www.ijemst.net

Using Video and Written Reflection to Assess Second-Grade Students' Design Thinking and Conceptual Understanding in an Engineering and Design Challenge

Joe P. Gaston
University of South Alabama, U.S.A.

Sarah K. Guffey-McCorrison
University of South Alabama, U.S.A.

Angela D. Rand
University of South Alabama, U.S.A.

To cite this article:

Gaston, J. P., Guffey-McCorrison, S. K., & Rand, A. D. (2023). Using video and written reflection to assess second-grade students' design thinking and conceptual understanding in an engineering and design challenge. *International Journal of Education in Mathematics, Science, and Technology (IJEMST), 11*(4), 820-843. https://doi.org/10.46328/ijemst.2746

International Journal of Education in Mathematics, Science and Technology

2023, Vol. 11, No. 4, 820-843 https://doi.org/10.46328/ijemst.2746

Using Video and Written Reflection to Assess Second-Grade Students' Design Thinking and Conceptual Understanding in an Engineering and Design Challenge

Joe P. Gaston, Sarah K. Guffey-McCorrison, Angela D. Rand

Article Info

Article History

Received:
23 July 2022
Accepted:
27 March 2023

Keywords

STEM
Design thinking
Engineering design
Flipgrid
Reflective thinking

Abstract

The purpose of this mixed methods triangulation convergence study (Creswell & Clark, 2017) was to identify the differences between expressions of understanding through writing versus video reflection in a second-grade science classroom. Specifically, we analyzed the reflective piece of the design thinking of the engineering and design process by focusing on the empathy and ideate themes of the d.school (2013) engineering and design thinking rubric. A total of 76 second grade science students participated in the study either through Flipgrid, video reflection, or written reflection. Results show that there were far fewer instances of empathy as opposed to ideate in both written and video responses. Additionally, there were far fewer instances of both present in the written responses compared to the video responses. Given that providing opportunities for students to engage in the inquiry-based learning of science with empathy may increase their capacity to learn science (Garner et al., 2017), learning how to effectively teach empathy could be beneficial in the teaching and learning of science.

Introduction

In response to the declining competitiveness of the United States in STEM fields (Science, Technology, Engineering, and Mathematics), the National Research Council (2012) proposed a new approach to K-12 science education which integrated engineering and design practices into science instruction. As a result, the ways in which educators teach science and the ways in which students learn science has drastically changed. The evolution of curriculum, standards reform documents, and standards-aligned assessments, have caused teachers to consistently evaluate the effectiveness of their teaching strategies and how their students learn. With the recent development of the three-dimensional *Next Generation Science Standards, NGSS* (NGSS Lead States, 2013), there's been an emphasis on teachers implementing engineering practices during instruction since *science and engineering practices* is one of the three dimensions. A major component within engineering practices is design thinking. As a result, there is an increasing call for teaching design thinking practices and providing guidance on developing a design thinking mindset in early elementary programs (Wu, Hu, & Wang, 2019). Changes in STEM education include a synthesized view of scientific inquiry and engineering as it is taught in K-12 classrooms

(Bybee, 2011). Part of this synthesis involves a change from focusing on methods to an adoption of process activities. According to Bybee (2011), science and engineering practices exist collaterally in STEM education. That is, science seeks evidence-based explanations about how the world works, and engineering proposes artifacts to solve human problems. Bybee (2011) used the example of old ways students used to memorize the steps to one in which they are actively involved in designing solutions to problems.

Design thinking offers a framework for structuring engineering design processes that includes taking an empathetic perspective. It provides scaffolding for engaging with a problem, an important consideration for teaching engineering design (Hatzigianni et al., 2021). Design thinking's human-centered, innovative, iterative approach proposes to incorporate the practice of design thinking as a systematic process (Goldman et al. 2012). It accomplishes this by attending to instructional strategies that promote design thinking as a standard way of thinking about problems. That is, it is a disposition to be acquired. Public education has the responsibility of preparing students to work innovatively and creatively, and to solve problems in a complex global environment. Design thinking skills include learning to think creatively while incorporating flexible thinking with the goal of satisfying human needs (Carroll et al., 2010). There is a need to teach and assess design thinking in K-12 schools and to measure the progress of design thinking skills acquisition (Bekker et al., 2015). As such, one goal of this study was to assess students' design thinking skills after engaging in an engineering and design challenge. To that end, students reflected on their understanding of the concepts and the engineering and design challenge through video and written responses. The study included an additional multiple-choice assessment of students' conceptual understanding of matter after engaging in the challenge.

Literature Review

In educational research, the mélange of design thinking theories and models characterizes a human-centered approach and related competencies that focus on empathy, collaboration, and creativity (Grammenos & Antona, 2018; Luka, 2020). Design thinking projects align with a socio-cognitive lens when students work in groups to discuss, propose next actions, and implement human-centered designs (Carroll et al., 2010). Design thinking refers to a problem-solving approach that is human-centered, generative, iterative, and evaluative (Brown, 2008; Leverenz, 2014) and "purposefully embeds creative and critical thinking skills" into a model for approaching problems and opportunities (Shively et al., 2018 p.152). Design thinking is important in engineering and business settings and across multiple disciplines because of its value in problem solving through creative thinking (McLaughlin et al., 2019; Razzouk & Shute, 2012; Shively et al., 2000).

Stanford's d.school identified five non-sequential, iterative modes of design thinking as empathy, define, ideate, prototype, and test (Hatzigianni et al., 2021). These stages are driven by the designer's continuous reflection and assessment of the process (Plattner, 2010). Design thinking is a developing pedagogical model used in K-12 and higher education to instruct in multiple disciplines (Luka, 2019; Panke, 2020; Scheer et al., 2012) because it offers a rich pedagogical framework promoting meaningful learning (Cook & Bush, 2018). As demonstrated in its inclusion in K-12 science core skills outcomes, design thinking is an important 21st century skill for our current education system (Cook & Bush, 2018). A notable characteristic of the evolving definition of design thinking is a focus away from actions and processes employed during the design of a product, toward consideration of the users and their use of the end product (Bybee, 2011).

Empathy in Design Thinking

The role of empathy in design thinking refers to a practice that constitutes a way of solving design problems through empathetic investigation of users' wants, needs, and satisfaction levels at the beginning and throughout the problem-solving process rather than as an added step of an iterated solution (Brown, 2008; Leverenz, 2014; Tellez & Gonzalez-Tobon, 2019). The process is expansive and considers environmental as well as current and future impacts. Thus, the design thinker is not merely the designer, but the problem solver too.

Ideation in Design Thinking

The role of ideation in design thinking is to advance generative thinking (Stempfle & Badke-Schaub, 2002). In this phase of design thinking an open-ended approach to possible solutions sets the stage for novel idea exploration in the absence of evaluation (Henriksen et al., 2017). Ideation differs from prototype phases because in the ideate phase, critical responses to brainstorming efforts are withheld (Simon & Cox, 2019). Ideation in design thinking is a space where students can bring conceptual understandings into the process of finding a solution within the constraints of disciplinary knowledge (Carroll et al., 2010). The value of situating STEM education strategies in design thinking frameworks are evident in Schon's (1992) knowing in action concept (Simon & Cox, 2019). During ideation students can refer to their models and artifacts as memory devices that prompt reflection related to knowing in action (Simon & Cox, 2019). Simon and Cox (2019) analyzed knowing in action in a classroom setting using mathematical modeling.

In the current study, we examined and coded student video and written responses to reflection prompts to capture knowing in action from the design thinking framework focusing on empathy and ideation. A second round of coding allowed us to discover responses for expressions of scientific conceptual understanding. The design thinking ideation phase, in which models are designed, proposed, discussed, and evaluated, can reveal students' understanding of mathematics concepts (Simon & Cox, 2019), and by extension, scientific concepts (Johns & Mentzer, 2016). Learning concepts through design thinking promotes exploration of science concepts over completion of a finished artifact (Simon & Cox, 2019). In this sense, ideation is an iterative process that does not seek a final solution, but rather exhibits evidence of conceptual understanding when learners communicate descriptions, explanations, and justifications for actions taken during ideation. Thus, ideation is a solution seeking space (Simon & Cox, 2019).

Conceptual Understanding & Misconceptions

The driving force behind the *NGSS* and the *Framework* is the nature of how students learn and understand science. The theories that drive students' understanding of science are intertwined and include constructivism (Wittrock, 1974; Osborne & Wittrock, 1983), conceptual understanding (Pines & West, 1986), misconceptions or alternative conceptions (Driver, 1981), and conceptual change (Ausubel, 1968; Posner et al., 1982). In science, constructivism views learners as “active agents struggling to make sense of their world” (Pines & West, 1986). All learners are believed to have private understandings of accepted concepts that are public knowledge (Sutton, 1981).

When students try to understand widely accepted science concepts through formal education (e.g., teaching), teachers may see a change in knowledge, however, this change may be rote memorization or a lack of meaningful understanding as opposed to conceptual understanding (Pines & West, 1986). Conceptual understanding has been defined in multiple contexts and across different disciplines but in science, it’s when students are able to think about and use a concept in areas that are different from which they learned. Additionally, students can describe the concept in their own words, identify analogies and metaphors, and build models of the concept (mental or physical) (Konicek & Keeley, 2015; Pines & West, 1986). In a way, students create their own language through understanding the new concept.

There are instances where students create their own understandings of concepts that are different from the accepted concepts that are public knowledge, which are called alternative frameworks or misconceptions (Driver & Erickson, 1983). Pine and West (1986) used Vygotsky’s (1962) vine metaphor to describe how students develop alternative frameworks and misconceptions. Students have two types of knowledge, spontaneous and formal, that originate from different sources and have their own “vine.” Spontaneous knowledge originates from the learner’s experiences in their personal environment (e.g., home; culture; family; friends, etc.) and is described as growing in an upward direction to represent the “organic growth of the learner” (Pine & West, 1986, p. 587). Formal knowledge originates from the learner’s academic career, or the classroom and is described as growing in a downward direction, “suggesting its imposition on the learner from the authorities above” (Pine & West, 1986, p. 587). The vines will meet at some point and will result in one of four situations, with the “conflict situation” being the most concerning for teachers. The conflict situation occurs when the learner’s spontaneous knowledge (vine) is deeply embedded (ingrained) in their scientific belief system which makes it difficult to accept the formal knowledge or accepted concepts being presented by the teacher (Pine & West, 1986). This is called a misconception or alternative conception and results in the teacher needing to provide instruction for the learner to experience conceptual change to unbind the misconception (Posner et al., 1982).

For a learner to experience conceptual change, they must be confronted with the discrepancies of their current understanding (the misconception) by experiencing an anomaly that results in dissatisfaction with their misconception (Posner et al., 1982). In the teaching of science, this is best achieved through discrepant events (Gonzalez-Espada et al., 2010; Konicek-Moran & Keeley, 2015), which involves the teacher setting up a demonstration that will cause students to expect a certain result (a misconception). The teacher asks students to predict what will happen and after observing the demonstration, students modify their predictions. In short, the discrepant event evokes dissatisfaction with the learner’s misconception or alternative framework, which is what Posner and colleagues (1982) suggested the learner experiences to undergo conceptual change. Considering the nature of each of these theories, it’s evident that their components were foundational in the development of the *NGSS* and the *Framework*.

Structure of the *NGSS* and Key Terms

In 2012, science content and education experts from the National Research Council (NRC), the National Science Teachers Association (NSTA), the American Association for the Advancement of Science (AAAS), and Achieve, Inc. developed a *Framework* to identify the science concepts that students should know upon the completion of their K-12 education (National Research Council, 2012). Once the *Framework* was created, Achieve, Inc. managed a state-led development of the K-12 *Next Generation Science Standards, NGSS* (NGSS Lead States, 2013). The *NGSS* includes three dimensions: 1) practices; 2) crosscutting concepts; and 3) disciplinary core ideas. The practices describe behaviors that scientists use "as they investigate and build models and theories about the natural world" (NRC, 2012, p. 30). The practices also describe the behaviors that engineers use as they design and build models and systems (NRC, 2012). In both contexts of science and engineering, the term "practices" is identified as the learner "engaging in scientific inquiry that requires coordination of both knowledge and skill simultaneously" (NRC, 2012, p. 41). The crosscutting concepts link and apply to the different domains of science (NRC, 2012). For example, the concept of density can be applied to physical science, life science, and earth and space science. The disciplinary core ideas focus K-12 science curriculum, instruction, and assessment on the most important concepts of science (Achieve Inc., 2013). Unlike previous standards documents, the goal of the *NGSS* isn't to teach all concepts within each discipline - there's simply too much information. Rather, the main purpose is to "prepare students with sufficient core knowledge so that they can later acquire additional information on their own" (NRC, 2012, p. 31).

In addition to the three dimensions of the *NGSS*, there are terms used throughout the *Framework* and the *NGSS* that are specific to this paper and should be described. The *Framework* defined science as the "body of knowledge that reflects the current understanding of the world. Knowledge is based on evidence from many investigations and is integrated into highly developed and well-tested theories that can explain bodies of data and predict outcomes of further investigations" (NRC, 2012, p. 26). Additionally, aligning with the constructivist theory, science is viewed as being "fundamentally a social enterprise, and scientific knowledge advances through collaboration and in the context of a social system with well-developed norms" (NRC, 2012, p. 27). Engineering is broadly defined in the *Framework* as "any engagement in a systematic practice of design to achieve solutions for particular human problems" (Appendix I, p. 1). The *Framework* describes a type of mutualistic relationship between science and engineering, where students experience the benefits of both at the same time through practices. While engaging in engineering, students utilize practical scientific methods and are provided a context to develop and evaluate their scientific understanding and apply it to real-world problems, and as a result, their understanding of and interest in science increases (NRC, 2012). While science and engineering are similar in nature, the biggest difference is their driving force. Engineering is a way of thinking to solve problems for a purpose (Moore et al., 2015). Alternatively, science is not always driven by immediate application but rather by a general curiosity in a specific concept (NRC, 2012). For example, the scientific theory of evolution helps learners understand that all species are related and change over time and while this is important to understand, it wasn't developed to solve a specific human problem, which is a key factor of engineering. The last term that is likely most significant for this paper is engineering design, which the *Framework* describes as an "iterative cycle of design that offers the greatest potential for applying science knowledge in the classroom and engaging in engineering practices" (NGSS Release Appendix I, 2013, p. 2).

***NGSS*, Engineering Design, and Design Thinking Connections and the Need for Empathy**

One of the theoretical underpinnings of the *Framework* and *NGSS* is social constructivism, which means students construct knowledge based on their experiences with their environment and their interactions with their peers and teachers (Driver et al., 1994; Lorsbach & Tobin, 1992; Palinscar, 1998; Rogoff, 1998; Vygotsky, 1980). In a way, the social aspect of learning science is just as important as the student's experiences with their environment. As students navigate through the engineering design process, they utilize aspects of design thinking and the disciplinary core ideas of the *NGSS*. Because the sole purpose of engineering is to solve problems that help people (Capobianco et al., 2011), it can be implied that empathy is a part of that human element of engineering. Additionally, in the *NGSS*, empathy is not explicitly defined but considering the human component of engineering and the social constructivist theoretical underpinnings of the *NGSS*, empathy can be implied. The desire for scientists and engineers to solve problems that make human life easier is directly connected to empathy, which is an important concept that could help reframe engineering education and possibly help recruit more women in engineering (Letourneau & Bennett, 2017). While it may be known that there is minimal research on empathy in STEM education (Cook & Bush, 2018), it's quite surprising that engineering as a discipline is also not explicitly defined in the *NGSS*; it's only described in terms of engineering practices (Pleasants & Olson, 2019).

Reflection as Assessment

An important skill students need to develop to be successful with engineering practices is flexible thinking. Flexible thinking helps students be more adept at seeking design solutions. One way to achieve this flexible thinking is through reflection. Reflection involves thinking about past events and experiences in order to come up with better solutions (Rodgers, 2002). Reflection has been found to help increase the flexible thinking process whether it is done individually or as a team (Bekker et al. 2015). Being able to capture student reflection can have several benefits. Student reflections can be used as a form of assessment to measure student progress and understanding. They can also be used to help direct future teaching and learning activities.

Capturing student reflections can also help students improve their ability to reflect. Reflective journals, for example, provide a way for students to express their learning externally, making the process of reflection more concrete (Walsh & Mann, 2015). Journaling can help with knowledge acquisition and the development of problem-solving skills (Li & Peng, 2018). Reflective journals also provide teachers with evidence of reflective thinking that can be used to determine growth or identify deficiencies (Kember, Jones, Loke, & McKay, 1999). For younger students who are still learning to write, however, journaling may not be the best method for capturing their reflective thinking. Lee's (2005) method looks at both the content and the depth of reflective thought. The content refers to the primary concerns of the students (in this study, empathy, ideation, and design). The depth of student reflective thinking is based on three characteristics: recall (what did you do), rationalization (what happened), and reflectivity (what should you do next). Because of the age of the participants (7-8 years old) and their inexperience with reflection, it was decided by the researchers that these three questions would be explicitly given to help guide student writing and video comments.

The Skill of Writing

Writing is a demanding and complex activity that requires numerous goals and processes to take place simultaneously (McCutchen, 1988). This complexity creates a high cognitive load for beginning writers (Cave, 2010) as they attempt to hold their ideas in short-term memory while also trying to remember letter sounds and formations to put those thoughts into print (Williams, 2017). Due to this complexity, it takes considerable time for students to become adequately competent in the types of writing expected in the school environment. (Rijlaarsdam et al., 2012). Students in the latter years of elementary school are still considered to be developing writers (Wijekumar et al., 2019). For younger students, attempting to demonstrate their understanding through the written word can often prove ineffective. Research has shown, however, that students need structured ways for documenting evidence that allow for reflection and analysis (Butcher et al., 2019). Without a structured way to document evidence, students tend to focus on rational that is often vague and generalized (Butcher et al, 2019). If younger students are not yet competent in their writing, other forms of documenting evidence should be considered. One alternative for capturing the reflective thinking of younger students is video.

Using Video to Capture Student Reflections

Shively, Krista, & Rubenstein (2018) suggest that video recordings can be used to capture student reflection to assess creative and critical thinking, but no specific study was conducted with this focus. In fact, our review of the literature found little research that used video as a reflection-capturing tool in the way it was used in the current study. Hartzigianni et al., (2021) used content analysis of transcripts of computer screen recordings to discover evidence of design thinking after second-grade students engaged in a design project. Other studies were found, however, that used video diaries to capture student thinking (Buckingham, 2009; Holliday, 2004; Lundstöm, 2013; Roberts, 2011).

In each of these studies, the students were at least teenagers and they used video cameras to record their thinking outside of school. Although the age of the students and the circumstances of their use of video were different from the current study, their results provide some insight into the benefits of the video platform. Lundström (2013) suggested that the teenagers in his study found using video to record their thinking to be easier than keeping a written diary. As he stated, “The skills of formulating themselves in writing may be an obstacle for some individuals, the video diary overcomes that obstacle” (p 7). Holliday (2004) and Roberts (2011) found that video provided a better platform for students to reflect over other record keeping methods. Buckingham (2009) suggests that when students can use video to record their thoughts, they are better able to express experiences that might be more difficult to do with the written word.

The purpose of this study was to analyze the written and video reflections of second grade students who participated in an engineering design challenge for evidence of empathy, ideation, and conceptual understanding. Specifically, we analyzed the reflective piece of the design thinking of the engineering and design process of the second-grade science standard "*Demonstrate and explain how structures made from small pieces (e.g., linking cubes, blocks, building bricks, creative construction toys) can be disassembled and then rearranged to make new and different structures*." (Alabama State Department of Education, 2015). Our research questions for this study were:

1) What are the outcomes of using video versus writing as a platform for capturing student reflection in the design thinking process?
2) What are the outcomes of using video versus writing on assessing students' understanding of science?

Research shows that some K-12 students have difficulty expressing their understanding through writing (Graham, Harris, MacArthur, & Schwartz, 1991). Additionally, research shows that students struggle with properties of material (Doran, 1972; Gomez-Zwiep, 2008).

Method

This study employed the mixed methods triangulation convergence model which involves collecting and analyzing complementary quantitative and qualitative data to better understand the research problem (Creswell & Clark, 2017). Quantitative data was used to help interpret the qualitative results. Figure 1 visually displays the research design.

Data Collection	Data Analysis	Results		
QUAN Procedures	**QUAN Procedures**	**QUAN Procedures**		
IXL Learning Pre- & Post Test n = 72	Reliability Descriptive statistics Paired *t*-Test ANOVA Post-Hoc	Identify misconceptions Compare student learning gains from three groups		
QUAN Products	**QUAN Products**	**QUAN Products**		
Numerical item scores	Cronbach alpha Mean, SD *p*-value Games-Howell	Content knowledge learning gains	Compare and contrast	Interpretation QUAN + QUAL
QUAL Procedures	**QUAL Procedures**	**QUAL Procedures**	QUAN + QUAL results	Reflection connects the two data points
Engineering lab n = 76	Interrater reliability Open coding	Compare and contrast themes from Flipgrid, video, and written responses Word count in reflections		
QUAL Products	**QUAL Products**	**QUAL Products**		
Student reflections Flipgrid (n = 13) Video (n = 22) Written (n = 76)	Identified themes using Engineering design rubric	Most effective way to reflect (Flipgrid, video, or written)		

Figure 1. Triangulation Design Convergence Model

Participants

The study took place at a suburban public elementary school in the southeastern United States during the fall of 2019. The school is part of one of the largest school systems in the region, with just under 60,000 students. The school consists of pre-kindergarten through second grade and had approximately 600 students enrolled in the 2019-2020 school year. The participants were second grade students from six different classes. A total of 76 students, across six classes, participated in the study.

Procedure

Prior to the design challenge, the participants took a content specific pre-test that was administered by each participant's classroom teacher. The design challenge took place in the school's science lab and was conducted by the cooperating teacher. At the beginning of the challenge, the students watched a presentation on different features and uses of towers. The students saw pictures of different kinds of towers such as fire towers, lighthouses, and the Eiffel Tower. Additionally, they talked about some of the features of the towers that made them stable, such as having a wide base and being made of various types of materials. At the end of the presentation, the cooperating teacher introduced the students to the design challenge, which consisted of working with a partner to build a tower that was at least 12 inches high and could support the weight of a toy cardinal using only the materials that were provided. Each team had to complete their tower within 15 minutes to have it tested with the toy cardinal. The supplied materials consisted of items such as straws, tape, paper cups of varying sizes, strips of cardboard and paper, and popsicle sticks. When the allotted time was complete, the cooperating teacher placed the toy cardinal on the towers that were at least 12 inches tall to determine if it could be supported. At the end of the challenge, all students were asked to reflect on what they did by answering three questions that were based on Lee's (2005) reflective thinking assessment method: 1) What did you do? 2) What happened? 3) What should you do next?

Two of the classes were selected to initially respond to the three questions using computer tablets and the educational platform, Flipgrid. Flipgrid is a social learning platform that allows students to record and upload videos into topics that have been created by their teacher. These videos are then visible by the teacher and the students in a "grid" layout. This platform was chosen for this study because it is free to use, the Flipgrid app can be downloaded onto a computer tablet, and the students were already familiar with how to use the application. We were also interested to see how the students would do with their video reflections if they were the ones holding the computer tablets and in charge of recording themselves. After answering the questions using the tablets and Flipgrid, the teams returned to their tables and wrote their responses to the same three questions.

Two other classes were selected to initially respond to the three questions verbally while one of the researchers recorded them with a video camera. This process allowed the students to still reflect verbally but without the additional task of recording themselves. Each team was recorded answering the questions and then returned to their tables to write their responses. The final two classes were selected to write their responses to the three questions first and then read those responses out loud while one of the researchers recorded them with a video camera. By video recording the students reading their responses, we were able to capture any additional comments they may have made beyond what was written. Table 1 provides an outline of the study.

Table 1. Overview of Classes and Interventions Received

	Group		
Intervention	**1**	**2**	**3**
Pre-test Science Knowledge Test	X	X	X
Engineering & Design Lab	X	X	X
Flipgrid Video Reflection	X		
Video Reflection		X	
Writing Reflection After Group Discussion	X	X	
Writing Reflection Before Group Discussion			X
Post-test Science Knowledge Test	X	X	X

Qualitative Data Collection and Analysis

Participants from two classes were selected to respond to three open-ended questions at the conclusion of the lab activity by first using tablets and the video platform, Flipgrid, to record their verbal responses. The questions were as follows: (1) What did you do (when building your tower); (2) What happened (when your tower was tested); and (3) What should you do next time? After completing their Flipgrid responses, the partners were instructed to return to their tables and write their responses on a sheet of paper that was supplied by the teacher. Participants from two different classes were selected to initially respond to the three questions verbally while one of the researchers used a video camera to record them. After answering the questions on camera, the participants wrote their responses on the paper supplied by the teacher. Participants from the final two classes were selected to first write their responses to the three questions and then read their responses out loud while one of the researchers recorded them with a video camera.

The researchers first coded the video and written transcripts using two of the five identified themes from the Stanford d. school rubric (d.school, 2013); namely, empathy and ideate. Empathy refers to expressions of human-centeredness regarding emotions or physical needs (d.school, 2013). Ideate refers to the generation of multiple ideas reflective of convergent and/or divergent thinking (d.school, 2013). For a comment to be coded as empathy, the participant needed to express concern for the cardinal's well-being rather than simply referring to it as an object. For example, one student stated, "we made a stable place where it (the cardinal) could sit", and another said, "When we put the cardinal on it (the tower) it stayed there safely, and nothing happened." Comments that were coded as instances of ideation had to either contain options for creating a plan or contain evidence of decision making. It was decided by the researchers that comments mentioning a change such as "make it sturdier" or "make it bigger" were not enough to qualify as evidence of ideation. Although these comments suggest decision making, they do not provide any specific evidence. For example, "and on the top we had to put cardboard on paper so it could hold the cardinal," shows evidence of a decision that was made for a specific purpose. When asked what they could do differently next time, one participant stated, "make a wider base so it won't fall down next time and put more cardboard; we only put three pieces." This statement contains both an option for creating a plan and evidence of decision making. All three of the researchers coded the responses independently, assigning one or more themes (empathy and/or ideate) to each unit of analysis. The researchers met to compare how themes were assigned and to resolve differences to achieve interrater reliability.

The video and written transcripts were also coded for evidence of conceptual understanding. To determine what evidence would constitute conceptual understanding, the researchers consulted with the cooperating teacher, the Alabama Learning Exchange website, and the National Assessment of Educational Progress (NAEP). The cooperating teacher stated that comments addressing how the students moved pieces around to see what worked, statements that addressed what did not work, properties of materials that built successful towers, and properties of materials that were not successful, were all indicators of conceptual understanding. The Alabama Learning Exchange website (ALEX), a website designed to provide lessons, activities, and in-depth explanations of the K-12 state standards, states that explanations that correlate with a demonstration in which characteristics of the new object or objects are described are indicators of conceptual understanding. The ALEX website also provided a statement from NAEP regarding the state standard: E4.6: *Some Earth materials have properties either in their present form or after design and modification that make them useful in solving human problems and enhancing the quality of life, as in the case of materials used for building or fuels used for heating and transportation*. An example coded for conceptual understanding states, "Next time we would, like, put some more stuff and make it stronger because if it's really light it'll just fall down." This example shows an understanding of properties and characteristics of a successful tower. The researchers coded the video and written responses independently and then met to resolve differences to achieve interrater reliability.

Quantitative Data Collection & Analysis

The assessment was created by the researchers and the classroom teacher in this study. Questions from *IXL Learning* (n.d.) were selected to create the assessment, which was administered by the classroom teachers to participants as a pre- and post-test in a classroom setting with paper and pencil. The pre-test was administered before the 5E engineering lesson was delivered and the post-test was administered at the conclusion of the 5E engineering lesson.

The identified questions assessed the concept of materials and were aligned with the *State Course of Study* Standard 3 for second grade (State Department of Education, 2015) and the *Next Generation Science Standards* Standard 2-PS1-3 (NGSS Lead States, 2013). The *State Course of Study* Standard 3 states, "Demonstrate and explain how structures made from small pieces (e.g., linking cubes, blocks, building bricks, creative construction toys) can be disassembled and then rearranged to make new and different structures" (State Department of Education, 2015). The *Next Generation Science Standards* Standard 2-PS1-3 states, "Make observations to construct an evidence-based account of how an object made of a small set of pieces can be disassembled and made into a new object [Clarification Statement: Examples of pieces could include blocks, building bricks, or other assorted small objects]" (NGSS Lead States, 2013). The 15-question assessment was multiple choice with the distractors being common misconceptions. Each question was worth one point therefore, a score of 15 points was equivalent to 100 percent.

A Cronbach's analysis was conducted to measure the reliability of the 15-question assessment (Field, 2013; Huck, 2012). Cronbach's alpha was calculated on both the pre- and post-assessment. Additionally, Field (2013) and Huck (2012) recommended conducting a correlation analysis between the pre- and post-assessment. To test for the normality of pre- and post-test scores, the recommended Shapiro-Wilk test was conducted (Razali & Wah, 2011; Shapiro & Wilk, 1965) and box plots were analyzed to visually assess normal distribution and identify outliers. Levene's test was conducted to assess the equality of variances. Johnson and Christensen (2019) recommended calculating the descriptive statistics, mean and standard deviation, for each of the three student groups (Flipgrid, Video, Writing).

To determine if students were at the same level academically, a one-way between-subjects analysis of variance was conducted to compare the mean pre-test scores of the three groups (Field, 2013; Huck, 2012). Levene's test was used to measure the assumption of homogeneity of variances and when this assumption was not met, the recommended Welch's *F* ratio was calculated (Field, 2013; Huck, 2012). Because the assumption of homogeneity of variances was not met, a post-hoc analysis using Games Howell's test was conducted to determine the significant differences between the mean pre-test scores of three student groups (Field, 2013; Huck, 2012). Additionally, a one-way between-subjects analysis of variance was conducted to compare the mean post-test scores of the three groups (Field, 2013; Huck, 2012). Levene's test was used to measure the assumption of homogeneity of variances and when this assumption was not met, the recommended Welch's *F* ratio was calculated (Field, 2013; Huck, 2012). Because the assumption of homogeneity of variances was not met, a post-hoc analysis using Games Howell's test was conducted to determine the significant differences between the mean post-test scores of three student groups (Field, 2013; Huck, 2012). A paired-samples *t*-test was conducted to determine significant differences in pre- and post-test scores.

Results

Qualitative Empathy, Ideation, and Conceptual Understanding

The first question asked of the participants was "What did you do?". In their answers, the students provided information about how they built their towers. A total of one hundred twenty-six (126) responses were collected for this question through a combination of video and written responses. Of these responses, three were coded as showing evidence of empathy, nineteen were coded for ideation, and thirteen were coded for conceptual understanding. Each of the examples of empathy and ideation were found in the transcripts from the video responses and none were found in the written responses, while twelve of the conceptual understanding examples came from the video responses and one came from a written response.

The second question was "What happened?". In their responses, the participants explained what happened when the teacher set the cardinal on top of their tower, or they explained that they were not able to test their tower with the cardinal because they did not meet the minimum height requirement for the tower in the allotted time. A total of one hundred twenty-three (123) video and written responses were collected for this question. Four responses were coded for empathy, four were coded for ideation, and two were coded for conceptual understanding. Each example of empathy, ideation, and conceptual understanding was identified in the video transcripts and not in the written responses.

The final question asked the students what they should do next time regarding their towers. Of the one hundred twenty-seven (127) responses collected, ten were coded for empathy, thirty-three were coded for ideation, and twenty-five were coded for conceptual understanding. Seven examples of empathy came from the video responses and three came from the written, while twenty-six of the video responses were coded for ideation compared to seven from the written. Sixteen video responses and nine of the written responses were coded for conceptual understanding. Table 2 displays the number and percentage of responses coded for each theme.

Table 2. Responses Coded for Each Theme

Question	Total Answers	Empathy (%)	Ideate (%)	Conceptual Understanding (%)
1 Video	55	3 (5%)	19 (35%)	12 (23%)
1 Written	71	0 (0%)	0 (0%)	1 (2%)
2 Video	56	4 (7%)	4 (7%)	2 (4%)
2 Written	67	0 (0%)	0 (0%)	0 (0%)
3 Video	63	7 (11%)	26 (41%)	16 (25%)
3 Written	64	3 (5%)	7 (11%)	9 (14%)
Total	376	17 (5%)	56 (15%)	40 (11%)

Table 3 displays the number of words written and spoken by each group.

Table 3. Comparison of Words Spoken and Written in Student Responses

Group	Words Written (%)	Words Spoken (%)	Total Words
1 (n = 27)	263 (12%)	1891 (88%)	2154
2 (n = 26)	526 (26%)	1525 (74%)	2051
3 (n = 16)	263 (34%)	507 (66%)	770
Totals (n = 69)	1052 (21%)	3923 (79%)	4975

A total of twenty-seven students from group 1 initially responded to the questions using the video platform, Flipgrid. After responding through Flipgrid, the students returned to their tables and provided written responses to the same three questions: 1) What did we do? 2) What happened? 3) What should we do next? The number of words spoken in the videos and written by the students to answer the questions totaled 2,154. Of this total, 1,891 words (88%) were from the videos, compared to 263 words (12%) from the written responses. Twenty-six students from group 2 answered the three questions verbally while one researcher video recorded their responses. These students then returned to their tables to write their answers. A total of 2,051 words were either spoken or written, with 1,525 (74%) being spoken and 526 (26%) being written. Group 3 consisted of twenty-one participants. The students wrote their answers to the questions first and then read their answers aloud while one researcher video recorded them. Of the 770 words used to answer the questions, 507 (66%) were spoken and 263 (34%) were written. Table 4 displays a verbal response and written response from the same student.

Table 4. Comparison of Verbal Response and Written Response from the Sample Student

Written response	Verbal response from same student
Our tower fell.	Well, we tried to put a cardinal on there and the cardinal couldn't stay balanced because we didn't have enough cardboard to put on there so it could have a little seat. So then, it found its balance once the teacher was holding it and then she tried to let it go and the tower fell. Well, half of it did.
Flat, tall, then flat.	And also, next time we're going to try to make it flat, some up (motioning with hands) and then make it flat again so the cardinal can get there and make it as tall as a ruler like this time.
Stack cups and put paper on top and tape.	She (his partner) would probably use the same tape because she uses a whole bunch of tape...And we're probably gonna just get some cardboard and stack it like that (motions with hands) like a diamond and put cardboard and keep on stacking with cups. It might work. And put more paper on the top. That might work...One right way, one upside down, and then we put a cardboard, and then one right way and one upside down, cardboard.

Quantitative

A Cronbach's analysis was conducted on the pre- and post-test. It was found that the pre- and post-test alpha levels were .534 and .341 respectively, which indicates that both tests did not have an adequate level of inter-item reliability. Further analyses found that deleting question four on the pre-test would increase the alpha level to .559 and deleting question one on the post-test would increase the alpha level to .501. It was found that the pre-test was positively correlated with the post-test, r (74) = .48, p < .01. Descriptive statistics were calculated for the mean pre- and post-test scores for each group (Table 5). Results from the Shapiro-Wilk test found that there were statistically significant differences between the pre- and post-test and the normal distribution, meaning the data was not normally distributed (p > 0.05). However, because the Analysis of Variance in SPSS is robust to violations in normality, we proceeded with ANOVA.

A one-way between subjects ANOVA was conducted to compare the mean scores of the three groups on the pre-test. An analysis of variance showed that there were no significant differences in mean pre-test scores of the three student groups, F (2, 69) = .273, p = .762. A one-way between subjects ANOVA was conducted to compare the effect of reflection type (video, written, video and written) on the post-test scores of the three student groups. An analysis of variance showed that the effect of reflection on the post-test was not significant, F (2, 69) = 1.037, p = .360. Additionally, a paired t-test was conducted to compare the pre- and post-test scores of each group and there was not a significant difference in pre- and post-test scores in the group 1 (t (23) = 1.277, p > 0.05); in group 2 (t (21) = -1.418, p > 0.05); or in group 3 (t (25) = 0.548, p > 0.05).

Table 5. Descriptive Statistics Pre- and Post-Test

	Group	Mean Score *Total Points (Percentage)*	SD	SE	95% Confidence Interval for Mean Lower Bound	Upper Bound
Pre-test score	1 Video (n = 24)	13.13 (87.5%)	1.361	.278	12.55	13.70
	2 Written (n = 22)	13.00 (86.7%)	1.51	.322	12.33	13.67
	3 Video & Written (n = 26)	13.35 (89%)	1.98	.388	12.55	14.15
	Total (n = 72)	13.17 (87.8%)	1.64	.193	12.78	13.55
Post-test score	1 Video (n = 24)	12.92 (86.1%)	1.06	.216	12.47	13.36
	2 Written (n = 22)	13.45 (89.7%)	1.10	.235	12.97	13.94
	3 Video & Written (n = 26)	13.15 (87.7%)	1.54	.302	12.53	13.78
	Total (n = 72)	13.17 (87.8%)	1.267	.149	12.87	13.46

Table 6. Analysis of Variance Pre- and Post-Test

		Sum of Squares	df	Mean Square	F	p
Pre-test Score	Between Groups	1.490	2	.745	.273	.762
	Within Groups	188.5	69	2.732		
	Total	190.0	71			
Post-test Score	Between Groups	3.328	2	1.664	1.037	.360
	Within Groups	110.7	69	1.604		
	Total	114.0	71			

Table 7. Paired t-Test Group 1 Pre- and Post-Test Scores

				95% Confidence Interval of Difference				
	Mean	SD	SE	Lower	Upper	t	df	p
Pretest - Posttest Score	.375	1.439	.294	-.233	.983	1.277	23	.214

Table 8. Paired t-Test Group 2 Pre- and Post-Test Scores

				95% Confidence Interval of Difference				
	Mean	SD	SE	Lower	Upper	t	df	p
Pretest - Posttest Score	-.455	1.50	.320	-1.12	.211	-1.42	21	.171

Table 9. Paired t-Test Group 3 Pre- and Post-Test Scores

				95% Confidence Interval of Difference				
	Mean	SD	SE	Lower	Upper	t	df	p
Pretest - Posttest Score	.192	1.789	.351	-.530	.915	.548	25	.589

Table 10. Paired t-Test All Groups Pre- and Post-Test Scores

				95% Confidence Interval of Difference				
	Mean	SD	SE	Lower	Upper	t	df	p
Pretest - Posttest Score	.000	1.52	.179	-.357	.357	.000	71	1.00

Discussion

Qualitative

This study analyzed the written and video reflections of second grade students who participated in an engineering design challenge for evidence of empathy, ideation, and conceptual understanding. The findings from this study are summarized in this section and organized according to the research questions.

Outcomes of using Video versus Writing as a Platform for Capturing Student Reflection in the Design Thinking Process

A total of 174 video reflections were recorded and transcribed for analysis. For this study, two components of the design thinking process, (e.g., empathy and ideation), were defined and coded. Out of the 174 video reflections, 14 comments were coded for empathy and 49 comments were coded for ideation. A total of 202 written responses were also collected from the participants. Out of the 202 responses, three comments were coded for empathy and seven were coded for ideation. Bekker et al. (2015) believes teachers need to measure the progress of design thinking skills acquisition. Our study found video reflection to be more effective in capturing instances of ideation and empathy compared to written reflection. This supports the findings from Holliday (2004) and Roberts (2011) that video provides a better platform for students to reflect when compared to other record keeping methods. Based on our study, video reflection may be an effective method for assessing the progress of students' design thinking skills.

Researchers are still in the early stages of finding ways to effectively incorporate design thinking into educational contexts (Scheer et al., 2012). The design challenge used in our study is one example of how design thinking can be infused into the elementary science curriculum. Although K-12 teachers implement engineering and design challenges, it is our position that empathy needs to be explicitly taught. Empathy is a component of the design thinking process and to date, there has been little research on empathy in STEM education (Cook & Bush, 2018). After completing a literature review of design competences for K-12 education, Rusmann and Ejsing-Duun (2021) suggest incorporating empathy into the design process by allowing students to work collaboratively and "become immersed in the social context of the problem" (p. 2074). As a result of the design process, Wells (2013) found that students develop awareness of their personal position while also being sensitive and responding to others' needs. There have also been studies that have found that the design thinking process helps students develop empathy towards peers, especially peers that are different from themselves (Ladachart et al., 2021; Tan & Wong, 2012). In addition to building self-awareness among students, it's suggested that empathy facilitates students' "ability to meta-cognitively assess one's own learning" (Rusmann & Ejsing-Dunn, 2021, p. 2076). There have been some studies where teachers are interested in teaching empathy as part of the design process (Retna, 2016). However, the results of our study found minimal examples of empathy, which supports the claim of Bekker et al. (2015) that design thinking needs to be taught in the K-12 environment.

Outcomes of using Video versus Writing on Assessing Students' Understanding of Science

A total of one hundred seventy-four (174) video comments were analyzed. Of those comments, thirty (30) were coded as indicators of conceptual understanding. A total of two hundred two (202) written responses were also analyzed, with ten (10) of them being identified as evidence of conceptual understanding.

Ideation in the design thinking process provides a space where learners can utilize existing content knowledge and conceptual understanding to help find solutions (Carroll, M. et al., 2010; Razzouk & Shute, 2012). The solutions are often a result of the student tapping into memories or drawing analogies from personal experiences (Charman, 2010). The ideation process can lead to sound prototypes but in design thinking this phase promotes exploration within the constraints of disciplinary knowledge (Carroll et al., 2010). Studies also suggest creating a classroom environment that stimulates ideation, allowing students to freely design solutions and solve problems free of their peers' judgement increases students' confidence and optimism in using creativity (Ladachart et al., 2021; Rusmann & Ejsing-Dunn, 2021; Tsai & Wang, 2021). Our study supports Carroll's (2010) and Razzouk and Shute's (2012) connection between ideation and conceptual understanding. Of the forty responses coded for conceptual understanding, twenty-eight (70%) of them were also coded separately for ideation.

The results of our study also indicate video reflection was a more effective platform for capturing students' conceptual understanding over written reflection. This again supports the findings of Holliday (2004) and Roberts (2011) that video provided a better platform for students to reflect when compared to other record keeping methods. We found three times as many examples of conceptual understanding in the video reflections as in the written reflections. This finding seems to support the work of Cave (2010) which states writing creates a high cognitive load for beginning writers, and Wijekumar et al. (2019) that claim students in the upper elementary school are still considered to be developing writers.

Quantitative

Quantitative results showed that there were no significant differences in mean pre- and post-test scores between any of the three groups, which suggests that the type of reflection in which students engaged did not affect students' conceptual understanding. Additionally, there were no significant differences in the overall mean pre- and post-test scores, which suggests that the engineering design activity did not impact students' conceptual understanding. Although we anticipated differences in both areas, there were a few possible reasons why there weren't any. First, the quantitative instrument we used did not have a reliable Cronbach's alpha level. Although the questions on the instrument were selected from the widely used source, *IXL Learning*, and were aligned to the identified standards, the instrument had not been used in previous studies. Developing an instrument was not the purpose of this study, however, using a reliable instrument or piloting the instrument we created before using it in this study may have yielded different results (Creswell & Creswell, 2017).

Second, the engineering and design thinking process is open-ended and ill-structured, meaning there is no clear or specific design to solve these types of problems (Jonassen, 2011; Lammi et al., 2018). Xing et al. (2021) found that it was difficult for instructors to formatively assess students' understanding of the design process. Because there is no clear answer, teachers may have difficulty in knowing exactly what they are assessing. This lack of understanding may also be related to teachers designing and implementing engineering challenges in hopes that their students learn science concepts. "Many engineering tasks lead students to focus more on the success of their construction than on learning the science content, which can hurt students' ability to learn and transfer scientific principles from them" (Malkiewich & Chase, 2019). In our study, emphasis was placed on constructing a tower to support the weight of the rescued bird. Students may have been solely focused on the construction of their tower which possibly resulted in them not making the connection to the concept of matter on the multiple-choice test.

Another issue that may have impacted our results was the use of a multiple-choice assessment, which is not always the best gauge for students' understanding of science (Dufresne et al., 2002; Martinez, 1999; Scully, 2017; Stanger-Hall, 2012). Traditionally, assessments have focused "on what students know (the science content), rather than what they can do with that knowledge (the science and engineering practices)" (Stephenson et al., 2020). Additionally, there are not many available assessments that fully align with the three-dimensional *NGSS* (Pellegrino, 2013; Wertheim et al., 2016). Pellegrino (2013) described that most of the assessments that are currently available are not designed for measuring three-dimensional understanding. He stated that:

> "Given the relative newness of the NRC Framework, it is no surprise that comprehensive sets of assessment examples that align completely with the NGSS performance expectations do not exist. Many of the tasks that have been used for classroom assessment, and those found in large-scale state, national, and international tests, focus primarily on science content or on aspects of scientific inquiry separate from content. With few exceptions, such assessments do not integrate core concepts and science practices in the ways intended by the NRC Framework or *NGSS* " (Pellegrino, 2013, p. 321-322).

As a result, teachers have attempted to create their own 3D assessments but struggle with this task (Pruitt, 2014). Pruitt (2014) highlighted several challenges that the science education community has faced since the development of the *NGSS.* Because not all states have adopted the new standards, the development of new materials that are aligned to the *NGSS*, including assessments, are scarce (Pruitt, 2014). There is the possibility that the assessment used in this study did not assess students' three-dimensional understanding of the content. The 5E lesson with which students engaged focused on the concept of matter by using the engineering and design process. However, the questions asked on the assessment focused only on the properties of matter.

Conclusion

The *Task Analysis Guide for Science* (TAGS) emphasizes the importance of students learning science by doing science (Tekkumru-Kisa, Stein, & Schunn, 2015). The new three-dimensional standards (*NGSS, State COS*), also state that more ways are needed to assess students' understanding of science concepts in the classroom. For these reasons, it's important for teachers to implement strategies and incorporate tasks and assessments that get their students to think more deeply about science (Tekkumru-Kisa, Stein, & Schunn, 2015). Multiple choice tests are not the only way for students to demonstrate their understanding. The results of this study found that the multiple-choice test coupled with written and video reflection provided greater insight into the student's conceptual understanding than through the multiple-choice test alone. Additionally, instead of written reflection, video reflection may be a more accurate way to capture elementary-aged students' thinking of science or conceptual understanding of science (Buckingham, 2009; Holliday, 2004; Lundstöm, 2013; Roberts, 2011). Future studies could be conducted to develop rubrics that help teachers streamline the analysis of student reflection videos.

A major goal of the teaching and learning of science is to help students become scientifically informed global members of a democratic society (McCurdy, Nickels, & Bush, 2020). The engineering and design rubric (d.school,

2013) is another tool that could be used to assess the process students go through and the artifacts that are generated when completing a task. This rubric could be used to not only evaluate students' conceptual understanding of science, but to also evaluate a teacher's own pedagogy. Additionally, learning how to implement, identify, and assess the empathy component of the rubric (d.school, 2013) could be helpful in teaching students about empathy, which is the human-centered factor of the engineering design process (McCurdy, Nickels, & Bush, 2020). Connecting STEM with empathy has been shown to impact students' interest in STEM (Gunkel & Tolbert, 2018; Wirkala & Kuhn, 2011). Providing opportunities for students to engage in the inquiry-based learning of science with empathy may increase their capacity to learn science (Garner et al., 2017).

References

Alabama State Department of Education. (2015). Alabama course of study: Science.

Bekker, T., Bakker, S., Douma, I., Van Der Poel, J.,Scheltenaar, K. (2015). Teaching children digital literacy through design-based learning with digital toolkits in schools. *International Journal of Child-Computer Interaction*, 5, 29-38.

Brown, T. (2008) Design Thinking. Harvard Business Review, 86, 84–92.

Buckingham, D. (2009). `Creative' visual methods in media research: possibilities, problems and proposals. *Media, Culture & Society*, *31*(4), 633–652.

Butcher, K. R., Larson, M., & Lane, M. (2019). Making critical thinking visible for student analysis and reflection: Using structured documentation to enhance effective reasoning and communication. *Science Scope*, 42(8), 44–53.

Carroll, M., Goldman, S., Britos, L., Koh, J., Royalty, A., & Hornstein, M. (2010). Destination, imagination, and the fires within: Design thinking in a middle school classroom. *International Journal of Art & Design Education*, *29*(1), 37-53.

Cave, A. (2010). Learning how to become a writer in elementary school: A review of the literature from cognitive, social cognitive, developmental, and sociocultural perspectives. *Journal on Educational Psychology*, *3*(4), 1-13.

Charman, H. (2010). Designerly learning: Workshops for schools at the design museum. Design and Technology Education, 15(3), 28-40.

Cook, K. L., & Bush, S. B. (2018). Design thinking in integrated STEAM learning: Surveying the landscape and exploring exemplars in elementary grades. *School Science and Mathematics, 118*(3–4), 93–103. https://doi.org/10.1111/ssm.12268

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Sage publications.

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

d.school (2013), bootcamp bootleg. [online] Stanford University. Available at: http://dschool.stanford.edu/wpcontent/uploads/2013/10/METHODCARDS-v3-slim.pdf (12-16-2016).

Doran, R. L. (1972). Misconceptions of selected science concepts held by elementary school students. *Journal of Research in Science Teaching*, *9*(2), 127-137.

Driver, R., Asoko, H., Leach, J., Scott, P., & Mortimer, E. (1994). Constructing scientific knowledge in the classroom. *Educational researcher*, *23*(7), 5-12.

Dufresne, R. J., Leonard, W. J., & Gerace, W. J. (2002). Making sense of students' answers to multiple-choice questions. *The Physics Teacher*, *40*(3), 174-180.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. sage.

Garner, P. W., Gabitova, N., Gupta, A., & Wood, T. (2018). Innovations in science education: infusing social emotional principles into early STEM learning. *Cultural Studies of Science Education*, *13*(4), 889-903.

Goldman, S., Carroll, M. P., Kabayadondo, Z., Cavagnaro, L. B., Royalty, A. W., Roth, B., ... & Kim, J. (2012). Assessing d. learning: Capturing the journey of becoming a design thinker. In *Design thinking research* (pp. 13-33). Springer, Berlin, Heidelberg.

Gomez-Zwiep, S. (2008). Elementary teachers' understanding of students' science misconceptions: Implications for practice and teacher education. *Journal of Science Teacher Education, 19*(5), 437-454.

González-Espada, W. J., Birriel, J., & Birriel, I. (2010). Discrepant events: A challenge to students' intuition. *The Physics Teacher*, *48*(8), 508-511.

Grammenos, D., & Antona, M. (2018). Future designers: Introducing creativity, design thinking & design to children. *International journal of child-computer interaction*, *16*, 16-24.

Gunkel, K. L. & Tolbert, S. (2018). The imperative toward a dimension of care in engineering education. *Journal of Research in Science Teaching, 55,* 938-961.

Huck, S. W. (2012). *Reading statistics and research: Sixth edition.* Pearson.

Karen R. Harris, Steve Graham, Linda H. Mason, & Bruce Saddler. (2002). Developing Self-Regulated Writers. *Theory Into Practice*, 41(2), 110–115.

Hatzigianni, M., Stevenson, M., Falloon, G., Bower, M., & Forbes, A. (2021). Young children's design thinking skills in makerspaces. *International Journal of Child-Computer Interaction, 27*(16), 1-11. DOI: 10.1016/j.ijcci.2020

Henriksen, D., Richardson, C., & Mehta, R. (2017). Design thinking: A creative approach to educational problems of practice. *Thinking skills and Creativity*, *26*, 140-153.

Holliday, R. (2004). Filming "The Closet": The role of video diaries in researching sexualities. *American Behavioral Scientist, 47* (12), 1597-1616.

IXL Learning. (n.d.). *Second grade science: Materials.* IXL Learning. https://www.ixl.com/science/grade-2

Johansson-Sköldberg, U., Woodilla, J., & Çetinkaya, M. (2013). Design thinking: past, present and possible futures. *Creativity and innovation management*, *22*(2), 121-146.

Johnson, R. B., & Christensen, L. (2019). *Educational research: Quantitative, qualitative, and mixed approaches*. Sage publications.

Jonassen, D. H. (2011). Design Problems for Secondary Students. *National Center for Engineering and Technology Education*.

Kember, D., Jones, A., Loke, A., & McKay, J. (1999). Determining the level of reflective thinking from students' written journals using a coding scheme based on the work of Mezirow. *International Journal of Lifelong Education, 18*(1), 18-30. HTTP://DX.DOI.ORG/10.1080/026013799293928

Konicek-Moran, R., & Keeley, P. (2015). *Teaching for conceptual understanding in science*. Arlington: NSTA Press, National Science Teachers Association.

Ladachart, L., Cholsin, J., Kwanpet, S., Teerapanpong, R., Dessi, A., Phuangsuwan, L., & Phothong, W. (2021).

Ninth-grade students' perceptions on the design-thinking mindset in the context of reverse engineering. International Journal of Technology and Design Education, 1-21.

Lammi, M., Denson, C., & Asunda, P. (2018). Search and review of the literature on engineering design challenges in secondary school settings. *Journal of Pre-College Engineering Education Research (J-PEER)*, *8*(2), 5.

Langer, J. A. (1986). Reading, Writing, and Understanding: An Analysis of the Construction of Meaning. *Written Communication, 3*(2), 219–267.

Lee, H. (2005). Understanding and assessing preservice teachers' reflective thinking. *Teaching and Teacher Education*, 21 699-715. https://doi.org/10.1016/j.tate.2005.05.007

Leverenz, C. S. (2014). Design thinking and the wicked problem of teaching writing. *Computers and Composition, 33*, 1-12.

Li, A., & Peng, T. (2018). Observing "myself" in the video: Fostering reflective practice in oral presentation training. *Advances in Language and Literary Studies* *9*(3), 138-144. http://dx.doi.org/10.7575/aiac.alls.v.9n.3p.138

Lorsbach, A., & Tobin, K. (1992). Constructivism as a referent for science teaching. *NARST Newsletter*, *30*, 5-7.

Lundström, M. (2013). Using video diaries in studies concerning scientific literacy. *Electronic Journal of Science Education*, *17*(3), 1-18.

Luka, I. (2019). Design thinking in pedagogy: Frameworks and uses. *European Journal of Education*, *54*(4), 499-512.

Malkiewich, L. J., & Chase, C. C. (2019). What's your goal? The importance of shaping the goals of engineering tasks to focus learners on the underlying science. *Instructional Science*, *47*(5), 551-588.

Martinez, M. E. (1999). Cognition and the question of test item format. *Educational Psychologist*, *34*(4), 207-218.

McCurdy, R. P., Nickels, M. L., & Bush, S. B. (2020). Problem-Based Design Thinking Tasks: Engaging Student Empathy in STEM. *The Electronic Journal for Research in Science & Mathematics Education*, *24*(2), 22-55.

McCutchen, D. (1988). "Functional automaticity" in children's writing: A problem of metacognitive control. *Written communication*, *5*(3), 306-324.

McElhaney, K. W., Basu, S., Wetzel, T., & Boyce, J. (2019). Three-dimensional assessment of NGSS upper elementary engineering design performance expectations. In *NARST Annual International Conference*.

NGSS Lead States. (2013). Next Generation Science Standards: For States, By States. Washington, DC: The National Academies Press.

Panke, S. (2020). Design thinking in education: Perspectives, opportunities and challenges. *Open Education Studies, 1*(1), 281–306. https://doi.org/10.1515/edu-2019-0022

Pellegrino, J. W. (2013). Proficiency in science: Assessment challenges and opportunities. *Science*, *340*(6130), 320-323.

Pruitt, S. L. (2014). The next generation science standards: The features and challenges. *Journal of Science Teacher Education*, *25*(2), 145-156.

Plattner, H. (Ed.) (2010). d.school bootcamp bootleg. Institute of Design at Stanford. Retrieved from http://dschool.stanford.edu/ wp-content/uploads/2011/03/BootcampBootleg

Pleasants, J., & Olson, J. K. (2019). What is engineering? Elaborating the nature of engineering for K-12 education. *Science Education*, *103*(1), 145-166.

Razali, N. M., & Wah, Y. B., (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics, 2*(1), 21-33.

Razzouk, R., & Shute, V. (2012). What is design thinking and why is it important? Review of Educational Research, 82(3), 330–348.

Retna, K. S. (2016). Thinking about “design thinking”: A study of teacher experiences. Asia Pacific Journal of Education, 36(sup1), 5-19.

Rijlaarsdam, G., Van den Bergh, H., Couzijn, M., Janssen, T., Braaksma, M., Tillema, M., et al. (2012). Writing. In K. R. Harris, S. Graham, & T. Urdan (Eds.), *APA educational psychology handbook* (Vol. 3, pp. 189–227). Washington, DC: American Psychological Association.

Roberts, J. (2011). Video diaries: a tool to investigate sustainability-related learning in threshold spaces. *Environmental education research*, *17*(5), 675-688.

Rogers, C. (2002). Defining reflection: Another look at John Dewey and reflective thinking. *Teachers College Record*, 104, 842-866. http://dx.doi.org/10.1111/14679620.00181

Rusmann, A., & Ejsing-Duun, S. (2021). When design thinking goes to school: A literature review of design competences for the K-12 level. International Journal of Technology and Design Education, 1-29.

Scheer, A., Noweski, C., & Meinel, C. (2012). Transforming constructivist learning into action: Design thinking in education. *Design and Technology Education: An International Journal, 17*(3).

Schon, D. A. (1992). Designing as reflective conversation with the materials of a design situation. *Research in Engineering Design*, *3*(3), 131-147.

Scully, D. (2017). Constructing multiple-choice items to measure higher-order thinking. *Practical Assessment, Research, and Evaluation*, *22*(1), 4.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (Complete samples. *Biometrika, 52*(¾), 591-611

Shively, K., Stith, K. M., & Rubenstein, L. D. (2018). Measuring What Matters: Assessing Creativity, Critical Thinking, and the Design Process. *Gifted Child Today, 41*(3), 149–158. https://doi.org/10.1177/1076217518768361

Simon, L. M., & Cox, D. C. (2019). The role of prototyping in mathematical design thinking. *The Journal of Mathematical Behavior*, *56*, 100724.

Stanger-Hall, K. F. (2012). Multiple-choice exams: an obstacle for higher-level thinking in introductory science classes. *CBE—Life Sciences Education*, *11*(3), 294-306.

Stempfle, J., & Badke-Schaub, P. (2002). Thinking in design teams-an analysis of team communication. *Design studies*, *23*(5), 473-496.

Tan, C., & Wong, Y. L. (2012). Promoting spiritual ideals through design thinking in public schools. International Journal of Children's Spirituality, 17(1), 25-37.

Tsai, M. J., & Wang, C. Y. (2021). Assessing young students’ design thinking disposition and its relationship with computer programming self-efficacy. Journal of Educational Computing Research, 59(3), 410-428.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Harvard university press.

Walsh, S., & Mann, S. (2015). Doing reflective practice: a data-led way forward. *ELT Journal*, *69*(4), 351-362. http://dx.doi.org/10.1093/elt/ccv018

Washington, S., Karlaftis, M. G., Mannering, F., & Anastasopoulos, P. (2020). *Statistical and econometric methods for transportation data analysis*. CRC press.

Wells, A. (2013). The importance of design thinking for technological literacy: A phenomenological perspective. International Journal of Technology and Design Education, 23(3), 623–636.

Wertheim, J., Osborne, J., Quinn, H., Pecheone, R., Schultz, S., Holthuis, N., & Martin, P. (2016). An analysis of existing science assessments and the implications for developing assessment tasks for the NGSS. *Palo Alto, CA: Stanford NGSS Assessment Project Team (SNAP)*.

Wijekumar, K., Graham, S., Harris, K. R., Lei, P. W., Barkel, A., Aitken, A., ... & Houston, J. (2019). The roles of writing knowledge, motivation, strategic behaviors, and skills in predicting elementary students' persuasive writing from source material. *Reading and Writing*, *32*(6), 1431-1457.

Williams, C. (2017). Learning to write with interactive writing instruction. *The Reading Teacher*, *71*(5), 523-532.

Wirkala, C., & Kuhn, D. (2011). Problem-based learning in K-12 education: Is it effective and how does it achieve its effects? *American Educational Research Journal*, *48*(5), 1157-1186. http://dx.doi.org/10.3102/0002831211419491

Wu, B., Hu, Y., & Wang, M. (2019). Scaffolding design thinking in online STEM preservice teacher training. *British Journal of Educational Technology*, *50*(5), 2271-2287.

Xing, W., Li, C., Chen, G., Huang, X., Chao, J., Massicotte, J., & Xie, C. (2021). Automatic Assessment of Students' Engineering Design Performance Using a Bayesian Network Model. Journal of Educational Computing Research, 59(2), 230–256. https://doi.org/10.1177/0735633120960422

Author Information

Joe P. Gaston

https://orcid.org/0000-0002-5959-6510

University of South Alabama

Department of Counseling & Instructional Design

U.S.A.

Sarah K. Guffey-McCorrison

https://orcid.org/0000-0002-9074-901X

University of South Alabama

Department of Leadership & Teacher Education

U.S.A.

Contact e-mail: *skguffey@southalabama.edu*

Angela D. Ran

https://orcid.org/0000-0002-6363-3839

University of South Alabama

University Library

U.S.A.