Australasian Journal of Educational Technology, 2013, 29(5).

Implementing the interactive response system in a high school physics context: Intervention and reflections

Ruey S. Shieh
Kainan University, Taiwan

Wheijen Chang
National Changhua University of Education, Taiwan

The interactive response system (IRS) has been widely used to promote student learning since 2003. It is an electronic system connected to handset devices allowing students to transmit their responses by pressing the desired buttons and meanwhile allowing the teacher to monitor and track individual students' answers anonymously and statistically. However, there is limited research examining the challenges teachers may encounter when designing IRS-based questions and providing mediations which may lead them to develop quality questions. The purpose of this study is to address this research gap by investigating one high school teacher's IRS implementation based on both the teacher's and students' teaching/learning experiences as well as presenting an intervention to help the teacher develop higher quality IRS questions. High quality questions denote questions that are able to help students engage in deeper thinking and eventually lead to comprehensive understanding of the concepts learned. The data sources consist of tests, classroom observations, interviews, face-to-face meetings, and email correspondence. The findings disclose that enhancing the teacher's content knowledge and capability of recognizing the students' learning pitfalls is the foundation to developing quality IRS questions. Collaboration established between the teacher and a university physics education expert appears to have effectively helped both participants gain insights and knowledge into designing quality questions aimed at identifying the students' learning bottlenecks.

Introduction

The interactive response system (IRS), also known as clicker, electronic response system (ERS), personal response system (PRS), classroom response system (CRS), and audience response system (ARS), was first introduced to the educational setting in the mid-1960s and has been widely used since 2003 (Kay & LeSage, 2009). It enables the instructor to elicit, assess and track students' responses to the discussed questions instantly, individually, and anonymously. It has also been found to help increase students' participation (Creese, 2011) and engagement (Bartsch & Murphy, 2011; Laxman, 2011), enhance class interactivity (Draper & Brown, 2004) and learning outcomes (Yourstone, Kraye, & Albaum, 2008, Henriksen & Angell, 2010), and promote higher-order thinking (Connor, 2011). In the area of physics, the device is regarded as a promising instructional tool in terms of engaging students in cognitive participation and facilitating their conceptual understanding (Beatty et al., 2006a; Beatty & Gerace, 2009). A number of research and educational institutes have further developed IRS-based pedagogy to help students acquire useful knowledge and/or problem-solving skills in science subject matters, such as Assessing-to-Learn (Dufresne & Gerace, 2004) or Question-Driven Instruction, designed by the University of Massachusetts Physics Education Research Group. Technology-enabled active learning (TEAL), initiated at Massachusetts Institute of Technology (MIT), also employs the personal response system to strengthen students' conceptual understanding in introductory physics courses more profoundly and analytically (Dori & Belcher, 2005). Likewise, Lin, Liu, and Chu (2011) developed an instructional model, called Clicker-Assisted Conceptual Change, to help students enhance their introductory physics concepts through the use of the clicker.

In Taiwan, the use of the IRS in educational institutions has also been increasingly prevalent in the past decade due to its decreasing cost (Yeh & Tao, 2012). In 2009, the Ministry of Education granted 50 high schools funding to construct TEAL studios on their campuses to promote a technology-enhanced teaching/learning environment (Shieh, 2012). TEAL, as an innovative pedagogical structure, emphasizes active, interactive, and collaborative learning (Dori & Belcher, 2005). The overall goal of TEAL is to establish a format that engages students in learning physics more meaningfully so that they can acquire a better understanding of the studied content. The IRS is a built-in feature for facilitating TEAL implementation. It is an electronic system connected to handset devices with a 10-digit numeric keypad. Via infra-red or radio signals, students can transmit their responses by pressing the desired buttons. The handsets are mapped to each student's ID number or name, which allows the teacher to monitor and track individual students' answers anonymously and statistically. Shieh (2012), in her study investigating the impact of TEAL on teachers and students, reported that implementation of TEAL helped foster the TEAL teacher's confidence in teaching physics courses, and the use of IRSs helped increase class interactivity and the students' study interest (Shieh, Chang, & Liu, 2011). However, whether and how the use of IRSs helped the students enhance their conceptual understanding was not examined. Although Lin et al. (2011) is one rare study addressing the students' conceptual change issue, the studied subjects were, however, a group of college students. The purpose of this study is therefore to address this research gap through exploring one high school teacher's IRS implementation based on both the teacher's and students' teaching/learning experiences. In addition, collaboration between the teacher and a physics education expert for designing IRS questions was also implemented. Whether and how such intervention helps improve the quality of the designed IRS questions is also reported in this study.

Design of IRS questions

The effectiveness of the IRS depends greatly on the quality of the questions asked (Beatty, Leonard, Gerace, & Dufresne, 2006b). Developing and designing effective questions is not only challenging, but it is also different from composing exam and homework problems (Beatty, Gerace, Leonard, & Dufresne, 2006a). Beatty et al. (2006a) suggested that a holistic perspective of question design, rather than traditional questions provided by textbooks, is necessary to enhance student learning. They suggested that questions be developed based on three goals: a content goal (focusing question cycle iterations on the foundational ideas and concepts), a process goal (making knowledge about physics useful in encountered situations), and a metacognitive goal (promoting students' beliefs about what and how to learn physics). They also identified several tactics for modification, such as removing nonessentials, comparing and contrasting, interpreting representations, and strategizing.

Mazur (1997) proposed IRS assisted peer instruction to enhance student learning. While answering the IRS questions, students are required to construct their individual answers first and then convince their peers of the correctness of their answer, which allows the instructor to assess the students' responses instantly and formatively. The feature of instant feedback helps reinforce the students' reflection, awareness, and ownership of their learning, and consequently expand their beliefs of what, how, and what for in learning physics (Baird, Fensham, Gunstone, & White, 1991). Beatty et al. (2006a, 2006b) suggested that a question cycle be adopted to more effectively implement the IRS-based instruction in teaching sciences. The purpose of the cycle is to (a) promote and focus student learning with question-driven instruction, that is, to highlight the key points and to retain concentration; (b) strengthen students' conceptual comprehension and scientific literacy with dialogical discourse, (c) use formative assessment to acquire better insights into students' learning pace and difficulties in order to grasp the optimal instructional decisions (Bell & Cowie, 2001), and (d) foster students' metacognitive skills, and cooperate in the learning process with meta-level interaction. In their view, while the instructor endeavours to implement agile teaching (teaching with a tight feedback pace, continually guiding students to examine their learning progress and bottlenecks), the students engage in active learning (learning with diligent, directed cognitive activity to absorb new ideas and form new knowledge). Beatty and Gerace (2009) developed the technology-enhanced formative assessment (TEFA) pedagogy for IRS-based science instruction, which stresses using formative assessment to narrow the achievement gap between students with various initial beliefs and knowledge, and assisting students to develop metacognitive skills. Although TEFA does not necessarily require a classroom response system, the researchers contended that the use of the technology does however help implement what TEFA is meant to accomplish. They argued that the tool allows students to answer the questions raised simultaneously and anonymously, while socially holding them accountable for the answers they provide. They also suggested that questions be built as sets to enhance students' understanding, and that the cycle be iterated three or four times in 50-60 minutes of TEFA instruction. Similarly, Hancock (2010) mentioned that three to six questions per one 75 minute session is appropriate.

Pedagogy and technology

When integrating technology such as IRSs into instruction, Buckenmeyer and Freitas (2006) argued that one must begin with the pedagogy of why (objectives), how (approaches), and what (content) one teaches, rather than beginning with the technology. Beatty et al. (2006b) emphasized that well-designed questions are merely a tool, and it is the pedagogy of how the instructor uses the questions to interact with the students that is more important. Some researchers (e.g., Mishra & Koehler, 2006; Koehler & Mishra, 2008) have asserted that three types of knowledge - technology, pedagogy, and content knowledge (TPACK) - must be closely connected to successfully integrate technology into instruction. They contended that effective teaching with technology demands an understanding of the representation via technology to illustrate the meanings of scientific concepts; pedagogical skills that employ technologies in the teaching sequence to enrich the classroom activities; knowledge of resolving the problems students encounter with the use of technologies; and knowledge of how technologies can be used to strengthen existing knowledge. In short, technology needs to cooperate with good teaching design and pedagogical skills in order to achieve the attempted instructional objectives.

Pedagogical content knowledge (PCK), initially proposed by Shulman (1986), includes the knowledge of students' prior understanding and potential learning difficulties (Magnusson, Krajcik, & Borko, 1999). To cope with students' learning difficulties, some researchers, such as Viennot, Chauvet, Colin, and Rebmann (2005), have argued that the crucial role of critical details in pedagogy may appear trivial to teachers but may actually be decisive for students in discerning the concepts taught. However, teachers may not be perceptive of such teaching subtlety in their teaching practice without clear guidance (Pint´o, 2005). To cope with this issue, Lijnse *(*1995) suggested that intervention from field experts be established to formulate a detailed description of didactical structures for a certain topic, based on science education research. Some researchers have contended that scaffolding provided by field experts, and partnerships between university faculty and teachers offer a venue to strengthen teachers' content knowledge, pedagogical knowledge, and content pedagogical knowledge (Viennot et al., 2005; Etkina, 2010). According to the National Research Council (Committee on Biology Teacher Inservice Programs, National Research Council, 1996), science education researchers are obliged to facilitate teacher professional development. When designing course content and questions, in addition to providing assistance to teachers, the science experts may, in the meantime, gain an opportunity to ascertain language and usages comprehensible to school pupils, which, as Lee, Ding, Reay, and Bao (2011) have argued, are normally ignored. Likewise, when developing IRS questions, partnerships between teachers and university science experts are regarded as being able to enhance the quality of the questions designed. High quality questions denote questions that are able to help students engage in deeper thinking and which, as earlier researchers have contended, eventually lead to comprehensive understanding of the learned concepts.

Although the aforementioned literature has provided useful guidance and/or theoretical underpinnings for designing IRS oriented instruction, there is limited research examining how school teachers design IRS questions, what challenges and benefits the participants may come across when using IRSs, and whether and how collaboration established between teachers and university scientists helps improve the quality of the questions designed. This study therefore addresses these research inquiries by exploring the following three research questions:

1. How does a teacher construct IRS questions?
2. What benefits and challenges do the teacher and the involved students encounter when teaching/learning with an IRS?
3. Whether and how does the partnership between the teacher and a physics education expert help strengthen the quality of the IRS questions developed?

Method

Research context and research design

The study took place at a senior high school in central-southern Taiwan. A physics teacher (Mr. Yen, a pseudonym) teaching at the school and the three 10th-grade classes he taught in 2010 (February – June) were involved in the study. The content of the course, basic physics, consisted of three units with three chapters forming a unit (Chapters 1-3, Chapters 4-6, and Chapters 7-9). Each chapter was scheduled to be taught over an average of two weeks of lessons (three 50-minute classes per week). Three uniform examinations, labelled as review tests, corresponding to the three units, were held in the pencil-and-paper format by the school at six-week intervals to assess all 10th-graders' overall understanding. Differing from other teachers at the school, who employed the traditional lecture approach, Mr. Yen used the TEAL studio to teach the course. In the first lesson of teaching a chapter, Mr. Yen would conduct a test, labelled as the pre-test, consisting of 10 to 15 questions with the use of IRS to examine the students' prior knowledge. The questions were structured in Microsoft PowerPoint format, with one question per slide, where four to five alternatives were provided for each multiple-choice question. After the pre-test, Mr. Yen used a PowerPoint presentation, 3D simulations, and hands-on experimental activities to teach the lesson. In the last lesson of the chapter, another IRS test, labelled as the post-test, also consisting of 10 to 15 questions was administered to the students to assess their learning outcomes.

To reflect the effect of the IRS implementation from the researchers' angle, another test, called the comprehensive test, was designed to further assess the students' learning results. An outsider (Prof. Pang, a pseudonym), an expert in developing introductory physics course content at a university in Taiwan, developed the test questions based on three aims: (a) examining the sophistication of the students' conceptual understanding, (b) confronting the prevalent alternative conceptions, and (c) promoting the demand for logical reasoning while answering the questions. The test covered the same topics as the second review test (Chapters 4~6), including wave, Newtonian mechanics, electrostatics, and electromagnetism. This unit was chosen for the comprehensive test because secondary students tend to encounter considerable difficulties in learning the covered topics (Maloney, O'Kuma, Hieggelke, & Van Heuvelen, 2001; Wittmann, Steinberg, & Redish, 1999). While the review test consisted of 40 multiple-choice questions, the comprehensive test only contained 20 questions, mainly revised from the review test questions. However, five of the questions were modified from the original multiple-choice questions into multiple-key questions. After the initial draft of the test questions was developed, Mr. Yen was invited to discuss the content of each question with Prof. Pang to assure its appropriateness in terms of wording and content. The test content was finalized after multiple exchanges of thoughts and feedback. The test was held one month after the review test. Differences between the review and comprehensive tests are described in the Results section.

Data collection and data analysis

Both quantitative and qualitative data were collected in the study. The quantitative data consisted of the aforementioned four tests: the IRS pre-tests, the IRS post-tests, the review test, and the comprehensive test. Five sources of qualitative data were gathered as follows.

1. Classroom observations: Two observations in each participating class were conducted. Seven semi-structured items were designed for observing the classes, including attendance, interaction between the teacher and the students, interaction in small-group discussions, number of IRS questions posed, correct ratios of the questions posed, student engagement in the use of IRS, and the overall atmosphere of the class.
2. Interviews with the teacher: two digitally recorded interviews were conducted, one at the beginning of the semester and the other at the end. A semi-open-ended interview protocol was used to compose interview questions, including the teacher's teaching experiences, instructional objectives of the course, methods used to design the IRS questions, the teacher's expectations of the students in terms of engaging in the IRS questions, and reflection on the overall IRS implementation.

3. Group interviews with the students: Three group interviews (one group per class) were conducted at the end of the course. Six students in each class were randomly selected. A list of six questions was included in the semi-structured interviews: attitudes toward learning physics, learning preferences, perceived class interaction, reactions to small-group discussions, attitudes toward using TEAL as well as IRS, and self-reflections on the course. The students took turns addressing each of the questions. Each interview lasted for 30-40 minutes, and all were digitally taped.
4. A face-to-face meeting with the teacher and the tertiary expert: the main purpose of the meeting was to discuss and exchange thoughts about the content of the comprehensive test questions and other test related tasks. The meeting lasted for a total of three hours.
5. Email correspondence with the teacher and the expert: email contacts, which frequently took place among the participants and the researchers, were gathered. The purpose of the contacts were to clarify earlier verbal conversations, verify the test design, confirm test schedules, follow up related activities, analyse and share the test results, and obtain needed information.

Statistical analysis of the quantitative data included correct ratio calculations of the test results and χ^2 tests. Comparisons between the IRS pre- and post-tests and between the review and comprehensive tests were conducted to assess the students' overall performance. Content analysis was used to analyse the question content. In addition, the theoretical framework of social construction proposed by Patton (2002) was adopted to analyse the qualitative data, particularly the interview data. To increase the reliability of the coding, a research assistant majoring in education and the first author conducted the coding separately. The two coders adopted a non-judgmental principle to code and analyse the data, which contributed to the validity of the information shared by the participants. The two sets of coding were compared and discussed until agreement was reached. Moreover, checks were administered by the teacher and the participating expert of the interview and the face-to-face conversation content. The classroom observations, interview analyses, and documents collected were triangulated to strengthen the reliability of the findings.

Results

According to Mr. Yen, he had more than 20 years of teaching experience at the school. He reported that he had used the IRS to teach courses since 2006 when he began to adopt the concept of TEAL to teach physics courses. He stated that it took him three solid years to become proficient in managing the multimedia facilities built in the TEAL studio. He was the only teacher at the school who adopted the technology-enhanced studio to teach courses as he firmly believed that integrating technology into instruction could more effectively help students acquire physics concepts. The three classes he taught in 2010 had a total of 128 students (43, 42, and 43, respectively). The participants' experiences of using the IRS, the sources Mr. Yen used to compose IRS questions, student performance, as well as the impact of the collaboration between Mr. Yen and Prof. Pang are described below.

Benefits and challenges of using IRS

Mr. Yen stated in the first interview that many students regarded physics as a difficult subject to learn, which was also echoed by the majority of the students interviewed. The adoption of IRS "clicking" activities and ensuing small-group discussions appeared to be Mr. Yen's approach to alleviate the students' adverse attitudes toward studying physics, and meanwhile, to increase their learning interest and class interactivity. The classroom observation data disclosed that not all students concentrated on listening to the teacher's lecture; a few even had their heads lowered, reading some other materials. Compared to the lecture presentation, the class atmosphere became much more active and interactive when the IRS activity took place. The interactive situation began when the teacher asked the students to engage in small-group discussion prior to responding individually to the raised IRS questions. Mr. Yen mentioned that through the students' responses to the posed questions, he was able to monitor their learning progress and then adjust his lesson content accordingly. In other words, the use of the IRS reportedly helped him fulfil the in-class formative assessment. Consistent with Mr. Yen's expectations, most of the students interviewed also mentioned that the use of IRS was "interesting and exciting". Many reported that they were eager to know whether they had outperformed their peers. Some stated that the IRS helped them grasp the key concepts of the lecture content; some expressed that small-group discussions helped them better comprehend the concepts discussed. One said that teaching group members the concepts embedded in the questions increased his sense of accomplishment.

Despite these benefits, however, several challenges associated with the use of IRS were also reported. First of all, Mr. Yen mentioned that the content of the IRS questions must be concise due to the limited space of the PowerPoint slide. That is, the length restriction prevented him from providing a clearer description of some sophisticated IRS questions. The students reportedly also lacked the patience to read through a question longer than 50 Chinese characters (equivalent to 30 or so English words). In addition, Mr. Yen stated that questions involving mathematical manipulation and requiring calculation were inappropriate as the students at various academic levels demanded differing lengths of time to respond to the questions. The classroom observation data indicated that the students were given one minute to answer concept-type questions but 2-3 minutes for those involving calculation, similar to those provided in Hancock (2010). The students were asked in the interview about the response time issue. It appears that copying group members' answers without verifying them was not uncommon among the students, especially when they were rushed to click an answer. Four of the students stated that the use of IRS was not of much benefit to them. The reasons were that they merely followed others to click an answer most of the time due to not quite comprehending the lesson content; the teacher did not always thoroughly explain the questions so that whatever they did not understand initially remained unknown afterward; and the content of the IRS tests was the same as that of the pencil-and-paper tests, so why not simply adopt pencil-and-paper tests. Some students mentioned that most of the IRS questions were associated with computation, rather than concepts. Some other students echoed this by stating that they preferred concept-type questions as they could quickly answer the questions by intuition.

Composition of the IRS questions

According to Mr. Yen, the IRS pre-test questions he posed in class were mainly extracted from questions listed in the textbook exercises or from some test banks provided by textbook publishers. He stated that the post-test questions were different from the pre-test questions, which he said were based on the students' learning responses in class. However, the questions also came from the same sources as those of the pre-test. Mr. Yen was once asked in the interview why he did not just compose questions on his own. He responded that the test banks already provided a variety of questions for him to choose from. In addition to composing questions from test banks, Mr. Yen reported that he also encouraged students to provide test questions (ten each), based on the question list he distributed to the class. He explained that the purpose of asking students to go through the provided questions was to help them actively reflect on what they had learned from the lessons. However, only two out of the 18 students interviewed indicated that they ever participated in the question composition task for the sake of gaining bonus points and challenging classmates' problem-solving ability by coming up with some difficult questions. Mr. Yen stated that the questions submitted by the students unexpectedly informed him of the students' prior knowledge attained in junior high school, which better helped him adjust his lesson content.

Student performance

The IRS test results show that student performance demonstrated in the pre- and post-tests varied considerably. In other words, there was no obvious pattern regarding whether or not the students improved their performance as a result of engaging in the IRS tests. Mr. Yen emphasized that conducting IRS tests was to promote the students' learning interest, track their learning progress, and enhance their conceptual understanding, rather than grade them. Therefore, the test variation was not unexpected to him. He added that occasional hardware failures in some classes also affected the test results. He was, however, rather satisfied with the students' performance achieved in the review test. According to some documents provided by Mr. Yen, the average score of his students' performance was slightly higher than that achieved by the counterpart classes. It is worth mentioning that, according to question content analysis, 20 out of the 40 questions in the review test were identical or highly similar to those in the IRS tests.

Mr. Yen's satisfaction decreased though when he learned of the test results his students achieved in the comprehensive test. Table 1 lists the ratios of the correct answers of the review and comprehensive test questions as well as the difficulty level of each question with a χ^2 test. It is noted that only ten of the questions which had highly similar content in both tests were compared. That is, the expanded questions and those with multiple-key items in the comprehensive test, as described in the Methodology section, are excluded. Table 1 reveals that the average ratios of correct answers in the review test and the comprehensive test are 71.6% and 49.3%, respectively. It is noted that average scores, rather than individual classes' scores, were used to conduct the comparison between the two tests as the ratios of correct answers of the two tests presented by each of the three involved classes were highly similar. The χ^2 test results indicate that seven out of the ten questions (excluding questions #4, #10, and #11) in the comprehensive test were significantly more difficult than those in the review test at the level of $p < 0.001$.

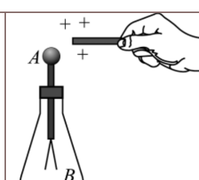
Table 1
The correct ratios in the review and the comprehensive tests

Review Q#[a]	1	3	9	10	28	29	35	37	38	39	Avg.
Correct %	77.3%	61.7%	82.0%	46.9%	82.8%	83.6%	44.5%	71.9%	97.7%	68.0%	71.6%
Comp. Q#[b]	1	2	3	4	9	10	11	17	18	19	Avg.
Correct %	51.6%	21.9%	26.6%	58.6%	51.6%	82.0%	84.4%	23.4%	68.8%	24.2%	49.3%
χ^2 test	48.5[c]	86.0[c]	26.7[c]	7.06	87.8[c]	0.23	82.3	149[c]	468[c]	113[c]	

Note. [a] The question number of the review test. [b] The corresponding question number in the comprehensive test. [c] The question in the comprehensive test is significantly more difficult than in the review test at $p < .001$.

To illustrate the different difficulty levels of the two tests, two examples of test questions are listed in Figures 1a and 1b and Figures 2a and 2b. As shown in Figures 1a and 1b, although 82% of the students correctly answered question #9 in the review test that electrostatic induction causes the gold leaves to spread apart (what happens), only 26.6% of the students chose the correct answer in the corresponding comprehensive test (question #3), indicating that the majority of the students did not grasp the electric charge distribution (why it happens) thoroughly. It is postulated that many students misconceived that electrostatic induction creates net opposite charges on the object due to nearby charges (option A, 39.1%). Similarly, in Figures 2a and 2b, although 68% of the students correctly selected frequency (option C) as the invariant quantity of sound wave embedded in wind blow in question #39 in the review test, less than a quarter (24.2%) chose the correct answer (option A) in the corresponding comprehensive test (question #19) via reasoning the three related quantities of sound wave (wave speed, wavelength, and frequency). In sum, the review test tended to assess the students' skills in recalling the phenomena (what happened), rather than understanding the related principles (why it happened). It is, however, noted that the percentages in Figures 1b and 2b do not add up to 100% due to not all students responding to the questions.

9. (E) A person shifts a rod with positive charge towards a neutral electroscope, as shown in the right figure. What will happen to the electroscope?
   (A) it will carry positive charges
   (B) it will remain unchanged
   (C) it will accumulate positive charges in part A
   (D) it will accumulate negative charges in part B
   (E) the gold leaves of part B will spread apart (82%*)



* The correct ratio

Figure 1a. The correct ratio of question #9 in the review test.

3. (C) A person shifts a rod with positive charge towards a neutral electroscope, as shown in the right figure. What will happen to the electroscope?
   (A) the electroscope carries negative charges, and the gold leaves of part B spread apart (39.1%)
   (B) the electroscope carries positive charges, and the gold leaves of part B spread apart (26.6%)
   (C) the electroscope remains neutral (with equal positive and negative charges), and the gold leaves of part B spread apart (26.6%)
   (D) the electroscope remains neutral (with equal positive and negative charges), and the gold leaves of part B move closer together (6.3%)

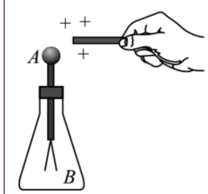


Figure 1b. The corresponding question (question #3) in the comprehensive test.

39. (C) A source transmits sound of frequency 1000 Hz and speed of 344 m/sec from the east to the west. The wind is blowing at a speed of 20 m/sec from the east towards the west. There are two people, A and B, sitting 100m to the east and west of the source, respectively. Which of the following quantities is the same for A and B?
   (A) wave speed
   (B) wavelength
   (C) frequency (68%*)
   (D) loudness
   (E) wave speed and wavelength

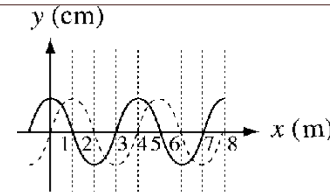
* The correct ratio

Figure 2a. The correct ratio of question #39 in the review test.

19. (A) A man sits at the east of a sound source, while the wind blows from the west to the east. How will the man observe change regarding the quantities of the sound?
   (A) wave speed increased, wavelength increased, frequency unchanged (24.2%)
   (B) wave speed unchanged, wavelength increased, frequency increased (6.3%)
   (C) wave speed increased, wavelength unchanged, frequency increased (25%)
   (D) wave speed unchanged, wavelength decreased, frequency increased (11.7%)
   (E) wave speed unchanged, wavelength unchanged, frequency unchanged (32%)

Figure 2b. The corresponding question (question #19) in the comprehensive test.

Among the ten compared questions, there is only one for which the students achieved a significantly higher average correct ratio (#11, 84.4%) at $p<.001$ in the comprehensive test than for the corresponding question (#35, 44.5%) in the review test (see Figures 3a and 3b for more details). The possible explanation is that question #35 was somewhat beyond the 10th graders' average ability to comprehend. Analysis of the question content discloses that this particular question required the students to identify three variables prior to being able to address the question correctly: (1) acquiring the wavelength from the graph first, (2) analysing multiple possibilities of the frequency from the graph, and then (3) adopting the formula "speed = wavelength x frequency" to obtain speed. However, when the students were tested one variable at a time, such as the wavelength, as demonstrated in question #11 in the comprehensive test, their performance increased significantly. That is, question #35 appeared difficult to the students mainly due to the complexity of the mathematics, rather than to a conceptual barrier.

35. (E) A transverse wave is travelling to the right, as shown in the figure. The rigid curve displays the wave at t=0, and the dotted curve indicates the wave at t = 2.0 sec. If the period (T) of the wave is given within the region of 1.5sec< T <2.5sec, determine the speed of the travelling wave. (A)34 m/second (B)20 m/second (C)0.5 m/second (D)2.0 m/second (E)2.5 m/second (44.5%*)	

* The correct ratio

Figure 3a. The correct ratio of question #35 in the review test.

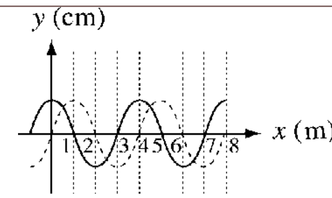
11. (C) A transverse wave is travelling to the right, as shown in the figure. The rigid curve displays the wave at t=0, and the speed of the wave is 20 m/sec. What is the wavelength of the transverse wave? (A) 2 m (9.4%) (B) 3m (1.6%) (C) 4m (84.4%) (D) 6m (0.8%) (E) 8m (3.9%)	

Figure 3b. The corresponding question (question #11) in the comprehensive test.

The interviewed students were also asked about their opinions of the review and comprehensive tests. Most of them said that the two test questions were different, though they appeared similar in the context. However, the majority of the students responded rather positively to the comprehensive test. They stated that the comprehensive test questions required more thinking to clarify the meaning of the core concept, might need to integrate multiple concepts prior to answering the questions, and were better able to tackle their "learning blind spots". A few students, however, commented that, based on conventional learning materials, the comprehensive test was too challenging for them.

The impact of collaboration between the teacher and the physics education expert

Prof. Pang had been carrying out research in physics education for 14 years. Helping students understand physics concepts more constructively and thoroughly had been her research focus for the past decade. Inviting her to develop the comprehensive test was initially meant to help the teacher assess the students' understanding level of fundamental physics concepts from an outsider's perspective. However, her collaboration with Mr. Yen had some unexpected impacts on the professional development of both parties. In the end-of-semester interview Mr. Yen stated that he benefited greatly from discussing and exchanging thoughts with Prof. Pang in the question developing process. He said that the students' relatively poor performance achieved in the comprehensive test helped him realize that they did not thoroughly acquire the knowledge he taught in class. This collaborative opportunity allowed him to extend his insights of knowledge regarding the students' learning difficulties and also the need to make his instruction clearer. He acknowledged that Prof. Pang's questions, composed with adept wording and from various perspectives, were better able to identify the students' misconceptions and learning pitfalls, underpinning Etkina's (2010) claim regarding the need for teachers' PCK development through coaching from university experts. Prof. Pang, on the other hand, also mentioned in email correspondence that the collaboration provided her with an opportunity to notice the reality of school science teaching, including the instructional challenges encountered by the teacher and the learning bottlenecks experienced by the students. To cope with the learning bottleneck issue, she stated that trivial concerns, such as selecting adequate words and terms intelligible to pupils in the test questions, mattered greatly. One example she illustrated was modifying "electromotive force" to "electric potential difference" to help the students understand the meaning of the targeted term. In addition, she emphasized that assuring that test content fell within the designed curriculum, and identifying commonly overlooked "critical details" were equivalently crucial to teachers' teaching and students' learning of physics, consistent with some earlier researchers' assertions (e.g. Viennot et al., 2005; Lee et al., 2011).

Conclusions

The teacher in this study was an experienced, highly self-motivated teacher and was also proficient in integrating technology into teaching. Through the use of the IRS, he effectively increased the students' participation and class interactivity. The student performance demonstrated in the review test initially appeared to have satisfied Mr. Yen's expectations. However, when the students were reassessed with a newly designed test, the comprehensive test, using different perspectives from the review test questions, their performance dropped significantly. Such a discrepancy seems to reflect that the practice effect of the IRS tests only helped the students attain a higher score but not to fully acquire fundamental understanding of the concepts taught. In other words, many of the students only attained knowledge about what happened (the phenomenon), rather than why it happened (the reasoning and principles) and the causal relations among the variables. Therefore, when questions demanded causal reasoning skills, rather than performing factual recalling skills, their performance suffered.

Beatty et al. (2006a) argued that designing IRS questions is different from composing exam and homework problems. The teacher in the present study, however, made up IRS questions merely by extracting questions provided in some test banks. Many test bank (traditional) questions were found to be unclear in their content description or were too complicated in their mathematical procedures for the 10th graders, like the ones illustrated in Figures 3a and 3b. That is, those questions did not aim at helping students acquire conceptual understanding or analytical skills; rather, they tended to emphasize mathematical operations. As indicated by Beatty et al. (2006a), designing effective IRS questions is a rather challenging task; the design of the questions must be based on the predetermined instructional goals. When using the four pedagogical purposes of adopting a question cycle suggested by Beatty et al. (2006b) to examine the current IRS implementation, it is found that the teacher only achieved one of the purposes; that is, increasing class interactivity and gaining the students' concentration. The other three purposes (enhancing the students' conceptual comprehension, using formative assessment to identify the students' learning pitfalls, and inspiring the students' meta-cognition of knowing what and how to learn physics) were achieved to a lesser degree. That is, Mr. Yen did not seem to have captured the principles of designing IRS questions. As a result, his use of the IRS did not actually help him fulfil the in-class formative assessment as he once claimed; nor did it help the students acquire physics concepts thoroughly.

It is quite common in Taiwan for teachers to adopt test bank questions available from textbook publishers for school test questions. Other than for convenience, this study reveals that it may reflect the teachers' incapability of developing quality questions on their own, no matter whether developing IRS-based or pencil-and-paper questions. Mr. Yen's reliance on outside sources for composing test questions was presumably associated with his lack of content knowledge. His proficiency in integrating technology into instruction did not assist him in developing quality IRS questions. As some researchers have contended, to successfully integrate technology into instruction, it must be coherent with pedagogy and content knowledge (Koehler & Mishra, 2008; Mishra & Koehler, 2006). The collaboration with the physics education expert appeared to provide him with an opportunity to enhance his content and pedagogical knowledge, in accordance with some previous research assertions that on-going professional development through partnering with university experts in the physics education domain is both feasible and effective (Viennot et al., 2005; Etkina, 2010). The communication process benefits not only the teachers in enriching their didactic knowledge, but also the experts in grasping high school students' knowledge background.

Suggestions

To cope with the reported issues of implementing IRSs, such as question content design, question type preferences, response time variations, and answer reproduction, the following tactics are suggested.

1. First of all, the overall goal of establishing TEAL is to reinforce students' understanding of physics concepts, and the use of IRSs is meant to help fulfil this goal, especially in tracking and assessing students' conceptual understanding formatively. Lin et al. (2011) also found that the interactive response system is not appropriate for students' responses to calculation questions. Therefore, the design of IRS questions must be geared toward enhancing students' ability of conceptual comprehension and logical reasoning, rather than testing their mathematical ability in a complicated manner. A well-designed, concept-oriented question can not only help clarify students' physics concepts, but can also engage students in higher-order thinking. For 10^{th} graders, the description of each question must not be too lengthy, approximately 50 Chinese characters, to capture their full attention. Provision of pictorial tools, such as photos, diagrams, or graphs, is also highly recommended to facilitate students in figuring out the question context. It is anticipated that the response time needed to answer concept-type questions will be more consistent than that of computation questions among different backgrounds of students. It also differentiates the purpose of IRS-based questions from pencil-and-paper questions.
2. Baird et al. (1991) stated that factual recall questions may decrease the students' meta-cognitive commitment and discourage their meaningful learning strategies. Similarly, repetition of the questions will only encourage rote learning and, in the meantime, reduce students' cognitive demand while resolving questions. To broaden the students' meta-cognitive beliefs about what and how to learn physics, the designed questions should involve not only sensible phenomena but also the principles that can be reasoned and explained in the phenomena, such as those illustrated in Figure 1b and Figure 2b. That is, physics knowledge involves not only factual recall and intuitive observation, but also logical reasoning and conceptual clarification.
3. To reinforce students' understanding of the key concept in a timely fashion, it is suggested that IRS questions be composed as a set, and that the cycle be iterated three or four times in a lesson, as proposed by Beatty and Gerace (2009) and Lee et al. (2011), rather than administering pre- and post-tests before and after a chapter. Conducting IRS tests as the lesson progresses also helps the teacher better assess the students' learning progress formatively.
4. The students in the study conducted by Shieh et al. (2011) reported that a question with a high correct ratio did not necessarily mean that all students comprehended the concept of the question raised. Baird et al. (1991) asserted that instant feedback helps students enhance their responsibility and beliefs of learning. It is thus suggested that the teacher review each IRS question immediately after the correct answer is displayed, with a brief explanation provided to questions with a high correct ratio, but more detailed explanation of those with a low correct ratio.
5. Henriksen and Angell (2010) found that students who had comprehended the physics concept taught should be able to verbally explain it, and vice versa. To prevent the students from simply following others in clicking an answer, and instead to foster their active thinking habits, it is suggested that the teacher randomly call on a few students to restate the key concept of the question discussed. Such a restating process can help the teacher ensure the students' reasoning skills and the level of understanding of the discussed question.
6. It is also suggested that the design of IRS questions be supported by teaching details. As asserted by Beatty et al. (2006a), teaching details must meet students' learning demands. Coverage of the details includes related variables, physics principles, and causality among the variables.

It is emphasized that provision of suggestions and guidance to teachers to design quality IRS questions is far from sufficient. It is pivotal that assistance, such as scaffolding or collaboration, be provided to the teachers as well, particularly at the beginning stage of the design process. Only when such a collaborative model is established between the two parties can it help the teacher develop IRS questions to effectively assist students to acquire actual understanding of the materials taught and also gain competitive scores in various types of administered tests.

This study investigated how one high-school teacher implemented an IRS and what benefits and challenges he and his students encountered. Although only one school site was examined, meaning that the results may not be generalized to other contexts, the study was, nonetheless, conducted based on three perspectives: the teacher, the students, and an expert outsider. In addition to triangulating multiple sources of data to depict the participants' learning/teaching experiences, the study further reveals that enhancing the teacher's content knowledge and capability of identifying the students' learning pitfalls is the key to developing quality IRS questions. Researchers in the associated fields, however, bear a high responsibility for building such a bridge to connect school teachers and university experts to fulfil the goal collaboratively. It is hoped that the overall findings of this study can provide useful insights to those who are interested in strengthening students' physics concepts through the use of IRSs.

Acknowledgments

The research was sponsored by the National Science Council, NSC 99-2511-S-424 -003 -MY2.

References

Baird, J. R., Fensham, P. J., Gunstone, R. F., & White, R. T. (1991). The importance of reflection in improving science teaching and learning. *Journal of Research in Science Teaching 28*(2), 163-182.

Bartsch, R. A., & Murphy, W. (2011). Examining the effects of an electronic classroom response system on student engagement and performance. *Journal of Educational Computing Research, 44*(1), 25-33.

Beatty, I. D., & Gerace, W. J. (2009). Technology-enhanced formative assessment: A research-based pedagogy for teaching science with classroom response technology. *Journal of Science Education and Technology, 18*,146-162.

Beatty, I. D., Gerace, W. J., Leonard, W. J., & Dufresne, R. J. (2006a). Designing effective questions for classroom response system teaching. *American Journal of Physics 74*(1), 31-39.

Beatty, I. D., Leonard, W. J., Gerace, W. J., & Dufresne, R. J. (2006b). Question driven instruction: Teaching science (well) with an audience response system. In D. A. Banks (Ed.), *Audience Response Systems in Higher Education: Applications and Cases* (pp. 96-115). Hershey, PA: Idea Group.

Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education 85*(5), 536-553.

Buckenmeyer, J., & Freitas, D. (2006). Is technology an effective tool to improve schools? Two Views. *Educational technology, 46,* 54-56.

Committee on Biology Teacher Inservice Programs, National Research Council. (1996). The role of scientists in the professional development of science teachers. Retrieved http://www.nap.edu/catalog.php?record_id=2310

Connor, E. (2011). Using cases and clickers in library instruction: Designed for science undergraduates. *Science &Technology Libraries, 30*, 244-253.

Creese, J. (2011). Self-and cohort-directed design in research training tutorials for undergraduate researchers: increasing ownership and relevance to improve learning outcomes. *The Journal Academic Librarianship, 37*(4), 327-332.

Dori, Y. J., & Belcher, J. (2005). How does technology-enabled active learning affect undergraduate students' understanding of electromagnetism concepts? *The Journal of the Learning Sciences, 14*(2), 243-279.

Draper, S. W., & Brown, M. I. (2004). Increasing interactivity in lectures using an electronic voting system. *Journal of Computer Assisted Learning, 20*, 81-94.

Dufresne, R. J., & Gerace, W. J. (2004). Assessing-to-learn: Formative assessment in physics instruction. *The Physics Teacher, 42*(7), 428-433.

Etkina, E. (2010). Pedagogical content knowledge and preparation of high school physics teachers. *Physical Review Special Topics: Physics Education Research, 6*(2)*,* 020110.

Hancock, T. M. (2010). Use of audience response systems for summative assessment in large classes. *Australasian Journal of Educational Technology, 26*(2), 226-237.

Henriksen, E. K., & Angell, C. (2010). The role of 'talking physics' in an undergraduate physics class using an electronic audience response system. *Physics Education, 45*(3), 278-284.

Kay, R. H., & LeSage, A. (2009). A strategic assessment of audience response systems used in higher education. *Australasian Journal of Educational Technology, 25*(2), 235-249. Retrieved from http://www.ascilite.org.au/ajet/ajet25/kay.html

Koehler, M. J., & Mishra, P. (2008). Introducing TPCK. In AACTE Committee on Innovation and Technology (Eds.)*, Handbook of technological pedagogical content knowledge (TPCK) for educators*, (pp. 3-30). New York: Routledge.

Laxman, K. (2011). A study on the adoption of clickers in higher education. *Australasian Journal Educational Technology*, *27*(Special Issue, 8), 1291-1303.

Lee, A., Ding, L., Reay, N. W., & Bao, L. (2011). Single-concept clicker question sequences, *The Physics Teacher,* 49*,* 385-389.

Lijnse, P. L. *(*1995). "Developmental research" as a way to an empirically based "didactical structure" of science. *Science Education, 79*(2), 189-199.

Lin, Y. C., Liu, T. C., & Chu, C. C. (2011). Implementing clickers to assist learning in science lectures: The clicker-assisted conceptual change model. *Australasian Journal Educational Technology*, *27*(6), 979-996.

Magnusson, S., Krajcik, J., & Borko, H. (1999). Nature, sources and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome & N. G. Lederman (Eds.), *Examining pedagogical content knowledge: The construct and its implications for science education* (pp. 95–133). Dordrecht: Kluwer Academic Publishers.

Maloney, D., O'Kuma, T., Hieggelke, C., & Van Heuvelen, A. (2001). Surveying students' conceptual knowledge of electricity and magnetism. *Physics Education Research, American Journal of Physics Supplement, 69*(7), S12-S23.

Mazur, E. (1997). Peer instruction: getting students to think in class. *AIP Conference Proceedings*, *399*(2), 981-988.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record, 108*(6), 1017-1054.

Patton, M. Q. (2002). *Qualitative research & evaluation methods*. Thousand Oaks, CA: Sage Publications.

Pint´o, R. (2005). Introducing curriculum innovations in science: Identifying teachers' transformations and the design of related teacher education. *Science Education, 89,* 1–12.

Shieh, R. S., Chang, W., & Liu, Z. F. (2011). Technology enabled active learning (TEAL) in introductory Physics: Impact on genders and achievement levels. *Australasian Journal of Educational Technology 27*(7), 1082-1099. Retrieved from http://www.ascilite.org.au/ajet/ajet27/shieh.html

Shieh, R. S. (2012). The impact of technology-enabled active learning (TEAL) implementation on student learning and teachers' teaching in a high school context. *Computers & Education, 59*, 206–214.

Shulman, L. S. (1996). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Viennot, L., Chauvet, F., Colin, P., & Rebmann, G. (2005). Designing strategies and tools for teacher training: The role of critical details, examples in optics. *Science Education, 89,* 13-27.

Wittmann, M. C., Steinberg, R. N., & Redish, E. F. (1999). Making sense of how students make sense of waves, *The Physics Teacher, 37*, 15–21.

Yeh, C. R., & Tao, Y. H. (2012). College students' intention to continue using a personal response system: Deriving a model from four theoretical perspectives. *Australasian Journal Educational Technology*, *28*(5), 912-930.

Yourstone, S. A., Kraye, H. S., & Albaum, G. (2008). Classroom questioning with immediate electronic response: Do clickers improve learning? *Decision Sciences Journal of Innovative Education, 6*(1), 75-88.

Corresponding author: Ruey S. Shieh, rueys99@gmail.com

Please cite as: Shieh, R. S., & Chang, W. (2013). Implementing the interactive response system in a high school physics context: Intervention and reflections. *Australasian Journal of Educational Technology, 29*(5), 748-761.