Eurasia Journal of Mathematics, Science & Technology Education, 2015, **11**(5), 1089-1104

Examining the Effects of Displaying Clicker Voting Results on High School Students' Voting Behaviors, Discussion Processes, and Learning Outcomes

Yu-Ta Chien, Yu-Hsien Lee, Tsung-Yen Li, & Chun-Yen Chang
National Taiwan Normal University, TAIWAN

•*Received 22 December 2014*•*Revised 3 March 2015* •*Accepted 17 March 2015*

This study explores the relationship between students' clicking behaviors, discussion processes, learning outcomes, and a prominent feature of clicker systems–the whole class' response results aggregated by clickers in real time. The results indicate that, while teaching Newton's laws of motion, displaying the real-time responses of the whole class to clicker questions can influence students' discussion processes and conceptual learning outcomes. The results have practical significance because that (1) the instructional design presented in this study (i.e., peer instruction) is widely used in clicker-integrated science instruction; and that (2) the effect sizes reported in this study are larger than the small magnitude. Implications for science teaching and technological development with clickers are discussed. A prototype of an advanced clicker system, developed based on the results of this empirical study, is presented at the end of this article.

Keywords: clicker, clicker-integrated instruction, collaborative learning, instant response system, peer instruction

INTRODUCTION

Clickers, formally called instant response devices, have gradually become an integral part of the science classroom. According to a survey by CNET News (Gilbert, 2005), it is estimated that millions of clickers are sold annually to schools around the world. These types of devices are basically signal transmitters, similar in size to television remotes, used to collect students' responses to teachers' questions in the classroom. Once the teacher poses a question, generally a multiple-choice type inquiry, students can click the buttons on their remote-like devices to vote on the answers they prefer. Students' votes are then transmitted to a central monitoring

Correspondence: Chun-Yen Chang,
Graduate Institute of Science Education, Department of Earth Sciences, National Taiwan Normal University, 88, Section 4, Ting-Chou Road, Taipei 11677, TAIWAN.
E-mail: changcy@ntnu.edu.tw
doi: 10.12973/eurasia.2015.1414a

ISSN: 1305-8223

system, typically through infrared or radio frequency signals; and, the central monitoring system then automatically aggregates a histogram, providing the teacher with the information about the answers from the entire class. By this means, every student in the classroom can express his/her thoughts instantly, and the teacher can get a rough picture of student learning progress just in time.

A substantial amount of science educators (e.g., Caldwell, 2007; Deslauriers, Schelew, & Wieman, 2011; Duncan, 2005; Smith et al., 2009) consider that clickers might be a useful tool to enhance science teaching because that (1) the use of clickers may nurture a sense of classroom participation and thus fosters social cohesion that makes students feel accountable to academic tasks; (2) the real-time data collected by clickers can assist teachers in tailoring feedback to timely address students' difficulties; and (3) if there is substantial disagreement among students' responses to the clicker questions, a productive social context might emerge to invite students to clarify the questions, infer information from the questions, and justify their own solution steps might emerge. Early reviews on the studies into clicker-integrated instruction (e.g., Caldwell, 2007; Fies & Marshall, 2006; Kay & LeSage, 2009; Lantz, 2010; MacArthur & Jones, 2008) have pointed out that students usually hold positive attitudes toward the use of clickers in the classroom. Our exhausted meta-analytic review (Chien, Chang, & Chang, in revision) further indicates that, overall speaking, the primary studies into clicker-integrated instruction have produced positive learning gains higher than those of conventional lectures.

State of the literature

- Several reviews published in this decade have indicated that clickers are gradually becoming an integral part of the science classroom.
- A common and promising instructional strategy to conduct clicker-integrated science instruction is peer instruction.
- Little empirical research has been done to examine specifically what features of clickers contribute to, or impede, peer instruction.

Contribution of this paper to the literature

- This study explores the relationship between students' clicking behaviors, discussion processes, learning outcomes, and a prominent feature of clicker systems—the whole class' response results aggregated by clickers in real time.
- This study clearly signals a sign that displaying the real-time responses of the whole class to clicker questions may influence students' discussion processes and conceptual learning outcomes.
- Implications for science teaching and technological development with clickers are discussed.

The most successful instructional strategy so far to conduct clicker-integrated science instruction might be peer instruction. The peer instruction technique was originally developed by Eric Mazur's group at Harvard University to improve teaching in physics (Crouch & Mazur, 2001; Mazur, 1997). Within the class conducted with the peer instruction technique, opportunities for students to discuss their own solutions to clicker questions should be amply provided. Once a clicker question is posed, students are asked to select their answers individually. Students' answers are instantly aggregated into a bar chart, indicating how many students voted for each possible answer. The bar chart is then shown to students, but no explanation for answers is given by the teacher. Rather, by initiating peer discussions, students are engaged in generating their own explanations to justify each choice and convince their partners. After that, students are given a chance to revote on answers. The teacher then explains why the answers are correct or incorrect. Several science educators have adapted Mazur's method and obtained positive outcomes from it. For instance, Deslauriers, Schelew, and Wieman (2011) applied the peer instruction technique with clickers to a large-enrollment college-level physics class. It was found that, compared with traditional lectures, the mean post-test score of the peer instruction class was more than twice as high. Smith et al. (2009) also adopted the peer instruction technique with clickers to teach an undergraduate introductory genetics course. It was found that students' conceptual understanding was enhanced, even when none of the students in a discussion group

 © 2015 iSER, *Eurasia J. Math. Sci. & Tech. Ed.*, **11**(5), 1089-1104

knew the correct answer. In their follow-up study, Smith. Wood, Krauter, and Knight (2011) found that either using peer instruction or giving the teacher's explanations alone was good for the students who were academically weak, but the effectiveness of peer instruction was slightly better. As for strong students, the teacher-only approach did not help at all, emphasizing the importance of peer instruction. Our meta-analysis (Chien et al., in revision) has further confirmed that peer instruction is a promising strategy to implement clicker-integrated instruction; the primary studies using peer instruction with clickers generally produced positive outcomes with a large mean effect size.

Peer instruction using clickers can be seen as an effective approach to promote active learning in the science classroom; it breaks up the passive learning format of a lecture by engaging students in thinking and discussing the solutions to teachers' questions (Caldwell, 2007; Crouch & Mazur, 2001; Deslauriers et al., 2011; Duncan, 2005; Mazur, 1997; Smith et al., 2009 & 2011). A prominent feature of clicker systems used with peer discussion, which can be found in previous paragraphs, is the display of the whole class' voting results. A common practice when using peer instruction is to show students the initial voting results prior to the group discussion (Caldwell, 2007; Deslauriers et al., 2011; Duncan, 2005; Smith et al., 2009 & 2011). The display, of voting results, is intentionally used to spur students' debates on which explanation/solution best fits a given question (Caldwell, 2007; Crouch & Mazur, 2001; Duncan, 2005). Since each student's vote counts and is represented in a concrete display, students are made aware that they are contributing to the academic task being excised in the class. As suggested by Hoekstra's case study (2008), the display of voting results seems to be a device to induce students' commitment to check their personal ideas; student voices can often be heard in communal expressions such as 'YES', 'Ughhh', and 'Ohhh!' as voting results are displayed. Students thus may have a greater sense of classroom participation and devote themselves more towards peer discussion (Caldwell, 2007; Duncan, 2005; Hoekstra, 2008).

It seems that the display of clicker voting results can be used as a learning device to engage students in peer discussion, and thereby enhance learning outcomes. However, although several comparative studies have been done in this research field (e.g., Deslauriers et al., 2011; Smith et al., 2009 & 2011), their control groups were traditional lectures, rather than peer instruction without showing voting results. Thus the research results could not reveal whether, or how, the display of voting results contributed to, or even impeded, peer discussion. Other research on this topic mainly relied on the case study approach to infer the possible effect of voting results on students' learning outcomes (e.g., Hoekstra, 2008). The power of inference was rather weak because the control group was absent. Recently, Perez et al. (2010) conducted a preliminary study to examine the usefulness of the display of voting results in a biology course. Perez et al. (2010) indicated that students were 30% more likely to switch from a less common to the most common answer if they saw voting results. They perceived that students' learning gains from clicker-integrated peer instruction might be biased by the display of voting results; students may simply shift their answers to the most chosen one without deliberation on the question. However, this opinion should be examined further because learning outcomes were evaluated only by students' final grades, rather than the difference between pre- and post-test scores. It was difficult to judge whether the display of voting results impeded learning because no solid information was provided to estimate how much students learnt from the course. Furthermore, the results were confounded with other instructional activities, as the final grades were scored based on students' performance across the whole course, rather than solely on the section of clicker-integrated peer instruction.

Therefore, we attempt to take a closer look at the relationship between the display of clicker voting results and students' learning outcomes in a more rigorous research setting. A comparative study was conducted to investigate the differences in Taiwanese 11th grade students' voting behaviors, discussion processes, and learning outcomes between two clicker-integrated physics classrooms. One of the classrooms showed students voting results before discussion whereas the other one did not. The results obtained from the study may advance researchers' and educators' understanding and practices of clicker-integrated science instruction. The results also shed light on the development of clicker technology; based on the insights obtained from this study, we have developed a more advanced clicker system to better support science teaching. Implications for science teaching and technology development with clickers are discussed. A prototype of an advanced clicker system, developed based on the results of this empirical study, is presented at the end of this article.

METHOD

Participants and instructional materials

Thirty 11th grade students from a public Taiwanese high school (School K) participated in this study. The main physics concepts for students to learn were Newton's laws of motion. From our point of view, peer instruction was basically a series of question-answer activities. Therefore, the revised Force Concept Inventory (FCI, Hestenes, Wells, & Swackhamer, 1992), a well-developed and validated test assessing students' understanding of the Newtonian concepts of force, was used as the main material to develop clicker-integrated instruction. All items of the revised FCI were multiple-choice questions, each having 5 possible answers. Two physics teachers, from different high schools in Taiwan, were invited to examine each of the revised FCI questions regarding (1) what concept(s) that students should have for solving the question, and (2) whether the question was suitable for 11th grade Taiwanese students to learn with reference to the curriculum standard issued by the Ministry of Education in Taiwan. The agreement between the two teachers was 100%. A total of 14 questions were chosen to develop research materials. The teachers were asked to categorize these 14 questions into 7 pairs to reveal the strategies to solve the questions. The inter-rater agreement in this stage was 71%. Any doubt in categorization was resolved by face-to-face discussion between the two teachers and the first author. We then followed the categorization results and modified FCI items to develop the conceptual questions—Part I (a total of 7 questions) and Part II (a total of 7 questions). Part I was used as the instructional material during the class. It was also used as the pre- and post-tests. Whereas Part II was used as the transfer test after the class.

Research design

As shown in Figure 1, all students were asked to take a pre-test to assess their prior knowledge about Newton's laws of motion. Each student was then assigned to participate either the display or non-display session. To minimize the possibility that the difference in students' prior knowledge levels may confound the study results, the assignment of students to sessions was implemented with a systematic grouping procedure. All students were grouped into 15 pairs, based on their scores on the pre-test. Then, within each pair, students were assigned to different sessions. The two instruction sessions were developed based on Mazur's peer instruction

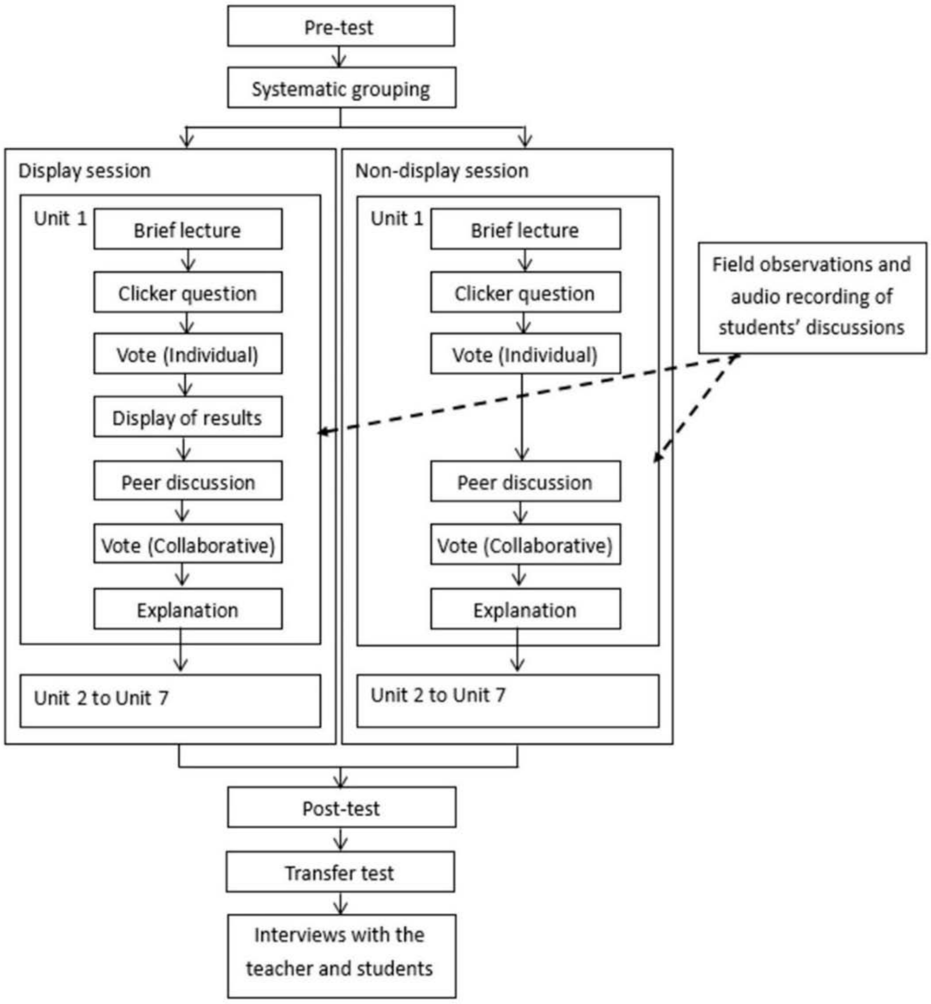

Figure 1. Design of the study

model (Mazur, 1997). Both instruction sessions were divided into 7 units, using the Part I conceptual questions as the central component. Within each unit, a multiple-choice clicker question of Part I was posed to students after a brief lecture. Students were then asked to vote answers individually by using clickers. After that, by initiating peer discussions, students were engaged in generating explanations to justify their own choices. Students were then given a chance to revote on their answers, followed by the teacher's explanations about why their answers were correct or not. The two instruction sessions (i.e., display vs. non-display) were basically the same as each other. The only difference between sessions was that one of them displayed voting results after individual voting (i.e., the display session), but the other did not (i.e., the non-display session). Each session was 4-hour long and lasted 2 weeks (2 hours per week), and taught by the same teacher. Two weeks after the instruction sessions, all students took a post-test, immediately followed by a transfer test.

Outcome variables

Four variables were defined as students' learning outcomes, including (1) the amount of individual correct responses, indicating how many in-class clicker questions per student correctly answered before peer discussion; (2) the amount of collaborative correct responses, indicating how many in-class clicker questions per student correctly answered after peer discussion; (3) the post-test score, indicating how many questions per student correctly answered on the post-test; and (4) the transfer test score, indicating how many questions per student correctly answered on the transfer test.

Instruments to obtain quantitative data

Quantitative data was obtained from the pre-test, post-test, transfer test, and the clicker system, to assist us in determining the relative instructional effectiveness between the display and non-display sessions. The items of the pre- and post-tests were exactly the same as the clicker questions presented in both sessions; they consisted of the Part I conceptual questions. These questions were pilot tested with 87 12th grades who were also from School K. The KR20 coefficient of the items was .70. The Part II conceptual questions, which were similar to but more difficult than Part I, were used as the transfer test. The KR20 coefficient of the transfer items was .63. The correlation between the post- and transfer tests was adequate, as evidenced by the spilt-half coefficient of .74. The feasibility and relevance of all instruments were asserted by 2 high school physics teachers who participated in the development of clicker-integrated instruction. All students' responses to the in-class clicker questions were automatically collected and recorded by the clicker system.

Instruments to obtain qualitative data

Qualitative data was obtained as a supplementary source of information to assist us in further understanding the differences in the relative instructional effectiveness between the display and non-display sessions, as well as the possible mechanisms underlying it. The qualitative data was obtained from field observations, interviews, and audio records. With permissions from the teacher and students, the first author entered the classroom to observe what the teacher and students were doing from the beginning to the end of both sessions. In order to take a closer look, the first author was allowed to walk around the classroom while students were doing peer discussions. However, further interactions between the researcher and students, such as exchanges of physics ideas or discussion about the clicker questions, were neither allowed nor happened during the class. The phenomena which were regarded as interesting or special were written down in field notes. Students' discussions were separately audio-taped for each group. Once the sessions were dismissed, the first author randomly interviewed with six students, three for each session, to gather information about students' perceptions toward clicker-integrated science instruction. The students were asked to talk freely about their own thoughts about the session that they had just experienced. The teacher, who conducted both the display and non-display sessions, was interviewed by the first author after both sessions were ended. The main topics of the interview were related to the teacher's (1) perceptions toward the effectiveness of clicker usages for science teaching; (2) reflection on both clicker-integrated sessions; (3) perceived difficulties in using clickers in real classrooms; and (4) suggestions for designing and implementing clicker-integrated science teaching.

Data analysis

The results and procedures of data analysis presented in this article focused on the quantitative data. Details of the qualitative part was shown because, in this article, we intended to use it as supplementary information to interpret and triangulate the quantitative results. In terms of quantitative data analysis, it was difficult to assess whether variables followed a normal distribution because the sample size was small. Therefore, the Mann-Whitney U test, a nonparametric test that does not require a particular sampling distribution, was used to examine possible differences in the outcome variables between sessions in this study. The hypothesis of testing was whether the outcome variable of the non-display session was systematically higher or lower than those of the display session. The significance level was set at .05. Medians (Mdn) and interquartile ranges (the first and third quartiles, Q1 and Q3) of the outcome variables were reported. Effect sizes were reported as r, which divided Z by the square root of the sample size. According to Cohen (1988), 0.1, 0.3, and 0.5 were deemed as small, medium, and large effect sizes for r. Excerpts of the transcriptions of field notes, interviews, and students' peer discussions were employed to interpret the instructional effectiveness of clicker-integrated instruction.

RESULTS AND DISCUSSION

Correct response rates improved by peer discussion

As shown in Table 1, the amounts of individual correct responses of the non-display session were neither better nor worse than those of the display session ($U = 110.50$, $Z = -0.09$, $p = .931$, $r = 0.02$). This provided us compelling evidence that the systematic grouping procedure did work. The difference in students' prior knowledge levels between sessions was intentionally controlled as expected. The variation in students' prior knowledge levels thus should not be a confounding variable to threat the research results. It was found that students in both sessions became more able to answer the in-class questions correctly after peer discussion, aligning with the previous literature that stressed the importance of peer instruction in science learning.

Individual—collaborative improvement magnitudes possibly influenced by voting displays

As shown in Table 1, though students of both sessions became more able to answer in-class questions correctly after peer discussion, the amounts of students' collaborative correct responses of the non-display session were systematically higher than those of the display session ($U = 49.00$, $Z = -2.71$, $p = .007$). The difference in individual—collab orative improvement magnitudes between sessions approximately reached a large size ($r = 0.49$). Students' individual and collaborative correct rates were visualized and juxtaposed to further examine how displaying voting results possibly contributed to the difference in improvement magnitudes between sessions. As shown in Figure 2, in terms of individual voting, the percentage of correct responses of the display session was nearly the same as that of the non-display session (37% [display] vs. 34% [non-display]). Moreover, once students answered correctly under the circumstance of individual voting, they tended to persist in their initial answers after peer discussion, regardless whether

Table 1. Distributions of in-class correct responses of non-display and display sessions

						Mann-Whitney U test			
Type	**Display**	**n**	**Mdn**	**Q_1**	**Q_3**	**Mean Rank**	**Sum of Ranks**	**U**	**Z**
Individual	No	15	2.00	2.00	3.50	15.37	230.50	110.50	-0.09
	Yes	15	2.00	2.00	3.50	15.63	234.50		
Collaborative	No	15	5.00	4.00	6.00	19.73	296.00	49.00	-2.71*
	Yes	15	3.00	3.00	4.50	11.27	169.00		

*Note: * p < .01*

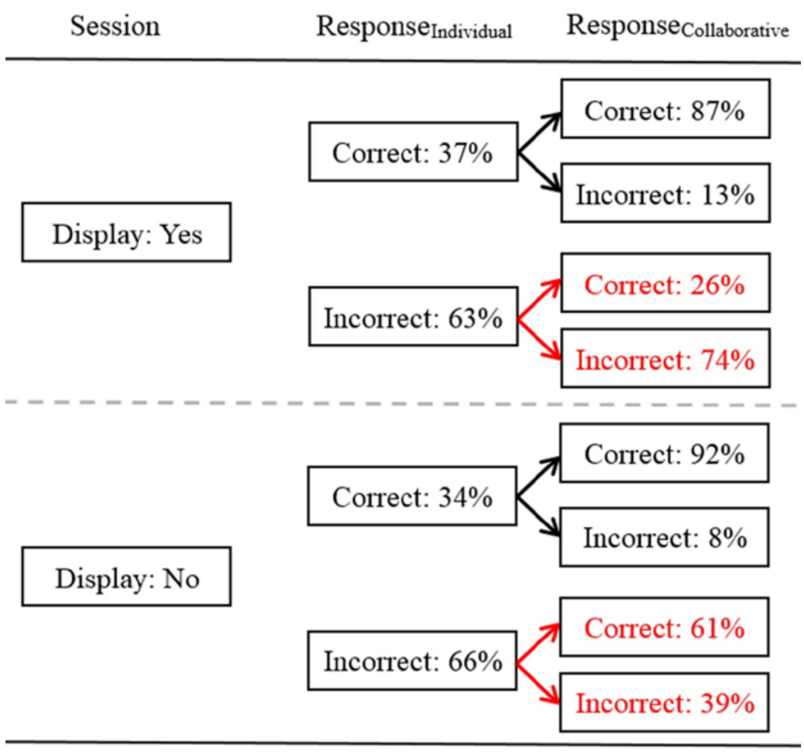

Figure 2. Students' individual and collaborative correct rates between sessions

Table 2. Distributions of test scores of non-display and display sessions

						Mann-Whitney U test			
Type	**Display**	**n**	**Mdn**	**Q_1**	**Q_3**	**Mean Rank**	**Sum of Ranks**	**U**	**Z**
Post	No	15	7.00	6.00	7.00	19.33	290.00	55.00	-2.47*
	Yes	15	5.00	4.00	5.50	11.67	175.00		
Transfer	No	15	3.00	2.00	4.00	18.83	282.50	62.50	-2.12*
	Yes	15	2.00	1.00	2.50	12.17	182.50		

Note: * $p < .05$

voting displays were shown or not (87% [display] vs. 92% [non-display]). However, when students answered incorrectly under the circumstance of individual voting, an interesting phenomenon emerged after peer discussion. If voting results were not shown, 61% of the incorrect responses were fixed through peer discussion. However, if voting results were shown, only 26% of the incorrect responses were fixed through peer discussion. The display of voting results seemed to have a significant impact on students' products of peer discussion when they did not know which answer was correct before discussion with peers.

Superior learning gains from the non-display session

As shown in Table 2, the post-test scores of the non-display session were systematically higher than those of the display session (U= 55.00, Z = -2.47, p = .013). The superior learning gain of the non-display session over the display session approximately reached a large size (r = 0.45). Moreover, on the transfer test, students of the non-display session systematically performed better than those of the display session (U= 62.50, Z = -2.12, p = .034). The superior transfer performance of the non-display session over the display session reached a medium size (r = 0.39). Over all, it might be concluded that, compared to those of the display session, students of the non-display session seemed to grasp a better understanding of Newton's laws of motion, and be more able to transfer what they had learnt in class.

Insights from qualitative data

Qualitative data provided us with information to interpret the impacts of displaying voting results on students' individual–collaborative improvement magnitudes and test performance. As shown in Figure 3, before peer discussion, students in both sessions initially did not know the correct answer to question #3; most of them preferred answer B or D (as indicated by the gray bars), but these answers were all wrong. It was noticed that, through field observations, students in the display session tended to use the information from voting results to initiate peer discussions; they usually started with the most commonly chosen answers. The transcriptions of students' discussions (additional information was added in parenthesis for better understanding) also revealed this behaviour, "which answers are chosen by most of our classmates? (D was one of the commonly chosen answers.) … Ah! I also choose D." It was further found that students preferred to work on the most commonly chosen answers, "Okay, let's take a little bit more time to think about it (referring to answer D)," and ignored other unpopular but might-be-right choices, "I choose E. … I am still thinking. …but E might be wrong (E in fact was the right answer)." Students then sticked with the commonly chosen answer, "Ok, D. Who, does who choose D as well? I choose D … you choose D … and they do, too. (We are) the same!," even when they did not figure out the scientific explanations supporting the answer (no scientific explanation was found in group #3's transcriptions although they reached a final agreement that the group answer was D). This strategy severely impeded students' thinking and led them to the dead end if the most commonly chosen answers were wrong, as provided in this case.

On the other hand, students in the non-display session just freely talked about their own ideas and explored the answers to the question as many as possible. For instance, group #3 of the non-display session examined all the answers and discussed whether each of them made sense or not. A student of group #3 started with answer A, "I choose A. … because (average) velocity is equal to distance divided by time," and then another student came out with an opposite idea that "You should not be fooled by timestamps 2 and 5 although the two blocks seem to be at the same places on these two timestamps. You have to do some calculations for solving this question." After a brief peer discussion on relative positions, average velocities, and

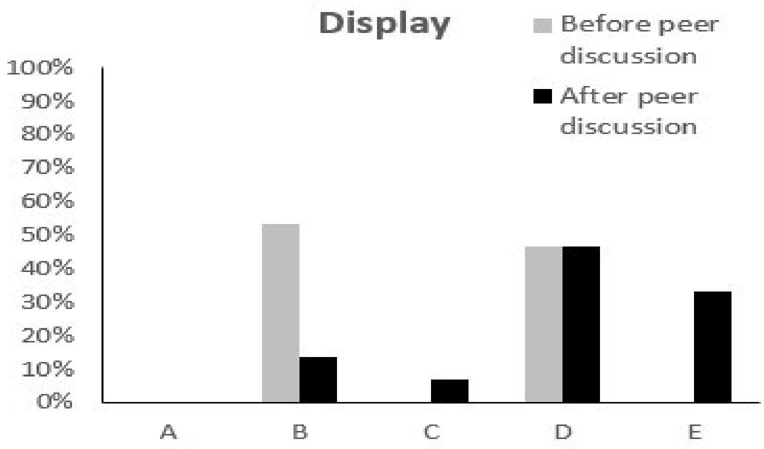

Figure 3. Students' responses to question #3 before and after peer discussion

instantaneous velocities, the whole group moved their attention to the concept of acceleration, "…look at that, the block above the line moves faster and faster, whereas the block below the line moves with a constant speed." They then tried to calculate the acceleration rate of the block above the line. Though this task was difficult to achieve by using the information of the questions, it made them rethink whether there were other strategies to solve this question. A student said that "At a certain point, the two blocks should have the same velocity because the block above the line keeps accelerating and it will move faster than the block below," and other students responded with that, "They will have the same velocity at a certain point," and that "If (their velocities are) the same, the block above the line will produce a displacement that equals to one of the below block's, I think." Students thus reexamined all the answers to see when the two blocks produced the same displacement, such as that "They indeed have the same velocity at a certain point so A is false," and that "B, C, and D must be wrong because they are indicating the same positions, not the same velocity." Finally, they successfully identified that the correct answer was E by examining all possible answers.

The observed sharp distinction between students' discussion strategies offered a possible explanation for the differences in students' learning gains between sessions; displaying voting results could guide/limit the directions of students' discussions. The voting results would induce students to focus on the popular answers while they were collaboratively constructing scientific explanations. However, if most of students went to the wrong direction, as led by the voting results, the outcomes just went bad because they did not make time to think about other possibilities. As a consequence, students would build improper explanations or even no explanation to the targeted scientific phenomena, as suggested by the post-test scores. They, therefore, had difficulties in applying what they had constructed to solve new but similar questions, as suggested by the transfer test scores.

The teacher's concerns

The teacher and student participants, in this study thought the clickers were useful and joyful for science learning. However, as informed by the teachers, it is impractical for schools to buy or rent clickers because the cost is still high at the present stage. The safekeeping of clickers in the classroom is also a great challenge to teachers. Furthermore, teachers have to spend a lot of time in installing the system and distributing clickers to students. In summary, the teacher held a very positive attitude toward the use of clicker for science teaching. He also did believe that he can make good use of clickers to enhance his teaching and let students be more willing to participate in the science class. However, he was dissatisfied with the design and cost of clicker systems currently available in the market.

IMPLICATIONS FOR TEACHING AND TECHNOLOGY DEVELOPMENT

This study clearly signals a sign that displaying the real-time responses of the whole class to clicker questions may influence students' discussion processes and conceptual learning outcomes. The results have practical significance because that (1) the instructional design presented in this study is widely used in clicker-integrated science instruction; and that (2) the effect sizes reported in this study are larger than the small magnitude. In a real classroom, displaying the real-time responses is indeed fun for both teachers and students; it makes learning just like participate in the Who Wants to Be a Millionaire TV show. However, the current study serves a warrant that science teachers should rethink whether displaying the real-time responses of the whole class is needed, especially when students cannot answer correctly under the circumstance of individual voting. A modest and quick

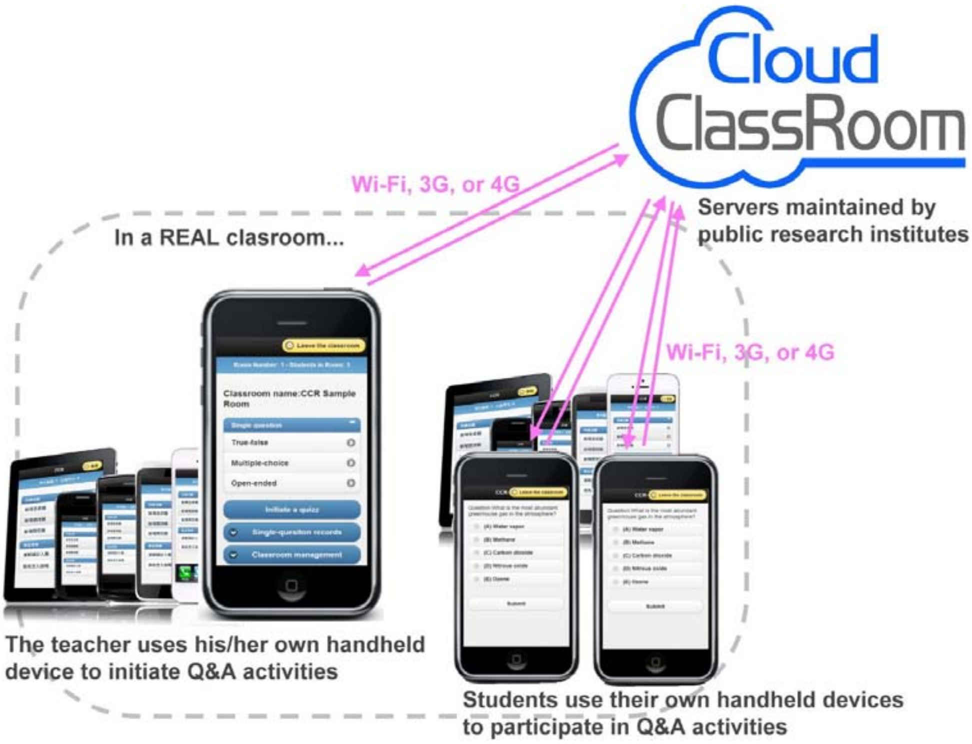

Figure 4. The conceptual framework of how CCR works

suggestion is as following: If most of your students answer the clicker question correctly before peer discussion, you can simply display the voting result so students might get an "I am doing great!" positive reinforcement. Otherwise, hiding the voting results might be a wiser decision to prevent students from being distracted and misled by the most popular but wrong ideas. Students might thus be more engaged in exploring ideas as many as possible, and in clarifying their own thoughts from multiple points of view.

The results from this exploratory study provide us with fresh and practical insights for improving the instructional design of clicker systems to support science teaching. The first insight from the study is about technical problems. The first, third, and corresponding authors of this article have come up with an alternative solution with emerging technologies to overcome the difficulties and concerns reported by the teacher: transforming teachers' and students' own smart handheld devices, such as smart phones and tablet computers, into clickers. According to a survey released by the Yahoo company in 2014 (Lo & Wu, 2014), the use of smart handheld devices for Internet connections is more extensive in Taiwan than in any other country in the world; it is estimated that one in three Taiwanese people has a smart phone or tablet computer. The high popularity and usage of smart handheld devices are mainly because of the wide coverage of Wi-Fi networks as well as unlimited Internet usage packages offered by local telecommunications. A web-based clicker system, called CloudClassRoom (CCR), thus has been developed, using HTML 5.0 and MySQL databases, with the aim to replace traditional clickers (Chien & Chang, in press). Such a technical design makes CCR work on every Internet-capable device without further software or plug-in installation. It also allows CCR to operate across-platforms, regardless of iOS, Android, or Windows being used; and it is compatible with a range of devices, such as smart phones, tablet computers, personal computers, and laptops. As shown in Figure 4, teachers and students are enabled to transform their own devices into clickers once they connect their devices with CCR on the Internet. The difficulties and concerns reported by the teacher thus might be substantially resolved due to the following reasons: (1) the coverage of Wi-

Fi networks is wide among Taiwan's schools, and it is continuing to get wider; (2) no hardware and software are required for purchase or installation because the only thing that teachers and students need to do is to connect their own devices to the Internet. In addition, CCR does not charge any fee to use because it is sponsored by public research institutes in Taiwan; and (3) no more safekeeping and distribution problems are imposed on teachers because students use their own devices as clickers. It is common that students in Taiwan bring their own smart handheld devices to schools and carry them everywhere.

Web-based clickers, such as CCR, have several strengths to overcome the weakness of traditional clickers. First, the price of traditional clickers is still unreasonably high (30 to 50 USD per clicker). In terms of implementing the active learning activities similar to the instruction described in this article, adopting web-based clickers is a more economical solution for schools, especially for those already have Internet access and Internet-capable devices. Such a solution will become more cost-effective as more and more schools start embracing the Bring-Your-Own-Device (BYOD) policy (Johnson, Adams Becker, Estrada, & Freeman, 2014); web-based clickers are well compatible with the devices that many students already own, such as laptops, smart phones, or tablets. Second, while using traditional clickers, the question formats are limited to true-false and multiple-choice questions because traditional clickers only accept numeric responses. However, the use of open-ended questions, compared to true-false and multiple-choice questions, is regarded as a more effective way to stimulate students' higher-order thinking (Brookhart, 2010). On the contrary, web-based clickers certainly can facilitate teachers' use of open-ended questions because textual responses are enabled through the use of PCs, laptops, smart phones, or tablets. Third, Students can upload photos, snapshotting either their drawing or others, to the web-based clicker system as their responses. Compared to solely requiring textual responses, asking students also to turn in photographic responses may make students more cognitively engaged and provide teachers with more information to evaluate students' understanding (Van Meter & Garner, 2005). And, finally, web-based clickers are a far better research tool than traditional clickers in terms of obtaining learning analytics in a large-scale manner. For instance, if a school tends to implement traditional clickers into 100 classrooms, each of the classrooms must be independently installed with a monitoring system to record students' responses. This kind of system design is highly inefficient for the school to investigate students' learning progression from a larger-scale of view; the school has to retrieve students' data from each classroom and then merge the data for further analysis. On the contrary, web-based clickers operate on clouds, and thus the school is enabled to use one central monitoring system to access all students' data, regardless the data is collected from which classroom. This system design also facilitates national or even global level research on learning analytics obtained from a group of real classrooms.

Other insights from the study are about how to make the clicker system become a more powerful pedagogical tool. The results from this exploratory study can be transformed as several research-informed instructional functions in CCR. First, we find that traditional clicker systems cannot provide teachers with adequate information to decide whether students' voting results should be publicly shown. This drawback is resulted from that traditional clicker systems project all information to the only one screen on the podium; if a teacher wants to see the distribution of the whole class' responses, he/she must project the whole class' responses to the screen on the podium. In other words, all students will see the whole class' responses at the same time once the teacher is trying to examine their responses. The teacher, in essence, has no opportunity to check the whole class' responses before he/she displays it publicly. CCR is thus embedded with a flexible broadcast function to overcome the aforementioned drawback of traditional clicker systems. As shown in the left panel of Figure 5, if students start answering the teacher's question, CCR will display their responses only on the teacher's device, rather than on students' devices. Once the teacher presses the "Broadcast the result" bottom, all students in the classroom will get a histogram on their own devices, as shown in the right panel of Figure 5, which automatically aggregates the answers from the entire class for their reference. This action totally depends on teachers' pedagogical decisions regarding when to broadcast voting results might be more productive for peer instruction. Teachers thus may have adequate information and time to decide whether and when students' voting results are publicly shown. Nonetheless, as recommended in the early part of this section, not to display the voting result might be a wiser decision to facilitate peer instruction when most of the students incorrectly answer the clicker question before discussing with peers. However, such a decision may decrease students' sense of classroom participation because no concrete information is presented as evidence for them to be aware of that they are contributing to the academic task being excised in the class.

Second, we are working on a more advanced CCR function to cope with the aforementioned catch-22 situation. The results of this study have indicated that displaying the real-time responses of the whole class to clicker questions may influence the directions and processes of students' discussions. Furthermore, it has been observed that in this study, students in the display session did use the information of whole-class voting results as a resource to initiate, and orchestrate, peer discussions. It might be reasonable to assume that manipulating the information of whole-class voting results, if properly used, can enhance the quality of students' discussions. As shown in the left panel of Figure 6, we are modifying CCR to provide teachers with a function to trim the distribution of students' responses to the clicker question. Once students respond to the clicker question, the teacher is thus enabled to reshape the distribution of students' responses, as shown in the right panel of Figure 6, by setting the percentage composition of answer

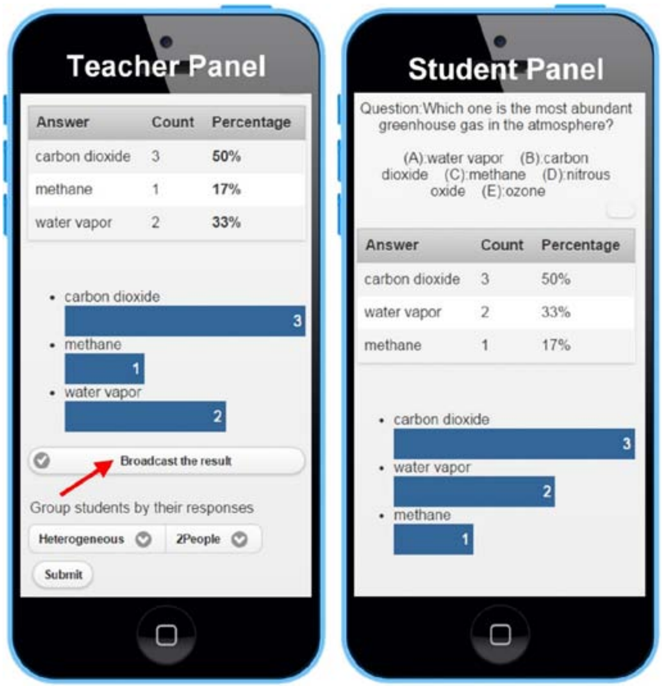

Figure 5. The flexible broadcasting function of CCR

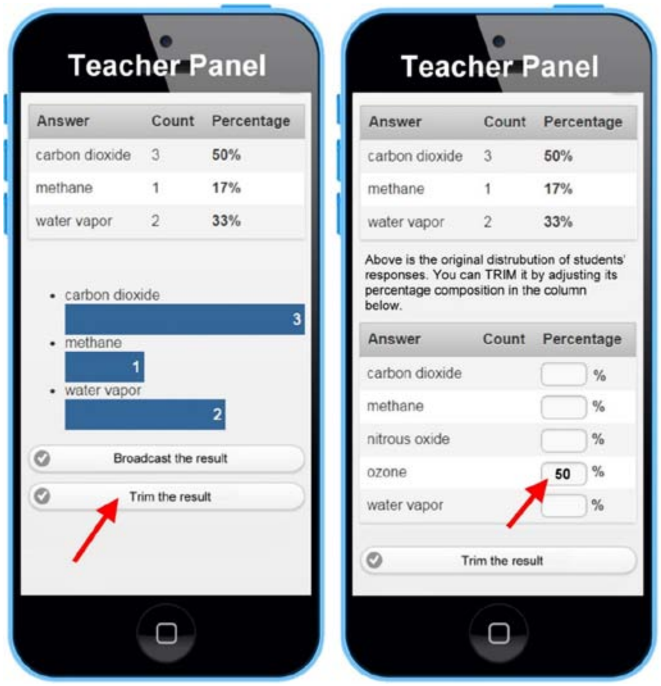

Figure 6. The real-time trimming function of CCR

options. In order to thoughtfully guide students' discussions, the teacher may make the distribution become flatter if most of students originally vote for one single but wrong answer. Students may start to explore other answers that they did not consider before when they receive the trimmed histogram. This instructional function might be an ideal and flexible solution to deepen students' discussions without decreasing students' sense of classroom participation.

LIMITATIONS AND FUTURE DIRECTIONS

Studies with larger sample sizes are required to verify the results. Fine-grained and moment-by-moment observations of what is happening in the clicker-integrated classroom are also warranted to understand further how students learn with clickers. The effectiveness of the instructional clicker functions, proposed in this study, should be further examined by conducting empirical studies, as what the core members of the CCR development team are undertaking in Taiwan. We are currently working on the intelligentization of the trimming function that aims to assist teachers in reshaping the distribution of voting results by simply clicking, rather than putting numbers into the system.

AUTHORS' NOTES

Part of this study has been accepted by the 2015 Annual International Conference of the National Association for Research in Science Teaching (NARST). Its abstract has been presented at the 2015 NARST conference. Sponsored by public research institutes in Taiwan, the CCR system is free to non-profit educational institutions all over the world. The system now has six different language versions, including Traditional Chinese, English, French, Japanese, Korean, and Turkish. For those who are interested in using CCR for non-profit educational purposes, we suggest you contact the corresponding author for more detailed information.

ACKNOWLEDGEMENTS

This research is partially supported by the Aim for the Top University Project of National Taiwan Normal University (sponsored by the Ministry of Education, Taiwan, R.O.C.), the International Research-Intensive Center of Excellence Program (sponsored by National Taiwan Normal University and the Ministry of Science and Technology, Taiwan, R.O.C., under grant no. MOST 104-2911-I-003-301), and the Ministry of Science and Technology, Taiwan, R.O.C. (under grant no. MOST 102-2511-S003-052-MY3). The authors gratefully acknowledge the assistance of Terrence Wong in editing the manuscript. Special thanks to the teachers and students who participated in this study.

REFERENCES

Boscardin, C., & Penuel, W. (2012). Exploring benefits of audience-response systems on learning: A review of the literature. *Academic Psychiatry, 36*(5), 401-407.

Caldwell, J. E. (2007). Clickers in the large classroom: Current research and best-practice tips. *CBE-Life Sciences Education, 6*(1), 9-20.

Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom.* Alexandria, VA : ASCD.

Chien, Y. T., & Chang, C. Y. (in press). Supporting socio-scientific argumentation in the classroom through automatic group formation based on students' real-time responses. In M. S. Khine (Ed.), *Science education in East Asia: Pedagogical innovations and research-informed practices.* Berlin: Springer.

Chien, Y. T., Chang, Y. H., & Chang, C. Y. (in revision). Do we click in the right way? A meta-analytic review of clicker-integrated instruction. *Educational Research Review.*

Cohen, J. (1988). *Statistical power analysis for the behavioral sciencies* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics, 69*(9), 970-977.

Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science, 332*(6031), 862-864.

Duncan, D. (2005). *Clickers in the classroom: How to enhance science teaching using classroom response systems*. San Francisco, CA: Pearson.

Fies, C., & Marshall, J. (2006). Classroom response systems: A review of the literature. *Journal of Science Education and Technology, 15*(1), 101-109.

Gilbert, A. (2005). New for back-to-school: 'Clickers'. Retrieved May 1, 2015 http://news.cnet.com/New-for-back-to-school-Clickers/2100-1041_3-5819171.html

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher, 30*(3), 141-158.

Hoekstra, A. (2008). Vibrant student voices: Exploring effects of the use of clickers in large college courses. *Learning, Media and Technology, 33*(4), 329-341.

Johnson, L., Adams Becker, S., Estrada, V., & Freeman, A. (2014). *NMC horizon report: 2014 K-12 edition.* Austin, TX: The New Media Consortium.

Kay, R. H., & LeSage, A. (2009). Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers & Education, 53*(3), 819-827.

Lantz, M. E. (2010). The use of 'Clickers' in the classroom: Teaching innovation or merely an amusing novelty? *Computers in Human Behavior, 26*(4), 556-561.

Lo, H. W., & Wu, J. (2014). Taiwan has world's biggest smartphone craze: Yahoo survey. Retrieved May 1, 2015 http://focustaiwan.tw/news/aeco/201408130034.aspx

MacArthur, J. R., & Jones, L. L. (2008). A review of literature reports of clickers applicable to college chemistry classrooms. *Chemistry Education Research and Practice, 9*(3), 187-195.

Mazur, E. (1997). *Peer instruction: A user's manual*. Upper Saddle River, NJ: Prentice Hall.

Perez, K. E., Strauss, E. A., Downey, N., Galbraith, A., Jeanne, R., & Cooper, S. (2010). Does displaying the class results affect student discussion during peer instruction? *CBE-Life Sciences Education, 9*(2), 133-140.

Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J. K., Guild, N., & Su, T. T. (2009). Why peer discussion improves student performance on in-class concept questions. *Science, 323*(5910), 122-124.
Smith, M. K., Wood, W. B., Krauter, K., & Knight, J. K. (2011). Combining peer discussion with instructor explanation increases student learning from in-class concept questions. *CBE-Life Sciences Education, 10*(1), 55-63.
Van Meter, P., & Garner, J. (2005). The promise and practice of learner-generated drawing: Literature review and synthesis. *Educational Psychology Review, 17*(4), 285-325.

◈◇◈