

Low-cost Audience Polling Using Computer Vision

Andrew Cross, Edward Cutrell, and William Thies

Microsoft Research India

{t-across, cutrell, thies}@microsoft.com

ABSTRACT

Electronic response systems known as “clickers” have demonstrated educational benefits in well-resourced classrooms, but remain out-of-reach for most schools due to their prohibitive cost. We propose a new, low-cost technique that utilizes computer vision for real-time polling of a classroom. Our approach allows teachers to ask a multiple-choice question. Students respond by holding up a *qCard*: a sheet of paper that contains a printed code, similar to a QR code, encoding their student IDs. Students indicate their answers (A, B, C or D) by holding the card in one of four orientations. Using a laptop and an off-the-shelf webcam, our software automatically recognizes and aggregates the students’ responses and displays them to the teacher. We built this system and performed initial trials in secondary schools in Bangalore, India. In a 25-student classroom, our system offers 99.8% recognition accuracy, captures 97% of responses within 10 seconds, and costs 15 times less than existing electronic solutions.

Author Keywords

Education; audience polling; low-cost; electronic response system; clickers; ICT4D

ACM Classification Keywords

H.4.1 [Information Interfaces and Presentation]: Group and Organization Interfaces; K.3.1 [Computers and Education]: Computer Uses in Education

INTRODUCTION

Contemporary research has shown that there is educational value in promoting interactive and collaborative learning in the classroom [2,8,11]. In addition to keeping students engaged with the material, interactive classroom exercises enable teachers to monitor the interests and understanding of students and to adjust the content and pacing of their lectures accordingly. Collaborative learning between peers can offer additional benefits, as students receive (and give) personalized attention and can discuss pertinent topics using their own language and metaphors.

A common requirement in many interactive and collaborative lessons is the ability to *poll* students regarding their views or comprehension of a subject. While there is no

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

UIST '12, October 7–10, 2012, Cambridge, Massachusetts, USA.

Copyright 2012 ACM 978-1-4503-1580-7/12/10...\$15.00.



Figure 1: Students respond to a multiple-choice question using *qCards*. Each card has a unique code that encodes the student’s ID. The rotation of the card encodes the student’s answer. Codes are recognized by a webcam (mounted at front) and aggregated for display on the teacher’s computer.

shortage of everyday polling techniques, each has its limitations. A show of hands is subject to peer pressure, where students who are shy or unsure may align themselves with popular opinion. Conversely, asking for feedback from one student at a time – either volunteered, or “cold-called” – does not paint a complete picture of the class. While a written test achieves the goals of accuracy and completeness, it is typically too slow to be administered and checked within the span of an interactive lesson.

To overcome these challenges, many classrooms have embraced *electronic response systems*, in which networked devices called *clickers* are distributed to students, allowing them to submit answers to multiple-choice questions (or occasionally, to submit richer data). Student responses are automatically aggregated and displayed to the teacher in real-time. Clickers have been commercially successful; the *i>*clicker company alone claims usage by 2 million students in over 1,000 institutions [17]. In terms of pedagogical benefits, reports are mixed, with reactions ranging from extremely positive to, at worst, distracting. However, in general, teachers trained to respond to their audience and prepared to alter a lesson based on feedback can use clickers to improve teaching where the worst result is no change in learning outcomes [1,4,7,15]. The consensus is that, when paired with appropriate pedagogy, clickers offer strong benefits by increasing interactivity, enjoyment, participation, attentiveness, and learning outcomes [1,4,7].

Despite their benefits, clickers remain out of reach for the vast majority of educational institutions due to their high cost. For example, the *i>clicker* handsets mentioned previously cost about \$30 each, plus \$200 for a central receiver. Thus, even if a classroom already has a computer, an additional \$950 is needed to equip 25 students with clickers. While systems such as Poll Everywhere aim to reduce device costs by leveraging students' cell phones as handsets, these systems face other challenges in the classroom. Cell phones distract students from the lesson, and are often banned from school for this reason. Also, they would impose usage fees (SMS or data connections) on students, and would disadvantage those who do not own appropriate phones.

In this paper, we propose a new approach to classroom polling that maintains the benefits of clickers while drastically reducing costs. Our system enables teachers to ask a multiple-choice question to students and receive their feedback without individual active components or a costly external receiver. Students respond by holding up a *qCard*: a sheet of normal paper that has a printed code, similar to a QR code. The code indicates the student's ID, while the rotation of the card indicates the student's answer (see Figure 1). Using a computer vision algorithm, an inexpensive, off-the-shelf webcam can automatically recognize and aggregate the students' responses for immediate evaluation by the teacher.

We have built this system and performed initial trials in government schools outside of Bangalore, India. Our results show that our system captures 97% of students' votes within the first 10 seconds of a poll, maintains 99.8% recognition accuracy, and costs at least 15 times less than alternative electronic solutions.

RELATED WORK

High-end electronic clickers have been commercialized by many companies, including *i>clicker*, Meridia, IML, Fleetwood, Qwizdom, and eInstruction. There is substantial literature that studies the benefits and drawbacks of electronic clickers as educational tools. As we do not aim to advance this discourse within the current paper, we refer readers to other excellent reviews on the subject [1,5,7,9].

Within the HCI community, there are two main thrusts of research surrounding audience polling. The first is in response to the CHI 2004 Student Design Competition, which challenged participants to design an audience voting system for the Olympic Games. Nineteen solutions were published as Extended Abstracts. Of these, the majority propose that each audience member use an active electronic device, not unlike a clicker. However, three submissions propose using computer vision to recognize printed signs, containing numeric scores, in the audience [6,12,16]. The PHOTOVOTE system also proposes to recognize colored signs, as a simple alternative to numbers [16].

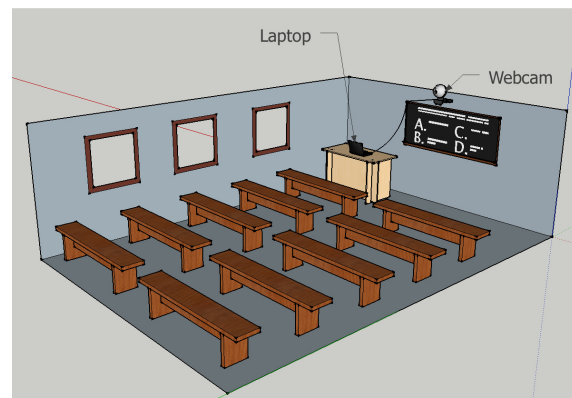


Figure 2: Physical components of our system in a classroom. A webcam is mounted above the chalkboard, tilted downward to see the entire class, and connected to a laptop for processing student responses.

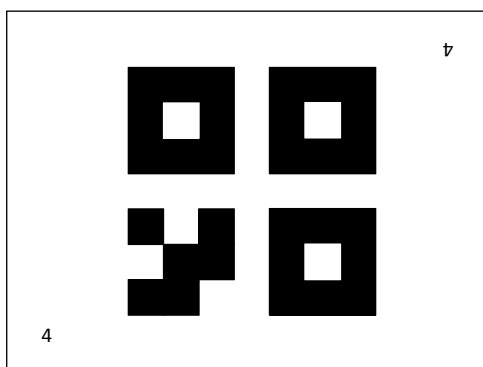
While these solutions have some resemblance to ours, there are three key differences. First, we are unaware of any prior proposal to use a machine encoding (such as QR codes) to represent a user's choice in an audience response system. In addition to encoding more bits than a simple gesture or printed number, a machine encoding preserves the privacy of the respondent relative to peers; his/her choice is undecipherable to everyone except for the computer. Second, prior systems utilizing computer vision encode each respondent's answer, but do not explicitly encode a unique ID number for the respondent. This lack of identity encoding prevents tracking the answers of individuals over time. Finally, we are unaware of any implementation or evaluation of an audience response system based on computer vision in a real-world classroom setting.

The second thread of research is in using computer vision to solicit other kinds of participation and feedback from the audience. Maynes-Aminzade et al. engaged audiences in creative ways, including a pong game controlled by the tilt of audience members' bodies and an audience poll where answers are submitted according to the focus of laser pointers [10]. BallBouncer is another game enabled by computer vision, in which audience members volley a virtual ball from side to side [14]. Cinematrix enables audience polling using two-sided reflective panels [18]. Unlike our system, these polling techniques do not encode the identity of each participant, or preserve participants' anonymity relative to peers.

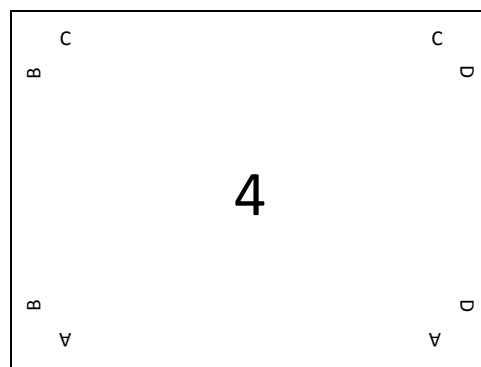
SYSTEM OVERVIEW

Our system can be instantiated in several different forms. The instantiation used for our experiments is illustrated in Figure 2. The fundamental elements of our design are as follows:

1. **A printed qCard** for each student. The qCard (see Figure 3) encodes the student's ID in a unique bit



(a) Front view. The three black squares are used to locate the card and determine its orientation. The student ID is encoded in the fourth quadrant. The human-readable labels are useful for classroom management of the cards, but are not recognized by the algorithm.



(b) Back view. Students rotate the card so their answer is at the top, cover their answer with their fingers, then raise the card with the front side facing the camera. The assignment of answers (A, B, C, D) to edges is randomized in order to further conceal a student's response from peers.

Figure 3: A sample qCard (for student 4).

pattern. On the back of the card, each orientation is labeled with an answer (A, B, C, or D). Students encode a response by rotating the card such that their intended answer is on top, and then showing the card to the camera. qCards are printed on A4 paper using a normal black-and-white printer.

- Camera-enabled computing device.** In our experiments, we utilize a laptop with a USB webcam (Logitech Pro 9000, 2.0 megapixel video). We envision that a mobile phone can substitute for the laptop and webcam in the future. (We have completed initial feasibility tests on phones, but do not report them in this paper.)
- (Optional) LCD projector.** If available, an LCD projector can be used to display the results of each poll to the class. We utilized such a projector in our initial prototyping, but not in our final tests, as the school where we ran our experiment did not have a projector in the classroom. For each question asked by the teacher, the options were written on the board prior to initiating a poll.

To utilize these elements, teachers need to follow the following basic procedure:

- Mount the camera** in an elevated location that can read all qCards simultaneously, without some students obscuring others. For our classroom experiments, we taped a webcam to the wall above the chalkboard (see Figure 2).
- Distribute qCards** to students. If the teacher desires to log each student's responses, then student names need to be mapped to qCard IDs in the computer software.
- (Optional) Calibrate the software** to the lighting conditions of the room. The key parameter is the

threshold for distinguishing black and white elements of the cards. It may help to adjust any irregular lighting conditions, e.g., to close or open window blinds.

- Ask students questions, and see the results.** Each polling period is started and stopped via a keypress on the computer. In order to ensure that students understand the multiple-choice options, we recommend writing each option on the board (or projecting on a slide).

While these steps cover the mechanical operation of the system, additional attention is needed to adjust the pedagogy to benefit from the technology. We briefly discuss such pedagogies near the end of the paper.

DESIGN AND RECOGNITION OF THE QCARDS

The central goal in designing the qCards is to encode information that can be quickly and robustly recognized by a low-end camera. A key parameter of interest is the number of bits encoded per card. As the number of bits increases, it becomes possible to support more students (each with a unique ID). However, the recognition accuracy also decreases, as the cards become more densely coded and thus more difficult to reliably decode. Our initial design encodes 9 bits of information, which would support up to 512 students – far more than we aspire to test in our initial trials. In settings with fewer students, the extra bits can be used for error correction.

Our starting point for the design was a 2-dimensional barcode, such as the QR code. However, traditional 2-dimensional barcode designs are optimized to encode high levels information (at least 231 bits, some used for error correction, in the case of QR codes). The primary use case for industry standard designs assumes the scanner is close to the code in relation to its size, and is targeted at a single

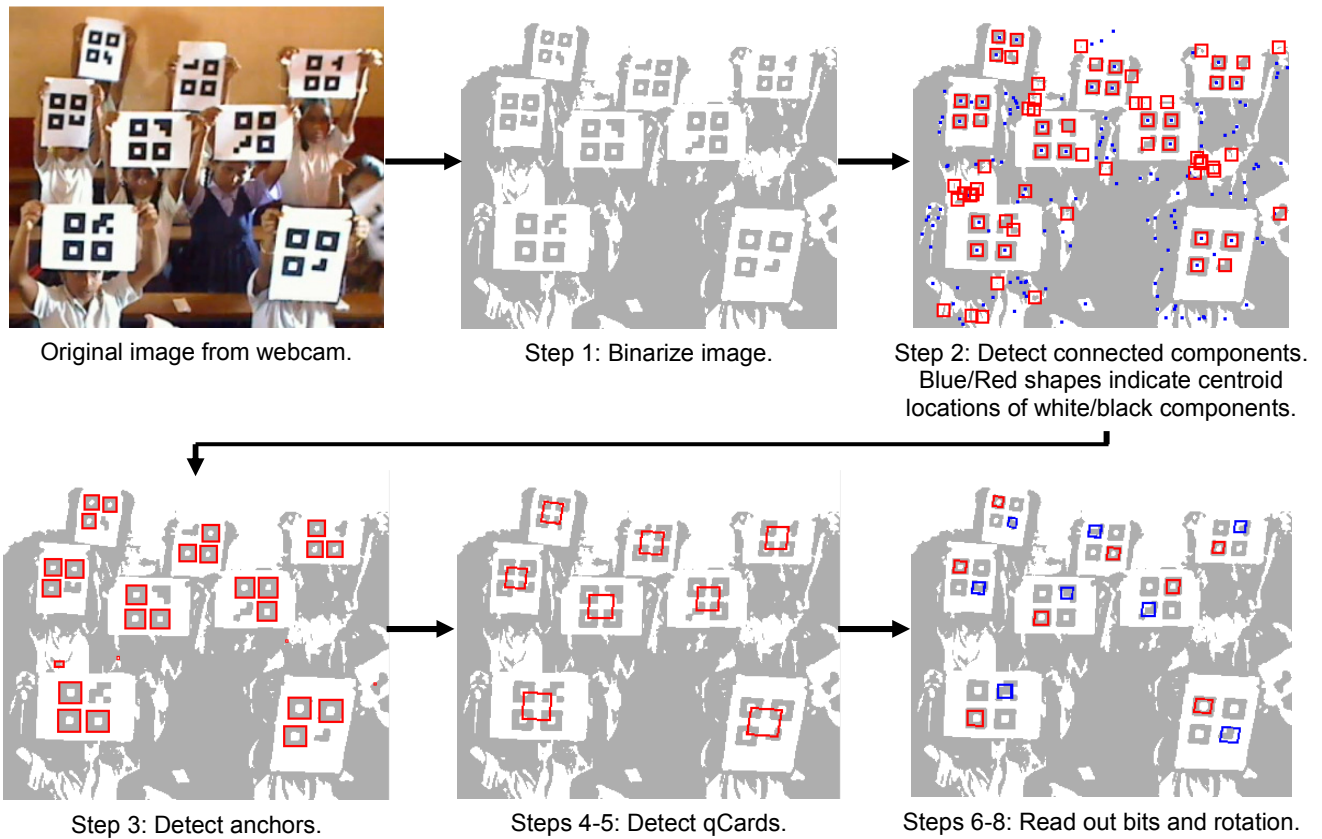


Figure 4: Example execution of qCard recognition algorithm.

code at any given time. Advanced algorithms exist to increase successful reading rates in low-resolution environments, but do not optimize for distance [13].

Thus, we designed a simplified QR code that increases the size of each symbol, thereby facilitating recognition by an ordinary webcam, even at a long distance or under imperfect lighting conditions. To further facilitate recognition accuracy, each code is printed on a full sheet of paper (in our experiments, we used A4 paper size).

Design Specification

The qCard design (see Figure 3) is a 7x7 square with a 1-unit quiet zone around the image. There are four regions of interest. The first three are 3x3 black areas with a 1x1 white center. These three “anchors” are used for locating the code in a frame, and determining orientation and rotation in 3-space. The fourth region consists of the 9-bit code.

Leveraging Video to Improve Accuracy

A distinct difference from other 2-dimensional barcode scanning algorithms is that the algorithm presented below does not require a “perfect” scan on a single photo frame. Instead, the system relies on a majority of reads over the duration of the poll in different frames of a video stream to account for false positives. The algorithm processes each frame individually, and does not require every qCard present to be recognizable in a single frame. It continually

scans images from the instantiation of a “poll” until its close, tabulating how many of each code it has read. If it reads more than one response for a given student, the code it read the most times is deemed the final answer.

Recognition Algorithm

Apart from its use of video to improve accuracy, our recognition algorithm follows the general principles of other barcode scanning algorithms in isolating regions of importance based on high-contrast locators in the image.

An example operation of our algorithm is illustrated in Figure 4. Each frame is processed in the following way:

1. Binarize the image according to a predetermined brightness which can be calibrated for the entire class or by region in difficult environments.
2. Scan the binarized image for connected components of black or white pixels. At the same time, compute the centroid of each connected component based on the number of pixels in the object, and their relative positions. This is the most computationally-intensive step of the algorithm; it is critical to use an incremental algorithm (linear in the number of pixels in the image) to compute the connected components and their centroids.
3. Pair black and white components whose centroids are closely aligned, pairing components separated by no

more than a given threshold. This step isolates the anchors – black squares with white centers – which in groups of three define the position and orientation of a card. Each anchor is assigned a “weight” determined by the number of pixels in the connected components of that anchor.

4. Group sets of three anchors based on location and weight to determine where a qCard likely exists. This grouping uses a nearest-neighbors algorithm.
5. Calculate the orientation and rotation in 3-space of the card from the distance and angle between the three anchors. The rotation encodes the student’s answer.
6. Calculate the location of the center of the fourth quadrant where the student ID is encoded.
7. Given the angles in 3-space at which the card is being held, calculate the locations of the centers of the 9 bits where the information is encoded.
8. Read the binary bits given the contrast of the region.
9. Combine the encoded bits and the card rotation to register a response for a given student.

IMPLEMENTATION

We implemented our system in C#. It consists of three parts: an image processing library that implements the algorithm detailed above; a GUI wrapper that maps found codes and orientations to students and answer choices, presents calibration options, and visually confirms card reads; and a webcam driver to pull in live images for processing. The image processing library consists of 1500 lines of non-comment, non-blank lines of C# code. The GUI wrapper consists of another 1600 lines of code, and the webcam drivers consist of another 250 lines of code.

The runtime performance of our system depends on the scene, lighting environment, and number of cards present. For the experiment environment, in a sunlit classroom with 25 students, the algorithm processes at about 12fps. For more varied environments, including multiple light sources, shadows, “border” threshold regions with many small connected components, or very large numbers of cards, our implementation can run as slow as 2fps.

The cost of the setup is the price of a laptop/netbook and the Logitech 9000, 2MP camera (available on Amazon.com for \$61). By way of comparison, it would cost about \$950 to provide i>clicker units for our classroom of 25 students. Thus, assuming that a computer is already available to utilize both systems, our approach represents about a 15-fold cost savings in this case. As the number of students increases, our system becomes even more economical, because our marginal cost is simply the cost of printed cards (rather than a \$30 device per student).



Figure 5: A simulated 100-person classroom. Our algorithm recognizes all 100 qCards in under 5 seconds from a webcam video stream, or in 85ms from this high-quality still image.

INITIAL EXPERIMENTS

We conducted three stages of early experiments culminating in a test of our system within a real classroom environment. Our goals were to assess the technical feasibility of accurately recognizing student responses as well as to gauge the usability and overall experience of students.

Stage 1: Laboratory Environment

For this stage of testing, we wanted to test the limits of the system design without the behavioral effects of students holding the cards. We set up 100 chairs in an auditorium (see Figure 5) to simulate a 100-person classroom. On each chair, we taped a qCard in alternating rotations within view of the webcam, mounted at the front of the room and looking down as it would in a classroom setting.

Our system correctly recognized all 100 qCards from the webcam video feed; however, this required processing 30 frames over a period of 4.4 seconds. Though the scene was static, processing multiple frames improved the recognition for two reasons. First, the contrast threshold was manually altered over the polling period to capture cards in different lighting conditions (the auditorium had rows of lights at fixed intervals, providing variable lighting for chairs). Second, due to the limited quality and resolution of the webcam, pixels on the border of a given threshold “flickered” from black to white across frames of the binarized image sequence. For larger cards in the front rows, single pixel changes do not affect the algorithm’s pairing of black and white components; however, with smaller cards toward the back, single pixel shifts have a more significant effect. Figure 6 shows the average time to recognize the cards for a given row, where the first four rows average about 250ms (recognized on the first frame) and the last row averages about 1.8 seconds.

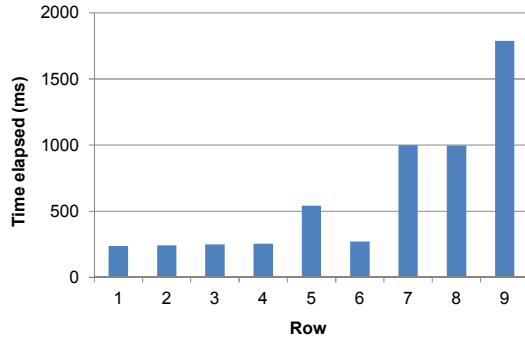


Figure 6: Average time needed to recognize qCards in each row of the simulated 100-student classroom (Figure 5).

The recognition speed can be improved by using a higher-quality camera. To demonstrate this, we took a high-quality still photo using a 14 MP camera (Sony NEX-3). Using this image, the algorithm recognizes all 100 cards in 518ms (a single frame). Recognition is also successful when the high-quality image is resized to 2 megapixels, corresponding to the resolution of the webcam. In this case, recognition requires only 85ms.

Stage 2: After-School Program

Our first field deployment was with a non-profit partner, the Children’s Lovecastles Trust (CLT India), which runs a “computer clubhouse” that hosts after-school enrichment activities for low-income children in peri-urban Bangalore. Students between the ages of 9 and 19 voluntarily come to the clubhouse to explore their own ideas (often utilizing computers) with support from peers and adult mentors. For our experiments, we interacted with students from both private and government schools. They were mostly in the 5th and 6th standard. While they were most comfortable speaking the local language (Kannada), they were also able to interact in English.

The goal of our experiments in the after-school program was to perform a simple litmus test: can students understand the concept of using encoded cards to submit an answer, and are they engaged by interactive activities enabled by the technology?

To engage with students, we created a participatory quiz show game that was themed after “Kaun Banega Crorepati” (the Indian equivalent of “Who Wants to Be A Millionaire?”) We utilized an LCD projector to show multiple-choice questions to a group (see Figure 7). The projector indicated which students have answered a given question in real-time, enabling them to lower their cards (decreasing congestion) once their answers were recognized. After each question, the system displayed the distribution of student responses, the correct response, as well as a leaderboard that listed the names of students with the highest score so far.



Figure 7: Experiments with a projected quiz show in an after-school program.

We conducted five quiz events over the course of three days. Between 12 and 19 students participated in each quiz (with significant overlap between sessions). While we did not attempt to rigorously evaluate these sessions, the feedback was overwhelmingly positive. Despite the low lighting conditions, qCards were recognized consistently and almost instantaneously. Students found the cards easy to use and were extremely engaged in the game. In fact, our most significant lesson learned was that the camera was almost too sensitive: it frequently read a card that was still in a student’s lap, as the student prepared to answer the question. We addressed this in two ways in advance of stage 3. First, we instructed students to shield their card from the camera, except when they intended to answer. Second, we changed the recognition algorithm to a majority vote, so that even if the student revealed an unintended answer, his or her final answer would appear for a longer time and thus would be prioritized by the recognition algorithm.

Stage 3: Government School

Encouraged by our experience in an informal setting, we proceeded to test the system in an actual classroom. With the help of CLT India, we made contact with a Kannada-medium government school and received their permission to use the system to administer and review a test that was pertinent to their curriculum. Our participants were 7th-grade students.

We performed trials on three separate days, in three different classes. Trials lasted for between 45 minutes and

1.5 hours. We limit our presentation to the third trial, as it contained the most refined software and experimental protocol. The first two trials contained similar procedures and results, but evaluated an earlier version of the qCards in which students encoded their answers by choosing one of four different cards, instead of rotating a single card into different orientations.

Study Goals

Our study was designed to answer two distinct questions:

Q1: Technology assessment. Once a student has decided on an answer, how do cards compare with traditional polling methods (clickers and hand-raising) in relaying that answer to the teacher? We are interested in two metrics: accuracy (does the teacher see the answer that the student intends) and speed (how long does it take for students to indicate their answer and have it recognized by the system).

Q2: Behavioral assessment. Does the polling technology affect which answer the student *intends* to submit? For example, with a show of hands, one might expect shy or unsure students to change their answer depending on answers observed from other students. This behavior is inhibited using clickers or cards, since answers are (mostly) private with respect to peers.

For clarity, we also specify a question that the current study does **not** intend to address, and that is whether cards offer pedagogical benefits in an Indian classroom. Though our ultimate goal is to impact educational outcomes, our scope in the current paper is to demonstrate that cards preserve the salient benefits of clickers (which have been demonstrated to impact education) while greatly reducing costs.

Experimental design

To answer the questions posed above, we structured our study as follows. We developed a multiple-choice test (details below), and administered each question as follows:

1. A facilitator wrote the question and multiple-choice answers on the board, and read them aloud.
2. Students indicated their answers using i>clickers.
3. Students indicated their answers again using qCards.
4. Students indicated their answer a third time with a show of hands. That is, the teacher asked students to raise their hand if their answer was A, followed by B, C, and D.

Steps 2 and 3 were interchanged on alternate questions so that clickers and cards were balanced with respect to ordering. Step 4 was always last, because showing hands revealed students' answers and could impact the answers gathered by clickers or cards. At each step, we waited for all students to respond before moving on. Students were told not to submit any answer until the teacher had finished reading the question and all the answer choices. Students were not allowed to talk during the exercise.

This experimental design allowed us to probe Q1 (technology assessment) by comparing the answers reported via the three polling technologies. The accuracy was assessed in two ways. First, we examined the answers to *stock questions*: questions on the exam that simply instructed students to respond with a given answer. These questions enabled us to compare the measured answers with what the student likely intended. Second, we examined the consistency of replies: the answers reported via clickers and cards should be identical for each student. Finally, the speed of the techniques was measured with a simple timer.

To probe Q2 (behavioral assessment), we compared the answers reported with a show of hands to the automated polling techniques. A systematic difference would suggest that students behave differently when everyone reports their answers openly.

Methodology

To obtain a test of appropriate difficulty, we used a test that had been given to a previous class and resulted in an average score of about 60%. It was prepared in English and Kannada, and presented to the class in Kannada. The test contained 5 "real" questions and 16 "stock" questions; the stock questions simply requested students to submit a particular answer, in order to test the voting technologies. The stock questions were inserted between the real questions, each time requesting students to answer A, B, C, and D, in sequence. All (real) questions included four possible answers.

The participants in this experiment were 25 students (13 female), all in the 7th standard. We utilized a medium-sized classroom, typical of government schools in Bangalore and the usual setting for this class. The lighting in the room was 100% natural, with sunlight coming through open windows on both sides. We closed some of the window blinds prior to the experiment to obtain more uniform lighting of the room. The room was organized into two sections, boys on one side and girls on the other side, with an aisle in between.

At the start of the experiment, each student received a clicker and a qCard. The teacher explained how to use the clickers and how to use the cards. Since the clickers supported 5 choices, students were told to ignore the "E" choice since there was no equivalent for the cards.

During the testing phase, responses were collected by a researcher who was monitoring the laptop running the qCard recognition software. Responses for clickers and cards were recorded in real time (and archived for later analysis). The researcher confirmed that all responses were received prior to initiating the next poll with students. In a handful of cases, a response was outstanding for a long period, and the researchers offered assistance to the student in question. Assistance included: replacing the expired battery of four i>clicker units, instructing two students to

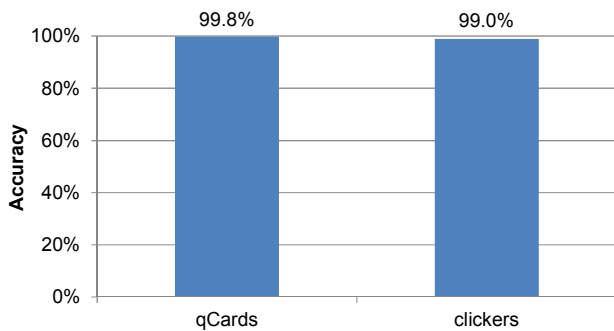


Figure 8: Accuracy of responses read via qCards and clickers for questions with known answers.

raise their qCards higher, and reminding a few students not to obstruct the qCard with their hands. Each poll continued until all students successfully submitted an answer.

Responses to the hand-raising questions were videotaped for later analysis. The responses were not counted in real-time, as we suspected this approach would be too slow to represent a viable model for real-time interaction.

Results: Technology Assessment (Q1)

As outlined previously, we assess the polling techniques according to their accuracy and speed. Accuracy is measured both in terms of correctness relative to known answers and consistency across techniques.

Figure 8 illustrates the absolute accuracy of clickers and cards for questions with “known” answers. For clickers, these correspond to the stock questions only. For cards, the analysis also encompasses the real questions, as we manually decoded students’ answers using the video logs. Clickers received 4 incorrect answers out of 400 data points, yielding an accuracy of 99.0%. All four errors came from the same student; we later verified that the clicker was not malfunctioning, so this likely represents an error on the student’s part. qCards decoded the correct answer on all of the stock questions, but made one error on a real question, leading to an overall accuracy of 99.8% (one error out of 525 cases). Inspection of video logs reveals that the mistake occurred for a student that changed their answer after raising the card. While such changes are typically detected by our algorithm, in this case the student also changed the card’s position such that the new answer was partially obscured by another student’s card. Thus, our algorithm recognized the first answer only.

Figure 9 compares the answers received by cards, clickers, and hands, encompassing 375 responses (25 students answering 5 real questions in 3 different ways). In the vast majority of cases (91.2%), all polling techniques report the same answer from a student. In cases where one technique led to a different answer, the most common outlier was cards (6.4%), followed by hands (0.80%) and clickers (0.67%). In only 0.80% of cases did all three techniques give a different answer. We are not sure why answers reported using qCards were a more frequent outlier

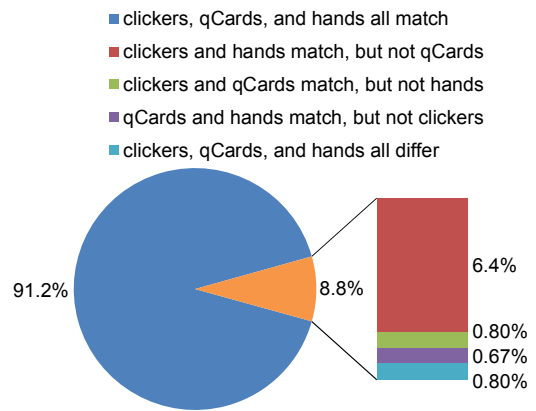


Figure 9: Comparison of responses received by clickers, cards, and a show of hands for 25 students across 5 questions.

compared to the other polling techniques, particularly since answers to stock questions were 100% accurate with cards. Inspection of video logs did not reveal any clear explanations for this trend, e.g., due to copying from other students, or confusion/carelessness with the cards. It is possible that using qCards is a bit more cognitively demanding than clickers or hand-raising, thus leading to these differences. We note that one of the outlying points corresponds to the decoding error for qCards; correcting this, the techniques are in agreement 92.0% of the time.

Given the high correspondence between the techniques, it is not surprising that students’ overall performance on the test is measured similarly with each method. The class average based on responses received by clickers is 31%; the average based on cards is 31%; and the average based on a show of hands is 33%. These scores are much lower than anticipated, and only marginally better than answering randomly. On further investigation, we believe that for these particular students, the material tested may not have been taught yet during this school year. However, in a trial of students during the preceding school year, the average was over 60%, and the correspondence between polling techniques remained high (ranging from 61% to 66%). This suggests that each method would provide similar guidance to the teacher regarding potential customization of lesson plans.

The speed of responses using clickers and cards is illustrated in Figure 10. Our goal was to measure the time taken for students to use the technology, rather than time spent thinking about the question at hand. Toward that end, we present timings for the stock questions only, and time each response relative to the *first* response received for that question. Results show that both clickers and cards are quite fast; after 10 seconds, 97.0% of qCard results are in, while 98.0% of clicker results are in. After 25 seconds, both techniques have read 99% of replies. Responses requiring longer than this represent exceptional situations, e.g., requiring a battery replacement on a clicker. Even then, no response required longer than one minute to read.

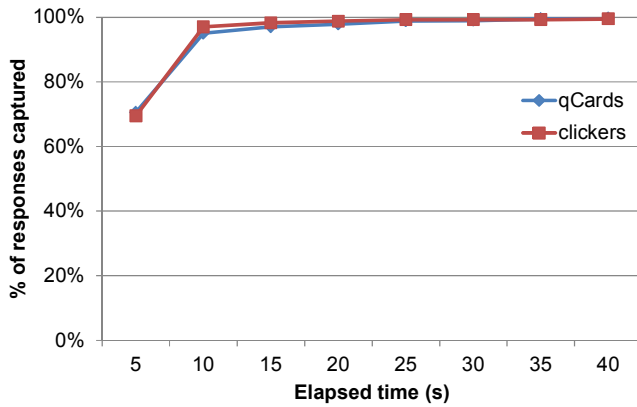


Figure 10: Speed of audience polling technologies: fraction of student responses captured over time.

We also analyzed the speed of hand raising, by analyzing the recorded video feed. We find that an entire hand-raising exercise (spanning all four answers) requires an average of 15 seconds per test question (stdev = 4.6 seconds), though this does not include the time needed to count (or estimate) the number of raised hands for each condition. Thus hand-raising is generally competitive with both clickers and cards as a way to survey the entire class; however, cards and clickers can give a partial estimate in less time.

Results: Behavioral Assessment (Q2)

Our second question of interest is whether students may have *intended* to submit different answers using different polling technologies. Though we conjectured that hand-raising would lead to perturbed results relative to cards and clickers, our data give no evidence for such an effect. The responses from hand-raising agree with responses from clickers 99% of the time, and they agree with qCards 93% of the time. In cases where clickers and cards agreed, the response with hand-raising was different only 1.2% of the time. While these cases could represent a deliberate change in behavior, additional study would be needed to further explicate this effect.

Discussion

We are pleased with the overall performance of qCards. As an initial exploration into this design space, an accuracy of 99.8% (one mis-read in 525 trials) is strong evidence for the feasibility of the system. Whether this accuracy is sufficient depends on the application. If the goal of the system is to give teachers a sense of where their students stand on a question, this accuracy seems well within the bounds of error. However, if the system is used as a graded assessment tool, then additional care may be needed to ensure that every single answer is read as intended.

We note that even the i>clicker system is not perfect; it admitted four errors on stock questions, leading to an accuracy of 99%. This illustrates the limits of accurate classroom polling. Also, i>clickers are subject to battery failure, which could conceivably impact the reports.

Regarding the behavioral assessment, we were surprised with students' consistency (and bravery!) in responding to the show of hands. On several occasions a lone student boldly raised his/her hand and kept it raised, even when it must have become immediately clear that the answer was incorrect. We conjecture that this bold behavior may have been influenced by certain aspects of our study, in particular, by the order of polls conducted. The students were told to vote first by clickers and cards, and then to vote publicly by hands. It is possible the students were more prone to answer honestly feeling the computer had already registered the first responses and that they were locked in. Another possibility is internal to the culture and environment of an Indian primary school classroom. Perhaps students in this environment are less inhibited by fear of standing out than researchers have documented in other primary-school contexts [3,15]. In either case, the influence of peers on student responses merits further research and exploration.

Pedagogies and Usage Scenarios

Due to the high cost of electronic clickers, their use to date has been tightly associated with rich environments. By reducing the cost of electronic response systems, we extend their traditional benefits to low-income environments, and also enable new capabilities that are uniquely tailored to low-income regions.

Traditional benefits of electronic response systems are documented elsewhere [1,4,7]. These include enabling interactive and collaborative pedagogies, as well as enabling teachers to customize lessons to the knowledge and understanding of the class. We believe that techniques to facilitate interactive lessons are especially valuable in developing-country contexts, as current lessons are often delivered by rote with minimal participation of students.

One usage scenario for qCards that may have particular relevance in low-income contexts is that of distance education. Distance-education programs aspire to bring quality teaching from urban centers to rural areas. The physical separation of teachers and students makes it more difficult for remote teachers to assess their students in real-time.

Our project originated with a visit to one such distance-education program, which is run by CLT India. With the help of a large corporation, CLT runs daily distance-education classes, connecting premier teachers in Bangalore with live audiences of students in rural areas. Though there is a high-bandwidth two-way video feed connecting the teacher and the classroom, it remains challenging for the teacher to monitor the progress of each student. The organization perceived that monitoring student progress is even more important in the distance-education scenario, since even the bright and attentive students may fall behind if there is any glitch in the technology (audio, video, volume, etc.) The system proposed in this paper is

especially well-suited to this scenario, because there is already a webcam mounted at the front of the class.

There are also several applications for low-cost audience polling outside of a classroom context. Examples include market research, audience choice awards, conference surveys, and anonymous polls for sensitive topics. We envision that the technology developed in this paper would apply equally well in these scenarios.

CONCLUSIONS

Electronic response systems such as clickers have had a demonstrated impact in making classrooms more interactive. However, to date these benefits have been restricted to wealthy environments, due to the prohibitive costs of clickers.

This paper proposes a replacement for clickers that utilizes computer vision to offer the same benefits at drastically lower cost. By replacing electronic handsets with sheets of printed paper, and replacing an expensive receiver with an ordinary webcam, our solution is about 15 times cheaper than clickers. This innovation makes real-time audience polling (and associated pedagogies) accessible to low-resource schools for the first time. While our study was performed in India, we believe the solution would also be valuable for schools in developed countries, where budget constraints often prohibit costly educational technologies.

Our field trials, though exploratory in nature, confirm the technical feasibility of our approach. Our system recognizes answers with almost 100% accuracy. Moreover, students immediately take to the system, and have no trouble interacting with the technology despite their limited exposure to hi-tech artifacts.

In the future, it will be important to undertake long-term trials to assess the learning outcomes that are enabled via usage of our system. We are also working on porting the system to a mobile phone, which could lower the price even further and simplify distribution in emerging markets.

ACKNOWLEDGEMENTS

We would like to thank Bhagya Rangachar and the entire CLT India team for helping us connect with the school for the experiment as well as giving us access to the after-school program for initial prototyping of the design. A special thanks to Jona, Hemalatha, Nireeksha, and Chethan for creating, translating, and administering the tests. Additionally, we would like to thank Manik Varma, P. Anandan, Samar Singh and especially Andy Wilson for helpful conversations about the project.

REFERENCES

1. Caldwell, J.E. Clickers in the large classroom: current research and best-practice tips. *CBE—Life Sciences Education* 6, 1 (2007), 9-20.
2. Cramer, M., Beauregard, R., and Sharma, M. An investigation of purpose built netbooks for primary school education. *8th International Conference on Interaction Design and Children*, 2009.
3. Crozier, W. The influence of shyness on children's test performance. *British Journal of Educational* 73, 3 (2003), 317-328.
4. Fies, C. and Marshall, J. Classroom response systems: A review of the literature. *Journal of Science Education and Technology* 15, 1 (2006), 101–109.
5. De Gagne, J.C. The impact of clickers in nursing education: A review of literature. *Nurse Education Today* 31, 8 (2011), e34 - e40.
6. Hawkey, K., Kellar, M., MacKay, B., Parker, K., and Reilly, D. From cookies to puppies to athletes: designing a visual audience voting system. *CHI*, (2004).
7. Kay, R.H. and LeSage, A. Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers & Education* 53, 3 (2009), 819-827.
8. Looi, C.-K., Chen, W., and Ng, F.-K. Collaborative activities enabled by GroupScribbles (GS): An exploratory study of learning effectiveness. *Computers & Education* 54, 1 (2010), 14-26.
9. MacArthur, J.R. and Jones, L.L. A review of literature reports of clickers applicable to college chemistry classrooms. *Chemistry Education Research and Practice* 9, 3 (2008), 187.
10. Maynes-Aminzade, D., Pausch, R., and Seitz, S. Techniques for interactive audience participation. *4th International Conference on Multimodal Interfaces*, IEEE Comput. Soc (2002), 15-20.
11. Nussbaum, M., Alvarez, C., Mcfarlane, A., Gomez, F., Claro, S., and Radovic, D. Technology as small group face-to-face Collaborative Scaffolding. *Computers & Education* 52, 1 (2009), 147-153.
12. Schalleck, R., Bober, M., and Drewes, H. Design of an audience voting system for the Olympic games. *CHI EA* (2004).
13. Shams, R. Bar code recognition in highly distorted and low resolution images. *Acoustics, Speech and Signal Processing*, (2007).
14. Sieber, J., McCallum, S., and Wyvill, G. BallBouncer: Interactive Games for Theater Audiences. *Crowd Computer Interaction (Workshop at CHI)*, (2009).
15. Stowell, J.R., Oldham, T., and Bennett, D. Using Student Response Systems ("Clickers") to Combat Conformity and Shyness. *Teaching of Psychology* 37, 2 (2010), 135–140.
16. Unger, P., Forsberg, K., and Jacobsen, J.H. PHOTOVOTE: Olympic judging system. *CHI Extended Abstracts* (2004).
17. i-clicker website. 2012. <http://www.iclicker.com/Products/iclicker/>.
18. Cinematrix website, 2012. <http://usa.cinematrix.info/>.