Computers & Education 157 (2020) 103944

Contents lists available at ScienceDirect

Computers & Education

journal homepage: http://www.elsevier.com/locate/compedu

Assessing learning in technology-rich maker activities: A systematic review of empirical research

Qiao Lin [a,*], Yue Yin [a], Xiaodan Tang [b], Roxana Hadad [a,c,d], Xiaoming Zhai [e]

[a] University of Illinois at Chicago, 1040 West Harrison St., Chicago, IL, 60607, USA
[b] Northwestern University, 625 Michigan Ave, Chicago, IL, 60611, USA
[c] Northeastern Illinois University, 5500 North St. Louis Avenue, Chicago, IL, 60625, USA
[d] University of California, 405 Hilgard Avenue, Los Angeles, CA, 90095, USA
[e] University of Georgia, 110 Carlton Street, Athens, GA, 30602, USA

ARTICLE INFO

Keywords:
Maker activities
Learning communities
Teaching/learning strategies
Interdisciplinary projects
Evaluation methodologies

ABSTRACT

Maker activities are drawing increasing attention in the field of science, technology, engineering and mathematics (STEM) education. Researchers have developed various assessments for maker activities to examine students' learning outcomes. However, a systematic review of research on such assessments is lacking. To fill this gap, we reviewed empirical studies on maker-based assessments in education. We systematically examined 60 studies regarding the overall features of the maker activities, the learning outcomes that were measured, the assessment formats, and the psychometric evidence of the assessments. Our review results indicate that more than 20 types of maker platforms have been employed in the activities, with e-textiles and LilyPad Arduino being the most popular. Five types of assessment tools have been used prevalently to examine students' diverse learning outcomes, specifically artifact assessments, tests, surveys, interviews, and observations. Most assessments are used in STEM-related maker activities, especially technology-centric activities, to measure STEM-related learning outcomes. Only 15% of the studies provide psychometric evidence of reliability and validity for the assessments. Based on the findings, we provide suggestions for future research which include developing more low-tech maker activities for students in lower-grades and with lower technology proficiency. In addition, future studies should improve rubrics for artifact assessment and explore more assessment tools for non-STEM subjects.

1. Introduction

Since the first Maker Faire in 2006, the maker movement has inspired educators with maker activities being integrated in both formal and informal learning settings to promote student learning (Bevan, Gutwill, Petrich, & Wilkinson, 2015; Peppler & Glosson, 2013; Rode et al., 2015). These environments where making and learning happen are commonly referred to as makerspaces. In physical makerspaces, communities with common interests design and create do-it-yourself (DIY) projects collaboratively or independently using technology and digital art (Halverson & Sheridan, 2014; Rivas, 2014) such as sewing, 3-D printing, laser cutting, and design. From a constructivism perspective (e.g., Piaget, 1950), these maker activities provide learners opportunities to construct knowledge through the development of mental representations and solving authentic problems. Thus, makerspaces have great

* Corresponding author. Department of Educational Psychology, University of Illinois at Chicago, 1040 W. Harrison St. Chicago, IL, 60607, USA.
E-mail addresses: qiaolin234@gmail.com (Q. Lin), yueyin@uic.edu (Y. Yin), xiaodan.tang@northwestern.edu (X. Tang), rhadad@ucla.edu (R. Hadad), zhaixia1@msu.edu (X. Zhai).

https://doi.org/10.1016/j.compedu.2020.103944
Received 24 December 2019; Received in revised form 3 June 2020; Accepted 6 June 2020
Available online 26 June 2020

potential to help students develop content knowledge, as well as the competencies of designing, innovating, understanding computational concepts, conducting collaborative work (Papavlasopoulou, Giannakos, & Jaccheri, 2017).

Appropriate assessments are necessary to understand students' learning outcomes and to examine whether students have developed the desired skills in maker activities. To achieve such goals, researchers have developed considerable assessments in multiple formats including paper-pencil tests, artifacts, surveys and observation protocols and used them as both formative and summative assessments in maker activities. However, a systematic review of these assessments is lacking in the field. We have very limited knowledge about what the attributes of the assessments are, how the assessments are integrated in the maker activities, and what the psychometric features of these assessments are. Without that knowledge, it is difficult to advance studies on designing, validating, and implementing maker-based assessments.

To fill this gap, this study aims to systematically review and synthesize empirical evidence on assessment of maker activities. We focus on the studies that employ specific assessment tools in maker activities to measure learning outcomes gained in diverse learning settings across educational levels. In particular, our review examines what materials and subject domains are involved in maker activities, what kind of assessments are used to measure what kind of learning outcomes. We also review the reliability and validity evidence that are reported for these assessments. As such, the study is intended to extends our knowledge on assessment of maker activities and directs future studies to develop and implement assessment using maker activities. Below, we provide a background of maker activities and their related assessments.

1.1. Making, Maker activity and maker education

Although the branding of the maker movement is relatively new, it has deep roots in the world of education. A significant role of making in education can be traced to the constructivist ideas of Dewey (1923), Piaget (1950), and Freire (1974). Harel and Papert (1991) built on these ideas with the theory of constructionism. Martinez and Stager (2013) credited Papert as "the father of the maker movement", because the theory of constructionism undergirds the maker movement's focus on problem solving and digital and physical fabrication. To further understand the promise of maker movement for education, Martin (2015) described three critical elements: (a) digital tools, which include platforms of rapid prototyping tools and low-cost microcontroller; (b) community infrastructure, which refers to online resources and in-person spaces such as museums, makerspace and events like maker faire; and (c) maker mindset and dispositions, which includes values (Dewey, 1929), beliefs (Elby & Hammer, 2001), and dispositions (Perkins, Tishman, Ritchhart, Donis, & Andrade, 2000). These elements have led to making being seen as part of a global movement for education reform (Sang & Simpson, 2019).

The core of the maker movement is making, focusing on designing, building, and modifying a real and/or digital "product". Specifically, making refers to creating objects not only with high-tech resources such as 3D fabrication, laser cutting, physical computing, and programming (Martin, 2015; Sheridan et al., 2014) but also with low-tech everyday materials such as broken toys, sewing materials, and craft supplies (Holbert, 2016; Parekh & Gee, 2018).

The wide range of materials and highly accessible technology have enabled the maker activities to become increasingly popular within the STEM education community in the past decade. By participating in hands-on investigations, students are more likely to develop knowledge-in-use to help them better understand how scientific knowledge is applied in solving problems in real life (Martin & Dixon, 2013; Martinez & Stager, 2013).

Given the great potential of maker activities to improve students' scientific understanding and skills in STEM (Kafai, Peppler, & Chapman, 2009; Resnick et al., 2009), mounting research has focused on integrating maker activities in both formal and informal education. Thus, maker education has drawn great attention as a way to facilitate invention and innovation in STEM. Martinez and Stager (2013) stated that maker education is a process of learning-by-doing that children gain knowledge from direct experience with materials. Likewise, Fleming (2015) claimed that maker education is about moving from consumption to creation and considered it as an equalizer within the modern education system by encouraging informal learning opportunities. Lundberg and Rasmussen (2018) defined maker education as a type of project-based learning where the learner produces a physical object or artifact by using newly learned concepts and skills. Fundamentally, the research on maker education focuses on investigating the promise of maker movement in education. For example, Halverson and Sheridan (2014) stated that the maker movement's role in education not only embraces the constructionist frame of progressive education to turn knowledge into action, but also provokes our thinking regarding the tension between formal and informal learning experiences. Taken together, maker education is not a defined pedagogy, but rather the amalgamation of educational approaches that emphasizes learning through making, such as project-based learning (e.g., Schneider, Krajcik, Marx, & Soloway, 2002) and problem-based learning (e.g., Schwartz, Mennin, & Webb, 2001).

Specifically, maker activities refer to the creative manufacturing processes that utilize various materials. In the maker-centered learning environments, students play a role as innovators to pose questions and find solutions through hands-on activities. This change of role catalyzes the demands for re-imagining curriculum and teaching. As such, maker education performs as a framework with significant value in guiding curriculum development. Following the guidance of maker education, the design of maker activities reflects the core concepts of learning-by-doing on individual growth and realize the values and effects of maker education.

Researchers, however, have not clearly differentiated maker education from maker activity in literature. For example, O'Brien, Hansen, and Harlow's (2016) used term of "maker education activity" to emphasize making in formal education. Yet, as this study intends to review assessment of maker activities in both formal and informal learning educational settings, we employ "maker activities," rather than "maker education" through the paper. As such, our study focuses on the assessment tools of maker activities that could mirror the effectiveness of maker education in either formal or informal contexts.

Although various maker activities have been used to promote learning and encourage individuals to be creators in daily life, gaps still exist in understanding the relationship between making and learning. For example, what kind of making activities can promote what kind of learning? What might be a more effective implementation method for promoting learning in maker activities? Essentially high-quality assessments are critical to examine these relationships and to fill in the associated gaps.

1.2. Challenges and importance of maker-based assessments

As lots of maker activities were implemented in informal educational settings or after-school programs in schools where the learning environment is less structured than formal educational settings, makerspaces have some features distinct from traditional learning environments: (a) Both the students and educators in makerspaces can change from day to day. In the informal educational settings, students typically come whenever they have time and different educators may rotate to manage the makerspaces. Students of different grades in the makerspace also have varying prior knowledge, interests, and backgrounds (Kurti, Kurti, & Fleming, 2014; Schrock, 2014). (b) Maker activities tend to be interdisciplinary. For example, e-textiles can involve physics and design (Moriwaki et al., 2012); 3D printing can involve programming and art (Schlegel et al., 2019); Makey Makey projects can involve programming, engineering, and/or product design (Cheng & Brown, 2015; Clapp & Jimenez, 2016). (c) The activities in which students are engaged may vary greatly depending the resources. As makerspaces provide various equipment and encourage creativity, students can have the freedom to choose whatever equipment that is available in the space to create their own products; or using the same equipment, they can create products with different designs/functions (Kafai, Fields, & Searle, 2014; Keune & Peppler, 2019). (d) The arrangements in makerspace can vary. Students can conduct maker activities independently, in pairs, or in groups (Fredrick, 2015; Halverson & Sheridan, 2014). (e) The learning outcomes may also vary greatly. As discussed earlier, students can be engaged in different activities with different arrangements. The different activities and arrangement can lead to different learning outcomes, which can be cognitive, non-cognitive, and/or social skills (Somanath, Morrison, Hughes, Sharlin, & Sousa, 2016).

With varying students, educators, activities, disciplines, arrangements, and learning outcomes, assessing students in makerspace can be a moving target. In addition, the making community can be critical of introducing assessments to the deep learning experience of a maker activity because assessments are sometimes seen as an interruption or a distraction that can harm motivation and understanding (Martinez & Stager, 2013).

Although it is challenging to use assessments in makerspaces, the development of maker-based assessments can be similar to the other typical assessments. Cun, Abramovich, and Smith (2019) designed an assessment matrix for library makerspaces based on the assessment needs of library makerspace participants and librarians. Researchers created five categories of assessment tools to address the needs of participants including visitor logs, self-assessment, one-on-one sessions, paper/digital surveys, and librarian observations. These assessment tools were developed based on the literature of summative and formative assessments. When we use high-quality summative assessment tools in makerspace, we can examine whether students have achieved the expected learning outcomes, whether the makerspace educators are effective, and/or whether a maker program/activity/professional development is helpful for students' learning. When we use high-quality formative assessment tools, we can use them to identify students' learning gaps, so that we can modify activities, instruction, and arrangements accordingly to improve learning. Therefore, it is critical to study and use assessments in makerspaces.

1.3. Gaps in maker-based assessment research

Previous studies have shown that final paper-pencil tests have been widely used as summative assessment tools to measure students' learning and academic achievement in traditional learning settings (e.g., Dixson & Worrell, 2016). In addition, instructional activities such as questioning, discussion, seatwork, self-assessment, peer assessment and homework assignments can be formative assessment tools (e.g., Yin, Tomita, & Shavelson, 2014). Although not used as frequently as those used in traditional learning environments, several instruments have been used to measure the impact of maker activities on students' cognitive achievement, affect development, and learning behavior. Survey and interview instruments have been widely used to assess students' self-reported attitudes and perceptions in maker activities (Lee & Fields, 2013). Artifact assessment has become a prevalent instrument to innovatively assess students' knowledge and skills (e.g., Barton, Tan, & Greenberg, 2016) and with the help of performance observation, researchers can analyze the process of artifacts construction and infer students' learning (Lee, Kafai, Vasudevan, & Davis, 2014). As a traditional assessment tool, tests have also been employed to measure both cognitive and non-cognitive learning outcomes (Chamrat, 2018). Using multiple forms, Cun et al. (2019) have proposed five categories of maker-based assessment tools: visitor logs, self-assessment, one-on-one sessions, paper/digital surveys, and librarian observations were regarded as useful types of assessments for makerspaces. In particular, visitor logs assess the fidelity in makerspaces by recording frequency of attendance; self-assessment enables educators to offer formative feedback to participants whose needs were recorded on the self-assessment forms in a timely manner; one-on-one sessions allow teachers to work individually with a student and administer assessment questions directly to understand students' mastery of technology in maker activities.

In addition to the maker-based assessment tools used in informal learning settings, researchers have proposed guidelines for assessment development in formal learning settings. Murai et al. (2019) have established four design principles for embedded assessment in school-based making: (a) it should be integrated within the learning environment or activities; (b) it should be construct-driven and clarify target outcomes that an activity intends to foster; (c) it should be evidence-centered, generating visible, tangible, and varied forms of evidence for the underlying constructs; and (d) it should involve students as active participants in the assessment process.

Although various assessments have been used in various maker activities to measure outcomes in different settings, no study has systematically reviewed the characteristics of these instruments. It is unclear what types of outcomes have been sufficiently or insufficiently measured, whether these instruments are aligned with learning objectives, as well as the psychometric qualities of these instruments. Our study is intended to summarize the past achievements and encourage and guide future endeavors. In particular, this systematic review is designed to fill these gaps and to answer the following research questions:

a. What are the overall features of maker activities that embedded assessments?
b. What learning outcomes have been measured in maker activities?
c. How are student learning outcomes assessed in maker activities?
d. What psychometric evidence has been gained from maker-based assessments?

By answering these questions, we expect to better understand what has been achieved and what is missing, as well as provide a comprehensive reference for researchers and educators to inform the design and use of assessments in maker settings.

2. Method

A systematic review evaluates and interprets all accessible research relevant to a theme, a set of research questions, or an event of interest (Kitchenham, 2004). To carry out this review study, we followed steps for conducting systematic reviews based on Kitchenham (2004). In the following section, we detail the search of the literature, inclusion and exclusion criteria, literature coding procedure, and data analysis.

2.1. Literature search procedures

To capture the assessments used in maker activities in this field, we collected both published and unpublished studies from a wide variety of online bibliographic databases: EBSCO (including ERIC), PsycINFO, Proquest, ScienceDirect, ACM Digital Library, and Google Scholar. We also searched key conferences such as *ACM CHI Conference on Human Factors in Computing Systems (SIGCHI)*, *Conference on Creativity and Fabrication in Education (Fablearn)* and *ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE)*. We first used three major maker-related terms to search articles: maker activity, maker movement and makerspace. Given that studies might employ maker activities but do not mention maker-related vocabulary in paper, we specified a set of keywords regarding widely used maker platforms (e.g., e-textile, 3D printing, laser cutting, Makey Makey, Arduino, Scratch, Lego) from the literature and combined these keywords with the terms “making” or “make” to do an advanced research. Therefore, the key search terms of platforms and making were “e-textile AND making”, “3D printing AND making”, “laser cutting AND making”, “Makey Makey AND making”, “Arduino AND making”, “Scratch AND making”, “Lego AND making”. We also used the combination of platforms and the term “make” in search, such as “e-textile AND make”, “3D printing AND make” etc. By doing so, we were able to include articles that employ different maker activities under the framework of maker theory as well as those contributed to maker-based assessment but did not mention assessment as a key term in the description.

2.2. Inclusion and exclusion criteria

We used the seven inclusion criteria to select articles: (a) articles include any one of terms: maker activity, maker movement, makerspace, making or make, in any part of the paper; (b) empirical studies rather than conceptual or theoretical papers; (c) participants were involved in the maker activities; (d) assessments were used to measure learning outcomes in the maker activities; (e) articles were written in English; (f) articles were published between 2006 and 2019; and (g) articles are available in full-text. In addition, we conducted a snowball search based on the articles that we have found. In particular, we searched the reference lists of the selected articles to find the relevant articles that we may have missed earlier. We also searched articles by authors who have published maker-based assessments. This snowball search resulted in 53 publications. By doing so, we were able to ensure that all the related articles are included. Reading the abstracts and full text when needed, we initially selected 341 eligible articles. The first author then downloaded all the initial selected articles.

After collecting the initial articles, we used two exclusion criteria to filter articles: (a) studies only focused on robotics, for example LEGO Mindstorms and robotics making were excluded; and (b) studies that involved pure programming learning without using any physical maker platforms or making any tangible artifacts, for example online computer game and application designs were excluded. We excluded robotics and pure programming studies because they can be the independent subject domains by themselves and have been widely studied, while maker activities focus more on tinkering projects and typically produce artifacts and are under-studied. Using the inclusion and exclusion criteria, we screened the initially selected articles and kept 60 studies.

2.3. Literature coding

Following the procedures of a content analysis (Fraenkel, Wallen, & Hyun, 2011), we built the coding rubrics and coded the literature systematically in three stages. First, we developed the coding system in an Excel spreadsheet, which we aligned with the research questions to capture the information from sample articles. Second, we refined the coding rubrics. The first three authors coded 13 related articles: For the first five articles, we coded each of them together while discussing article coding and further refined our coding system in three ways: (a) adding or removing codes, (b) clarifying the meaning of the codes, and (c) modifying the values of codes. For the other eight articles, each researcher first coded the same article independently, and then we compared our coding, resolving our coding discrepancies by discussing and further refining the coding system in the three ways above. Third, after we finalized our coding system and received sufficient coding training, we tested inter-rater agreement. Using the coding system, authors 1 and 3 further independently coded 10 articles. The inter-rater agreement on all the codes is 0.94. When author 1 and author 3 differed on their codes, they invited author 2 to discuss with them, the three authors reached a consensus, and used it as the final code for that entry. After we achieved a satisfactory interrater agreement, author 1 coded the rest of the articles. Appendix A lists the articles included in this study, characterized by part of our finalized coding system.

3. Results

We imported the coded data to SPSS and analyzed it to answer the research questions. We organize our findings in four sections corresponding to the research questions.

3.1. Overall features of the maker-based activities

We examine the overall features of the maker-based learning from the following aspects: (a) platforms and materials for maker activities, (b) subject domains, (c) age range, and (d) learning settings (see Table 1).

Platforms and materials for maker activities. Among the 60 reviewed studies, 29 different platforms have been used to create various artifacts. The most frequently used platforms for maker activities are e-textiles and LilyPad Arduino, covering 40% of the reviewed studies. 3D printing (20%) is popular as well, followed by Scratch (12%) and Makey Makey (12%). Appendix B illustrates seven of the most commonly used platforms. For example, e-textiles is a type of fabric containing electronic elements that helps students better understand the mechanisms of circuits and how it is influencing people's daily life (Buechley, Peppler, Eisenberg, & Kafai, 2013). Makey Makey is a basic microcontroller, which allows users to control computer programs by everyday objects (Silver, Rosenbaum, & Shaw, 2012). Using Makey Makey students can use conductive materials to replace the keys on the keyboard, and in the process, learn about conductivity of everyday objects like fruits, Play-Doh, and aluminum foil (Kafai & Vasudevan, 2015a). The more complex Arduino platform makes it possible for young people to engage in programming and creative development of physical computing, such as robots and interactive installations (Giannakos & Jaccheri, 2013), which help students develop knowledge of computer science (CS) and engineering.

In the review, we found that more than 90% of the platforms are technology-based, such as laser cutting, Scratch, LilyPad, and Arduino, while some traditional platforms and everyday materials such as LEGO, Play-Doh, wood and fabric have also been used in innovative ways. Besides these, other particular platforms were developed and used less frequently, for example, (a) Talkoo (1), a kit consist of physical computing plug-and-play modules as part of a visual programming environment with prototyping materials that allow beginners to get started with building electronics; (b) EarExplorer (1), an interactive tangible system where students can manipulate and connect parts of the auditory system to rebuild a functional structure; and (c) Maker Theater (1), a basic battery-switch-resistor-LED circuit that was carefully designed for children to tell a 'lighted story'.

Subject domain. We coded the subject domain in each study and found that STEM have been covered in almost all the studies. Compared to the majority studies that focus on technology and engineering, fewer studies (2) focus on subjects of science and mathematics. Schneider, Bumbacher, and Blikstein (2015) employed platform EarExplorer to introduce knowledge of biology and Torralba (2019) used a mixed-methods approach to investigate mathematics learning in a maker-based activity. Besides, eight studies involved non-STEM subject domain such as design and art to encourage personal and creative self-expression in the making process (eg., Jacobs & Buechley, 2013). In particular, Moriwaki et al. (2012) expanded STEM to Science, Technology, Engineering, Arts, Math, and Design (STEAMD), adding art and design to emphasize the comprehensive nature of making.

Age range. Regarding the age range of participants in the maker activity, we found that middle school and high school students were the most involved. A total of 39 (65%) studies involved participants age from 11 to 18 years old. Fourteen studies engaged elementary students and 11 studies engaged university students of 18-years-old or older. Three studies didn't have age restrictions so that participants of a broad age range from kindergarten to adult could participate in the maker activities. In general, maker activities are not age-specified, participants in different age and level can always find suitable platforms and gain knowledge from it.

Table 1
Overview features of the maker activity studies.

Variables	Categories	Numbers	Percent
Platforms	E-textiles and LilyPad Arduino	24	40%
	3D printing	12	20%
	Scratch	7	12%
	Makey Makey	7	12%
	Others	21	35%
Subject Domains	STEM	52	87%
	Non-STEM	8	13%
Age Range	K - elementary	14	23%
	Middle and high school	39	65%
	College and beyond	11	18%
Educational Setting	Formal	33	55%
	Informal	31	52%

Learning settings. Among the 60 studies reviewed, 55% of the studies were conducted in a formal setting such as a classroom or school workshop. The remaining half of the studies (31) were conducted in an informal setting such as a library, after-school program, independent lab, or summer camp. The arrangement of mentors and participants in the learning settings were diverse. In some informal learning settings like libraries, materials were provided and no mentor directed the activities (Lotts, 2016), and participants independently choose the materials to build what they are interested in. In some other informal settings such as after-school programs, mentors work with researchers to design the maker activities (Bevan et al., 2015) and provided help with technical tools and design decisions for participants in the making process (Barton et al., 2016). In both formal and informal learning settings, participants could either work independently to complete a personal artifact (Qi & Buechley, 2014), or work with other people in a group collaboration (Kostakis, Niaros, & Giotitsas, 2015). In the review, we found that no single form arranged the mentors and students in learning settings.

Originated from maker culture and being popular in informal settings like makerspaces, maker activities have also been increasingly valued in the formal educational context. Although some researchers described the challenges of teaching and learning with maker technology in formal schooling (Godhe, Lilja, & Selwyn, 2019), in the review we found a similar number of studies conducted in both formal and informal educational settings. This pattern shows that maker activities, being versatile and flexible, have great potential to be used for teaching and learning.

3.2. Learning outcomes measured in maker activities

Research has shown that by engaging participants in constructing creative artifacts, maker activities provide an opportunity for students to learn content knowledge (e.g., Brady et al., 2017), cultivate positive feelings and attitudes about making (e.g., Fordyce, Heemsbergen, Mignone, & Nansen, 2015), and to enhance engagement and communication with others (e.g., Schwartz, DiGiacomo, & Gutierrez, 2013). Thus, to understand the role of making in learning, we classified learning outcomes in three categories: a) cognition – understanding and construction of content knowledge; b) affect – feelings and attitudes that are fostered in the making process, and c) others – mainly including engagement and collaboration in maker activities. Table 2 shows the learning outcomes measured in the reviewed studies. In general, 11 studies measured all three learning outcomes (cognitive, affect, and others), while the rest of the studies identified one or two learning outcomes in maker activities.

Cognitive outcomes. Based on the cognitive domain, we further classified student learning outcomes into three sub-categories: (a) STEM-related content knowledge that includes knowledge of science, technology, engineering and mathematics; (b) programming knowledge that focus on computational principles and/or programming concepts; and (c) skills and competence, which are comprised of different thinking skills and abilities in problem solving.

A total of 48 (around 80%) studies documented cognitive learning outcomes. The high percentage reflects the importance of cognitive learning outcomes in maker activities. Regarding each sub-category, 28 studies measured STEM-related content knowledge. In particular, knowledge of electronics and circuitry have been the focus of 13 studies. For example, Peppler and Glosson (2013) introduced an afterschool workshop to understand if youth developed a conceptual understanding of simple circuitry while constructing electronic artifacts using an e-textiles toolkit. The analysis of both quantitative and qualitative data showed that youth significantly improved their ability to diagram working circuits as well as their understandings of core circuitry concepts (e.g., current flow, connections, polarity) after the 20-h workshop. In addition, science knowledge has also been integrated in maker activities. Schneider et al. (2015) designed EarExplorer to improve students' biology knowledge of the human hearing system. The EarExplorer is an interactive tangible system that consists of a tabletop interface with 3D-printed tangibles tagged with markers. After connecting all the tangibles in the correct sequence, the users can learn about the different organs and how the sound waves travel in ears. Using a between-subject experimental design, the researchers grouped students into two conditions: a) students rebuild the hearing system by self-driven discovery; b) students were required to follow the step-by-step instructions of a video-teacher. A learning test was used to measure students' biology knowledge gains, which show that students in the condition of making and exploring learned significantly more knowledge of human hearing system than those in the condition of watching and listening.

Compared with STEM-related content knowledge, about one third studies (20) assessed programming knowledge in maker activities, focusing on how the students understood computational concepts and used programming language. Makey Makey, Scratch, and Arduino were the most commonly used platforms in helping students learn computing and programming. Kafai and Vasudevan (2015b) connected on-screen programming Scratch with hands-on crafting Makey Makey to introduce students to the computational concepts. The researchers examined students' projects and approaches to computing and crafting in their on-screen and off-screen designs. They found

Table 2
Learning outcomes measured in maker activities.

Learning Outcomes	Sub-categories	Numbers	Percent
Cognition	STEM-related content knowledge	28	47%
	Programming knowledge	20	33%
	Skills and Competence	15	25%
Affect	Confidence	10	17%
	Attitude	7	12%
	Feeling and Perceptions	19	32%
Other	Engagement	16	27%
	Collaboration	11	18%

Note: Some studies assessed more than one kind of learning outcomes so the total percentage was more than 100%.

that in the process of integrating video and board game elements, students learned computational concepts such as sequences, events, loops, and conditionals, while refining their board game designs. In addition to employing a single maker platform, some studies combined several platforms in maker activities and e-textiles have been widely combined with Arduino to promote both STEM and programming knowledge simultaneously. For instance, Litts, Kafai, Lui, Walker, and Widman (2017) used LilyPad Arduino and e-textiles to expand students' learning about circuitry. Pretest and posttest results demonstrated that students not only understood functional circuits but also improved their ability to read, design, and remix code for controlling circuits. Besides, some computer platforms have been creatively developed and used in programming learning. For example, Jacobs and Buechley (2013) developed a preliminary computational-design programing tool, Codeable Objects, to work in conjunction with digital-fabrication machines. The results suggested that the students constructed a deep understanding of computation by participating in the workshop.

Among the studies that assess cognitive learning outcomes, fewer studies (15) focused on developing student skills (e.g., abstract thinking skills, computation thinking skills) and competence (e.g., creative competence) in maker activities. In the review, we found that the more advanced technology of 3D printing had been extensively brought in maker activities to facilitate the development of thinking skills and competence. For example, Huang and Lin (2017) employed 3D printing to increase mental rotation and visualization components of spatial ability. The results indicated that different diagrams and 3D printed solid models improved different spatial abilities. Similarly, by combining laser cutting and 3D printing, Mellis and Buechley (2012) mixed the learning of technical skills with the construction of a finished working product. Saorín et al. (2017) measured students' creative competence after participating a workshop to promote creativity in students using 3D scanning tools, digital edition of 3D meshes, and 3D printers. The results showed that activities with digital editing tools and 3D printing are valid for the development of creative competence. Instead of using advanced technology platforms, Lotts (2016) described the benefits of using the low-cost, traditional maker platform of LEGO to improve creative problem-solving skills, abstract thinking, and conceptual thinking skills in the process of building LEGO models.

Affective outcomes. Affective learning outcomes in this study refer to confidence, self-efficacy, attitudes, and feelings that are generated and/or changed during maker activities. Research suggested that students would change their dispositions toward making or a specific subject domain when they are engaged in hands-on activities (e.g., Searle & Kafai, 2015). A total of 17 studies in our review measured participants' confidence level (10) and attitude changes (7). In particular, computing and programming have been the center in these affect assessments. For instance, Jacobs and Buechley (2013) developed Codeable Objects, a programming and digital fabrication tool that allow novices to produce personal and functional objects and measured practitioners' confidence in programming and design. Specifically, some studies focus on examining female students' confidence and attitude toward making and computing. Nation and Durán (2019) examined how young Latinx women actively build their confidence in making as well as establishing intersecting identities as makers, group members, and contributors to their community after participating a Critical Maker program. Four studies in our review assessed students' attitudes toward STEM disciplines and CS/computing related careers, mostly focusing on the attitude toward CS/computing and they found that the students' intention to study programming in the future increase after participating the maker activities.

More studies examined students' feelings and perceptions about the creative activities than confidence and attitudinal learning outcomes. Giannakos, Jaccheri, and Leftheriotis (2014) and Chu, Angello, Saenz, and Quek (2017) assessed personal feelings such as enjoyment, happiness, and anxiety in the process of making activities and they confirmed the positive effects of happiness and the negative effects of anxiety. Other studies were more concerned with the perceptions and perspectives on computing itself. Kafai et al. (2013) found that students expressed personal relevancy of computing at large after participating in a 10-week e-textile curriculum module workshop. In our review, a total of 19 studies examined personal feelings and perceptions as part of the outcomes.

Other outcomes. In addition to the cognitive and affective outcomes, about half of the studies (31) also examined some other outcomes. Except for cognitive and affective outcomes, we summarized all other learning outcomes into the last category and found engagement and collaboration have been emphasized in 27 studies. Engagement has long been identified as an important analytical dimension of learning in both formal and informal settings (Hidi & Renninger, 2006; Humphrey & Gutwill, 2017; Sawyer, 2005). In our review, quite a few studies (16) emphasized engagement and participation in the making process. Searle, Tofel-Grehl, and Breitenstein (2019) integrated e-textiles into core content classes to investigate the engagement change of girls from non-dominant communities. Moreover, engagement was defined and measured in different components and levels. Barton et al. (2016) developed an afterschool program to support youth in sustained engagement in engineering. They defined and measured three dimensions of engagement - critical, connected, and collective, which play an important role in supporting youth from underrepresented communities in learning STEM. Critical engagement was when students were involved in the issues that framed their lives and shaped the ways that they find the problems and develop solutions. Connected engagement referred to how youth connect their making with their community and broader social issues. Collective engagement meant they work together on problems that are defined through interactions with others and leverage others' experiences and struggles. Bevan et al. (2015) identified the points of transition or choice in engagement, which emerge naturally in the open-ended nature of making programs. The authors analyzed audio/video data to identify the engagement in terms of spending time in tinkering activities, displaying motivation or investment through affect or behavior, and requesting or offering help in solving problems.

Maker activities also promoted collaboration by encouraging students to share their views and work with others to construct the artifact. Qi and Buechley (2014) conducted an observation on participants' working process and found participants collaborate with each other to create circuits, shared and traded their circuit sketchbooks. In addition to the peer collaboration, maker activities are helpful in promoting child-parent collaboration. Through observation, Moriwaki et al. (2012) found that parents adopted a translation tactic to deepen their children's understanding when kids encountered difficulty in a maker activity.

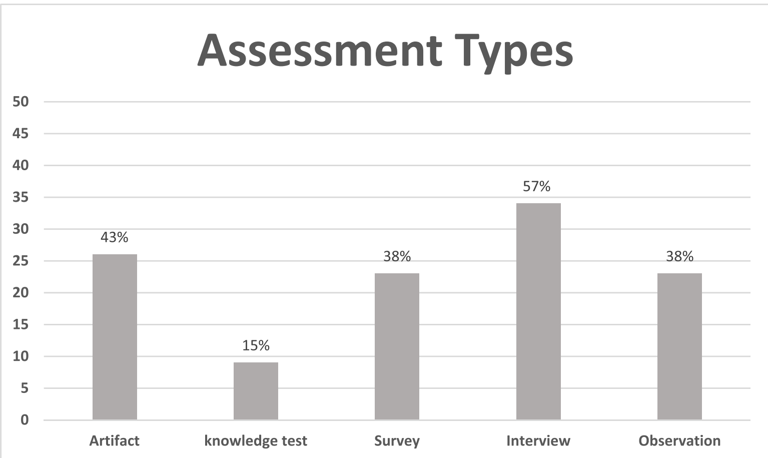
3.3. Assessments used in maker activities

To answer the third research question, we mainly investigated (a) what type of assessments are used and (b) how the assessments are integrated in maker activities. Our review shows that artifacts, tests, surveys, interviews and observation are frequently used assessment tools in maker activities (see Fig. 1). Appendix C provide sample questions from these assessment tools except observation because observation is more likely to be conducted via video record or filed nots.

Artifact assessment. One of the most salient findings in our review is that artifact assessment has become a popular tool in maker activity research. A total of 42 studies (70%) encouraged students to build artifacts as part of participating in maker activities. Among them, some studies expected students to build a specific artifact based on the learning topics while other studies let students create freely with the materials provided. For example, using e-textiles in conjunction with LilyPad Arduino, Kafai et al. (2014) encouraged students to design and create their own e-textiles projects. Beginning with paper and pencil designs, students drew circuit schematics, then sewed and crafted their designs with textile materials and finally programmed the LilyPad Arduino. The final artifacts were diverse and included a hat inspired by anime symbols, a light-up tote bag with light-sensing handles, and a belt inspired by the double-helix structure of DNA. In the review, we found that the studies involving e-textiles, Makey Makeys and programming activities tended to use artifacts assessment tools more than the other platforms, possibly because these platforms provide students with more flexible hands-on opportunities to explore the knowledge and mechanism in the activities. Thus, building their own artifact provide an effective way for educators to evaluate what students have learned in the making process.

Although 70% of the studies required students to construct artifacts, most of them did not analyze the artifacts thoroughly. Only 26 studies (43%) evaluated the finished artifacts to understand what learning outcomes were demonstrated in the artifacts. In one of the studies that assess the final product, Torralba (2019) implemented a semester-long maker-based course to investigate how participation in maker activities impacts students' understanding of grade-level proportional reasoning in mathematics. To evaluate students' understanding of proportional reasoning and proficiency level, researchers analyzed their artifacts, final products, and oral presentation by using a coding framework. This coding framework included two major codes for artifact and oral presentation: Proportional Computation (PC) and Proportional Relationships (PR). With each code, four labels for levels of proficiency were applied: Below Expectations (B), Developing (D), Knowledgeable (K), and Advanced (A). Similarly, Nemorin (2017) conducted a research to understand how 3D printing technologies and practices shape learning within a school context. A set of detailed assessment criteria was developed to evaluate the final 3D artifacts. The assessment rubric includes statements of research and design criteria, as well as criteria related to production, functionality, and aesthetic appeal of the final artifact. Students could also refer to the assessment rubric as a formative assessment tool in the process of design and construction. Moreover, Lui, Fields, and Kafai (2019) designed portfolio that asked students to provide three sections about the artifact: a) documentation of the final projects design; b) discussion of the process of making the final project; and c) reflection of the learning throughout the unit. In these studies, artifact assessments were implemented through developing coding schemes, rubrics, or portfolios to understand students' learning.

Test. In the review, we found that tests were used in nine studies to measure students' learning in making, demonstrating the feasibility of traditional assessment tools in educational maker settings. The content of the tests was both cognitive or non-cognitive, depending on the subject domains involved in the maker activities. Among them, STEM-related content knowledge and programming concepts are the core of many knowledge tests. For example, Chamrat (2018) selected 11 Ordinary National Educational Test (O-NET) items to assess students' gains in knowledge of electricity and circuits after the students participated in five modules of science camp and tinkering activities.



Note: Some studies employed more than one kind of assessment tool so the total percentage was more than 100%.

Fig. 1. Assessment types.
Note: Some studies employed more than one kind of assessment tool so the total percentage was more than 100%.

Moreover, tests have also been used to measure skills and abilities such as creativity and spatial ability. Saorín et al. (2017) used the Abreaction Test of Creativity to measure the development of creativity value at the beginning and the end of the workshop. Huang and Lin (2017) developed the Mental Rotation Test which consists of 25 questions over 10 min and used Guay's Purdue Spatial Visualization Test to measure the development of the mental rotation and spatial visualization components of spatial ability in a 3D program.

Survey. A total of 23 studies employed surveys with items measuring not only performance in maker activities, but also dispositions toward making and STEM subjects domains. Surveys were commonly used to measure the self-report confidence level in making. For example, Blikstein et al. (2017) developed and validated the Exploration and Fabrication Technologies Instrument (EFT) to measure students' confidence in EFT skills and assess how confidence is related to actual task performance. With generalizable and psychometrically sound index, the EFT instrument provided a solid scale to capture a new and distinct set of technology literacies that arise within making settings. In addition, 14 studies used surveys to assess the affective learning outcomes including attitude toward computing and STEM disciplines, as well as feeling and motivations to participate in maker activities in future. For example, Tofel-Grehl et al. (2017) provided students an abbreviated version of the Is Science Me (ISM; Gilmartin, Li, & Aschbacher, 2006) survey, which contains a wide range of questions that attempt to capture both the students' feelings toward STEM disciplines and careers as well as their perceptions of the beliefs of their families and peers. The sample questions about teachers, family, and peers include "My teacher cares if I think science is interesting"; "My family would be happy if I chose to pursue a career in science, technology, or engineering"; and "How many of your close friends like science, technology, or engineering?"

Interview. As a qualitative measure, interviews have been widely used in more than half of the studies (34). With variations of pre-post interview, informal interview, semi-structured interview, and transcribed interview, interviews examine students' performance and learning in making. For instance, Lee and Fields (2013) designed and developed a clinical interview protocol, the "Interactive Toy Interview" (ITI), which was used to assess prior knowledge and learning outcome among undergraduate students. Using three different toys, the interviewer asked students interactive questions about their capabilities associated with a university-level "Craft Technologies" course to understand their thinking and learning progress. In many studies, instead of developing specific protocols, we found that debriefing interviews have been used to understand students' learning process and perceptions after they participated in the maker activities. Searle, Fields, Lui, and Kafai (2014) conducted extensive debriefing interviews with high school participants in a 10-weeks e-textile unit as part of an elective CS class at their high school. Within these interviews, they focused on students' perceptions of the computational process and CS as a field. Interviews have been a useful tool in assessing learning process in making because it can capture the performance details that are easily omitted in the traditional test. The drawback of interviews is that analyzing the interview data is labor-intensive, involving transcribing, coding, and analyzing the video and/or audio recording. Interviews can be used to collect data encapsulating rich information but it is not practical for educators to use due to the time-cost of analysis. Supplemented with other assessment approaches, interviews can be used in focus groups to deepen the understanding of learning outcomes.

Observation. Observation has also been commonly used in maker activities to assesslearning outcomes. More than one third (23) of reviewed studies employed observation as a qualitative measure. Taking the field notes, researchers analyze the phenomenon they observed and infer what students have learned in the making process. Rode et al. (2015) conducted weekly club sessions to introduce children to computational thinking. All club sessions were documented in short session observations that were extended to provide ethnographic field notes. During the ongoing project, the researchers observed different problems occurred and strategies students used to deal with these challenges. The observation notes were coded and analyzed in several rounds to identify cognitive skills demonstrated in each session and produce a grounded model of computational making. Observation can be flexibly used by both researchers and educators to understand students' learning process. A systematic coding scheme may be established to fully interpret the field notes and generalized patterns.

Combination of assessment tools. As maker activities can produce various learning outcomes, many of the studies combined multiple assessment tools. For example, Peppler and Glosson (2013) used both pre-posttest and other instruments including surveys, interviews, artifacts, and observations to examine youths' understanding of circuitry in regard to current flow, connections and polarity, and electrical circuits diagram. In a study on using maker activities to promote computational thinking (CT) and engineering learning, Yin, Hadad, Tang, and Lin (2020) measured students' CT skills and engineering learning using knowledge tests and artifacts. In addition, they used a survey to measure self-reported behaviors (such as frequency of using CT skills), dispositions that enhance CT (such as attitudes towards ambiguity).

Similar to the knowledge test, observation has also been combined with other assessment tools, especially artifact assessments (12), to assess diverse learning outcomes. Ryoo, Kali, and Bevan (2016) conducted observations and made field notes in a weekly afterschool making program where students constructed their paper circuit artifact that served as the evidence of individual pathways and problem solving through various making processes. Researchers used observation methods to describe students' everyday experiences of learning and identify the value of making to students learning. Combining with artifact assessment, observation data could provide detailed evidence on how the learning was reflected in the final product.

3.4. Psychometric evidence

To evaluate the assessment tool, we examined whether any psychometric evidence of reliability and/or validity was provided. In our review, only nine studies (15%) reported reliability evidence and six studies (10%) reported validity evidence. The reliability evidence mainly includes Cronbach's alpha for items' internal consistency and Cohen's Kappa for inter-rater reliability. Cronbach's alpha was applied to assess the reliability of the survey scales and when several factors were involved in the measurement. For example, Blikstein et al. (2017) employed Cronbach's alpha to assess the reliability of the three subscales of the survey instrument–General computing familiarity, exploration and fabrication technologies (EFT); information and communication technologies (ICT). They reported that

Cronbach's alpha of ICT, which increased from 0.59 to 0.53 to 0.83 and 0.73 for the revision of the instrument. Cohen's Kappa was reported for multiple coders' reliability in coding qualitative data. After collecting the students' interviews, Giannakos and Jaccheri (2013) read all responses and coded important keywords until categories emerged from similar codes. The researchers discussed and reached consensus in categories. Cohen's Kappa (0.703) was used to evaluate the inter-rater agreement.

Validity evidence was reported in different ways. Yin et al. (2020) validated an internal computational thinking test by correlating it with an external Bebras test. Meanwhile, the convergent validity was examined by correlating scores of different scales that all measure Computational Thinking related constructs. Giannakos and Jaccheri (2013) used three procedures to assess the convergent validity of the measure: (a) Composite reliability of each construct, (b) Item reliability of the measure, and (c) Average Variance Extracted. Bevan et al. (2015) developed and validated a Tinkering Learning Dimensions Framework by reviewing and discussing the video of learners in the Tinkering Studio. Their researchers and practitioners' team improved the construct validity of their measures through iterative discussion and revision.

Instead of providing specific evidence of reliability and/or validity, some studies assumed their assessment tools to be acceptable by referring to the well-developed standards and large-scale tests. However, 85% studies (51) did not report any reliability or validity evidence.

4. Discussion

We conducted this systematic literature review to understand maker-based assessments regarding four aspects. We discuss the major findings and gaps in the following section.

4.1. Features of maker platforms

Our review shows the assessments are most frequently used in maker platforms of e−textiles and LilyPad Arduinos, probably because these platforms can be flexibly integrated into learning objectives (e.g., electric circuits) and curriculum design in the making process. In addition, we found that most platforms have a high demand for technology, (e.g., LilyPad Arduino, 3D printing, and laser cutting). Compared with maker activities with low technology demands, the high demand for technology may make these maker activities less accessible for students in a lower socioeconomic status or in lower grades and more intimidating to the students without sufficient experience in or confidence with technology.

In contrast, platforms based on everyday-life materials or low-technological resources are insufficiently studied or assessed. For example, LEGO, a flexible tool that can be as complex as desired, has been widely used to promote creative thinking and concepts learning long before the maker movement became popular (e.g., Hadjiachilleos, Avraamidou, & Papastavrou, 2013; Özgün-Koca Edwards, & Chelst, 2015). The prevalent use of LEGO in education has philosophical roots in constructionism, in which students are encouraged to acquire and construct knowledge through their experience. Therefore, previous studies have introduced LEGO in learning and teaching based on theoretical framework of constructionism and the effects on learning have been well studied. On the other hand, emerging from the same roots in constructionism, the maker movement is a growing phenomenon that emphasizes digital and physical fabrication and supporting learners in creating personally relevant artifacts. As a result, maker-related studies employ more "high-tech" platforms in maker activity. In our initial search, we found considerable articles using LEGO platforms only (excluding LEGO Mindstorms) to assess learning outcomes, but very few studies were constructed on the basis of maker movement theory that is the core of this study. If well-designed, however, these traditional "low-tech" maker platforms would have great potential to facilitate not only higher order thinking such as creativity and abstract thinking skills, but would also promote maker mindsets and skills such as computational thinking and problem-solving skills in makerspaces (Lotts, 2016). To be scaled up for use, more studies are needed on low-tech maker activities, especially to explore the innovative applications in the current state of maker education as well as to facilitate student engagement, especially for those in lower-grades, with lower socio-economic status and lower technology proficiency.

The majority of studies involved middle school and high school students rather than elementary students or college students. This pattern might be due to the popularity of maker activities in middle school and high school in general. Some maker activities may be too challenging for younger elementary students or too easy for college students. However, maker activities vary in difficulty and complexity and different degrees of scaffolding can also be provided, we suggest that more maker activities and assessments can be used in elementary school and college. For example, teacher guided 3-D printing can be used in elementary schools to improve and assess math and geometry, or Arduino design with less scaffolding can be used among college students to improve and assess programming and engineering skills.

4.2. Domains of learning outcomes

Various learning outcomes have been measured in maker studies, including content knowledge, skills, feeling, attitude, engagement, and collaboration. Among the learning outcomes facilitated in making, cognitive learning outcomes have received more attention than other learning outcomes. In the cognitive skills, we found the STEM-related content knowledge have been measured more than non-STEM area. Only six studies introduced design and art in making, and researchers suggested that current trends in art and design practice can promote the development of non-STEM learning (Clapp & Jimenez, 2016; May and Clapp, 2017). Maker activities can provide fertile ground for non-STEM learning, but its potential has not been fully explored. Thus, future studies should expand to diverse subject areas and explore more assessments to measure learning of non-STEM subject in maker activities as well.

In addition, many studies demonstrated that engaging in maker activities would improve students' confidence and perceptions of making (e.g., Rode et al., 2015). We found that most of the reviewed studies focused on assessing students' affect and attitude toward

STEM/computing more than other subject domains. The popularity of STEM subject domains in the affect assessment might be due to the fact that technology platforms are more commonly used in maker activities than other ones, as discussed earlier. Moreover, due to the over-emphasis on STEM subjects in the current maker studies, attention may be paid to the demographic differences in confidence and attitude, given that female students and students from low socioeconomic status families may be exposed to various technologies less frequently than their counterparts (Margolis, 2010; Rich, Perry, & Guzdial, 2004). School girls typically show less interest in CS topics, which later deters them from studying and attaining a CS career (Tai, Liu, Maltese, & Fan, 2006). Giannakos et al. (2014) has pointed that women more often than men choose disciplines like linguistics, cultural studies, and the arts. Thus, more initiatives should be taken to help underrepresented groups cultivate positive mindset in programming and increase their intention to adopt STEM in the future.

4.3. Utilization of assessments

In the reviewed studies, assessments of artifacts, tests, surveys, interviews, and observations are the most widely-used tools to measure learning outcomes in maker activities. In particular, they are mainly used to measure STEM related knowledge and affective learning outcomes such as students' attitude toward computing (e.g., Tofel-Grehl et al., 2017; Jacobs & Buechley, 2013). In contrast, knowledge tests were under-utilized. Only nine studies employed tests to measured students' knowledge construction in making activities. The lack of knowledge test might be due two reasons: (a) the maker-based assessment primarily target higher-level cognitive abilities such as creativity rather than lower-level abilities such as content knowledge. (b) The arrangements of makerspaces typically lack structure; therefore, it is difficult to administer a traditional knowledge test. However, as stated by Saorín et al. (2017), knowledge construction is an essential part of higher-level cognitive ability such as creative ability and abstract thinking. Therefore, it is helpful to measure knowledge construction and examine how it is associated with students' creativity. A brief knowledge test can also be given to students as an entrance and exit test to understand whether students have sufficient prior knowledge to start making and whether students gain the related knowledge in making activities. The results can be used to evaluate the maker activities and educators' effectiveness. Also, if it shows that maker activities can improve students' knowledge and skills that are valued in traditional tests, the teachers in formal settings, who are overwhelmed by many other mandatory obligations, would also be more encouraged to incorporate maker activities in their teaching.

Comparing to the lack of knowledge test in makerspace, we found that artifact assessments have become prevalent in maker activity, which is not surprising as makerspaces emphasize making products. Those studies, however, evaluated the artifacts as the final products using general guidelines instead of analytical rubrics. A general guideline usually lists the expected outcomes without elaborating on how learning skills or competence were improved in the making process. For example, Vasudevan, Kafai, and Yang (2015) conducted a descriptive analysis of students' final wearable controller and their Scratch code using a simplified framework of categorization to understand how students approached the artifact designs. Researchers use the framework as a general guideline to distinguish between simple and creative designs and evaluate artifacts. Although the framework provides a basic indicator on artifact assessment, it is too general to capture the multiple dimensions of students' on-and off-screen designs.

In contrast, an analytic rubric provides detailed performance criteria to explain how to score each part of the artifact and evaluate the maker's performance comprehensively. For example, Lin, Yin, Tang, and Hadad (2018) created an analytic rubric for e-textile projects and broke down the characteristics of the artifact into several dimensions: the functions of circuit, appearance of the product, and connection to the community. For each dimension, students' products are given a score with multiple points. This type of analytic rubric allows the scorer to itemize and examine what parts are strong and what parts need improvement. In addition, an analytic rubric could not only help researchers and educators measure students' learning outcomes and making ability with more details, it also guides students to make their artifacts. However, in the reviewed studies, general guidelines are lacking in artifacts assessments and very few studies employ analytic rubrics. Thus, it is unclear how students' skills and competence were demonstrated in their artifacts. Although it is challenging to develop full-scale assessment for artifacts due to its diverse forms and usage of different materials, it is important to construct comprehensive guideline and/or rubrics that can help researchers and educators evaluate the learning outcomes demonstrated in artifacts.

4.4. Psychometric evidence

A majority of the studies reported limited evidence for test reliability and validity. When the measurement tool consists of several factors and scales, internal consistency should be examined to ensure the scores are reliable. When multiple raters are involved to score students' maker performance, inter-rater agreement such as Cohen's Kappa, or interrater reliability should be examined. However, about 85% of the reviewed studies did not report reliability or validity evidence. As the maker assessments are emergent tools in the field and maker activities are expected to improve multiple and complex learning outcomes, it is essential to systematically document the test reliability and validity, so that the learning outcomes from the maker activities can be measured with rigor.

5. Conclusion and implications

5.1. Conclusion and suggestions

Through reviewing the assessment studies of maker activities, our study provides a comprehensive review to inform the research, design, and use of assessments in making settings. We identify the achievements and gaps associated with assessment of maker activities and provide suggestions for future research. Below we conclude with our major suggestions.

First, we suggest that more assessments should be used in maker activities. Although we initially found hundreds of papers related to making and makerspace, when we narrowed it down to assessments used in maker activities, only 60 were left. Most of the studies focus on the philosophy and design of the maker movement and makerspaces instead of investigating what students have learned from the maker activities. More empirical studies with assessments should be conducted to examine the impact of maker activities on students' learning, which can provide formative assessments to improve those activities and implementation.

Second, learning outcomes assessed in maker activities should be expanded to non-STEM-related ones (such as art and design) and non-cognitive learning outcomes (such as affective measures and social skills), which are expected to develop in makerspaces as well. Maker activities provide an environment and opportunity to enrich multidimensional learning. Therefore, assessments should be designed and administered accordingly to capture the rich learning outcomes.

Third, assessments with various forms should be used to measure the diverse learning outcomes in maker activities. Knowledge tests may be used more in makerspaces to examine students' prerequisite knowledge and gained knowledge in maker activities. Performance assessments could also be used more to examine students' problem-solving procedures. Artifact assessment is preferable in maker activities for it provides the opportunity for educators to understand the learning demonstrated in the maker products. In addition, more specific scoring rubrics should be developed to evaluate students' final products and guide students' making process.

Four, more psychometric evidence of reliability and validity should be provided for the maker assessment to ensure its rigor, especially if the assessment tools are used to evaluate the effectiveness of maker activity design and implementation.

5.2. Implications for research and practice

The implication of this systematic review is threefold. First, the overview of maker platforms in this study can help researchers and educators better understand the features of platforms as well as their implementations in making and learning. The 60 reviewed studies reflect the current usage of platforms in maker activities. In particular, more than 50% of the prevalent platforms are computer-based such as LilyPad Arduinos, Makey Makey and Scratch, and nearly all of the maker platforms are related to STEM subject. Thus, educators can adopt appropriate maker activity platforms for their STEM teaching, especially computer and engineering. As Martin (2015) indicated, digital technology could substantially lower the barriers to engaging students in physical computing as well as offering robust pathways for them to learn about engineering and programming. As maker activities tend to produce more tangible products than pure computer programming, they can serve as an ideal stepping stone for students to overcome their initial fear of computer programming.

Second, this systematic review reveals that varying types of learning outcomes can be achieved using maker activities, which answers "what can be learned in making," the question raised by researchers (e.g., Sheridan et al., 2014). The review shows that the making processes can greatly facilitate cognitive learning outcomes and non-cognitive learning outcomes (such as self-efficacy) related to STEM, especially in the fields of computing and engineering, educators may develop maker activities and corresponding assessments to target these learning outcomes. In addition, researcher may also consider assessing the learning outcomes that are promoted in make activities but insufficiently measured in literature (e.g., communication skill and decision-making ability).

Third, this study shows that both traditional and innovative assessment tools can be adapted in maker activities to assess and improve learning. It addresses researchers' concerns that grading students' work is likely to interrupt the learning processes in making (Martinez & Stager, 2013). The utilization of varying assessment tools (Fig. 1) implies that researchers and educators can still design appropriate assessments and actively employ them in maker activities to evaluate students' learning outcomes. Like all the other assessments, the measurement of learning outcomes in maker activities can be used to provide feedback to students/mentors, improve maker activity design/implementation, and improve maker programs. Therefore, assessment tools would be invaluable if embedded in maker education, which aligns with the findings in Murai et al. (2019) that assessment tools enable the teachers to identify students' different learning styles and provide immediate support for them.

To fully realize the potential of maker activities, researchers need to integrate diverse platforms on all kinds of subject domains, develop systematic learning framework and indicators, and employ appropriate and rigorous assessments to measure learning outcomes. As assessments can be used to evaluate educational programs and teaching methods (summative) and provide constructive feedback to educational programs, teachers, and students (formative), it is critical to further the development of assessments to evaluate and understand the learning process in making. Future research needs to pay attention to not only developing and using suitable assessments, but also validating the assessments to make them benefit more researchers, educators, and students.

Funding

This work was supported by the National Science Foundation (NSF) under grant with an ID of 1543124.

Declaration of competing interest

None.

CRediT authorship contribution statement

Qiao Lin: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing - original draft. **Yue Yin:** Funding acquisition, Supervision, Conceptualization, Data curation, Writing - review & editing. **Xiaodan Tang:** Formal analysis, Writing - review & editing. **Roxana Hadad:** Supervision, Project administration, Funding acquisition, Writing - review & editing. **Xiaoming Zhai:** Conceptualization, Writing - review & editing.

Acknowledgement

This material is based upon work supported by the National Science Foundation (NSF) under grant with an ID of 1543124. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

Appendix A. The reviewed maker activity assessment studies

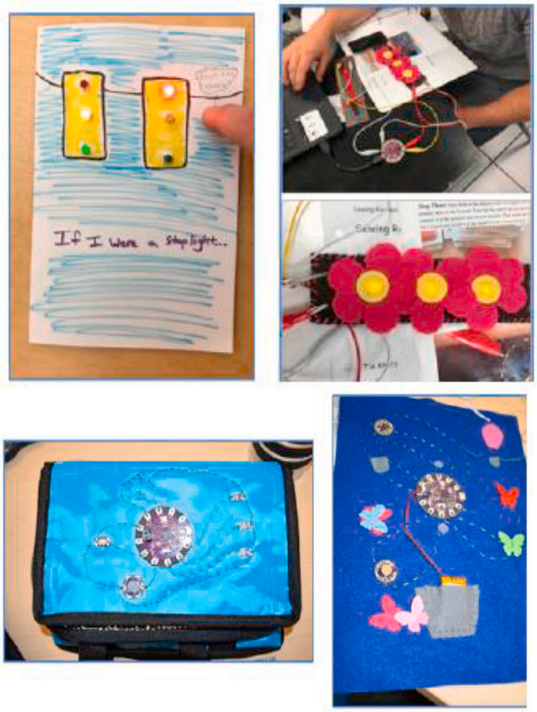
Authors	platform used for learning	Subject domain involved	Setting	age range	Assessment tools
Barton et al. (2016)	e-textile	STEM	informal	ages 10-14	artifacts, interviews
Bevan et al. (2015)	tinkering materials	STEM	informal	a broad age range	observation
Blikstein et al. (2017)	Fablabs	Technological literacy	formal	grade 4-12	survey
Brady et al. (2017)	PartSims, agent-based modeling	CS	formal	7th to 12th grades	survey, interviews artifacts
Buechley, Eisenberg, & Elumeze (2007)	e-textile, Arduino	electronics, programming	formal	14–17	test, artifacts
Buechley, Eisenberg, Catchen, & Crockett (2008)	LilyPad Arduino	electronics, programming	informal	10–14	survey, artifacts
Chamrat (2018)	electrical circuits modules	programming	informal	middle school students	test
Chu, Quek, Bhangaonkar, Ging, & Sridharamurthy (2015)	Maker Theater	electronics	informal	elementary-school (aged 8 to 11)	survey, interviews
Chu et al. (2017)	tinkering materials	electronics	formal	8–11 years old	survey, observation
Lui et al. (2019)	E-textile	electronics	formal	high school students	artifacts
Flores (2018)	3D printing	STEM	formal	5th and 6th grader	survey, interviews, observation
Fordyce, Heemsbergen, Mignone, and Nansen (2015)	3D printing	STEM	formal	university students	semi-structured interviews, surveys
Giannakos and Jaccheri (2013)	Scratch, Scratch for Arduino	programming	informal	12–18 years old	interview. Observation, survey
Giannakos et al. (2014)	Scratch, Arduino	programming	informal	12 years old	survey
Tofel-Grehl et al. (2017)	e-textile	electronics	formal	grade 8	pre-posttest, surveys, observation, interviews
Holbert (2016)	graphics design software, laser cutter	design	informal	9–10 years old	artifacts, survey, interviews, observation
Huang and Lin (2017)	3D printing	STEM	formal	college students	test
Hughes & Morrison (2018)	e-textile, LilyPad Arduino	electronics	formal and informal	8–13 years old	interviews, observation
Jacobs and Buechley (2013)	Codeable Objects	programming	informal	graduate students; 11–17 years old	pre-post surveys, interviews, artifacts
Kafai and Vasudevan (2015a)	Makey Makey, Scratch	programming	formal	11–14 years	artifacts, interview
Kafai and Vasudevan (2015b)	MaKey MaKey, Scratch	programming	informal	13–15 years old	artifacts, observation
Kafai et al. (2013)	e-textiles, Lilypad Arduino	electronics	formal	16–18 years old	artifacts, semi-structured pre-post interviews, portfolios
Kafai et al. (2014)	e-textile, LilyPad Arduino	electronics	formal	16–18 years old	pre post interview, performance, artifacts, portfolio assessment
Kafai et al. (2014)	e-textile, LilyPad Arduino	programming	semiformal	14–15 years old	observation, interviews
Kafai, Searle, Martinez, & Brayboy (2014)	e-textile, LilyPad Arduino	electronics, programming	formal and informal	12–15 years old youth	artifacts, interviews
Katterfeldt, Cukurova, Spikol, & Cuartielles (2018)	Talkoo toolkit	CS, programming	formal and informal	14–18 years old	survey, interviews, observation
Kostakis et al. (2015)	3D design and printing	technological literacy	formal	15-16 years-old	survey, semi-structured interviews, observation, artifact
Lee and Fields (2013)	e-txtile, LilyPad Arduino	STEM and programming	formal	undergraduate students	performance, interview
Lee et al. (2014)	MaKey Makey, Scrath	programming	formal	10–12 years	

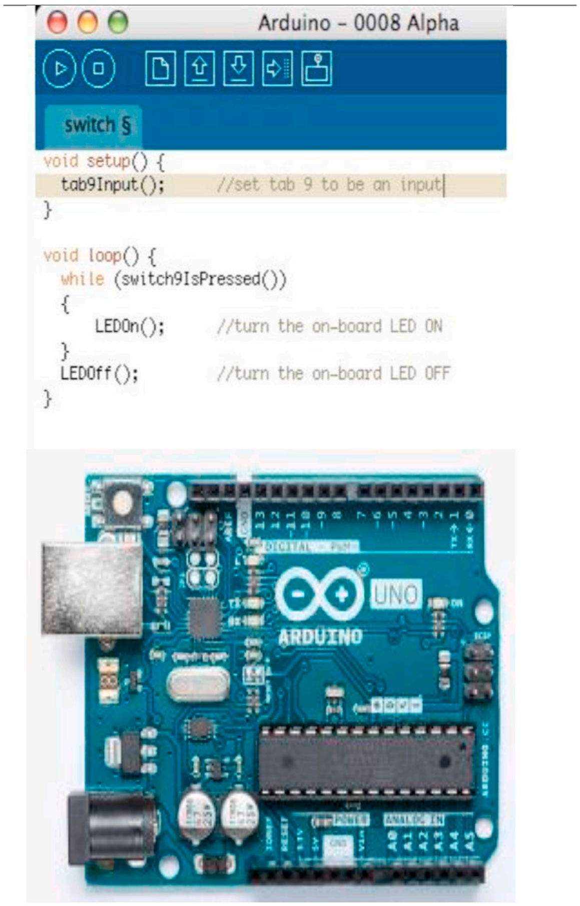
(continued on next page)

(*continued*)

Authors	platform used for learning	Subject domain involved	Setting	age range	Assessment tools
					observation artifacts, interview
Lille & Romero (2017)	SmartCity Maker	educational technologies	informal	college students	artifacts
Litts et al. (2017)	LilyPad Arduino board, e-textile	CS and electronics	formal	16–17 years old	interview, artifacts
Lotts (2016)	LEGO	No subject domain	informal	adults	survey observation
Mellis and Buechley (2012)	laser cutter, 3D printing	STEM	informal	adult	pre-post survey
Leduc-Mills & Eisenberg (2011)	3D printing, UCube	STEM	informal	12–14 years old	observation, usability test
Moriwaki et al. (2012)	Squishy Circuits	STEM and art	informal	4–11 years old	observation, interview
Mylonas et al. (2019)	educational lab kit and IoT infrastructure of GAIA	STEM	formal	11–15 years old	survey, interviews
Nation and Durán (2019)	circuits with microprocessor and conductive ink	electronics	informal	12–15 years old	interviews
Nemorin (2017)	3D printing, Google ShetchUp	STEM	formal	9th grader	artifacts, interviews, observation
Parekh and Gee (2018)	everyday materials	STEM	informal	4–12 years old	interviews
Peppler and Glosson (2013)	e-textile	electronics	informal	9–10 years old	pre posttest, surveys, interviews, artifacts, observation
Qi and Buechley (2014)	ATTINY85	electronics	informal	19-21, 18–26, 18–37, 15–50 years old	artifacts, observation
Qiu, Buechley, Baafi, & Dubow (2013)	Modkit, LilyPad ProtoSnap, Arduino	CS	informal	13–16 years old, 12–17 years old	pre-post survey
Rode et al. (2015)	e-textiles, LilyPad Arduino, Ardublock	electronics and programming	formal	8–10 years old	observation
Ryoo et al. (2016)	paper circuits	electronics	informal	5th grade	artifacts, observation, interview
Saorín et al. (2017)	3D printing	STEM	formal	freshman	test, survey
Schlegel et al. (2019)	3D design, block-based programming	STEM	formal	8–11 years old	survey
Schneider et al. (2015)	EarExplorer	biology	formal	22.5 years old	pre-mid-posttest
Schwartz et al. (2013)	Minecraft	STEM	formal	undergraduates, grade K2-5, grades 6-7	observation
Searle and Kafai (2015)	e-textile, LilyPad arduino	electronics and programming	formal	12–14 years old	artifacts interview, portfolio assessment
Searle et al. (2014)	LilyPad Arduino, e-textile	electronics	formal	16–18 years old	pre-post interview
Searle, Litts, & Kafai (2018)	e-textiles	STEM	formal and informal	high school freshmen	artifacts, interviews, observation
Searle et al. (2019)	e-textiles	electronics	formal	13–14	observation, interviews, artifacts
Sheridan et al. (2014)	3-D printing laser cut	programming	informal	adult; 8–19 years old	observation, interviews, artifacts
Smith, Iversen, & Hjorth (2015)	3D printer, Arduino, Makey Makey	digital fabrication and design	formal	11–15 years old	artifacts, observation, interviews
Somanath et al. (2016)	LilyPads, Makey-Makey, and Arduino Starter Kits	electronics, programming	informal	12–14 years old	survey, interview, observation
Torralba (2019)	Arduino	mathematics	formal	7th grade	artifacts
Vasudevan et al. (2015)	MaKey MaKey, Scratch	electronics, programming	informal	middle school youth	artifacts
Wargo & Alvarado (2020)	squishy circuits, analogue technologies	early literacy	formal	3–9 years old	interviews
Perner-Wilson, Buechley, & Satomi (2010)	e-textile	electronics	informal	20–57	survey
Yin et al. (2020)	E-textile, Makey Makey, Arduino	STEM	informal	high school student	test, survey, artifacts

Appendix B. Commonly Used Maker Activity Platforms

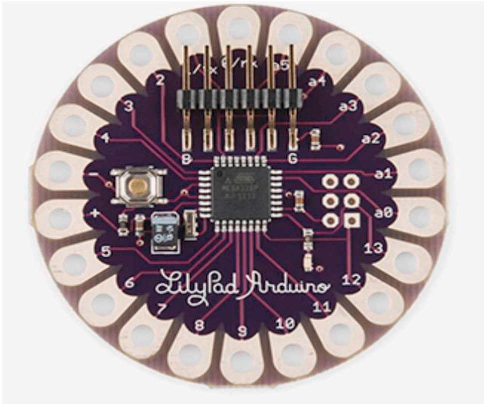
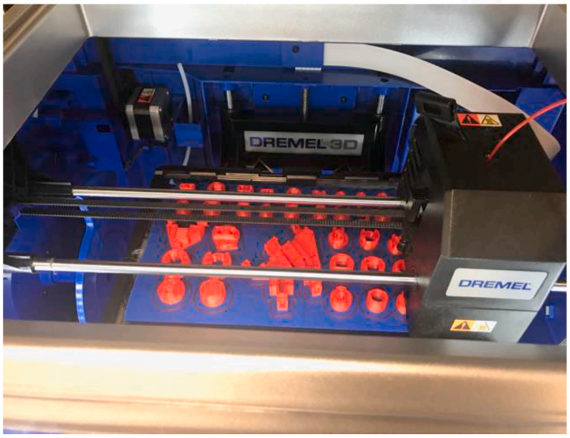
<table>
<tr><th>Platform</th><th>Figure</th><th>Description</th></tr>
<tr><td>E-textile</td><td>

(Searle et al., 2019)</td><td>Electronic textile (E-textile) is a type of fabric that contains electronic elements. E-textiles materials consist of a small, flat, programmable microcontroller; a power source (battery); and a variety of sewable sensors and actuators, such as LED lights; using a needle and conductive tread.</td></tr>
<tr><td>Arduino</td><td>

(https://www.arduino.cc/)</td><td>Arduino is an open-source electronics platform based on easy-to-use hardware and software. Its major steps include hardware design/implementation and software design/implementation that is similar to C language.</td></tr>
</table>

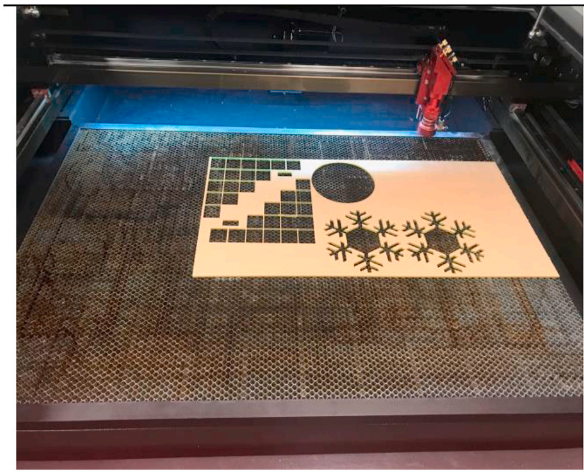
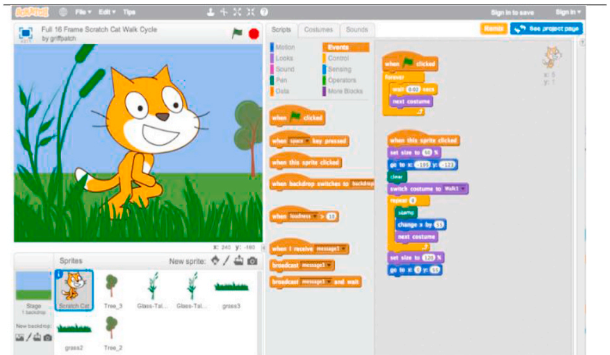
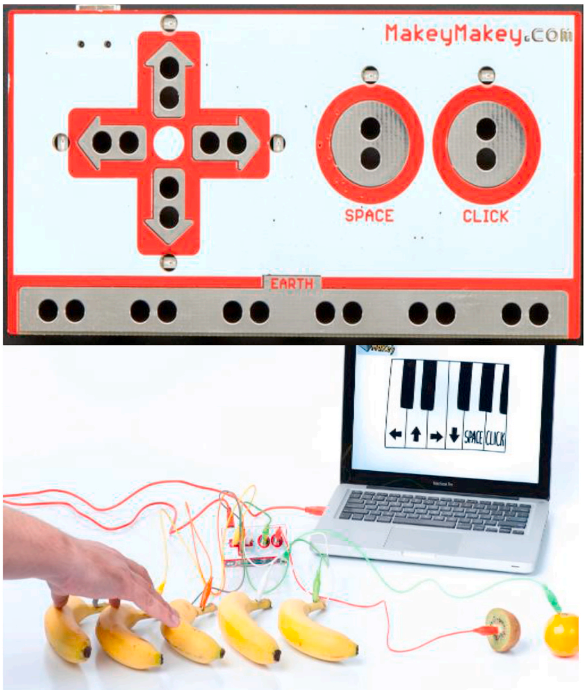
(continued on next page)

(*continued*)

Platform	Figure	Description
LilyPad	 (https://www.arduino.cc/en/Main/ArduinoBoardLilyPad/)	The LilyPad Arduino is Construction Kit that is designed for e-textiles and wearables projects. It can be sewn to fabric and similarly mounted power supplies, sensors and actuators with conductive thread. Using Arduino software, participants can program the LilyPad microcontroller to manage sensor and output modules employed in the artifacts.
3D Printing	  (https://makerspace.uic.edu/)	3D printing, also known as additive manufacturing, is a process of making three dimensional solid objects from a digital file. In 3D printing, an object is created by laying down successive layers of material until the object is created. Each of these layers can be seen as a thinly sliced horizontal cross-section of the object.

(*continued on next page*)

(*continued*)

Platform	Figure	Description
Laser Cutting	 (https://makerspace.uic.edu/)	Laser cutting is a technology that uses a high-power laser optics to cut materials. The laser optics and CNC (computer numerical control) are used to direct the material or the laser beam generated.
Scratch	 (https://www.aace.org/review/prepare-for-fun-scratch-3-0-is-coming/)	Scratch is a block-based visual programming language that was developed by MIT Media Lab, which primarily targets at children. Users can create interactive stories, games, and animations using a block-like interface and share in the online community.
Makey Makey	 (https://makeymakey.com/)	Makey Makey is an electronic invention kit that allows users to connect everyday conductive objects to create a tangible user interface that controls any software running on a computer. Using a circuit board, alligator clips, and a USB cable, Makey Makey uses closed loop electrical signals to send the computer either a keyboard stroke or mouse click signal.

Appendix C. Sample Questions for Each Assessment Tool

I. Artifact Assessment

Statements of the assessment criteria for the final project of 3D car:

The task: Design and manufacture an electronic car that is functional, accurate, and visually appealing.

Function: Car must travel 20 m.

Time: 7 weeks to produce.

Cost: Not exceeding a budget of $20 to produce.

Design criteria: Should be in color and wood or 3D (using basic sketching techniques). Components can be designed using SketchUp (CAD program).

Product criteria: Must use 3 of the main processes:

- Measuring and marking out.
- Sawing and/or cutting.
- Shaping and/or forming.
- Joining.
- Sanding and/or finishing.

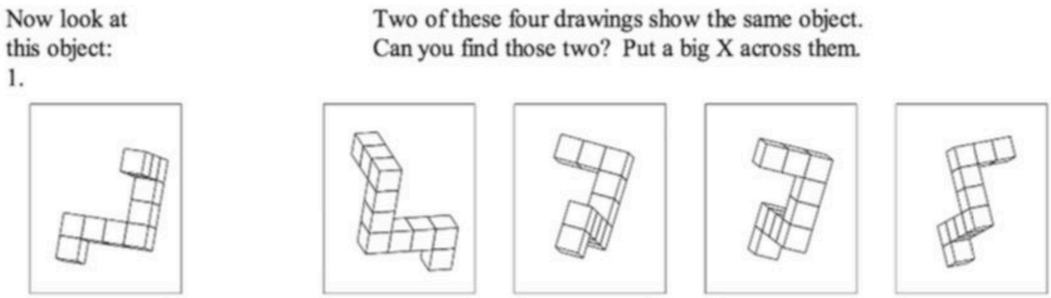
Size criteria: Separate components must be no larger than 120 mm 9120 mm 9120 mm.

Materials criteria: Any of the following—balsa wood, 3D printed resin, glue, metal, and acrylic sheet.

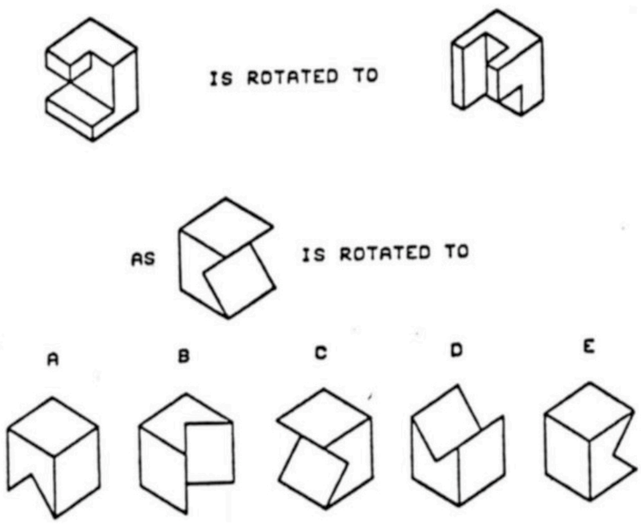
Source: Nemorin (2017).

II. Test

Sample items from the mental rotation test (MRT).



Sample items from the Purdue Spatial Visualization Test: Rotations (PSVT: R).



Source: Huang and Lin (2017).

III. Survey

Teacher.

1. My teacher cares if I think science is interesting.
2. My teacher cares if I learn science.

Family.

3. It is important to my family that I try my best in school.
4. My family would be happy if I chose to pursue a career in science, technology, or engineering.

Peer.

5. How many of your close friends: like science, technology, or engineering?
6. How many of your close friends: think science, technology, or engineering is cool?
7. How many of your close friends: get good grades in science, technology, or engineering courses?
8. How many of your close friends: care about your grades in school?
9. How many of your close friends: encourage you to do well in class?

Note: It's a 4-point Likert-type scale that ranged from 1 = agree strongly to 4 = disagree strongly.
Source: Tofel-Grehl et al. (2017).

IIII. Interview

Sample Interview Questions:
McKell and the Duck.

1. My question is how does it do that?
2. What do you mean by a chemical in the product?
3. Does that involve chemicals too because you were just talking about it with the hand warmers?
4. I have this rubber duck toy … [that lights up] My question for you is how does it do that thing [light up when placed in water]?
5. You said there was sensors [sic] and the water conducts. Explain what you mean by conductive.

McKell and the Elephant.

6. What sorts of materials do you think you would need to take the elephant and make it do that stuff? [light up when the paws are touched]
7. How do they talk?
8. What types of material do you think I would need in order to make that happen?
9. Where would the LilyPad go?

Source: Lee and Fields (2013).

References

Barton, A. C., Tan, E., & Greenberg, D. (2016). The makerspace movement: Sites of possibilities for equitable opportunities to engage underrepresented youth in STEM. *Teachers College Record, 119*(6), 11–44.
Bevan, B., Gutwill, J. P., Petrich, M., & Wilkinson, K. (2015). Learning through STEM-rich tinkering: Findings from a jointly negotiated research project taken up in practice. *Science Education, 99*(1), 98–120.
Blikstein, P., Kabayadondo, Z., Martin, A., & Fields, D. (2017). An assessment instrument of technological literacies in makerspaces and fabLabs. *Journal of Engineering Education, 106*(1), 149–175.
Brady, C., Orton, K., Weintrop, D., Anton, G., Rodriguez, S., & Wilensky, U. (2017). All roads lead to computing: Making, participatory simulations, and social computing as pathways to computer science. *IEEE Transactions on Education, 60*(1), 59–66.
Buechley, L., Peppler, K., Eisenberg, M., & Kafai, Y. (2013). Textile messages: Dispatches from the world of E-textiles and education. *New Literacies and Digital Epistemologies, ume 62* (ERIC).
Buechley, Leah, Eisenberg, Mike, & Elumeze, Nwanua (2007). Towards a curriculum for electronic textiles in the high school classroom. *Proceedings of the 12th annual SIGCSE conference on Innovation and technology in computer science education*.
Buechley, Leah, Eisenberg, Mike, Catchen, Jaime, & Crockett, Ali (2008). The LilyPad Arduino: using computational textiles to investigate engagement, aesthetics, and diversity in computer science education. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.

Chamrat, S. (2018). The Science Camp Model based on maker movement and tinkering activity for developing concept of electricity in middle school students to meet standard evaluation of ordinary national educational test (O-NET). In *Paper presented at the AIP conference proceedings*.

Cheng, M. F., & Brown, D. E. (2015). The role of scientific modeling criteria in advancing students' explanatory ideas of magnetism. *Journal of Research in Science Teaching, 52*(8), 1053–1081. https://doi.org/10.1002/tea.21234.

Chu, Sharon Lynn, Quek, Francis, Bhangaonkar, Sourabh, Ging, Amy Boettcher, & Sridharamurthy, Kumar (2015). Making the Maker: A Means-to-an-Ends Approach to Nurturing the Maker Mindset in Elemen- tary-Aged Children. *International Journal of Child-Computer Interaction, 5*, 11–19.

Chu, S. L., Angello, G., Saenz, M., & Quek, F. (2017). Fun in Making: Understanding the experience of fun and learning through curriculum-based Making in the elementary school classroom. *Entertainment Computing, 18*, 31–40.

Clapp, E. P., & Jimenez, R. L. (2016). Implementing STEAM in maker-centered learning. *Psychology of Aesthetics, Creativity, and the Arts, 10*(4), 481.

Cun, A., Abramovich, S., & Smith, J. M. (2019). An assessment matrix for library makerspaces. *Library & Information Science Research, 41*(1), 39–47.

Dewey, J. (1923). *Democracy and education: An introduction to the philosophy of education*. Macmillan.

Dewey, J. (1929). *The quest for certainty: A study of the relation of knowledge and action*. New York, NY: Minton, Balch and Company.

Dixson, D. D., & Worrell, F. C. (2016). Formative and summative assessment in the classroom. *Theory into practice, 55*(2), 153–159.

Elby, A., & Hammer, D. (2001). On the substance of a sophisticated epistemology. *Science Education, 85*(5), 554–567.

Fleming, L. (2015). *Worlds of making: Best practices for establishing a makerspace for your school*. Corwin Press.

Flores, Christa (2018). Problem-based science, a constructionist approach to science literacy in middle school. *International Journal of Child-Computer Interaction, 16*, 25–30.

Fordyce, R., Heemsbergen, L., Mignone, P., & Nansen, B. (2015). 3D printing and university makerspaces: Surveying countercultural communities in institutional settings. *Digital Culture & Education, 7*(2).

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2011). *How to design and evaluate research in education*. New York: McGraw-Hill Humanities/Social Sciences/Languages.

Fredrick, K. (2015). Making it matters. *School Library Monthly, 31*(7), 22–24.

Freire, P. (1974). *Pedagogy of the oppressed Hammondsworth*. Penguin Books.

Giannakos, M. N., & Jaccheri, L. (2013). What motivates children to become creators of digital enriched artifacts?. In *Paper presented at the proceedings of the 9th ACM conference on creativity & cognition*.

Giannakos, M. N., Jaccheri, L., & Leftheriotis, I. (2014). Happy girls engaging with technology: Assessing emotions and engagement related to programming activities. In *Paper presented at the international conference on learning and collaboration technologies*.

Gilmartin, S. K., Li, E., & Aschbacher, P. (2006). The relationship between interest in physical science/engineering, science class experiences, and family contexts: Variations by gender and race/ethnicity among secondary students. *Journal of Women and Minorities in Science and Engineering, 12*(2-3).

Godhe, A.-L., Lilja, P., & Selwyn, N. (2019). Making sense of making: Critical issues in the integration of maker education into schools. *Technology, Pedagogy and Education*, 1–12.

Hadjiachilleos, S., Avraamidou, L., & Papastavrou, S. (2013). The use of lego technologies in elementary teacher preparation. *Journal of Science Education and Technology, 22*(5), 614–629.

Halverson, E. R., & Sheridan, K. (2014). The maker movement in education. *Harvard Educational Review, 84*(4), 495–504.

Harel, I. E., & Papert, S. E. (1991). *Constructionism: Ablex publishing*.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127.

Holbert, N. (2016). Leveraging cultural values and "ways of knowing" to increase diversity in maker activities. *International journal of child-computer interaction, 9*, 33–39.

Huang, T.-C., & Lin, C.-Y. (2017). From 3D modeling to 3D printing: Development of a differentiated spatial ability teaching model. *Telematics and Informatics, 34*(2), 604–613.

Hughes, Janette, & Morrison, Laura (2018). The Use of E-Textiles in Ontario Education. *Canadian Journal of Education*.

Humphrey, T., & Gutwill, J. P. (2017). *Fostering active prolonged engagement: The art of creating APE exhibits*. Routledge.

Jacobs, J., & Buechley, L. (2013). Codeable objects: Computational design and digital fabrication for novice programmers. In *Paper presented at the proceedings of the SIGCHI conference on human factors in computing systems*.

Kafai, Y., Fields, D., & Searle, K. (2014). Electronic textiles as disruptive designs: Supporting and challenging maker activities in schools. *Harvard Educational Review, 84*(4), 532–556.

Kafai, Y., Peppler, K., & Chapman, R. (2009). *The computer clubhouse: Constructionism and creativity in youth communities*. Technology, Education–Connections: ERIC.

Kafai, Y., Searle, K., Kaplan, E., Fields, D., Lee, E., & Lui, D. (2013). Cupcake cushions, scooby doo shirts, and soft boomboxes: E-Textiles in high school to promote computational concepts, practices, and perceptions. In *Paper presented at the Proceeding of the 44th ACM technical symposium on Computer science education*.

Kafai, Y., & Vasudevan, V. (2015a). Constructionist gaming beyond the screen: Middle school students' crafting and computing of touchpads, board games, and controllers. In *Paper presented at the proceedings of the workshop in primary and secondary computing education*.

Kafai, Y., & Vasudevan, V. (2015b). Hi-lo tech games: Crafting, coding and collaboration of augmented board games by high school youth. In *Paper presented at the proceedings of the 14th international conference on interaction design and children*.

Kafai, Yasmin, Lee, Eunkyoung, Searle, Kristin, Fields, Deborah, Kaplan, Eliot, & Lui, Debora (2014). A Crafts-Oriented Approach to Computing in High School: Introducing Computational Concepts, Practices, and Perspectives with Electronic Textiles. *ACM Transactions on Computing Education (TOCE), 14*.

Kafai, Yasmin, Searle, Kristin, Martinez, Crîstobal, & Brayboy, Bryan (2014). Ethnocomputing with Electronic Textiles: Culturally Responsive Open Design to Broaden Participation in Computing in American Indian Youth and Communities. *Proceedings of the 45th ACM technical symposium on Computer science education*, 241–246.

Katterfeldt, Eva-Sophie, Cukurova, Mutlu, Spikol, Daniel, & Cuartielles, David (2018). Physical computing with plug-and-play toolkits/ Key recommendations for collaborative learning implementations. *International Journal of Child-Computer Interaction, 17*, 72–82.

Keune, A., & Peppler, K. (2019). Materials-to-develop-with: The making of a makerspace. *British Journal of Educational Technology, 50*(1), 280–293.

Kitchenham, B. (2004). *Procedures for performing systematic reviews* (Vol. 33, pp. 1–26). Keele, UK: Keele University, 2004.

Kostakis, V., Niaros, V., & Giotitsas, C. (2015). Open source 3D printing as a means of learning: An educational experiment in two high schools in Greece. *Telematics and Informatics, 32*(1), 118–128.

Kurti, R. S., Kurti, D. L., & Fleming, L. (2014). The philosophy of educational makerspaces part 1 of making an educational makerspace. *Teacher Librarian, 41*(5), 8.

Leduc-Mills, Ben, & Eisenberg, Michael (2011). The UCube: a child-friendly device for introductory three-dimensional design. *proceedings of the 10th international conference on interaction design and children*, 72–80.

Lee, V., & Fields, D. (2013). *A clinical interview for assessing student learning in a university-level craft technology course*.

Lee, Kafai, Y., Vasudevan, V., & Davis, R. L. (2014). Playing in the arcade: Designing tangible interfaces with MaKey MaKey for scratch games. In *Playful user interfaces* (pp. 277–292). Springer.

Lille, Benjamin, & Romero, Margarida (2017). Creativity Assessment in the Context of Maker-based Projects. *Design and Technology Education: an International Journal, 22*, 32–47.

Lin, Q., Yin, Y., Tang, X., & Hadad, R. (2018). *Assessing circuit-based artifacts in maker activities*. San Francisco: Paper presented at the American Psychological Association.

Litts, B., Kafai, Y., Lui, D., Walker, J., & Widman, S. (2017). Stitching codeable circuits: High school students' learning about circuitry and coding with electronic textiles. *Journal of Science Education and Technology*, 1–14.

Lotts, M. (2016). On the road, playing with Legos®, and learning about the library: The rutgers university art library lego playing station, Part Two. *Journal of Library Administration, 56*(5), 499–525.

Lui, D., Fields, D., & Kafai, Y. (2019). Student maker portfolios: Promoting computational communication and reflection in crafting E-textiles. In *Paper presented at the proceedings of FabLearn 2019*.

Lundberg, M., & Rasmussen, J. (2018). Foundational principles and practices to consider in assessing maker education. *Journal of Educational Technology, 14*(4), 1–12.

Margolis, J. (2010). *Stuck in the shallow end: Education, race, and computing*. MIT Press.
Martin, L. (2015). The promise of the Maker Movement for education. *Journal of Pre-College Engineering Education Research (J-PEER), 5*(1), 4.
Martin, L., & Dixon, C. (2013). *Youth conceptions of making and the maker movement*. New York: Paper presented at the Interaction Design and Children.
Martinez, S. L., & Stager, G. (2013). *Invent to learn. Making, tinkering, and Engineering in the classroom. Torrance, Canada: Construting modern knowledge*.
May, S., & Clapp, E. P. (2017). Considering the role of the arts and aesthetics within maker-centered learning. *Studies in Art Education, 58*(4), 335–350.
Mellis, D. A., & Buechley, L. (2012). Case studies in the personal fabrication of electronic products. In *Paper presented at the proceedings of the designing interactive systems conference*.
Moriwaki, K., Brucker-Cohen, J., Campbell, L., Saavedra, J., Stark, L., & Taylor, L. (2012). Scrapyard Challenge Jr., Adapting an art and design workshop to support STEM to STEAM learning experiences. In *Paper presented at the integrated STEM education conference (ISEC), 2012 IEEE 2nd*.
Murai, Y., Kim, Y. J., Martin, E., Kirschmann, P., Rosenheck, L., & Reich, J. (2019). Embedding assessment in school-based making: Preliminary exploration of principles for embedded assessment in maker learning. In *Paper presented at the proceedings of FabLearn 2019, New York, NY, USA*.
Mylonas, Georgios, Amaxilatis, Dimitrios, Pocero, Lidia, Markelis, Iraklis, Hofstaetter, Joerg, & Koulouris, Pavlos (2019). An educational IoT lab kit and tools for energy awareness in European schools. *International Journal of Child-Computer Interaction, 20*, 43–53.
Nation, J. M., & Durán, R. P. (2019). Home is where the heart is: Latinx youth expression and identity in a critical maker project. *Mind, Culture and Activity, 26*(3), 249–265.
Nemorin, S. (2017). The frustrations of digital fabrication: An auto/ethnographic exploration of '3D making'in school. *International Journal of Technology and Design Education, 27*(4), 517–535.
O'Brien, S., Hansen, A. K., & Harlow, D. B. (2016). Educating teachers for the maker movement: Pre-service teachers' experiences facilitating maker activities. In *Paper presented at the proceedings of the 6th annual conference on creativity and fabrication in education*.
Özgün-Koca, S. A., Edwards, T. G., & Chelst, K. R. (2015). Linking LEGO and algebra. *Mathematics Teaching in the Middle School, 20*(7), 400–407.
Papavlasopoulou, S., Giannakos, M. N., & Jaccheri, L. (2017). Empirical studies on the maker movement, a promising approach to learning: A literature review. *Entertainment Computing, 18*, 57–78.
Parekh, P., & Gee, E. R. (2018). Zooming into a tinkering project: The progression of learning through transitional objects. *Interdisciplinary Journal of Problem-Based Learning, 12*(2).
Peppler, K., & Glosson, D. (2013). Stitching circuits: Learning about circuitry through e-textile materials. *Journal of Science Education and Technology, 22*(5), 751–763.
Perkins, D., Tishman, S., Ritchhart, R., Donis, K., & Andrade, A. (2000). Intelligence in the wild: A dispositional view of intellectual traits. *Educational Psychology Review, 12*(3), 269–293.
Perner-Wilson, Hannah, Buechley, Leah, & Satomi, Mika (2010). Handcrafting Textile Interfaces from A Kit-of-No-Parts. *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*.
Piaget, J. (1950). *The psychology of intelligence*. New York, NY: Harcourt Brace.
Qi, J., & Buechley, L. (2014). Sketching in circuits: Designing and building electronics on paper. In *Paper presented at the proceedings of the SIGCHI conference on human factors in computing systems*.
Qiu, Kanjun, Buechley, Leah, Baafi, Edward, & Dubow, Wendy (2013). A Curriculum for Teaching Computer Science Through Computational Textiles. *Proceedings of the 12th international conference on interaction design and children*.
Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., & Silverman, B. (2009). Scratch: Programming for all. *Communications of the ACM, 52*(11), 60–67.
Rich, L., Perry, H., & Guzdial, M. (2004). A CS1 course designed to address interests of women. In *Paper presented at the ACM SIGCSE bulletin*.
Rivas, L. (2014). Creating a classroom makerspace. *Educational Horizons, 93*(1), 25–26.
Rode, J. A., Weibert, A., Marshall, A., Aal, K., von Rekowski, T., El Mimouni, H., et al. (2015). From computational thinking to computational making. In *Paper presented at the proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*.
Ryoo, J. J., Kali, L., & Bevan, B. (2016). Equity-oriented pedagogical strategies and student learning in after school making. In *Paper presented at the proceedings of the 6th annual conference on creativity and fabrication in education*.
Sang, W., & Simpson, A. (2019). The maker movement: A global movement for educational change. *International Journal of Science and Mathematics Education*, 1–19.
Saorín, J. L., Melian-Díaz, D., Bonnet, A., Carrera, C. C., Meier, C., & De La Torre-Cantero, J. (2017). Makerspace teaching-learning environment to enhance creative competence in engineering students. *Thinking Skills and Creativity, 23*, 188–198.
Sawyer, R. K. (2005). *The Cambridge handbook of the learning sciences*. Cambridge University Press.
Schlegel, R. J., Chu, S. L., Chen, K., Deuermeyer, E., Christy, A. G., & Quek, F. (2019). Making in the classroom: Longitudinal evidence of increases in self-efficacy and STEM possible selves over time. *Computers & Education, 142*.
Schneider, B., Bumbacher, E., & Blikstein, P. (2015). Discovery versus direct instruction: Learning outcomes of two pedagogical models using tangible interfaces. In *Paper presented at the exploring the material conditions of learning: Opportunities and challenges for CSCL," the proceedings of the computer supported collaborative learning (CSCL) conference*.
Schneider, R. M., Krajcik, J., Marx, R. W., & Soloway, E. (2002). Performance of students in project-based science classrooms on a national measure of science achievement. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching, 39*(5), 410–422.
Schrock, A. R. (2014). "Education in Disguise": Culture of a hacker and maker space. *InterActions: UCLA Journal of Education and Information Studies, 10*(1).
Schwartz, L., DiGiacomo, D., & Gutierrez, K. (2013). Diving into practice with children and undergraduates: A cultural historical approach to instantiating making and tinkering activity in a designed learning ecology. In *Paper presented at the international conference of the learning sciences (ICLS), boulder CO*.
Schwartz, P., Webb, G., & Mennin, S. (2001). *Problem-based learning: Case studies, experience and practice*. Psychology Press.
Searle, K., Fields, D., Lui, D., & Kafai, Y. (2014). Diversifying high school students' views about computing with electronic textiles. In *Paper presented at the Proceedings of the tenth annual conference on International computing education research*.
Searle, K., & Kafai, Y. (2015). Boys' needlework: Understanding gendered and indigenous perspectives on computing and crafting with electronic textiles. In *Paper presented at the international computing education research (ICER), omaha, Nebraska*.
Searle, K. A., Tofel-Grehl, C., & Breitenstein, J. (2019). Equitable engagement in STEM: Using E-textiles to challenge the positioning of non-dominant girls in school science. *International Journal of Multicultural Education, 21*(1), 42–61.
Searle, Kristin A., Litts, Breanne K. Litts, & Kafai, Yasmin B. (2018). Debugging open-ended designs_ High school students' perceptions of failure and success in an electronic textiles design activity. *Thinking Skills and Creativity, 30*, 125–134.
Sheridan, K., Halverson, E. R., Litts, B., Brahms, L., Jacobs-Priebe, L., & Owens, T. (2014). Learning in the making: A comparative case study of three makerspaces. *Harvard Educational Review, 84*(4), 505–531.
Silver, J., Rosenbaum, E., & Shaw, D. (2012). Makey Makey improvising tangible and nature-based user interfaces. In *Paper presented at the ACM tangible embedded and embodied interaction, Kingston, Ontario, Canada*.
Smith, Rachel Charlotte, Iversen, Ole Sejer, & Hjorth, Mikkel (2015). Design thinking for digital fabrication in education. *International Journal of Child-Computer Interaction, 5*, 20–28.
Somanath, S., Morrison, L., Hughes, J., Sharlin, E., & Sousa, M. C. (2016). Engaging'At-Risk'Students through maker culture activities. In *Paper presented at the proceedings of the TEI'16: Tenth international conference on tangible, embedded, and embodied interaction*.
Tai, R. H., Liu, C. Q., Maltese, A. V., & Fan, X. (2006). Planning early for careers in science. *Science, 312*(5777), 1143–1144.
Tofel-Grehl, C., Fields, D., Searle, K., Maahs-Fladung, C., Feldon, D., Gu, G., et al. (2017). Electrifying engagement in middle school science class: Improving student interest through e-textiles. *Journal of Science Education and Technology*, 1–12.
Torralba, J. (2019). A mixed-methods approach to investigating proportional reasoning understanding in maker-based integrative STEAM projects. In *Paper presented at the proceedings of FabLearn 2019*.

Vasudevan, V., Kafai, Y., & Yang, L. (2015). Make, wear, play: Remix designs of wearable controllers for scratch games by middle school youth. In *Paper presented at the Proceedings of the 14th international conference on interaction design and children*.
Wargo, Jon M., & Alvarado, Jasmine (2020). Making as worlding: young children composing change through speculative design. *Literacy, 54*, 13–21.
Yin, Y., Hadad, R., Tang, X., & Lin, Q. (2020). Improving and assessing computational thinking in maker activities: The integration with physics and engineering learning. *Journal of Science Education and Technology, 29*(2), 189–214.
Yin, Y., Tomita, M. K., & Shavelson, R. J. (2014). Using formal embedded formative assessments aligned with a short-term learning progression to promote conceptual change and achievement in science. *International Journal of Science Education, 36*(4), 531–552.