/3568813.3600138

[22] Ryan Lenfant, Alice Wanner, John R. Hott, and Raymond Pettit. 2023. Project-Based and Assignment-Based Courses: A Study of Piazza Engagement and Gender in Online Courses. In *ACM ITiCSE*. 138–144. https://doi.org/10.1145/3587102.3588833

[23] Soohyun Nam Liao, Kartik Shah, William G. Griswold, and Leo Porter. 2021. A Quantitative Analysis of Study Habits Among Lower- and Higher-Performing Students in CS1. In *ACM ITiCSE*. 366–372. https://doi.org/10.1145/3430665.3456350

[24] Dastyni Loksa, Lauren E. Margulieux, Brett A. Becker, Michelle Craig, Paul Denny, Raymond Pettit, and James Prather. 2022. Metacognition and Self-Regulation in Programming Education: Theories and Exemplars of Use. *ACM Trans. Comput. Educ.* 22, 4 (2022), 39:1–39:31. https://doi.org/10.1145/3487050

[25] Kara A. Makara and Stuart A. Karabenick. 2013. Characterizing sources of academic help in the age of expanding educational technology: A new conceptual framework. In *Advances in help-seeking research and applications: The role of emerging technologies*, Stuart A. Karabenick and Minna Puustinen (Eds.). 37–72. https://www.infoagepub.com/products/Advances-in-Help-Seeking-Research-and-Applications

[26] Amogh Mannekote, Mehmet Celepkolu, Aisha Chung Galdo, Kristy Elizabeth Boyer, Maya Israel, Sarah Heckman, and Kristin Stephens-Martinez. 2022. Don't Just Paste Your Stacktrace: Shaping Discussion Forums in Introductory CS Courses. In *ACM SIGCSE TS*. 1164–1164. https://doi.org/10.1145/3478432.3499110

[27] Anna May, Johannes Wachs, and Anikó Hannák. 2019. Gender differences in participation and reward on Stack Overflow. *Empirical Software Engineering* 24 (2019), 1997–2019. https://doi.org/10.1007/s10664-019-09685-x

[28] Diba Mirza, Phillip T. Conrad, Christian Lloyd, Ziad Matni, and Arthur Gatin. 2019. Undergraduate Teaching Assistants in Computer Science: A Systematic Literature Review. In *ACM ICER*. 31–40. https://doi.org/10.1145/3291279.3339422

[29] Vidushi Ojha, Leah West, and Colleen M. Lewis. 2024. Computing Self-Efficacy in Undergraduate Students: A Multi-Institutional and Intersectional Analysis. In *ACM SIGCSE TS*. 993–999. https://doi.org/10.1145/3626252.3630811

[30] Aadarsh Padiyath, Xinying Hou, Amy Pang, Diego Viramontes Vargas, Xingjian Gu, Tamara Nelson-Fromm, Zihan Wu, Mark Guzdial, and Barbara Ericson. 2024. Insights from Social Shaping Theory: The Appropriation of Large Language Models in an Undergraduate Programming Course. In *ACM ICER*. 114–130. https://doi.org/10.1145/3632620.3671098

[31] José Carlos Paiva, José Paulo Leal, and Álvaro Figueira. 2022. Automated Assessment in Computer Science Education: A State-of-the-Art Review. *ACM Trans. Comput. Educ.* 22, 3 (2022), 34:1–34:40. https://doi.org/10.1145/3513140

[32] Leah Perlmutter, Jean Salac, and Amy J. Ko. 2023. "A field where you will be accepted": Belonging in student and TA interactions in post-secondary CS education. In *ACM ICER*. 356–370. https://doi.org/10.1145/3568813.3600128

[33] Piazza. 2025. https://piazza.com/.

[34] Thomas W. Price, Zhongxiu Liu, Veronica Cateté, and Tiffany Barnes. 2017. Factors Influencing Students' Help-Seeking Behavior while Programming with Human and Computer Tutors. In *ACM ICER*. 127–135. https://doi.org/10.1145/3105726.3106179

[35] Adnan Qayyum. 2018. Student help-seeking attitudes and behaviors in a digital era. *International Journal of Educational Technology in Higher Education* 15, 1 (2018), 1–16. https://doi.org/10.1186/s41239-018-0100-7

[36] Pooja Sankar, Jessica Gilmartin, and Melissa Sobel. 2015. An examination of belongingness and confidence among female computer science students. *SIGCAS Comput. Soc.* 45, 2 (2015), 7–10. https://doi.org/10.1145/2809957.2809960

[37] Christian Schlusche, Lenka Schnaubert, and Daniel Bodemer. 2023. Understanding students' academic help-seeking on digital devices - a qualitative analysis. *Research and Practice in Technology Enhanced Learning* 18 (2023), 17. https://doi.org/10.58459/rptel.2023.18017

[38] Mrinal Sharma, Hayden McTavish, Zimo Peng, Anshul Shah, Vardhan Agarwal, Caroline Sih, Emma Hogan, Ismael Villegas Molina, Adalbert Gerald Soosai Raj, and Kristen Vaccaro. 2023. Engagement and Anonymity in Online Computer Science Course Forums. In *ACM ICER*. 48–62. https://doi.org/10.1145/3568813.3600121

[39] Naaz Sibia, Angela M. Zavaleta Bernuy, Tiana V. Simovic, Chloe Huang, Yinyue Tan, Eunchae Seong, Carolina Nobre, Daniel Zingaro, Michael Liut, and Andrew Petersen. 2024. Exploring the Effects of Grouping by Programming Experience in Q&A Forums. In *ACM ICER*. 206–221. https://doi.org/10.1145/3632620.3671107

[40] Naaz Sibia, Angela M. Zavaleta Bernuy, Joseph Jay Williams, Michael Liut, and Andrew K. Petersen. 2023. Student Usage of Q&A Forums: Signs of Discomfort?. In *ACM ITiCSE*. 33–39. https://doi.org/10.1145/3587102.3588842

[41] Naaz Sibia, Giang Bui, Bingcheng Wang, Yinyue Tan, Angela M. Zavaleta Bernuy, Christina Bauer, Joseph Jay Williams, Michael Liut, and Andrew Petersen. 2024. Examining Intention to Major in Computer Science: Perceived Potential and Challenges. In *ACM SIGCSE TS*. 1237–1243. https://doi.org/10.1145/3626252.3630843

[42] James Skripchuk, John Thomas Bacher, and Thomas W. Price. 2024. An Investigation of the Drivers of Novice Programmers' Intentions to Use Web Search and GenAI. In *ACM ICER*. 487–501. https://doi.org/10.1145/3632620.3671112

[43] James Skripchuk, Neil Bennett, Jeffrey Zhang, Eric Li, and Thomas W. Price. 2023. Analysis of Novices' Web-Based Help-Seeking Behavior While Programming. In *ACM SIGCSE TS*. 945–951. https://doi.org/10.1145/3545945.3569852

[44] Stack Overflow. 2025. https://stackoverflow.com/.

[45] Nick A Stites, Edward Berger, Jennifer Deboer, and Jeffrey F Rhoads. 2019. A cluster-based approach to understanding students' resource-usage patterns in an active, blended, and collaborative learning environment. *International Journal of Engineering Education* 35, 6 (2019), 1738–1757. https://www.ijee.ie/1atestissues/Vol35-6A/16_ijee3834.pdf

[46] Adrian Thinnyun, Ryan Lenfant, Raymond Pettit, and John R. Hott. 2021. Gender and Engagement in CS Courses on Piazza. In *ACM SIGCSE TS*. 438–444. https://doi.org/10.1145/3408877.3432395

[47] Bogdan Vasilescu, Andrea Capiluppi, and Alexander Serebrenik. 2014. Gender, representation and online participation: A quantitative study. *Interacting with Computers* 26, 5 (2014), 488–511. https://doi.org/10.1093/iwc/iwt047

[48] Mickey Vellukunnel, Philip Sheridan Buffum, Kristy Elizabeth Boyer, Jeffrey Forbes, Sarah Heckman, and Ketan Mayer-Patel. 2017. Deconstructing the Discussion Forum: Student Questions and Computer Science Learning. In *ACM SIGCSE TS*. 603–608. http://doi.org/10.1145/3017680.3017745

[49] Elizabeth Wirtz, Amy Dunford, Edward Berger, Elizabeth Briody, Gireesh Guruprasad, and Ryan Senkpeil. 2018. Resource usage and usefulness: academic help-seeking behaviours of undergraduate engineering students. *Australasian Journal of Engineering Education* 23, 2 (2018), 62–70. http://doi.org/10.1080/22054952.2018.1525889

[50] David Wong-Aitken, Diana Cukierman, and Parmit K. Chilana. 2022. "It Depends on Whether or Not I'm Lucky" How Students in an Introductory Programming Course Discover, Select, and Assess the Utility of Web-Based Resources. In *ACM ITiCSE*. 512–518. https://doi.org/10.1145/3502718.3524751

[51] Stephanie Yang, Hanzhang Zhao, Yudian Xu, Karen Brennan, and Bertrand Schneider. 2024. Debugging with an AI Tutor: Investigating Novice Help-seeking Behaviors and Perceived Learning. In *ACM ICER*. 84–94. https://doi.org/10.1145/3632620.3671092

[52] Matthew Zahn, Lina Battestilli, and Sarah Heckman. 2022. Academic Help Seeking Patterns in Introductory Computer Science Courses. In *ASEE Annual Conference & Exposition*. https://peer.asee.org/41526

[53] Barry J. Zimmerman. 2001. Theories of self-regulated learning and academic achievement: An overview and analysis. In *Self-regulated learning and academic achievement: Theoretical perspectives*, Barry J. Zimmerman and Dale H. Schunk (Eds.). 1–37. https://doi.org/10.1007/978-1-4612-3618-4