Chemistry Education Research and Practice

Factors Affecting Individuals' Cognitive Engagement during Group Work in General Chemistry: Timing, Group Size, and Question Type

Journal:	*Chemistry Education Research and Practice*
Manuscript ID	RP-ART-10-2023-000279.R2
Article Type:	Paper
Date Submitted by the Author:	25-Mar-2024
Complete List of Authors:	El-Mansy, Safaa; Reed College, Chemistry Stephens, Alexandra; Portland State University, Chemistry Mortensen, Abigale; Portland State University, Applied Linguistics Francis, Joan; Portland State University, Applied Linguistics Feldman, Shayna; Portland State University, Applied Linguistics Sahnow, Cecilia; Portland State University, Department of Biology Barbera, Jack; Portland State University, Department of Chemistry Hartig, Alissa; Portland State University, Applied Linguistics

Factors Affecting Individuals' Cognitive Engagement during Group Work in General Chemistry: Timing, Group Size, and Question Type

Safaa Y. El-Mansy, Alexandra Stephens, Abigale Mortensen, Joan M. Francis, Shayna Feldman, Cecilia A. Sahnow, Jack Barbera, and Alissa J. Hartig

Abstract

Understanding how individual students cognitively engage while participating in small group activities in a General Chemistry class can provide insight into what factors may be influencing their level of engagement. The Interactive-Constructive-Active-Passive (ICAP) framework was used to identify individual students' level of engagement on items in multiple activities during a General Chemistry course. The effects of timing, group size, and question type on engagement were investigated. Results indicate students' engagement varied more in the first half of the term, and students demonstrated higher levels of engagement when working in smaller groups or subsets of larger groups when these groups contained students with similar levels of knowledge. Finally, the relation between question type (algorithmic versus explanation) and engagement depended on the activity topic. In an activity on Solutions and Dilutions, there was a significant relation where algorithmic items had higher occurrences of Interactive engagement. The implications of this work regarding teaching and research are discussed.

Introduction

Active learning (AL) has become an increasingly prevalent teaching pedagogy due to the positive effect on achievement outcomes, particularly for marginalized student populations (Freeman et al., 2014; Haak et al., 2011; Harris et al., 2020). Furthermore, a meta-analysis which looked at multiple research studies that used a variety of different AL techniques in chemistry classrooms showed the effect of AL on achievement outcomes can vary greatly based on the AL technique being implemented (Rahman & Lewis, 2020). For example, one result of this analysis showed that across multiple studies which used Process-Oriented Guided Inquiry Learning (POGIL), the outcomes of POGIL implementation on academic performance ranged from no effect to a medium effect size. One factor that may contribute to this result is cognitive engagement, which has been defined as the effort students put forth towards learning and mastering new material (Fredricks et al., 2004). Therefore, understanding how students cognitively engage while participating in small group activities, and more specifically, identifying what factors may be influencing engagement, could be important in optimizing the positive effect of group work on student performance.

One factor that has been shown to affect student engagement is how long students have been in school. For example, longitudinal studies have investigated student engagement from one to three years and have found fluctuations in the level of engagement students exhibit (Bruce et al., 2010; Kahu et al., 2020). A study among Chinese university students measured student engagement using surveys over a two-and-a-half-year period and found an increase in engagement across this time (Guo et al., 2023). While most longitudinal studies have investigated engagement over multiple years, Kahu et al. analyzed engagement over a single year through narratives provided by interviews and found engagement for first year university students fluctuated throughout the year due to factors such as self-efficacy and sense of belonging (Kahu et al., 2020). Therefore, it can be expected that student's engagement may not even be consistent across a single term as students adapt to their course schedule and become settled into a routine. Additionally, most of these studies have investigated engagement at an institutional level rather than the course level. Students' engagement within a single course may also show variation due to factors related to the specific course content or environment. Therefore, looking at individual students' cognitive engagement across the course of a single term in chemistry classes may offer insights that could provide instructors with actions they could use to improve student engagement.

A second factor that may influence student engagement is group size. Research has shown that group size can have an effect on an individual's learning outcomes, team performance, and learning satisfaction. A review of the effect of group size for elementary, secondary, and post-secondary students showed a negative relation between the number of students in a group and learning outcomes (Wilkinson & Fung, 2002), and that the optimal group size for learning is three to four students (Lou et al., 1996). Work done among secondary school physics students showed that students progressed further in their reasoning when working in groups of four versus pairs (Alexopoulou & Driver, 1996). A study conducted in an undergraduate marketing class demonstrated that group performance increased with number of students in a group up to five students and then decreased (Treen et al., 2016), while a second study investigated the differences between two, three and four person teams on team performance and found that four person teams showed higher performance than two or three person teams (Cossé et al., 1999). Research has also shown that college engineering students who worked in groups of two to four students showed stronger learning satisfaction than those who worked in groups of five to seven students (Chou & Chang, 2018). In summary, previous research indicates that optimal group size may range from three to five students based on its effect on learning outcomes, performance and satisfaction. It may also be dependent on education level (e.g., secondary versus higher education) and subject matter. Although the impact of group size on academic performance and outcomes has been investigated in the literature, we were unable to find similar research on the relation between group size and engagement. However, since research has shown that both engagement and group size can influence learning outcomes, it is possible that group size may also affect student engagement. Therefore, investigating the effect of group size on individuals' cognitive engagement in General Chemistry may provide valuable insight that can be used to optimize student engagement.

The type of question asked in an activity could also contribute to the mode at which students engage. Previous research has shown that achievement outcomes vary depending on whether students were asked to perform a calculation or use a pre-determined set of procedures versus if they were asked questions that were more conceptual in nature (Cracolice et al., 2008; Surif et al., 2014; Zoller et al., 2002). Additional work indicated that questions more focused on calculations promoted lower-order thinking skills whereas questions focused more on concepts and explanations promoted higher-order thinking (Zoller et al., 2002). While question type seems to have an effect on both achievement outcomes and the level of thinking skills students exhibit, it may also be related to the degree to which students engage with the question; therefore, the relation between question type and engagement should be further investigated.

To investigate the effect of the previously mentioned factors on engagement, a way to measure individual students' cognitive engagement is needed. The Interactive-Constructive-Active-Passive (ICAP) framework provides a model which can be used to measure the mode at which students cognitively engage by looking at overt behaviors that students display (Figure 1) (Chi et al., 2018). This framework provides an ideal tool to measure cognitive engagement during group work by examining the content of the group conversation as well as non-verbal behaviors (El-Mansy et al., 2022). During group work, in the lowest mode, Passive engagement, students display behaviors which demonstrate that they receive information but do not physically manipulate the information in any way, e.g., nodding in agreement with statements made by members of the group but not writing anything down. In the Active mode, students physically manipulate information but do not generate any new information. For example, students may nod in agreement but also write their answer on their worksheet. For the Constructive mode, students generate new information beyond that which is presented to them. During group work, this may include making statements that demonstrate independent generation of information. At the highest mode, Interactive, students co-generate information through dialogue between students or between students and instructors. This may be posing a question which results in generation of information by another student or answering a question posed by a student.

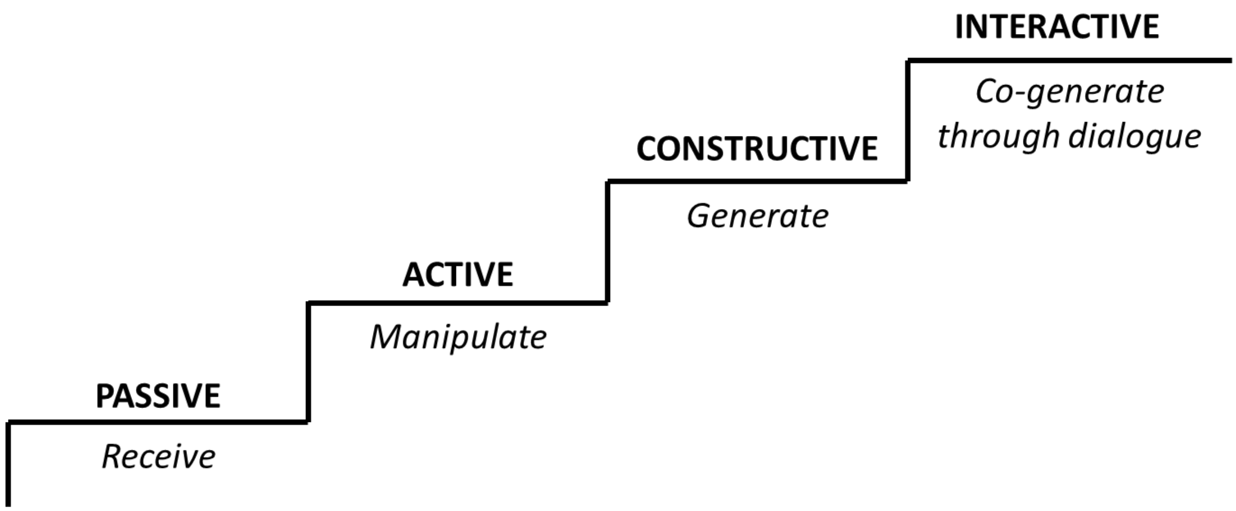


Figure 1: ICAP cognitive engagement modes (bold) and characteristic behaviors (italics) (based on Chi et al., 2018)

Research Questions

The purpose of this study is to investigate how individual students within a first term General Chemistry course cognitively engage when working in groups on activity worksheets and to identify factors which may be influencing the level at which they engage. To do so, we used the ICAP framework and deductive coding to answer the following research questions:

1. How does individual students' cognitive engagement vary across activities?
2. What is the effect of group size on individual students' cognitive engagement?
3. What relation is observed between the type of question asked in the activities and students' level of cognitive engagement?

Methods

Context of Study

Students from two sections of the first term of a General Chemistry course at Portland State University (PSU) in the Pacific Northwest of the United States participated in this study. The course was conducted during the Fall 2022 term (a 10-week quarter), and each section was taught by a different instructor and contained approximately 200 students. The course was taught twice a week for 110 minutes. One day of the week was a "lecture" day where the instructor presented the material and engaged students through the use of clicker questions, and the other day was an "activity" day where students generally worked in groups of 3-5 students to complete an activity worksheet during class time. The purpose of these worksheets was to introduce students to new concepts in an active learning format. The activity worksheets were developed in-house at PSU, and both instructors were involved in the planning and development of the activities and worked together to build their other course materials and assessments around the activities, ensuring that each section provided a similar learning environment to students.

Each activity worksheet consisted of models which presented the topic material and/or data. The worksheets consisted of Key Questions (KQ), Exercises (E), and Problems (P). Key Questions generally asked students to identify information explicitly present in the model. In Exercises, students applied the information from the model to perform a calculation or determine an answer to a conceptual question, and Problems asked students to perform a multi-step calculation or apply the information from the models in a novel way. All students participated in the activity worksheet, and most students worked in groups of three to five during an activity day. Groups were facilitated by the instructor and learning assistants who moved between groups, answering questions and monitoring the progress of the groups. The activity worksheets were not turned in or graded and an activity key was posted to the course learning management system within 24 hours of an activity day. While the worksheets themselves were not graded, quiz and exam question content were directly tied to the material from each preceding activity.

Data Collection

The data collected for this analysis was part of a larger research study which was approved by PSU's Institutional Review Board (HRRP# 217370-18). Students were recruited by author S.Y.E. approximately one week prior to each activity being observed. Consenting students filled out a Qualtrics survey where they provided demographic information. From this initial pool, students were selected to maximize racial and gender diversity. These students were then randomly divided into one of the groups for observation. Due to equipment constraints, a maximum of two groups per activity per section were recorded. Groups were capped at five students with the goal of fostering conversation among all group members. Group sizes varied across the activities and sections based on the number of consenting students who showed up for class on data collection days. Twenty-three students participated across both sections of the course, and six of the twenty-three participated in more than one observed activity. All student names used in this manuscript are pseudonyms.

Three activities were observed during the 10-week term: Solutions and Dilutions in week 3 contained 18 items, Periodic Trends in week 7 contained 37 items, and Molecular Polarity in week 10 contained 24 items. Each group was audio and video recorded for each activity using two video cameras placed on opposite sides of the group to maximize recording all members' behaviors (e.g., gaze direction and when students wrote their response). These groups were placed at the back of the lecture hall in opposite corners to minimize non-study group recordings. The recordings were transcribed verbatim using an automated transcription service. The transcripts were then reviewed and edited by author C.A.S., and pertinent physical actions such as nodding in agreement or pointing to a particular item on the worksheet were added. The completed activity worksheets were collected by author S.Y.E. at the end of the class period. The worksheets were scanned as an additional resource which could be used to aid in the identification of engagement modes. The scanned copies were returned to students the same day as the activity. Additionally, semi-structured interviews were conducted with consenting students approximately one week after the activity day. The purpose of these interviews was 1) to gain insight into the students' perception of the effectiveness of both the activity and the dynamics within their group and 2) as a second data source to triangulate the results obtained from analysis of the recorded observations.

Data Analysis

Individual Coding

Development of the codebook for individual students' cognitive engagement began with a codebook that had been previously developed using ICAP to identify the engagement mode of the group (El-Mansy et al., 2022). In that study, the highest observed engagement mode during group response to an item was identified as the engagement mode of the group for that item. That codebook was applied to group work that was conducted remotely over Zoom and focused primarily on participants' verbal contributions since most of their non-verbal behaviors were not visible in the recording and the overall level of engagement for the group as a whole could usually be determined based on their verbal contributions alone.

During the Winter 2022 term, author S.Y.E. began by applying this prior codebook to each individual student's statements within a group response to a specific item to determine their cognitive engagement. The codebook was first applied to data collected from one group participating in the Molecular Polarity activity during the Fall 2021 term. As coding progressed, S.Y.E. found statements or behaviors that did not align with the code descriptions given in the codebook. To develop the codebook to be more focused on individual students' engagement, during the Spring 2022 term, S.Y.E. met weekly with author A.J.H., an applied linguist, and two undergraduate applied linguistics students, authors A.M. and J.M.F., to analyze conversation excerpts that were difficult to code. The applied linguistics students focused primarily on features relevant to multimodal conversation analysis (Mondada, 2019; Sacks et al., 1974), such as response timing, gaze direction, and intonation (e.g., whether students' tone indicated a question or statement). Through these meetings, the definitions of each engagement code were expanded and other sources of evidence were considered, including where students were looking, when students wrote their answers relative to when the group conversation occurred, and their written response on their worksheet. These meetings continued throughout the Spring 2022 term until all ambiguous excerpts were coded to consensus. Video and audio recordings were also collected during a Physical Chemistry class in Spring 2022. This data was used to further refine the codebook during weekly meetings held in the Fall 2022 term. This process involved discussing ambiguous excerpts with A.J.H. and an applied linguistics master's student, author S.F., and coding the excerpts to consensus. S.Y.E. then used the final codebook (Table 1) to code all remaining transcripts from the Fall 2022 General Chemistry classes.

To investigate the effect of when the activity occurred during the term, codes were assigned to each student for each item completed in each activity. For students that participated in more than one activity, the number of codes assigned to a student for each ICAP category were summed, and the distribution of these "summed" ICAP codes was graphed for each activity in which they participated. Trends in these distributions across multiple activities for a single student were then explored.

Table 1: Codebook for individual engagement

UNENGAGED (U)

• Student does not appear to be working on the activity.
  ○ Makes statements unrelated to activity.
  ○ Off-task use of their phone.
  ○ Other behaviors unrelated to working on the activity.

PASSIVE (P)
• Student's gaze is directed towards other group members or activity worksheet. • Student may nod in agreement with the conversation regarding the activity but does not orient to the content of the activity in any visible way (e.g., they do not write down any information, do not point to specific parts of the activity).
ACTIVE (A)
• Repeating content from the activity or repeating a statement made by another student in reference to a simple idea (e.g., identifying and writing down something specific from an equation or model). • Statements or gestures of agreement with an answer provided by another student (e.g., nodding while writing down an answer, giving a "thumbs up").

<table>
<tr><td>
• Other evidence of orienting to the content of the activity but without generating new information (e.g., asking what number they are on).
• Students listen to statements/conversation occurring among other students (i.e., orienting to the question) and then write their answer on their worksheet, and/or their written response on the worksheet reflects the content of the conversation.
</td></tr>
<tr><td align="center">CONSTRUCTIVE (C)</td></tr>
<tr><td>
• Student provides an answer to question, generating new information independent of other members in the group.
  ○ This may take the form of a statement or a request for confirmation. Other students may agree with the original student's statement and provide evidence of their own generation of information which does not add any new information to the group.
  ○ If a single student's knowledge increases through dialogue, but the group's does not, this would be CONSTRUCTIVE for that student. (If one student has the answer and "teaches" and others are learning or "catching up", the original student is CONSTRUCTIVE).
  ○ If the content of the conversation is about the process for solving the problem but does not address the actual answer to the problem, and the students write after the discussion, this implies that each student is independently generating information to answer the item because the answer itself has not been co-generated through group conversation.
  ○ Student may be writing their answer to a specific item prior to conversation about that item among group members and does not alter their written response after the discussion.
  ○ Student's written response on their worksheet may show information not present in the group conversation, suggesting independent generation of knowledge.
</td></tr>
<tr><td align="center">INTERACTIVE (I)</td></tr>
<tr><td>
• Student's contribution co-generates new information with other students or instructor/learning assistants to answer the item. Neither party shows evidence of generating the answer independently.
  ○ If the conversation occurs between students, knowledge of more than one group member increases through this conversation.
  ○ If the conversation occurs between a student and an instructor, the conversation results in co-generation of information for the students engaged in the conversation or for the entire group.
• Student's statements add to information that has been previously contributed by another student.
  ○ Student provides new information to answer a question posed by another student.
  ○ Student poses a question that leads to generation of new information later in the conversation about the question.
  ○ The generation of new information by the first student prompts a second student to provide additional information.
</td></tr>
</table>

To investigate the effect of group size, the number of codes in each ICAP category for *each item* in an activity for a specific group was determined. For example, for a single item answered by a four-person group, one student was Active, one was Constructive, and two were Interactive. This distribution was determined for every group for every item in every activity. The number of codes for each ICAP category for each group were then summed together. The distribution of these "summed" ICAP codes was plotted across different group sizes, and trends were observed and analyzed.

Question Type Coding

Items in the activities were defined as either “algorithmic” or “explanation.” Algorithmic items were defined as those requiring a set procedure or series of steps to determine the answer. Such items may involve a mathematical calculation to determine a numerical solution or require students to recall or apply basic knowledge of a theory. Explanation items were defined as those requiring descriptive explanations, manipulation of algebraic expressions using variables to provide conceptual explanations, or synthesizing multiple pieces of knowledge together to determine an answer.

To investigate the relation between question type and engagement, instances of Constructive and Interactive engagement for each item for each student were tabulated based on question type. While there are four engagement modes, Passive and Active engagement were not investigated in this analysis because students do no generate new information at these lower modes. The Constructive and Interactive modes require students to generate new information, and the difference between these modes is based on whether that generation occurs independently or through dialogue. Therefore, investigating the relation of question type with these modes could provide insight into how and/or why question type promotes dialogue. Each item from each of the three activities was coded as either algorithmic or explanation. The activities were analyzed separately for the relation between question type and engagement mode. This was done to reduce variation caused by the fact that the tasks required by the items for each activity were quite different.

A 2 x 2 contingency table (Figure 2) was used to determine if a significant relation exists between the type of question being answered and the Constructive or Interactive engagement mode individual students showed in their response to each item in each activity.

	Algorithmic	**Explanation**
Constructive		
Interactive		

Figure 2: Contingency Table to compare student engagement with question type

Since a single student is represented multiple times in the dataset, because they answered multiple items, the two categorical variables are not completely independent. Therefore, McNemar’s chi-squared test was used to determine if a statistically significant relation exists.

Data Cleaning

Analysis of the timing and group size factors did not require data cleaning. Because analysis of the question type factor required a statistical test, data cleaning was required. This is because two sources of variation occurred due to the fact that 1) the students across a single activity completed differing numbers of items, and 2) some items had a large number of students who demonstrated Passive or Active engagement or did not answer the item at all. To reduce these sources of variation, the data was cleaned in three steps. First, for each activity, any item that a student did not answer or engaged at the Active or Passive mode was removed for that student. Second, the total number of items that each student answered at the Constructive or Interactive mode was tabulated. If this total was less than 50% of the items in the activity, all of that student's responses were removed from the data because by completing such a small part of the activity, these students would not be representative of group work across an entire activity. Third, for each item within an activity, the total number of students who answered the item was tabulated. If this total was less than 50% of the students who participated in the activity, the item was removed from the data because such a small sample of student responses to a specific item may not reflect how most students in the group would engage with that type of item. After cleaning, the data consisted of 11 students each in the Solutions and Dilutions and Periodic Trends activities, and 6 students in the Molecular Polarity activity. Six items were removed from the Solutions and Dilutions activity, leaving 12 items for analysis; 3 items were removed from Periodic Trends, leaving 34 items; and 3 items were removed from Molecular Polarity, leaving 21 items.

Trustworthiness

For the individual codes, trustworthiness was established by using investigator triangulation to determine credibility (Korstjens & Moser, 2018; Lincoln & Guba, 1985). This was accomplished through iterative revision of the codebook by authors S.Y.E., A.M., J.M.F., S.F., and A.J.H. until saturation was reached. The remaining data was coded by author S.Y.E. in consultation with author A.J.H. on ambiguous excerpts, and these excerpts were coded to consensus. For question type codes, the codebook was developed by author S.Y.E., and all items on all three activities were coded to consensus with a secondary coder (author A.S.). Data triangulation was also used to assess credibility, with student interviews providing a second data source. The interview responses were used to confirm observed trends of individual engagement across the course of the term.

Results and Discussion

Engagement Across Activities

Six of the twenty-three students who consented to this study participated in more than one activity. Figure 3 shows the distribution of each individual's engagement based on the activity they participated in. Tammy, Adriana, Mai, and Melissa participated in both the Solutions and Dilutions (conducted in week 3) and the Periodic Trends (conducted in week 7) activities. All the students except Adriana showed an increase in Interactive engagement; Tammy increased from 50% to almost 90%, Mai increased from 25% to 35%, and Melissa increased from 50% to over 70%. Tammy, Adriana, Melanie, and Molly participated in both the Periodic Trends activity and Molecular Polarity activity which was conducted in week 10; all the students except Adriana showed relative consistency in their Interactive engagement. Tammy's was close to 90% for both activities, Melanie's was approximately 65%, and Molly's was 40%.

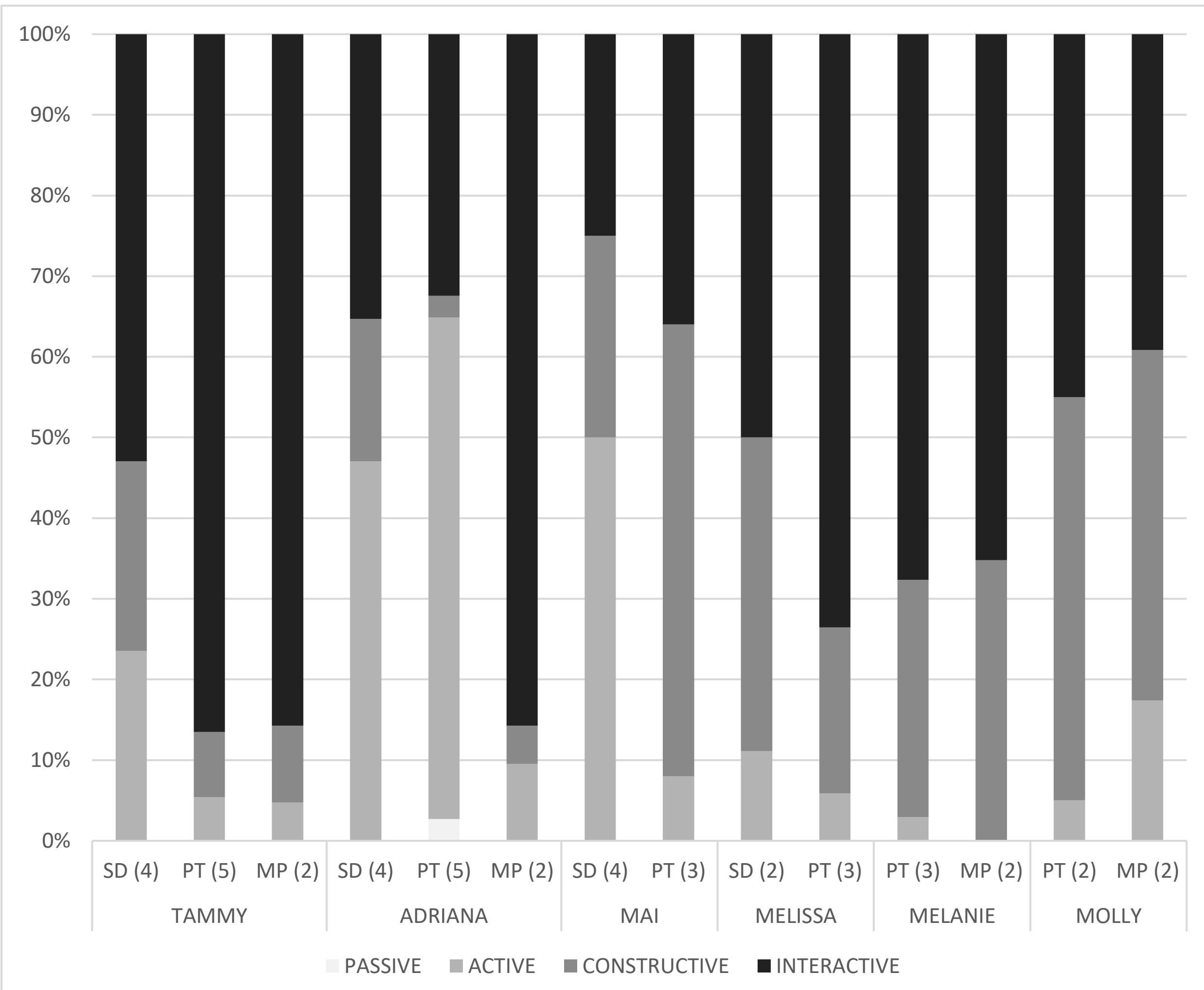


Figure 3: Individual students' engagement across activities. SD = Solutions and Dilutions, PT = Periodic Trends, MP = Molecular Polarity. Number in parentheses refers to group size.

However, Adriana did not follow this trend. She participated in all three activities and her Interactive engagement stayed stable around 30% for the first two activities and increased dramatically to approximately 85% for the third activity. For the first two activities, Adriana was in large groups (four people for Solutions and Dilutions and five people for Periodic Trends) and in a small two-person group with Tammy for the Molecular Polarity activity. During an interview after the Solutions and Dilutions activity, Adriana discussed how working in a group with students whose understanding varied resulted in students working at different speeds. She said (key ideas are in bold):

"…and I think it matters **what kind of group you're in, what people's levels are**. It is nice to have the variety, the range of like, somebody who doesn't know very much maybe and then somebody who knows more 'cause **if everybody knows a little bit, you can work it together**. But if you are in a group where myself, or the person feels like the others are way ahead, then that gets challenging 'cause you do feel like you're slowing everything down…So having, having a group that's kind of, I don't know if it's better, but **working with people that are a little bit in your range of knowledge or speed of understanding matters** because you don't want to feel like you're the one who's holding the group back from moving onto questions because you still don't understand it or you're just a little slower to understand all the concepts."

Because Adriana felt that the different levels of understanding created pressure on her as a student who did not work as fast (i.e., she did not want to hold up the group from moving forward), this may be what led her to engage more at a lower (e.g., Constructive or Active) mode during the first two activities. Example 1 shows the response of the five-person group consisting of Tammy, Adriana, Anita, Walt, and Kim to Key Question 19 from the Periodic Trends activity (PT-KQ19).

Example 1: Key Question 19 from the Periodic Trends activity (PT-KQ19) and the group response between Anita, Tammy, Walt, Kim, and Adriana

Describe how Boron, Aluminum, and Gallium are similar and different from one another.

2519 WALT: They all have the exact same number of valence electrons.

2520 TAMMY: Yep. Same valence electrons.

2521 ANITA: Also aluminum and gallium are metals. And boron is a metalloid.

2522 TAMMY: It is a metalloid, correct.

2523 ADRIANA: Say it another more time?

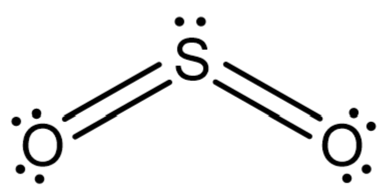
2524 ANITA: (speaking to Adriana) So aluminum and gallium are metals. Uh, boron is a metalloid.

Tammy and Walt discussed the fact that the three elements all had the same number of valence electrons (lines 2519 and 2520), and Anita added that gallium and aluminum are metals and boron is a metalloid (line 2521). Adriana began writing her answer after these statements were made. In addition, she asked for Anita to repeat her answer (line 2523). Furthermore, her written response contained only the information that was discussed in the group conversation. All these pieces of evidence suggest Adriana demonstrated Active engagement because she does not generate new information but instead manipulates information provided by other group members.

However, in the Molecular Polarity activity, Adriana worked with Tammy in a two-person group, and both students showed approximately 85% Interactive engagement. Example 2 shows their response to Exercise 2a (MP-EX2a). In this excerpt, Tammy and Adriana work together to determine the bond dipoles for the SO_2 molecule.

Example 2: Tammy and Adriana's group response to Exercise 2a from the Molecular Polarity activity (MP-EX2a)

The Lewis structure of SO_2 is provided below. Its molecular geometry is bent. (Note: sulfur is an exception to the octet rule.) Draw in the bond dipole moments.*



448 TAMMY: So now moving on to 2 with, SO_2, the molecular, it already gives us the molecular geometry, it's bent and sulfur is an exception to the octet rule.

449 ADRIANA: Correct.

450 TAMMY: Draw in the bond dipole moments. So S and O.

451 ADRIANA: So S, I have to look at this and-

452 TAMMY: I have to look up the electronegativity of S.

453 ADRIANA: (points to the periodic table) But I, I think- what's the rule, with the table? Is- it's, it's low to high?

**We acknowledge that there is disagreement in the literature regarding prioritizing the octet rule over reducing formal charge in writing Lewis structures* (Suidan et al., 1995).

454 TAMMY: I think so-

455 ADRIANA: So-

456 TAMMY: Sulfur is.

457 ADRIANA: Sulfur is less?

458 TAMMY: 2, sulfur's about uh, 2.58 and oxygen I believe...hold on. Oxygen's going to be 3.44.

459 ADRIANA: So it is more.

460 TAMMY: It is way more.

461 ADRIANA: So it should go, oh wait no the other direction.

462 TAMMY: So we want it to be, um, 3.44 minus the 2.58 is 0.86. So the difference is 0.86 and I think they're about equal for each side, right?

463 ADRIANA: Mm-hmm.

464 TAMMY: So one is going 0.86 and it's going away from the central atom and then the other way, 0.86 away from the central atom. Okay, so those are the dipole. So kind of like circle that, that's the dipole moments.

Tammy first recognized that she needed to know the electronegativity of sulfur, and Adriana built upon this idea by mentioning that lower electronegativity values are found in the lower rows on the periodic table (lines 452-453). Adriana also recognized the direction the bond dipoles will point, and Tammy then added on with the numerical electronegativity difference and the fact that the bond dipoles are equal on both sides (lines 461-462). Because Adriana and Tammy seemed to be at the same level of knowledge and were working at the same speed, they both displayed Interactive engagement because they successfully worked together by each contributing pieces of information and combining these pieces to generate the final answer.

Additional insights into how a student engages with both the activity and group members were gained through interview data. For example, during Adriana's interview after the Solutions and Dilutions activity, she was asked about how the activity helped her understanding of the material, and she talked about the importance of understanding why problems were solved in a specific way, not just how:

"Yeah, because in the moment [during group work], I still felt like I had done the problems, but **I still didn't fully understand where we were grabbing numbers from or why we were putting them in certain orders and what equations we were using**. And so it was one of those where I just copied, I see we're just grabbing numbers, we're placing them in equations, cool, but **I didn't understand the concept behind it and the idea of why we are putting those numbers there and why they should be put there**. So that didn't make sense, I just knew that's how I had to do it. So I was like, cool, I know how I have to do it, now I'll go home and figure out why I have to do it like that."

Tammy was also interviewed after the Solutions and Dilutions activity, and she discussed her positive opinion of group work saying:

"I've always liked group environment. I like talking things out, I am an auditory learner. I feel that if I am able to talk to someone and hear back, we just converse, and especially if **I'm able to teach it and teach it correctly, then that means I actually understand the concepts**…I don't like just teaching, I like to learn from others as well, like that kind of give and take, that back and forth, so you know, for the majority of the activity, I was in my wheelhouse, I knew what I was doing, so I was kind of leading it, but then as **we were going for more the conceptual things that was where they were coming in, they were teaching me**. I really appreciate it."

Both Tammy and Adriana mentioned the importance of conceptual understanding in their interviews. This attitude is shown in Example 2 by how they worked together to identify the steps and pieces of information needed to determine the bond dipole moments for the SO_2 molecule. The focus on a deeper understanding displayed by both students may have contributed to the development of a strong rapport between them. This may have also resulted in a higher comfort level for Adriana which caused her to more frequently engage at the Interactive mode with Tammy during the Molecular Polarity activity. The high amount of Interactive engagement could also be due to the small group size, which is discussed in the next section.

Group Size

Figure 4 shows the distribution of individual engagement codes for all students in a group based on group size, where groups consisted of two to five students. The figure shows variation in engagement levels across group sizes. Given that previous research suggests a positive correlation between academic outcomes and higher modes of engagement (Chi & Wylie, 2014; Menekse et al., 2013), exploring what aspects of group conversations for different group sizes lead to higher engagement could give insights into how to structure effective groups.

Across the five two-person groups, Groups 1-4 showed Interactive engagement of approximately 50% or less. Only Group 5 showed much higher Interactive engagement of approximately 85%. This group consisted of Tammy and Adriana and, as mentioned in the previous section, the high amount of Interactive engagement was likely due to their similar levels of knowledge and goals regarding group work. Since Group 5 was the only group to show such a high level of Interactive engagement, it seems likely that the high Interactive engagement was due to rapport between Tammy and Adriana based on their similar knowledge level and perception of group work, not necessarily the small group size.

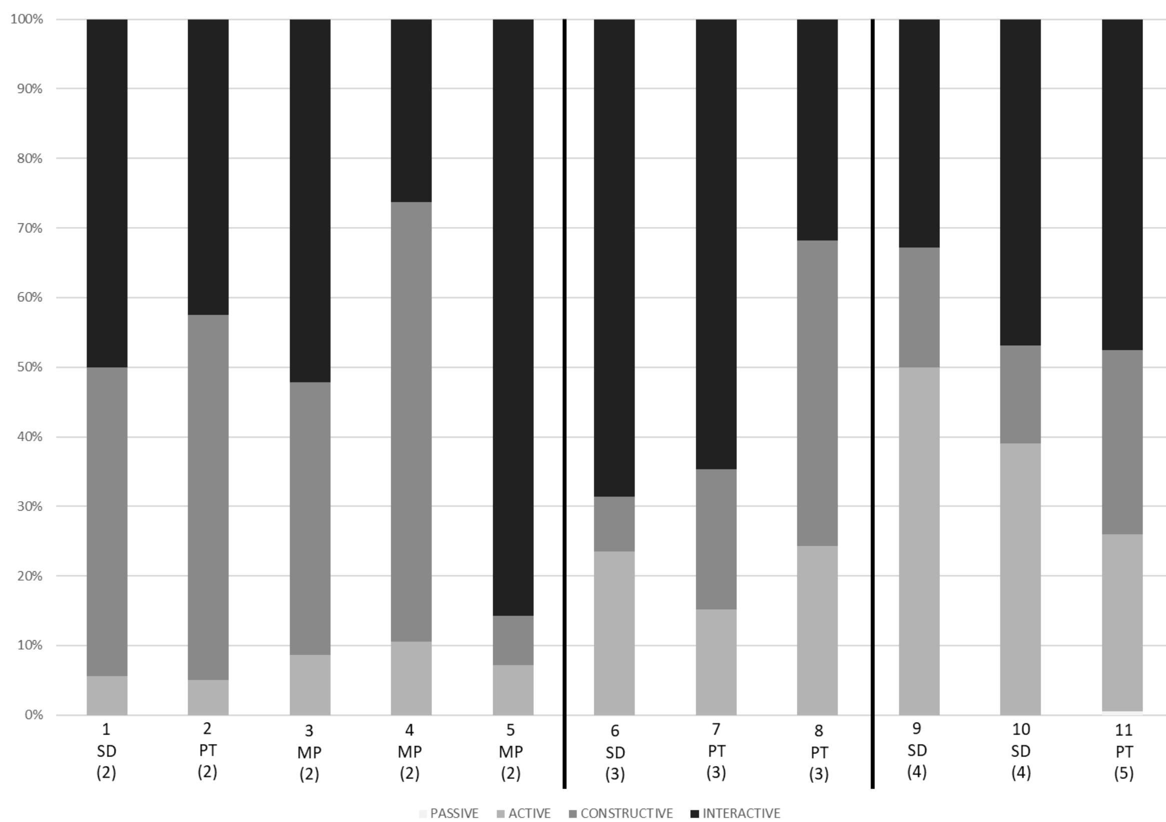


Figure 4: Distribution of individual engagement codes based on group size. Activity is identified for each group with SD = Solutions and Dilutions, PT = Periodic Trends, and MP = Molecular Polarity. Number in parentheses refer to number of students in a group.

In the three-person groups, Groups 6 and 7 showed between 60% and 70% Interactive engagement while Group 8 was much lower (approximately 35%). Groups 7 and 8 both worked on the Periodic Trends activity, and examination of conversation excerpts indicates that in Group 7, the three students seemed to be working at the same pace and knowledge level whereas this was not the case in Group 8. Example 3 shows an excerpt where Group 7, consisting of Mike, Melissa, and Melanie, collaborated to answer Exercise 7 from the Periodic Trends activity (PT-EX7). The students worked together by each contributing information and putting the pieces together to determine the final answer.

Example 3: Group 7's response to Exercise 7 from the Periodic Trends activity (PT-EX7)

What are the characteristics of an electron configuration when IE_1 is high?

1928 MELISSA: Okay. So what are the characteristics of an electron configuration when the, uh, ionization ener-, first ionization energy is high. So it has a free electron, right? It has a-, it has one valence electron.

1929 MELANIE: No, there's, there's none.

1930 MIKE: It's all noble gases.

1931 MELISSA: Oh, when the energy is high. Sorry. Yes.

1932 MELANIE: Yeah.

1933 MELISSA: Yep. Yeah. They're noble gases. So the, the valenc-, the shells are full.

1934 MIKE: Mm-hmm.

1935 MELANIE: Yeah. The shells are full

1936 MELISSA: And the atomic radius is small.

1937 MELANIE: They are happy. And they don't wanna be separated. Snug as a bug in a rug.

In this excerpt, Melissa began by incorrectly stating that an atom with high ionization energy would have a free electron, causing both Mike and Melanie to correct her by offering additional pieces of information; Melanie stated there would be no free electrons and Mike identified these atoms as being the noble gases (lines 1928-1930). Melissa then built on this by recognizing that this meant the valence shell would be full (line 1933). Additionally, the video recording shows that all three students do not write their answers until the conversation is over, indicating that no single student seemed to be working ahead and that the students were all working at the same knowledge level.

In contrast, in Group 8, which consisted of Henry, Rachel, and Mai, the lower amounts of Interactive engagement may be partially due to the disparate levels of knowledge of Henry and

Rachel. Example 4 shows part of Group 8's response to Key Question 11, which asks students to describe the trend in ionization energy as one moves down a group of the periodic table. In line 408 below, Henry gives a detailed description of how the ionization energy changes, using the idea of electron shells and referencing the s orbitals. However, the concept of orbitals is not introduced until the next model, suggesting that Henry had prior knowledge of this idea prior to answering this item.

Example 4: Portion of Group 8's response to Key Question 11 from the Periodic Trends activity (PT-KQ11)

Summarize how the first ionization energy changes as you move down a group (column) on the periodic table.

408 HENRY: So I know why it increases or, uh, how do you say the ionization, the energy decreases as you go down a group. As you move farther down, the sub shell count starts increasing. (HENRY uses hands to demonstrate) So like 1s, $1s^2$, and like $2s^2$, like when you reach like, uh, an element like xenon for example, there's a lot more electron configuration that you're gonna have to write down, (RACHEL and MAI nod heads) which means that the sub shell, there's gonna be a lot more sub shells within that element, which makes it bigger, but it doesn't make it more covalent in terms of a noble gas. But every other element, as it moves down, ionization energy decreases because the radius is increasing.

409 RACHEL: Mm-hmm.

410 HENRY: Because the amount of sub shells are increasing. As you move farther down.

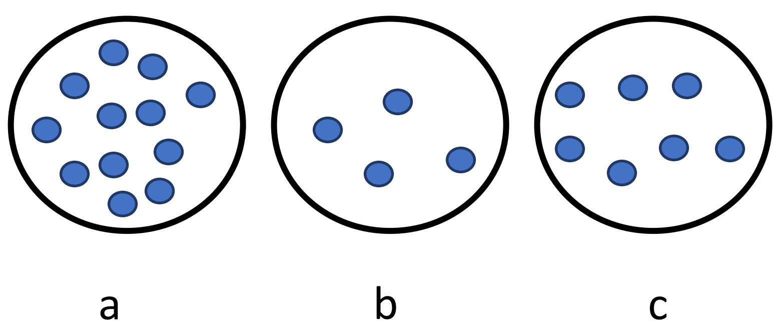
411 RACHEL: Mm-hmm. And that's like pulling apart the electron.

Additionally, Rachel seemed to struggle with understanding the concepts in this activity. For example, Exercise 1 was a multi-part question which asked students to determine between a pair of elements which one had the larger atomic radius. In the first two parts of this item, Rachel made statements such as, "I was confused" or "I don't understand". Such statements suggest that Rachel did not have prior knowledge coming into this activity, and this may have contributed to her high amount of Active engagement (33%) compared to Henry (9%). The discrepancy between Rachel's lack of prior knowledge and Henry's more advanced level of knowledge may partially explain the lower levels of Interactive engagement demonstrated by this group due to the fact that Rachel demonstrated higher amounts of Active engagement because she waited for someone to "give" her the answer while Henry demonstrated higher amounts of Constructive engagement (42%) because he used his higher level of knowledge to independently answer items or teach his fellow group members concepts as needed. In summary, analysis of conversation excerpts from small and medium groups supports the idea that grouping students with a similar amount of prior knowledge may foster an increase in Interactive engagement.

All three of the large groups (Groups 9-11) showed less than 50% Interactive engagement. This may be partially due to the fact that there are cases where not all group members are engaging at the Interactive mode when answering a specific item. For example, Group 10 was a four-person group consisting of Tammy, Amy, Mai, and Zoey. Example 5 shows their response to Key Question 5 from the Solutions and Dilutions activity (SD-KQ5). This excerpt shows that even though Tammy and Amy demonstrate Interactive engagement, Zoey showed Constructive engagement, and Mai engaged at the Active mode.

Example 5: Group 10's response to KQ5 in the Solutions and Dilutions Activity (SD-KQ5)

*The images below represent the same small volume within three different solutions and the spheres represent solute particles (solvent particles are not shown). Which solution has the <u>lowest concentration</u>? Circle your response and **explain** why you chose it.*



396 TAMMY: Okay, cool. Moving on! Number five. Images below represent the same small volume within three different solutions and spheres. And the spheres represent, sorry, solute particles, the solvent particles are not shown. Which solution has the lowest concentration? Circle the response and explain why.

397 AMY: Okay.

398 AMY: Say b, right?

399 TAMMY: I wanna say b.

400 AMY: Yeah.

401 TAMMY: Yeah, 'cause we just established in the last one that concentration means more of whatever the substance is.

402 AMY: Right, yeah. And explain what-. Yeah. There there's fewer...solute particles.

403 TAMMY: Yes. Not more, less. Sorry, less solute.

404 AMY: Mm-hmm.

Both Tammy and Amy were coded as Interactive because they co-generated information to produce the answer. In line 398, Amy said the initial answer, and in lines 401-402, Tammy and Amy co-generated the explanation for that initial answer. Mai was coded as Active because in the video, she looked at Tammy and Amy during their conversation and then looked at Zoey's worksheet prior to writing her answer. Additionally, on her worksheet, she circled option b and wrote, "There are fewer solute particles," nearly replicating Amy's exact wording from line 402. Both the timing of when Mai wrote her response and the content of what she wrote suggest that she simply manipulated information she received from her group members. On the other hand, in the video, Zoey wrote her answer before any conversation occurred. Additionally, on her worksheet, she wrote, "b has less because there are fewer substances inside than the others." This statement is phrased differently than what was said during the conversation, further supporting the interpretation that Zoey independently generated her answer and did not modify it based on Tammy and Amy's conversation; therefore, Zoey was coded as Constructive.

A second factor that may contribute to the lower amount of Interactive engagement is the idea of group splitting. In this group, Zoey and Mai were sitting next to one another and Tammy and Amy were sitting beside each other. The two pairs were across and slightly diagonal from each other. Each "split" group displayed their own engagement pattern, where Zoey and Mai showed higher amounts of Active and Constructive engagement, similar to the other two person groups that were observed (Figure 4). Tammy and Amy showed higher amounts of Interactive engagement, which was similar to what was observed between Tammy and Adriana. The larger group size and where the students were seated relative to each other may have been a contributing factor to the high level of interaction between Tammy and Amy and the lower engagement modes from Zoey and Mai. Since the engagement modes shown in Figure 4 were determined by summing the engagement modes for all students in the group, this can result in a lower amount of Interactive engagement for the group.

The results of this analysis did not identify an “optimal” group size; instead, analysis of group response excerpts at the different group sizes identified aspects of group dynamics that seemed to facilitate higher modes of engagement. Although large group size and group splitting may have contributed to the lower levels of Interactive engagement in groups of four and five students, in general, results of this study indicate that groups which contained students with similar knowledge bases demonstrated higher levels of Interactive engagement. Previous research on group dynamics suggests that one of the primary sources of problems in group environments is the presence of dominant and quiet students in the same group (Ahmed, 2014; Hendry et al., 2003). These studies defined a dominant student as someone who talks a lot and controls the direction of the conversation whereas a quiet student is one who rarely contributes to the conversation. Hendry et al. and Ahmed both hypothesized that the presence of dominant and quiet students in the same group led to tension in the group because the overbearing nature of the dominant student hindered the learning of the other students in the group since they felt inferior if their opinions differed from the dominant voice. This may have led to quieter students being unwilling to voice their opinions. Additionally, Hendry et al. and Ahmed suggested that the quiet students in the group were erroneously perceived to be simply “sponging” information from more vocal members, again leading to feelings of tension and conflict (Ahmed, 2014; Hendry et al., 2003). However, results of this study suggest that the dominant and quiet demeanors that students adopt in a group may stem from differing incoming knowledge bases. In the Example 4 excerpt above, Henry dominates the conversation possibly due to the fact that he had a higher level of knowledge than his groupmates. Therefore, our results suggesting that the observed engagement may be due to the differing levels of knowledge may also suggest an alternative reason for observed group dynamics between dominant and quiet students.

Question Type

Each activity was analyzed for the relation between the students’ engagement mode and the question type. For the Solutions and Dilutions activity, 16 of the 18 items (89%) were coded as algorithmic; for the Periodic Trends activity, 16 of the 37 items (43%) were coded as algorithmic; and for the Molecular Polarity activity, 11 of the 24 items (46%) were coded as algorithmic. The results of McNemar’s chi-squared tests indicate there is a significant relation between question type and engagement for the Solutions and Dilutions activity but not for the Periodic Trends or Molecular Polarity activities (Table 2).

Table 2: Results of McNemar's test. Numbers in parentheses are the percent of items at the level of engagement for a question type. Bold p-values indicate a significant relation.

		Algorithmic	**Explanation**	
Solutions and Dilutions	**Constructive**	14 (15)	6 (35)	n = 112 $\chi^2 = 64.655$, df = 1 **$p < 0.01$**
	Interactive	81 (85)	11 (65)	
	Total	95	17	
Periodic Trends	**Constructive**	45 (38)	64 (38)	n = 285 $\chi^2 = 0.59124$, df = 1 $p = 0.44$
	Interactive	73 (62)	103 (62)	
	Total	118	167	
Molecular Polarity	**Constructive**	23 (46)	17 (30)	n = 107 $\chi^2 = 2.2727$, df = 1 $p = 0.13$
	Interactive	27 (54)	40 (70)	
	Total	50	57	

The McNemar's test shows that a significant relation between question type and engagement only exists for the Solutions and Dilutions activity items. This result indicates that students are more likely to engage at the Interactive mode on algorithmic items over conceptual items during this activity. Students engaged at the Interactive mode 81 times on algorithmic items (85%) and 11 times on explanation items (65%). Previous research showed that students use higher order thinking skills on conceptual items and lower order thinking skills on algorithmic items (Zoller et al., 2002); therefore, we initially hypothesized that students may be more likely to engage at a lower mode for algorithmic items. However, our results suggest the opposite trend in this activity. Zoller et al. defined algorithmic questions as those which require the use of memorized procedures for their solution and defined lower-order cognitive skills questions (LOCS) as knowledge questions which required simple recall or application of known information and are solvable using a set of processes that can be applied through practice (Zoller et al., 2002). In this study, we defined algorithmic items more broadly, as including items which required a mathematical calculation and/or items which required students to use a set of procedural steps to determine an answer. Initially, we began by defining "calculation" items, which were narrowly defined as mathematical problems. However, through the coding process, we found that there were questions that did not fit as calculation but were not conceptual in nature either. This occurred specifically in the Periodic Trends activity and to a lesser degree in the Molecular Polarity activity. We then decided to widen our definition to include items where students followed a certain set of steps and renaming these “algorithmic” to alleviate this problem. Using this definition, the Solutions and Dilutions activity contained primarily algorithmic items which required students to perform a mathematical calculation whereas the Periodic Trends and Molecular Polarity activities contained procedural-based algorithmic items (e.g., write an electron configuration or draw a Lewis structure). Therefore, the items in the Solutions and Dilutions activity are algorithmic in accordance with Zoller’s definition. However, the items in the Periodic Trends activity such as drawing Lewis structures or determining electron configuration require student to follow a set of known steps, e.g., determine the total number of electrons, calculate how many electrons would go into each orbital. Using such steps falls under Zoller’s definition of LOCS, which may account for the differences in the McNemar’s test results.

In the Solutions and Dilutions activity, the algorithmic items where conversations had mostly Interactive engagement focused in two areas: 1) students working together to correctly associate numerical values with the correct variables in the dilution equation ($M_CV_C = M_DV_D$), and 2) determining the correct significant figures for their answer. For example, Group 10’s (Figure 4) conversation related to Exercise 4 from the Solutions and Dilutions activity (SD-EX4) illustrates this pattern (see full excerpt in the Appendix). SD-EX4 asked students to determine the volume of a stock solution needed to produce a known volume of a dilute solution at a known concentration. Tammy and Amy spent a large portion of the conversation attempting to identify what they were solving for and what the V_C and V_D variables referred to. Mai helped alleviate some of their confusion by recognizing that the concentration of the stock solution should be pulled from Model 2. Once the group had identified values for all the variables, the conversation shifted into a discussion of significant figures in which Zoey, Tammy, and Amy worked together to determine that they would need three significant figures. This analysis indicates that multiple facets of calculation-based items, i.e., correct association of numerical values to their appropriate variables and application of significant figures, can promote higher occurrences of Interactive engagement as students work together to complete these tasks. This group showed lower amounts of Interactive engagement in Key Question 5 from the same activity (Example 5) due to differing levels of engagement of each student and group splitting whereas their response to SD-EX4 showed higher amounts of Interactive engagement and did not show evidence of group splitting. This suggests that question type may contribute to the higher engagement on SD-EX4 versus SD-KQ5; therefore, question type may also influence behaviors that resulted in group splitting observed on SD-KQ5.

Conclusion

While there are many factors that may affect students' cognitive engagement while participating in small group AL activities, this study investigated some factors that are related to the group environment and the activity itself. Specifically, we looked at the effect of timing, group size, and question type. The three factors were explored through the following research questions:

How does students' cognitive engagement vary across activities?

Analysis of six students across three activities throughout the term indicates that in general, students' Interactive engagement increases during the first half of the term but stabilizes during the second half. However, this was not consistent for every student, and we cannot draw any firm conclusions regarding the effect of timing on engagement based on this analysis.

Nevertheless, it is possible that other factors, such as students' perception of group work and their individual personal goals, may impact their mode of engagement. For example, Tammy had a very positive perception of group work and displayed high levels of Interactive engagement across all activities. In contrast, Adriana's opinion of group work was more reserved, and she mentioned that the success of group work was dependent on the type of group and specifically, people's level of understanding and speed at which they worked. Accordingly, her Interactive engagement remained low in the first two activities where group members worked faster than her but increased in the third activity, where she and Tammy worked at similar speeds with similar goals.

Although students' engagement varied across activities over the course of the term, it is also possible that the topic being presented and the type of questions being asked in the activity may also affect students' engagement, which was explored through the next research question.

What relation is observed between the type of question asked in the activities and students' level of cognitive engagement in General Chemistry?

The results of McNemar's chi-squared test showed a significant relation between question type and engagement mode for the Solutions and Dilutions activity, but not for the Periodic Trends or Molecular Polarity activities. In the Solutions and Dilutions activity, algorithmic items were associated with higher occurrences of Interactive engagement. Algorithmic items in this activity generally asked students to perform a mathematical calculation to determine the answer, and analysis of student conversations indicated that students engaged at the Interactive mode to correctly associate numerical values with the appropriate variables and to correctly apply significant figures. Although this result was significant, 89% of the items on the Solutions and Dilutions activity were algorithmic, resulting in a skewed dataset. In comparison, the Periodic Trends and Molecular Polarity activities had a more balanced distribution of algorithmic and explanation items (43% and 46% algorithmic items, respectively). Since there were very few explanation items in the Solutions and Dilutions activity, this analysis should be repeated with a more balanced spread of algorithmic and explanation items to determine if a significant relation would be obtained again.

It should be noted that this study broadly defined algorithmic items as those which required a mathematical calculation and/or asked students to use a set of procedures to answer the question whereas other studies (Zoller et al., 2002) defined algorithmic more narrowly. In addition to having a better balance between the different question types, the algorithmic items in the Periodic Trends and Molecular Polarity activities did not involve calculations but instead asked students to apply a set of procedural steps to complete a task. Although the raw data in Table 2 does show higher occurrences of Interactive engagement on algorithmic items for these activities, the relation is not statistically significant. It is possible that these different definitions of algorithmic may have contributed to these observed differences. Zoller's study indicated higher performance on algorithmic items compared to LOCS items which may correlate to the observed higher engagement on the items in the Solutions and Dilutions activity since these items align with Zoller's definition of algorithmic (Zoller et al., 2002). Repeated analysis of student engagement on these items and on the items on the Periodic Trends and Molecular Polarity activities which could be considered LOCS may improve our understanding of the relation between student engagement and question type.

Since timing and question type both seem to be related to engagement, there may be conflation between these factors. For example, the Periodic Trends and Molecular Polarity activities had similar proportions of explanation items and they both occurred during the second half of the term. For students who showed an increase in Interactive engagement from the Solutions and Dilutions activity to either the Periodic Trends or Molecular Polarity activity, this may be due to a combination of the effects of timing and question type.

What is the effect of group size on individual students' cognitive engagement?

The results of this analysis do not definitively suggest an "optimal" group size; however, the analysis did indicate that higher amounts of Interactive engagement occurred in groups that contained two or three students when the students in these groups had similar levels of prior knowledge. In larger groups of four or five students, lower amounts of Interactive engagement were observed. This is because only a portion of the group typically demonstrates Interactive engagement, while the remaining students show lower modes of engagement. In these larger groups, students may be more likely to split into sub-groups where each sub-group would have a different group dynamic. Depending on multiple factors affecting group dynamics, including perceptions of group work and students' level of prior knowledge, these smaller groups could resemble the previously observed engagement distributions for smaller groups.

Further analysis into the effect of group size on engagement found that behaviors such as group splitting were not consistent within a single group on an activity. This result suggests that there may be interplay between the various factors investigated in this study.

Limitations

This study investigated the effect of timing, group size, and question type; however, individual student characteristics described in previous literature, such as gender identity and students' aspirations (Fullarton, 2002) and academic capability (Lee et al., 2022), may also influence a student's engagement and may contribute to the observed results. In addition, other student-level variables which may affect engagement, such as GPA or SAT/ACT math scores, were not considered when selecting students for participation in this study. As this is a qualitative study with a small sample size from a single term of a General Chemistry I course, these results may not be generalizable to other courses or activities. Additionally, the ICAP framework assumes that the overt behaviors students display are reflective of their internal cognitive engagement; however, this may not always be the case. For example, students may independently generate information (Constructive engagement) while their conversation may show only Active modes. In future studies, additional reflective interviews with students while they are reviewing the group interaction video, i.e., stimulated recall, may be able to address this (Dempsey, 2010).

In addition, the observed groups in this study did not remain the same across multiple activities. Therefore, for each activity, students were working with new peers for the first time and had to learn how to communicate and work together. Since students may interact with one another differently with a different set of group members, the engagement of individual students may have been affected. It is also possible that factors related to group composition could influence engagement. In this study, after students were selected into the initial student pool to maximize racial and gender diversity, the students were assigned to groups randomly. As a result, if a student felt marginalized in their group based on gender, race, or other identity-related characteristics, this may have affected their comfort level and engagement. Although our analysis of timing, group size, and question type suggested that higher modes of engagement were observed when students in the groups had similar knowledge bases, this study was not designed to investigate the effect of ability level on engagement. Our results are based on analyses of written transcripts and observations of behavioral interactions in the video recordings; however, no *a priori* measurement of prior knowledge was used.

Finally, this work looked at unstructured groups, and the results may not be applicable to highly structured groups, such as those used in POGIL. While POGIL groups are larger, generally containing 4-6 students, each student is assigned a specific role (e.g., manager, recorder, reflector) (Farrell et al., 1999). The duties of each assigned role may affect that student's engagement; for example, the recorder may show lower modes of engagement using ICAP as their primary role is to record the group's thoughts and answers, and as a result, they may be less likely to verbally contribute information to the conversation.

Implications for Instructors

The results of the analysis of students' engagement across multiple activities suggest that there may be opportunities for instructors to influence students early in the term. The stabilization of engagement in the second half of the term (during the Periodic Trends and Molecular Polarity activities) suggests that students may have established their academic habits and may be less willing to change. Therefore, we would encourage instructors to continually emphasize the benefits of group work and specifically the type of conversations in which students should be engaging. While it is possible that changes in engagement across the term may not be solely due to timing, there may be conflation with the type of questions being asked. Instructors could address this by giving examples of what productive conversations would look like for different question types. For example, instructors may want to encourage students to talk through the specific steps of an algorithmic item requiring a mathematical calculation or clearly discuss their thought processes behind their answer when asked to make a prediction on a conceptual item.

The group size analysis suggests that students should work in smaller groups and that students with similar knowledge levels should be grouped together to enhance productive conversation and Interactive engagement. While it may be difficult to determine which students have similar knowledge levels and group them accordingly, we encourage instructors to have students form smaller groups whenever possible and continually emphasize the importance of *all* students participating in the group conversation, regardless of size. We would also suggest that instructors continually discuss the idea that group work is intended to improve the understanding and knowledge of all students and that each student may be able to bring different insights or perspectives to the activity. The instructors can highlight that this will occur through conversations with fellow group members.

Implications for Research

This study analyzed the engagement of individual students in a group environment and found that factors such as timing, group size, and question type may affect individual students' engagement. Although previous research found a correlation between higher modes of engagement and improved achievement outcomes (Chi & Wylie, 2014; Menekse et al., 2013; Wiggins et al., 2017), further research is needed to explore the relation between these factors which influence engagement and learning outcomes. Additionally, previous research investigating the effect of same versus mixed-ability groups on academic performance indicated that mixed-ability groups benefit lower-attaining students; however, there are mixed results regarding the effect on high-attaining students (Lejk et al., 1999; Linchevski & Kutscher, 1998; Venkatakrishnan & Wiliam, 2003). An exploration of the effect of grouping by ability level on engagement could provide deeper insights into the relation between ability level, engagement, and academic performance. Furthermore, research into individual students' engagement when group composition remains constant may provide additional insight into the effect of other factors on engagement, such as students' sense of belonging, active learning environment, and instructor support (Craft & Capraro, 2017; De Loof et al., 2021; Gasiewski et al., 2012; Struyf et al., 2019; Wilson et al., 2015). Finally, the analysis of Tammy's and Adriana's engagement suggests that students' perceptions of group work and individual student goals regarding the activities may influence their engagement as well and should be explored.

Conflicts

There are no conflicts of interest to declare.

Acknowledgements

This study is supported by the National Science Foundation under Grant No. 2120843. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Appendix

SD-EX 4) What volume of the stock solution in Model 2 would you need to prepare 20.0 mL of a dilute solution with [$C_{12}H_{22}O_{11}$] = 0.1406 M?

1726 AMY: So you want the-

1727 TAMMY: The volume of the stock solution, right?

1728 AMY: Yeah. You need the volume of diluted. Is that V_D we're finding?

1729 TAMMY: So volume of the stock solution is gonna be concentrate cause of the Vc.

1730 AMY: Mm-hmm.

1731 TAMMY: Concentrated solution, and it says here, volume of the stock solution, right here on this side says Vc. (points to the model)

1732 AMY: Okay. Would you need to prepare 20? Oh, okay.

1733 TAMMY: Okay then.

1734 AMY: So we're solving for Vc.

1735 TAMMY: Solving for Vc. So then you wanna isolate the equation that way, right?

1736 AMY: Mm-hmm.

1737 TAMMY: So, and it already gives it right the back. What that is, right here. (points to equation in the model)

1738 AMY: Dang.

1739 TAMMY: So M_D…

1740 AMY: Alright, that's nice.

1741 TAMMY: M_D is V_D over Mc.

1742 AMY: Right, okay.

1743 TAMMY: Okay.

1744 MAI: So what would our Mc be?

1745 TAMMY: That is a good question. [TAMMY laughs]

1746 MAI: So are we using Model 2?

1747 AMY: Yeah.

1748 MAI: To like, we're gonna replace…Basically the top what I'm understanding and then we use 0.5625 as-, Uh, I'm sorry, my brain is moving too fast.

1749 AMY: Which one is which?

1750 TAMMY: Okay, so hold on. Oh, you want volume of the stock solution?

1751 AMY: Mm-hmm.

1752 TAMMY: Would you need to prepare 20 milliliters of the diluted solution-

1753 AMY: So that's-

1754 TAMMY: Of the molarity.

1755 AMY: Alright. So that's V_D and M_D that they give us.

1756 TAMMY: So molarity of the diluted solution is M_D.

1757 TAMMY: And then there's the molarity of the dilution. So the molarity is going to be 0.1406 M, okay?

1758 AMY: Yeah, yeah.

1759 TAMMY: That's that one. Okay. So that's M_D, then V_D volume of the dilute solution-

1760 AMY: That'd be the 20 milliliters, right?

1761 TAMMY: Okay. So 20 milliliters of that solution. That makes sense to me. Cause that's the volume of the dilute solution. Okay.

1762 AMY: So that means we just copy over the Mc.

1763 TAMMY: Right? What is the Mc?

1764 AMY: Oh cause it's a stock solution of .565 moles.

1765 TAMMY: Yeah. Yep, yep, yep, yep.

1766 AMY: That makes sense.

1767 TAMMY: That makes sense, so that is the stock solution. So Mc equals 0.5625 molarity. Okay. Then you just plug those in.

1768 AMY: And the unit should cancel that for-

1769 TAMMY: Yes. Yes. Absolutely. Yes. Um, we need to convert milliliters to liters.

1770 AMY: You're so right.

1771 TAMMY: Yes. You need to convert that cause otherwise, um, it's gonna be the wrong answer.

1772 AMY: Yeah.

1773 TAMMY: So I need to do that before, milliliters.

1774 MAI: Do we need to convert it to milliliters?

1775 TAMMY: Yes. Because molarity will always be moles over liters, always

1776 MAI: But we're not solving for molarity, we're solving for volume.

1777 AMY: Oh, you're right. So we only need milliliters, yeah.

1778 MAI: We don't need to convert that-

1779 AMY: Okay, and in the example they, they kept it as milliliters.

1780 TAMMY: Oh, all right. Thank you.

1781 AMY: Like one less step.

1782 TAMMY: Perfect. And then Mc is 0.5625. Perfect and then you just do math from there.

1783 AMY: Yeah.

1784 AMY: Did you get 4.999?

1785 ZOEY: Yeah but I'm thinking since it's a sig fig, or like do we need to do that?

1786 AMY: Oh yeah.

1787 TAMMY: Yep. How many sig figs would we have? Three.

1788 AMY: You're right. I was thinking the decimals.

1789 AMY: So 4.99.

1790 MAI: So, I was thinking cause it's like nine and nine, right?

1791 AMY: Oh.

1792 MAI: So would it be like five?

1793 ZOEY: 5.-

1794 TAMMY: 5.00?

1795 AMY: Yeah. Just throw in extra zeros to make it.

1796 MAI: Okay. So I just wanted to make sure.

References

Ahmed, Z. (2014). Problems of Group Dynamics in Problem Based Learning Sessions. *J Ayub Med Coll Abbottabad*, *26*(2), 230–234. http://www.ayubmed.edu.pk/JAMC/26-2/Zafar.pdf

Alexopoulou, E., & Driver, R. (1996). Small-Group Discussion in Physics: Peer Interaction Modes in Pairs and Fours. *Journal of Research in Science Teaching*, *33*(10). https://doi.org/10.1002/(SICI)1098-2736(199612)33:10<1099::AID-TEA4>3.0.CO;2-N

Bruce, M., Omne-Pontén, M., & Gustavsson, P. J. (2010). Active and emotional student engagement: A nationwide, prospective, longitudinal study of Swedish nursing students. *International Journal of Nursing Education Scholarship*, *7*(1). https://doi.org/10.2202/1548-923X.1886

Chi, M. T. H., Adams, J., Bogusch, E. B., Bruchok, C., Kang, S., Lancaster, M., Levy, R., Li, N., McEldoon, K. L., Stump, G. S., Wylie, R., Xu, D., & Yaghmourian, D. L. (2018). Translating the ICAP Theory of Cognitive Engagement Into Practice. *Cognitive Science*, *42*(6), 1777–1832. https://doi.org/10.1111/cogs.12626

Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Chou, P. N., & Chang, C. C. (2018). Small or Large? The Effect of Group Size on Engineering Students' Learning Satisfaction in Project Design Courses. *Eurasia Journal of Mathematics, Science and Technology Education*, *14*(10). https://doi.org/10.29333/ejmste/93400

Cossé, T. J., Ashworth, D. N., & Weisenberger, T. M. (1999). The Effects of Team Size in a Marketing Simulation. *Journal of Marketing Theory and Practice*, *7*(3). https://doi.org/10.1080/10696679.1999.11501844

Cracolice, M. S., Deming, J. C., & Ehlert, B. (2008). Concept learning versus problem solving: A cognitive difference. *Journal of Chemical Education*, *85*(6), 873–878. https://doi.org/10.1021/ed085p873

Craft, A. M., & Capraro, R. M. (2017). Science, Technology, Engineering, and Mathematics Project-Based Learning: Merging Rigor and Relevance to Increase Student Engagement. In *Electronic International Journal of Education, Arts, and Science* (Issue 6). http://www.eijeas.com

De Loof, H., Struyf, A., Boeve-de Pauw, J., & Van Petegem, P. (2021). Teachers' Motivating Style and Students' Motivation and Engagement in STEM: the Relationship Between Three Key Educational Concepts. *Research in Science Education*, *51*, 109–127. https://doi.org/10.1007/s11165-019-9830-3

Dempsey, N. P. (2010). Stimulated recall interviews in ethnography. *Qualitative Sociology*, *33*(3), 349–367. https://doi.org/10.1007/s11133-010-9157-x

El-Mansy, S. Y., Barbera, J., & Hartig, A. J. (2022). Investigating small-group cognitive engagement in general chemistry learning activities using qualitative content analysis and the ICAP framework. *Chemistry Education Research and Practice*, *23*(2), 335–347. https://doi.org/10.1039/d1rp00276g

Farrell, J. J., Moog, R. S., & Spencer, J. N. (1999). A guided-inquiry general chemistry course. *Journal of Chemical Education*, *76*(4), 570–574. https://doi.org/10.1021/ed076p570

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Fullarton, S. (2002). *Student engagement with school: Individual and school-level influences*. ACER.

Gasiewski, J. A., Eagan, M. K., Garcia, G. A., Hurtado, S., & Chang, M. J. (2012). From Gatekeeping to Engagement: A Multicontextual, Mixed Method Study of Student Academic Engagement in Introductory STEM Courses. *Research in Higher Education*, *53*(2), 229–261. https://doi.org/10.1007/s11162-011-9247-y

Guo, J. P., Lv, S., Wang, S. C., Wei, S. M., Guo, Y. R., & Yang, L. Y. (2023). Reciprocal modeling of university students' perceptions of the learning environment, engagement, and learning outcome: A longitudinal study. *Learning and Instruction*, *83*. https://doi.org/10.1016/j.learninstruc.2022.101692

Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science*, *332*(6034), 1213–1216. https://doi.org/10.1126/SCIENCE.1204820/SUPPL_FILE/HAAK.SOM.PDF

Harris, R. B., Mack, M. R., Bryant, J., Theobald, E. J., & Freeman, S. (2020). Reducing achievement gaps in undergraduate general chemistry could lift underrepresented students into a "hyperpersistent zone." *Science Advances*, *6*(24), 1–9. https://doi.org/10.1126/sciadv.aaz5687

Hendry, G. D., Ryan, G., & Harris, J. (2003). Group problems in problem-based learning. *Medical Teacher*, *25*(6), 609–616. https://doi.org/10.1080/0142159031000137427

Kahu, E. R., Picton, C., & Nelson, K. (2020). Pathways to engagement: a longitudinal study of the first-year student experience in the educational interface. *Higher Education*, *79*(4), 657–673. https://doi.org/10.1007/s10734-019-00429-w

Korstjens, I., & Moser, A. (2018). European Journal of General Practice Series: Practical guidance to qualitative research. Part 4: Trustworthiness and publishing. *European*

Journal of General Practice, *24*(1), 120–124. https://doi.org/10.1080/13814788.2017.1375092

Lee, J., Park, T., & Davis, R. O. (2022). What affects learner engagement in flipped learning and what predicts its outcomes? *British Journal of Educational Technology*, *53*(2), 211–228. https://doi.org/10.1111/bjet.12717

Lejk, M., Wyvill, M., & Farrow, S. (1999). Group assessment in systems analysis and design: A comparison of the performance of streamed and mixed-ability groups. *Assessment and Evaluation in Higher Education*, *24*(1), 5–14. https://doi.org/10.1080/0260293990240101

Linchevski, L., & Kutscher, B. (1998). Tell Me with Whom You're Learning, and I'll Tell You How Much You've Learned: Mixed-Ability versus Same-Ability Grouping in Mathematics. *Journal for Research in Mathematics Education*, *29*(5), 533–554. https://doi.org/10.2307/749732

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry*. SAGE.

Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & D'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research*, *66*(4), 423-458. https://doi.org/10.3102/00346543066004423

Menekse, M., Stump, G. S., Krause, S., & Chi, M. T. H. (2013). Differentiated overt learning activities for effective instruction in engineering classrooms. *Journal of Engineering Education*, *102*(3), 346–374. https://doi.org/10.1002/jee.20021

Mondada, L. (2019). Contemporary issues in conversation analysis: Embodiment and materiality, multimodality and multisensoriality in social interaction. *Journal of Pragmatics*, *145*, 47–62. https://doi.org/10.1016/j.pragma.2019.01.016

Rahman, T., & Lewis, S. E. (2020). Evaluating the evidence base for evidence-based instructional practices in chemistry through meta-analysis. *Journal of Research in Science Teaching*, *57*(5), 765–793. https://doi.org/10.1002/tea.21610

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, *50*(4), 696-735. https://doi.org/10.2307/412243

Struyf, A., De Loof, H., Boeve-de Pauw, J., & Van Petegem, P. (2019). Students' engagement in different STEM learning environments: integrated STEM education as promising practice? *International Journal of Science Education*, *41*(10), 1387–1407. https://doi.org/10.1080/09500693.2019.1607983

Suidan, L., Badenhoop, J. K., Glendening, E. D., & Weinhold, F. (1995). Common Textbook and Teaching Misrepresentations of Lewis Structures. *Journal of Chemical Education*, *72*(7), 583–586. https://doi.org/10.1021/ed072p583

Surif, J., Ibrahim, N. H., & Dalim, S. F. (2014). Problem Solving: Algorithms and Conceptual and Open-ended Problems in Chemistry. *Procedia - Social and Behavioral Sciences*, *116*, 4955–4963. https://doi.org/10.1016/j.sbspro.2014.01.1055

Treen, E., Atanasova, C., Pitt, L., & Johnson, M. (2016). Evidence From a Large Sample on the Effects of Group Size and Decision-Making Time on Performance in a Marketing Simulation Game. *Journal of Marketing Education*, *38*(2), 130-137. https://doi.org/10.1177/0273475316653433

Venkatakrishnan, H., & Wiliam, D. (2003). Tracking and mixed-ability grouping in secondary school mathematics classrooms: A case study. *British Educational Research Journal*, *29*(2), 189–204. https://doi.org/10.1080/0141192032000060939

Wiggins, B. L., Eddy, S. L., Grunspan, D. Z., & Crowe, A. J. (2017). The ICAP Active Learning Framework Predicts the Learning Gains Observed in Intensely Active Classroom Experiences. *AERA Open*, *3*(2), 1–14. https://doi.org/10.1177/2332858417708567

Wilkinson, I. A. G., & Fung, I. Y. Y. (2002). Small-group composition and peer effects. *International Journal of Educational Research*, *37*(5), 425-447. https://doi.org/10.1016/S0883-0355(03)00014-4

Wilson, D., Jones, D., Bocell, F., Crawford, J., Kim, M. J., Veilleux, N., Floyd-Smith, T., Bates, R., & Plett, M. (2015). Belonging and Academic Engagement Among Undergraduate STEM Students: A Multi-institutional Study. *Research in Higher Education*, *56*(7), 750–776. https://doi.org/10.1007/s11162-015-9367-x

Zoller, U., Dori, Y. J., & Lubezky, A. (2002). Algorithmic, LOCS and HOCS (chemistry) exam questions: Performance and attitudes of college students. *International Journal of Science Education*, *24*(2), 185–203. https://doi.org/10.1080/09500690110049060