# WIP: Faculty Learning Community Collective Action: Investigating the Impact of Metacognition, Transparency, and Grade Interpretation on Student Engagement and Performance in STEM Classrooms

**Dr. Stephany Santos, University of Connecticut**

Stephany Santos serves as the Executive Director for the Vergnano Institute for Inclusion in the College of Engineering at UConn. Dr. Santos is also a leader and thought partner in the Science Education Campaign for Research, Equity and Teaching (SECRET), as well as the International and Comparative Advancement of Racial Equity for Social Justice (ICARE4Justice). She leads dozens of programs, initiatives and classes that are designed to support, love on, and grow the community. She believes everyone who wants to be an engineer on in STEM deserves to be here and pursue their dreams. At UConn. Dr. Santos is faculty in Biomedical Engineering, and Affiliate Faculty for Engineering for Human Rights, and for the Engineering Education PhD Program. She is a leader nationally, including serving as the Incoming Chair for the CDEI Commision for the American Society for Engineering Education. She advocates for equality and for creating support networks and pathways forward so that as many students as possible persist and push through. Her research interests are in engineering identity and leadership development, engineering thriving, and understanding how programs and interventions can create positive, dignity-affirming experiences in engineering.

**Dr. Hasan Baig, University of Connecticut**

Dr. Hasan Baig currently serves as an Assistant Professor at the University of Connecticut's Stamford Campus, USA. Prior to joining the Computer Science and Engineering (CSE) department, Dr. Baig was a Postdoctoral Research Fellow in the Mendes Research Group at UConn Health, USA, where he focused on developing cloud-based computational tools for systems biology.

Dr. Baig is award winning educationist who has developed numerous EdTech tools for students and instructors. Other than educational technology, his research spans a diverse range of interdisciplinary areas, including Genetic Design Automation (GDA), Cloud Computing, Embedded Systems, Computer Architectures, Fault-Tolerant Systems, and Human-Machine Interactions.

**Dr. Sylvanie Wallington, University of Connecticut**
**Umut Ozuguzel, University of Connecticut**

# WIP: Teaching the Students We Have:
# A Faculty Learning Community for First-Year Engineering


## Abstract

This work-in-progress paper examines a multi-course Faculty Learning Community (FLC) coordinating first-year instruction in programming, chemistry, physics, and introductory engineering at a regional campus. The FLC uses Communities of Practice principles to translate classroom anecdotes and evidence into feasible, week-scale adjustments that better support the learners enrolled. We present three complementary sources of descriptive evidence that were planned for and collected during the FLC: an early-semester student voice survey (n = 60), midterm micro-surveys administered across first-year courses (physics n = 74; chemistry n = 87), and categorized interactions from PyPal, a faculty-developed artificial intelligence (AI) assistant used throughout the semester. These sources provided early, mid-course, and cumulative insight into students' learning preferences, affect, and approaches to problem solving. The FLC used this evidence to implement practical instructional adjustments while strengthening cross-course coherence. We describe how these decisions were informed by theory and how the FLC used ongoing feedback to refine instructional design across the first-year experience.


## Introduction

First-year engineering students traverse multiple courses simultaneously and often encounter different expectations, strategies, and rhythms for learning. A Faculty Learning Community (FLC) offers a structure through which instructors can coordinate across courses, share artifacts, and collaboratively refine practice. In this setting, the FLC supported instructors in identifying students' strengths and understanding their lived experiences: where students struggled, where expectations were unclear, and where small but meaningful instructional adjustments could better align with how students learn.

Over two years of weekly meetings, the FLC developed a shared approach for collecting and interpreting evidence about students' experiences in ways that were practical during a busy semester. After the first year, the FLC intentionally curated a set of evidence sources designed to capture different moments in the student learning journey, including early expectations, mid-course affect and intentions, and cumulative patterns of engagement. These sources helped the community better understand how students made sense of material, how they used support structures, and how instructional designs could be tuned in response.

## Theoretical Frameworks and Framing

This study is grounded in three complementary frameworks. Communities of Practice (CoP) theory conceptualizes learning as participation in a shared enterprise with evolving practices and artifacts, which justifies treating the FLC itself as the primary site of faculty learning and coordination [1], [2]. Design-Based Research (DBR) provides a methodological and theoretical frame for iterative design, enactment, analysis, and redesign in authentic settings, which supports the work-in-progress nature of the study and the focus on usable knowledge for practice [3]. The ICAP framework (Interactive, Constructive, Active, Passive) distinguishes four modes of cognitive engagement and supports cross-course interpretation of how instructional designs position students to think and participate [4].

Additional perspectives inform specific design moves and interpretations but are not treated as grounding frameworks. These include Transparency in Learning and Teaching (TILT) for clarifying purpose, tasks, and criteria [5]; Cognitive Load Theory (CLT) for sequencing worked examples and managing extraneous load [6]; Self-Determination Theory (SDT) for interpreting motivation and help-seeking [7]; and the Awareness, Monitoring, Evaluation, Regulation, and Transfer (AMERT) metacognitive hierarchy, which helps name patterns observed in planning and debugging behaviors [8].

## Methods

### *Participants and Faculty Learning Community Context*

Participants were four instructors teaching the required first-year sequence in programming, chemistry, physics, and introductory engineering at a campus enrolling approximately 130 engineering students and 180 liberal arts and sciences students. The shared aim was to use practical evidence to tune scaffolds, not to standardize syllabi. Equity and feasibility shaped what could be sustained by instructors and students.

The FLC met weekly for two years, bringing together instructors from programming, chemistry, physics, and introductory engineering. Meetings emphasized collaborative review of student evidence, discussion of emerging instructional challenges, and planning of small feasible adjustments that could be implemented within the semester. After the first year, the FLC intentionally curated three recurring evidence sources that captured beginning-of-semester expectations, midterm affect and intentions, and cumulative semester-long engagement patterns. These discussions supported shared understanding of student experiences and helped faculty align instructional designs across courses.

### *Research Questions*

Guided by this aim, we examined the following research questions:

**RQ1.** How can student voice and feedback impact instructors’ in-class or out-of-class pacing, support, and resources in early STEM courses?

**RQ2.** What might AI assistant interactions reveal about students, and how should the FLC interpret or adjust?

**RQ3.** How can a faculty learning community that includes instructors from multiple first-year foundational STEM courses better support student learning, motivation, engagement, and retention?

## Artifacts and Evidence from the FLC

The FLC supported instructors in both individual and collective endeavors. The community intentionally curated evidence from students at three points in the semester, each serving a distinct instructional purpose.

### *Early Understanding: Beginning-of-Semester Survey*

At the beginning of the semester, the introductory engineering course administered a Week 1 student voice survey that yielded 60 responses. The purpose of this instrument was to surface students' initial expectations, learning preferences, and prior understandings of their intended pathways so that instructors could align early scaffolds with the needs of the learners enrolled. Students were asked why they chose the institution, how they currently understood their intended major or career interests, and how they believed they learned best.

Within the Faculty Learning Community, these Week 1 responses were treated as early-semester evidence used to guide initial instructional planning and to ensure that early course activities reflected students' stated learning approaches rather than assumed ones.

### *Midsemester Micro-Survey (All First-Year Courses)*

Shortly after midterm grades were released, instructors in programming, chemistry, physics, and introductory engineering administered a brief micro-survey to gather timely, student-centered feedback on affect, planned study behaviors, and perceived barriers to engagement. The physics micro-survey yielded 74 responses across Physics 1401 ($n = 48$) and Physics 1402 ($n = 26$). The chemistry micro-survey yielded 87 responses across Chemistry 1127 ($n = 62$) and Chemistry 1124 ($n = 25$). The instrument was intentionally low-burden and consisted of a small set of items asking students how they were feeling about the course at that moment, what actions they intended to take in the next week, and which existing supports they had or had not been using. Across courses, the survey was framed as an invitation for students to reflect on their own learning while giving instructors actionable insight into how students were experiencing the midpoint of the semester.

Within the FLC, the midsemester micro-survey served as a form of human-centered formative feedback for instructors, rather than as an evaluative measure of students. In weekly meetings, faculty used the responses to understand student affect, identify points where activation energy to seek help might be high, and consider whether support structures were accessible and well-signaled. The FLC treated this evidence as a prompt for instructional adjustment during the ongoing semester, using the midcourse timing to identify small, feasible modifications that could lower barriers to participation, clarify expectations, and strengthen cross-course coherence.

Several instructors noted that some students expressed surprise and appreciation for being asked these questions, reinforcing the survey's role in fostering relational trust and demonstrating instructional care.

### *PyPal Analytics (AI-Supported Learning)*

PyPal is a faculty-developed artificial intelligence (AI) learning tool designed to provide sustained, course-embedded support to students throughout the semester in the introductory programming course. The tool was developed specifically for this context to help students make progress during out-of-class work by offering on-demand explanations, debugging guidance, and planning assistance. PyPal logs each student interaction in a minimally intrusive way, capturing the prompt category (for example, conceptual explanation, debugging and evaluation, project planning or task decomposition), the timing of the request, and the length of the exchange. No personally identifying information beyond course enrollment is collected. The intent of the tool is support first, analytics second: PyPal was created as a mechanism for students to receive help whenever they needed it, not as a monitoring instrument.

Within the FLC, PyPal interactions were treated as cumulative, course-embedded evidence about how students engaged with material across the semester. During weekly meetings, faculty reviewed category-level summaries to understand where students were seeking clarification, what kinds of debugging challenges they encountered, and how often they requested help with planning or decomposing project tasks. These patterns were used not to evaluate individual students, but to guide instructional decision-making across courses by identifying where conceptual anchors, worked examples, or planning scaffolds might be most beneficial. The PyPal data complemented the early-semester student voice survey and the midsemester micro-survey by offering a third, longitudinal perspective on how students approached learning over time.

## Results

In this paper, we focus on how the FLC discussed and interpreted results from their interventions. We present brief themes of student results below, with selected descriptive statistics from the surveys reported in the Appendix.

### *Week 1 Student Voice Survey*

The Week 1 survey (n = 60) revealed several patterns relevant to instructional design. When asked why they chose the institution, the most frequently cited reasons were proximity to home or familiarity with the area, followed closely by affordability, and then the reputation or availability of their intended program. Several students also cited family members who had attended or who worked in related fields. These motivations are commonly seen among first-generation or low-income students, and they suggested that for some students, the decision to pursue engineering was tied to economic security, geographic accessibility, and a clear return on investment.

When asked why they chose their major, a clear majority described genuine interest, curiosity, or passion for the field. About a quarter cited career and financial motivations, and a smaller group mentioned problem solving, building, or creating things. Others entered engineering with limited prior exposure and used the introductory course to understand what engineers do. Many expressed strong interest in real-world problem solving, curiosity about specific domains, and pride in being the first in their family to pursue a technical degree.

Regarding learning preferences, the most frequently mentioned approaches were visual learning and hands-on practice, each cited by roughly a third of respondents, followed by learning through worked examples, mentioned by about one in five. Smaller numbers mentioned interactive or collaborative formats, structured notes, and listening or lecture-based learning. These self-reported preferences converged on a clear pattern: students expected to learn by seeing, doing, and working through concrete examples before attempting problems independently. This helped the FLC design early course experiences that validate diverse goals, clarify pathways, and support students who bring different forms of prior knowledge.

*Midsemester Micro-Survey: Affect and Grade Distributions*

The midsemester micro-survey provided a snapshot of students' experiences after viewing their midterm grades. Grade distributions differed markedly across courses. In physics, the distribution skewed higher: about a third of respondents reported midterm grades in the A range, and roughly one in ten fell below 60. In chemistry, the pattern was strikingly reversed: more than a third reported midterm grades below 60, and fewer than one in ten reached the A range. This cross-course disparity was a significant finding for the FLC, as it meant the same cohort of students was simultaneously experiencing success in one course and discouragement in another.

Affect responses reflected these grade differences. In physics, more than three quarters of students reported feeling motivated, while discouragement was uncommon. In chemistry, motivation dropped to just over half, and nearly a third reported feeling discouraged. Across both courses, however, a majority of students reported that their midterm grade was an accurate reflection of their effort. This pattern of self-aware attribution where students acknowledge the link between effort and outcome was an important signal for the FLC. This suggested that students were not primarily externalizing blame but were reflecting on their own contributions to their performance.

Even among students earning below 60, motivation persisted: roughly half of the lowest-performing students in both physics and chemistry still agreed or strongly agreed that they felt motivated. This coexistence of discouragement and motivation, namely short-term discouragement paired with renewed determination to improve, was a recurring theme in the open-ended responses. Students reported mixed emotions, and several described a turning point at midterm where they resolved to change their approach. Many expressed appreciation that instructors asked for their input, which reinforced a sense of care.

*Midsemester Micro-Survey: The Intent–Behavior Gap*

A central finding from the micro-survey was a consistent gap between students' stated intentions to increase engagement and their actual baseline usage of available supports. The pattern was most visible in office hours attendance. In physics, 70% of respondents reported never attending office hours, yet a majority of those same non-attenders indicated they intended to do more of it in the second half of the semester. A similar pattern appeared in chemistry, where nearly half reported never attending and a comparable share of non-attenders expressed intent to increase. Tutoring center use showed the same gap: large portions of both cohorts had never visited, yet most intended to start.

Students also planned to adopt more structured strategies for the second half of the semester, including targeted practice, systematic review of notes, rewatching short concept videos, and attending office hours or tutoring with prepared questions. Others intended to start work earlier, form study groups, or revisit previous assignments to identify recurring errors. Grade expectations further illustrated this optimistic framing: in physics, about a third of students expected a higher final grade than their current midterm, while in chemistry, where midterm grades were substantially lower, a clear majority expected to finish higher. This asymmetry suggested that the chemistry cohort, facing steeper deficits, was expressing even stronger aspirational intent, a finding that underscored the need for accessible, low-friction supports to help students translate intention into action.

These intentions pointed to opportunities for low-friction supports, timely examples, and clear participation cues that could help students follow through. However, faculty observations over the remainder of the semester indicated that these stated intentions did not translate into visible behavioral changes at the scale students described.

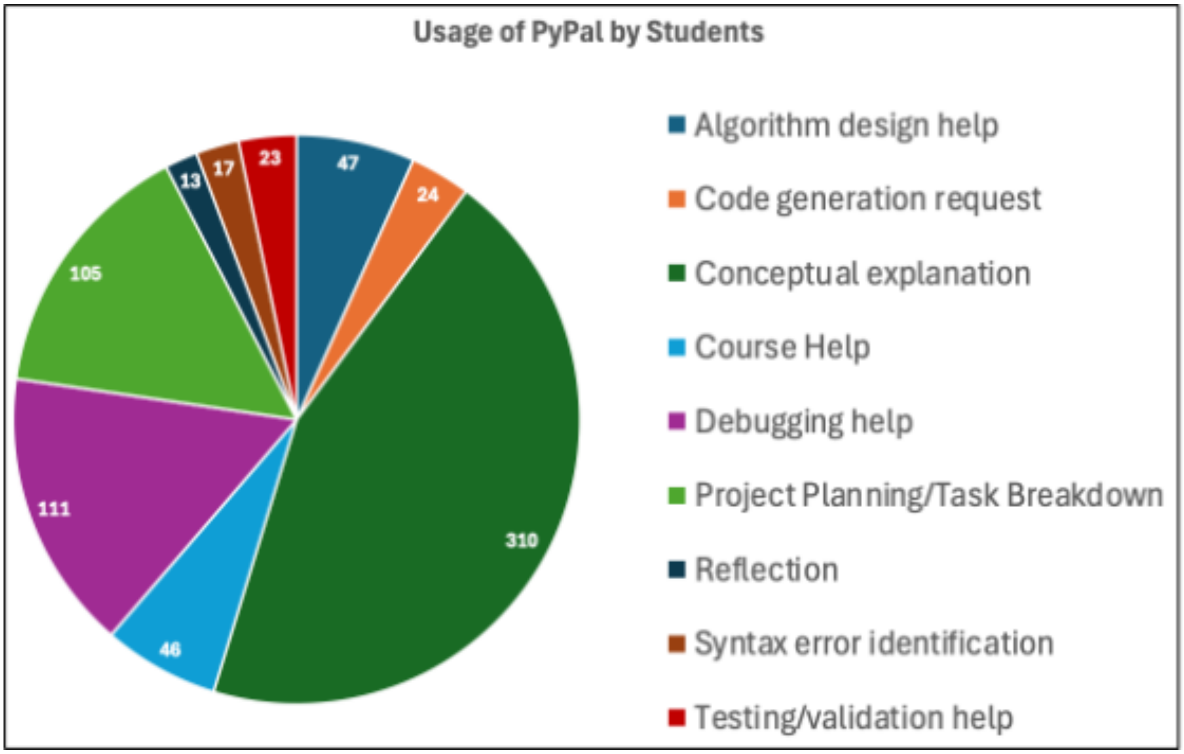
*Midsemester Micro-Survey: AI Usage Patterns*

The micro-survey asked separately about using AI as a study tool and using AI to get quick homework answers, revealing a nuanced pattern. In both courses, a majority of students reported at least occasional use of AI for studying, with chemistry students reporting somewhat higher rates than physics students. Use of AI for quick homework answers was substantially lower in both courses, with most students reporting they rarely or never used it for that purpose.

Notably, students' intended changes showed self-regulatory awareness. In both physics and chemistry, close to half of students intended to reduce their use of AI for quick answers in the second half of the semester. Intended changes for AI as a study tool were more stable, with the majority in both courses planning to maintain their current level. This distinction, where students chose to maintain AI for conceptual support while intending to reduce its use as a shortcut, paralleled the PyPal finding that students used AI primarily for understanding rather than bypassing.

*PyPal Interaction Patterns*

PyPal interactions offered cumulative evidence of how students sought help across the semester (Figure 1). Across 700 randomly sampled prompts, the most frequent category was conceptual explanation, suggesting students were trying to make sense of underlying ideas before coding. Debugging-related requests showed where students encountered errors they could not resolve alone, and planning questions indicated when they needed help organizing multi-step tasks. The pattern of use suggested that students often turned to PyPal when they wanted to persist independently before seeking in-person support.




[Figure 1: Distribution of student PyPal usage across 700 randomly sampled prompts collected over the semester, categorized by type of academic support requested.]

*Incorporating Student Voice at Multiple Time Points*

Table 1 summarizes how the FLC incorporated student voice and classroom evidence at several key moments in the semester to better support student learning, motivation, engagement, and retention (RQ3). The table brings together four sources of information collected at the beginning, middle, and across the full semester and highlights how each source prompted faculty discussion and potential adjustments within and across courses.

**Table 1.** Using student voice and course evidence to drive near-term instructional adjustments.

| Evidence source | Observed pattern | FLC Discussion and Potential course adjustment | ICAP interpretation and supporting lens |
|---|---|---|---|
| Intro engineering student voice (Week 1) | Students requested visual and hands-on learning, examples before independent work, and structured notes. | Guided attempt and debrief; post structured notes; add short authentic design prompts aligned to interests. | Shift from Passive to Active and Constructive engagement. Support with transparency of purpose, task, criteria (TILT) [5]. |
| Midterm micro-survey (all first-year classes), representative physics and chemistry | Intentions to increase practice and participation exceeded current use of supports. Seven in ten physics students never attended office hours, yet a majority of non-attenders intended to increase. | Inform students of survey results; observe changes in student behaviors in class engagement, or attending office hours; time management help. | Lower barriers to Active engagement and create conditions for Constructive engagement. Support with worked examples to manage cognitive load (CLT) [6]. |
| PyPal categorized interactions (programming) | High conceptual explanation and debugging requests; distinct need for planning and task breakdown. | AI-supported learning over shortcuts; short concept-clarify segments; planning template before projects and revision after first tests. | Use metacognitive prompts aligned to awareness, monitoring, evaluation, regulation, transfer (AMERT) [8]. |
| Cross-course comparison for same cohort | Students experienced inconsistent rhythms across courses; the same cohort reported high motivation in physics but markedly lower motivation in chemistry. Uncertainty about how to participate productively. | Attempt, evaluate, revise; distributed responsibility for fundamentals (like how to use calculators). | Increase opportunities for Constructive engagement across contexts and reduce hidden curriculum through transparency (TILT) [5]. |

*Student Intent versus Observed Behavior*

Across physics and chemistry, the micro-survey indicated that student intentions to increase practice and participation outpaced faculty's observed use of resources or supports such as office hours or tutoring. To summarize, students stated that they would change behaviors toward high-impact practices like engaging with optional videos, attending office hours proactively, and forming study groups, but FLC participants did not observe a noticeable change in any of the following behaviors beyond traditional spikes near exams: required assignment completion; optional engagement with online tools such as video views; recovery attempts such as submitting assignments for bonus points; response to questions in class or participation in in-class activities; academic performance as reflected in grades; academic effort as displayed in work or thought process; and professor interaction such as engagement in office hours for help.

A quote from one FLC participant captures this tension:

> *I am convinced these are helping with student performance, mostly because many, many students have told me how helpful they are. My teaching evaluation survey comments routinely cite the videos, and when I recommend the videos to the class multiple people will attest to their effectiveness. Can I really measure how well they work, though? I looked at the viewing statistics, and there does not seem to be any particular pattern. Probably since I give a weekly quiz, students are motivated to watch videos all semester long.*

The persistence of this intent–behavior gap suggests that motivation alone is insufficient without structural supports that lower the activation energy for follow-through. FLC participants discussed how to better capture changes in these domains qualitatively or quantitatively in a low-stakes way that does not burden an instructor. This remains a work in progress.

## Discussion

Interpreted through the ICAP framework [4], evidence across courses suggested that many learners were initially positioned for Passive or minimally Active engagement, rather than for Constructive or Interactive sense-making. Midterm survey responses, help-seeking patterns, and AI assistant interactions indicated that students frequently wanted to understand underlying concepts but were uncertain about how to begin or how their efforts would be evaluated. The FLC interpreted this pattern not as a lack of motivation, but as a misalignment between instructional intentions and the cognitive engagement students were positioned to enact. Across these moves, transparency supported engagement by clarifying how students could participate productively and by reducing the hidden curriculum around what counts as good work, which has implications for equity and persistence [5].

*Addressing the Research Questions*

Week 1 student voice informed early scaffolds and sequencing in line with RQ1. The finding that students most frequently identified as visual learners and preferred hands-on practice provided concrete direction for designing course activities that began with demonstrations and guided examples rather than independent problem sets. Midsemester micro-survey evidence supported week-scale modifications that lowered activation energy for help-seeking and practice, also consistent with RQ1. The cross-course comparison of affect, where the same cohort reported substantially higher motivation in physics than in chemistry, helped the FLC consider how course-level grade distributions shaped students' emotional experience of the first-year sequence.

PyPal interactions revealed where students devoted effort to conceptual sense-making, debugging, and planning, which guided targeted demonstrations and templates across courses, addressing RQ2. The parallel finding from the micro-surveys, where nearly half of students across physics and chemistry intended to reduce AI use for quick answers while maintaining it for studying, suggested that students were developing a self-regulatory distinction between productive and unproductive AI use, a finding that supported the FLC's approach of naming productive uses of AI in service of learning.

Finally, by coordinating evidence and adjustments across programming, chemistry, physics, and introductory engineering, the FLC advanced RQ3 by using shared language and timing to support student learning, motivation, engagement, and retention.

*Adjustment of Expectations and Equity-Oriented Reframing*

A recurring theme in FLC discussions was the need to revisit assumptions about what students already knew when entering the first-year sequence. Students arrived with varied levels of preparation in areas such as calculator use, unit conversion, and scientific notation. These skills were sometimes treated as prerequisites but that not all students had been equally exposed to. The FLC responded by distributing responsibility for these foundational competencies across courses rather than assuming any single prior course had covered them. Instructors also became more intentional about explaining disciplinary jargon and scientific terms explicitly, rather than relying on passive exposure. This shift acknowledged that students' past educational experiences were diverse and that equitable instruction required making implicit expectations visible.

The Week 1 survey data reinforced this reframing. The prevalence of affordability, proximity, and family-first reasoning among students' motivations for attending the institution suggested that many students were making decisions under constraint. The FLC discussed how acknowledging these realities—rather than treating all students as though they had chosen from among equally accessible options—fostered a more empathetic and accurate understanding of the student population. This did not mean lowering standards; rather, it meant coupling high expectations with the scaffolding, transparency, and care needed for all students to meet them.

*Assumptions about Students and Reframing through the FLC*

Early FLC discussions surfaced common faculty assumptions, including the beliefs that students are not interested or that many primarily seek shortcuts with artificial intelligence tools. Evidence collected through the Week 1 survey, the midsemester micro-survey, and PyPal interactions suggested a different picture. Students expressed positive intent, care about learning, and a desire to understand concepts, even when they felt uncertain or discouraged after midterm grades. PyPal queries were often requests for conceptual explanation, debugging guidance, and help planning multi-step work, indicating an interest in learning how to proceed rather than bypassing the process. The micro-survey data reinforced this: across both courses, the majority of students intended to reduce AI use for quick answers while maintaining it as a study tool, suggesting a self-regulatory awareness that faculty had not initially expected.

This reframing helped the FLC customize instruction to the students they have by pairing transparent expectations with low-friction supports, designing assignments that value explanation and process, and naming productive uses of AI in service of learning.

*How This Changes What Faculty Know about Students*

Throughout the FLC discussions and data collected, the participants learned the following about their students:

Students are motivated to understand and seek conceptual coherence first. The dominance of conceptual explanation requests in PyPal, the prevalence of genuine interest and curiosity as the primary reason students chose their major, and the persistence of motivation even among students earning below 60 at midterm all validate spending more time on explicit concept models, attention to invariants, and brief "why it works" narratives rather than leaning primarily on procedural drill. In ICAP terms, the FLC interpreted this as a need to design for Constructive engagement earlier in the learning sequence, so that students generate understanding rather than only imitate steps [4].

Students want help closing the loop from idea to implementation. Planning and debugging patterns indicate a need for visible bridges among concept, decomposition, and incremental testing. Instruction should deliberately move learners from monitoring and evaluation to regulation and transfer by naming these moves and building them into tasks and criteria, consistent with AMERT [8].

Students respond to transparent expectations. The emphasis on sense-making over answer-generation aligns with environments where purpose, tasks, and criteria are explicit. The FLC therefore continued pairing example-driven criteria with structured notes and short acceptance checks, consistent with TILT and supportive of reduced extraneous load during problem solving [5], [6]. This matters for equity because clearer entry points reduce reliance on implicit disciplinary norms and help more students engage constructively.

Students are self-aware but many face structural barriers to follow-through. Most students in across courses agreed their grades reflected their effort; yet, strong stated intentions, and limited observable behavioral change points to an intent–behavior gap that is not primarily about motivation but about barriers, time management, and unclear pathways to help. This finding reframes the instructional challenge: rather than asking how to motivate students, the FLC shifted toward asking how to make it easier for already-motivated students to act on their intentions.

**Limitations and Next Steps**

Analyses remained descriptive and centered on community functioning and instructional decision-making. Sample sizes varied across instruments (60 for the Week 1 survey, 74 for physics, 87 for chemistry), and the micro-surveys were voluntary and anonymous, which limits the ability to link responses across courses or to track individual students over time. The intent–behavior gap was identified through faculty observation rather than systematic measurement, and finer-grained tracking would strengthen this finding.

The next phase of this work will transition to a full research study. We will develop an Institutional Review Board protocol and continue to design and collect data about the FLC to examine how faculty reasoning, collaboration, and instructional choices evolve over time. A key priority is to build mechanisms that better capture the intuitive questions, classroom observations, and informal noticing that instructors naturally generate but rarely document. Future work will refine the three time-point evidence strategy while adding structures that make instructor noticing visible, traceable, and analytically meaningful. Future iterations should also explore linked, longitudinal designs that connect student affect and intentions at midterm to end-of-semester outcomes across courses.

**Conclusion**

A small FLC can improve coherence in the first-year experience by treating student voice and classroom evidence as design requirements and by privileging feasible, well-named changes. As instructors share artifacts and iterate in synchrony, students encounter clearer invitations into disciplinary thinking and more accessible routes to practice and feedback. Grounding interpretation in ICAP supports cross-course alignment around engagement quality, while transparency supports equitable access to that engagement by clarifying how students can participate productively. The cross-course comparisons made possible by the FLC, in particular, the finding that the same students experienced markedly different levels of success and affect across physics and chemistry, underscore the value of coordinated instructional design in the first-year sequence.

**References**

[1] E. Wenger, Communities of Practice: Learning, Meaning, and Identity. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[2] E. Wenger, R. McDermott, and W. M. Snyder, Cultivating Communities of Practice: A Guide to Managing Knowledge. Boston, MA: Harvard Business School Press, 2002.

[3] S. A. Barab and K. D. Squire, "Design-Based Research: Putting a stake in the ground," Journal of the Learning Sciences, vol. 13, no. 1, pp. 1–14, 2004.

[4] M. T. H. Chi and R. Wylie, "The ICAP framework: Linking cognitive engagement to active learning outcomes," Educational Psychologist, vol. 49, no. 4, pp. 219–243, 2014.

[5] M.-A. Winkelmes, A. Boye, and S. Tapp, Transparent Design in Higher Education Teaching and Leadership. New York, NY: Routledge, 2019.

[6] J. Sweller, P. Ayres, and S. Kalyuga, Cognitive Load Theory. New York, NY: Springer, 2011.

[7] R. M. Ryan and E. L. Deci, "Self-Determination Theory and the facilitation of intrinsic motivation, social development, and well-being," American Psychologist, vol. 55, no. 1, pp. 68–78, 2000.

[8] J. Willison, C. Draper, L. Fornarino, M. Li, T. Sabri, Y. Shi, and X. Zhao, "Metacognitively ALERT in science: A hierarchical framework for metacognition and preliminary evidence of its viability," Studies in Science Education, vol. 60, no. 2, pp. 153–189, 2024.

**Appendix: Selected Survey Results**

**Table A1.** Midterm grade distributions by course.

| Grade Range | Physics (n = 74) | Chemistry (n = 87) |
|---|---|---|
| 100–90 | 25 (34%) | 6 (7%) |
| 90–80 | 13 (18%) | 16 (18%) |
| 80–70 | 15 (20%) | 15 (17%) |
| 70–60 | 13 (18%) | 17 (20%) |
| Below 60 | 8 (11%) | 33 (38%) |

**Table A2.** Affect at midterm: percentage agreeing or strongly agreeing.

| Affect Item | Physics (n = 74) | Chemistry (n = 87) |
|---|---|---|
| Proud | 46% | 33% |
| Motivated | 77% | 55% |
| Discouraged | 12% | 30% |
| Relieved | 43% | 30% |
| Puzzled/Surprised | 22% | 32% |
| Accurate reflection of effort | 54% | 49% |

**Table A3.** Intent–behavior gap: selected indicators by course.

| | Physics | Chemistry |
|---|---|---|
| Report never attending office hours | 70% | 48% |
| Of non-attenders, intend to do more | 62% | 55% |
| Intend to reduce AI for quick answers | 43% | 46% |
| Expect higher final grade than midterm | 36% | 61% |