

ECDemocratized: A Democratization of Educational Assessment

Yoon Jeon Kim, Florida State University, yjkim.fsu@gmail.com
Peter Wardrip, University of Pittsburgh, psw9@pitt.edu
Benjamin Stokes, University of Southern California, bstokes@usc.edu
Adam Ingram-Goble, DePaul University, adamaig@gmail.com
R. Benjamin Shapiro, University of Wisconsin, ben@getdown.org
Russell Almond, Florida State University, ralmond@fsu.edu
James Paul Gee, Arizona State University, James.Gee@asu.edu

Abstract: How might educational assessment be democratized and made more meaningful and engaging for students? Too often, assessment is separated from the pleasures of learning, like a painful test after the game, or worse, a boring class. We propose a symposium to introduce a design experiment with iPad graphics and user narratives called ECDemocratized and initiate discussion around democratization of educational assessment. ECDemocratized is designed to blend the principled assessment design framework, Evidence-Centered Design, with a power shift to give students new agency over their assessment. While ECDemocratized is not a game, it will draw on lessons from games about mediating feedback and gathering data as part of player engagement. Can we democratize current educational assessment practices by introducing a digital media tool that incorporates new learning and assessment theories? Let's find out.

Introduction

High stakes testing has led to the proliferation of pedagogical approaches that treat the student as a vessel to be filled with knowledge, evidenced by the increased pressure on schools to teach to tests. This has the effect of increasingly rendering knowledge inert (Whitehead, 1929), consisting of facts about a domain, with distant and diminishing utility for authentic use. In this assessment regime, tests are treated as the best—and potentially the only—option for collecting and evaluating evidence for monitoring the learning of children, and the performance of schools.

Simply put, the current practice of educational assessment is based on the outdated learning theories and a psychometric tradition that often emphasized students' content knowledge without context. In contrast, contemporary cognitive and learning science theories support sociocultural and situative perspectives of learning where students interact with their social and cultural environments (i.e., activities, resources, teachers, and peers) to develop knowledge and understanding of the world. Thus many strongly argue that we need to move beyond testing content to incorporate more complex aspects of learning (Moss, Pullin, Gee, Haertel, & Young, 2008; Pellegrino, Chudowsky, & Glaser, 2003).

Another problem of current assessments is that it alienates the most important stakeholder in learning: the student. Stiggins (2002) once noted that,

We are a nation obsessed with the belief that the path to school improvement is paved with better, more frequent and more intense standardized testing. The problem is that such tests, ostensibly developed to "leave no student behind," are in fact causing major segments of our student population to be left behind because the tests cause many to give up in hopelessness -- just the opposite effect from that which politicians intended. (p. 2)

It is very difficult for students to feel ownership in the learning process when they are assessed by others' standards and rubrics—especially when they are not given support to understand how the assessments are really being made. Yet blaming these standards is premature when the testing infrastructure and designs are focused so purely on serving administrators rather than students. To students, it is not the standards but the assessment *experience* that can be punitive and lacking in formative guidance.

Addressing these problems can help shape educational reform at large. As Gee and Shaffer (2010) strongly argue, changing assessment is essential to making any deeper changes in what and how students learn. Yet assessment theory alone can be incremental and reactive. We propose a design

research intervention to illustrate what assessment can look like as an important component of 21st century education.

The goal of this symposium is twofold. First, it is intended to open up discussion about what democratized assessment is and how we can accomplish it using digital media technology. Second, this symposium intends to ground often-abstract discussions of assessment in a specific context by introducing an iPad application called *ECDemocratized*. Particularly, we focus on the potential for overlap between student engagement and assessment rigor. That is, our design research is a way to investigate if digital media can provide a middle-ground with both psychometric rigor *and* the flexibility to engage and empower students. We will also highlight how such an application can be embedded within school curriculum, and discuss core design features of the application using an example of Rube Goldberg projects. The panelists of the symposium include Adam Ingram-Goble, Ben Shapiro, Benjamin Stokes, Yoon Jeon Kim, and Peter Wardrip who are currently developing *ECDemocratized*; and two assessment experts as respondents, James Paul Gee (Arizona State University) and Russell Almond (Florida State University). The resulting symposium will be provocatively concrete, include responses from former teachers, and aims to start a new conversation about the goals and democratic possibilities of assessment in schools.

Theoretical Frameworks

Situativity Theory

Drawing on situativity theory (Brown, Collins, & Duguid, 1989; Greeno, 2006), we believe that we can engage students and teachers in the collaborative development of assessments. In the process, we hope to increase the transparency of the act of assessment, to alter the relationship between work and the assessment, and to help teachers to make the assessment process more iterative and formative. Fredericksen and Collins (1989) utilized the word “transparency” to stress the importance of students understanding the criteria by which they will be assessed. As Lave and Wenger (1991) point out, participation in the activity cycle of work is as important to the learning process as any one task. Indeed, the activity, context, and purpose of work are inseparable in traditional non-school based sites of learning. Informal learning contexts are increasingly relevant models, especially as we seek to make learning more authentic to applied domains of expertise and relevant to student passions in order to maintain engagement.

There are three problems of educational assessment that are emphasized by our design research. Note that by “democratizing” we refer broadly to a state of affairs whereby those involved in the assessment process—from students to teachers and parents—have more power over the things that matter to them. First, assessment is too often executed as *separate from learning*, as illustrated by specialized and separated testing as well as the abstract language and processes that are opaque to students, and in some cases teachers. Such separate assessments distract from why learners are in school, which is quite simply to learn (not to be tested). Our design is based on the assumption that the contexts and activities in which people learn are fundamental to what they learn (Greeno, Collins & Resnick, 1996). Separate assessments are subsequently perceived as a burden on teachers and students without a return on the learning they care about.

Second, assessment is *rarely done with students*, even though the results have an impact on the student’s life—from college admissions to employment. When the stakes are high, students care about the outcomes, and yet they often have no power over how assessments are constructed or administered. *ECDemocratized* positions students as co-creators of the assessment with teachers, thereby changing the power dynamics of the classroom, and thus helping students gain ownership over the learning process (in the model of Shepard, 2000). By democratizing assessment, students will gain a greater sense of agency. The development of agency should be intrinsic to the learning environments we design (Greeno, 2006).

Third, current assessments *rarely empower students with assessment skills*. This leaves students at a disadvantage, unable to articulate the value of their own learning, and unable to meaningfully apply assessment for their own purposes. Especially in the modern information economy, our work only counts when we can articulate why it has value as a competency, and students need the skills to pick the value system within which they will be judged. *ECDemocratized* forces students to debate their assessment criteria with their peers, offering an opportunity for students to cultivate their identity in relation to the disciplinary vocabulary. Great teachers already spur such discussions, but not as part

of formal assessment. We propose that a digitally-mediated assessment process could also spur conversations for reflection and identity formation, as part of the student becoming a member of a disciplinary community (Greeno, 2003).

In sum, the ECDemocratized seeks to empower students to participate in the full ecological cycle of their own learning and assessment.

Evidence-Centered Design (ECD)

ECDemocratized is based on ECD, and its approach to making “evidentiary argument.” ECD itself is an assessment design framework that supports collaborative design with rigor (Mislevy, Steinberg, & Almond, 2003). Specifically, ECD fosters rigor by emphasizing the coherence and fit between evidence collection and the interpretation of results for the learning goals. Coherence is fundamental in order to make broader claims about students. Importantly, achieving this coherence comes through social negotiation between stakeholders. This negotiation is an excellent content for democratizing power relations between actors, beginning with conversations about what counts as evidence and why they think that evidence is important.

Three conceptual models of ECD are echoed in the design of ECDemocratized:

- *Competency Model:* A competency model concerns what is the claim that we want to make about a student at the end of the assessment. A given assessment is meant to support making claims or inferences about students at different levels of competency, and variables in the competency model usually describe the set of knowledge, skills, ability, and other attributes of students on which inferences are to be based.
- *Evidence Model:* An evidence model expresses how the student’s interactions with, and responses to a given problem constitute evidence about competency model variables. The evidence model (EM) attempts to answer two questions: (a) What specific behaviors reveal targeted competencies?, and (b) What is the functional or mathematical connection between those behaviors and the CM variable(s)? Simply put, an evidence model lays out the argument about why and how the observations in a given task or situation constitute evidence about CM variables.
- *Task Model:* A task model (TM) characterizes and constructs situations with which a student will interact to provide evidence about related competencies. That is, TM specifies what the student will be asked to do and what kinds of responses will be submitted. Tasks are the most obvious part of an assessment, and their main purpose is to elicit evidence (which is observable) about competencies (which are unobservable).

ECDemocratized will embody the very fundamental idea of the ECD approach by engaging students in collecting evidence, choosing between assessment criteria, arguing how the collected evidence does or does not satisfy the criteria, and structuring the collection of evidence as an argument for project completion.

Games + Learning for Assessment

Finally, our approach to assessment design is also heavily influenced by the growing literature on digital games for learning (e.g., Gee & Shaffer, 2010). Well-designed games are inherently engaging, with embedded ongoing assessment mechanics that are tied to the game tasks. Also, games provide implicit and explicit feedback about players’ performance, and players must use that information to be able to succeed in the game and enjoy it. Similarly, productively “failing” in games is a normalized experience for players, which promotes opportunities to learn by iterating through solutions and strategies. All these activities are not separated but instead are seamlessly merged together in games. Even though ECDemocratized is not a game, it transforms assessment as a *game-like* activity because (a) it leads students to actively and continuously interact with the assessment process as part of what makes it pleasurable, (b) it makes students’ accomplishments and failures explicit and visible in ways that are useful for the learner to proceed, (c) it mediates the feedback in real-time, yet is based on countless hours of testing to ensure that the assessment is meaningful to the participants themselves.

How ECDemocratized Works

ECDemocratized is both an iPad application and a curriculum, grounded in the psychometric rigor of ECD. ECDemocratized works by guiding students over several weeks to use their iPads to make evidentiary arguments in the form of pictures they take, audio they record, and text they write. Students progressively link together their multimedia to make evidentiary arguments for their competence in a domain of expertise.

At the same time, ECDemocratized will give students some power to choose the learning goals themselves. This process is semi-structured, both by the app and by how the teacher implements it. The underlying ECD theory is very compatible with this approach, since ECD is fundamentally about aligning assessment goals with the evidence available. This is unusual—normally assessment goals are not allowed to evolve, especially in response to student demands. ECDemocratized scaffolds this process by first providing the students with a complete assessment model (as set by their teacher), before allowing them to suggest modifications to the underlying rubric. Through this scaffolding, the app avoids overwhelming students with a blank-slate of unlimited assessment options (this is a problem with entirely “open” portfolio systems), and instead seeks to provide them with what game designers would call *‘meaningful choices’* between assessment possibilities.

ECDemocratized is designed as a supplement to existing classroom projects, especially those that take place over several weeks and lend themselves to photographic documentation. We find that engineering labs have particular promise as test contexts.

The simple act of taking pictures becomes a reflection activity, in part because the act of preservation implies a future use, and presumes a future audience. As a photographer, the student is forced to be proactive, choosing moments that matter and ignoring those that distract. This process of curation is a natural and ongoing reminder about the underlying goals for the activity, and the presence of rubrics that will be used to indicate success.

Concretely, consider an ECDemocratized physics lab where the students are asked to demonstrate systems thinking around simple machines in Rube Goldberg projects. Rube Goldbergs are a classic middle school physics lab wherein students must use engineering machines like ramps and levers to accomplish an amusingly simple task, like turning off a light bulb. As they build the physical Rube Goldberg device, the student group will also use their iPad to capture evidence, and structure evidence into arguments about their mastery of physics concepts (like the conservation across kinetic and potential energy), ultimately argue for their competence as systemic thinkers.

Of course, systems thinking has many possible indicators. Working from the notion that a system is a set of things interconnected in a way that produces their own pattern of behavior (Meadows, 2008), students would first be challenged to “define” the concepts of systems thinking salient to their work. By using drag-and-drop on their iPad, they would select three to five concepts to emphasize. For example, students may highlight the constitutive parts of their Rube Goldberg and highlight the multiple interconnections that impact the Rube Goldberg carrying out its task. The result is a personalized competency model for systems thinking. This competency model could also include the students’ design skills for building their Rube Goldberg. This might include planning by storyboarding or the generation and analysis of alternative methods of movement for their Rube Goldberg device.

Once their model is complete, they would similarly choose several possible indicators for each component in the competency model. The irony is that this general process is nothing new—many teachers must already tackle similar logic chains to align with national standards—but it is rare that students are given the power to not only gather their own evidence, but to refine the model, and see how particular models fit with the kind of evidence they iteratively collect. We hypothesize that this *process* will be empowering and will not only improve outcomes on traditional standardized tests, but it will also lead to new meta-cognition.

As students progress in the project, they may collect new and better evidence for the criteria. They also engage their small team in reviewing how they have been assessed. This could change the group configuration, where a team of 3 students may assign one person at a time to be in charge of “documentation” while the other two are building. Throughout, the teacher will have access to the evidence collected by each group, creating opportunities to engage students in their thinking about evidence and the criteria, both online and in direct discussion. This offers a level of transparency to

the assessment process, both for what the teacher is expecting of the students and the extent to which the process is facilitating the growth of the learner. At the end of the day or week, student groups will “swap” their portfolio with another team, asking for their input on whether specific pictures and videos are convincing for the argument they are trying to make. This social negotiation among peers serves at least two learning goals. First, it exposes the students to their peers’ use of evidence in their own learning process to critique and observe. Second, it enables the students to check how their own choice of evidence supports their argument with different audiences.

The process is iterative (see Figure 1) involving weekly (or at any interval that is appropriate for the given project) formal reflection on the evidence that has been amassed and the claims it supports. With each cycle, students’ abilities in collecting and arranging evidence will grow and become more sophisticated. Several weeks are likely to be necessary to see the growth in meta-cognition around assessment, and to justify the additional investment that the classroom teachers must make in establishing the system and dedicating scarce classroom time.

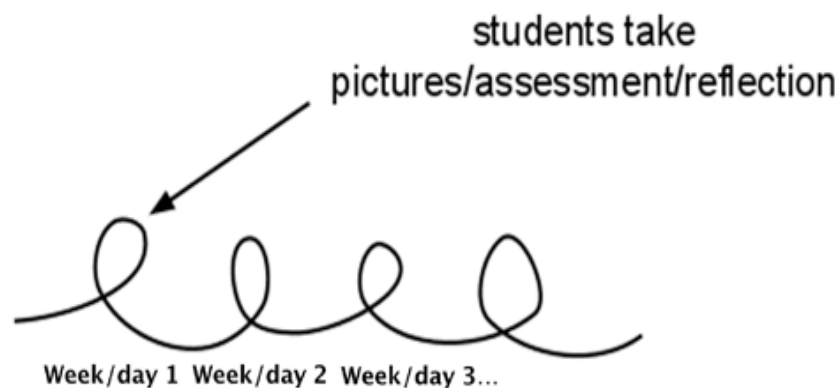


Figure 1: Assessment cycles embedded in student work

As they develop their thinking about how assessment works, students will develop the assessment skills and dispositions that are increasingly necessary for careers that require constant learning—where the criteria for success are complex and articulating that success is vital. Rather than primarily serving to help sort students as might be construed from the test-driven assessment paradigm, assessment with ECDemocratized aims to support students by fostering skills and motivation to leverage assessment for their own lifelong learning by treating assessment as a 21st century skill.

Key to treating assessment as a 21st century skill is our explicit view of epistemological pluralism, or “...accepting the validity of multiple ways of knowing and thinking” (Turkle & Papert, 1992, p. 3). This not only makes explicit the varied types of evidence from which the students may draw in order to richly defend their evidentiary arguments. In addition, this potentially engages them in epistemological conflicts with their peers around what constitutes sound evidence for a particular argument.

ECDemocratized ultimately seeks to develop a participatory assessment system. Here, students not only collect their own evidence, but also take turns evaluating each other. In this way, assessment becomes an iterative social conversation, always toward making improvements in the product of labor rather than passing judgment on the laborer. Moreover, the evidence of learning and competency is collected by students in order to form arguments. This is similar to the way that storyboards are used in legal courts, giving the evidence narrative coherence. Ultimately, the collection of evidence is made transparent to teachers, who see a “classroom view” as well as a “student view.” The teacher can zoom in to see the evidence and arguments that each student is making, and which learning standards this evidence is connecting to.

As a Provocation to the Field

As a provocation, ECDemocratized aims to stir a conversation among education scholars, well before it launches as a product in 2013. Here are some contributions it aims to make:

- *Building theories*: ECDemocratized explores the intersection of assessment and design. The process of co-assessment serves as an additional learning opportunity with implications for the students' identity construction as well as their feeling of empowerment.
- *Assessment as 21st century skill*: ECDemocratized will offer a clear model for situated assessment that builds meta-reflection about the very nature of assessment itself. Identifying learning goals, making claims about learning with self-selected evidence, critiquing one's own as well as one's peers' assessment choices and evidence serve as valuable components of building lifelong learning skills.
- *Open and extensible platform*: ECDemocratized can guide other projects to create their own iPad and mobile tools for learning by incorporating embedded assessment and documentation into the design.
- *Assessment policy connection*: Projects such as the Gordon Commission's report on the future of assessment, and the PARCC and Smarter Balanced consortia for the development of common core state standards assessment are currently investigating the role of assessment for 21st century schools. In addition, there are broader efforts of foundations like Gates and MacArthur that are investigating how to embed and situate assessment into digital media and learning. These policies will benefit from a rich example of assessment on mobile/iPad media embedded in science classrooms.

Discussion by James Gee: Implications for Education Reform

In these remarks, I will react to the specific and unfolding design of ECDemocratized from the perspective of school reform. Where appropriate, my remarks will provide some historical context for reform efforts, but they will primarily focus on where we can go next, and how this application does (and does not) hold promise as an exemplar.

Over the last decade, the issue of testing has become central to school reform efforts in the United States (McNeil, 2000). Demands for the "accountability" of schools has led to an emphasis on basic skills (e.g., reading, math, and science), and teaching to the tests; this has sacrificed opportunities for active and critical learning. However, current work in sociolinguistics, cognitive science, and literacy studies suggests that a more complicated view of learning, assessment, and equity is required. Most importantly the primary stakeholders of these maligned assessments—the teachers and students—remain voiceless on what and how they are assessed, and how results of assessments are used. A tool like the ECDemocratized can provide a "sneak peek" to what accountability might look like under a more democratic paradigm for assessment.

I will especially analyze how ECDemocratized might provide an opportunity for students to learn reflexivity and to synthesize—especially how their knowledge and skills fit together, and how this knowledge is situated. Such synthesis is necessary for the knowledge to be authentic, and for its future application. We might call this systems thinking, which requires understanding of how the knowledge is situated in a community and semiotic domain. For equity in education, students must engage in critical learning, which means going beyond understanding of how to produce meanings in that domain, and, in addition, how to think about the domain at a "meta" level as a complex system of inter-related parts.

My provocation will address the kind of inter-related parts that are implied by ECDemocratized, and I will challenge the audience to reflect with us on what might be missing, and how it might be addressed.

Discussion by Russell G. Almond: Informal versus Formal Models of Assessment Design

Evidence-centered assessment design owes much of its structure, including its name, to Schum's 1994 book, *The Evidential Foundations of Probabilistic Reasoning*. While ECD works at a qualitative

level, which is where this App aims, it does so in large part because the informal qualitative argument can be mapped onto the more formal probability model. In the first part of this discussion, I will explore to what extent the mapping between the qualitative arguments captured in the App can be mapped back onto the more formal mathematical models of ECD.

Zapata-Rivera and Greer (2004a,b) describe a similar system. Here students use a graphical representation of the proficiency model to start a dialog with their instructor about their levels of proficiency. The second part of my commentary will be to draw parallel and lessons from that work to the current proposal.

The third issue is that a large part of “democratizing” ECD involves sharing ECD models. There are several technical (e.g., common data formats), practical (e.g., common meta-data for classifying models), and legal (e.g., restrictions caused by copyright, patent and distribution rules in the Apple App store) which could prove challenging.

References

- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32-42.
- Fredericksen, J. & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher*, 18(9), 27-32.
- Gee, J. P. (2000-2001). Identity as an analytic lens for research in education. *Review of Research in Education*, 25, 99-125.
- Gee, J. P., & Shaffer, D. W. (2010). The right kind of GATE: Computer games and the future of assessment. In M. Mayrath, D. Robinson, & J. Clarke-Midura (Eds.), *Technology-Based Assessments for 21st Century Skills: Theoretical and Practical Implications from Modern Research*. New York, NY: Springer-Verlag.
- Greeno, J. G. (2003). Situative research relevant to standards for school mathematics. In J. Kilpatrick, W. G. Martin, and D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 304-332). Reston, VA: National Council of Teachers of Mathematics.
- Greeno, J. (2006). Authoritative, accountable positioning and connected general knowing: Progressive themes in understanding transfer. *Journal of the Learning Sciences*, 15(4), 537-547.
- Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and learning. In D. Berliner and R. Calfee (Eds.), *Handbook of Educational Psychology* (pp. 15-41). New York, NY: MacMillan.
- Lave, J. & Wenger, E. (1991). *Situated Learning: Legitimate peripheral participation*. New York: Cambridge University Press.
- Meadows, D.H. (2008). *Thinking in systems: A primer*. White River Junction, VT: Chelsea Green.
- McNeil, L. M. (2000). *Contradictions of school reform: Educational costs of standardized testing*. New York: Routledge.
- Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). *On the structure of educational assessment. Measurement: Interdisciplinary Research and Perspective*, 1(1) 3-62.
- Moss, P.A., Pullin, D. C., Gee, J. P., Haertel, E. H, & Young, L. J. (Eds.). (2008). *Assessment, equity, and opportunity to learn*. Cambridge: Cambridge University Press.
- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.
- Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.
- Schum, D. A. (1994). *The evidential foundations of probabilistic reasoning*. New York: Wiley.
- Stiggins, R. J. (2002). Assessment Crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83(10), 758-765.
- Turkle, S. & Papert, S. (1992). Epistemological pluralism and the revaluation of the concrete. *Journal of Mathematical Behavior*, 11 (1), 3-33.
- Whitehead, A. N. (1929). *The aims of education and other essays*. New York, NY: The Free Press.
- Zapata-Rivera, J. & Greer, J. (2004a). Inspectable Bayesian student modeling servers in multi-agent tutoring systems. *International Journal of Human-Computer Studies*, 61(4), 535-563
- Zapata-Rivera, J. D. & Greer, J. E. (2004b). Interacting with inspectable Bayesian student models. *International Journal of Artificial Intelligence in Education*, 14(2), 127-163