

ARTICLE

Investigating Patterns of Student Engagement during Collaborative Activities in Undergraduate Chemistry Courses

Received 00th August 2021,
Accepted 00th xx 20xx

DOI: 10.1039/x0xx00000x

Joshua W. Reid^a, Zubeyde Demet Kirbulut Gunes^b, Shaghayegh Fateh^a, Adan Fatima^e, Michael Macrie-Shuck^c, Hannah T. Nennig^d, Fabrizio Quintanilla^e, Nicole E. States^d, Ahmad Syed^e, Renee Cole^d, Gregory T. Rushton^a, Lisa Shah^e, and Vicente Talanquer^{c*}

Several studies have highlighted the positive effects that active learning may have on student engagement and performance. However, the influence of active learning strategies is mediated by several factors, including the nature of the learning environment and the cognitive level of in-class tasks. These factors can affect different dimensions of student engagement such as the nature of social processing in student groups, how knowledge is used and elaborated upon by students during in-class tasks, and the amount of student participation in group activities. In this study involving four universities in the US, we explored the association between these different dimensions of student engagement and the cognitive level of assigned tasks in five distinct general chemistry learning environments where students were engaged in group activities in diverse ways. Our analysis revealed a significant association between task level and student engagement. Retrieval tasks often led to a significantly higher number of instances of no interaction between students and individualistic work, and a lower number of knowledge construction and collaborative episodes with full student participation. Analysis tasks, on the other hand, were significantly linked to more instances of knowledge construction and collaboration with full group participation. Tasks at the comprehension level were distinctive in their association with more instances of knowledge application and multiple types of social processing. The results of our study suggest that other factors such as the nature of the curriculum, task timing, and class setting may also affect student engagement during group work.

Introduction

A major focus of chemistry education in recent years has been on helping students develop a more solid and integrated understanding of central ideas, core practices, and ways of thinking in the chemical sciences (National Research Council, 2012). Several reports have highlighted the advantages of “active learning” strategies in supporting this type of learning through meaningful student engagement in course activities (National Research Council, 2012; Freeman *et al.*, 2014; Järvelä and Renninger, 2014; Theobald *et al.*, 2020). However, active learning is a broad and ill-defined construct that means different things to different people in various domains (Lombardi *et al.*, 2021). This often leads to a lack of fidelity in the implementation of active learning strategies (Stains and Vickrey, 2017), resulting in differences in student engagement and performance. In active learning environments, student engagement

has often been used as a metric of quality in terms of student participation and interactions (Kahu, 2013), and has been linked to positive learning outcomes (Sinatra *et al.*, 2015). Several factors such as the nature of in-class tasks, student-teacher interaction, and organization of the learning environment have previously been identified as influencers of student engagement (Zepke and Leach, 2010; Groves *et al.*, 2015). Previous research has highlighted that effective activity design is critical for fostering high-quality engagement of students with the task as well as their peers and instructors in active learning environments (Chi and Wylie, 2014; Roberson and Franchini, 2014; Lombardi *et al.*, 2021). However, designing productive tasks is challenging for teachers and instructors at all educational levels. More research is needed to identify critical features of effective task design and implementation that support the productive engagement of students in different contexts. Thus, this study was designed to explore and characterize student engagement in different learning environments where students were expected to actively engage in a variety of classroom tasks. Our goal was to better characterize differences in student engagement in diverse class settings and explore how this was affected by the expected cognitive level of in-class activities. Accordingly, this study addresses the following research questions:

1. What patterns of student engagement characterize different college general chemistry learning environments?

^aTennessee STEM Education Center, Middle Tennessee State University, Murfreesboro, TN 37131, USA

^bDepartment of Mathematics and Science Education, Harran University, 63300, Sanliurfa, Turkey

^cDepartment of Chemistry and Biochemistry, University of Arizona, Tucson, AZ 85721, USA.

^dDepartment of Chemistry, University of Iowa, Iowa City, IA 52242, USA

^eDepartment of Chemistry, Stony Brook University

*Corresponding author: vicente@arizona.edu

2. How does the expected cognitive level of in class-activities affect student engagement in different college general chemistry learning environments?

Looking into Classrooms

Several calls for reforming undergraduate chemistry education recommend the adoption and use of active learning instructional techniques (National Research Council, 2012; Freeman *et al.*, 2014; Theobald *et al.*, 2020). In their highly cited meta-analysis, Freeman *et al.* (2014) found that active learning environments increased student performance compared to students in lecture-based courses. Students enrolled in traditional course environments were 1.5 times more likely to fail than students in active-learning courses. In addition to improving student performance, active learning can help narrow achievement gaps among underrepresented student groups. Theobald *et al.* (2020) found the achievement gap between underrepresented groups of students and well-represented students was reduced by 33% when students engaged in active learning.

While different studies have demonstrated the benefits of so-called active learning strategies and learning environments, the construct of “active learning” has been contested as being ill-defined (Lombardi *et al.*, 2021). Different interpretations of what active learning is or how it is effectively implemented can lead to inconsistencies in teaching practice that have differential effects on student engagement and performance (Stains and Vickrey, 2017). For example, fidelity in the implementation of active learning strategies is a significant moderator in reducing achievement gaps among underrepresented student groups (Theobald *et al.*, 2020).

Several factors have been identified as mediating the effect of active learning environments on student outcomes. These include: task design and implementation, frequency of active learning opportunities, course structure, and fidelity of implementation (Theobald *et al.*, 2020; Lombardi *et al.*, 2021). Student engagement in group tasks is not only affected by course design and instruction but also by group composition and individual students’ background and beliefs about how to learn (Deslauriers *et al.*, 2019; Hancock *et al.*, 2019; Liyanage *et al.*, 2021). Students’ buy-in for reformed teaching practices has a significant impact on their engagement (Prather *et al.*, 2009). Students who believe that they learn best from lecture are less likely to participate in active learning activities (Deslauriers *et al.*, 2019). Additionally, the personal and societal relevance of course work may also affect student engagement (Hancock *et al.*, 2019). Thus, there is growing evidence that the creation of active learning opportunities is necessary, but often not sufficient for promoting productive student engagement and fostering meaningful learning (Stains and Vickrey, 2017).

Student engagement

Lampert *et al.* (2009) described learning as an interaction between individuals mediated by an intellectual or social activity. Interactions allow individuals to refine their understandings. From this perspective, learning environments must be designed to foster opportunities for students to socially engage with each other. This is often accomplished through collaborative activities in which students work together, typically in organized groups, to complete a

task. In these environments, student engagement often refers to the degree of participation, attention, and intellectual involvement of students in a group while completing assigned tasks (Kuh, 2009). It is expected that higher levels of engagement will result in higher levels of learning and stronger student performance overall.

Student engagement has been conceptualized as a multifaceted and complex construct to help explain student outcomes (e.g., persistence, success, achievement; Kahu, 2013) and as such is often considered a proxy for quality of student participation (Kuh, 2009). Scholars have characterized student engagement through behavioral (e.g., collaborative participation in learning activities), emotional (e.g., presence of interest), cognitive (e.g., engaging in strategy use), and agentic (e.g., engaging in constructive contribution to the instruction) aspects (Fredricks *et al.*, 2004; Reeve and Tseng, 2011). This body of research suggests at least three different dimensions of analysis in the characterization of student engagement in classroom activities: degree or amount of participation (Pike *et al.*, 2011), knowledge use and elaboration (knowledge dynamics) (Ford, 2008; van Last, 2009), and the nature of the social interactions among students (social processing) (Kumpulainen and Kaartinen, 2003).

Pike *et al.* (2011) have highlighted the contingent relationship between student engagement and active participation in learning activities. High levels of participation in classwork seem to be associated with higher levels of cognitive engagement and increased use of higher-order thinking (Pike, 1999; Zhao and Kuh, 2004). Increased student participation often results in and manifests through a larger number of student-student and student-instructor interactions (Pike, 1999; Inkelas *et al.*, 2003). Thus, the analyses of these different types of interactions can be used as a first measure of student engagement in active learning environments.

Researchers in the area of knowledge management in groups (Nonaka *et al.*, 2006; Ford, 2008) conceptualize knowledge as a dynamic process emerging from human interaction and consider it important to characterize how knowledge is shared, used, and created during collaborative activity (knowledge dynamics). In a community engaged in group work, knowledge can be transmitted between people (knowledge sharing) to facilitate the completion of tasks. Knowledge sharing involves the introduction of information and ideas without much attention to their origin, interpretation, or evaluation (van Last, 2009). Group members can also engage in the application of shared knowledge in a rather systematic manner (knowledge application) to achieve task goals. Individuals or groups can also construct understandings as they work on an activity by engaging in the interpretation and evaluation of information, sharing, testing, and critiquing ideas, and actively seeking to make sense of situations and problems (knowledge construction). Knowledge construction is often associated with deep learning as it may lead to significant changes in knowledge structure and approaches to problem solving and decision making (Biggs, 1987).

The nature of student-student interactions in a group may affect student participation and knowledge dynamic (Forman, 1989). Thus, Kumpulainen and Kaartinen (2003) consider it important to characterize the social relationships and types of participation in peer groups (social processing) when analyzing the complex dynamics of group work. These authors have identified different modes of social processing in groups, including collaborative (participants attempt to reach a common understanding of a

problem or situation), tutoring (some participants assist others in comprehending or completing a task), individualistic (participants work individually on an activity), domination (an individual or group of individuals direct the work of others), and confusion (group members express lack of understanding of the task or associated concepts). The emergence of different types of social processing is affected by group and task characteristics. For instance, these authors found that domination was a key social process exhibited in student groups when students did not have a shared understanding of a task or the solution to a problem.

Learning environments and task design

It is expected that the nature of the learning environment and in-class tasks used to scaffold and foster student understanding will affect student engagement during group work (Lombardi *et al.*, 2021). Several types of pedagogies of engagement (e.g., collaborative group work, POGIL, project-based learning, peer-led team learning) have been used in undergraduate science and chemistry courses to create more student-centered learning environments (Eberlein *et al.*, 2008). The goal of these environments is to promote student active engagement with the content and support students in constructing knowledge about central ideas (Järvelä and Renninger, 2014; Arthurs and Kreager, 2017).

While didactic approaches to teaching tend to organize instruction around the presentation of disciplinary content, instructors in student-centered environments often orchestrate instruction around a sequence of tasks that help learners to develop more expert ways of reasoning and acting in a domain. Effective task design and implementation are thus critical for fostering meaningful learning in these environments (Roberson and Franchini, 2014). Learning tasks are the vehicle through which understanding is expected to develop as students analyze information, apply and construct ideas, make decisions, and build arguments and explanations (Doyle and Carter, 1984).

Teachers and instructors at all educational levels struggle to design and implement tasks that engage students cognitively. Many of them use classroom activities based on passive modes of engagement in which learners receive information from diverse types of instructional materials (e.g., reading a text or watching a video without doing anything else) (Chi, 2009). Instructors who introduce more active elements in their classrooms often create tasks with low cognitive demand. Some, for example, simply ask students to physically manipulate different resources (e.g., copy solution steps or underline text in a reading). Other instructors engage students in answering questions, but the questions posed are frequently designed to test whether students can remember a definition, apply a formula, or reproduce a schema.

Research has shown that students benefit more from participating in “constructive” activities that have two basic characteristics: a) lead to the production of outputs, and b) these outputs are not presented in the learning materials (Chi and Wylie, 2014). Examples of these types of activities include self-explaining (Chi *et al.*, 1994), drawing concept maps (Biswas *et al.*, 2005), comparing and contrasting cases (Schwartz and Bransford, 1998), drawing analogies (Chinn and Malhorta, 2002), and making predictions (Klahr and Nigam, 2004). Learning benefits are enhanced

when the construction of new outputs involves productive interactions with others (Chi, 2009). These types of “interactive activities” often involve participating in instructional dialogues with more knowledgeable others or in joint dialogues with peers. Within instructional dialogues, learners may participate in guided-construction activities such as responding to scaffolded questions and revising errors from corrective feedback. Within joint dialogues, students could co-construct ideas through arguing and defending a position or building and elaborating on a partner’s contribution.

Wang *et al.* (2019) found that the cognitive level of tasks (e.g., descriptive, relational, reasoning) can differentially affect students’ knowledge acquisition. The cognitive level of a task determines the mental operations or thinking skills that are likely to be deployed to complete the activity. For instance, recall or retrieval tasks require lower-order cognitive skills such as memorization to be completed. In contrast, analysis tasks require higher-order cognitive skills such as interpreting information and applying knowledge (Zoller, 1993; Crowe *et al.*, 2008). Different frameworks have been developed to characterize the cognitive level of tasks, such as Bloom’s (Anderson *et al.*, 2001) and Marzano’s (Marzano and Kendall, 2007) taxonomies. In this latter framework, the cognitive system is assumed to engage at four major sublevels: a) retrieval, which involves the activation and transfer of knowledge from permanent memory to working memory; b) comprehension, which requires integrating information and creating symbolic representations (linguistic or imagery); c) analysis, which may involve identifying differences and similarities, organizing knowledge into meaningful categories, analyzing errors, generalizing, and transferring, and d) knowledge utilization, which demands decision making, problem solving, experimenting, and investigating.

Methods

The study described in this paper is part of a larger project involving four research sites across the United States. The overarching goal of this project is to characterize the features of tasks and facilitation, including design and implementation, in collaborative learning environments that promote productive student engagement and, by extension, meaningful learning. In this contribution, however, we focus our attention on the analysis of the impact of one aspect of in-class task design (expected cognitive level of the task) on student engagement across five different learning environments.

Research settings and data collection

Data were collected in five different learning environments across four different universities. Key features of each environment are described in the text descriptions below and summarized in Table 1. All data collection was approved by the Institutional Review Board (IRB) at each institution (approved protocol numbers are listed in Table 1) and written consent was obtained from all participants.

Stony Brook University (SBU): At SBU, data were collected from two of eight first-semester General Chemistry I discussion sections, each with an enrollment of ~150 students. These discussion sections were a co-requisite offering to a large-lecture General Chemistry I course with approximately 1100 students across two sections. Students within this lecture course completed graded online homework assignments and took 3 midterm exams and one

Table 1. Key features of each of the five learning environments observed as part of this study.

Research Site (IRB Protocol Number)	Learning Environment	Identifier	Class Size	Number of Groups Observed	Number of Students Observed	Number of Analyzed Tasks (unique questions)	Number of Episodes of Student Engagement
Stony Brook University (917004)	Discussion (POGIL)	SBU-D	~150	6	18	84	128
University of Iowa (201309825)	Discussion (Traditional)	UI-D	~24	2	6	72	86
	Lecture (Traditional)	UI-L	250	2	6	62	104
Middle Tennessee State University (19-2253)	Lecture (POGIL)	MTSU-L	24	2	9	87	178
University of Arizona (1905584616)	Lecture (Chemical Thinking)	UA-L	220	10	45	114	517
TOTALS				22	84	419	1,013

cumulative final exam that included multiple choice items. Co-requisite discussion sections met once per week for 80 minutes for 15 weeks. The course followed a traditional curriculum, but students engaged in POGIL activities where they worked collaboratively on questions and problems related to lecture earlier in the week under the supervision of a graduate teaching assistant (TA). Students used an online platform to access and respond to activity questions and received immediate electronic feedback about the accuracy of their responses. These responses were not graded for completion or accuracy. Rather, students were graded on whether they actively worked on these activities as a group for the duration of the 80-minute session. Six student groups with 3-4 students per group were selected for this study. Data were collected from three discussion sessions, which included 84 different tasks. Video recordings of whole class interactions and observational field notes were collected during the discussion periods.

University of Iowa (UI): At UI, data were collected from a first-semester introductory chemistry course with an enrollment of ~700 students and consisting of both a discussion and lecture component. Students within this course completed graded online homework assignments and took 3 midterm exams that included multiple-choice and free response questions, and a final exam that included only multiple-choice questions. Both components of this course were analyzed in this study. Discussion sections led by graduate teaching assistants (each with a typical enrolment of 20-26 students) met for 50 minutes each week for 15 weeks in a classroom with seven square tables. Two groups of students from two different discussion sections, each composed of three students, were selected for investigation. During the discussion students worked in their groups to complete a guided inquiry activity that addressed lecture material for a given week through the lens of real-life phenomena. These real-life topics were given as pre-class readings or pre-class video assignments to ensure everyone had some background on the topic before working on the activities. Activities were turned in by student groups, but no grading was completed. Audio data of student collaboration and written data of students' work were collected and analyzed. The groups' audio and written work were collected using a white board app on an iPad. Video recordings of whole class interactions were also collected.

There were two lecture sections for this course, one with 250 students and one with 450 students. In the lecture section, students

attended the course three times a week for 50 minutes for 15 weeks in a large auditorium with stadium style seating. Two groups of students from the smaller lecture section, each composed of three students, were selected for this study. The course was team-taught by three instructors who rotated in the classroom across semester topics. Three graduate teaching assistants were present in the classroom to aid with answering student question during questions. The classroom lecture included a conventional curriculum, during which each instructor lectured from the front of the room using PowerPoint slides. Students were periodically asked to answer questions using a student response system after example problems or explanations were discussed. These questions were graded by completion and if students actively participated in all problems before an exam, they received bonus points for that exam. Data was collected in a similar way as the University of Iowa discussion section (see above) with the addition of observational field notes.

Middle Tennessee State University (MTSU): At MTSU, data were collected from a first-semester General Chemistry I lecture course with an enrollment of 24 students that met two times a week for 90 minutes for 15 weeks. Two student groups (4-5 students per group) were observed in this study. This was a conventional curriculum taught through POGIL activities. Each student received a hard copy of the POGIL activity to work on and each group had an iPad for reporting the answer for grading purposes. Students within this course completed graded online homework assignments and took 12 weekly tests that included free response and multiple-choice questions. The small group discussions were video recorded and written work conducted on the iPad was recorded to capture students' discussions and the writing of their group responses.

University of Arizona (UA): At UA, data were collected from a first-semester General Chemistry I lecture course with an enrollment of 220 students. This class met 3 times a week for 50 minutes for 15 weeks in a collaborative learning space with 60 tables. Students within the course completed online reading assignments before every class and weekly online homework. They took 4 midterm exams that included multiple-choice and free-response questions, and a final that only included multiple choice questions. Eight learning assistants were present in the classroom to facilitate student learning. Ten student groups (3-4 students per group) were

selected to participate in the study. This course followed an alternative curriculum (Chemical Thinking) that actively engages students in constructing and applying chemical concepts and ideas to analyze, discuss, reflect upon, and propose reasonable explanations and solutions to relevant problems and phenomena (Talanquer and Pollard, 2010). In-class tasks were interspersed with mini-lectures and whole-class discussions every lecture session. In most activities, students were expected to work collaboratively with their groups but submit individual responses through a response system when task time was over. Data was collected from 26 class sessions, which included 114 different tasks. Each group was assigned a camera and two audio recorders and recorded for the whole duration of a class.

Data Analysis

The project team developed a codebook to characterize the expected cognitive level of tasks and the nature of student engagement across all learning environments. The codes were adapted from previously established frameworks, described and referenced in a prior section of this paper, that were refined through their application to several tasks and video/audio recordings of students working in small groups. The final codebook included four variables, each with several codes. One of the variables (task level) was associated with task characteristics, and the other three (knowledge dynamics, social processing, and amount of participation) helped us characterize student engagement. Once the larger project codebook had been established, the research team discussed issues related to specific code definitions. The research team worked together to apply these codes to several tasks and episodes of student engagement across all sites. Application of codes were compared until a consensus had been reached for all codes.

Task characteristics. To characterize the tasks across the different learning environments, data from all sites was compiled for analysis. As summarized in Table 1, 419 unique tasks were analyzed across all learning environments. Each of these activities was evaluated for the expected cognitive level of the task (task level) using Marzano's taxonomy (Marzano and Kendall, 2007). While Marzano's taxonomy includes four cognitive levels (1: Retrieval, 2: Comprehension, 3: Analysis, 4: Knowledge Utilization), none of the tasks for which student discourse was recorded were classified as requiring Level 4. While a few discussion activities included questions at this level toward the end of the worksheet, students did not complete these questions during class time. See Appendix A for specific examples of tasks at each of these levels.

Student engagement. To investigate patterns of student engagement during collaborative learning tasks, we analyzed group behaviors and interactions for each identified episode. We defined an episode as a period of time in which a small student group worked to complete a single task. Each episode was characterized based on its knowledge dynamics, social processing, and amount of student participation as described below. For a more complete description and exemplification of each code see Appendix A.

- **Knowledge dynamics:** Four main codes were applied to capture how students use knowledge as a group to complete a task: knowledge sharing, knowledge application, knowledge construction, and not observable. Knowledge sharing was used when students primarily shared information and ideas to

complete the task without much discussion of ideas among peers. Knowledge application was used when students applied formulas or concepts to answer the question and explained their approach. Knowledge construction was used when groups of students engaged in conversations that led them to answer a question or solve a problem by critiquing or building on each other's ideas. A not observable code was applied to episodes in which there were no observable student interactions.

- **Social processing:** The codes applied in this dimension of analysis included: individualistic, confusion domination, tutoring, and collaboration. The individualistic code was used when students worked independently on an activity. Confusion was used when students explicitly manifested doubts on how to proceed or about their understanding of relevant concepts. The domination code was used when a single group member answered or completed the task without significant input from others. Tutoring was used for instances of dialogic interaction between group members in which one individual mainly answered the questions from others. The collaboration code was used when students collectively worked on a task. The non-interactive code was applied when there was no interaction or evidence of individual student work on a task. Multiple types of social processing were observed in some episodes and thus a "multiple social processing" code was applied to these cases.
- **Amount of participation:** Each episode of student engagement was also coded for the fraction of students in a group that participated in completing the task. Three levels of participation were coded: minimal, partial, and full. Minimal participation occurred when students worked individually or only one student was explicitly engaged. Partial participation referred to groups that had two or more (but not all) students interacting to work on the task. Full participation referred to every group member engaging in the discussion.

Out of the total number of 1,091 episodes, 1,013 episodes that could be coded in all areas (i.e., task level, knowledge dynamics, social processing, and amount of participation) were selected for further analysis in our study. Episodes were excluded from this study if they did not have audio due to technical issues ($n=14$) or researchers were not able to capture all variables of interest for an episode ($n=64$). As shown in Table 1, the number of episodes varied between learning environment due to differences in total number of tasks and groups observed. Due to the nature of the variables under analysis (categorical and not normally distributed) non-parametric statistical analyses were performed. Patterns in task characteristics and student engagement were identified and compared across the five different types of learning environments. Chi-square tests of independence were performed using R studio software v1.2.5033 (RCoreTeam, 2021) to identify significant differences between these environments. Furthermore, we used Chi-square tests of independence to determine whether task level was significantly associated with student engagement variables. Due to the number of levels within each variable, a Bonferroni correction was applied for each comparison. Post-hoc analysis of residuals was used to identify major contributors to statistical significance. Items with standardized residuals greater than 2 occurred at frequencies higher-than-expected, while items with values less than -2 occurred at frequencies lower-than-expected (Appendix B includes the results from these analyses).

Results

In this section, we first present a descriptive analysis of differences in expected cognitive level of task and student engagement (knowledge dynamics, social processing, amount of participation), as well as the interaction between these variables in each of the learning environments observed in our study. We look at similarities and differences between settings to better characterize the effect of various factors (e.g., classroom setting, curriculum) on student engagement. We follow this description with a summary and discussion of key findings.

Task level

As shown in Table 2, comprehension tasks were the most common in each of the observed learning environments, with the exception of UA-L where analysis tasks comprised more than half of observed in-class activities. No knowledge utilization tasks were enacted in any of the sites. A Chi-square test of independence revealed that there were significant differences in the relative numbers of tasks at each level of understanding observed in the different learning environments ($\chi^2 = 108.9$, $df = 8$, $p\text{-value} < 0.001$). Our analysis indicated (see Appendix B) that the significant difference mostly stemmed from the higher-than-expected number of tasks at the analysis level at UA-L compared to MTSU-L, where comprehension tasks were predominant, and the higher-than-expected number of retrieval tasks at both UI-D and UI-L compared to UA-L.

Knowledge dynamics

The analysis of knowledge dynamics in the different episodes revealed knowledge sharing to be the most common dynamic observed in each of the observed classes (see Table 3). Nevertheless, statistical analysis of these data revealed a significant dependence of knowledge dynamic on the learning environment ($\chi^2 = 402.14$, $df = 12$, $p\text{-value} < 0.001$). Our analysis showed (see Appendix B) that the significance mostly stemmed from higher-than-expected numbers of

Table 2. Frequency of tasks by task level and learning environment.

Task Level	SBU-D	UI-D	UI-L	MTSU-L	UA-L
Retrieval	7 (8.3%)	22 (30.5%)	19 (30.6%)	17 (19.5%)	5 (4.4%)
Comprehension	62 (73.8%)	40 (55.6%)	28 (45.2%)	70 (80.5%)	50 (43.9%)
Analysis	15 (17.9%)	10 (13.9%)	15 (24.2%)	0 (0.0%)	59 (51.7%)
Total	84	72	62	87	114

Table 3. Knowledge dynamics in observed discursive/interactive episodes across the different learning environments.

Knowledge Dynamic	SBU-D	UI-D	UI-L	MTSU-L	UA-L
Not Observable	5 (3.9%)	3 (3.5%)	42 (40.4%)	23 (12.9%)	0 (0%)
Knowledge Sharing	70 (54.7%)	55 (64.0%)	58 (55.8%)	106 (59.6%)	391 (75.6%)
Knowledge Application	52 (40.6%)	18 (20.9%)	1 (1.0%)	28 (15.7%)	19 (3.7%)
Knowledge Construction	1 (0.8%)	10 (11.6%)	3 (2.9%)	21 (11.8%)	107 (20.7%)
Total	128	86	104	178	517

“not observable” episodes in the groups from UI-L, “knowledge sharing” episodes in groups from UA-L, “knowledge application episodes” in both SBU-D and UI-D, and “knowledge construction” episodes in UA-L.

We also analyzed the knowledge dynamics in the different student groups in the various learning environments in relation to the level of the task in which they were engaged. Major trends in this area can be observed in Figure 1 where we represent the relative frequency of each type of knowledge dynamic observed when

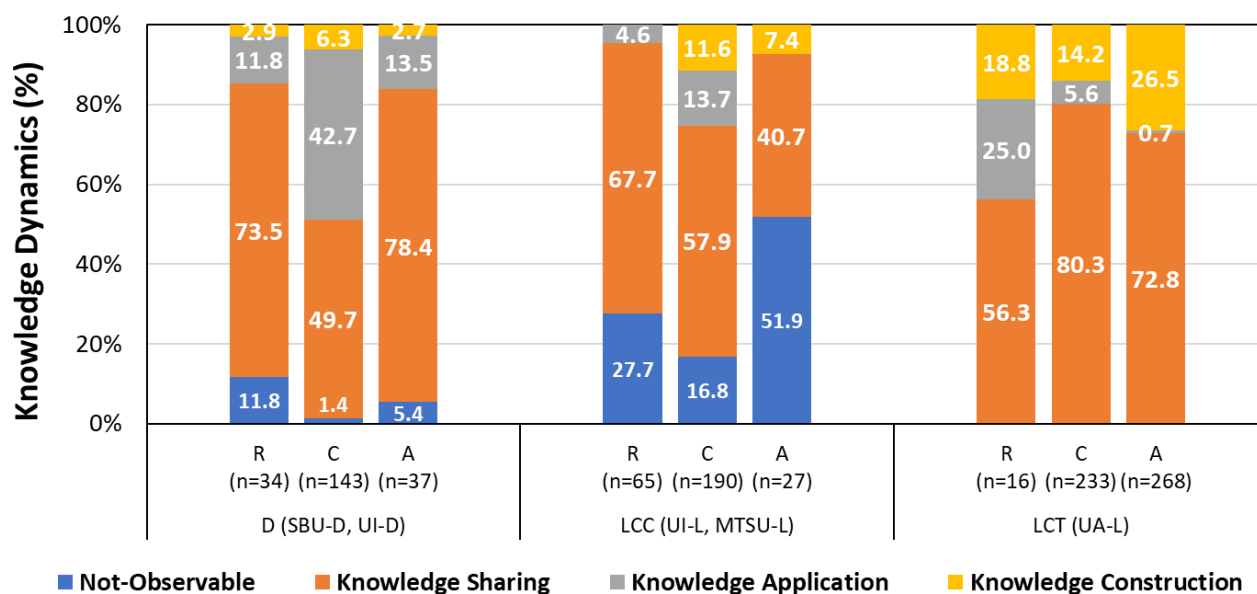


Figure 1. Distribution of knowledge dynamics by task level, Retrieval (R), Comprehension (C), and Analysis (A), in the three categories of sites: Discussions (D), Lecture with Conventional Curricula (LCC), and Lecture using Chemical Thinking (LCT).

students worked on tasks targeting different levels of understanding. In this graph, we have grouped the different learning environments into three different types: Discussions (D) including the SBU-D and UI-D sites, lectures using a conventional general chemistry curriculum (LCC), which included the classes UI-L and MTSU-L, and the lecture at UA-L using the alternative chemical thinking curriculum (LCT). Statistical analysis of these data indicated a significant difference in the knowledge dynamics for different level tasks that was distinct in each of the categories of sites. In particular, there were a higher-than-expected number of “knowledge application” episodes associated with comprehension tasks in the discussion classes (UI-D, SBU-D), a higher-than-expected number of “not observable” episodes during analysis tasks in LCC sites (particularly at UI-L), and a higher-than expected number of “knowledge construction” episodes and a lower-than-expected number of “knowledge application” cases for analysis tasks at UA-L.

Social processing

Analysis of social processing interactions in the different sites revealed diverse patterns in each of the observed learning environments as summarized in Table 4. Collaboration and Multiple Social Processing were, however, often the most prevalent forms of social processing in most sites. Many episodes were characterized by more than one social processing interaction such as tutoring and confusion or tutoring and domination. Chi-square analysis of the data revealed a significant association between learning environment and social processing ($\chi^2 = 813.1$, $df = 24$, $p\text{-value} < 0.001$). This analysis indicated (see Appendix B) that the significance was mostly linked to a) a larger-than-expected number of non-interactive episodes in groups from UI-L, b) a higher-than-expected number of individualistic episodes in groups from UI-L and MTSU-L and lower in SBU-D and UI-D; c) a higher-than-expected number of confusion episodes at UA-L and lower in UI-L and MTSU-L; d) a lower-than-expected number of domination social processing in UA-L; e) a lower-

Table 4. Social processing in observed discursive/interactive episodes across the different learning environments.

Social Processing	SBU-D	UI-D	UI-L	MTSU-L	UA-L
Non-interactive	0 (0.0%)	0 (0.0%)	50 (48.1%)	0 (0.0%)	0 (0.0%)
Individualistic	3 (2.3%)	5 (5.8%)	29 (27.9%)	42 (23.6%)	64 (12.4%)
Confusion	5 (3.9%)	2 (2.3%)	0 (0.0%)	0 (0.0%)	53 (10.3%)
Domination	25 (19.5%)	16 (18.6%)	10 (9.6%)	34 (19.1%)	53 (10.3%)
Tutoring	17 (13.3%)	2 (2.3%)	1 (1.0%)	14 (7.9%)	63 (12.2%)
Collaboration	39 (30.5%)	30 (34.9%)	14 (13.5%)	36 (20.2%)	284 (54.9%)
Multiple Social Processing	39 (30.5%)	31 (36.1%)	0 (0.0%)	52 (29.2%)	0 (0.0%)
Total	128	86	104	178	517

than-expected number of tutoring episodes in UI-D and UI-L; f) the larger-than-expected number of collaboration episodes in UA-L and lower in UI-L and MTSU-L, and g) the larger-than-expected number of episodes with multiple types of social processing in groups from SBU-D, UI-D, and MTSU-L and lower at UI-D and UA-L.

As described in the case of knowledge dynamic, we also analyzed the association between task level and social processing in the observed groups. The results are summarized in Figure 2 where we show the relative frequency of each type of social processing for each task level for the three categories of sites (D, LCC, LCT). Our analysis indicated that instances of multiple social processing increased with task level in the discussion sites, although the effect was not statistically significant. In LCC sites, instances of individualistic work during retrieval tasks and episodes of non-interactive processing

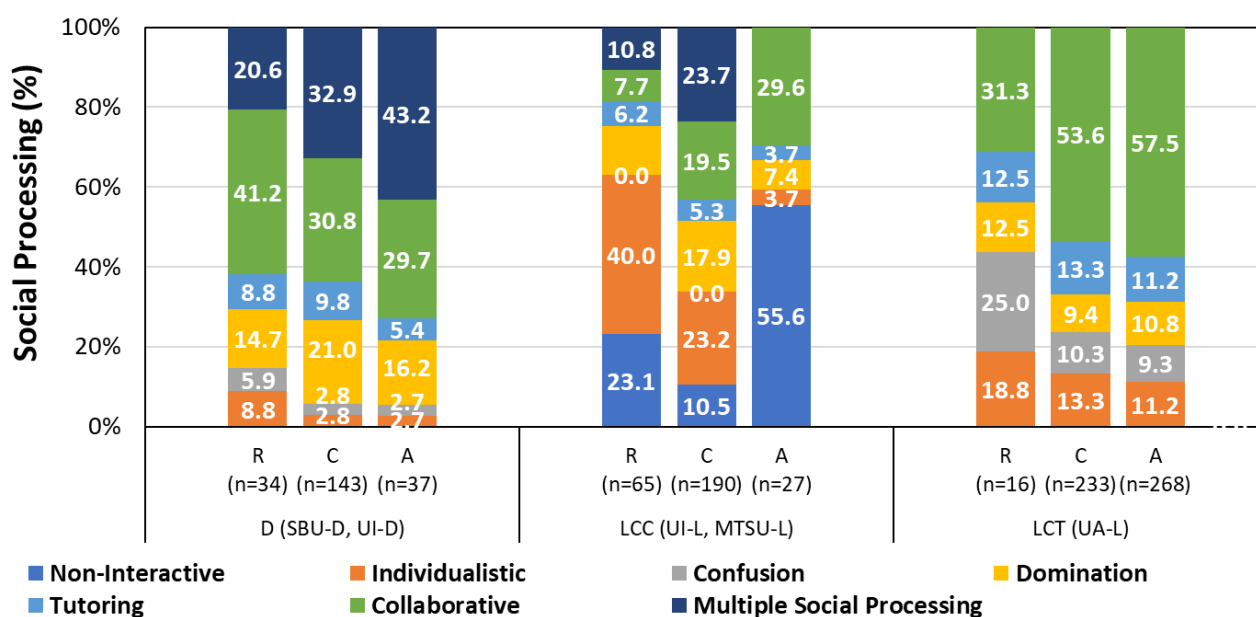


Figure 2. Distribution of social processing by task level, Retrieval (R), Comprehension (C), and Analysis (A), in the three categories of sites: Discussions (D), Lecture with Conventional Curricula (LCC), and Lecture using Chemical Thinking (LCT).

during analysis tasks were higher-than-expected. At the UA-L site, instances of collaborative processing increased with task level although this effect was not statistically significant.

Amount of participation

The analysis of amount of participation revealed that the majority of episodes analyzed were characterized by having student groups fully (SBU-D, UI-D, UA-L) or partially (UI-L, MTSU-L) participating in the assigned tasks. Specific results for each site are summarized in Table 5. Further analysis of these data showed that the amount of participation had a statistically significant association with learning environment ($\chi^2 = 222.3$, $df = 8$, p -value < 0.001). The significance mostly stemmed from a) the higher-than-expected number of instances of full participation at UI-D and UA-L and lower at UI-L and MTSU-L; b) the higher-than-expected number of episodes with partial student participation at MTSU-L and lower at UI-D, and c) the higher-than-expected number of episodes of minimal participation at UI-L and lower at SBU-D and UA-L.

Table 5. Amount of participation in observed discursive/interactive episodes across the different learning environments.

Amount of Participation	SBU-D	UI-D	UI-L	MTSU-L	UA-L
Minimal Participation	0 (0.0%)	2 (2.3%)	36 (34.6%)	19 (10.7%)	22 (4.2%)
Partial Participation	52 (40.6%)	13 (15.1%)	43 (41.4%)	114 (64.0%)	188 (36.4%)
Full Participation	76 (59.4%)	71 (82.6%)	25 (24.0%)	45 (25.3%)	307 (59.4%)
Total	128	86	104	178	517

We also explored whether the amount of student participation was associated with expected task level. Figure 3 shows the relative frequency of different amounts of participation for each task level in

the different categories of sites. The relationship between amount of participation and task level was quite similar in the discussion sites and the UA-L. In both cases, minimal participation was higher-than-expected when working on retrieval tasks. The relationship between these two variables was quite different in the LCC sites where higher-than-expected instances of minimal participation were observed when working on both retrieval and analysis tasks.

Summary and discussion of key findings

Our task analysis showed that comprehension tasks were the most common in the majority of the observed learning environments, except at UA-L where an alternative general chemistry curriculum is followed. Students at this latter site were more frequently engaged in analysis tasks. Differences in task level across sites can be expected to reflect differences in both curriculum and instruction. For both UI-L and UA-L, tasks were primarily in the form of single questions interspersed during lecture. While these two settings had similar deployment of tasks, the overall focus of the questions was very different with UI-L having significantly more retrieval questions and UA-L more analysis questions. In the other three environments (MTSU-L, SBU-D, and UI-D), students completed worksheets consisting of a series of questions, which were typically designed to start with retrieval and comprehension questions to guide students through thinking about a problem or concept. The high degree of scaffolding in these activities led to a higher proportion of lower-level questions. No knowledge utilization tasks were enacted in any of the observed classrooms. These types of tasks typically demand a considerable amount of time to implement and often require out-of-class preparation. They were present in some course materials but were never completed by students in the observed classes, often due to lack of time. The prevalence of lower-level tasks observed in most sites is similar to that identified in the analysis of end-of-chapter questions in traditional general chemistry textbooks (Dávila

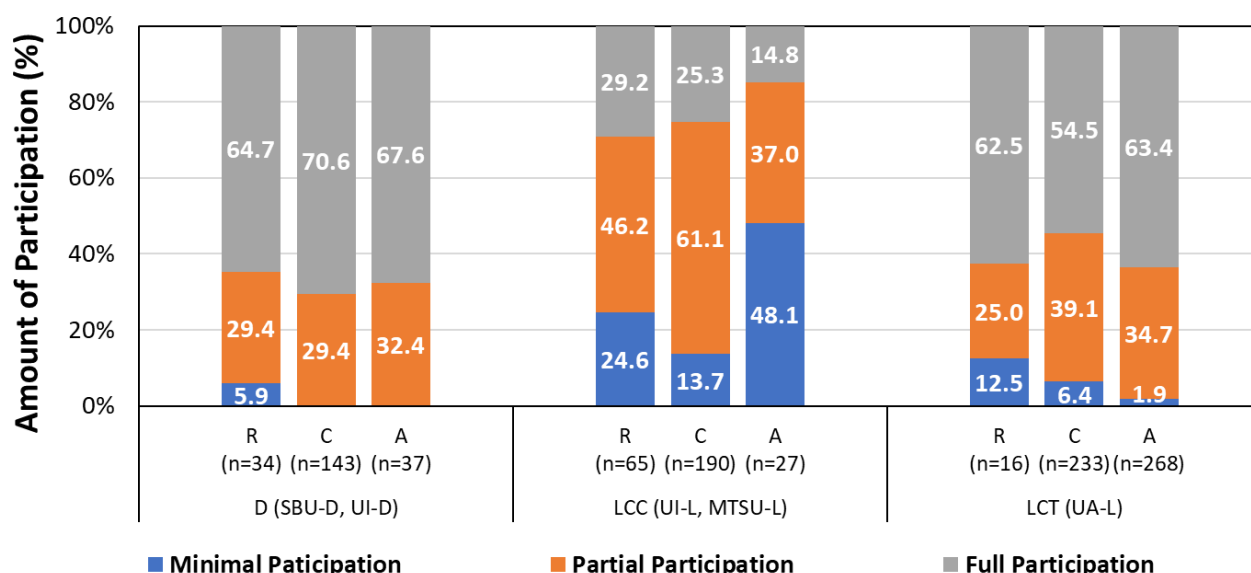


Figure 3. Distribution of amount of student participation by task level, Retrieval (R), Comprehension (C), and Analysis (A), in the three categories of sites: Discussions (D), Lecture with Conventional Curricula (LCC), and Lecture using Chemical Thinking (LCT).

and Talanquer, 2010). Our results thus highlight the impact that conventional curricula and educational resources may have on the opportunities students have to engage in higher-order thinking (Zoller, 1993).

Our analysis revealed a significant association between task level and student engagement as characterized by knowledge dynamic, social processing, and amount of participation of students in observed groups in the five learning environments. Retrieval tasks, which require low levels of cognitive processing for completion, often led to a significantly higher number of instances of no interaction between students and individualistic work, and a lower number of knowledge construction and collaborative episodes with full or partial student participation across all sites. On the other hand, analysis tasks, which require higher cognitive levels, were significantly linked to more instances of knowledge construction and collaboration with full group participation at the UA-L site, although they were linked to higher instances of no interaction at the sites taught using a conventional general chemistry curriculum in auditorium-style classrooms. These results reinforce findings from previous studies that highlight the effect that different types of tasks have on the extent and level of cognitive processing in which students working in groups engage (Chi and Wylie, 2014) and the importance of creating educational experiences that are challenging to students to encourage higher levels of engagement (Zepke and Leach, 2010).

Tasks at the comprehension level in our investigation were distinctive in their significant association with more instances of knowledge application and multiple types of social processing in the discussion sites. We also found that the frequency of knowledge application decreased with analysis questions across all types of sites. While this may seem counter-intuitive, it is likely due to the fact that students tended to explain the procedures used to complete comprehension tasks more often than when completing analysis tasks which demanded more than an algorithmic approach. Knowledge application requires students to connect the conceptual ideas, formulas, and methods explicitly and verbally to their task completion (see Appendix A).

Given the association between task level and student engagement, it is not surprising that the significantly higher fraction of retrieval tasks at UI-L contributed to a higher number of episodes of non-interactive and individualistic work, with lower instances of knowledge construction and full student participation that characterized this latter site. Differences in the levels of the tasks implemented at the UA-L compared to other sites are likely linked to the use of an alternative curriculum (Chemical Thinking) that seeks to foster conceptual understanding over algorithmic problem solving and integrated versus fragmented learning (Talanquer & Pollard, 2010). Reformed curricula in introductory science course developed in recent years strongly emphasize the need for implementing higher-order in-class activities and assessment tasks that demand students to integrate central ideas, scientific practices, and crosscutting concepts (Laverty et al., 2016). Our results support the positive effects on student engagement of curricula aligned with these ideas.

Nevertheless, our results show major differences in all aspects of student engagement for tasks at the same level implemented in the different sites. For example, analysis tasks at UA-L were linked to a

higher number of instances of knowledge construction, collaborative work, and full student participation in this setting. Tasks at the analysis level were associated with significantly more frequent instances of non-interactive processing and minimal student participation at the LCC sites (UI-L in particular). These results suggest that other differences between the observed learning environments, such as the layout of the classroom (e.g., lecture hall versus collaborative learning space), the management of class time, the nature of the instruction, or the scaffolding of tasks likely affected various aspects of student engagement in group work. For example, learning environments in which students worked in small groups on guided inquiry activities for most of the class time (SBU-D, UI-D, MTSU-L) had a significantly higher number of episodes in which students engaged in multiple types of social processing. This was likely due to longer periods of sustained group work and the scaffolded nature of the worksheets compared to the shorter in-class activities implemented at UI-L and UA-L. In these two cases, the class layout may also have been responsible for observed differences in student engagement. While the UA-L class was taught in a flat collaborative learning space designed to facilitate students working in groups of 3 to 4 students, the UI-L class was taught in a traditional auditorium where interactions between more than two students were difficult. This difference in classroom set up likely contributed to the higher number of episodes of no interaction, individualistic work, and minimal participation observed at UI-L. These findings align with those of other authors who have analyzed how the nature of the implementation of evidence-based teaching practices (Stains and Vickrey, 2017) or the type of classroom setting (Cotner et al., 2013; Talbert and Mor-Avi, 2019) affect student engagement.

In our study we also observed that instances of knowledge application were significantly higher in the two learning environments that corresponded to discussion sessions (SBU-D, UI-D). In these cases, students typically worked on tasks related to topics already covered in the associated lectures, while in the other learning environments tasks tended to be completed as new concepts and ideas were being introduced in class. In these latter environments we saw an uptick in not observable (UI-L), knowledge sharing (MTSU-L, UA-L), and knowledge construction (UA-L) dynamics. This suggests that the timing of a task during the learning process (i.e., during or after concepts are introduced) may affect the type of knowledge dynamics in which students engage.

Independent of the type of learning environment, the most common knowledge dynamic in the observed groups was knowledge sharing (see Table 3). In these types of episodes, the most common type of social processing was collaboration, but a majority of instances of knowledge sharing (>50%) involved a wide range of other forms of interaction (see Figure 2). Student engagement in knowledge construction was much less frequent but when it happened, it was associated with more instances of collaboration and fewer episodes of domination by a single student in a group.

Although each learning environment was designed to leverage collaborative activity in some capacity, collaboration was observed in fewer than half of the analyzed episodes for all sites. There was, however, wide variation from one learning environment to the other. While the distribution of observed types of social processing was quite similar between the two discussion environments (SBU-D, UI-D), the observed interactions were quite distinct in the three lecture

sites, going from primarily non-interactive and individualistic at UI-L to primarily collaborative and tutoring at UA-L. Social processing at MTSU-L was quite diverse, with no dominant form of interaction in the observed groups (see Table 4).

The amount of student participation in group work was also similar between the discussion sites, with full participation in over two thirds of all episodes in both cases, and more diverse among lecture sites (see Table 5). The different factors already described and discussed likely contributed to the observed minimal, partial, and full participation observed at UI-L, MTSU-L, and UA-L, respectively.

Limitations

The generalization of our findings is limited by the small number of student groups observed at each of the participating sites. Although observed student groups worked on a large number of different in-class tasks, all of the activities correspond to the first semester of general chemistry and thus may not be representative of all types of tasks students encounter in a chemistry curriculum. Our results are based on the analysis of students' explicit actions and conversations during group work. These are only proxies for their actual level of intellectual engagement with the in-class activities. Additionally, the presence of video cameras and recorders may have affected students' behaviors and expressed thoughts. Observations were carried out in classes taught by different instructors with varied approaches in the design, implementation, and monitoring of student work. Some of these approaches were not necessarily representative of best practices as described in the collaborative learning research literature. The relationship between the variables analyzed in this study can be expected to depend on the quality of a teacher's planning and instruction.

Observed differences in tasks characteristics and student engagement across sites are likely due to a complex combination of factors including institution type, student population, instructional style, and curricular materials to name a few. We focused on differences in curricular factors that we characterized as part of the study but acknowledge that other differences in the institutions (described in the research settings) may also play a role.

Implications

Our results indicate that in-class tasks in those learning environments that followed a traditional general chemistry curriculum tended to target the comprehension and retrieval levels, while the fraction of group activities at the analysis level was significantly larger in the site using an alternative curriculum purposefully designed to foster students' conceptual understanding (Talanquer and Pollard, 2010). A larger number of episodes of collaborative group work and knowledge construction were observed in this latter site. These results suggest that chemistry instructors using conventional curricula should carefully analyze the level of cognitive processing that their different in-class tasks demand, and work to diversify these activities to include higher level tasks if their goal is to engage all students more actively in knowledge construction.

Our study also elicited the effect that other factors of learning space, task design, and implementation, besides the curriculum, may have on student engagement in group work. The layout of the

classroom, for example, seemed to have a major effect on differences observed between the class taught in a traditional auditorium-style lecture hall and the class taught in a learning space designed to foster student collaboration. This result points to the importance of creating physical conditions that support student active engagement. Our findings also suggest that the timing of a task during a learning sequence could affect how students approach it. Tasks completed post-lecture in discussion/recitation settings more often resulted in knowledge application dynamics than tasks used during lecture as new concepts were explored or introduced. This result needs to be further explored to better understand whether this effect is actually caused by task timing rather than other factors such as students' differential perception and behavior in lecture versus discussion sessions.

Despite the association between higher-level analysis tasks and more knowledge construction interactions in some sites, all types of activities across all observed learning environments were dominated by a knowledge sharing dynamic. Knowledge construction is often cited as an instructional goal for active learning settings in general chemistry (Chi and Wylie, 2014; Lombardi *et al.*, 2021), but we saw relatively little of it across our sites. Although more frequent engagement in knowledge construction may be desirable, this result invites us to investigate what balance between knowledge sharing and knowledge construction dynamics might be best to foster meaningful learning during in-class tasks, and what ratio might be actually feasible to accomplish in large-enrollment courses with many diverse students. Given that we did not observe students working on knowledge utilization tasks (the highest level in Marzano's taxonomy) in any of the sites, it would also be important to explore how working on these types of tasks affects different aspects of student engagement. However, knowledge utilization tasks are often about planning and completing activities; this is often difficult to accomplish in a single class period. Knowledge utilization tasks require significantly more time and attention for successful implementation. Overall, our results speak to the importance of designing tasks that lead students to meaningfully engage with each other's ideas and of supporting them in these processes (Chi, 2009).

Although collaborative activity was the most common type of social processing in the observed groups, it amounted to less than half of the total number of episodes with wide variation across sites. This result suggests that, independently of the curriculum and nature of in-class tasks, instructors should carefully evaluate and reflect on their planning and implementation of group work. While group work is a common instructional strategy, it is often implemented with little structure or feedback and students may not be aware of effective ways to interact. The literature on collaborative learning provides insights on best practices to foster active, equitable, and productive interactions among group members that were not always implemented in the different classes observed in our investigation (Eberlein *et al.*, 2008). As part of our overall project, we plan to investigate how the implementation of these best practices with consistency and fidelity across different sites affects student engagement as characterized in this study.

Appendix A

The coding system used in our study is described and exemplified in the following tables:

- A.1: Tasks characteristics
- A.2: Knowledge dynamics
- A.3: Social processing

Table A.1 Coding for tasks characteristics

Level	Definition	Example
Retrieval	Involves the simple recognition, recall, or execution of knowledge, including rote calculations. Tasks of this level ask a learner to reiterate or identify information in almost the exact way it was introduced.	How is atomic radius defined?
Comprehension	Involves the integration and symbolic representation of knowledge, generally with a focus on key features and organization of information.	Draw the Lewis structures of O ₃ and O ₂ .
Analysis	Involves examining knowledge in detail and generating new conclusions.	Consider substances made up of the following atoms and molecules: He, CH ₄ , Ne, C ₂ H ₆ . Arrange the substances in order of increasing boiling point and clearly justify your rankings.
Knowledge Utilization	Requires that students apply or use knowledge in specific situations and almost always includes a component of justification. These tasks will include decision making between two or more alternatives, problem solving that includes accomplishing goals for which an obstacle exists, experimenting, or investigating.	Based upon everything you have learned; do you think that solar geoengineering should be an option for combating climate change? Justify your answer.

Table A.2 Coding for knowledge dynamics ("S#" indicates student number # in a group; "I" indicates instructor); "Q" indicates the question posed).

Category	Definition	Example
Knowledge Sharing	The focus of the group interactions is based on sharing information to answer the question without questioning of the utterances presented.	<p>Q: 300.00 grams of ethanol at 10.0°C is heated with 14,640 J of energy. What is the final temperature of the ethanol?</p> <p>S1: We're gonna be looking for delta T.</p> <p>S2: So why don't we just divided by fourteen thousand six hundred forty by the 300 grams of ethanol?</p> <p>S1: Delta T will be equal to 20. So if we go from 10, the heat should be equal to 30 degrees Celsius.</p> <p>S2: So just be 30 degree Celsius?</p> <p>S1: Right, so...</p> <p>S2: Cuz if you're looking for delta T, you're gonna have 300...</p> <p>S1: Times 2.44, right?</p> <p>S2: Yeah.</p> <p>S1: Times delta T is going to be equal to 14.640</p> <p>S2: And when you work that out, you get 30 or 20 degree Celsius...</p> <p>S1: you get 20.</p> <p>S2: So you add that to and you get 30 degree Celsius.</p> <p>S1: That's right.</p>
Knowledge Application	The focus of the group interactions is based on applying a formula/method/concept and relating that to a clear understanding of how it relates to the explanation or process of solving the problem.	<p>Q: How can you use the periodic table to help you determine the most stable oxidation state?</p> <p>S2: Let's see. How can we use periodic table to help you determine the most stable oxidation state.</p> <p>S2: Umm. How can...[Mumbles bits of the activity prompt]...state. We're on?</p> <p>S1: We're on C right now.</p> <p>S2: C?</p> <p>S1: Yeah.</p> <p>S2: The most stable oxidation state. Um, elements. I don't know how to say this... How can you help the periodic table...um... elements are trying to lose..or actually, elements are trying to..</p> <p>S3: It goes back to the orbitals being filled.</p> <p>S2: Yeah, they're trying to get back to a core electron. They want to be like in a stable state. So just be like, the atoms want to be in a stable state and so they're gaining or losing electrons.</p> <p>S3: Even if you had orbitals filled with one arrow, that would still be its stable state, right?</p> <p>S2: Um...If you had one arrow, you'd have like 1s¹. What do you mean by like one arrow?</p> <p>S2: Okay.</p> <p>S3: Because like, let's say you have the 3p orbital and like all the orbitals are filled with one arrow.</p> <p>S3: Like let's say all of them.</p> <p>S2: Okay. So all of them. So you've got...um.. you've got 3p³, P. What, would that be five? So you'd have...1,2...1,2...1,2,3,5 which would be iron. And you'd want to... I don't know, it's a transition metal. I think you'd want to ah,you're not stable there. So you're going to want to have to oxidize or reduce. I'm not sure exactly. I'm just as confused as you are pretty much. I swear they try to make this confusing during lectures so nobody understands it.</p> <p>S3: [Inaudible]</p> <p>S2: Yeah.</p> <p>S1: Okay, I think I got everything.</p>
Knowledge Construction	The focus of group interactions is based on sharing information and building upon the ideas of others by questioning or critiquing the ideas presented.	<p>Q: when dry ice (solid CO₂) sublimates, what forces are overcome? (select all that apply)</p> <p>A. covalent bonds between carbon atoms and oxygen atoms</p> <p>B. hydrogen bonding between carbon dioxide molecules</p> <p>C. dipole-dipole forces between carbon dioxide molecules</p> <p>D. dispersion forces between carbon dioxide molecules</p> <p>S2: Dip dip</p> <p>S1: Yeah</p> <p>S2: [Inaudible] degrees Celsius</p> <p>S1: Yeah and then it's gotta be at its gotta be at that specific pressure</p> <p>S2: 611.2 Pascals of mercury</p> <p>S2: [Inaudible] mmHg</p> <p>S1: 4.58 mmHg</p> <p>I: So, we want to know when dry ice or solid carbon dioxide sublimates, what kind of forces are basically broken?</p> <p>S1: Dipole dipole right?</p> <p>S4: Is it dipole dipole? There is only a change of like .3, does that count?</p> <p>S3: Its dipole dipole between the molecules not between the</p> <p>S1: Yeah</p> <p>S3: So like [Inaudible]</p> <p>S1: Yeah it doesn't matter if the bonds are polar it just matters if the molecules are polar</p> <p>S1: Dispersion forces are between nonpolar</p> <p>S4: Dispersion forces between [Inaudible]</p> <p>S1: Yeah, but also but yeah so just</p> <p>S2: So those are also carbon</p> <p>S1: Oh yeah</p> <p>S2: No, wait,</p> <p>S1: I don't think it's breaking the bond</p> <p>S2: No, it's not</p> <p>S3: No, it's not</p> <p>S1: Fight me</p> <p>S2: Cause the molecule is pol</p>
Not Observable	No knowledge dynamic is seen due to a lack of student interaction.	

Table A.3 Coding for social processing ("S#" indicates student number # in a group; "Q" indicates the question posed).

Category	Definition	Example
Collaborative	Students are co-constructing ideas and generating products together.	Q: <i>How do the kinetic and potential energy of electrons change when atoms bond?</i> S1: "Kinetic goes down." S2: "Potential goes up." S1: "And attractive go up." S3: "Wait, wait, wait. Kinetic goes down and potential goes?" S2: "Potential attractive goes up." S3: "It goes up?" S2: "Yea it goes up and potential repulsive goes down." S3: "And then all potential goes down." S1: "next, shouldn't it go down if they are staying in the same place and everything else is moving?" S3: "It would have to go down because they are further away." S4: "I have the potential energy of the electrons not delocalizing decreasing." S1: "I think attractive is increasing because they are further apart." S4: "Then I think the answer is option four."
Tutoring	One or more students ask questions that another student responds to by guiding the students through the problem asking for the tutees ideas or just by explaining their reasoning without asking input from the student who asked the question.	Q: <i>Consider the bonds H-H and C-C: Which chemical bond is longer, and which is stronger? Why?</i> S1: "I am just so lost right now." S2: "Okay, so the first thing would be... you don't need to know the whole electron and proton thing right now. When they are far enough apart, they attract towards each other then they hit this perfect balance point between repulsive and attraction. Then there is this part over here which is super repulsive but generally when you get a bond, like in water they stay in this perfect range between attraction and repulsion and that's what the graph shows. If the get really close they shot apart from each other but if they get to that valley that's were they have the lowest potential energy and bonded." S2: "So by longer do is it talking about distance?" S1: "Yea, so the C-C bond is going to be longer than the H-H bond. Because they have a stronger pull, they also have a stronger repulsion so it will be longer." S2: "So is C-C also stronger?" S1: "I think so, because they will have stronger forces, both repulsive and attractive."
Individualistic	Students are working independently and are not having conversation about the question products.	Q: <i>If two particles each have 8 protons and 8 electrons but, one is larger, which is more polarizable?</i> S1: "I put A." S2: "Okay."
Domination	One individual construct the response on their own without significant input from others.	Q: <i>What will be observed when these substances are mixed: $Cl_2(aq)$ and $Na_2CO_3(aq)$?</i> S1: "Well, one of the things is going to be copper carbonate and the other is going to be sodium chloride. So you are going to see a change in the color composition as the things move around. One of them would participate out, I think the copper carbonate." S2: "Okay." S3: "alright."
Confusion	Students are too confused to really generate the expected product or make confident progress for a question.	Q: <i>Can you infer how many bonds C-H bonds are present in the molecule based on the IR spectrum?</i> S1: "I don't understand" S2: "I have no idea how to even begin." S3: "Should we just guess?" S4: "I'm putting C then."
Non-Interactive	Students are not having any conversation, but there is no proof of individualistic work.	
Multiple Social Processing	Students engage in more than one social processing.	

Appendix B

Tables B.1 through B.4 present the results of the post hoc analysis of residuals performed across different dimensions in our study.

Table B.1. Residuals: Task level and learning environment (Table 2 data)

Task Level	SBU-D	UI-D	UI-L	MTSU-L	UA-L
Retrieval	-1.89	-2.88	2.69	0.647	-3.22
Comprehension	1.68	-0.452	-1.48	2.51	-2.19
Analysis	-1.09	-1.70	0.092	-4.53	6.18

Table B.2. Residuals: Knowledge dynamics and learning environment (Table 3 data)

Knowledge Dynamic	SBU-D	UI-D	UI-L	MTSU-L	UA-L
Not Observable	-1.39	-1.28	12.6	2.84	-6.10
Knowledge Sharing	-1.72	-0.359	-1.41	-1.23	2.36
Knowledge Application	9.61	2.52	-3.19	1.60	-5.31
Knowledge Construction	-4.00	-0.592	-3.03	-0.791	4.06

Table B.3 Residuals: Social processing and learning environments (Table 4 data)

Social Processing	SBU-D	UI-D	UI-L	MTSU-L	UA-L
Non-interactive	-2.51	-2.06	19.8	-2.96	-5.05
Individualistic	-3.55	-2.05	3.74	3.37	-1.05
Confusion	-0.938	-1.37	-2.48	-3.25	4.04
Domination	1.81	1.25	-1.11	1.98	-2.08
Tutoring	1.36	-2.17	-2.84	-0.737	1.92
Collaboration	-1.67	-0.720	-4.26	-4.14	5.46
Multiple Social Processing	6.01	6.41	-3.54	6.60	-7.89

Table B.4. Residuals: Amount of participation and learning environment (Table 5 data)

Amount of Participation	SBU-D	UI-D	UI-L	MTSU-L	UA-L
Minimal Participation	-3.16	-1.82	9.79	1.37	-2.88
Partial Participation	0.027	-3.70	0.140	-4.94	-1.47
Full Participation	1.20	3.98	-3.93	-4.91	2.42

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

We would like to thank all students and instructors who allowed us to observe their work in classrooms. We are also grateful for the support from the National Science Foundation through the collaborative projects DUE-IUSE 1914510, 1914813, and 1915047.

Author Contributions

RC, GTR, LS, and VT contributed to funding acquisition, study conceptualization, project administration and supervision, and writing (review & editing). SF, AF, MMS, HTN, FQ, NES, and AS contributed to the investigation through data collection, curation, and formal analysis, as well as writing (review & editing). JR, ZDKG, and VT contributed to formal overall analysis of collected data and writing of the original draft, including visualization of main findings.

References

- Anderson, L.W. (ed.), Krathwohl, D.R. (ed.), Airasian, P.W., Cruikshank, K.A., Mayer, R.E., Pintrich, P.R., Rath, J. and Wittrock, M.C., (2001), *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of educational objectives* (complete edition), New York: Longman.
- Arthurs, L. A. and Kreager, B. Z., (2017), An integrative review of in-class activities that enable active learning in college science classroom settings, *Int. J. Sci. Educ.*, **39**(15), 2073-2091.
- Biggs, J., (1987), *Student approaches to learning and studying*. Hawthorne, Victoria: Australian Council for Educational Research.
- Biswas, G., Leelawong, K., Schwartz, D., Vye, N. and Teachable Agents Group at Vanderbilt, (2005), Learning by teaching: A new agent paradigm for educational software, *Appl. Artif. Intell.*, **19**, 363-392.
- Chi, M. T. H., (2009), Active-Constructive-Interactive: A conceptual framework for differentiating learning activities, *Top. Cogn. Sci.*, **1**, 73-105.
- Chi, M. T. H., de Leeuw, N., Chiu, M. H. and Lavancher, C., (1994), Eliciting self-explanations improves understanding, *Cogn. Sci.*, **18**, 439-477.
- Chi, M. T. H. and Wylie, R., (2014), The ICAP framework: Linking cognitive engagement to active learning outcomes, *Educ. Psychol.*, **49**, 219-243.
- Chinn, C. A. and Malhorta, B. A., (2002), Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks, *Sci. Educ.*, **86**, 175-218.
- Cotner, S., Loper, J., Walker, J. D. and Christopher-Brooks, D. (2013), "It's not you, it's the room" —Are the high-tech, active learning classrooms worth it? *J. Coll. Sci. Teach.* **42**, 82-88.
- Crowe, A., Dirks, C., & Wenderoth, M. P. (2008), Biology in bloom: implementing Bloom's taxonomy to enhance student learning in biology, *CBE Life Sci. Educ.*, **7**, 368-381.
- Dávila, K. and Talanquer, V. (2010), Classifying end-of-chapter questions and problems for selected general chemistry textbooks used in the United States, *J. Chem. Educ.* **87**, 97-101.
- Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K. and Kestin, G., (2019), Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom, *PNAS, USA*, **116**(39), 19251-19257.
- Doyle, W. and Carter, K., (1984), Academic tasks in classrooms, *Curric. Inq.*, **14**, 129-149.
- Eberlein, T., Kampmeier, J., Minderhout, V., Moog, R. S., Platt, T., Varma-Nelson, P. and White, H. B., (2008), Pedagogies of engagement, *Biochem. Mol. Biol. Educ.*, **36**, 262-273.
- Ford, M. J., (2008), "Grasp of practice" as a reasoning resource for inquiry and nature of science understanding, *Sci. & Educ.*, **17**(2&3), 147-177.
- Forman, E., (1989), The role of peer interaction in the social construction of mathematical knowledge, *Int. J. Educ. Res.*, **13**(1), 55-70.
- Fredricks, J. A., Blumenfeld, P. C. and Paris, A. H., (2004), School engagement: Potential of the concept, state of the evidence, *Rev. Educ. Res.*, **74**(1), 59-109.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H. and Wenderoth, M. P., (2014), Active learning increases student performance in science, engineering, and mathematics, *PNAS, USA*, **111**(23), 8410-8415.
- Groves, M., Sellars, C., Smith, J. and Barber, A., (2015), Factors affecting student engagement: A case study examining two cohorts of students attending a post-1992 university in the United Kingdom, *Int. J. High. Educ.*, **4**(2), 27-37.
- Hancock, T. S., Friedrichsen, P. J., Kinslow, A. T. and Sadler, T. D., (2019), Selecting socio-scientific issues for teaching, *Sci. & Educ.*, **28**(6-7), 639-667.
- Inkelas, K. K., and Weisman, J. L., (2003), Different by design: An examination of student outcomes among participants in three types of living-learning programs, *J. Coll. Stud. Dev.*, **44**, 335-368.
- Järvelä, S. and Renninger, K., (2014), Designing for learning: Interest, motivation, and engagement, in R. Sawyer (ed.), *The Cambridge Handbook of the Learning Sciences* (Cambridge Handbooks in Psychology), Cambridge: Cambridge University Press, pp. 668-685.
- Kahu, E. R., (2013), Framing student engagement in higher education, *Stud. High. Educ.*, **38**(5), 758-773.
- Klahr, D. and Nigam, M., (2004), The equivalence of learning paths in early science instruction, *Psychol. Sci.*, **15**, 661-667.

- Kuh, G.D., (2009), What student affairs professionals need to know about student engagement, *J. Coll. Stud. Dev.*, **50**, 683-706.
- Kumpulainen, K. and Kaartinen, S., (2003), The interpersonal dynamics of collaborative reasoning in peer interactive dyads, *Int. J. Exp. Educ.*, **71**(4), 333-370.
- Lampert, M., and Graziani, F., (2009). Instructional activities as a tool for teachers' and teacher educators' learning, *Elem. Sch. J.*, **109**, 491-509.
- Laverty, J. T., Underwood, S. M., Matz, R. L., Posey, L. A., Carmel, J. H., Caballero, M. D., Fata-Hartley, C. L., Ebert-May, D., Jardeleza, S. E. and Cooper, M. M. (2016), Characterizing college science assessments: The three-dimensional learning assessment protocol. *PLoS One*, **11** (9), No. e0162333.
- Liyanage, D., Lo, S. M. and Hunnicutt, S. S., (2021), Student discourse networks and instructor facilitation in process oriented guided inquiry physical chemistry classes, *Chem. Educ. Res. Pract.*, **22**(1), 93-104.
- Lombardi, D., Shipley, T. F. and Astronomy Team, Biology Team, Chemistry Team, Engineering Team, Geography Team, Geoscience Team, and Physics Team, (2021), The curious construct of active learning, *Psychol. Sci. Public Interest*, **22**(1), 8-43.
- Marzano, R. J. and Kendall, J. S., (2007), *The new taxonomy of educational objectives*, Thousand Oaks, CA: Corwin Press.
- National Research Council, (2012), *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*, Washington, DC: The National Academies Press.
- Nonaka, I., von Krogh, G. and Voelpel, S.C. (2006), Organizational knowledge creation theory: evolutionary paths and future advances, *Organ. Stud.*, **27** (8), 1179-208.
- Pike, G. R., (1999), The effects of residential learning communities and traditional residential living arrangements on educational gains during the first year of college, *J. Coll. Stud. Dev.*, **40**, 269-284.
- Pike, G.R., Kuh, G.D. and McCormick, A.C., (2011), An investigation of the contingent relationships between learning community participation and student engagement, *Res. High. Educ.*, **52**, 300-322.
- Prather, E. E., Rudolph, A. L., Brissenden, G. and Schlingman, W. M., (2009), A national study assessing the teaching and learning of astronomy. Part I. The effect of interactive instruction, *Am. J. Phys.*, **77**(4), 320-330.
- R Core Team, (2021), R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Reeve, J. and Tseng, C. M., (2011), Agency as a fourth aspect of students' engagement during learning activities, *Contemp. Educ. Psychol.*, **36**(4), 257-267.
- Roberson, B. and Franchini, B., (2014), Effective task design for the TBL classroom, *J. Excell. Coll. Teach.*, **25**(3-4), 275-302.
- Schwartz, D. L. and Bransford, J. D., (1998), A time for telling, *Cogn. Instr.*, **16**, 475-522.
- Sinatra, G. M., Heddy, B. C. and Lombardi, D., (2015), The challenges of defining and measuring student engagement in science, *Educ. Psychol.*, **50**(1), 1-13.
- Stains, M. and Vickrey, T., (2017), Fidelity of implementation: An overlooked yet critical construct to establish effectiveness of evidence-based instructional practices, *CBE Life Sci. Educ.*, **16**(1), rm1.
- Talanquer V. and Pollard J., (2010), Let's teach how we think instead of what we know, *Chem. Educ. Res. Pract.*, **11**(2), 74-83.
- Talbert, R. and Mor-Avi, A. (2019), A space for learning: An analysis of research on active learning spaces, *Heliyon*, **5**, e029672.
- Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., ... and Freeman, S., (2020), Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math, *PNAS, USA*, **117**(12), 6476-6483.
- van Last, J., (2009), Distinguishing knowledge-sharing, knowledge construction, and knowledge-creation discourses, *Comput.-Support. Collab. Learn.*, **4**, 259-287.
- Wang, Y., Chen, A., Schweighardt, R., Zhang, T., Wells, S. and Ennis, C., (2019), The nature of learning tasks and knowledge acquisition: The role of cognitive engagement in physical education, *Eur. Phys. Educ. Rev.*, **25**(2), 293-310.
- Zepke, N. and Leach, L., (2010), Improving student engagement: Ten proposals for action, *Act. Learn. High. Educ.*, **11**(3), 167-177.
- Zhao, C. M. and Kuh, G. D., (2004), Adding value: Learning communities and student engagement, *Res. High. Educ.*, **45**(2), 115-138.
- Zoller, U., (1993), Are lecturing and learning compatible? Maybe for LOCS: Unlikely for HOCS, *J. Chem. Educ.*, **70**, 195-197.