JOURNAL OF CHEMICAL EDUCATION

pubs.acs.org/jchemeduc Article

Tracking Student Argumentation Skills across General Chemistry through Argument-Driven Inquiry Using the Assessment of Scientific Argumentation in the Classroom Observation Protocol

Kathryn N. Hosbein, Meghan A. Lower, and Joi P. Walker*

Cite This: https://doi.org/10.1021/acs.jchemed.0c01225 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

Downloaded via EAST CAROLINA UNIV on May 4, 2021 at 14:51:33 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

ABSTRACT: Current research suggests that students often fail to focus equally on the three aspects of scientific argumentation (cognitive, epistemic, and social) when they are given an opportunity to engage in argumentation. This study examined student argumentation within a two-semester general chemistry laboratory sequence at East Carolina University to explore how the three aspects of argumentation change over time with repeated exposure through the Argument-Driven Inquiry (ADI) instructional model for laboratory instruction. Video recordings of group argumentation across five investigations were transcribed and coded using the Assessment of Scientific Argumentation in the Classroom (ASAC) observation protocol. A positive increase was seen in the total ASAC scores for each of the experiments. A significant increase was seen within each of the three subcategories of the ASAC observation protocol, cognitive, epistemic, and social, over the two-semester sequence. These results support the idea that increased opportunities to engage in argumentation improves an essential scientific practice.

KEYWORDS: *First-Year Undergraduate/General, Chemical Education Research, Laboratory Instruction, Problem Solving/Decision Making, Inquiry-Based/Discovery Learning, Student-Centered Learning, Kinetics*

FEATURE: Chemical Education Research

Argumentation plays a central role in the development, evaluation, and validation of scientific knowledge and is an important practice that makes science different from other ways of knowing.[1−3] The Next Generation Science Standards (NGSS) state that engaging in argument from evidence is a key practice within science and engineering.[4] Accordingly, the Framework for K-12 Science Education,[5] which serves as the foundation for the Next Generation Science Standards, emphasizes that students learn to argue from evidence. The argumentation and analysis that relate evidence and theory are essential features of science; scientists need to be able to examine, review, and evaluate their own knowledge and ideas and critique those of others.[4]

The chemistry laboratory provides a setting where students can engage in argumentation repeatedly with various experiments. In chemistry laboratory settings, argumentation can include evaluation of data quality, modeling of scientific theories, development of new testable questions from the experiment design, and modification of the design as evidence indicates.[2] Current research suggests that students often fail to equally focus on the three aspects of scientific argumentation (cognitive, epistemic, and social) when they are given an opportunity to engage in argumentation, whether it be proposing, supporting, challenging, or revising claims with their peers in a classroom setting.[6]

■ THEORETICAL FRAMEWORK

This study defined scientific argumentation as comprised of three components: the *claim*, the supporting *evidence*, and the *justification*.[7] The claim is the students' answer to the problem or question. The evidence consists of the collected data and the analysis of the data, which are then used to construct the argument. The justification is the explanation of the significance of results as well as the connection of data and analysis to the overarching concepts of the experiment. Engaging in the subjects of science and engineering should generate the need for argumentation through the need to defend an idea or explain a phenomenon.[3] Students should be engaged in the argumentation practice to the point where they argue for their proposed claims and explanations, defend their analysis and justification of the data, and advocate for the designs they propose.[3]

There have been various frameworks developed to facilitate instructional activities and research using scientific argumentation,[2,8−12] some of which have been critiqued in terms of their assumptions and context dependence.[13] This study defines the term *scientific argumentation*, as described by Sampson et al.,[15] as a social and collaborative process of proposing, supporting, evaluating, and refining ideas in an effort to make sense of a complex or ill-defined problem or to advance knowledge in a manner that is consistent with conceptual structures, cognitive

Received: September 23, 2020
Revised: March 29, 2021

ACS Publications © XXXX American Chemical Society and Division of Chemical Education, Inc.

processes, epistemological commitments, and the social norms of science (p 239).

We consider the science laboratory an *Epistemic Community of Practice* which provides opportunities to apply disciplinary knowledge to engage in an epistemic practice within a social setting. In argument construction students need to use conceptual structure and cognitive processes that are critical to science, i.e., the *need-to-know* content promoted in traditional science curricula. The use of scientific knowledge to explain the natural world facilitates moving from beliefs or opinions to a reliance on science. In order to participate in science as practice, the focus shifts to the *need-to-do*. This requires proficiency with epistemic frameworks that characterize science, such as the use of evidence to support a claim and the evaluation of the evidence with scientific knowledge. Finally, social norms of the science community pertaining to knowledge formation and communication should be followed when participating in scientific argumentation. This requires not just a discussion but also the uptake of criticism, tolerance for dissent, and changing views.[14] Frameworks, such as Toulmin's Argument Pattern, tend to ignore social interactions during an episode of argumentation.[8,15] Therefore, successful engagement in argumentation should include the simultaneous incorporation of three integrated domains: *conceptual* structures and *cognitive processes*, *epistemic* frameworks, and *social* processes and contexts.[16] It is essential to understand and track the progress of how students attend to each of these aspects of argumentation and how they contribute to the students' overall capacity to engage in this essential scientific practice over time. This process can be difficult for researchers because argumentation is often nonlinear in nature and the various components of a verbal argument (e.g., data, warrants, and backings) are difficult to identify.[7]

■ ARGUMENT-DRIVEN INQUIRY

The Argument-Driven Inquiry (ADI) instructional method is designed to elicit scientific practices, such as argumentation and its three aspects. ADI is comprised of a four week cycle for a single experiment (Figure 1). The first week for each experiment consists of a *prelaboratory activity* in which students are able to prepare for the inquiry investigation, become familiar with the equipment and techniques used, and gather information to complete their investigation proposal. The *proposal* is completed at the end of the prelaboratory activity where students plan their investigation for the following week to answer the guided question provided for the particular experiment. Week two begins with students conducting their investigation and collecting data. After completing the investigation, students analyze data and generate a *tentative argument* that answers the guiding question provided for the experiment. Their claim, evidence, and justification are displayed on a student-generated whiteboard. Groups rotate throughout the classroom and engage in an *argumentation session*, where they critique and share their findings and claims. Students then generate *investigation reports* that go through a double-blind *peer review* process during week three. Following peer review, students are able to *revise their report and submit* it to be graded by the instructor. The submission of the report is completed outside of class after the peer review process and thus does not occupy an extra week in the laboratory schedule. The four week cycle begins for a new experiment in the same laboratory session as the previous experiment's double-blind peer review proceeds. After the peer review session is complete, students engage in a prelab activity for the next experiment and generate a corresponding proposal. Experiment cycles overlap to where only 3 weeks are used for each experiment in the laboratory.

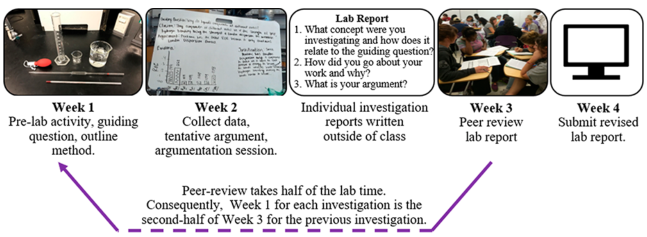

Figure 1. Argument-driven Inquiry (ADI) instructional model expressed through a four week cycle for each experiment, as shown in the top row. Adapted with permission from ref 22 (p 437, Figure 1). Copyright 2019 American Chemical Society and Division of Chemical Education, Inc.

There have been several studies that highlight the impacts of ADI in the general chemistry laboratory within both a community college and a minority-serving comprehensive university setting. Within community college settings, students participating in ADI have shown a higher increase in positive attitudes toward chemistry,[17] student use of evidence and reasoning,[17] student use of scientific practices,[18] and quality of argumentation[19] when compared to students participating in laboratories with traditional instruction. Quality of both oral and written argumentation has been shown to increase across the course of a semester.[7] Additionally, ADI has been shown to improve student ability to write in science.[20,21] Within the minority-serving university setting, the three aspects of argumentation (cognitive, epistemic, and social) have been studied over a two-semester general chemistry course implementing ADI to qualitatively investigate what elements were challenging for students. Students struggled the most with changing a claim when noticing inconsistent information (cognitive) as well as using theories or models to make sense of phenomena (epistemic).[22] While there have been several studies highlighting the impact of ADI on various outcomes, there is still a need for studies that explore these outcomes, such as the change in the three aspects of argumentation across the general chemistry sequence, within different contexts.

■ ASAC OBSERVATION PROTOCOL

The Assessment of Scientific Argumentation in the Classroom (ASAC) observation protocol was designed to measure the conceptual and cognitive, epistemic, and social aspects during a session of scientific argumentation.[15] Specific criteria used to evaluate the three aspects according to the ASAC observation protocol can be seen in Figure 2. Change in Conceptual and Cognitive Aspects of Argumentation consists of six items that describe the framework and theory of argumentation. These items include focusing on advancing understanding or solving a problem, discussing and evaluating alternative claims, being skeptical of ideas, providing reasoning when supporting or challenging ideas, and modifying claims or explanations when necessary. Change in the Epistemic Aspects of Argumentation is intended to evaluate how consistent the process is with the culture of science. There are six items that target how the group determines what counts as acceptable or valid. Items include using evidence to support or challenge ideas,

https://doi.org/10.1021/acs.jchemed.0c01225
J. Chem. Educ. XXXX, XXX, XXX−XXX

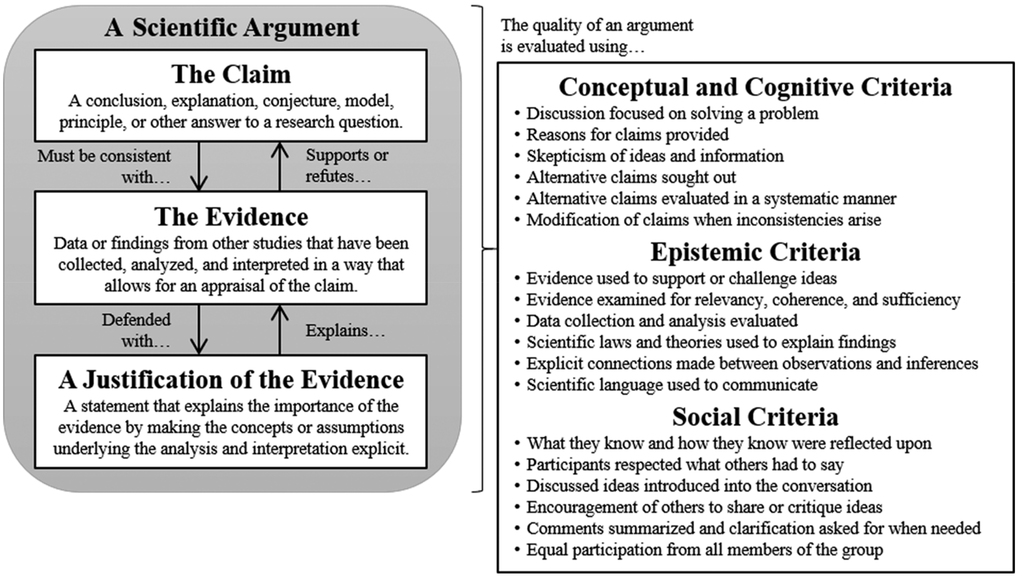

Figure 2. Key components of a scientific argument along with criteria for evaluating the aspects of argumentation. Adapted with permission from ref 7 (p. 564, Figure 1). Copyright 2013 Wiley Periodicals Inc.

examining the relevance, coherence, and sufficiency of the evidence, evaluating how data were gathered and analyzed, using scientific laws and theories to explain the findings of the investigation, connecting to observations and inferences, and using the language of science to communicate. Change in the Social Aspects of Argumentation is centered around group dynamics and how participants interact with each other. The six items in this section address being reflective about their discussions, respecting their peers, discussing ideas when introduced into the conversation, encouraging peers to share or critique their ideas, asking for elaboration or clarification on comments, and observing whether equal participation in the argumentation session occurred. The ASAC observation protocol has been shown to distinguish between groups of students with different argumentation skills ranging from high school students to graduate students.[15] Undergraduate students included in the ASAC development were measured at the beginning and end of a semester of a chemistry laboratory that implemented ADI, showing that the ASAC is a useful tool for the evaluation of argumentation through ADI.

RATIONALE AND RESEARCH QUESTIONS

It is important that students experience multiple opportunities in multiple contexts to engage in science learning that brings together science content and practices to enhance students' abilities to transfer what they learn to new situations.[23,24] This study monitored the development of student proficiency with scientific argumentation, through the three aspects of argumentation, as their exposure to argumentation-based laboratories increased over the course of one academic year. Following the progression of each of the argumentation aspects, this study was able to determine if repeated exposure to argumentation-based laboratories improved the students' overall proficiency with scientific argumentation. In addition, it is important to note that this study did not focus on the individual but rather looked at groups of students and their capacity to exhibit the norms of scientific discourse through repeated opportunities to engage in scientific argumentation. This research study was guided by the following research questions.

How does proficiency with oral argumentation change with repeated opportunities for argumentation within General Chemistry I and II in terms of

(1) overall argumentation?
(2) the cognitive, epistemic, and social aspects of argumentation?

METHODS

Setting and Participants

At East Carolina University, General Chemistry I and II lecture courses were accompanied by the corresponding 1-credit laboratory course, which was a co-requisite of the lecture course. All laboratory sections of General Chemistry I and II at the university followed the ADI instructional method. The investigations that involved argumentation sessions and took place across General Chemistry I and II are listed in Table 1. The ADI curriculum for General Chemistry I laboratory consisted of five argument-generating experiments, along with a molecular model activity that did not follow the four week ADI flow. The General Chemistry II laboratory curriculum consisted of three argument-generating experiments and one real-world application experiment that did not follow the four week ADI flow. Thus, argumentation sessions occurred five times during the General Chemistry I laboratory course and three times during General Chemistry II.

While there were eight total argumentation sessions throughout General Chemistry I and II laboratories, only five were considered for analysis: GC1.1, GC1.3, GC1.5, GC2.1, and GC2.2. The acid−base titration within GC1.4 was designed to find an answer to a guiding question through acid−base titration and compare the answer to a standardized value. The design of this investigation did not warrant much argumentation, and this was apparent through observing the argumentation sessions. For this reason, GC1.4 was omitted from further analysis. With the omission of GC1.4, the authors additionally omitted GC1.2 in order to study differences in argumentation between the beginning, middle, and end of the semester. Investigation GC2.3 had issues that were not discovered until the lab was performed. The experiment used

https://doi.org/10.1021/acs.jchemeduc.0c01225
J. Chem. Educ. XXXX, XXX, XXX−XXX

Table 1. ADI Investigation Descriptions within General Chemistry I and II[a]

Investigation	Concept	Guiding Question	Description
GC1.1	Physical properties	Are these objects made of the same material?	Students are given three objects of varying shapes and sizes. At least one of the objects does not fit in a graduated cylinder for water displacement which requires use of a spill can.
GC1.2	Chemical reactions	What are the products of a chemical reaction (or what iron compound is formed)?	Students are asked to decide what iron compound is produced, $FeSO_4$ or $Fe_2(SO_4)_3$, through the reaction of aqueous $Cu(SO)_4$ and solid Fe.
GC1.3	Solutions and molarity	How could you make 1-L of Purple Pirate dye solution?	Students have to develop a procedure to make a purple solution that matches the color of a given purple solution using Blue 1 and Red 3 dyes and the relation between concentration and absorbance.
GC1.4	Acid−base titrations and neutralization reactions	What is the concentration of acetic acid in commercial vinegar?	Students perform an acid−base titration to determine the concentration of acetic acid in a commercial vinegar solution and compare their value to the commercially reported concentration.
GC1.5	Thermochemistry and calorimetry	Which salt and in what quantity should be used to make an effective but economical cold pack?	Students are tasked to develop an inexpensive cold pack that can reach 2 °C when a salt and 60 mL of water are combined. Students are given four salts, NH_4Cl, NH_4NO_3, $Na_2S_2O_3$, and $MgSO_4$, and corresponding prices to conduct a cost/benefit analysis for each salt.
GC2.1	Intermolecular forces	Why do liquids evaporate at different rates?	Using a temperature probe, students determine the rate of evaporation for seven liquids: pentane, acetone, methanol, ethanol, propanol, butanol, and water. The strength of intermolecular forces is then inferred.
GC2.2	Kinetics	How fast does crystal violet decolorize?	Students collect absorbance data of various concentrations of crystal violet with excess sodium hydroxide to determine the rate law for the reaction.
GC2.3	Equilibrium	What is the formation constant (K_f) for the $FeSCN^{2+}$ system?	Students collect absorbance data from standard and equilibrium solutions of $Fe(SCN)^{2+}$ to determine the formation constant for the purpose of developing a fake blood-like solution for a theater company.

[a]Investigations included in this study are in bold.

within this investigation ended up taking much more time than expected, and consequently, students did not have time to participate in a meaningful argumentation session. For this reason, GC2.3 was omitted from further analysis.

Participants eligible to participate in this research study were students enrolled in the General Chemistry I and II laboratories at the university during the 2017−2018 academic year. Two sets of students were used for data collection. The first set of participants included students from a single section of General Chemistry I. The second set of participants included students from a single section of General Chemistry II. The instructors for the two sections had the same previous experience with facilitating ADI in the lab. The students enrolled in the selected sections that agreed to participate in the study did so through an informed consent that was approved by the university's Internal Review Board.

All laboratory sections had a capacity of 48 students, creating a possibility of 12 groups of 4 students per section. The General Chemistry I section that was selected for the study was at capacity, totaling 48 students. Over the course of the semester, 4 students dropped the laboratory course, leaving only 11 groups to record for the last three of five experiments. In the General Chemistry II section that was selected, the class enrollment began at 30 and therefore 8 groups were formed. There was one student who did not agree to participate in the study as well as two students who dropped the course during the semester. These situations did not affect the number of argumentation groups. All students who were enrolled in the General Chemistry II section had previously participated in ADI during General Chemistry I and no students were repeating the laboratory course. Within the final rosters for General Chemistry I and II, there were only two students who were enrolled in both sections within this study.

Each laboratory section in this study was taught by a lead instructor with two semesters of experience teaching ADI and an undergraduate teaching assistant (UTA). The UTAs are required to take a one-semester course on laboratory instruction which is coupled with a semester shadowing an experienced instructor. All new instructors attend a workshop that orients them to ADI but more importantly to student-centered instruction. Examples of student- vs instructor-centered instruction are provided and demonstrated for each stage of the ADI instructional model. For the argumentation session, instructors are trained to remain on the peripheral of the discussion, asking guiding questions to direct the conversation if students are overlooking a significant issue. Instructors are discouraged from participating in the argumentation, as this has been found to derail the argumentation process. At the close of the argumentation session, instructors are encouraged to lead a "wrap-up" discussion where they may make suggestions for the written arguments in the laboratory reports. The description of ADI, development of the whiteboard, and prompts for the argumentation session provided to the students in the Laboratory Manual are included in the Supporting Information.

Data Collection

Students participating in the study were video recorded during argumentation sessions. A video recorder was attached to each whiteboard, facing the audience. Within each group, one group member was selected to remain with the whiteboard to present the group's argument. The remaining group members became

https://doi.org/10.1021/acs.jchemed.0c01225
J. Chem. Educ. XXXX, XXX, XXX−XXX

the travelers and rotated to other whiteboards to engage in argumentation sessions. Each video recorder typically recorded 2−3 group rotations at a single whiteboard, followed by a postargumentation discussion with the host group. This equated to 12−16 students participating in each video. The argumentation rotations and postargumentation discussion captured on a single video recorder were considered a single argumentation session. Figure 3 illustrates the movement of students for a whole class participating in argumentation sessions.

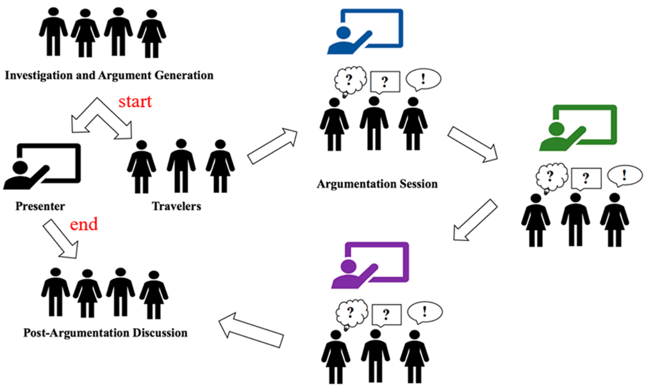

Figure 3. Argumentation session illustration. Adapted with permission from ref 22 (p 438, Figure 2). Copyright 2019 American Chemical Society and Division of Chemical Education, Inc. At the start of the argumentation session, student groups break into a presenter and travelers. Travelers then participate in argumentation with two to three presenters from different student groups. At the end of the argumentation session, travelers reunite with their own group presenter and participate in a postargumentation discussion.

Argument Analysis

In an effort to capture argumentation, researchers video recorded the argumentation sessions, transcribed the videos, and open coded the corresponding transcripts using the ASAC observation protocol. The complete ASAC protocol is provided in the Supporting Information. For each of the 18 items on the observation protocol, a score of 0, 1, 2, or 3 was given. This score was determined from the frequencies of the item as coded in the transcript in NVivo (Version 10). A score of 0 or *Not at All* indicated the item was not observed or coded in the transcript, a score of 1 or *Rarely* indicated that the item was observed and coded once, a score of 2 or *Sometimes* indicated that the item was observed and coded 2−3 times in the transcript, and a score of 3 or *Often* indicated that the item was observed and coded four or more times in the transcript. Each of the three sections, with six items a piece, could have a maximum score of 18, for a maximum score of 54 for each argumentation transcript. Inter-rater reliability was established between author M.A.L., author J.P.W., and an additional undergraduate researcher trained in chemistry education research. IRR was considered established with a Cohen's Kappa above 0.70.[25]

ANOVA and pairwise comparisons were used to investigate the significance of mean differences of ASAC scores between investigations. Student argumentation group composition within each respective course stayed the same throughout the semester and changed between semesters due to student course enrollment. As a result, different types of ANOVA and pairwise statistics were used depending on what comparisons were being made. Details of these decisions and assumptions made per statistic are included in the Supporting Information.

■ RESULTS AND DISCUSSION

Proficiency with Oral Argumentation: How Does It Change with Repeated Opportunities for Argumentation within General Chemistry I and II in Terms of Overall Argumentation?

Table 2 depicts the calculated mean score for ASAC total scores for each of the argumentation sessions included within this study. The overall quality of argumentation increased between GC1.1 and GC2.2 as evidenced by the significant increase of total ASAC scores over the course of the academic year with a large effect size of $d = 1.07$ (Figure 4). This

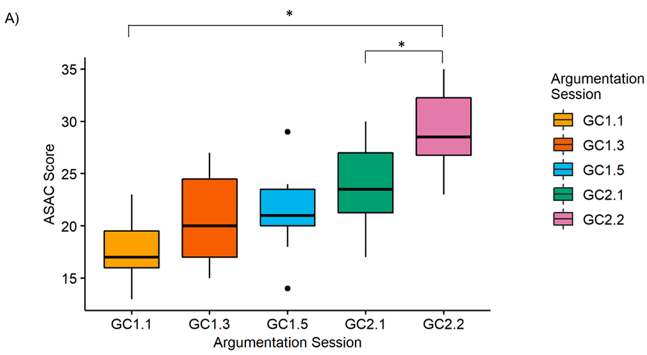

Figure 4. Boxplots of ASAC total scores across the five argumentation sessions. Asterisks (*) mark significant differences between scores denoted with brackets.

evidence suggested that argumentation quality could be improved with multiple exposures to argumentation sessions within ADI. While argumentation quality did improve, the average ASAC total score on GC2.2 was 29 out of 54 possible points during the last recorded argumentation session, showing that there is still room for argumentation improvement. This was to be expected, as the ASAC was designed to distinguish

Table 2. Descriptive Statistics of ASAC Scores for Each Argumentation Session

			Mean ± Standard Deviation			
Argumentation Session	Course[a]	*N*	Total	Cognitive	Epistemic	Social
GC1.1	GC I	11	17.8 ± 2.86	3.27 ± 1.68	7.36 ± 2.16	7.18 ± 1.47
GC1.3	GC I	11	20.8 ± 4.12	6.09 ± 2.55	6.27 ± 1.27	8.46 ± 1.86
GC1.5	GC I	11	20.5 ± 3.80	5.64 ± 1.63	7.91 ± 2.30	8.36 ± 2.84
GC2.1	GC II	8	23.9 ± 4.70	7.00 ± 2.27	8.75 ± 1.39	8.12 ± 2.30
GC2.2	GC II	8	29.1 ± 3.98	9.12 ± 2.95	10.0 ± 2.00	9.62 ± 2.07

[a]GC = General Chemistry.

https://doi.org/10.1021/acs.jchemed.0c01225
J. Chem. Educ. XXXX, XXX, XXX−XXX

argumentation levels ranging from high school students to graduate students. For detailed descriptions of all comparisons tested, see the Supporting Information.

In the following sections we address research question two, "How does proficiency with oral argumentation change with repeated opportunities for argumentation within General Chemistry I and II in terms of the cognitive, epistemic, and social aspects of argumentation?", by discussing the change in subcategories within the ASAC protocol over time between argumentation sessions GC1.1 and GC2.2. Before discussing the change in argumentation between experiments, it is necessary to provide a more in-depth description of the GC1.1 and GC2.2 investigations.

Detailed Description of GC1.1 and GC2.2

The density investigation, GC1.1, has been previously described in detail,[26] but for this work it is important to consider the specific content. Density is a physical property that students have observed previously without knowing the role of density as a mass to volume conversion factor in chemistry. The mathematics necessary in finding density is a simple ratio, mass divided by volume. The ADI investigation introduces some uncertainty into a relatively simple investigation by using objects that do not fit in the supplied graduated cylinder. An overflow can (spill can) is provided, which is a fairly unreliable method of determining displacement by volume. Previous research found that this investigation resulted in argumentation focused on methodology.[7,22]

The kinetics investigation, GC2.2, uses colorimetry to determine a rate law. Figure 5 diagrams the process students

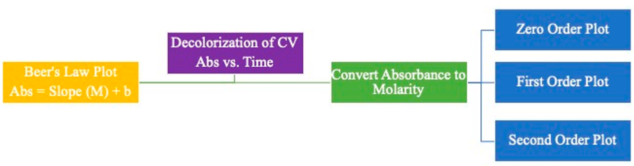

Figure 5. Kinetics investigation flowchart.

follow over 2 weeks. First, Beer's Law and serial dilution are used to determine the relationship between absorbance and molarity for crystal violet. The decolorization reaction ($CV^+ + OH^- \rightarrow CV$) is then conducted in the colorimeter while collecting absorbance and time data. The absorbance data are subsequently converted to molarity, and integrated rate law plots are used to determine the order with respect to crystal violet. Students are given that the reaction is first order with respect to hydroxide. The final piece of this complex problem requires understanding the flooding technique, i.e., excess hydroxide ion, which forces the rate to rely on the crystal violet but also means that the slope of the integrated rate law is a pseudo-k value. Depending on the number of absorbance vs time data points students used, e.g., waiting too long to start data collection or stopping data collection too soon, they will either claim zero or first order with respect to crystal violet. The correct order with respect to crystal violet is one. This kinetics investigation is substantially more complex than the density investigation.

Proficiency with Oral Argumentation: How Does It Change with Repeated Opportunities for Argumentation within General Chemistry I and II in Terms of the Cognitive, Epistemic, and Social Aspects of Argumentation?

Table 2 depicts the calculated mean scores for ASAC subcategories for each of the argumentation sessions included within this study. In addition to significant differences in total ASAC scores between GC1.1 and GC2.2, there were significant differences between all three subcategories of the ASAC (Figure 6). While there was a significant change in

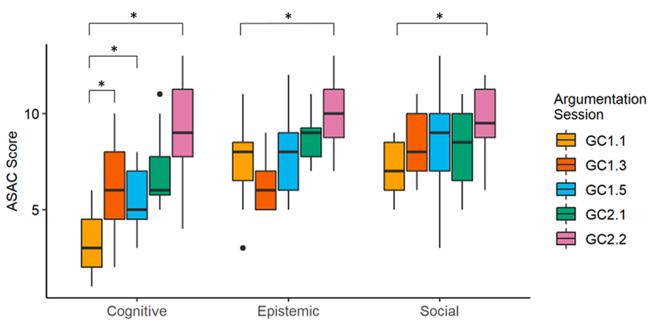

Figure 6. Boxplots subcategory scores across the five argumentation sessions. Asterisks (*) mark significant differences between scores denoted with brackets.

argument quality between the second-semester investigations, GC2.1 and GC2.2, as indicated by the significant difference in the total ASAC scores (Figure 4), there were no significant changes within any of the ASAC subcategories (Figure 6) between these two investigations. For detailed descriptions of all comparisons tested, see the Supporting Information.

Change in Conceptual and Cognitive Aspects of Argumentation

The cognitive aspects of argumentation improved between argumentation sessions GC1.1 to GC1.3 (large effect size, d = 0.87), GC1.5 (large effect size, d = 1.05), and GC2.2 (large effect size, d = 2.56) (Figure 6). Changes between individual cognitive codes across GC1.1 and GC2.2 can be seen in Figure 7. Changes between individual cognitive codes for the remaining investigations are included in the Supporting Information. Cognitive code descriptions are available in

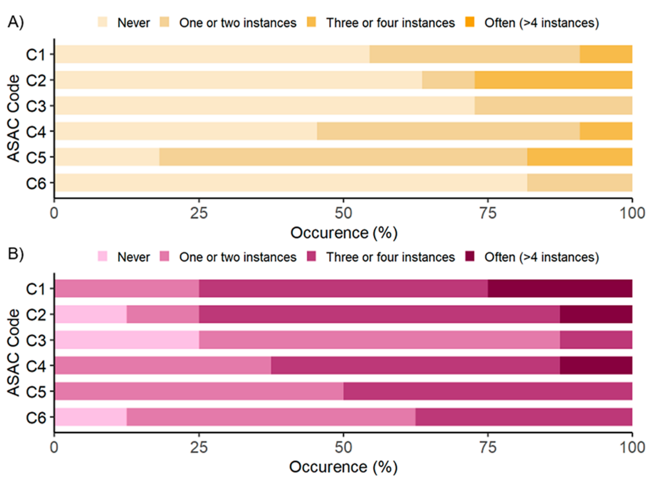

Figure 7. Individual cognitive ASAC code distributions for argumentation sessions (A) GC1.1 and (B) GC2.2.

https://doi.org/10.1021/acs.jchemed.0c01225
J. Chem. Educ. XXXX, XXX, XXX−XXX

Table 3. The examples of poor argumentation tied to codes of the cognitive subscores can be seen in the following examples

Table 3. Conceptual and Cognitive ASAC Codes and Their Definitions

Code	Definition
C1	The talk of the group was focused on solving a problem or advancing understanding.
C2	The participants sought out and discussed alternative claims or explanations.
C3	The participants modified their explanation or claim when they noticed an inconsistency or discovered anomalous information.
C4	The participants were skeptical of ideas and information.
C5	The participants provided reasons when supporting or challenging an idea.
C6	The participants attempted to evaluate the merits of each alternative claim or explanation in a systematic manner.

from a single argumentation session within GC1.1. Students within these excerpts are denoted by numbers indicating what order they appeared on a single video recording. Numbers were kept as identifiers during a string of excerpts because this analysis does not focus on individual students, but rather the group as a whole.

In this excerpt, the presenter first introduced their claim that two out of the three objects consisted of the same material on the basis of evidence from their two indirect methods of finding density, water displacement and use of a ruler and scale. The displacement method produced densities of 1.36, 1.37, and 1.41 g/cm^3, while the ruler method provided densities of 0.799, 0.444, and 1.28 g/cm^3. After discussing the two different methods that were used to find the density of the objects, the two groups realized that they had conflicting values for both methods and both groups used different equations for calculating density on the basis of the ruler method.

- **Speaker 3**: Yeah. Our numbers are definitely different, but we used a different equation for using the ruler.
- **Speaker 1**: What'd do?
- **Speaker 3**: Pi r squared time height.
- **Speaker 1**: That sounds confusing....
- **Speaker 2**: So why do we have the same number can, but our pieces are the different sizes?
- **Speaker 3**: Same material, different size should still have the same density if they're the same.
- **Speaker 4**: So did you guys get similar densities for your object?
- **Together**: No.
- **Speaker 3**: I mean, they were not that far off. Because we were—I think we got 1.2, 1.4, 1.6 when we used the water displacement. But when we went with the ruler method, it was still in the one-point range. And that is because we used a different equation. I think we got 1.7, 1.7, and 1.4.

The groups provided one reason as to why the objects they had, which were suspected to be the same material, were different sizes; even if they were different sizes, if they had the same densities, they should be the same material. This statement provided an example of code C5; the participants provided reasons when supporting or challenging an idea. Even though the density values between the two groups were in disagreement, neither group investigated these differences beyond surface-level mistakes in calculations for the ruler method. Neither group tried to reconcile which calculation was correct before the first traveling group moved on.

After the second group arrived, the conversation quickly derailed after the presenter discussed their argument, without changes, and asked the traveling group about their results.

- **Speaker 1**: What did you guys [get]?
- **Speaker 4**: We had a rectangle, a cylinder, and···
- **Speaker 9**: Did you have biology last year?
- **Speaker 1**: Mm-mm.
- **Speaker 9:** Have you ever had a lab before.
- **Speaker 1**: No.
- **Speaker 9:** Usually, we sit down. I do not want to stand up for 3 h.

The lack of focus on solving the task at hand resulted in a score of zero for this group on code C1, the talk of the group was focused on solving a problem or advancing understanding. The groups eventually got back on track but only to discuss that other groups had the same claim: that two objects of the three were made of the same material.

- **Speaker 9**: I wrote something like it is basically the same thing. They're saying the same thing. They have similar densities, so they're made of the same material.
- **Speaker 4:** Yeah, but we do not know what others...
- **Speaker 9**: No, but I'm saying that is what every group is saying.
- **Speaker 4**: There's two [other groups] now.
- **Speaker 9:** Even that [group] said the same thing.

The groups were not providing evidence to justify confidence in their original claim, and this decision seemed to be based on others' answers despite the conflicting information they had received from the first traveling group and the uncertainty in the correct calculation used in the ruler method. This level of argumentation resulted in a score of zero for code C3; the participants modified their explanation or claim when they noticed an inconsistency or discovered anomalous data, because the groups were explaining away their discrepancies. This mirrors a previous finding where students were hesitant to modify their claims after finding inconsistencies between their claim and their evidence.[22] This excerpt also provides an example of the level of argumentation that resulted in a score of zero for C6; the participants attempted to evaluate the merits of each alternative claim or explanation in a systematic manner, as there was no systematic breakdown of the reasoning behind their argument. The way that Speaker 9 shut down the conversation by citing other group results but not specifying why the other groups may be correct is an example of taking an escape hatch,[27] where students do not push the conversation further when tension in the group arises. This may be due to cognitive load issues or to social discomfort.

While the presenter's group did have the correct claim, the students within this argumentation session did not participate in quality conceptual and cognitive aspects of argumentation because the argument never developed into negotiating meaning of the conflicting evidence beyond incorrect calculations, such as reasons as to why the first traveling group had conflicting densities for their objects on the basis of the displacement method. Missing within the argument was also a sense of skepticism about ideas and information (code C4). This group scored a zero for this code. Only 55% of the groups had a score above zero for code C4. An example of skepticism within this specific session during GC1.1 a student

https://doi.org/10.1021/acs.jchemed.0c01225
J. Chem. Educ. XXXX, XXX, XXX−XXX

asking "*Why is that a good idea*?" in response to a classmate making a statement that they measured volume of their objects three times.

All cognitive codes within the ASAC increased between GC1.1 and GC2.2 (Figure 7). Within GC1.1, 55% of groups had a score of zero within code C1 and the talk of the group was focused on solving a problem or advancing understanding. There were no groups within GC2.2 with a score of zero on this code. This suggested that students within GC2.2 were more focused on the task at hand or less likely to use an escape hatch when tension arose[27] within the group when participating in argumentation as compared to GC1.1.

It is important to note that the correct order with respect to CV^+ for this investigation was one. The order of a reaction with respect to a reactant can be determined from analyzing the respective plots of absorbance vs time for zero, first, and second order reactions ([A] vs time, ln[A] vs time, and 1/ln[A] vs time, respectively). The most linear relation supports the "correct" order for the reaction. Within the experimental portion of this investigation, there was potential for variation with student data collection. Some students did not start data collection soon enough or stopped data collection too early. This artifact resulted in a more linear relation (R^2 closest to one for the zero order reaction instead of the first order reaction when these plots were analyzed. Some of the students overheard the teaching assistant and the instructor say that the order with respect to CV^+ was supposed to be one. Students were more likely to seek out and discuss an alternative claim or explanation (code C2) within GC2.2, such as when a student posits that the order of the reaction with respect to crystal violet (CV^+) could be one instead of zero after overhearing a TA say the correct order was one, "*But I'm confused because Jenny said that she got zero, but it was supposed to be first or something like that.*" Speaker 2 then uses justification from their evidence, "*No, I think that it's the entire reaction that supposed to be first. But this specific one,* [crystal] *violet, it's just zero.*" to argue against the alternative claim.

The evidence students had collected conflicted with the correct order being one, but students from different groups still tried to come up with reasons to support the alternative claim (code C5) that the reaction order could be one:

- **Speaker 7**: And really, how would you like get that? You know what I mean? What if you overdo change [in absorbance]? Like if you added more points, I feel you could just make that line straighter. But you never know.
- **Speaker 3**: ...what happened with ours is we got a zero-order rate for the reaction, but she said it was supposed to be first. Why would—What in the data? Do you think if we graphed more points it would have been? Does that make sense?

Students posited that if they had used more data points, it may have changed their results, but they were quick to dismiss this idea. The statements above also provide examples of students being skeptical of ideas and information (code C4); for example, when Speaker 7 asks the question "*How would you get that*?" Even though some students were explicitly told that the reaction order with respect to CV^+ was one, groups were still reluctant to change their claim based on this information due to the lack of evidence within the data they had collected. While instructor interaction is not taken into account within the ASAC observation protocol, it is important to note that instructor interaction may have an undesirable effect on student argumentation. There was only one group within GC2.2 that scored a two on the ASAC code C3; the participants modified their explanation or claim when they noticed an inconsistency or discovered anomalous data. The remaining seven groups scored either a zero or one. The claim changes that did occur involved revising the rate for the decolorization reaction via a calculation error. The revision of claims via calculation errors also reflected the increase in score for code C6; the participants attempted to evaluate the merits of each alternative claim or explanation in a systematic manner. Students within GC2.2 often calculated different rates for the decolorization of CV^+. In order to work through these conflicting claims, students would work together to explain the calculations performed in a systematic way.

Change in the Epistemic Aspect of Argumentation

The change in the epistemic subscore was significant between GC1.1 and GC2.2 with a large effect size of $d = 1.26$ (Figure 6). All epistemic codes on the ASAC protocol increased between argumentation sessions GC1.1 and GC2.2 (Figure 8). Changes between individual epistemic codes for the remaining investigations are included in the Supporting Information. Definitions of epistemic codes can be seen in Table 4.

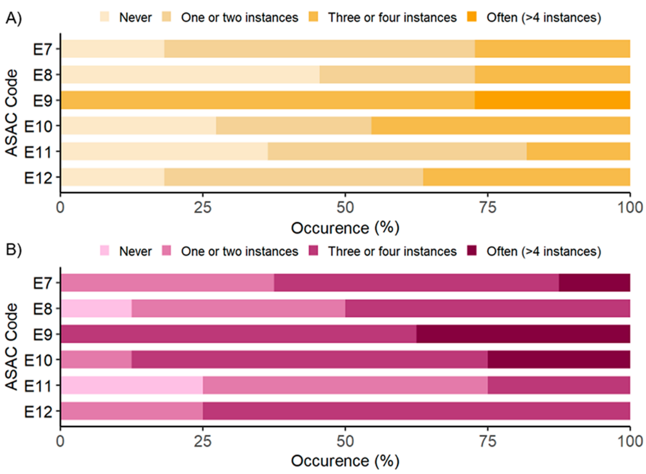

Figure 8. Individual epistemic ASAC code distributions for argumentation sessions (A) GC1.1 and (B) GC2.2.

An explanation for low scores on the epistemic codes within argumentation session GC1.1 can be seen within the following examples.

- **Speaker 2:** We had two cylinders and a cube. Our claim was that they're all made of different materials, once we

Table 4. Epistemic ASAC Codes and Their Definitions

Code	Definition
E7	The participants used evidence to support and challenge ideas or to make sense of the phenomenon under investigation.
E8	The participants examined the relevance, coherence, and sufficiency of the evidence.
E9	The participants evaluated how the data were gathered, analyzed, or interpreted.
E10	The participants used scientific theories, laws, or models to support and challenge ideas or to help make sense of the phenomenon under investigation.
E11	The participants made distinctions and connections between inferences and observations explicit to others.
E12	The participants used the language of science to communicate ideas.

https://doi.org/10.1021/acs.jchemed.0c01225
J. Chem. Educ. XXXX, XXX, XXX−XXX

did our two different methods. Method one, we used a graduated cylinder and spill can [displacement method]. The densities of the cube and the purple cylinder were very similar [1.2 g/mL and 1 g/mL vs 2.1 g/mL]. But then once we did a completely different method [ruler method], they were two completely different densities [1.7 g/cm^3 and 1.4 g/cm^3 vs 2.5 g/cm^3]. So with that, we thought that the ruler and formulas would be more accurate than the graduated cylinder and spill can. So we went with that they're all different materials because of their densities.

After this introduction by the speaker, the students did not critique the evidence presented, even though the presenting group had conflicting results from the two methods. The investigation within GC1.1 was designed with the intention of students generating unreliable data using the spill can method in order to facilitate argumentation.[26] In this excerpt, while students discussed frustration with using the spill can for the displacement method, they did not evaluate the accuracy of either method. The traveling students also claimed the three objects they had were different materials but did not compare their evidence with the presenter. This lack of discussion explains the score of zero for the codes E7, the participants used evidence to support and challenge ideas or to make sense of the phenomenon under investigation, and E8, the participants examined the relevance, coherence, and sufficiency of the evidence. The students provided conflicting evidence and never explained why one method was favored over another.

Later in the argumentation session, Speaker 6 posed a question about the surface features of the objects.

- **Speaker 6**: When you felt the objects, you picked them all up, did they seem like they were the same...?
- **Speaker 2**: Yes and no because they're all really light and you could not really tell a huge difference. It is not that big of a difference.
- **Speaker 6**: Yeah. We had one that was a huge difference. And then the other two were almost the same mass too. They're different shapes.
- **Speaker 2**: Well, the cube has more mass than the other two. The other two were really close [in weight]. But then we do the density, they're completely different.

The students were making sense of the difference between mass and density. While two objects felt similar in weight, the densities were different, and therefore the students concluded that the materials were different. This shows the only example, for this group, of the code E11, the participants made distinctions and connections between inferences and observations explicit to others.

The group within in the following example, scored a zero on the code E10, the participants used scientific theories, laws, or models to support and challenge ideas or to help make sense of the phenomenon under investigation. The focus on surface-level features in argumentation has been documented in past research[22] and occurs again, in the following example.

- **Speaker 2**: So the cube was a different shape, but the cube and the thing had the same density.
- **Speaker 3**: So it was just made out of the same material.
- **Speaker 2**: We told them that they were all different because it was a different size. But thinking about it, it is probably the same material.
- **Speaker 3**: Yeah, it is the same material.
- ...
- **Speaker 3**: So basically, our two cylinders are not made out of the same material of the cubes because it is gold and our cylinders are yellow.

This group reasoned that two of the three objects had the same density, and therefore were the same material, but they never explained why density would explain this. In addition, after deciding that density was evidence for their claim, speaker 3 reverts back to a surface-level feature, color, as the explanation for the cylinders to be the same material instead.

Students did not use scientific language (code E12) as frequently throughout argumentation session GC1.1, as compared to argumentation session GC2.2. Within GC1.1, students sometimes used scientific terms such as "accuracy" and "physical properties" but did not always use scientific language when speaking, such as when a student says "*For the spill can, we did it twice for each object, just to make sure that our measurements were right*." Instead of using the term "accurate", the student suggested that there was a "right" answer. While scientific language was more prevalent in GC2.2 overall, no group scored a three on E12. This highlights an area that may need to be emphasized throughout both semesters in order to foster argumentation that involves more scientific language.

Although students posed questions about how data were collected and analyzed within argumentation session GC1.1, there was little discussion about the quality of evidence (code E8) as compared to argumentation session GC2.2. Within GC2.2, after presenting the group claim that the rate of the reaction with respect to CV^+ was zero because their R^2 for the zero order plot was the closest to one (example of code E7), Speaker 1 calls into question their evidence because they realized they were supposed to collect more data points.

- **Speaker 1**: ...So you are supposed to take data like every one second, so that you are marking points...and it would be more accurate.

This is an example of code E8 because Speaker 1 is questioning the quality of another group's evidence. Additionally, students within GC2.2 more frequently discussed scientific theories, laws, or models to support ideas or make sense of the phenomenon under investigation (code E10). For example, when a student discussed the relational model within their zero order plot,

- **Speaker 1:** We used the slope of that graph that represented our k-prime value, k-prime being the pseudo k.

when finding the rate law for the decolorization of CV^+.

While the code E9, the participants evaluated how the available data was interpreted or the method used to gather the data, did show a slight increase between GC1.1 and GC2.2, students did not score below a two for either argumentation session. Student lab manuals contained a guide to aid with asking questions during their first experience with an argumentation session. These questions were mostly reflective of the epistemic construct and were likely the reason students scored high on code E9. Example question prompts included "What did your group do to analyze the data, and why did you do it that way? Did you check your calculations?" and "What did your group do to ensure that the data you collected are reliable? What did your group do to decrease measurement error?". While these guides were available during future argumentation sessions, there was no evidence that they

https://doi.org/10.1021/acs.jchemed.0c01225
J. Chem. Educ. XXXX, XXX, XXX−XXX

were continually used. This is evidenced by the questions students used to evaluate data interpretation and methods in GC2.2. Questions students asked during argumentation session GC2.2 became more specific to the respective investigation such as "*You find the rate by multiplying the k value 0.5, right*?".

Change in the Social Aspect of Argumentation

The social subscore of the ASAC was significantly different between GC1.1 and GC2.2 with a large effect size of $d = 1.40$. The codes reflecting the social aspects of argumentation occurred more frequently within argumentation session GC2.2 compared to GC1.1 except for one code, S13: The participants were reflective about what they know and how they know (Figure 9). Changes between individual social codes for the remaining investigations are included in the Supporting Information. Descriptions of the social codes are available in Table 5. Within GC1.1, students would ask questions that prompted reflection of what they know and how they know, such as "*Does that make sense for everybody?*". This type of questioning occurred in GC2.2 but occurred less frequently. The conversations within GC2.2 were less focused on the reflection of group knowledge.

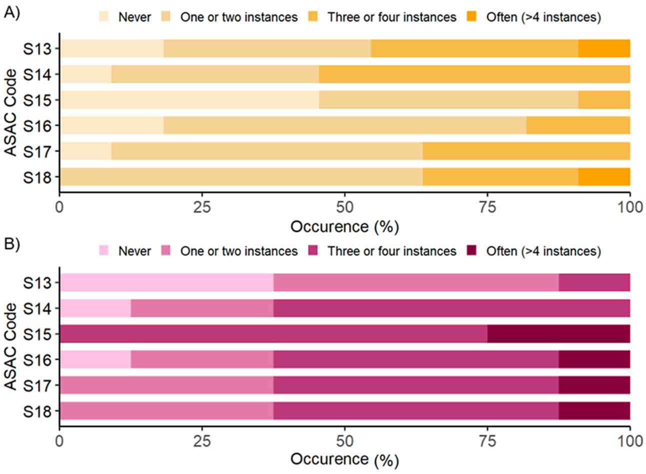

Figure 9. Individual social ASAC code distributions for argumentation sessions (A) GC1.1 and (B) GC2.2.

Table 5. Social ASAC Codes and Their Definitions

Code	Definition
S13	The participants were reflective about what they know and how they know it.
S14	The participants respected what each other had to say.
S15	The participants discussed an idea when it was introduced into the conversation.
S16	The participants encouraged or invited others to share or critique ideas.
S17	The participants restated or summarized comments and asked each other to clarify or elaborate on their comments.
S18	There was equal participation from all members of the group.

While students may have been less explicitly reflective of what they knew and how they knew it, there was an increase in the overall social aspect of argumentation between argumentation sessions GC1.1 and GC2.2. The following is an example of code S15, the participants discussed an idea when it was introduced into the conversation, in GC2.2 where a student introduces the idea that the slope of their graph should be negative because the slope is visually decreasing.

- **Speaker 1**: ...this was our slope. Negative 10. Or negative 1×10^{-7}. I think the slope should be negative because it is decreasing.

After introducing this idea, another student described how the pseudo-*k* equation had a negative sign within it, which canceled out the negative slope, which made the overall value positive. Instead of ignoring the idea that Speaker 1 presented, the other group members addressed the confusion. Addressing student ideas or comments within GC1.1 occurred as well, such as when a student explicitly acknowledged an idea introduced by another student, "*And I think the thing that* [another student] *kind of pointed out that the bigger the objects...*", but these types of statements were less common compared to those of GC2.2.

Questions that were reflective of the code S16, the participants encouraged or invited others to share or critique ideas, such as "*What did you guys get? What order was yours?*" and "*How did you guys find* [your reaction order]?", were prevalent throughout argumentation session GC2.2. These types of questions were present in GC1.1 but occurred less frequently. The code S14, the participants respected what each other had to say, stayed relatively consistent between GC1.1 and GC2.2, only increasing by 2%. Throughout both experiments, there were instances of respect shown between students. For example, in GC2.1 when Speaker 1 stated they were confused.

- **Speaker 1**: I'm confused about our k and k primes.
- ...
- **Speaker 1**: We solved for k but I'm pretty sure k prime was just the rate because k prime is your slope that you get from your equation. So it is that. So that is actually your k prime.
- **Speaker 3**: Yeah, that would be k prime.
- **Speaker 1**: Yeah, and then you divide that by 0.05.
- **Speaker 2**: And then that is how you get that.
- **Speaker 3**: Yep, you are right.
- **Speaker 2**: Thank you. The other group did not teach us.

Speakers 2 and 3 helped Speaker 1 through their confusion as opposed to ridiculing them for something they did not know.

Students within argumentation session GC2.2 restated or summarized comments and asked each other to clarify or elaborate on their comments (code S17) more than students within GC1.1. An example of this from GC2.2 occurred after Speaker 1 presented their claim. Speaker 2 asked, "*So you said that the reaction was a zero order. Do you mean just the CV^+ was a zero order?*" and Speaker 1 clarifies that "*CV^+ was a zero order.*" Group participation increased between GC1.1 and GC2.2 as evidenced by the increase in code S18, there was equal participation from all members of the group. This suggested that more students were participating in argumentation, and therefore ASAC scores may have been more reflective of group argumentation in GC2.2 as opposed to a few individuals who were vocal during each session. While group participation increased overall between the two argumentation sessions, 38% of groups still had a low score of one on code S18 within GC2.2 (compared to 64% of groups in GC1.1). Entire group participation is a portion of the social aspect of argumentation that still needs improvement to ensure that ASAC scores are

https://doi.org/10.1021/acs.jchemed.0c01225
J. Chem. Educ. XXXX, XXX, XXX−XXX

reflective of entire group participation rather than a few vocal students.

LIMTATIONS AND FUTURE WORK

The two sections selected for data collection in General Chemistry I and II consisted of different students but were still representative of the student population enrolled in these laboratory courses. All students in the General Chemistry I laboratory course passed and were able to move to the laboratory course for General Chemistry II. The two selected classes also consisted of two different instructors, but both were experienced in teaching ADI-based laboratories in order to minimize variance among the different classes. The nature of student enrollment in courses at the university precluded all attempts to maintain consistency across two semesters.

We investigated student engagement in argumentation holistically, i.e., an entire classroom. Each video scored the discourse that took place between 12 and 16 students. Sometimes this score was reflective of a vocal few within a group as compared to the entire group, as seen in the various scores within code S18, but the likelihood that an individual student would impact the ASAC score is nominal in these sessions. Future studies will aim to tie these group argumentation scores to individual written arguments in order to clarify the relation between group and individual argumentation development.

Evidence from previous studies suggest instructor facilitation plays a role in student argumentation.[28,29] Peer to peer interactions are intended to be the main source of argumentation generation within ADI argumentation sessions. Consequently, the ASAC observation protocol does not take into account how interactions between the instructor and students affect the trajectory of argumentation. When instructors tell students the theoretically supported answer for the investigation, it can cause dissonance in students and throw off the process of argumentation, as seen in argumentation session GC2.2 when instructors tell students that the reaction order with respect to CV^+ is one. While it was positive that students stuck with the evidence they had collected as opposed to changing their claim on the basis of hearsay, this artifact suggests that the quality of student argumentation may not be high enough to analyze all of the variables to investigate contradictions in their answers at this point. This is an artifact that should be further explored in future studies.

CONCLUSIONS AND IMPLICATIONS

Students often fail to focus equally on the three aspects of scientific argumentation (cognitive, epistemic, and social) when they are given an opportunity to engage in argumentation. This study examined student argumentation within a two-semester general chemistry laboratory sequence at East Carolina University to explore how the three aspects of argumentation change over time with repeated exposure through ADI. Successful engagement in argumentation should include the simultaneous incorporation of three integrated domains: *conceptual* structures and *cognitive processes*, *epistemic* frameworks, and *social* processes and contexts.[16] As supported by the change in ASAC total and subscores, students' argumentation proficiency significantly increased from the beginning of General Chemistry I to the end of General Chemistry II. This increase in quality of argumentation occurred despite an increase of complexity in investigation topics. Previous research has found that complex content does not necessarily generate complex argumentation.[30] This suggests that students are able to manage a more complex investigation topic while increasing their quality in argumentation. The observed increase in each of the three ASAC subcategories related back to the three aspects of argumentation: cognitive, epistemic, and social. During scientific argumentation, students were able to use conceptual structure and cognitive processes, such as scientific theories or models. The increase in the quality of argumentation was accompanied by a greater understanding and use of epistemic frameworks that characterize science. The final dimension, social norms of the science community pertaining to knowledge formation, communication, and argumentation, was again observed to grow and expand with repeated opportunity for engagement in argumentation.

The increase within the cognitive subcategory over time suggested that students were more often trying to negotiate meaning; e.g., students were more focused on the task at hand, provided reasons for their statements, and attempted to evaluate explanations in a systematic manner. Students went from knowing they had an incorrect equation but not searching for the correct equation in GC1.1 to systematically walking through rate constant calculations as a group in GC2.2. Within the epistemic component of argumentation, students argued in a way that was more consistent with the culture of science, e.g., using scientific theories or examining the relevance of data more often. While students improved over time, the final ASAC score for the epistemic subcategory was 10 out of 18. Students did not always further evaluate their evidence and often took the evidence at face value, an observation that has been previously seen in high school students.[31] Within the social component of argumentation, students interacted with others more often; e.g., they invited others to share ideas and discussed ideas when they were introduced, and there was an increase in equal participation from group members. This could be a reflection of students becoming more comfortable with one another but would need further exploration. The significant increase in ASAC subscores suggested that students' capacity to attend to the cognitive, epistemic, and social aspects of argumentation increased with experience participating in this scientific practice. It is important to note that the GC2.2 ASAC score average was 55% of the maximum score possible after seven exposures to argumentation. The concurrent validity that was established for the ASAC ensured that the instrument would distinguish between groups. The participants in this study were novices, and we would expect to see growth but not expert argumentation. The results from this study support argumentation as a practice, i.e., an exercise that needs to be repeatedly performed in order to move students toward mastery. As students move through a science curriculum, we should continue to provide opportunities to engage in argumentation to reinforce and expand student proficiency with this key scientific practice. Overall, this research provided evidence that repeated exposure to argumentation in a laboratory setting improves the quality of this essential scientific practice. Argument-Driven Inquiry is one way to implement repeated exposure to argumentation in the chemistry laboratory.

https://doi.org/10.1021/acs.jchemed.0c01225
J. Chem. Educ. XXXX, XXX, XXX−XXX

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available at https://pubs.acs.org/doi/10.1021/acs.jchemed.0c01225.

Associated content includes the complete ASAC protocol, the description of ADI, development of the whiteboard, and prompts for the argumentation session provided to the students in the Laboratory Manual, individual ASAC code distributions for GC1.3, GC1.5, and GC2.1 and more specific information about ASAC total and subscore comparison statistics (PDF, DOCX)

■ AUTHOR INFORMATION

Corresponding Author

Joi P. Walker − *Department of Chemistry, East Carolina University, Greenville, North Carolina 27858, United States;* orcid.org/0000-0001-7783-4706; Email: walkerjoi15@ecu.edu

Authors

Kathryn N. Hosbein − *Department of Chemistry, East Carolina University, Greenville, North Carolina 27858, United States;* orcid.org/0000-0001-6771-3660

Meghan A. Lower − *Department of Chemistry, East Carolina University, Greenville, North Carolina 27858, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jchemed.0c01225

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was supported by a grant from the National Science Foundation (DUE No. 1725655). We thank the students and instructors involved in the study for allowing video recording during laboratory time. We also thank the STEM Collaborative for Research in Education (STEM CoRE), ECU Department of Chemistry, and Undergraduate Research and Creativity Activity (URCA) Award for financial support of the study.

■ REFERENCES

(1) Hogan, K.; Maglienti, M. Comparing the Epistemological Underpinnings of Students' and Scientists' Reasoning about Conclusions. *J. Res. Sci. Teach.* **2001**, *38* (6), 663−687.

(2) Sandoval, W. A. Conceptual and Epistemic Aspects of Students' Scientific Explanations. *J. Learn. Sci.* **2003**, *12* (1), 5−51.

(3) Driver, R.; Asoko, H.; Leach, J.; Mortimer, E.; Scott, P. Constructing Scientific Knowledge in the Classroom. *Educ. Res.* **1994**, *23* (7), 5−12.

(4) NGSS Lead States. *Next Generation Science Standards: For States, By States*; The National Academies Press: Washington, DC, USA, 2013; DOI: 10.17226/18290.

(5) National Research Council. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*; The National Academies Press: Washington, DC, USA, 2012, DOI: 10.17226/13165.

(6) Ryu, S.; Sandoval, W. A. Improvements to Elementary Children's Epistemic Understanding from Sustained Argumentation. *Sci. Educ.* **2012**, *96* (3), 488−526.

(7) Walker, J. P.; Sampson, V. Learning to Argue and Arguing to Learn: Argument-Driven Inquiry as a Way to Help Undergraduate Chemistry Students Learn How to Construct Arguments and Engage in Argumentation during a Laboratory Course. *J. Res. Sci. Teach.* **2013**, *50* (5), 561−596.

(8) Toulmin, S. E. *The Uses of Argument*; Cambridge University Press: Cambridge, U.K., 1958.

(9) Zohar, A.; Nemet, F. Fostering Students' Knowledge and Argumentation Skills through Dilemmas in Human Genetics. *J. Res. Sci. Teach.* **2002**, *39* (1), 35−62.

(10) Kelly, G. J.; Bazerman, C. How Students Argue Scientific Claims: A Rhetorical-Semantic Analysis. *Applied Linguistics* **2003**, *24* (1), 28−55.

(11) Lawson, A. The Nature and Development of Hypothetico-Predictive Argumentation with Implications for Science Teaching. *International Journal of Science Education* **2003**, *25* (11), 1387−1408.

(12) Duschl, R. Quality Argumentation and Epistemic Criteria. In *Argumentation in science education: Perspectives from classroom-based research*; Erduran, S., Jimenez-Aleixandre, M., Eds.; Springer: Dordrecht, The Netherlands, 2007; pp 159−175, DOI: 10.1007/978-1-4020-6670-2.

(13) Sampson, V.; Clark, D. Assessment of Argument in Science Education: A Critical Review of the Literature. In *ICLS '06: Proceedings of the Seventh International Conference of the Learning Sciences - Making a Difference*; S. A. Barab, K. E. H., Hickey, D. T., Eds.; Erlbaum: Mahwah, NJ, USA, 2006; pp 655−661.

(14) Kelly, G. Discourse in Science Classrooms. In *Handbook of Research on Science Education*; Abell, S. K., Lederman, N., Eds.; Lawrence Erlbaum Associates: Mahwah, NJ, USA, 2007, 443−469.

(15) Sampson, V.; Enderle, P. J.; Walker, J. P. The Development and Validation of the Assessment of Scientific Argumentation in the Classroom (ASAC) Observation Protocol: A Tool for Evaluating How Students Participate in Scientific Argumentation. In *Perspectives in Scientific Argumentation: Theory, Practice and Research*; Kline, M., Ed.; Springer: New York, 2012; pp 235−264, DOI: 10.1007/978-94-007-2470-9_12.

(16) Duschl, R. Science Education in Three-Part Harmony: Balancing Conceptual, Epistemic, and Social Learning Goals. *Review of Research in Education* **2008**, *32*, 268−291.

(17) Walker, J. P.; Sampson, V.; Grooms, J.; Zimmerman, C.; Anderson, B. Argument-Driven Inquiry in Undergraduate Chemistry Labs: The Impact on Students' Conceptual Understanding, Argument Skills, and Attitudes toward Science. *J. Coll. Sci. Teach.* **2012**, *41* (4), 74−81.

(18) Walker, J. P.; Sampson, V.; Southerland, S.; Enderle, P. J. Using Laboratory to Engage All Students in Science Practices. *Chem. Educ. Res. Pract.* **2016**, *17* (17), 1098−1113.

(19) Grooms, J. A Comparison of Argument Quality and Students' Conceptions of Data and Evidence for Undergraduates Experiencing Two Types of Laboratory Instruction. *J. Chem. Educ.* **2020**, *97*, 2057−2064.

(20) Sampson, V.; Grooms, J.; Walker, J. P. Argument-Driven Inquiry as a Way to Help Students Learn How to Participate in Scientific Argumentation and Craft Written Arguments: An Exploratory Study. *Sci. Educ.* **2011**, *95* (2), 217−257.

(21) Walker, J. P.; Sampson, V. Argument-Driven Inquiry: Using the Laboratory To Improve Undergraduates' Science Writing Skills through Meaningful Science Writing, Peer-Review, and Revision. *J. Chem. Educ.* **2013**, *90* (10), 1269−1274.

(22) Walker, J. P.; Van Duzor, A. G.; Lower, M. A. Facilitating Argumentation in the Laboratory: The Challenges of Claim Change and Justification by Theory. *J. Chem. Educ.* **2019**, *96* (3), 435−444.

(23) Grooms, J.; Sampson, V.; Enderle, P. How Concept Familiarity and Experience with Scientific Argumentation Are Related to the Way Groups Participate in an Episode of Argumentation. *J. Res. Sci. Teach.* **2018**, *55* (9), 1264−1286.

(24) Kuhn, D.; Arvidsson, T. S.; Lesperance, R.; Corprew, R. Can Engaging in Science Practices Promote Deep Understanding of Them?: SCIENCE PRACTICES. *Sci. Educ.* **2017**, *101* (2), 232−250.

(25) Hallgren, K. A. Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in quantitative methods for psychology* **2012**, *8* (1), 23−34.

(26) Walker, J. P.; Wolf, S. F. Getting the Argument Started: A Variation on the Density Investigation. *J. Chem. Educ.* **2017**, *94* (5), 632−635.
(27) Sohr, E. R.; Gupta, A.; Elby, A. Taking an Escape Hatch: Managing Tension in Group Discourse. *Sci. Educ.* **2018**, *102* (5), 883−916.
(28) Simon, S.; Erduran, S.; Osborne, J. Learning to Teach Argumentation: Research and Development in the Science Classroom. *International Journal of Science Education* **2006**, *28* (2−3), 235−260.
(29) Stanford, C.; Moon, A.; Towns, M.; Cole, R. Analysis of Instructor Facilitation Strategies and Their Influences on Student Argumentation: A Case Study of a Process Oriented Guided Inquiry Learning Physical Chemistry Classroom. *J. Chem. Educ.* **2016**, *93* (9), 1501−1513.
(30) Berland, L. K.; McNeill, K. L. A Learning Progression for Scientific Argumentation: Understanding Student Work and Designing Supportive Instructional Contexts. *Sci. Educ.* **2010**, *94* (5), 765−793.
(31) Sandoval, W.; Millwood, K. The Quality of Students' Use of Evidence in Written Scientific Explanations. *Cognition and Instruction* **2005**, *23* (1), 23−55.

https://doi.org/10.1021/acs.jchemed.0c01225
J. Chem. Educ. XXXX, XXX, XXX−XXX