Enhancing Learning with Digital Ink: Graphical Real-Time Formative Assessment via Mobile Devices

Susan E. Kowalski[1], Frank V. Kowalski[2]
[1,2]Colorado School of Mines, [2]Red Rocks Community College
(United States)
[1]skowalsk@mines.edu, [2]fkovalsk@mines.edu

Abstract

Learning can be enhanced when real-time formative assessment is facilitated by the use of mobile devices. Formative assessment has many well-established pedagogical advantages: students are actively engaged with their learning, metacognition is increased, misconceptions can be identified and corrected, and instruction can be efficiently paced. However, the logistics of collecting formative assessment can be cumbersome, and instructor response must be timely in order to be most effective. Although clickers (a.k.a. personal response devices) overcome some of these logistical hurdles, many educators are frustrated by their limitations, particularly in probing higher-level thinking skills, in assessing student understanding of complex material, and in nurturing skills such as creativity.

This study further explores another promising technology-based solution for collection of real-time formative assessment. The instructor poses open-format questions (not multiple choice!). Students then use the digital ink of various mobile devices (iPads, tablets, some smartphones) to reveal their thinking through graphs, sketches, equations, proofs, etc. This is facilitated by software (InkSurvey) that is free worldwide; since the software is browser-based, there is nothing to download locally and students can use either their own or school-owned devices.

The instructor receives the graphical responses instantaneously, providing him/her with immediate feedback on the efficacy of instruction, a seamless avenue for scaffolding learning, and timely opportunities to refine student understanding.

This paper presents data collected over 4 semesters in an introductory biology classroom at a U.S. community college. First, the average learning gains achieved (per pre-test/post-test question) when this pedagogical model was used during learning were assessed to establish its viability (n=27). For the following three semesters (n=68), the study compared learning gains achieved for each student when this model was implemented vs. those achieved by a mixture of other active learning instructional strategies. Data indicate statistically significant enhancement of learning gains when graphical real-time formative assessment was collected using mixed mobile devices.

The paper also includes examples of questions posed to the students, some actual student responses, and a brief discussion of some of the challenges of implementation. Although this study involved a community college biology course, the pedagogical model is appropriate for any discipline and any level.

1. Introduction to Formative Assessment

Embedded formative assessment allows students to reveal their understanding during the learning process and provides an opportunity for the instructor to reinforce or refine that understanding in subsequent, appropriately paced instruction. Widely accepted advantages of formative assessment include active engagement of learners and increased metacognition (students' awareness of what they do and do not understand). In a meta-analysis of 250 research articles, Black and Wiliam [1] found solid evidence that seamlessly embedding formative assessment can produce significant and often substantial learning gains. They emphasize, however, that every student must be involved in the process, and the interval between student input and instructor response must be short. This combination presents a great clerical challenge to the instructor: a class session creates large amounts of student-generated learning artifacts that must be submitted, evaluated, and responded to quickly.

Fortunately, technology can lower hurdles for the instructor. Personal response systems ("clickers") are inexpensive devices that afford many of the advantages of formative assessment in real-time, even in large lecture halls, by allowing students to electronically respond to questions posed by the instructor. Similarly, students can text responses via their cell phones and software such as Poll Everywhere, or cloud-based applications (*e.g.*, Top Hat).

Recent widespread adoption of these methods to quickly and easily gather real-time formative assessment has also made educators aware of their limitations. Restricted to multiple-choice or very short answer questions, it is difficult to probe higher-level thinking skills valued in the 21st century. Also, this format may not reveal student understandings of complicated or visually-rich concepts, and it furthermore implies that there is only one correct response. To overcome these shortcomings, instructors can effectively pose open-format questions and students can use pen-enabled mobile devices to prepare and submit their responses.

2. Using Digital Ink for Graphical Real-Time Formative Assessment

Although the use of digital ink for graphical real-time formative assessment can be facilitated by various avenues, this research project used *InkSurvey.* The browser-based software, developed expressly for this purpose, is available for free worldwide (ticc.mines.edu).

The instructor can prepare before class meetings all of the formative assessment questions to be used. At appropriate times during class, the questions are activated to become available on the student devices. Learners use digital ink to prepare their response, by drawing, writing, solving equations, graphing, etc. Upon submission, the instructor receives the responses instantaneously. With this glimpse into student thinking, the instructor is thus informed for subsequent, well-paced instruction that reinforces and refines student understanding. Similarly, students are "primed" and receptive to the instructor's response to their input. Selected student submissions can be anonymously displayed, providing springboards for class discussions that can include valuable practice in critical thinking. Such graphical input can be extremely rich in revealing student misconceptions and the open-format questions are well suited for probing student understanding of complex concepts, higher level thinking skills, and creativity.

Earlier research on the effectiveness of using digital ink for real-time formative assessment has been conducted at a prestigious engineering university, where this pedagogical model was effective at enhancing learning in physics/engineering physics [2] and chemical engineering [3] students, independent of learning styles [4].

3. Experimental Design

To investigate the transportability of this teaching model to other learners, studying other disciplines at different institutions of higher education, this experiment was conducted in an introductory biology course at a community college in the US. Broadly, this setting includes students more diverse than those in previous studies in many ways: by gender (67% females, compared to <15% previously), by age, by ethnicity, by socio-economic status, by prior academic experience and success, by major fields of study, and by aspiring career paths. The experiment was conducted over 4 semesters with a single instructor.

The intervention for this experiment involved the collection of and response to graphical real-time formative assessment embedded in the classroom instruction. Students were equipped with a variety of pen-enabled devices (institution-owned HP Tablet computers and iPads and personal devices including iPads, tablets, and Smartphones). For each concept targeted for the intervention, there were 1-5 real-time formative assessment questions facilitated by *InkSurvey*. Sometimes these questions asked students to apply their new understanding, and other times they were scaffolding in nature, designed to help students "connect the dots." To measure learning gains for each concept targeted for the intervention, students completed a brief paper and pencil pre-test question at the beginning of the class session, and the same question as a paper and pencil post-test at the end. Thus only learning gains achieved during the class session were measured and gains achieved outside of class were excluded. No other modifications were made to content, pedagogical techniques, and classroom procedures.

The first semester's design explored the viability of the intervention in this pedagogical environment and provided an opportunity for the instructor to gain experience in effective implementation. The experimental design was:

PRE-TEST → LEARNING with intervention → POST-TEST

time

where the pre-test and post-test are identical and measured student understanding for the targeted concept. The entire procedure was completed within one class session (100 minutes). This design was implemented for a total of 10 concepts interspersed throughout the semester.

The following 3 semesters used a slightly different experimental design to compare the efficacy of the intervention with other pedagogical models. Pre- and post-tests were modified to assess learning gains not only for the concepts targeted for the intervention, but also for concepts taught by other methods during the same class session, using various active learning instructional strategies that the instructor had found to be effective with previous classes at the same institution. For the final 3 semesters, the experimental design was:

PRE-TEST — [LEARNING with intervention and LEARNING without intervention] — POST-TEST

time

where the pre-test and post-test are identical; each measured student understanding of one concept delivered via the intervention and one concept delivered using other active learning models. Since not all concepts in a course are equally difficult to master, care was taken to randomly target concepts for intervention; the order of instructional delivery techniques and pre- and post-test questions were also randomly alternated. The entire procedure was implemented within one class session (100 minutes) on a total of 20 separate class days over 3 semesters. Each day yielded 4 points of data for each student: pre- and post-assessment of student mastery for one concept taught using the intervention and one concept taught without the intervention.

Only data for students present for both the pre-test and the post-test of a concept were included in the results. Although the set of students providing data varies from concept to concept according to attendance, on any given day the identical set of students served as both the control (no intervention) and experimental (intervention) groups (modified from Shadish [5]).

All pre-tests and post-tests were scored on a 0-3 point scale. Then, for each concept assessed for each student, the "learning gain" was calculated (after Hake [6]) to compare how much learning was achieved (or at least demonstrated) with how much *could* have been achieved:

$$\frac{(\text{Post-test} - \text{Pre-test})}{(3 - \text{Pre-test})} \times 100\% = \text{Learning Gain}$$

4. Results and Discussion

Students in the first semester of the experiment (n=27) demonstrated an average learning gain of 44% for the 10 concepts targeted for instruction with this intervention. These results were considerably better than "doing no harm," even though the pedagogical model had not been used previously by either instructor or students. To illustrate the richness of student responses, figure 1 shows representative examples of student work from the first semester of this experiment.

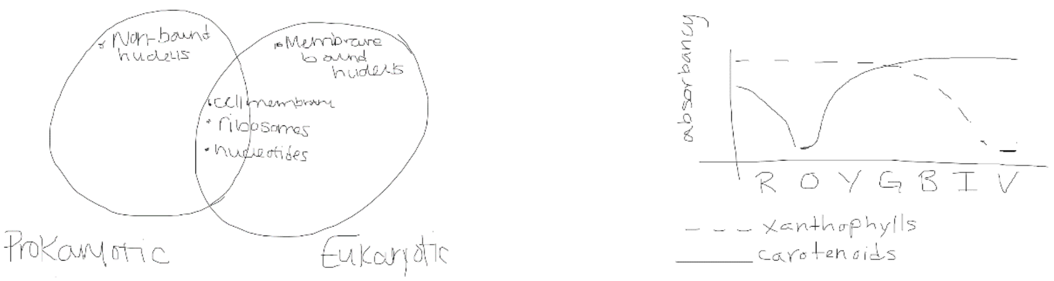

Fig.1. *(left) Student response on InkSurvey to the question: "Make a Venn diagram for the characteristics of prokaryotic and eukaryotic cells." (right) Student response on InkSurvey to the question: "Make a graph showing your predicted absorbance for carotenoids and xanthophylls*."

These results indicate that the teaching model is transportable to this particular learning environment and paved the way for the following 3 semesters of data collection, comparing this intervention to other active learning models (n=68 students total).

The average learning gains achieved for the 20 concepts taught with the intervention were 54.8%; for the 20 concepts taught without the intervention, average student learning gains were 41.3%.

Cohen's kappa coefficient, a measure of inter-rater agreement calculated using ReCal OIR [7], for this data is 0.954; a value of 1 is perfect agreement (all graders producing exactly the same score).

To determine the statistical significance of these results, a 2-tailed Student's t-test compared *each* student's learning gains achieved, with and without the intervention. That is, it is a comparison for each of the 68 students of the total learning gains they individually demonstrated for all concepts taught using the intervention vs. the total learning gains they achieved otherwise. The result of the Student's t-test was 0.00428; generally, if the value is less than 0.05, the difference is significant.

Finally, the 3 semesters of data were statistically analyzed for effect size, a quantitative measure of the strength of the intervention's effect. Cohen's d value was calculated to be 0.53, which is considered a moderate effect size [8].

The technology presented no barrier in implementation and the students, even those who could not be considered digital natives, readily and enthusiastically adapted. Student acceptance of this model appeared to be enhanced by class discussions of the low-stakes, non-threatening nature of real-time formative assessment.

As in other studies of this type, one shortcoming of this research is that there was no measurement or evaluation of the quality of instructor response to student input, even though instructor response has been described as the "linchpin" that links together all the parts of the formative assessment process [9]. However, every time the intervention was implemented, the instructor discussed the student input in class and used the formative assessment to inform subsequent instruction.

5. Conclusions

Statistically significant results confirm that real-time formative assessment with digital ink can enhance learning in introductory biology students at a community college. Students achieved large learning gains when graphical real-time formative assessment was collected using mixed mobile devices; these gains are significantly greater than those achieved when other active learning techniques were used.

This study provides encouraging evidence that the pedagogical model may be applicable to enhancing learning in a wide variety of educational environments, including those diverse in student and instructor populations, discipline, and institution.

6. Acknowledgements

This material is based upon work supported by the U.S. National Science Foundation under Grant #DUE1044255 and the HP Catalyst Initiative program.

References

[1] Black, P. & D. Wiliam. "Inside the Black Box: Raising Standards through Classroom Assessment," *Phi Delta Kappan*, 80 (2), Nov. 1998.

[2] Kowalski, F.V. & S.E. Kowalski. "Enhancing Classroom Learning by Coupling Interactive Computer Simulations with Real-Time Formative Assessment through Technology-Mediated Open-Format Questioning," *Proceedings of the AAEE 2014 Conference*, Wellington, New Zealand, Dec. 2014.

[3] Gardner, T.Q., S.E. Kowalski, & F.V. Kowalski. "Interactive Simulations Coupled with Realtime Formative Assessment to Enhance Student Learning," *Proceedings of the ASEE 2012 Conference and Exposition*, San Antonio, TX, June 2012.

[4] Kowalski, F.V. & S.E. Kowalski. "The Effect of Student Learning Styles on the Learning Gains Achieved When Interactive Simulations Are Coupled with Real-Time Formative Assessment via Pen-Enabled Mobile Technology," *ASEE/IEEE Frontiers in Education Conference,* Seattle, WA, Oct. 2012.

[5] Shadish, W.R., T.D. Cook, and D.T. Campbell*. Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* Belmont, CA: Wadsworth Cengage Learning, 2002.

[6] Hake, R.R. "Interactive-Engagement versus Traditional Methods: A Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses." *American Journal of Physics*, 66, 1998.

[7] Freelon, D. Online tool for reliability calculations at *http://dfreelon.org/utils/recalfront/*

[8] Cohen, J. *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates, 1988.

[9] Brookhart, S.M., C. M. Moss, & B.A. Long. "Teacher Inquiry into Formative Assessment Practices in Remedial Reading Classrooms," *Assessment in Education: Principles, Policy and Practice*, 17, pp. 41-58, 2010.