What are Middle School Students Talking About During Clicker Questions? Characterizing Small-Group Conversations Mediated by Classroom Response Systems

Lauren A. Barth-Cohen¹ · Michelle K. Smith^{2,3} · Daniel K. Capps⁴ ·
Justin D. Lewin^{2,3} · Jonathan T. Shemwell^{2,5,6} · MacKenzie R. Stetzer^{2,6}

Published online: 11 August 2015

© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract There is a growing interest in using classroom response systems or clickers in science classrooms at both the university and K-12 levels. Typically, when instructors use this technology, students are asked to answer and discuss clicker questions with their peers. The existing literature on using clickers at the K-12 level has largely focused on the efficacy of clicker implementation, with few studies investigating collaboration and discourse among students. To expand on this work, we investigated the question: Does clicker use promote productive peer discussion among middle school science students? Specifically, we collected data from middle school students in a physical science course. Students were asked to answer a clicker question individually, discuss the question with their peers, answer the same question again, and then subsequently answer a new matched-pair question individually. We audio recorded the peer conversations to characterize the nature of the student discourse. To analyze these conversations, we used

a grounded analysis approach and drew on literature about collaborative knowledge co-construction. The analysis of the conversations revealed that middle school students talked about science content and collaboratively discussed ideas. Furthermore, the majority of conversations, both ones that positively and negatively impacted student performance, contained evidence of collaborative knowledge co-construction.

Keywords Classroom response systems · Small-group discussions · Collaboration

Introduction

Clickers are a type of instructional technology that allows instructors to gauge students' real-time performance in the classroom. Typically, instructors pose conceptual questions in multiple-choice format at several points in the class period and students respond to the question or vote on the answers using clickers. Clicker questions are thought to support learning because they break up periods of lecture into smaller chunks, provide students with opportunities to practice solving problems and monitor their understanding during class, and can serve as a formative assessment tool for instructors. In one commonly used mode known as peer instruction, students answer a question individually, discuss the question with their peers, and answer the question a second time (Mazur 1997). The instructor then leads a class discussion eliciting answer choices and asking students to justify their chosen responses without revealing which choices are correct. Finally, the instructor explains the question and often shows a histogram of all student responses, which gives both instructor and student immediate feedback on how well a concept is understood.

✉ Lauren A. Barth-Cohen
l.barthcohen@miami.edu

¹ Department of Teaching and Learning, University of Miami, Coral Gables, USA

² Maine Center for Research in STEM Education (RiSE Center), University of Maine, Orono, USA

³ School of Biology and Ecology, University of Maine, Orono, USA

⁴ Department of Mathematics and Science Education, University of Georgia, Athens, USA

⁵ College of Education and Human Development, University of Maine, Orono, USA

⁶ Department of Physics and Astronomy, University of Maine, Orono, USA

In recent years, clicker use has become common in K-12 and university science courses; however, the majority of the research to date has focused on university-level instruction (e.g., Crouch and Mazur 2001; Crossgrove and Curran 2008; Smith et al. 2009). Although it seems plausible that a technology that promotes peer discussion in one setting may do so in another setting, little is known about clickers at the K-12 level and how they may be used to promote discussion and student learning in precollege science classrooms. To date, the work in the K-12 setting has explored teacher impressions of using a clicker system in the classroom (Stewart and Stewart 2013) and student and teacher perceptions of their effectiveness (Penuel et al. 2007; Kay and Knaack 2009).

Use of Clickers at the University Level

Clicker Effectiveness Studies

Clicker studies in university-level science courses have largely focused on student performance in classes taught with and without clickers or other audience polling devices (e.g., colored cards for different answer choices). While some studies show differences in student achievement when clicker questions are asked (e.g., Crouch and Mazur 2001; Crossgrove and Curran 2008; Fagen et al. 2002; King and Joshi 2008; Lasry et al. 2009), others have not (Addison et al. 2009; Doty et al. 2006). Conflicting results may be due to the difficulty in performing these types of comparison studies, which can be confounded by differences in class size, cognitive level of questions, class composition, and instructor (Crossgrove and Curran 2008). Nonetheless, studies of student attitudes show that students have positive feelings about using clickers (e.g., Addison et al. 2009; Caldwell 2007; Crossgrove and Curran 2008), and assigning participation points to encourage students to answer clicker questions improves class attendance and provides structure for at-risk students (Freeman et al. 2007).

Learning Benefits of Peer Discussion During Clicker Questions

Researchers have also investigated whether undergraduate students improve their performance on clicker questions after discussing questions with peers. When instructors use the peer instruction approach described above, the frequency of correct answers often increases after peer discussion (Mazur 1997; Crouch and Mazur 2001; Knight and Wood 2005; Smith et al. 2009). Two commonly discussed hypotheses could explain this observation: (1) active engagement of students during peer discussion with peers leads to increased conceptual understanding, or (2) students

simply choose the answer most strongly advocated by adjacent peers who seem to know the correct answer. In order to distinguish between these hypotheses, one experimental approach, which was also used in this study, has students first answer a multiple-choice concept question individually, engage in a discussion with peers, answer the same question again, and then answer a second matched-pair question individually (Smith et al. 2009, 2011). Matched-pair clicker questions are question sets that ask about the same concept, but in slightly different contexts. Studies using matched-pair questions have shown strong support for the first hypothesis (Smith et al. 2009; Porter et al. 2011); furthermore, students show the greatest gains in performance if they are able to first engage in a peer discussion and then immediately listen to an instructor explanation (Smith et al. 2011). This result is consistent with previous findings, showing that student engagement in a learning activity, such as answering questions, predisposes them to learn from a subsequent lecture (Schwartz and Bransford 1998).

In addition to monitoring student performance on clicker questions, recent studies have focused on the nature of the peer discussion. For example, James and Willoughby (2011) recorded peer conversations in introductory university-level astronomy courses to determine whether students had “standard conversations,” which were defined as involving at least one student in a peer group discussing a minimum of one multiple-choice answer and the student answers were aligned with the articulated ideas of the group. They chose this definition because it matched the idealized productive conversations imagined by the course instructors. The authors labeled conversations that did not meet this definition as “nonstandard” and found this category described the majority of the conversations (62 %). Nonstandard conversations included: (1) incorrect ideas that were not anticipated by the instructors, (2) statistical feedback that misrepresents student understanding, such as searching for cues in the questions phrasing, and (3) conversation pitfalls. The conversation pitfalls included instances in which one student passively deferred to another student (5.3 % of the conversations) and instances in which the answer appeared self-evident such that the conversation concluded prematurely (8.9 % of the conversations). The results of the study raise a question: Are these “nonstandard” conversations unproductive or is there something redeeming within them? Answering this question requires an in-depth examination of the character of peer discussion, with an emphasis on the collaboration of multiple students as they discuss science content.

In another study, Knight et al. (2013) investigated peer discussions among students in an upper-division university-level biology course. Specifically, this study focused on turns of talk, coding student conversations for elements

of arguments: claim, reasoning, questioning, and background (Driver et al. 2000; Osborne et al. 2004). The authors of this study also coded the conversations based on the quantity of warrants exchanged on a scale from no reasoning statements to multiple students exchanging reasoning. They showed that the majority of students exchanged reasoning and provided evidence for their ideas. Furthermore, students were more likely to have high-quality discussion of the questions when the instructor provided reasoning-centered cues such as “Discuss your answers with your table, and focus on the reasons for your answers. Then, I’ll ask you to share your reasons.”

Knowledge Co-construction in Peer Discussion

To build on previous work examining peer discussion during clicker questions and extend it into a new educational setting, we analyzed middle school student performance on clicker questions and simultaneously examined the nature of their peer conversations. When examining their conversations, we focused on the collaborative way scientific knowledge can be supported using this technology, building on calls for more rigorous investigations of pedagogical approaches to implementing clicker questions in K-12 classrooms (Beatty and Gerace 2009).

Specifically, we analyzed the peer discussion using a Vygotskian perspective in which assumed peer interactions would contribute to the co-construction of knowledge. During peer interactions, students may build on their own understanding or on the idea of a peer, and there may be some negotiation and/or co-construction of ideas (Hatano 1993). Through this process, students may clarify their thinking and respond to, add to, and connect with each other’s ideas in order to collaboratively build knowledge that may not have been available before they began working together (Hogan et al. 1999). The result of this process may be a change in students’ thinking. Our use of knowledge co-construction builds on previous work where this perspective has been applied to classroom discourse and open-ended interview settings (e.g., Berland and Lee 2012; Engle and Conant 2002; Roschelle 1992).

We focused on applying knowledge co-construction to peer discussions during clicker questions because the conversations are short, often on the order of a few minutes, and are sparked by conceptually challenging multiple-choice questions (Mazur 1997). The questions are typically constructed with appealing plausible alternatives that align with common conceptual difficulties or misconceptions. Furthermore, students know that at the conclusion of the conversation, they will need to choose the answer they think is correct, and there is a reasonable expectation that they will receive immediate feedback from the instructor. Built into the task is the assumption that peer discussion

facilitates a change in thinking. Indeed, after the conversations, students often have an opportunity to re-answer the prior question thereby explicitly needing to make a decision about choosing the same or a different answer. Finally, these conversations are typically characterized by student discourse for an allocated period of time that is unstructured and unmediated by teacher input. As a result of these characteristics of peer discussion, students are expected to be collaborating and co-constructing ideas, suggesting that this framework is a reasonable for this setting.

Given the constraints associated with peer conversations, we aimed to characterize the kind of collaborative knowledge co-construction that occurred in this middle school environment. In this setting, we anticipated that the conversations would pertain to the content of the multiple-choice questions posed and that the conversations would contain contributions by multiple individuals. Specifically, we anticipated discourse moves on the part of the students that contributed to the knowledge co-construction. Furthermore, we expected these discourse moves to contribute to changes in thinking about the relevant science content.

In addition, we investigated two concerns about using clickers in the K-12 classroom. First, groups of students might be off task during peer discussion, which is problematic because it implies an unproductive use of classroom time (Lee et al. 2012). Second, groups of students discussing clicker questions might have one individual who tells peers the answer without engaging them in a productive discussion. For instructors looking to implement peer conversations within the constraints of K-12 classrooms, further research is needed on how often these pragmatic concerns are realized.

Research Questions

Broadly, we investigated whether middle school science students have productive peer discussions during clicker use. Our particular research questions were: (1) Does the percentage of correct responses on clicker questions improve after students discussed questions with their peers? (2) What are the characteristics of knowledge co-construction in student peer conversations during clicker questions, and do certain characteristics relate to student performance? (3) Are some of the previously mentioned concerns about clicker use, namely off-topic conversations or one student telling the others the answer without engaging in a conversation, common in a middle school setting?

The first research question is addressed in Analysis I and is important as a baseline for how these middle school students performed and allows for a comparison with the existing results in the literature. The second research

question is addressed in Analysis II, in which we examine peer conversations in terms of contributions to the science content from multiple individuals, acknowledgment of ideas, asking questions, and revision of ideas, and connect these features to student performance. The third research question is addressed in Analysis III, in which we examine whether the previously mentioned concerns about peer conversations materialized. Addressing these research questions is important because of the potential to inform and refine the pedagogical approaches employed when implementing clicker questions in K-12 classrooms. In addition, this work will help move the field beyond questions about clicker efficacy and toward questions about how to encourage productive peer discussion among students at a variety of educational levels.

Methods

Data Collection: Participants

Data were collected from three teachers' classrooms, representing eight sections of middle school physical science. The school was located in rural, northern New England. Forty-eight percent of students in this school district were eligible for free and reduced lunch, higher than the state average. Two of the teachers transferred to other schools after the first year of the study; therefore, we collected two academic years worth of data from one teacher and only one year from the other two teachers. All three teachers were using the Project Based Inquiry Science Curriculum (Kolodner et al. 2010) for physical sciences. The data collection occurred after students had been formally taught the material, and thus, the intent was for the questions to cover material that was familiar to the students.

One-hundred and thirty-four students (70 male, 64 female) participated in the study. They answered between two and four sets of multiple-choice clicker questions individually for a total of 250 data points. During peer discussion opportunities, students worked in groups of 2–5, and a total of 72 small-group peer conversations were recorded. All student groups were designated by the teacher. We did not specify how teachers should form the groups, but found they did so largely to minimize disruptions in the classroom. Since we were interested in what the students could learn from one another, we asked teachers to only intervene in student discussions if needed to maintain order.

Study Design

We used both quantitative and qualitative approaches in order to examine student performance on clicker questions

and evidence for co-construction during peer discussion. A matched-pair clicker question design was used (e.g., Smith et al. 2009), and all student peer discussions during clicker questions were audio recorded.

For the matched-pair question study design, we wrote pairs of clicker questions on the concepts of kinetic energy, electrical energy, thermal energy, and forces. Teachers reviewed the questions prior to the study, and we revised the questions based on their comments (see Appendix for questions). The experimental design followed a preexisting methodological approach (e.g., Smith et al. 2009, 2011; Fig. 1). First, students answered one clicker question individually (Q1), and then, they participated in a conversation with peers about that question. Following the discussion, they had an opportunity to answer the same question again (Q1 After Discussion, Q1_{AD}). Finally, they were presented with a second matched-pair question (Q2) that was conceptually similar to the first question. The students answered Q2 individually without consulting their peers. After Q2 had been answered, the teachers discussed Q1/Q1_{AD} and Q2 with the class and showed the students' voting frequency graphs. Neither the answers nor graphs that show the frequency of Q1/Q1_{AD} or Q2 answers were revealed to students until after voting for Q2 was completed.

Questions (Q1/Q1_{AD} and Q2) were randomized using a coin flip to decide which question would be asked initially and which would be asked as a matched-pair question (Q2). Throughout this paper, these questions are referred to by the science content topic: thermal energy, kinetic energy, electrical energy, and forces. There were no incentives given to students for participating in the voting and discussions associated with these questions; however, students were accustomed to discussing questions as part of the Project Based Inquiry Science Curriculum (Kolodner et al. 2010), but were new to the use of clicker questions.

A digital voice recorder was placed at each desk to record the student peer conversation about Q1. The clickers were assigned to each student, so when listening to the audio files, we knew the number of students sitting together, and each student's clicker performance.

For the quantitative analysis, all statistical analyses were performed with SPSS (IBM, Armonk, NY) or Excel (Microsoft, Redmond, WA).

Qualitative Analysis

To analyze peer conversations during clicker questions, we used a grounded approach (e.g., Glaser and Strauss 2009; Shkedi 2005) that focused on evidence for collaborative knowledge co-construction. To conduct this analysis, we transcribed all of the peer discussions, read through each conversation, and generated initial elements of

Fig. 1 Description of the experimental design

collaborative knowledge co-construction in each conversation. The unit of analysis was the group discussion related to each of the four clicker questions. There were 72 conversations, with an average length of time on task of 57 ± 36 s. We then discussed the initial elements with the research team, and through these discussions, consensus emerged about documenting similar elements of co-construction including: contributions to the science content from multiple individuals, acknowledgment of ideas, asking questions, and revision of ideas. Each of these elements of co-construction is described in more detail below.

Contributions to the Science Content from Multiple Individuals

Prior research has acknowledged the benefits of collaborative group discussion in which multiple individuals contribute ideas (Dillenbourg 1999; Stahl 2006). In collaborative group discussions, ideas are shared, explored, modified, improved, and expanded upon by multiple individuals. By emphasizing contributions to the science content from multiple individuals in the analysis, we explicitly distinguish among situations in which one student tells peers an answer (correct or incorrect) and situations in which multiple students contribute to the science content. For this element, we are not evaluating the accuracy of any individual's contribution or when individuals tell their peers incorrect answers. Rather, when coding for this element, we focused on whether or not multiple students made verbal statements about the science content. It was not sufficient for multiple students to say only "yes" or "no" or to mention their choice for the multiple-answer question. Here is an example of a conversation about the forces on a propeller car moving at constant speed in which two students discussed the science content: "*Student 1*: I think its C. *Student 2*: Yeah. *S1*: Because there's friction and like the wind and the air. *S2*: I just always pick C when I don't know the answer. *S1*: Same here. *S2*: And it can't be equal [Answer E]. *S1*: Yeah. *S2*: Because it's moving forward so there is more force forward. *S1*: Yup. *S2*: What are you thinking? *S1*: Yes."

Acknowledgment of Ideas

For knowledge co-construction to be successful, students need to listen to one another and not talk past each other; students need to *acknowledge their own ideas and those of peers*. Students need to listen to each others ideas with the

aim of eventually incorporating these ideas into their own thinking. Less important is evaluating the correctness or incorrectness of other ideas. Practically, in the analysis, we were interested in instances when a student mentioned his or her own ideas or a peer's ideas. Mentioning other ideas could involve evaluating the correctness or incorrectness of a statement or stating some uncertainty or a hole in one's thinking. For instance, in this conversation about the same propeller car question, the second student in this excerpt points out a hole in her thinking while acknowledging the first student's idea: "*Student 1*: I said D because it's going at a constant speed. [Laughter] *Student 2*: Oh, I forgot about that." And, in this next statement, also about the propeller car, a student acknowledges a hole in her own thinking: "*Student 1*: Yeah, I thought it was going to be the overall forward force, but it's not speeding up so ..." In these examples, students acknowledged peers' or their own ideas and this could instigate contributions to the science content from multiple students.

Asking a Question About Science Content

Additionally, for successful knowledge co-construction, students need to ask questions. Asking questions is one of the eight practices highlighted as essential elements in K-12 science curriculums (NGSS Lead States 2013). Questioning helps develop habits of mind and is an important part of scientific literacy and helpful in the growth of scientific knowledge. Asking questions can lead to modifications in a student's understanding and, in the case of peer discussion, can lead to a change in answer choice. Less important is the nature of the question, open or closed, and whether the question is on or off the main science content topic. In the analysis for this element, we focused on questions explicitly about the science content (e.g., "Did you pick number two because of the temperature?"). If a student simply asked about answer choice (e.g., "What did you pick?" "Why did you pick C?"), the question was not counted in this element because our goal was to capture questions that could support contributions to the science content from multiple students.

Revision of Ideas

Finally, for knowledge co-construction in this setting, it was insufficient for there to simply be discourse in which multiple people contribute ideas; there needed to be some

revision of ideas. Revision is important due to the nature of clicker conversations. Students are typically first asked to answer a question individually, and then, after the peer discussion, they are given an opportunity to stick with their initial answers or revise their answers. Therefore, a fundamental instructional goal of peer discussion is for students who may have initially chosen an incorrect answer to revise their answers based on what they learned (Mazur 1997). We expected that there will be both revision in the direction of an incorrect idea and revision in the direction of a correct idea. A recent study, using a similar experimental design, found that of the students who changed their clicker question answer, the majority of changed answers were from the wrong answer to the correct answer (Miller et al. 2015). However, given that we are less interested in correctness or evaluation, we focused on the existence of revision not necessarily the direction to the correct answer. Practically, we focused on verbal statements that implied a change in any direction about either the correct answer or the science content. We included both verbal revision of a multiple-choice answer (e.g., “I used to think it was A, now I think it’s D.”) and verbal revision of an explanation for the science content (e.g., “I used to think only temperature influences amount of thermal energy, but now I think both mass and temperature influence thermal energy.”).

These four elements capture the kind of knowledge co-construction seen in these settings given the inherent constraints. Once the elements were defined, two researchers then independently coded all of the conversations and differences were resolved through discussion. In several cases, the substance of the disagreement was due to ambiguities in the transcript and in those cases three of the researchers listened to the audio together to resolve the disagreement.

After coding all the conversations, we also investigated whether different co-construction elements were present in peer discussion where student groups improved their performance from Q1 to Q1_{AD}. To determine which groups showed improvement, we looked at mean Q1 and Q1_{AD} scores for each student group. Groups were labeled according to whether their mean Q1 to Q1_{AD} scores increased ($Q1 < Q1_{AD}$), decreased ($Q1 > Q1_{AD}$), or remained the same ($Q1 = Q1_{AD}$). If the group’s mean Q1 was equal to 100 %, the group was labeled “ceiling.”

Finally, to address potential concerns about clickers, namely conversations that are off-topic and conversations in which one student told the rest of the group an answer with no discussion, we further examined the conversations that included none of the co-construction elements described above. We systematically read through those conversations and examined whether these conversations included any mention of science content, and if they did, how the

other students in the group responded to a student offering an explanation to the clicker question.

Analysis and Results

The analysis is presented in three parts. Analysis I investigates student performance on the clicker questions. Analysis II investigates which combination of the four elements that characterize knowledge co-construction may contribute to learning gains. Analysis III examines pragmatic concerns about using clickers in a middle school setting such as off-topic conversations or a student stating an answer without engaging peers in a conversation.

Analysis I: Gains from Individual and Pooled Data

Students’ Individual Performance

We examined student performance on Q1 (individual vote), Q1_{AD} (vote after discussion), and the matched-pair question Q2 (individual vote). Using pooled data from all four questions, the mean scores on Q1, Q1_{AD}, and Q2 improve and are significantly different from each other (repeated measures ANOVA, $p < 0.05$, $n = 250$ answers per question; Fig. 2), indicating that students benefitted from the discussion of Q1. These results are similar to other studies at the undergraduate level showing that if students engage in peer discussion, their scores improve from Q1 to Q1_{AD} (e.g., Smith et al. 2009, 2011; Crouch and Mazur 2001) and remain high on Q2, a second matched-pair question (e.g., Smith et al. 2009, 2011). In these previous studies and this study, the question order is randomly assigned and the improvement is in the absence of instructor feedback, so the higher performance on Q2 suggests that students are

Fig. 2 Student performance on individually answered Q1, Q1_{AD} after peer discussion, and individually answered matched-pair question Q2. Error bars represent standard error of the mean (SEM)

transferring the knowledge they gain discussing Q1 to the matched-pair question Q2 (Smith et al. 2009, 2011).

Aggregate Peer Conversations by Gains

To further explore changes in student performance, we also examined the average group performance on the questions. As described in the methods section, we sorted the student groups into four categories: (1) groups with scores that *increased* from Q1 to Q1_{AD} ($n = 26$); (2) groups with scores that *decreased* from Q1 to Q1_{AD} ($n = 10$); (3) groups with scores that *remained the same* from Q1 to Q1_{AD} ($n = 33$), including three groups in which no students had the correct answer on Q1 and Q1_{AD}; and (4) groups with scores that were at the *ceiling*, starting, and ending with 100 % correct on Q1 and Q1_{AD} ($n = 3$). Comparing across groups, we found that the groups with scores that increased and decreased start with similar Q1 scores (MANOVA $p > 0.05$). However, there were significant differences in Q2 performance between the increase and decrease groups (MANOVA $p < 0.05$; Fig. 3).

Analysis II: Qualitative Analysis of Peer Discussions

In Analysis II, we examined the elements of knowledge co-construction during peer discussion and how the presence of these elements related to student performance. We first examined student peer discussion in terms of the four elements of knowledge co-construction: contributions to the science content from multiple individuals, acknowledgment of ideas, asking a question about science content, and revision of ideas. Eighty-two percent of the student peer discussion included at least one of these elements, and the most common elements were multiple individuals

contributing to science content and acknowledgment of ideas (Table 1).

We were also interested in whether groups of students who improve their performance from Q1 to Q1_{AD} included different elements of knowledge co-construction in their conversations (Table 1). For this analysis, we combined groups that decreased and remained the same from Q1 to Q1_{AD} because the peer discussion did not improve the overall performance of the group. When comparing the increase and decrease/remaind the same groups, the largest difference was in revision. Eighty-five percent of the conversations that showed an increase from Q1 to Q1_{AD} included revision, whereas only 26 % of the conversations that decreased or remained the same from Q1 to Q1_{AD} included this element. This difference is potentially important as revision might capture the essence of peer conversations.

Furthermore, when conversations showed an increase from Q1 to Q1_{AD}, those conversations tended to contain more co-construction elements compared to conversations that decreased/remained the same from Q1 to Q1_{AD} (Fig. 4). To further illustrate the differences in the conversations, Fig. 5 shows an example of a conversation in which all four elements were used and the student performance increased from Q1 to Q1_{AD}. Figure 6, on the other hand, shows an example conversation in which the students included only two elements and performance decreased from Q1 to Q1_{AD}.

Comparing these two annotated examples, there were differences as the peer discussion leading to an increase in performance contained four elements, while the peer discussion leading to a decrease in performance contained only two elements. While these two conversations do not represent the entire corpus of conversations leading to either performance increases or decreases, they do serve as

Fig. 3 Performance of peer groups calculated from normalized learning gains from Q1 to Q1_{AD}: increase, decrease, remained the same, and at ceiling. Error bars represent standard error of the mean (SEM)

Table 1 Relative frequency of each element of knowledge co-construction across all conversations

Co-construction elements	Total (<i>n</i> = 72)	Increase (<i>n</i> = 26)	Decrease/remained the same (<i>n</i> = 43)	Ceiling (<i>n</i> = 3)	Definition
Contributing to science content from multiple individuals	81 %	92 %	74 %	66 %	Statements suggesting that multiple students contributed to the science content
Acknowledgment of ideas	78 %	92 %	70 %	66 %	Statements that contained instances when a student mentions his or her own ideas or a peer's ideas
Asking a question about science content	24 %	31 %	21 %	0 %	Statements that included questions explicitly about the science content, not questions about answer choice
Revision	49 %	85 %	26 %	66 %	Statements that implied a change in any direction about either the correct answer or the science content

Fig. 4 Percentage of peer discussions by number of elements for both decrease/same and increase conversations**Fig. 5** Annotated example of a peer conversation containing all four elements. In this example, student scores increased from Q1 to Q1_{AD}

<p>Student 1: I had D. Student 2: I did E. Student 3: I said E. S2: I said E. S1: I feel dumb. S2: Because in picture one you can see that one has a higher mass and they are both the same temperature. And in picture two you can see that they have two different temperatures and picture three you see they have bigger mass and more temperature. S3: And more temperature. S1: I feel dumb. S3: You are not dumb. S2: Did you pick number two because of the temperature? S1: Yeah. S2: Yeah. S3: It still makes sense though. S2: It does. S3: They all had the same one. If you just didn't look. Didn't remember that it was mass and temperature. S2: You were 33 percent correct. S3: Ish. S2: Ish. S3: She was still right. S2: Yeah, it was right, but, you just had to look at the other pictures and understand. S1: I did. I'm just not that smart today. S2: You have to remember it's temperature and mass. S3: It is hard to understand sometimes. S1: I'm just not smart today. S2: Cause you're looking at it and it's like wait wait wait. That has more whatever. It looks white so we're going to say it's milk. [...side conversation about milk...] S2: But yeah it would be all three. S3: All three, cause, you just have to remember that temperature and mass. S2: Is the. S3: The thermal energy. S2: Yeah. S3: So if it gives you one, you can kind of tell. S2: Factors or indicators or which ever. S3: You can, what's the word? S3: Infer. S2: Information. S2: It's inference. S3: Inference.</p>	Revision (S1)
	Contributions to the science content from multiple individuals (S2 & S3)
	Question about science content (S2)
	Revision (S1)
	Acknowledgment of ideas (S2 & S3)
	Acknowledgment of ideas (S2 & S3)

illustrative examples showing how conversations with more elements could be richer.

Analysis III: Concerns About Peer Discussion

Next, we examined whether some of the previously mentioned concerns about peer discussion, namely that students are off-topic or conversations involve one individual who tells peers the correct answers (Lee et al. 2012), materialized in this middle school setting. For this analysis, we investigated the 13 conversations that included none of the co-construction elements.

Fig. 6 Annotated example of a peer conversation containing only two elements. In this example, student scores decreased from Q1 to Q1_{AD}

<p>Student 1: E. Student 2: I put E. S1: Yeah cause it's a, yeah it's a, yeah. S2: <u>Negative positive battery it won't touch all of them.</u> S1: <u>Yeah, cause it's a series circuit.</u> Student 3: <u>Bulb one would stay on.</u> S1: <u>No, it's a parallel circuit.</u> S3: <u>Bulb one would stay on.</u> S1: Nope. Yeah, yeah bulb one that's why I did A. Student 4: I did A. S1: Yeah we did A. S4: No you didn't. S1: Yeah I did. S4: No you didn't. S4: You said 1A. S1: Yeah I said A.</p>	<p>Contributions to the science content from multiple individuals (S2, S1 & S3)</p> <p>Acknowledgment of ideas (S2, S1 & S3)</p>
---	--

Three of these 13 conversations had no meaningful conversation about science content at all, meaning that out of the 72 conversations, only three of them had no meaningful science content. For example:

Student 1: D, what did you get? *Student 2:* I did D. *S1:* I did D too. *S2:* Is that what you did? *S1:* So, D. *S2:* I like how we don't even discuss a lot. It's diff. *S1:* Obvious.

In the remaining 10 conversations, we found a reoccurring pattern where a student stated an answer and explanation to the question and the other students did not engage in a conversation beyond noting that the student said something. For example:

Student 1: I said C because the propeller is moving so like it would because of like friction and everything like the propeller moving and the wheels moving would make it go forward. *Student 2:* Yeah.

Importantly, only 3 of these 10 conversations resulted in groups improving their performance from Q1 to Q1_{AD}. Therefore, the concern that increases in student performance are simply due to one student telling peers the correct answer is not supported in our data set. This result is promising given that these students had no training in effective peer conversations during clicker questions, the teachers were not involved in the conversations, and the students were not given any incentives.

Discussion and Conclusion

In this study, we found that middle school students improve their performance on clicker questions when they engage in peer discussion (Fig. 2). These results are consistent with other studies that show similar trends in undergraduate science classrooms (e.g., Smith et al. 2009, 2011; Crouch and Mazur 2001). Given that these students had no clear incentives for participation (e.g., academic credit) and were new to this type of peer instruction, our results suggest that there is considerable potential for productive clicker use in the middle school environment. Indeed, one might expect even better results if clicker use were more integrated into the classroom culture in an ongoing manner that rewarded students for their participation.

In addition, we found that middle school student groups typically used multiple knowledge co-construction elements, even when their conversations are not associated with

increases in clicker performance, although there were differences in the frequency of the elements (Table 1; Fig. 4). Across clicker performance, there was a noticeable difference in revision, and examining the source of this difference may be a useful direction for future work given the importance of revision to peer instruction. Moreover, we found that the majority of conversations included some meaningful science discussion and therefore can be considered to be on topic. We also found that the previously discussed concern about performance increases solely due to one student telling peers the answer was not supported by our data set. These two concerns have often been voiced as reasons for not implementing peer conversations in the classroom (Beatty and Gerace 2009). Results from the present investigation provide evidence that should mollify these concerns as we have shown that middle school students, even those without prior experience using clickers, had conversations that were on topic and there was some co-construction of knowledge.

Given our data corpus and analysis methods, there are a few limitations. Our analysis did not take into account the changes in conversations over time, the quality of knowledge co-construction, and the cross-pollination of ideas from one group to another. Similarly, we did not analyze issues of timing and whether some peer conversations were on topic for more or less time. Also, given the study design, there is the possibility of transfer from Q1 to Q2, which may explain why groups whose mean remained the same from Q1 to Q1_{AD} gained from Q1 to Q2. These are important issues to consider, and we would encourage future work to delve deeper into the nature of revision within peer conversations as connected to issues of timing.

Importantly, this work raises a question about the role of clicker questions in supporting learning through knowledge co-construction in classroom environments. Although classroom factors were not a central part of our study, we would encourage future work to examine classroom factors that may contribute to learning gains and knowledge co-construction. Specifically, what teachers could do to scaffold knowledge co-construction in this setting and other aspects of the classroom environment. Teachers could focus on conversation quality, the temporal dimension of these conversations, and the cross-pollination of ideas, all of which may contribute to knowledge co-construction and learning gains. Given the larger literature in science education examining student discourse in small-group settings

that involves collaborative knowledge co-construction (e.g., Hogan et al. 1999), this future research direction would also involve examining how peer conversations are similar to and different from other types of classroom knowledge co-construction. We would suggest that these discussions are likely not unique, and it is important to probe the ways in which they are similar and different from other kinds of knowledge co-construction in classroom environments.

Finally, for K-12 teachers who might consider using clickers in their classrooms, results from this study can be interpreted as promising. Our data showed that middle school students with little to no prior experience with clickers, and little pedagogical support (besides encouraging them to talk to one another), learned as a result of conversation with their peers. Furthermore, in the 72 conversations we studied, we observed minimal evidence of students off task and/or circumventing discussion by simply saying the correct answer. This result suggests that most of the conversations were productive. Given that the results were obtained with students that had little to no experience with clickers, and the instructional intervention was minimal, we posit that we would observe even greater engagement and performance if students had experience engaging in clicker-mediated discussions and if teachers were actively involved in the process and knew how to best promote peer discussion. We suggest that future work in this area could investigate the role of the teacher in supporting clicker-mediated discussion as well as ways to

design learning environments to maximize the potential for student learning.

Acknowledgments The authors thank the three teachers who allowed us to collect data in their classrooms and their students for agreeing to participate in the study. We would also like to thank Bailey Brewster for her assistance with the data analysis. This material is based upon work supported by the National Science Foundation under Grant No. 0962805. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Informed Consent Informed consent was obtained from all individual participants included in the study, and the appropriate institutional review board approved this research.

Ethical Standard All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix: Four clicker questions and their matched-pair questions

Thermal Energy Q1: Each picture has two glasses of milk. In which picture(s) can you tell that one glass of milk has more thermal energy than the other?

- A. Picture #1 only
- B. Picture #2 only
- C. Picture #3 only
- D. Pictures #2 and #3 only
- ☒ E. All three pictures

Kinetic Energy Q1: Two *identical* pennies are launched using rubber bands made of *different* materials.

Right after the pennies are launched, which statement about the penny from band 1 is true?

- A. It has *more* kinetic energy than the penny from band 2.
- B. It has *less* kinetic energy than the penny from band 2.
- C. It has *the same* kinetic energy as the penny from band 2.
- ☒ D. There is not enough information to tell if answer A, B, or C is correct.

Electric Energy Q1: Three identical light bulbs are connected in the circuit below. All three bulbs are lit. If you cut the wire where you see the scissors, which bulb(s) will stay on?

- ☒ A. Only bulb #1 will stay on.
- B. Only bulb #3 will stay on.
- C. Only bulbs #1 and #3 will stay on.
- D. Only bulbs #2 and #3 will stay on.
- E. All three bulbs will turn off.

Forces Q1: A propeller car is moving at a constant speed across a floor. The propeller is spinning. Which of the following is true about the horizontal forces acting on the car?

- A. No forces are acting on the car.
- B. There is only one force acting on the car.
- C. More than one force is acting on the car, but the overall forward force is larger.
- ☒ D. More than one force is acting on the car, but the overall forward force is equal to the overall backward force.

Thermal Energy Q2: What can you say about the thermal energy of these steel marbles?

- A. Marbles 1 and 3 have the same
- B. Marbles 2 and 3 have the same
- C. Marble 2 has more than Marble 3
- ☒ D. Marble 3 has more than Marble 1
- E. All of the marbles have the same thermal energy

Kinetic Energy Q2: A student wants to launch two identical marbles from springs made of *different* materials.

If the student wants to launch both marbles with the same kinetic energy, which of the following is correct?

- A. Spring 1 must be squeezed a greater distance.
- ☒ B. Spring 2 must be squeezed a greater distance.
- C. Both springs must be squeezed the same distance.
- D. It is not possible to launch the marbles with the same kinetic energy.

Electric Energy Q2: Three identical light bulbs are connected in the circuit below. If you cut the wire where you see the scissors, which bulb(s) will stay on?

- A. Only bulb #1 will stay on.
- B. Only bulbs #1 and #2 will stay on.
- ☒ C. Only bulbs #1 and #3 will stay on.
- D. Only bulbs #2 and #3 will stay on.
- E. All three bulbs will turn off.

Forces Q2: Is it possible for an object to move at constant speed when there is at least one force acting on it?

- A. No, there can be no forces acting on the object.
- B. Yes, if there is one constant forward force.
- ☒ C. Yes, as long as the forward forces are balanced by backward forces.
- D. Yes, as long as the forward forces are larger than the backward forces.

References

- Addison S, Wright A, Milner R (2009) Using clickers to improve student engagement and performance in an introductory biochemistry class. *Biochem Mol Biol Educ* 37(2):84–91
- Beatty ID, Gerace WJ (2009) Technology-enhanced formative assessment: a research-based pedagogy for teaching science with classroom response technology. *J Sci Educ Technol* 18(2):146–162
- Berland LK, Lee VR (2012) In pursuit of consensus: disagreement and legitimization during small-group argumentation. *Int J Sci Educ* 34(12):1857–1882
- Caldwell JE (2007) Clickers in the large classroom: current research and best-practice tips. *CBE-Life Sci Educ* 6(1):9–20
- Crossgrove K, Curran KL (2008) Using clickers in nonmajors-and majors-level biology courses: student opinion, learning, and long-term retention of course material. *CBE-Life Sci Educ* 7(1):146–154
- Crouch CH, Mazur E (2001) Peer instruction: ten years of experience and results. *Am J Phys* 69:970–977
- Dillenbourg P (1999) What do you mean by collaborative learning? In: Dillenbourg P (ed) *Collaborative-learning: cognitive and computational approaches*. Elsevier, Oxford, pp 1–19
- Doty M, Griffin J, Sutherlin AL, Sutherlin GR (2006) The effect of immediate feedback on the achievement of introductory chemistry students using a student handheld response system. Harding University, Searcy, AR and Abilene Christian University, Abilene, TX. Paper presented at the annual meeting of the MidSouth Educational Research Association's annual meeting, Birmingham, AL
- Driver R, Newton P, Osborne J (2000) Establishing the norms of scientific argumentation in classrooms. *Sci Educ* 84(3):287–312
- Engle RA, Conant FR (2002) Guiding principles for fostering productive disciplinary engagement: explaining an emergent argument in a community of learners classroom. *Cogn Instr* 20(4):399–483
- Fagen AP, Crouch CH, Mazur E (2002) Peer instruction: results from a range of classrooms. *Phys Teach* 40(4):206–209
- Freeman S, O'Connor E, Parks JW, Cunningham M, Hurley D, Haak D, Dirks C, Wenderoth MP (2007) Prescribed active learning increases performance in introductory biology. *CBE Life Sci Educ* 6(2):132–139
- Glaser BG, Strauss AL (2009) *The discovery of grounded theory: strategies for qualitative research*. Transaction Publishers, New Brunswick
- Hatano G (1993) Time to merge Vygotskian and constructivist conceptions of knowledge acquisition. In: Forman E, Minick N, Stone CA (eds) *Contexts for learning: sociocultural dynamics in children's development*. Oxford University Press, New York, pp 153–166
- Hogan K, Nastasi BK, Pressley M (1999) Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cogn Instr* 17(4):379–432
- James MC, Willoughby S (2011) Listening to student conversations during clicker questions: what you have not heard might surprise you! *Am J Phys* 79(1):123–132
- Kay R, Knaack L (2009) Exploring the use of audience response systems in secondary school science classrooms. *J Sci Educ Technol* 18(5):382–392
- King DB, Joshi S (2008) Gender differences in the use and effectiveness of personal response devices. *J Sci Educ Technol* 17(6):544–552
- Knight JK, Wise SB, Southard KM (2013) Understanding clicker discussions: student reasoning and the impact of instructional cues. *CBE-Life Sci Educ* 12(4):645–654
- Knight JK, Wood WB (2005) Teaching more by lecturing less. *Cell Biol Educ* 4(4):298–310
- Kolodner JL, Krajcik JS, Edelson DC, Reiser BJ, Starr ML (2010). *Project based inquiry science*. It's About Time
- Lasry N, Charles E, Whittaker C, Lautman M (2009) When talking is better than staying quiet. In: Sabella M, Henderson C, Singh C (eds) *Proceedings of the Physics Education Research Conference*. American Institute of Physics, 1179, pp 181–184
- Lee H, Feldman A, Beatty ID (2012) Factors that affect science and mathematics teachers' initial implementation of technology-enhanced formative assessment using a classroom response system. *J Sci Educ Technol* 21(5):523–539
- Mazur E (1997) *Peer Instruction*. Prentice-Hall, Upper Saddle River
- Miller K, Schell J, Ho A, Lukoff B, Mazur E (2015) Response switching and self-efficacy in Peer Instruction classrooms. *Phys Rev Spec Top-Phys Educ Res* 11(1):010104
- NGSS Lead States (2013) *Next generation science standards: for states, by states*. The National Academies Press, Washington, DC
- Osborne J, Erduran S, Simon S (2004) Enhancing the quality of argumentation in school science. *J Res Sci Teach* 41(10):994–1020
- Penuel WR, Boscardin CK, Masyn K, Crawford VM (2007) Teaching with student response systems in elementary and secondary education settings: a survey study. *Educ Tech Res Dev* 55(4):315–346
- Porter L, Bailey Lee C, Simon B, Zingaro D (2011, August) Peer instruction: Do students really learn from peer discussion in computing? In: *Proceedings of the seventh international workshop on Computing education research*. ACM, pp 45–52
- Roschelle J (1992) Learning by collaborating: convergent conceptual change. *J Learn Sci* 2(3):235–276
- Schwartz DL, Bransford JD (1998) A time for telling. *Cogn Instr* 16(4):475–5223
- Shkedi A (2005) *Multiple case narratives: a qualitative approach to studying multiple populations*, vol 7. John Benjamins Publishing, Philadelphia
- Smith MK, Trujillo C, Su TT (2011) The benefits of using clickers in small-enrollment seminar-style biology courses. *CBE Life Sci Educ* 10(1):14–17
- Smith MK, Wood WB, Adams WK, Wieman C, Knight JK, Guild N, Su TT (2009) Why peer discussion improves student performance on in-class concept questions. *Sci* 323(5910):122–124
- Stahl G (2006) *Group cognition: computer support for building collaborative knowledge*. MIT Press, Cambridge
- Stewart S, Stewart W (2013) Taking clickers to the next level: a contingent teaching model. *Int J Math Educ Sci Technol* 44(8):1093–1106