American Institutes for Research

Academy for Educational Development

Aga Khan Foundation

CARE

Discovery Channel Global Education Fund

Education Development Center

International Reading Association

The Joseph P. Kennedy, Jr. Foundation

Juárez and Associates, Inc.

Michigan State University

Sesame Workshop

Save the Children Federation, USA

University of Pittsburgh

World Education

USAID FROM THE AMERICAN PEOPLE

EQUIP1 Educational Quality Improvement Program Classrooms • Schools • Communities

Pilot of Digital Data Collection Options for Diagnostic Assessments in Honduras

Produced by

Mauricio Estrada and Mariano Solorzano
with Jeffrey Lansdale and Cassandra Jessee
American Institutes for Research

December 2011

U.S. Agency for International Development
Cooperative Agreement No. GDG-A-00-03-00006-00

Pilot of Digital Data Collection Options for Diagnostic Assessments in Honduras

by
Mauricio Estrada and Mariano Solorzano
with Jeffrey Lansdale and Cassandra Jessee

Mejorando el Impacto al Desempeño Estudiantil de Honduras/
American Institutes for Research
(MIDEH/AIR)
Under the Education Quality Improvement Program 1 (EQUIP1)

December 2011

CONTENTS

Introduction 1

Contextual Summary 1

Pilot Rationale 3

Pilot Study of Data Collection Options for Diagnostic Testing 6

Descriptions of the Test Administration Sessions 7

Results of the Pilots 8

Key Findings 8

Conclusions 9

INTRODUCTION

By integrating technology into the student assessment process, the *Mejorando el Impacto al Desempeño Estudiantil en Honduras* (MIDEH—"Improving Student Achievement in Honduras") project has shown in previous pilots that data collectors could reduce the data collection timeframe streamline the entry and analysis process, and report the results much faster.[1]

During 2009, MIDEH conducted a pilot study to compare two data collection technologies, one using portable computers and the other ResponseCards. The purpose of the pilot was to identify cost-effective data collection technologies that could be used to administer and process the results of diagnostic, formative, and end-of-grade assessments. Researchers determined that ResponseCards provided the best technology option for data collection of end-of-grade assessments in the Honduran context. With the support of the Education Quality Improvement Program (EQUIP1), MIDEH purchased another ResponseCard kit to continue with the next phase of the pilot study, using this technology to collect data during diagnostic testing. During the implementation of the pilot study using the ResponseCard devices, MIDEH personnel decided to explore options using other technology, such as Microsoft (MS) Excel. In this report, the use of ResponseCards for the administration and processing of diagnostic tests are compared to the effectiveness of a worksheet designed in MS Excel by MIDEH personnel for processing the results of the diagnostic tests.

CONTEXTUAL SUMMARY

In Honduras, the American Institutes for Research (AIR), with the Honduran Ministry of Education and other partners, has implemented a series of projects called *Mejorando el Impacto al Desempeño Estudiantil en Honduras* (MIDEH—"Honduras Improving Student Achievement Project"), funded by the U.S. Agency for International Development (USAID) as part of the Honduras Education for All–Fast Track Initiative (EFA-FTI). The projects were designed to build a national standards-based reform program, and since 2004, have been able to create a sustainable educational reform structure. The nationwide approach was based on the development of internationally comparable academic standards, teacher pacing guides for implementing the standards, standardized student assessments (diagnostic, formative, and summative), and in-service training programs.

In 2002, the Honduran Ministry of Education launched a new curriculum for kindergarten (K) through Grade 9, called the *Diseño Curricular Nacional Básico* (DCNB). To support the implementation of the new curriculum in the classroom throughout the country, MIDEH, in close coordination with the Ministry of Education and education leaders from every department of the country, developed a national testing system based on content standards in mathematics and Spanish language. Materials developed by MIDEH/AIR include content standards that define what should be taught and learned in Grades K through 11 in math and Spanish language; pacing guides that organize the standards to be taught and mastered month by month; monthly formative tests aligned with the standards, with a teacher's manual describing the way to administer and score the tests, and interpret the results; and a software package that assists teachers and school directors to record formative test scores on a monthly basis. These materials are all aligned to the new curriculum and the new official Ministry textbooks for Math and Spanish. (See Figure 1.)

---

[1] Estrada, M., Zimmerman, R., & Lansdale, J. (2009). *Pilot of digital datacollection options for Honduras' Education Assessments.* Washington, DC: American Institutes for Research.

Figure 1: DCNB and aligned educational materials

MIDEH/AIR developed end-of-grade summative tests, which were administered over a few years to an increasingly larger national sample of schools spanning Grades 1 through 9. Using the test results and input of teachers from across the country with subject area and grade-level specialization, MIDEH/AIR facilitated the defining of performance standards in math and Spanish language. Analysis of the results from the end-of-grade summative tests has been used by Ministry of Education senior authorities to guide policy adjustments and fine-tune teacher-training initiatives in response to weaknesses identified by test results.

While the project successfully closed in July 2011, the June 28, 2009, coup d'état had a negative impact on the education sector during the 2009 and 2010 school years. The period was marked by teacher strikes, extended periods of missed classes, automatic student promotion, and suspension of donors' assistance. This challenged the administration of the assessments and the piloting of technology to support it.

PILOT RATIONALE

The MIDEH/AIR team considered the following benefits of using technology in data collection:

- Using technology can accelerate data collection. Data can be collected, analyzed, and reported in a significantly shorter timeframe.
- Data security is greater with technology than with paper. Data backups can be established with the initial data transmission, and all the data can be backed up at a secure designated location.
- Technology provides real-time monitoring so that staff can monitor the progress of the data collection efforts, validate the integrity of the data, and identify issues with regard to collection instruments or collectors while the teams are still in the field.
- Using technology can decrease some of the logistical issues associated with paper- based instruments, although different technologies create their own logistical concerns.
- While technology has an up-front cost, the logistical and recurring costs of printing, shipping, and labor with paper-based assessments are greater.

Within Honduras, data collection technology can help overcome the many challenges involved in transporting and handling paper surveys, which for summative assessments, represents a significant barrier to timely processing, since Honduras' remote schools are geographically dispersed in mountainous terrain. Even with diagnostic and formative assessments, large class sizes make paper-based assessment time consuming and cumbersome. Timely processing is needed to ensure that results can be effectively used to improve education. For example, in the 2008 summative assessment administration, it took nearly 3 months to transport, unpack, and collate all the tests, and conduct quality control and digitize them through optical scanning. This delayed the presentation of results and report findings. Both central and local education authorities need to have these results prior to the subsequent school year in order to be responsive to the learning needs of students and training needs of teachers.

While not as challenging as summative assessments, for diagnostic and formative assessment, challenges of paper-based collecting become more problematic for teachers of Grades 7 to 11 because they specialize in a subject area and have to teach several sections of as many as 50 students in each section. For example, it is a common for a high school teacher to teach as many as 7 grade sections with 40 students each. This means having to grade as many as 280 test booklets each month, as well as having to produce the results analysis report. Grading diagnostic or formative tests becomes a lengthy and tedious process, which in turn may deter teachers from administering the tests to their students and/or taking the time to analyze the results in order to provide remediation for students who need it.

This pilot builds on several previous digital data collection pilots conducted to test strategies to better support student assessments. Following is a brief background on the various technologies explored and then a brief description of the earlier pilots.

Turning Technologies offers Audience Response Systems (ARS), a technology that allows recording real-time responses of an audience to a series of questions. The answers are recorded by remote control and are electronically connected to a computer with a radio receiver and software to process the responses. The ARS can be used with different handheld devices called ResponseCards (clickers). These clickers allow data collectors to enter data at the site of collection. (See Table 1 for a picture and specification of the two types of clickers piloted.)

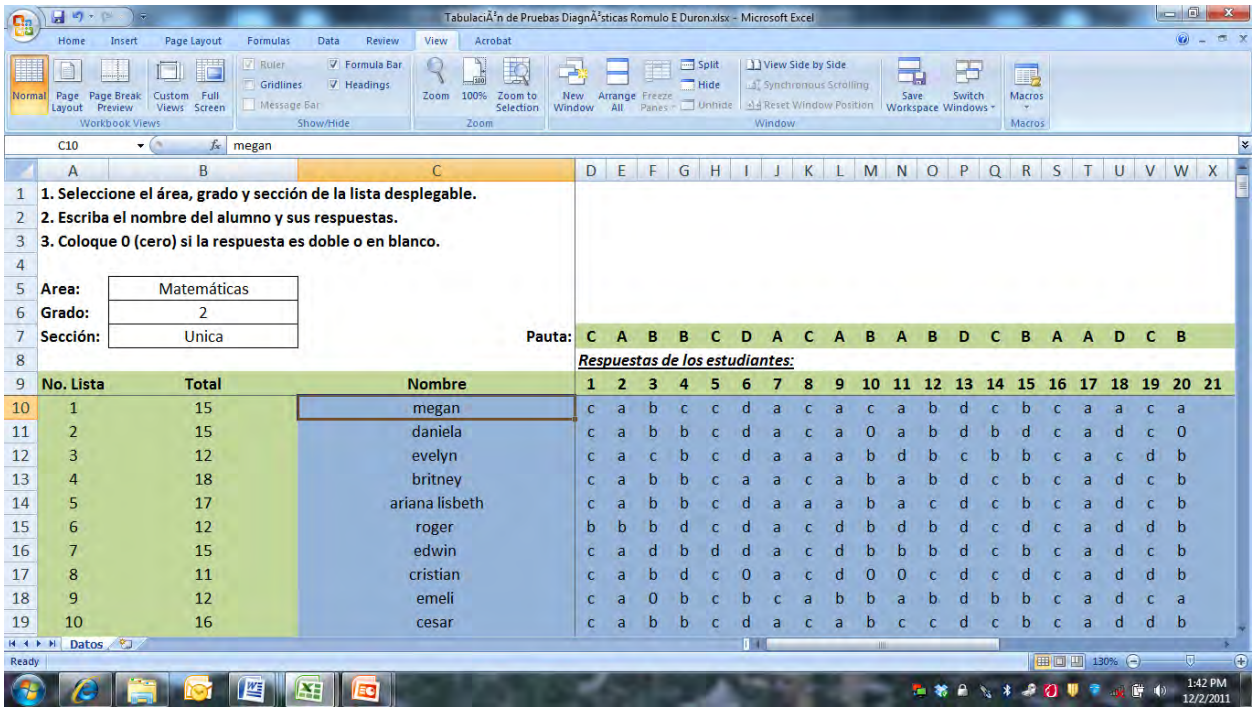
Specifications	ResponseCard RF	ResponseCard XR
LCD Screen	No	Yes
Self-paced data entry	No	Yes
Batteries	Coin cell	AAA
Languages	No	Language pack via updates
Memory	No	Stores up to 5,100-question test
Alphanumerical responses	No	Yes
Storage of user data	No	Yes
Response confirmation before sending	No	Yes
Successful transmission indicator	Yes, via blinking LED	Yes, on screen message
Multiple digit responses	No	Yes

Figure 2. Comparison of ResponseCard RF and ResponseCard XR

Another technology that was included in the earlier pilot was the EarthWalk mobile computer labs. These computer labs allow the use of laptop computers with long-life external batteries, and also provide the possibility of setting up a wireless network to exchange data and share an Internet connection. EarthWalk wireless mobile computer labs can simultaneously charge all regular laptops; power wireless access points, as well as all additional peripherals from one standard wall outlet; and integrate with existing networks and/or Internet access, VSAT, and satellite communication. (See Figure 3 and www.earthwalk.com.)

Figure 3. Image of an EarthWalk lab

MIDEH/AIR staff developed MS Excel worksheets. (See Figure 4.) These worksheets require manual input of data collected from the students' tests. The scores for each student and the analysis for all the students in every grade are calculated automatically by the MS Excel program.



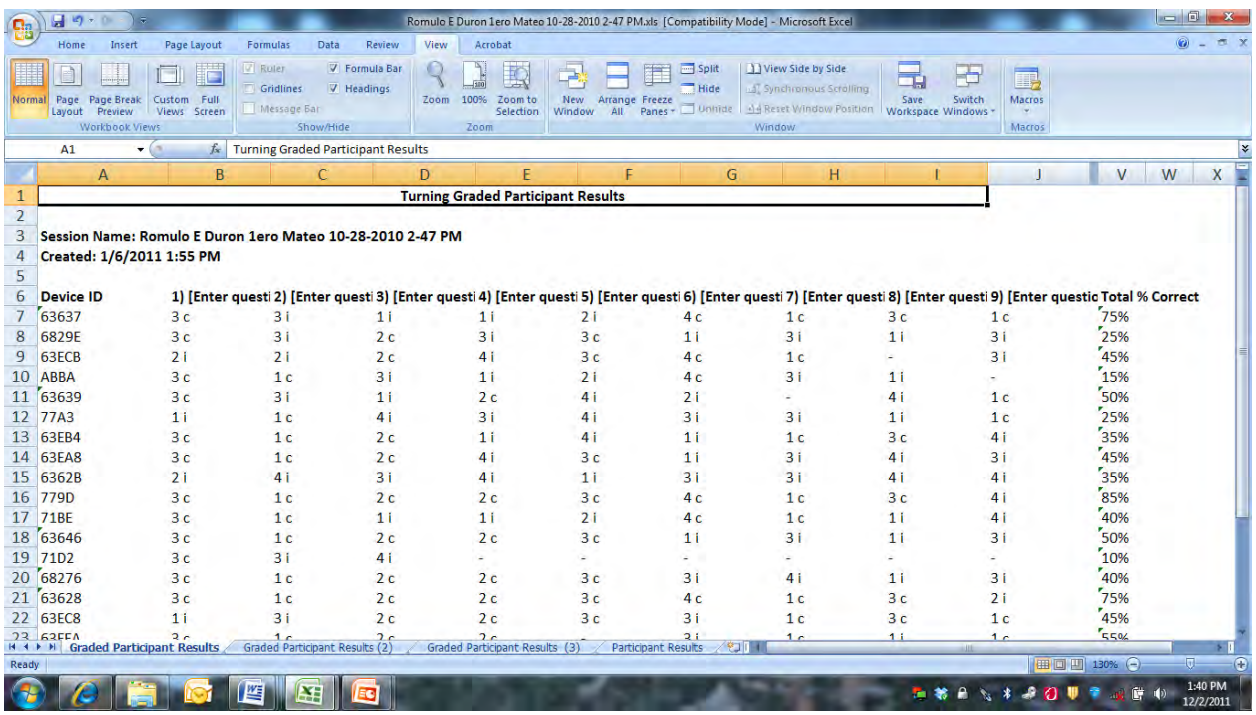
1. Seleccione el área, grado y sección de la lista desplegable.
2. Escriba el nombre del alumno y sus respuestas.
3. Coloque 0 (cero) si la respuesta es doble o en blanco.

Area: Matemáticas
Grado: 2
Sección: Unica

Pauta: C A B B C D A C A B A B D C B A A D C B

Respuestas de los estudiantes:

No. Lista	Total	Nombre	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	15	megan	c	a	b	c	c	d	a	c	a	c	a	b	d	c	b	c	a	a	c	a	
2	15	daniela	c	a	b	b	c	d	a	c	a	0	a	b	d	b	d	c	a	d	c	0	
3	12	evelyn	c	a	c	b	c	d	a	a	a	b	d	b	c	b	b	c	a	c	d	b	
4	18	britney	c	a	b	b	c	a	a	c	a	b	a	b	d	c	b	c	a	d	c	b	
5	17	ariana lisbeth	c	a	b	b	c	d	a	a	a	b	a	c	d	c	b	c	a	d	c	b	
6	12	roger	b	b	b	d	c	d	a	c	d	b	d	b	d	c	d	c	a	d	d	b	
7	15	edwin	c	a	d	b	d	d	a	c	d	b	b	b	d	c	b	c	a	d	c	b	
8	11	cristian	c	a	b	d	c	0	a	c	d	0	0	c	d	c	d	c	a	d	d	b	
9	12	emeli	c	a	0	b	c	b	c	a	b	b	a	b	d	b	b	c	a	d	c	a	
10	16	cesar	c	a	b	b	c	d	a	c	a	b	c	c	d	c	b	c	a	d	d	b	



Turning Graded Participant Results

Session Name: Romulo E Duron 1ero Mateo 10-28-2010 2-47 PM
Created: 1/6/2011 1:55 PM

Device ID	1) [Enter questi	2) [Enter questi	3) [Enter questi	4) [Enter questi	5) [Enter questi	6) [Enter questi	7) [Enter questi	8) [Enter questi	9) [Enter questic	Total % Correct
63637	3 c	3 i	1 i	1 i	2 i	4 c	1 c	3 c	1 c	75%
6829E	3 c	3 i	2 c	3 i	3 c	1 i	3 i	1 i	3 i	25%
63ECB	2 i	2 i	2 c	4 i	3 c	4 c	1 c	-	3 i	45%
ABBA	3 c	1 c	3 i	1 i	2 i	4 c	3 i	1 i	-	15%
63639	3 c	3 i	1 i	2 c	4 i	2 i	-	4 i	1 c	50%
77A3	1 i	1 c	4 i	3 i	4 i	3 i	3 i	1 i	1 c	25%
63EB4	3 c	1 c	2 c	1 i	4 i	1 i	1 c	3 c	4 i	35%
63EA8	3 c	1 c	2 c	4 i	3 c	1 i	3 i	4 i	3 i	45%
6362B	2 i	4 i	3 i	4 i	1 i	3 i	3 i	4 i	4 i	35%
779D	3 c	1 c	2 c	2 c	3 c	4 c	1 c	3 c	4 i	85%
71BE	3 c	1 c	1 i	1 i	2 i	4 c	1 c	1 i	4 i	40%
63646	3 c	1 c	2 c	2 c	3 c	1 i	3 i	1 i	3 i	50%
71D2	3 c	3 i	4 i	-	-	-	-	-	-	10%
68276	3 c	1 c	2 c	2 c	3 c	3 i	4 i	1 i	3 i	40%
63628	3 c	1 c	2 c	2 c	3 c	4 c	1 c	3 c	2 i	75%
63EC8	1 i	3 i	2 c	2 c	3 c	3 i	1 c	3 c	1 c	45%
63EFA	3 c	1 c	2 c	2 c	-	3 i	1 c	1 i	1 c	55%

Figure 4. MS Excel worksheet examples

MIDEH/AIR conducted the digital data collection technology pilot study in four phases. Phase 1, was a proof of concept workshop for the ResponseCard RF. The main objective was to determine whether the ResponseCard devices could be used for data collection procedures. In Phase 2, MIDEH/AIR piloted a different ResponseCard device, the ResponseCard XR, to see which ResponseCard was more appropriate. In Phase 3, the chosen MIDEH tested ResponseCard, XR in a near-to-reality pilot study and

compared with a portable computer lab technology from EarthWalk. While the EarthWalk portable lab was faster, the cost, training, and other factors made the Response Card XR the best choice for collecting summative testing data. Phase 4—and the focus of this study—is the comparison of ResponseCard and a simple MS Excel worksheet for collecting diagnostic testing data.

In 2010, diagnostic tests in Spanish language and mathematics were developed for Grades 2 through 11. There were concerns about how teachers would respond to the additional task of having to administer, grade, and analyze the results of the diagnostic tests. MIDEH/AIR saw this as an opportunity to investigate the use of the ResponseCard devices in support of the administration of the diagnostic tests. To see if digital technology could minimize the time required to grade the diagnostic and formative tests, MIDEH/AIR piloted the ResponseCard XR device, which was previously selected as the technology of choice for end-of-year testing, and compared its performance against an MS Excel worksheet designed to score the tests and process the results after the teacher input the students' responses. To carry out the pilot study, technical staff with the MIDEH/AIR project, conducted 6 sessions during which diagnostic tests were administered to students. The pilot covered a range of settings, grades, and school types. The pilot did not cover formative assessments because, as a consequence of the 2009 political turmoil, the formative test booklets were not distributed to schools in 2010 and the distribution of the 2011 booklets was delayed. Moreover, the formative tests include constructed response items that are difficult to assess with the ResponseCards, which work best with multiple-choice items.

PILOT STUDY OF DATA COLLECTION OPTIONS FOR DIAGNOSTIC TESTING

MIDEH/AIR selected four institutions in the department of Francisco Morazán to participate in the pilot study. Since the study was conducted close to the end of the 2010 school year, teachers known to the researchers were contacted to ensure collaboration at a time when teachers were busy with administrative responsibilities related to closing out the school year. Even though tests are available for both Spanish language and mathematics, only the math diagnostic tests were administered. Math tests require more time for analysis than the Spanish tests because the different blocks and standards evaluated in math require more time for manual calculations and data analysis.

To pilot the devices in a variety of settings, researchers conducted test administration sessions at university, high school, and elementary school levels, at public and private schools in rural and urban environments. (The sessions are detailed Table 1.)

Institution	Grade level	Urban/rural	Public/Private
Universidad Pedagógica Nacional Francisco Morazán	Undergraduate mathematics program	Urban	Public
Instituto Polivalente San Juan de Flores	7th and 9th grades	Rural	Public
Instituto San Miguel	6th grade	Urban	Private
Escuela Rómulo E. Durón	1st grade	Rural	Public

Table 1. Test administration locations

DESCRIPTIONS OF THE TEST ADMINISTRATION SESSIONS

Universidad Pedagógica Nacional Francisco Morazán

The Universidad Pedagógica Nacional Francisco Morazán (UPNFM) is the national university at which students are trained to become high school teachers. A group of 35 students enrolled in the mathematics program took the 10th-grade diagnostics tests. The tests were administered, and when all students had finished, they were given instructions on how to use the ResponseCard devices. The students input their names and their responses. Their information was automatically sent to a central computer in the classroom. The teacher analyzed the compiled test results to quickly produce the final report of individual and classroom results. Overall, the participants felt comfortable with the technology, and the whole process took less than 1 hour to complete.

Instituto Polivalente San Juan de Flores

The Instituto Polivalente San Juan de Flores is a public high school located in the municipality of San Juan de Flores, 40 kilometers from Tegucigalpa. At this high school, students from the 7th and 9th grades took the diagnostic tests, (28 students in Grade 7 and 27 students in Grade 9). Both sections performed well during the activity. Researchers gave the students instructions on entering data into the ResponseCards, which lasted 15 minutes, and then students took the test and sent their responses to the computer as they answered the test questions. As with the session at the UPNFM, the activity lasted 1 hour.

Instituto San Miguel

The Instituto San Miguel is a private school located in Tegucigalpa. Thirty-four students from sixth grade took the diagnostic test. As with the other groups, researchers provided instructions on how to use the ResponseCards in 15 minutes, and then administered the test and collected answers to test items in real time, using the ResponseCards. The participants felt comfortable with the devices, and the activity lasted 1 hour.

Escuela Rómulo E. Durón

The Escuela Rómulo E. Durón is a public school located in Mateo, a village 12 kilometers from Tegucigalpa. At this location researchers asked first graders to take a test and then input data into the ResponseCard devices. The researchers did not ask students to input their names because of the many steps required. The first grade class was difficult to manage because of their young age. The students used the ResponseCard devices as if they were toys. The ResponseCard was too great a distraction. It was not possible to get the whole classroom to pay attention and follow instructions.

RESULTS OF THE PILOTS

On average, the test administration sessions lasted 1 hour: 15 to 30 minutes to explain how to use the ResponseCard devices, and 30 to 45 minutes to answer the test questions and input the data using the devices. The results were available immediately, but the teacher had to manually create the report for the class. The results for each group of students were presented only after the last student completed the test. In the classrooms where this study was administered, researchers asked teachers to score the results manually, to compare time and accuracy against the ResponseCard devices and MS Excel worksheet. When interviewed, teachers said it took them, on average, 2.5 hours to complete the scoring by hand, whereas teachers obtained results immediately when using the ResponseCards and took only about 10 minutes to compile into a report. For the MS Excel worksheet, MIDEH/AIR processed the actual test responses in the MIDEH/AIR office. It took 35 to 45 minutes to input data for a section of 40 students.

KEY FINDINGS

1. The ResponseCards required less time for analysis and reporting, but required sufficient training for students to be able to use the devices appropriately. Even with the training, some students made mistakes inputting their names or the test responses into the device.
2. ResponseCards are more appropriate for older students. The pilot conducted in the first grade showed that these young students used the ResponseCards as toys and were not able to follow directions.
3. Even though the ResponseCards are a fast alternative, teachers still need to do additional calculations to produce the reports. Also, the hardware is expensive and a computer needs to be available in the classroom at the time of the test administration.
4. The MS Excel worksheet provides faster scoring and production of reports than manual scoring, but is slower than the ResponseCards. However, MS Excel is a cheap option relative to the ResponseCards because there is no additional hardware, aside from a computer with access to MS Excel.
5. Manually scoring and reporting is only recommended when there is no access to computers, since it takes an average of 2.5 hours to score and analyze data for a class of 40 students.

Table 2 summarizes the findings across the 3 processing methods.

	ResponseCard	MS Excel worksheet	Manual input
Computer in the classroom	Yes	No (but need access in the school)	No
Additional hardware	Yes	Yes (if don't have a computer associated with class)	No
Time required for reporting	10 minutes after the test session	35–45 minutes after the test session	150 minutes after the test session
Cost	$319.00 for a 40-student kit + computer + batteries	Free (if there is access to a computer with MS Excel)	Free

Table 2. Comparison of ResponseCard, MS Excel Worksheet, and Manual Input Results

CONCLUSIONS

MIDEH/AIR considered the ResponseCard devices a viable technology for processing end-of-grade summative test results, but are not as viable for diagnostic testing. Using ResponseCards decreases the time between test administration and digitization, as well as reducing paper loads and transportation logistics, an incredibly challenging undertaking for summative testing of 100,000 or more students nationwide. Further, by allowing immediate data file transfer, this platform requires less overall data manipulation; using ResponseCards allows for implementation of quality control measures during digitization rather than making teachers wait until after all data are scanned—especially important for national summative assessments. Finally, the implementation of a decentralized approach facilitated by the clickers allows for faster problem-solving strategies during administration and data processing. While further piloting of data collection strategies for summative assessments was planned in 2009, the political turmoil made the piloting impractical at that time. Consequently, MIDEH/AIR personnel made other plans for the use of the ResponseCard devices.

In 2010, diagnostic tests in Spanish language and mathematics were developed, giving teachers additional responsibilities to administer, grade, and analyze the results of the diagnostic tests. While data collection challenges for diagnostic and even formative testing are not as insurmountable as those involved in summative assessments, this was an opportunity to test the use of technology in support of the administration of the diagnostic tests.

The conclusions of this pilot study are that the ResponseCard technology is not appropriate for administering and processing diagnostic tests for the following reasons:

1. The diagnostic tests are administered nationally at the beginning of the school year, in every classroom in the country. Providing ResponseCard devices to every school in the country is too expensive to be a viable.
2. Teachers increasingly have access to computers, either their own laptops or the school's, in the computer lab. With access to a computer with MS Excel, teachers can input the letter results of each student's diagnostic test in the program developed by the MIDEH/AIR staff and have immediate access to the results of all the students, grouped by performance standards, aligned to curricular components, and presented in graphs. In teacher trainings in which this program was presented, teachers responded enthusiastically.

However, ResponseCard technology could be appropriate in the following scenarios:

3. While not piloted, MIDEH/AIR technical staff expect that ResponseCard devices used during formative test administration could help the teachers motivated by technology to better monitor the students' performance item by item on those assessments. The teacher may take the time required for remedial feedback based on those results, especially if a significant number of students submit an incorrect answer..
4. Also, while not piloted, ResponseCards may be useful for monitoring school enrollment by principals and area supervisors and attendance by teachers, principals, and area supervisors. If used by multischool supervisors, the costs of the technology would not be such a strong barrier and using the ResponseCards could allow for more accurate reporting school level data.
5. ResponseCard devices are useful for processing end-of-grade assessments based on pilots conducted previously.
6. Use for EMIS data collection is undetermined, but this is a possible area to explore in the future.