FACT: Helping teachers develop complex mathematical thinking and practices in their classrooms
March, 2014
© Kurt VanLehn & the FACT project

Abstract: The Mathematics Assessment Project has developed over 70 classroom Formative Assessment Lessons (FALs) that address the mathematical practices of the Common Core State Standards for Mathematics. The main objective of our Formative Assessment using Computational Technology (FACT) project is to enhance these FALS by having students work on tablets instead of paper so that the teacher’s expertise can be supplemented with the computerized expertise of other teachers and experts. In particular, when FACT is completed, a teacher monitoring the FALS work of students in the classroom will receive suggestions in real time from FACT on what to say and do, and the suggestions will be based on comparing the students’ work with hundreds of other students’ work that have been annotated by experts. The activities and software are being developed in a series of design experiments and usability studies. This report describes the system and the initial usability studies’ findings.

The major issues addressed
The issue addressed here is to supplement through technology the teacher’s capacity for supporting students as they work on complex, rich tasks for learning math. The proposed combination of classroom activities and technology is called Formative Assessment with Computational Technologies (FACT). A formative assessment (as opposed to a summative assessment, which is a test used for placement, advancement and grading) provides information about students’ performance during their learning activities that is then used to help improve learning (that’s when it becomes formative). For formative assessment, the information needs to be qualitative and fairly detailed, not summative scores or grades. The information at each stage may be given to the teacher or the students. Students can get responses to their ideas in real time – e.g., “other students who tried this approach found this idea particularly helpful” – and teachers can access a library of productive responses to common student performances. The FACT system is socially intelligent, i.e., the system will obtain some of its wisdom from a community of teachers as well as students, and thereby enhance the classroom’s social structure while building the teaching community. The main criteria for success are:

- FACT should promote the adoption of classroom practices that have been shown to improve long-term learning outcomes, such as:
  - rich formative feedback followed by re-engagement with tasks (Black & Wiliam 1998)
  - collaborative learning, discussion and reflection (Swan, 2006).
  - more emphasis on rich problem-solving tasks (Schoenfeld, 1992).
- FACT should apply to a range of tasks and activities, not just a few selected ones.
- FACT should never restrict student performances just so that it can understand them – a challenge most computer-based tutoring systems fail to meet. In particular:
  - Students must do most of their work on the computer, and yet their work cannot be distorted by a clumsy or unfamiliar user interface.
  - Students must be encouraged to converse freely, and yet the computer must also monitor progress even though it cannot understand their speech well.
  - FACT must be able to understand important aspects of student performances even though these activities are complex enough that novel performances, both good and bad, are common.

FACT is built on the Classroom Challenges developed by Mathematics Assessment Project (MAP; http://map.mathshell.org). These detailed, paper-based lessons both assess and promote students’ mathematical understanding and performance as set out in the Common Core State Standards for Mathematics (CCSSM). The lessons encourage students’ integrating bits of mathematical knowledge they possess into the substantial chains of reasoning required by the CCSSM Mathematical Practices. The challenge of electronically capturing and analyzing such chains, and then providing useful feedback to students and teachers, is the core of this project.

The CCSSM and other standards require activities that elicit complex mathematical practices and thinking. MAP’s Classroom Challenges provide such activities. They engage students in exploring a single challenge over approximately 90 minutes spread over several class meetings. Students work in pairs, individually and as a whole class. The teachers’ role is to observe the students’ work and offer feedback, usually in the form of questions that help the students re-engage with their work to improve and extend it. Given the complexity of the students’ work, teachers find this difficult to do, especially when they are walking around the classroom and must analyze a piece of work almost instantly. When teachers cannot understand
what the students are doing, they have an unfortunate tendency to suggest an approach that the teacher understands. Too much guidance undermines the intended exploration. The major issue addressed by this project is developing technology that will supply teachers with a real-time analysis of students’ work that allows them to extend the students’ work rather than quash it. Moreover, the technology should support existing classroom processes rather than requiring changes to them.

**Contextualization**

Currently, MAP has about 30 Classroom Challenges (CC) for middle school and 40 for high school. Each CC’s materials comprise printable worksheets and cards for students, slides for projection and a slim teacher’s manual. The CCs are freely available online, with over a million downloads of individual lessons to date.

Every CC starts, prior to the main lesson, with students spending about 15 minutes working on a task. These tasks are designed to be engaging but hard, with few students expected to produce a strong solution at the first attempt. Teachers review the work overnight with the aid of a list of common issues and suggested formative feedback included in the teachers’ manual. The main lesson is divided into several activities including whole-class discussions and work in pairs or small groups. For instance, students may spend 15 minutes working in groups, then 5 minutes sharing their work with other groups. The teacher then convenes the whole class to have selected pairs display and explain their solutions. The teacher may then hand out work from “other students” (included in the lesson materials) and have the pairs critique them.

There are two main types of CC. **Problem Solving CCs** are built around a single unstructured problem, such as the one shown in Figure 1, for which there are many mathematically distinct and interesting approaches. Students may initially be able to produce a partial solution with simple mathematics (such as trial and improvement) and the activities in the main lesson help them improve this solution. The use of carefully designed “other students’ work” is key in exposing them to more sophisticated and generalizable approaches using mathematics that they should know, but had not thought of applying (a major shortcoming of content-focused curricula). **Concept Development CCs** concentrate on a particular mathematical topic (such as percents). In these CCs, the initial pre-lesson task is more like a conventional math assignment, but is designed to expose common errors and misconceptions. The activities during the main lesson then explore the topics through multiple representations, applications and connections with other topics. For example, Figure 2 shows a part of the outcome of a card-based group/pair activity that helps students connect percents, decimals and fractions and understand, for instance, why reducing something by a percentage is not the inverse of increasing it by that percentage.

![Figure 1: A problem solving CC.](image1)

![Figure 2: A part of a concept development CC](image2)

Observation of teachers and students using the CC have suggested that there are several areas where teachers require additional help, and where technology can appropriately fill this gap. The first is responding to students’ pre-assessments, such as the one shown in Figure 1. The goal is not to “correct” or “score” the pre-assessments, but to raise questions that will elicit deeper reasoning from the student. The teachers’ manual has a table that lists typical student performances and suggests how to respond to each. However, as Figure 1
illustrates, it is not easy to determine what students were thinking given the inscriptions they have made. Teachers have only a limited time for responding to pre-assessments, which should be handed back the day after they are collected. Teachers often omit the pre-assessment activity or omit providing individualized formative assessments. Given a large collection of pre-assessments that have been annotated by experts, it should be feasible to assist teachers by matching the students’ pre-assessments against the collection in order to provide suggestions to the teacher for how to respond. So responding to students’ pre-assessments is one way that FACT should help teachers.

A second occurs during the main lesson when students are working in pairs. Teachers circulate among the pairs posing questions intended to move the students’ reasoning forward. Although teachers know to provide only questions that lead students to consider new possibilities, they are often confronted with a complex solution like the one in Figure 1. They have to analyze it as the group waits for a response. If teachers cannot see an appropriate question to ask, they might fall back on old habits and start telling the students how to solve the problem. Moreover, when a teacher visits every group several times, it can be confusing to keep track of which group got which advice. It should be feasible to help teachers by offering them an analysis as they approach a group. This is the same analysis as used on the pre-assessments, but it would need to operate in real-time, keeping track of the states of many pairs of students working simultaneously.

In addition to these computationally challenging ways of helping teachers, there are mundane ways that FACT can help, such as (1) recording the chronology of students solutions for replaying by the teacher; (2) allowing the teacher to instantly project any pair’s work for the whole class to see; (3) offering students a covert way to “raise a hand” and ask the teacher for help; (4) expediting the workflow, which can get garbled in paper-based CCs that require multiple large posters, sets of cards and worksheets are being passed around.

### Theoretical and methodological approaches being pursued

Despite the aforementioned areas where teachers need help, MAP’s CCs appear to be a stunning success. Not only are they being adopted widely, observations in multiple classrooms suggest that their implementations hew close to MAP’s recommendations and that students are profiting from them (Foster, 2013). Thus, our theoretical approach is do no harm! Assuming this can be achieved, then FACT’s theoretical approach is the same as MAP’s theoretical approach, which in turn is based on Black and Wiliam’s (1998) theory and practice of formative assessment, enhanced by MAP’s 15 years of design experiments.

In order to do no harm to students’ practices, students should continue to do exactly what they are doing now even though their work is being monitored in real time. Moreover, because the CC are used infrequently, perhaps once a month on average, the technology must be walk-up-and-use—very little user interface training should be needed. These design constraints rule out keyboard and mouse interaction, since this fundamentally transforms the act of writing mathematical formulae and sketching diagrams. Students need the freedom and familiarity of pens for inscription. Although using electronic pens with special paper meet these constraints, they do not allow easy monitoring of students’ movement of cards. Thus, FACT has students use tablets with styluses/pens.

However, tablets have a small screen but students need to work on a large surface. Some of the CC activities use a poster that is about 1 m x 0.5 m and has dozens of small (90 cm x 50 cm) cards on it. The other activities use multiple printed sheets of letter-sized paper. Thus, the basic user interface concept is that a CC activity is done on a virtual worksheet of any size that may have immobile read-only text and images and/or movable cards on it. Students can use a pen to write and draw on the worksheet or the cards. They can enter text on cards using either handwriting recognition or a virtual keyboard. The tablet screen acts as a viewport on the worksheet. Pinching does zooming. Dragging on the worksheet causes scrolling. Dragging on a card moves it around the worksheet. Students can have their own private worksheet, and a group of students can use their tablets to work on the same, shared worksheet. Although all changes to a shared worksheet are instantly made visible on the tablets of every member of the group, the location and magnification of the tablet viewports can differ so that, for example, one student can work on one part of the worksheet while another student works on a different part of the worksheet. Alternatively, a pair of students can leave one student’s tablet zoomed out to show the whole worksheet while they work together on the other student’s tablet, which is displaying only a small portion of the worksheet.

In order to do no harm to teacher’s practices, teachers need to be able to see students work not only in an office or home setting as they annotate the pre-assessments, but as they walk around the classroom helping students. For portability, FACT has teachers work on tablets. Moreover, teachers need a walk-up-and-use interface for the same reasons that students do—in frequent usage. Thus, the teachers’ tablet’s main function is to act as a viewport on a student worksheet, and it has the same easily discovered gestures as the students’ tablet.

Working at home or in an office, the teacher can annotate student pre-assessments using ink and cards. These annotations are in color that only teachers can use. Currently, teachers can create annotation cards that contain pre-fabricated questions from the table in the CC’s teacher’s manual. When FACT is able to interpret
student’s work, then it may suggest annotation cards as well. Teachers can edit the pre-fab cards or create
annotation cards of their own.

When walking around the classroom, the teacher can view a group’s worksheet prior to approaching
the group. When FACT is able to interpret the students’ work, it may pop up suggestions for questions to ask
the group. The teachers’ tablet has a few workflow tools as well. One is display of icons for every individual
and group worksheet, with colors that indicate which ones need attention either because the student or group has
“raised its hand” or because the FACT analysis system has detected a serious issue. Another tool is a button
that causes the teacher’s view of a worksheet to be projected onto a screen at the front of the classroom.
Another is a “pause” button that temporarily disables all the student tablets.

Thus, our methodological approach is to replace the posters, cards and worksheets with one tablet per
student, one tablet for the teacher, a portable server and a dedicated wireless access point. These fit into two
rolling suitcases that can be wheeled from one room to another. The classroom’s data projector is plugged into
the server.

This particular electronic media appears to be novel. Because multiple students use it face-to-face, it is
similar to multi-touch tables (Higgins, Mercier, Burd, & Hatch, 2011). Because a small viewport is used with a
large worksheet, it is similar to a shared whiteboard (e.g., awapp.com, twiddla.com). Because it supports
movable cards, it is similar to Group Scribbles (Lin, Chen, Yang, Xiet, & Lin, 2014), trello.com and other
shared card systems. However, FACT’s particular combination of media features makes it difficult to predict
from the literature how students will use it.

**Substantiation**

The project goals are (1) to develop technology that supports existing classroom practices rather than requiring
changes to them, (2) to develop and apply methods for observing, measuring and understanding classroom
processes and student performances, (3) to develop analysis methods for CC solutions that are as accurate as
expert analyses, and (4) to develop a community of teacher-expert practitioners that invents and shares methods
for utilizing these analyses in support of students’ mathematical development. At this writing, we are well
along on goal 1 (electronic media), midway on goal 2 (observation methods), and just starting goal 3 (CC
solution analysis).

With respect to goal 1 (electronic media), the FACT system consists of a Java app for Android tablets
and Java server. The tablet app has four modes: student alone, student in group, teacher in office, and teacher
in classroom. All four modes are operational, and testing in classrooms has begun. Initial reactions suggest that
even without its analysis capabilities, the FACT system might be competitive with the paper-based CCs. For
instances, students like that when they run out of room while writing on a card, they can make the card bigger.
They like to assign colors to group members, so that ink and cards bear the colors of their “owner.” One teacher
liked that students no longer have to waste time cutting cards out of photocopied pages from the teachers’
manual. Another teacher liked that the system stores posters electronically, making it easier to maintain records
and portfolios.

The MAP designers are also excited about the new affordances offered by the FACT electronic media.
For instance, with paper cards, the initial arrangement of cards is not under the control of the designers, but with
FACT, the initial arrangement is controlled by designers and can invite certain mathematically productive
mistakes. As a second example, it is impractical to have both individual and group posters made from paper,
but the easy creation, duplication and comparison of electronic posters invites revision of some CC activities.

With respect to goal 2 (observation methods), when FACT goes to more formal trials, the main
question to be answered by observation is whether it is doing harm to the classroom processes. The main
measure of classroom processes will be based on the TRU Math classroom analysis schema (Schoenfeld, in
press). MAP has been developing this scheme for several years, starting with a review of eight existing
schemes for analyzing classroom process and from a theoretical analysis of goal-oriented human decision
making, especially as it occurs in classrooms (Schoenfeld, 2010). The development process involved cycling
between coding of existing videos of MAP lessons and revising the coding scheme in order to make it more
accurate, more complete, simpler and more reliable. The current schema consists of five dimensions:
Mathematics, Cognitive demand, Access, Agency (authority and accountability) and Uses of assessment. Each
dimension has three levels. For assessing individual learning, we plan to use the pre-assessment and post-
assessment tasks that are part of the activities of many CC. Although assessment of collaboration would be
interesting, many of the methods used in CSCL research (Gress, Fior, Hadwin, & Winne, 2008) cannot be
applied in our classroom contexts.

Our third goal is to develop analytic methods that can interpret data from the students’ tablets and offer
suggestions to teachers about questions to ask students that will extend their mathematical reasoning. We are
just starting work on this goal, and approaching it both top down and bottom up. Top down: for each of the
suggested questions in the teacher’s manual, we are trying to determine if there are ad hoc feature detectors that
can be developed. For instance, if a pair of students has not drawn a table (tables should be easily detected in
the handwriting data), and they have not made many inscriptions in the last 2 minutes which suggests they are stuck or off-task, then one suggestion the teacher could make is to draw a table. In addition, many of the group activities involve matching and arranging cards, and have been deliberately designed so that particular groupings of cards are indicative of well-known mathematical misconceptions – detecting these arrangements should be straightforward. Bottom up: In collaboration with other Bill & Melinda Gates Foundation grantees, we have surveyed open source and commercial handwriting recognition software, and located one that we will start testing in-house. Our tentative plan is to build detectors based on dynamic Bayesian networks. Until we get classroom data, we will calibrate the networks with data from a large corpus of paper-based solutions that we will re-enter onto the tablets.

Generality
The CCs were invented specifically to elicit sophisticated mathematical practices in a classroom context. However, the CC and perhaps any other activity designed to elicit such practices creates a practical problem for teachers: How to understand students’ behavior deeply enough that the teacher can facilitate the students’ practices rather than undermine them. The FACT project hypothesizes that off-loading this analysis onto technology will allow teachers to concentrate on effective and affective communication and thus increase the chance that students will become practitioners.

FACT’s underlying analysis technology will be adapted from intelligent tutoring systems (VanLehn, 2006). One way to conceptualize this project is that it extends that technology by (1) applying it to open-ended, creative problem solving, (2) applying it in a collaborative setting and (3) having it send advice to the teachers instead of the students.

The key challenge is whether the changes that the technology makes to classroom processes are beneficial and acceptable to participants. Although FACT is just beginning to be tested in classroom settings, focus groups, usability lab studies and pull-out studies all suggest that it will meet this challenge successfully.

Acknowledgements
This research is supported by the Bill & Melinda Gates Foundation under grant OPP10612881.

References