An Assessment Instrument of Technological Literacies in Makerspaces and FabLabs

Paulo Blikstein[a], Zaza Kabayadondo[b],
Andrew Martin[a], and Deborah Fields[c]

[a]Stanford University, [b]Smith College ,[c]Utah State University

Abstract

Background As the maker movement is increasingly adopted into K-12 schools, students are developing new competences in exploration and fabrication technologies (EFT). This study assesses learning with these technologies in K-12 makerspaces and FabLabs.

Purpose Our study describes the iterative process of developing an assessment instrument for this new technological literacy, the Exploration and Fabrication Technologies Instrument, and presents findings from implementations at five schools in three countries. Our index is generalizable, psychometrically sound, and permits comparison between student confidence and performance.

Design/Method Our evaluation of distinct technology skills separates general computing, information and communication technology (ICT), and exploration and fabrication technologies (EFT) into non-overlapping areas of technological expertise required to perform their respective sets of tasks. The instrument also tracks student confidence in EFT skills and assesses how that confidence relates to actual task performance.

Results EFT constitutes a new and distinct set of technology literacies arising from fabrication settings. The EFT instrument weighs students' self-reported confidence against their performance on complex design tasks and demonstrates that, for students, exposure to general computing and information and communication technology (ICT) tools differs from exposure to exploration and fabrication technologies (EFT) tools.

Conclusion The EFT instrument captures a new and distinct set of technology literacies that arise within fabrication settings and are independent of both general computing and digital content production skills.

Keywords: makerspaces; fablabs; design practice; assessment tools; project-based learning; constructivism, constructionism.

Introduction

Makerspaces and FabLabs bring together people, tools, skills, and resources in a single physical location for the purposes of designing, prototyping, and do-it-yourself manufacturing. These spaces can serve individuals, for-profit companies, nonprofit organizations, or educational institutions. They are collective organizations that share knowledge and provide open access to equipment, community, and education, in what is often referred to as "the democratization of invention" (Blikstein, 2013; Gershenfeld, 2008; Mikhak et al., 2002). In recent years, K-12 students have been invited to participate in these spaces in their schools, in afterschool programs, and at nearby universities or libraries. Participation in makerspaces and FabLabs promises to turn students into makers, creators, and innovators, and to increase interest in science, technology, engineering, arts, and mathematics (STEAM) fields. Martinez and Stager (2013) characterize these spaces as making possible a "technological and creative revolution" (p. 1), which has introduced difficult technological subjects such as engineering and design to students who traditionally might not have enrolled in courses in electronics, programming, or other technical fields. These spaces also pave multiple pathways to learning twenty-first-century skills by facilitating collaborative and iterative projects, and constitute a space where constructionist learning (Papert, 1980) can take place with unprecedented sophistication

Here we call these diverse spaces "digital fabrication facilities." They deploy a range of toolkits and machines for the pursuit of learning through making and construction, and are designed to encourage a broad range of social practices and pedagogical approaches that utilize contemporary fabrication technologies for design and prototyping (Blikstein, 2013; Kafai & Peppler, 2011; Remold, Verdugo & Michalchik, 2013; Worsley & Blikstein, 2014; Smith, Cheruvelil & Auvenshine, 2013). In this article, we describe our iterative development of and early findings on an assessment instrument intended to capture the learning that occurs in these settings.

Historical perspective
Fabrication settings inherit a rich theoretical tradition that reaches back to Dewey's practice-driven ideals (1916) and Piaget's constructivism (1929; 1974). Inspired by constructivist ideas, educators have advocated the creation of learning spaces and toolkits with the premise that learning is optimized when guided by student interests and abilities that correspond to their respective stages of development. Learners' ways of doing and thinking evolve over time, and fabrication activities and toolkits can be scaffolded to match these evolving stages. Papert's constructionism (1991) builds on Piaget and emphasizes the significance of externalizing one's ideas into objects that can be refined, debugged, shared within a community of practice. In this framework, self-directed learning occurs with the assistance of others and the availability of expressive tools, and mediated learning occurs when learners engage in rich conversations with others and with artifacts. Fabrication settings reflect the values of constructionism (Papert, 1991; Kafai, 2006) through their emphasis of the importance of creating and sharing personally meaningful artifacts in the context of a supportive community. By blurring the boundaries between disciplinary subjects, digital fabrication facilities introduce students to the expression, exploration, and design processes central to domains such as engineering, science, art, and crafts (Blikstein, 2013; Martinez & Stager, 2013).

Assessment of Technological Literacies in Makerspaces and FabLabs
Blikstein, Kabayadondo, Martin, & Fields (2017)

The rising popularity of digital fabrication facilities as sites of learning, exploration, and expression constitutes an acknowledgement within the curriculum of a new area of emergent technology skills that students must master. The acquisition of these skills is accompanied by the kinds of cognitive apprenticeship (Collins, Brown, & Holum, 1991) afforded by fabrication settings—such as evaluation, revision and creative planning. Lewis, Petrina and Hill (1998) state that such skills add an increment of creativity to technological problem solving. Through laser cutting, 3D printing, molding plastics, soldering and sewing circuits, students can learn to solve problems and troubleshoot, and they are asked to consider specific technical tasks, identify relevant scientific and design principles to be applied to the tasks, and devise solutions that can be implemented and modified in multiple iterations.

The ethos of personal creativity and learner-centeredness in digital fabrication facilities creates a dilemma for assessment. How are we to objectively assess work that is open-ended and creative? Traditional assessments of science and technology neither capture the particular types of learning in which students engage in fabrication settings nor reflect that learning of such settings is grounded in developing competence with digital fabrication tools. As Velayutham, Aldridge, and Frasier (2011) pointed out in the areas of science education, motivation is contextual (see also Bandura, 2006; Zimmerman, 1998; Barron, 2006). Educators who support the integration of fabrication settings into school do so because they find that such settings are providing transformative experiences for students (Blikstein & Worsley, 2016). They cannot, however, rely on existing determinations of learning outcomes to establish how well fabrication settings function pedagogically. Specific instruments are needed to quantify student confidence in their mastery of fabrication technologies, so these instruments must go beyond measures in science currently employed in general computing and in information and communication technologies (ICT). Designed to meet these complex requirements, our proposed instrument captures the relationship between students' confidence with technology and their performance on technology tasks specific to fabrication settings.

New Frameworks

While digital fabrication facilities are spreading to schools, several of the national frameworks for technological literacy do not offer appropriate models for tracking or accounting for the learning that goes on in these settings. Existing technology assessment instruments pose a threat to the future of teaching and learning in fabrication settings by compelling teachers and principals to align their activities with inadequate standards. Fabrication settings are opportunely situated at the intersection of national initiatives in support of engineering and technology education and local manifestations of constructionist learning, for which reason it is necessary, as other researchers have shown, to expand what constitutes evidence of learning in these settings (Petrich, Wilkinson, & Bevan, 2013).

We developed an assessment specifically designed for exploration and fabrication technologies (EFT): technologies centered on fabrication (activities oriented towards invention, construction and design) and those centered on exploration (activities oriented towards expression, tinkering, learning and discovery). Building on our ongoing research on the implementation of fabrication settings in schools (Blikstein, 2013), we designed two measures in a single survey to capture how fabrication settings shape students' understandings of and competencies with EFT. Having tested the survey in five different schools, we report on its development, its statistical validation, and our initial findings. The resulting instrument is intended for use by educators, researchers, and policymakers whose work must negotiate

between the requirements of national frameworks for technology literacy and a sensitivity to the novelty and uniqueness of the kinds learning that emerge within digital fabrication facilities.

Literature Review

Technology Literacy

In this first section, we provide a selective historical review of how learning and digital technology have been framed. Our intention is to focus on the evolution of how technology has been understood in the United States and the implications of these understandings have for our definitions of technology-relevant practices and competences.

Over the last two decades, the general consensus on technology literacy has evolved from quantifying student exposure to different digital software and hardware implementations to more comprehensive measures of fluency in technology. In our analysis, the development of technological literacy must begin with (1) exposure to technologies as a necessary starting point; (2) once students have gained sufficient exposure to a technology, they develop confidence, or self-efficacy, in their abilities to manipulate and work with that technology in the pursuit of other learning; finally, (3) students who have attained confidence must be able to demonstrate it through objective performance tasks. Even though the discourse on technology in the US academy has come to regard exposure, confidence, and performance as three important milestones in the achievement of technological literacy, they have often been conflated into a single measure. Also, as we will show in the following sections, the prevailing view of technology literacy and assessments treats all digital technologies as though all were equivalent. One of our broader goals with the EFT instrument is to investigate alternative frameworks for understanding the relationship between exposure, confidence, and performance, and how they could be different for diverse types of technologies. Our instrument makes separate assessments of confidence and performance to better articulate how the two are related to each other.

National Frameworks

As early as 1996, the national agenda on educational progress and student preparation for the twenty-first century placed technology at the center of its discourse. The International Technology and Engineering Educators Association (ITEEA) first prioritized the need to expose students to engineering and design practices for engagement in critical thinking through the design and development of "products, systems, and environments to solve practical problems" (ITEEA, 1996, p. 1). The Next Generation Science Standards (NGSS) identified a competency requirement for engineering design and systematic problem solving (NGSS Lead States, 2013). Most recently, the National Assessment of Educational Progress report (NAEP, 2014) has changed its goals from *technology literacy* to *technology and engineering literacy*, and made a case for integrating engineering practices and competencies into the assessment of technology fluencies. The report recommends a new national framework that adds emphasis on design and systems in technology.

FIGURE 1 HERE

The NAEP Technology and Engineering Literacy framework (NAEP TEL for short) emphasizes three distinct areas of technology and engineering practices: technology and society; information and communication technology (ICT); and design and systems (NAEP, 2014). The third category, design and systems, is the most relevant to fabrication settings because of its focus on engineering design, engineered systems, the physical characteristics and constraints of

technologies, and technology troubleshooting. This framework points to a need for assessments that identify students' ability to think about technology at the level of engineering systems. According to the NAEP TEL framework, students who have mastery of design and systems knowledge and practices should be able to critique, rationalize, and communicate design choices; plan, construct, and test prototypes; as well as troubleshoot malfunctions in technological products. Such competences are central to the activities in fabrication settings, yet assessments for them are still largely undeveloped.

Existing Technology Assessments

The NAEP TEL report proposes first to account for three different dimensions: technology *use*, *effects*, and *design,* in order to mitigate the discrepancies in the terminology used in various national and state policies. However, technology literacy assessments that have been built around the new NAEP TEL framework do not adequately account for the diverse range of practices and competencies that apply to different technology settings and bypass an entire category of tools capable of fostering engineering and design competence (i.e., exploration and fabrication technologies).

A major concern for national programs on technology literacy has been to account for the rapid proliferation of and continuously shifting contexts for technology use (for a comprehensive review of general technology assessment tools, see National Academy of Engineering and National Research Council, 2006). Our selective review of existing assessments will provide a more fine-grained discussion of what an EFT-specific literacy assessment should capture. In the end, we argue that ICT and digital media literacies involve different skills than those drawn from fabrication technologies and related practices and require their own sophisticated assessments.

Early ICT assessments were focused on self-reported knowledge of computer hardware use and a few programs such as email clients and web browsers. The technological scope of these assessments has been expanded to include web searches for information retrieval, spreadsheets for computation, and algorithmic skills (Chu, Tse, & Chow, 2011a; Conde et al., 2011), and three general skills areas have emerged: general computing tasks (saving files, adjusting privacy settings), communication and information retrieval (web search, videoconferencing), and digital content production (video and photography, spreadsheets).

Expanding the set of technologies considered for literacy has improved the ways that ICT assessments can evaluate students' literacy and have led to a more holistic approach to ICT focused on digital literacies (Green, Yu, & Copeland, 2014; Litt, 2013; Gençtürk, Gökçek, & Güneş, 2010; Werner, Denner, & Campe, 2006; Hargittai, 2005). This expansion has also provided the further benefit of redefining student learning goals as becoming "media literate" (Zylka, Muller & Martins, 2011). The discourse on technology literacy has focused mainly on capturing overlaps in the skill sets used in ICT with other fluencies, such as creativity, innovation, communication, collaboration, critical thinking, problem solving, and decision making (ISTE, 2007; NAEP, 2014). In their review of disciplinary and methodological issues in media and ICT, Livingstone, Van Couvering, & Thumin (2008) point to a convergence of media and ICT competencies that results from a rapidly expanding variety of digital tools for communication and information dissemination.

While some of the new assessments for ICT digital literacies may include technical skills such as critical thinking and understanding the social implications of technology, they do not focus on the tasks required for the actual design and construction of technological systems. The practices in fabrication settings involve developing a dynamic understanding of how computer

programming, design, materials, tools, and processes of fabrication merge in ways that generate rich opportunities for decision making, critical evaluation, and problem solving. Learning expresses itself in the creativity and diversity of student prototyping, troubleshooting, and iteration practices, and these are exactly the kind of competencies a technology literacy assessment should measure.

Technology literacy has been described as evolving along three axes: knowledge, critical thinking/decision making, and capabilities (Pearson, 2007). The development of technology assessments has occurred within two broad categories: (1) self-reported instruments on task values such as curiosity, perceived relevance, and confidence; and (2) the evaluation of performance in technology-driven tasks. We are chiefly interested in the future directions assessments could take to evaluate all dimensions of technology literacy, including confidence and performance.

One example of a more comprehensive confidence assessment was the iterative development of a survey of computer programming by Werner, Denner, & Campe (2006). In their survey of middle school girls in an afterschool computing program, the teachers reported that their students attained the highest fluency in communication and visualization technologies and the lowest in computer programming. A second example of a confidence assessment, a survey by Denner and Werner (2011) of computing interest among middle school Latina girls, examined frequency of use, parent and school support, computer confidence, perceived relevance of computing, and technological curiosity. Results showed that girls valued computing but expressed low confidence in future success and low technological curiosity. Instruments such as these can reveal disparities in youth's confidence with different types of technologies (e.g., communication technologies versus computer programming), as well as gender and socioeconomic status gaps that will result in unequal access to opportunities that require high levels of technological self-efficacy.

Several performance assessments on ICT and programming skills provide some promising models for developing EFT performance assessments. The most prominent example of a performance assessment is the Educational Testing Services (ETS) iSkills test (Lei, Shen, & Johnson, 2014). This test assesses objective task performance on critical thinking skills applied to digital information management. One of its main contributions is to remove the bias common to self-reporting on ability. Performance assessments have traditionally prompted students with computer science tasks such as programming, understanding of abstraction or modeling, and algorithmic thinking (Denner & Werner, 2011; Denner, Werner, & Ortiz, 2012; Huggins, Ritzhaupt, & Dawson, 2014). However, we could hardly find assessments of this level of sophistication created to measure students' performance in designing engineering systems (Miller, Webster, Dauer, & Anderson, 2013).

Our review showed one attempt to assess students' understanding of the interaction between digital software and hardware systems. Lebeaume and Perez (2012) asked a sample of Parisian students to describe the operation of two commonly encountered systems: a transit card that uses radio frequency identification and an interactive whiteboard. In their performance assessment, most students lacked the scientific and engineering vocabulary to describe the interaction of sensors, software programs, and actuators. The open-ended student responses revealed that, regardless of age, most students could identify, overall, only the most readily visible components of these technological systems. This instrument indicates a fruitful direction that EFT skills assessments could take.

While EFT confidence assessments are more common, some assessments, such as the one described above, have started to include EFT performance, but still none address both dimensions. Consequently, existing assessments do not provide common ground for determining the relationship between confidence and performance. As a corrective, our EFT instrument measures confidence and performance in parallel. The combination of our EFT performance assessments with EFT, ICT, and general computing confidence assessments represent a promising direction forward: our EFT instrument elaborates on design and engineering systems as a component of technological literacy, while also providing a platform for the comparison of the differences between general computing, ICT and the less understood EFT.

We acknowledge that our instrument and approach err, at times, on the side of technocentrism, in the sense put forward by Papert (1991). That is, some of our questions ask students directly about their familiarity with specific machines and technologies (e.g., laser cutters and 3D printers). We are aware that we could have asked questions that dealt more with projects that include laser cutting or more holistic activities that integrate more machines and technologies. Still, our rationale was that our first instantiation of the research should aim for simpler, technology-specific measures to support more complex measures in future versions of our instrument. Further, other assessments of maker-type knowledge have found it prudent to be context-specific. For instance, in creating an assessment for student understanding of basic circuitry (simple, parallel, and series circuits) in electronic textiles, Peppler and Glosson (2013) found it sensible to utilize images of the exact parts with which students had contextual experience. Using stickers with images of the sewable lights and battery holders, students drew working circuits. Our instrument is sensitive to the specificity of students' expertise with particular tools while also probing for more generalizable problem-solving skills.

Methods

The EFT instrument was developed iteratively and administered in four versions between fall 2012 and spring 2014 (see Table 1). In Versions 2, 3, and 4, the instrument was formulated as two separate measures: the confidence measure and the performance measure.

To permit generalizability, we have administered our instrument to a diverse set of students drawn from children in five schools. For the selection of the venues, we contacted various schools that had implemented some type of a digital fabrication environment. These venues are in three school districts within the United States, and the remaining two are in Mexico and Australia, both purposefully recruited as cross-cultural checks. The school-based digital fabrication programs included a wide age range of students, at the elementary and secondary levels, with varying socio-economic statuses. Some of these schools had been working with our research team in these implementations, and others had not. Schools implemented fabrication settings in different ways: from the integration of the spaces into traditional curricula to the provision of digital fabrication as an elective class. One reason for the diversity of schools was to ensure that the EFT instrument would work consistently across these differing conditions and would be more likely to function well in future applications at other school systems, both in the United States and abroad.

TABLE 1 HERE

At School A we tested three versions (1, 2, and 3), each during a different school year (2012, 2013, and 2014). Given the differences between the versions of the instruments utilized across this time span, we did not choose to match across administrations (although it would be possible for many students). At School B we tested Version 2 twice, administering the survey to the same seventh and eighth grade classes. This resulted in a matching for 71 students over both administrations of the instrument.

The EFT instrument was administered using both paper and electronic formats (the electronic version used the Qualtrics software platform). Student participation was not mandatory: students could opt-out, although we do not know how many chose this option for the paper survey administrations at Schools C and D. Survey completion ranged from 94% in School B and A (in 2013) to 77% in School E and 74% in School A in 2014. Our discussions with the classroom teachers administering the instrument with higher drop-out rates suggested that many students returned to complete the instrument at a later date, starting at the beginning each time. For example, at School A, grades 8 to 10 had 164 completions in 2013 and 162 in 2014, even though drop-out rates were much greater in 2014.

Completion of the instrument took between 20 and 45 minutes, depending on student grade level and variations in the number of questions in the different versions of the instrument. After approval from the administrators of participating schools, classroom teachers responsible for the fabrication settings at each school administered the instrument from written instructions included in the instrument package and with limited assistance from the research team. Question ordering was fixed for each version of the of the instrument. However, in the electronic versions of the instrument we did employ random orderings of items for questions that asked students to identify long lists of items (for examples, see below in Figure 2 and Appendix B).

Instrument Design

Our EFT instrument combines two measures: the EFT confidence and EFT performance measures. The first, EFT confidence measure, asks students to report on their familiarity and skills in key exploration and fabrication technologies (EFT) as well as common information and communication technologies (ICT). We deliberately included technologies from both EFT and ICT in our instrument because the relationship between the two remains largely unexplored and may be more clearly characterized with an instrument measuring confidence in both areas. The second measure, which assesses EFT performance, asks students first to identify major parts of commonly encountered electronic devices, redesign them, and finally to troubleshoot them. We developed an objective score from these in order to assess what students know about the design and systems interaction of components of everyday devices. The instrument was designed to be generalizable and to have sound psychometric properties. Table 2 summarizes the category of questions in the survey.

TABLE 2 HERE

The EFT instrument also includes a number of additional questions about student background to assist in validation and to facilitate comparison across the participating schools.

Background information on family socio-economic status and parents' educational attainment were incorporated into the instrument, as discussed earlier (Table 1). Student-level background questions included queries on grade level, gender, STEM classes taken, participation in after-school STEM programs, career aspirations and post-school plans, as well as estimates of time spent on different student activities, such as studying, sports, and reading. Below, we outline how the iterative design of the two measures was guided by validity testing throughout the different versions of the assessment.

EFT Confidence Measure

Iterative Instrument Design
The EFT confidence measure captured students' confidence and familiarity with a variety of digital technologies. Respondents were asked to rate their competence with items on a 6-point Likert-type scale, where 1 indicated "I know nothing about it" and 6 indicated "I could be a professional or teach other people about it" (see Figure 2). In the online version of the instrument, the order of questions was randomized where possible to reduce ordering effects.

FIGURE 2 HERE

After administering Version 2 of the survey at School A (spring 2013), we ran an exploratory factor analysis of student responses in the EFT confidence measure. Figure 3 provides the results of the factor analysis. The magnitude of eigenvalues (y-axis) from the instrument compared with randomly reshuffled data suggests that the instrument consists of three related factors. To check whether these factors were closely related or could be treated as independent, we next ran a factor analysis with a nonorthogonal rotation (Oblimin) and a principal component analysis (PCA) employing an orthogonal (Varimax) rotation. With the nonorthogonal rotation of the data, we found nonnegligible correlations between the factors, which indicates the factors are strongly correlated. Distribution into three correlated factors was consistent across subsequent versions of the EFT confidence measure. The three factors could be clearly mapped with distinct types of technologies, thus we named them *general computing tools*, *ICT production tools*, and *EFT production tools*.

FIGURE 3 HERE

To refine the EFT confidence measure in subsequent versions, we added and deleted items in two steps. The first step was prioritized because of our moderate sample sizes for the different versions of the measure (ranging from 80 to 300 students in the five schools). Following the convention of loading factors greater than 0.4 (Yoon Yoon et al., 2014; Ro et al., 2015), we flagged items for deletion if they weighed less than 0.4. If any item loaded >0.35 on more than one of the three factors, it was removed from the instrument to ensure that the remaining items were unique to each category. (The table of loading values for the second and third versions of the instrument can be found in Appendix A.) We found that computer programming tended to load strongly on both the ICT production tools and EFT production tools – a result we attribute to the intrinsic complexity of the programming process. Like reading or writing, programming underlies a number of complex skills and tasks that have a wide-range of applications in production activities, such as programming scripts for robots and webpages,

which would be relevant as both EFT and ICT production skills, respectively. The latest version of the measure specifies programming for EFT production with the item "programmable microcontrollers" that has consistently loaded on only EFT production tools.

The second refinement step involved adding more items, specifically production tools, to the three subscales. Earlier versions of the confidence measure had fewer items than subsequent revisions because we did not initially anticipate three subscales. As an improvement, we added new features that included confirmatory factor analysis (CFA) and classical test theory measures for reliability. We used weights from the primary principal component or core of each subscale to screen the new items and employed Cronbach's alpha (Table 3) to assess the reliability of the three factors. Given that our instrument is not intended for ability testing, reliability values of $\alpha < 0.7$ were deemed poor, 0.7 or 0.8, acceptable, and those greater than 0.8 (DeVellis, 2016), or in the range 0.85 to 0.90, excellent (DeVellis, 2016). The revision of the instrument brought up consistency in ICT production tools from 0.59 and 0.53 to 0.83 and 0.73. EFT production tools remained consistently high at 0.87 and 0.85.

TABLE 3 HERE

Instrument validation is a never-ending process; however, several findings make us confident that the EFT confidence measure is assessing what we intended. We adopted an approach to validity that recognizes that many potential threats to validity arise during test application and development (Crooks, Kane, & Cohen, 1996). The use of Likert-type scales, the online format, and scoring, along with our efforts to make administration easy and consistent, all deal with methodological threats that come from the survey administration and scoring protocol in our methods. Our use of diverse trial populations along with systematic psychometric methods also gives us confidence that threats to validity arising from aggregation and reliability have also been adequately addressed.

Validity Testing
We evaluated the validity of the EFT confidence measure by correlating it to other tests and also confirmed that scores change with variations related to grade level and curriculum exposure (Borsboom, Mellenbergh, & van Heerden, 2004; Lissitz & Samuelsen, 2004). For example, School A has a substantial engineering curriculum already in place that utilizes a digital fabrication facility. We predicted that the 34 students (out of 308) who were enrolled or had completed the engineering course, the robotics club, or computer science courses would score higher on confidence with EFT and ICT tools because these students rely heavily on the school's digital fabrication facility to support their club and class projects. Throughout this article, we employ a robust version of two-way ANOVA with an alpha of 0.05 because these estimates tolerate moderate departures from normality and homoscedasticity while retaining good power (Koller & Stahel, 2011). After controlling for grade level, the ANOVA confirms that students with engineering course experience at School A reported higher EFT and ICT confidence than their grade counterparts at School A (effect of 1.21 points with $p < 0.0001$ and effect of 0.38 points with $p = 0.056$, respectively).

TABLE 4 HERE

Version 2 of the confidence measure was administered as a pretest to the seventh and eighth grades at School B at the beginning of the school year. Version 3, which contained some additional items, was administered as a posttest shortly after the students' completion of a digital fabrication curriculum intervention that academic year (see Table 4). A paired Wilcoxon test (with the Holm multiple comparison correction applied and alpha of 0.05) suggests that the student's increasing confidence with EFT production tools of one point in our 6-point scale for both grades is statistically significant. Again, we chose this test over the more common paired t-test because it is nonparametric and robust to departures from normality and homoscedasticity. Students in both grades made significant gains on the EFT confidence instrument and the curriculum in either case involved design or engineering related to digital fabrication. However, we were surprised at the significant gains made by the eighth grade as measured by the General Computing and ICT confidence items. Our investigation suggests that these results may be due to curricular interventions at School B, namely the distribution of iPads to both grades and, in the case of the eighth grade, a media literacy curriculum. Perhaps more importantly, the different gains by students demonstrate that the confidence measure as a whole captures differences between student confidence in three areas of technology literacy that previously were not well differentiated. Although we did not find universal improvement in all three dimensions simultaneously, we did observe that it is possible for students to improve only in the EFT dimension while demonstrating no improvement in the two other dimensions.

EFT Performance Measure

Design

As with the confidence measure, the questions in the EFT performance measure underwent several revisions in order to improve its test properties. A number of challenges presented themselves, in part because there were still only limited examples to model from. We needed to develop entirely new questions to assess learning that may have been facilitated by a variety of technologies in different curricular environments. These questions could not take more than a few minutes to complete, and because of the diversity of technical infrastructure in the schools, the instrument had to work both on paper and electronically. As a result, developing questions that went beyond probing simple factual knowledge turned out to be quite complex. Iterative changes to the instrument included not only alterations in wording but required wholesale question additions or deletions as well to produce a set of questions that met basic validity thresholds. We tried many different kinds of questions. Some did not appear to measure EFT skills and knowledge consistently, some were not correctly understood by students, and some were difficult to score consistently. One example of an unsuccessful question involved an open-ended question about fixing a broken coffee maker that prompted students for their solution of how to fix it. It was a difficult question to score reliably, and the scoring was very time consuming and thus unfeasible for large-scale assessments. Answers to this broken coffee maker question never correlated with other performance measures, nor did they pass measures of external validity. After testing that question through two revisions at three schools, we finally removed it.

We developed three overarching questions, each with many component parts. In order to simplify and improve the reliability of scoring, we developed a combination of short-answer and binary-choice questions in contrast with the open-ended approach used by Lebeaume and Perez (2012). The questions we developed used two different strategies: some questions required

students to analyze an item and identify key component parts; other questions prompted students to explain which parts and tools were needed to create a specified electrical appliance or electronic device. Both question types invited students to demonstrate their knowledge of common electronic devices as simple engineering systems.

Validity Testing

Self-selecting groups such as the engineering course students from School A are valuable for estimating the validity of the EFT performance measure. As predicted, School A's engineering students outperformed their non-engineering course peers on total scores for each question. Student engineering background was a statistically significant predictor after controlling for grade-level of performance on both the key-fob and the blender questions, per an ANOVA procedure similar to that used in validating the EFT confidence instruments (effect size of 1.35 points for the key-fob and 0.39 points for the blender with $p=0.0168$ and $p=0.0176$ respectively). These results suggest that the questions are sensitive enough to determine differences in student understandings of basic electronics.

Results

EFT Confidence
Measure Scores
Students were generally quite confident in their self-reported ability to use general computing tools, but reported lower confidence in the ICT and EFT production tools, as indicated in the mean confidence scores for the eighth grade in schools A, B, and E (Table 5). The results validate our purposeful differentiation between confidence in general computing, ICT, and EFT.

TABLE 5 HERE

Test scores revealed that students lacked confidence in EFT production tools compared with their confidence in general computing and ICT. We attribute the high general computing and ICT scores to a pervasive social emphasis on information consumption and ICT content, especially because all three schools (A, B, and E) have well-funded technology programs and serve students from high SES backgrounds. Despite the maker movement's increasing popularity and intentional efforts by the schools to develop curricula for digital fabrication labs and engineering courses, student confidence with EFT production tools are overshadowed by their confidence with more ubiquitous digital tools (general computing and ICT). It is unclear why School B has a higher EFT score than the others. Whatever the reasons for the differences, the confidence measure captured differences between schools, which indicates that it should allow for a more systematic evaluation of student confidence, and capable of supporting deeper research comparisons of different technology-based programs.

EFT Performance
Measure Scores
In analyzing the EFT Performance measure scores, we consider responses to two key questions. First, the key-fob question was first introduced in the second version of the instrument and was revised heavily in the third and tested on the large sample of 324 students from School A. In order to map out the correlation between confidence and performance, we juxtapose the results

of the key-fob question with the results from the EFT confidence measure and with a second EFT Performance question (the blender question, also introduced in the third version of the performance measure and also tested on School A). The blender and key-fob questions were the most "dialed-in" performance-based questions that we developed through our iterative design-validate-and-refine process.

Key-Fob Question The key-fob question begins by presenting students with a photo of a car key-fob with four buttons such as lock and unlock (see screenshot of question in Appendix C). A key-fob comprises a battery-powered circuit board, an antenna printed directly on the circuit board, a radio transmitter, and a microcontroller, along with some mechanical switches that are activate when buttons are pressed. The question then lists a number of mechanical and electronic parts and asks students (1) whether or not each part is present in the key fob, (2) how sure they are of their answer on a scale of 1 to 10, and (3) if they know what that part is. The "sureness" rating allowed us to diagnose item difficulty and to track changes in student confidence over time, and across different groups (not be confused with the EFT confidence instrument). Asking whether students know a part helps to distinguish between knowledgeable responses and guesses. Each item listed was worth a single point, but the point was only awarded if the student both recognized the component and correctly indicated whether or not it was actually present on the device. The sum of these scores was the overall question score. The combination of predicting the contents of the key fob and recognizing electronic components increased the difficulty of the question and led to a wider variation in total scores.

TABLE 6 HERE

The results of the key-fob question suggest that most students have only a rudimentary knowledge of how the device functions. They understood the need for power from batteries, and they knew lasers and microphones were not necessary components. While two-thirds of the students recognized the need for a radio transmitter, few realized the importance of an antenna or microcontroller (see Table 6). The most difficult items show a much higher percentage of students correctly answering whether it was present or absent in the key fob (57% for the microcontroller) versus indicating that they recognize the part and its function (19%). The discrepancy in item scores allows us discriminate which students answered that a key fob has a microcontroller by guessing.

FIGURE 4 HERE

Analyzing both students' knowledge of the component, and the correctness of the answer illuminates some meaningful issues. From Figure 4, it is noticeable that the components which are closer to everyday experience, such as microphones, batteries, motors, switches, screws, lasers, have the lowest discrepancy between students' knowledge about them and correct identification—sometimes very close to zero. At the same time, components that are more hidden from everyday experience, such as microcontrollers and transistors, exhibit a very high discrepancy (38% and 25% respectively). If our measure had only asked about the presence of the component in the key-fob we would be led to believe that 50% of students actually knew what a microcontroller is. However, since we also asked if students knew the function of the component and whether or not it was present in the device, we could determine that a large percentage of students guessed correctly. To a lesser degree, the same happens with transistors:

we would be led to believe that 72% of students know what transistors are, but one third of those, in fact, merely may have been guessing.

The discrepancies in answers to these questions, however, do not always deliver such clear differentiations between mere guesswork and a basic understanding of engineering principles. For example, 88% of students properly understood the nature of an antenna, but only 19% correctly answered the question. We suspect that students are conceptualizing antennas in terms that are familiar to them: as cylindrical metallic objects that protrude from electronic devices. Since the key-fob is small and does not have any such protrusion, most students got that part wrong. One implication is that students might conceptualize electronic devices by their appearance rather than their functionality. For example, they may not have associated the fact that the key-fob has to transmit a wireless signal to the necessity of an antenna's presence in the device. After an electronics or digital fabrication workshop, the key-fob question can measure to what degree students have relevant knowledge of key components of typical engineering systems and how these components work together in the system—regardless of students visually identifying the classical versions of these components in the device (such as an antenna.)

Blender Question The second question was about a blender, which prompted for semi-structured responses focused on electrical engineering systems (see Appendix D). The question required students to describe the parts and fabrication steps involved in the construction of a simple kitchen blender. The first section asked students to list the major parts: motor, blades, pitcher, lid, and power cord. Each named part was worth 0.2 points. The next part of the question asked students to describe the fabrication steps needed for the construction of the plastic base of the blender – 3D printing, vacuum forming, or casting a mold were methods students could describe to score a point. An extra point was awarded for procedural knowledge of specific techniques; for instance, if a student specified that 3D printing first required designing a model, or that a mold must be pre-shaped to specification. The final section asked students to correctly identify soldering as the ideal method for connecting the circuitry. The results of the scoring of the blender question are discussed in the next section (Patterns across Measures).

Patterns across Measures
Taken together, the confidence measure and the two main questions on the performance measure suggest three main trends. First, we notice low levels of EFT literacy, especially regarding student understanding of the use of tools in the construction of elementary mechanical devices controlled by programmable electronics. Second, students consistently rated their own confidence with EFT tools considerably lower than general computing tools or ICT tools. Finally, students could identify surface features of devices and name the parts required for the common devices to functions, but they demonstrated limited awareness of how such devices actually work.

TABLE 7 HERE

Table 7 is the pairwise correlation matrix between the key-fob and blender questions and three factors from the EFT confidence measure. Student performance was not strongly correlated to students' self-assessed skills (confidence). Previous research has already found that across many different domains, survey respondents are poor judges of their own ability in tasks that are

moderately or very difficult (Lichtenstein & Fischloff, 1977; Koriat, Lichtenstein, & Fischloff, 1980; Griffin & Tversky, 1992), and assessments have moved away from self-reporting. In the Discussion section, we propose interpretations of the relationship between student confidence and performance.

Discussion

Accessibility to Technologies

One finding from the analyses of our EFT Instrument is that differences in the accessibility and nature of various technologies are reflected in our data. First, our statistical analysis demonstrates that exposure to general computing and to ICT tools is fundamentally different from exposure to EFT tools. Additionally, self-reported confidence in ICT was not a good predictor of student EFT task performance. These data contradict both the notion that "digital natives" are without qualification well versed in all types of technology and the notion that a sequential or spontaneous relationship exists between exposure, confidence, and performance for all types of technology.

One implication of this is a need to reexamine how confidence and performance might vary for different types of technologies. The way youth get exposed to general computing, ICT, and EFT technologies differ greatly. For many, general computing and, more recently, ICT have become a part of everyday activities. There are many opportunities for daily exposure to general computing and ICT tools in activities such as those involving mobile phone, tablet, and computer use, and the school's role in exposing students to these tools is less crucial. Students encounter these devices in a variety of settings, when they are not carrying one with them at all times. Furthermore, these tools have benefitted from decades of improvements in usability, human-computer interaction, the scaffolding of user-generated content, and peer-to-peer support, so their mastery requires little formal instruction. Students can discover and learn on their own and move from exposure to confidence and to performance more easily. On the other hand, familiarity with fabrication technology (EFT) is deeply contextualized and constrained to digital fabrication facilities. While EFT tools have moved into libraries and are increasingly accessible for some students at commercial 3D print shops and maker supply stores, for most, there are few opportunities for exploration with EFT. Thus for many, school remains the only place that could offer sustained access to EFT technologies. Likewise, since the software and hardware resources are relatively new products and their usability and support are still rudimentary, scaffolding and instruction remain crucial for developing competence. Since it is much more challenging for novices to master such technologies on their own, facilitation and help from more able peers or teachers is still crucial.

General or Specific Technologies?

Another common concern with technology assessments is insuring their relevance in the face of rapid technology obsolescence. This means that effective assessments should be neither too general nor too specific to the technologies currently popular. The challenge with many existing technology literacy assessments has been finding the middle ground between generality, which does not capture the range of technology literacies (such as the difference between EFT and ICT) and specificity, which tends to increase the chances an assessment will quickly become obsolete as the popularity or relevance of certain technology tools phases out over time. The struggle to

find the right level of generality for items in our assessment was one of the major reasons for its multiple revisions.

For example, a very early version of the assessment contained an item that asked students to rate their familiarity with "electronics." When we administered the question to students (and to teachers, as part of preliminary prototyping of the instruments) we found that for students, electronics correlated closely to robotics and shop tools, while for teachers, electronics correlated with smartphones and tablets. It was clear that the electronics category, much like the term technology, was too broad. We removed the item and instead focused on specific skills and tools, such as soldering and using multimeters, which are needed to build and diagnose electronic circuits. The risk of our approach is too much specificity. Recall the key-fob question where 57% of students correctly stated that a microcontroller was present but only 19% of students actually felt confident that they knew what a microcontroller was. We suspect that using more recognizable phrasing such as "computer chips" might allow students to be more confident in their answers and to correctly determine a microcontroller's presence without resorting to guessing.

Given the risk of obsolescence, the challenge of finding the balance between generality and specificity pointed us to a larger issue concerning the ability of technology literacy frameworks to keep up with the emergence of new technologies. Since EFT tools are evolving rapidly, there is a constant need to update EFT assessments, including our own instrument. For example, recently, popular media have proclaimed that 3D printing's projected societal importance make it a key technology that should be taught in schools. Therefore, developing an instrument to measure 3D design and printing skills would follow as a natural part of a technology assessment questionnaire. However, since its commercialization and entry into the mass market in 2013, 3D printing has proved to be less reliable and more complex than expected. Software environments for novice users are still not well developed, and the introduction of 3D printing capabilities in schools is still experimental with no universal standards for implementation or evaluation. It remains unknown how important 3D printing will be for a K-12 technology assessment instrument.

Generating categories that are specific enough to have consistent meaning while not being so specific that they make the instrument irrelevant in the future is a challenge. Using a midlevel approach by sticking to classes of devices while focusing specific processes of technology skills was successful for our EFT instrument. Our multiple design cycles allowed us to detect and address the problem for the electronics item, and we expect to be able to do the same for microcontrollers and computer chips in future iterations.

Confidence versus Performance
Our analysis showed a marked difference between students' confidence in EFT and their performance in EFT. One plausible reason that students' EFT confidence may be high while performance is relatively low is that there are a number of different EFT technologies currently available. Because of the wide range of design projects students take on and their differing approaches to prototyping, different levels of familiarity and specialization with specific technological tools becomes a factor. Consider, for example, a design challenge for students to build a functional floating vessel. Students may start with blueprints and cardboard prototypes before building and testing their designs. While one student might plan to build the vessel structure by laser cutting acrylic sheets; another might consider polymer casting; while yet another might 3D print a model. Each approach will lead to more experience with the specific

planning, tools, and materials related to the respective EFT production tools. Each particular specialization emphasizes mastery of one particular subset, frequently to the neglect of others that are less useful for a project. Thus, students' actual expertise may be limited to specific areas, leading to students performing solidly on some performance tasks, but not on all.

At the same time, because they are part of a community in a digital fabrication facility, students' physical presence in fabrication labs means that they are exposed to a variety of production tools, at least passively. They may become aware of the expertise that others are developing, and are also exposed to the variety of machines and general techniques second-hand, by debugging and corroborating with peers. This general knowledge of the range of machines, techniques, and skills used in exploration and fabrication technologies may be one factor that leads to student confidence being high while their performance is comparatively low. It may reflect the difference between knowing about technology and being able to use it effectively.

Distributed Expertise

While the variations in learning among students presents a challenging scenario for an assessment instrument, such differences constitute an exciting development for educators because they provide opportunities for peer pedagogy and models of leadership where everyone has knowledge to share (Fields, 2009; Gee, 2004). In these settings, the nature of learning is not located in a single individual but across individuals as they share knowledge and solve problems. Thus, distributed expertise, rather than being an inconvenience for assessment, should be seen as possibly a new standard by which we could evaluate learning when students acquire overlapping but different knowledge. These peer learning interactions challenge researchers to look not only at what students can accomplish on their own but what they can achieve among themselves (Vygotsky, 1978) when they become resources for each other in accomplishing their creative work. As much as the distributed nature of expertise disrupts the conventional schooling notion of each student learning the same thing at the same time, it is a crucial and urgent issue for researchers to deal with.

The distributed nature of expertise suggests that the way students become conscious of their familiarity and expertise in EFT tools is more complex than a linear path. Expertise is acquired through a dynamic and emergent process that relies on both tools and peers that are engaged in multiple iterative cycles of debugging, exploration, and the development of mastery. Thus, future versions of our instrument should also find ways to assess students' perceived difficulty in acquiring expertise in EFT tools, and not just the current level of their expertise. We believe there is still much work to be done to improve our EFT instrument, and we hope to report further results in our future research.

Conclusion

Technology and Engineering Literacy

This article opened with a discussion of the shortcomings of the perception and framing of technology in the national discourse on STEM education. While there is a new emphasis in the national goals of education to consider "technology and engineering literacy" (NAEP, 2014), the implementation of technology in the curriculum is largely restricted to teaching general computing and ICT, with little attention to EFT. Yet fabrication settings have become an integral part of many (though far from most) schools around the United States and the world. The attitude towards experimentation and exploration engendered in them has the potential to transform how

schools approach STEM learning. However, given the limited view of the technology and its complex possibilities in the current assessments available, which largely focus on general computing and ICT, the kinds of learning opportunities that arise in digital fabrication facilities cannot be properly measured and characterized in terms of their essential contributions to overall technology and engineering literacy.

Assessments play a key role in helping educators to align fabrication activities and opportunities with explicit learning goals. The development of the EFT instrument that we presented in this paper marks one effort toward distinguishing the specialized knowledge and practices necessary in learning with EFT tools from that of general computing tools and ICT production tools. In reflecting the kind of learning that actually takes place in fabrication settings, our instrument provides schools with strategies to measure learning and incorporate the results into improved curricula and guidelines for learning. Ideally our instrument will become one of a set of assessments that can provide feedback in setting national policy and curricula regarding the role of fabrication settings in technology and engineering literacy.

Broader Implications

The EFT instrument we developed captures a new and distinct set of technology literacies that arise within digital fabrication facilities and are independent of both general computing and digital content production skills. Our instrument also weighs students' self-reported confidence against their performance on complex design and systems tasks and provides insight into the relationship between confidence and performance. The instrument has also set us on the path to a more detailed understanding of the different ways that expertise develops across different types of technology: an agenda that must be carried forward in future research. For technologies involving basic computer use, which in many respects are already integrated into everyday life, the school's role may not be as crucial as it once was. However, new and emerging EFT tools are far from being common objects in children's lives. EFT tools are still in their "exposure" phase when children are being made aware of the range of tools and what they can be used for. Children build confidence only when they are deeply involved in environments in which these technologies are made accessible to them; and settings such as digital fabrication facilities provide opportunities for broad exposure and the acquiring of specific expertise, as students use tools to collaboratively create projects.

Alongside the need for greater exposure to EFT tools is the necessity of providing means to develop competent performance with these tools. Our instrument raises questions about whether performance naturally develops as a natural consequence from simply gaining exposure and progressing in confidence, a commonly held belief in regard to the progression of technology literacy. A common belief is that children learn to interact with ICT technology as autodidacts. We hypothesize that autodidactism is only possible today because of a combination of the ubiquity of ICT tools in everyday life, and decades of advances in human-computer interaction and interface design. It would be a mistake to assume that children's autodidactism would automatically extend to EFT technologies. Therefore, until the machines, kits, and software used in digital fabrication facilities reach the same level of usability as modern apps and software, educators should budget extended periods of time for students to become proficient with EFT tools with the help of experienced facilitators.

A new technology assessment instrument requires several cycles of development, and longitudinal data. We believe that this article contributes to the creation of assessment tools that better characterize and assess the learning that takes place in digital fabrication facilities. Our

Assessment of Technological Literacies in Makerspaces and FabLabs
Blikstein, Kabayadondo, Martin, & Fields (2017)

EFT instrument allows an examination of the outcomes of the many different kinds of activities occurring in such spaces with a careful attention to unique aspects of this emerging technology that would be missed entirely by assessments designed for other practices, technologies, and types of learning.

Appendix A

Loadings For Factors Second Version of the Instrument at School A.

2013 factor loadings from all items included in second version. Individual items were also evaluated on a single factor with their neighbors as a further check on validity. Example, blogging and video editing lie on a factor explaining 73% of their common variance.

Item	Factor 1	Factor 2
Computers	0.26	0.76
Smartphones	0.07	0.89
Tablets	0.34	0.77
Blogging	0.19	0.3
Editing digital videos	0.23	0.3
Programming robots	0.9	0.07
Building robots	0.93	0.11
Computer programming	0.67	0.11
Soldering	0.84	0
Fixing stuff	0.71	0.05
Shop tools	0.79	0.06
Laser cutters	0.78	0.12

Assessment of Technological Literacies in Makerspaces and FabLabs
Blikstein, Kabayadondo, Martin, & Fields (2017)

Appendix B

2014 Factor Loadings from Third Version of Instrument at School A

Not all items in list used to score instrument or to score sub-scales such as carpentry or the item about upgrading a hard drive or RAM

Item	Factor 1	Factor 2	Factor 3
PC or Mac computers	0.88	–0.01	–0.2
Smartphones	0.83	0.03	–0.22
Tablets	0.66	0.09	–0.09
Blogging	0.52	–0.23	0.28
Programming robots	–0.05	0.81	0
Building robots	0.05	0.89	–0.11
Shop tools	–0.11	0.27	0.69
Laster cutters	0.04	0.61	0.07
Carpentry	–0.01	0.22	0.57
3D printers	0.11	0.64	–0.01
Soldering	–0.11	0.4	0.44
Creating webpages	0.42	0.16	0.08
Word processing	0.65	–0.07	0.01
Using spreadsheets	0.52	0.17	0.06
Creating presentations	0.65	–0.08	0.14
Cordless drill	–0.19	0.18	0.79
Editing digital videos	0.58	–0.11	0.03
Multimeter	–0.02	0.54	0.2
Upgrading a hard drive or RAM	0.39	0.17	0.13

Appendix C

**Screenshot of Part of the Key-Fob Question Formatted
for Online Administration**

FIGURE APP C HERE

Appendix D

Screenshot of the Blender Question Formatted for Online Administration <mark>of the Performance Instrument</mark>

FIGURE APP D HERE

Acknowledgments

Assessment of Technological Literacies in Makerspaces and FabLabs
Blikstein, Kabayadondo, Martin, & Fields (2017)

References

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307–337). Greenwich, CT: Information Age.

Barron, B. (2006). Interest and self-sustained learning as catalysts of development: A learning ecology perspective. *Human Development, 49*, 193–224.

Blikstein, P. (2013). Digital fabrication and 'making' in education: The democratization of invention. In J. Walter-Herrmann & C. Büching (Eds.), *FabLabs: Of machines, makers and inventors* (pp. 203-221). Bielefeld: Transcript Publishers.

Blikstein, P**.** & Worsley, M. (2016). Children are not hackers: Building a culture of powerful ideas, deep learning, and equity in the Maker Movement. In K. Peppler, E. Halverson, and Y. Kafai (Eds.), *Makeology: Makerspaces as Learning Environments (Volume 1)* (pp. 64-79), New York, NY: Routledge.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological review*, *111*(4), 1061-1071.

Chu, S. K. W., Tse, S., & Chow, K. (201la). Using collaborative teaching and inquiry project-based learning to help primary school students develop information literacy and information skills. *Library & Information Science Research, 33*(2), 132-143.

Collins, A., Brown, J. S., & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American Educator*, *15*(3), 6-11.

Conde, M., Miguelanez, S., Molina, M., Abad, F., & Riaza, B. (2011). Informational literacy and information and communication technologies use by secondary education students in Spain: a descriptive study. *Contemporary Issues in Education Research (CIER), 4*(4), 1-12.

Crooks, T. J., Kane, M. T., & Cohen, A. S. (1996). Threats to the valid use of assessment. *Assessment in Education, 3*(3), 265-285.

Denner, J. & Werner, L. (2011). *Measuring computational thinking in middle school using game programming*. American Educational Research Association Conference, New Orleans, LA.

Denner, J., Werner, L., & Ortiz, E. (2012). Computer games created by middle school girls: Can they be used to measure understanding of computer science concepts? *Computers and Education, 58*(1), 240-249.

DeVellis, R. F. (2016). *Scale development: Theory and applications* (Vol. 26). Sage Publications.

Dewey, J. (1916). *Democracy and Education.* New York, NY: The Free Press.

Fields, D. A. (2009). What do students gain from a week at science camp? Youth perceptions and the design of an immersive research-oriented astronomy camp. *International Journal of Science Education, 31*(2), 151-171.

Gee, J. P. (2004). *Situated language and learning: A critique of traditional schooling.* New York, NY: Routledge.

Gençtürk, E., Gökçek, T., & Güneş, G. (2010). Reliability and validity study of the technology proficiency self-assessment scale. *Procedia-Social and Behavioral Sciences, 2*(2), 2863-2867.

Gershenfeld, N. (2008). *Fab: the coming revolution on your desktop--from personal computers to personal fabrication.* New York, NY: Basic Books.

Green, J., Yu, S., & Copeland, D. (2014). Measuring critical components of digital literacy and their relationships with learning. *Computers and Education,* 76, 55-69.

Griffin, D. & Tversky, A., (1992) The weighing of evidence and the determinants of confidence. Cognitive Psychology, 24, 411-435.

Hargittai, E. (2005). Survey measures of web-oriented digital literacy. *Social Science Computer Review, 23*(3), 371-379.

Huggins, A. C., Ritzhaupt, A. D., Dawson, K. (2014). Measuring information and communication technology literacy using a performance assessment: Validation of the student tool for technology literacy (ST^{2}L). *Computers and Education, 77,* 1-12.

International Society for Technology in Education. (2007). National Education Technology Standards for Students. www.iste.org/Content/NavigationMenu/NETS/ForStudents/NETS_for_Students.htm.

International Technology Education Association. (ITEA/ITEEA). (1996). Technology for all Americans: A rationale and structure for the study of technology. Reston, VA.

Kafai, Y. B. (2006). Constructionism. In R. K. Sawyer (Ed.). (2006). *The Cambridge Handbook of the Learning Sciences* (pp. 35-46). New York: Cambridge University Press.

Kafai, Y. B., & Peppler, K. A. (2011). Youth, technology, and DIY developing participatory competencies in creative media production. *Review of Research in Education*, *35*(1), 89-119.

Koller, M., & Stahel, W. a. (2011). Sharpening Wald-type inference in robust regression for small samples. *Computational Statistics and Data Analysis*, *55*(8), 2504–2515.

Koriat, A., Lichtenstein, S., & Fischloff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 107-118.

Lebeaume, J. & Perez, W.G. (2012). How do the interactive white board and the radio frequency identification and tracking system work? Exploration of pupils' spontaneous knowledge and didactical proposals for technology education. In *PATT 26 Conference; Technology Education in the 21st Century; Stockholm; Sweden;* pp. 293-300. Linköping University Electronic Press.

Lei, J., Shen, J., & Johnson, L. (2014). Digital technologies and assessment in the twenty-first-century schooling. *Contemporary Trends and Issues in Science Education, 41,*185-200.

Lewis, T., Petrina, S., & Hill, A. M. (1998). Problem Posing--Adding a Creative Increment to Technological Problem Solving. *Journal of Industrial Teacher Education*, *36*(1), 5-35.

Lichtenstein , S. & Fischhoff , B. (1977). Do those who know more also know more about how much they know? The calibration of probability judgments. *Organizational Behavior and Human Performance, 20*, 159-183.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The Concept of Validity. *Psychological Review, 111*(4), 1061-1071.

Litt, E. (2013). Measuring users' internet skills: A review of past assessments and a look toward the future. *New Media & Society, 15*(4), 612-630.

Livingstone, S., Van Couvering, E., & Thumin, N. (2008). Converging traditions of research on media and information literacies. *Handbook of research on new literacies*, 103-132.

Martinez, S. L., & Stager, G. S. (2013). *Invent to learn: Making, tinkering and engineering in the classroom.* Torrance, CA: Constructing Modern Knowledge Press.

Mikhak, B., Leon, C., Gonton, T., Gershenfeld, N., McEnnis, C., & Taylor, J. (2002). Fab Lab: An alternative model of ICT for development. *ThinkCycle: Development by Design (dyd02)*. Bangalore, India.

National Assessment of Educational Progress. (2014). Technology and engineering literacy framework for the 2014 NAEP. https://www.nagb.org/publications/frameworks/technology/2014-technology-framework.html

National Academy of Engineering and National Research Council. (2006.) Tech Tally: Approaches to Assessing Technological Literacy. Washington, DC: The National Academies Press.

NGSS Lead States. (2013). Next Generation Science Standards: For States, By States. http://www.nextgenscience.org/

Papert, S. (1980). *Mindstorms: children, computers, and powerful ideas.* New York: Basic Books.

Papert, S. (1991). Situating constructionism. In S. Papert & I. Harel (Eds.), *Constructionism* (pp. 1–11). Cambridge, MA: MIT Press.

Pearson, G. (2007). Assessment of technological literacy in the United States. *Design and Technology: An International Journal, 12*(2):87-97.

Peppler, K. & Glosson, D. (2013). Stitching circuits: Learning about circuitry through e-textile materials. *Journal of Science Education and Technology, 22*(5), 751-763.

Petrich, M., Wilkinson, K., & Bevan, B. (2013). It looks like fun, but are they learning? In M. Honey & D. Kanter, (Eds.), *Design, make, play: Growing the next generation of STEM innovators* (pp. 50-70). New York, NY: Routledge.

Piaget, J. (1929). *The child's conception of the world.* New York, NY: Harcourt, Brace and Co.

Piaget, J. (1974). *To understand is to invent.* New York, NY: Penguin Books.

Remold, J., Verdugo, R. & Michalchik, V. (2013) California tinkering afterschool network pilot year evaluation report. SRI, Center for Technology and Learning. Retrieved from http://www.sri.com/sites/default/files/publications/ca-tinkering-afterschool-network-evaluation-report-final_web_1.pdf

Ro, H. K., Merson, D., Lattuca, L. R., & Terenzini, P. T. (2015). Validity of the contextual competence scale for engineering students. *Journal of Engineering Education, 104*(1), 35-54.

Smith, J.J., Cheruvelil, K.S., & Auvenshine, S. (2013). Assessment of student learning associated with tree thinking in an undergraduate introductory organismal biology course. *CBE Life Sciences Education, 12*(3), 542-552.

Velayutham, S., Aldridge, J. & Fraser, B. (2011) Development and validation of an instrument to measure students' motivation and self-regulation in science learning. *International Journal of Science Education, 33*(15), 2159-217.

Vygotsky, L. S. (1978). *Mind in society*. M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds. Cambridge, MA: Harvard University Press.

Werner, J., Denner, L., & Campe, S. (2006). IT fluency from a project-based program for middle school students. *Journal of Computer Science Education Online, 2*, 205-6.

Worsley, M. & Blikstein, P. (2014). Analyzing engineering design through the lens of learning analytics. *Journal of Learning Analytics*, *1*(2), pp. 151-186.

Yoon Yoon, S., Evans, M. G., & Strobel, J. (2014). Validation of the Teaching Engineering Self-Efficacy Scale for K-12 teachers: A structural equation modeling approach. *Journal of Engineering Education, 103*(3), 463-485.

Zimmerman, B. J. (1998). Academic studying and the development of personal skill: A self-regulatory perspective. *Educational Psychologist, 33*, 73–86.

Zylka, J., Muller, W., & Martins, S. W. (2011). Media literacy worldwide: Similarities and differences of theoretical and practical approaches. Proceedings of the *Global Engineering Education Conference (EDUCON),* IEEE, 726-733.

Authors

Paulo Blikstein is assistant professor at Stanford University Graduate School of Education and (by courtesy) in the Computer Science Department, 520 Galvez Mall, Room 102, CERAS Building, Stanford, CA 94305; paulob@stanford.edu.

Zaza Kabayadondo is co-director of the Design Thinking Initiative at Smith College, 25A Henshaw Avenue, Capen Annex, Northampton, MA 01063; zaza.kabayadondo@gmail.com.

Andrew P. Martin is an independent statistical consultant and former Stanford graduate student, 390 Serra Mall, Stanford, CA 94305; am1982@stanford.edu.

Deborah A. Fields is an educational research consultant and a temporary assistant professor at Utah State University, 2830 Old Main Hill, Logan, UT 84322; deborah.fields@usu.edu.