



Guardian Angels of Our Better Nature: Finding Evidence of the Benefits of Design Thinking

Dr. Luke David Conlin, Stanford University

Luke Conlin is a postdoctoral scholar in the Graduate School of Education at Stanford University. His work focuses on the learning of engineering and science in formal and informal environments.

Doris B Chin Ph.D., Stanford Graduate School of Education

Dr. Kristen Pilner Blair, Stanford University

Dr. Maria Cutumisu, Stanford University

Prof. Daniel L Schwartz, Stanford University

Dr. Schwartz studies human learning, especially as it applies to matters of instruction.

Guardian Angels of Our Better Nature: Finding Evidence of the Benefits of Design Thinking

INTRODUCTION

In the field of engineering education and beyond, there is a widespread and growing belief in the importance of teaching the disciplinary practices of engineering design, i.e., “design thinking”. Many have attempted to teach design thinking in the service of learning content, especially in math and science¹; others have emphasized its importance for training future designers and engineers². Despite the energy and enthusiasm of proponents, there has been relatively little research conducted on (1) whether design thinking is teachable, i.e. whether students will apply it somewhere else besides the immediate context, and (2) whether design thinking is beneficial for learning in these new contexts.

In our view, one of the most important features of design thinking is that, similar to the scientific method, it invites a focus on ways of thinking that can protect against common pitfalls when learning, solving problems, or doing creative work. For example, often times for fear of failure people will avoid exposing their ideas to constructive criticism, allowing them to blindly pursue their first idea without considering plausible alternatives. The scientific method includes strategies to counteract this tendency, such as exposing hypotheses to empirical test. Design thinking also includes strategies to prevent premature commitment to ideas. Many of these strategies are expressed in mantras such as “Fail early, fail often!” Instruction incorporating design thinking has the potential to benefit students because of its emphasis on arming individuals against the affective repercussions of failure, a difficult but necessary part of the learning process. It may be especially helpful for lower achieving students, for whom instruction tends to focus on remediation of content rather than the disciplinary strategies and practices of engineering.

In this paper, we report on preliminary results from a twelve classroom study in which we taught design thinking to middle school students (N~200) for five weeks, integrating it into the existing curriculum throughout several subject areas (Math, Social Studies, and Science). Instruction was the same in each condition except for a few key activities, which focused on a condition-specific design thinking strategy. Half of the classes were assigned to the Stakeholder Design condition. These students were taught design thinking in a way that emphasized the stakeholder relationship, specifically the importance of seeking critical feedback³. The other half of the classes were assigned to the Parallel Design condition, where the emphasis was on parallel prototyping⁴, specifically the benefits of coming up with many initial ideas and sorting them on key characteristics to induce patterns.

At the end of instruction, all students played two computer-based assessment games, each of which targeted only one of the design thinking strategies (parallel design or seeking critical feedback) emphasized within each condition. As opposed to self-report measures or paper-and-pencil tests, the games provide process data of the choices students make while learning in new contexts and within subject matter domains. Thus, they are a form of assessment that provides insight into students’ preparation for future learning⁵.

In what follows, we report on the preliminary analysis of the game-based assessment of instruction of two design thinking strategies: seeking critical feedback and parallel prototyping. We seek to answer two main research questions:

1. Did instruction of design thinking strategies influence students' learning choices in a new context?
2. Did these choices improve learning within a new context?

METHODS: TEACHING AND ASSESSMENT

The participants were 6th grade students (N~200) in a high SES public middle school. The study employed a 2 (Parallel vs. Stakeholder Design Conditions) x 3 (Math Placement: Regular Math, Advanced Math, Double-advanced Math) x 2 (Posterlet vs. Photolet) design. Specifically, twelve intact classes were randomly assigned to one of two conditions, Parallel Design or Stakeholder Design, stratified by the school's three levels of math tracking (Regular, Advanced, and Double Advanced). Instruction took place over about 20 class sessions spanning a period of five weeks, covering three main topic areas in succession: Math, Social Studies, and Science. At the end of instruction, students in both conditions played both the Posterlet and Photolet assessment games (described below).

In both conditions, the lessons integrated the subject matter content within the context of a series of design projects. In Math, design projects included designing a house or a candy box. In Social Studies, the project was to design a fair system for making class-level decisions. In Science, the project was to design a game that would teach people about factors contributing to the stability of ecosystems. Throughout the instruction, we introduced and referenced a simplified, three-phase "design cycle" (Figure 1). As the projects progressed, we made explicit references to where we were in the design cycle.

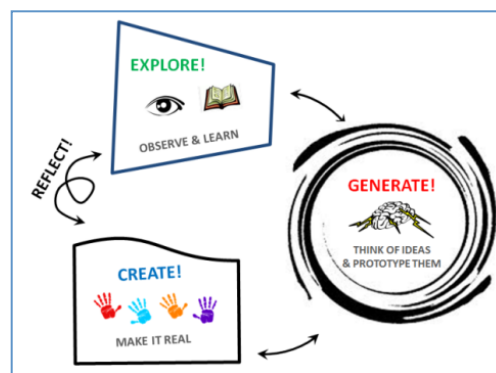


Figure 1 – A simple design cycle we used to organize activities into three phases: *Explore!*, *Generate!*, and *Create!*

The difference in instruction across conditions was a focus on one specific design method (Parallel Design or Stakeholder Design) at a few key points of a design cycle (Figure 2). Typically in the *Generate!* phase, in one condition the instruction focused on coming up with

multiple prototypes initially and comparing them (Parallel Design), whereas in the other the focus was on seeking constructive criticism from multiple stakeholders (Stakeholder Design).

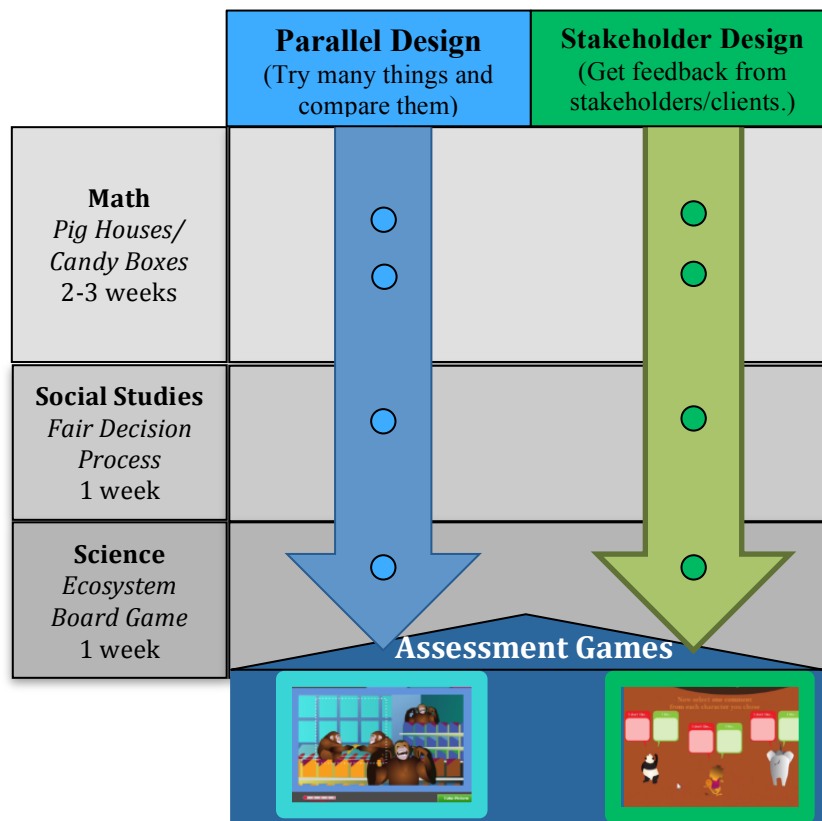


Figure 2 – Study design. Instruction was largely the same in each condition, except for key activities (dots), which emphasized one design thinking strategy specific to treatment. At the end of instruction, students in both conditions played both assessment games.

Instructional Sequence: Science

Here we describe the activities of the Science unit, which occurred during the last week of the instructional intervention, to illustrate both the similarities and differences between conditions. The science unit lasted four days, during which the overarching design project in both conditions was for students to design a card or board game that would teach someone about ecosystems. The central content goal was to understand what contributes to the stability of an ecosystem based on how energy flows through the various trophic levels (producers, primary consumers, secondary consumers, etc.). In the first two days of the unit, students played and then modified a few pre-existing games, which were based on aspects of ecosystem dynamics. On the third and fourth days, students designed their own ecosystem game, often using the first few games as inspiration.

In both conditions, students began in the *Explore!* phase of the design cycle by playing an ecosystem game based on the card game War. In Eco-War, each deck of cards contained organisms at several trophic levels, including producers (plants, trees, etc.), consumers (mice, squirrels, etc.), and secondary consumers (wolves, jaguars, etc.). A pair of students would each draw a card from the deck, and the player who drew the higher trophic level would keep both cards. For example, if one player drew a wolf, and the other drew a mouse, the student with the wolf would take both cards. The game would continue until one player ended up with all of the cards.

After playing a few rounds of Eco-War, students entered the *Generate!* phase in which they played a modified version of War that capitalized on a particular design thinking strategy. At this point, the conditions differed based on which specific strategy was emphasized. In the Stakeholder Design condition, students were asked to generate new rules to make the game into a “Peace” game, where the goal was to keep the game going as long as possible to simulate a stable ecosystem. Each student pair wrote down a new set of rules to handle the range of possible match-ups, how they would handle each scenario, and how it makes sense in terms of how ecosystems actually work. They then swapped rule sheets with another pair and played with each other’s set of rules. They wrote down feedback to improve the games, and shared the feedback with the people who came up with the rules. Feedback included constructive criticism on gameplay (e.g., missing rules, or rules that made the game too hard to play) and science content (e.g., unrealistic rules that did not correspond to their sense of how organisms or ecosystems would behave). Once the feedback was delivered back to the rule-writers, there was a brief discussion of the importance of seeking constructive criticism from stakeholders.

In the Parallel Design condition, students played a slightly different game based on a combination of War and Musical Chairs. Here, each student within a large group chose a single card and played that card against different partners in successive turns. With each round, students generated data on their whole group’s ecosystem, using a data sheet to keep track of how many cards at each trophic level they started off with, and how many turns the game lasted before there was only a single player left. After several rounds of the game, students got into new groups in which they sorted all of the data on multiple dimensions in order to find out what makes an ecosystem more stable. Upon sorting their data sheets, they discovered that “bottom-heavy” ecosystems, with more producers than consumers, were more stable. There was a brief discussion on the importance of generating multiple prototypes (in this case, multiple varieties of ecosystems including top-heavy, bottom-heavy, etc.) and sorting all of the data to find patterns.

After these initial game-playing activities, instruction for both conditions was the same. Students in both conditions played a board game emphasizing factors that contribute to the stability of an ecosystem, and then spent two class periods in the *Create!* phase, designing their own game that could teach others about ecosystems. A quick content assessment after the science unit found that students in both conditions understood that to create a stable ecosystem, you should have more producers than consumers.

We described the science unit not as a model of instruction, but to characterize what was the same and what was different across conditions. Most of the instruction was the same, with one or two condition-specific activities within each subject matter unit included to target either the

Stakeholder Design or the Parallel Design strategy. This is not a study to demonstrate the effectiveness of this type of instruction, but rather to establish an existence proof that design thinking methods can be (1) learned in ways that are useful in new situations, and (2) assessed. Next we discuss the assessments of the design thinking methods.

Assessment: Choice-based assessment games

At the end of the five weeks of instruction, all students played both choice-based assessment games⁶. These assessments do not test recall of content learned during the instruction. Rather they assess students' approaches to learning. They are a principled way of collecting process data that can be used to assess how students choose to learn in a new situation and within a new content domain. The purpose is to determine whether the design-based instructional experiences influenced how students would approach a new learning task. Specifically, in a new context, would students who went through the Stakeholder Design experiences be more likely to choose negative feedback to help them? Similarly, would students who went through the Parallel Design experiences be more likely to explore multiple ideas and look for patterns?

The games are designed to not look like assessments, so that they capture everyday learning behaviors, not "maximal performance" test behaviors^{7,8}. Additionally, the games give students multiple paths to success, allowing them to freely choose among multiple ways of learning *during* the game⁵. This allows the researcher to compare the choices students make in the game with measures of learning (both internal and external to the game) to establish construct validity.

Posterlet

Posterlet is a game in which players can learn how to design posters that adhere to basic graphical design principles. The players are volunteers for the school's Fall Festival for which they are asked to design posters to advertise the various activity booths (Figure 3). Their job is to choose an activity booth (Step 1), design a poster for it (Step 2), and "market test" the poster with a focus group (Step 3). Students then choose either positive ("I like...") or negative feedback ("I don't like...") (Step 4) from each of three characters (see Figure 3.B). After reading the feedback (Step 5), students can choose to revise (Step 6) or submit their poster (Step 7). Upon submission, students can see their booth's "ticket sales" reflecting the quality of their poster design. There are three rounds of posters to design, thus nine opportunities to choose positive or negative feedback. The key measure is whether the students choose negative feedback. In previous studies⁹, we have found that choosing negative feedback in Posterlet correlates with learning within the game (poster quality score & post-test scores), as well as learning outside the game (school grades and standardized test scores).

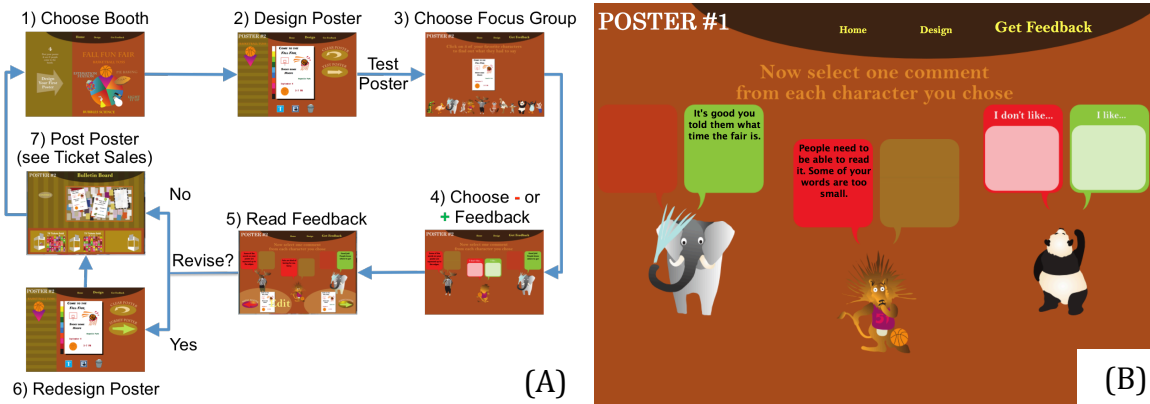


Figure 3. (A) Flowchart of Posterlet game mechanics and (B) the selection of positive or negative feedback.

On the back end, the game has a set of 21 graphical design principles, which are used to evaluate the poster, generate feedback, and calculate poster scores. These principles encompass what information to include (e.g., date and time), readability (e.g., high contrast between font and background color), and aesthetics (e.g., proper use of available space). They were generated in consultation with a graphic artist. We make sure that the positive and negative feedback has an equal chance of being informative, although negative feedback is more actionable in this context (Table 1).

Category	Positive	Negative
Information	It's nice that the poster says how much the booth costs.	You didn't say how much the booth was.
Readability	Your poster has big letters. Really easy to read.	People need to be able to read it. Some of your words are too small.
Space Use	Great job on getting your pictures away from the edge.	Your picture is squished up next to the edge.

Table 1. Examples of positive and negative feedback phrases for each type of feedback (informative or uninformative) and each category (information, readability, or space use)

After completing three posters, the students take a separate post-test, which probes their learning of the graphical design principles. The post-test comprises two parts. The first part asks students to rank the quality of two similar posters, which differ only by a few design mistakes, and repeat this process three more times for other poster pairs. The second part asks students to evaluate the good and bad properties of a model poster that we provided, by checking items from a checklist of graphical design principles.

Photolet

Photolet is a game in which players can learn how to adjust camera settings to take clear pictures (Figure 4). The player's character is a photographer for a newspaper for which they receive several assignments. The assignments are to take pictures of specific animals amid scenes depicting several animals and objects in interaction. In a sequence of assignments of increasing difficulty, the assigned animal may be in the foreground or background, or may be moving. To capture them clearly the students must learn to use the proper camera settings (lens setting, shutter speed, and aperture). The students can learn about the camera settings in multiple ways. They can take many pictures while adjusting the camera settings, and sort picture along different dimensions, such as brightness. They can also choose to look at brief animations that explain the physics behind how different camera features work. We expected students in the Parallel Design condition to generate photos using more settings and to be more likely to sort them to look for patterns, as opposed to only taking photos until the first suitable photo was taken. The key measure for this game is whether students make many changes to the camera settings, in line with the Parallel Design approach of generating lots of initial ideas and looking for patterns.

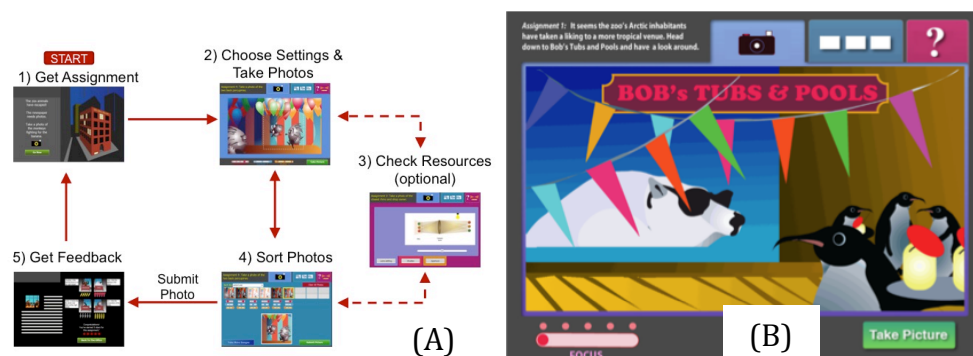


Figure 4. (A) Flowchart of Photolet game mechanics, and (B) screenshot of a photo assignment

After completing Photolet, the students took a post-test to assess how well they had learned the ways camera settings affect the appearance of photos. They saw seven pairs of real photos and they had to determine what camera setting had changed between the two. They also received four multiple-choice questions that probed their declarative understanding of camera setting effects.

RESULTS

Here we report on the results of a preliminary analysis of the game-based assessments, which the students took after five weeks of design thinking instruction across the three content areas. There is good reason to think that design thinking will benefit the lower achieving students most, since instruction for them tends to focus on a remediation of content rather than on the disciplinary strategies we taught. We therefore analyze the results of each assessment broken out by the level of math placement (Regular, Advanced, and Double Advanced).

In the research design, the two separate conditions (Stakeholder vs. Parallel Design) take *both* assessments at the end of instruction. Since each game measures only one of the targeted design

thinking strategies (choosing critical feedback vs. choosing to make multiple prototypes), each condition should look better on one assessment game while acting as a control on the other. We discuss the results of each assessment in turn.

Posterlet results

The key dimension for the poster design game-based assessment was the number of times they chose negative feedback, of which there were a total of nine in-game opportunities (three for each poster). We explored the variation in choices to seek negative feedback by condition and math placement (Figure 5). For both conditions, the high-tracked students chose more negative feedback than the other students. However, the lower-tracked students in the Stakeholder Design condition (treatment) chose more negative feedback than the lower-tracked students in the Parallel Design condition (control).

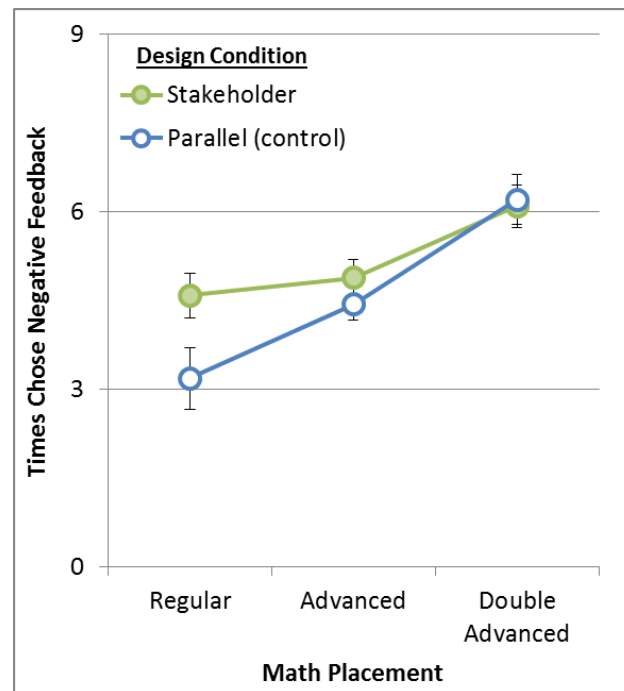


Figure 5: Results for Posterlet. The lower-tracked students in the Stakeholder Design condition chose more negative feedback than the lower-tracked students in the control condition.

We also examined whether choosing more negative feedbacks led to improved learning of the graphical design principles. We found that the number of negative feedbacks chosen correlated with the post-test score, $r(139) = .21$, $p < .05$, suggesting that this was a helpful strategy for learning the graphical design principles.

Photolet results

The key dimension for the photography game-based assessment was the number of different camera settings tried. We expected this assessment to pick up effects of the Parallel Design

learning experience, thus the Parallel Design condition serves as the treatment for this assessment, and the Stakeholder Condition the control. We analyzed the average number of camera settings tried in the game by condition (Stakeholder vs. Parallel) and math placement (Figure 6). The highest-tracked students in both conditions tried more camera settings than the middle- and lower-tracked students. However, the lower-tracked students in the treatment condition (Parallel Design) tried more camera settings than the lower-tracked students in the control (Stakeholder Design).

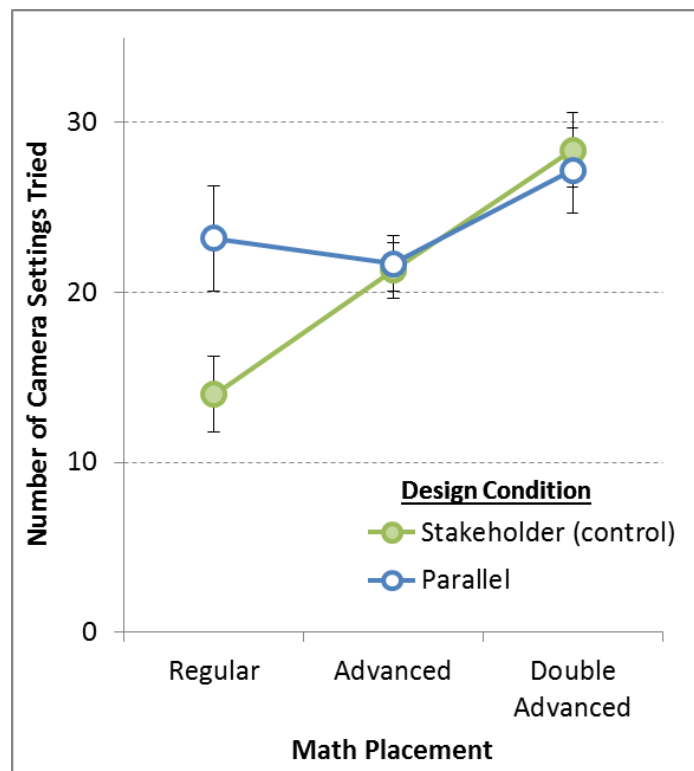


Figure 6: Results for Photolet. The lower-tracked students in the Parallel Design condition tried more camera settings than the lower-tracked students in the control condition.

We also examined whether trying more camera settings led to improved learning of how camera settings work. We found that the number of camera settings correlated with the post-test score, $r(174) = .24, p < .01$. This suggests that trying many camera settings was a helpful strategy for learning how camera settings affect the appearance of photos.

In summary, for the lowest-tracked students we saw significant differences between conditions for both measures, in the expected directions. The Stakeholder Design condition chose more negative feedback than the Parallel Design condition, while the Parallel Design condition chose more combinations of camera settings than the Stakeholder Design condition. In both cases, the main source of condition difference was for the lower-tracked students. This suggests that the higher-tracked students already knew how to apply these strategies, while the lower-tracked students benefited from instruction of the design thinking strategies.

Through correlations with post-game measures, we also found evidence that the strategies were helpful for learning in their respective games. Moreover, performance on one assessment game was uncorrelated with performance on the other game, providing divergent validity that we were indeed measuring separable design thinking skills that could independently contribute to learning. Taken all together, these results suggest that teaching design thinking strategies can help prepare students to learn in new situations.

DISCUSSION

The primary purpose of this research was to determine whether we could measure behavioral effects of design thinking experiences when students are placed in a new learning situation. While design thinking is increasingly incorporated into learning in school and out of school, and many people find it intuitively compelling, there is a general lack of empirical evidence of the benefits of teaching design thinking. We taught design thinking strategies to provide middle schoolers with tools to overcome the fear of failure: seeking negative feedback and parallel prototyping. Our results show that students learned to apply these skills in new contexts, and that the strategies helped them learn in these new contexts. Moreover, we have also shown that we can detect the benefits of teaching design thinking with game-based assessments of their choice behaviors.

More broadly, our results suggest that design thinking strategies can help all students learn in new situations. High-tracked students are often characterized as being overly focused on providing correct answers, and therefore as being afraid of being wrong, but their high performance on both assessment games regardless of condition suggests that they know how to apply these design thinking strategies related to “failing early and often” when given a context in which that is valued. Our results also offer an existence proof that the lower-tracked students can learn these strategies as well, and moreover that they benefit the most from learning design thinking strategies.

REFERENCES

- [1] Kolodner, J.L., Camp, P.J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., Puntambekar, S., & Ryan, M. (2003). Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting Learning by DesignTM Into Practice. *The Journal of the Learning Sciences*, 12(4), 495-547.
- [2] Dym, C.L., Agogino, A.M, Eris, O., Frey, D.D., Leifer, L.J. (2005). Engineering design thinking, teaching, and learning. *Journal of Engineering Education*, 94(1), 103-120.
- [3] Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological bulletin*, 119(2), 254-284.
- [4] Dow, S. P., Glassco, A., Kass, J., Schwarz, M., Schwartz, D. L., & Klemmer, S. R. (2010). Parallel prototyping leads to better design results, more divergence, and increased self-efficacy. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 17(4), 1-24.
- [5] Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review Of Research In Education*, 24, 61-100.
- [6] Schwartz, D. L., & Arena, D. (2013). *Measuring what matters most: Choice-based assessments for the digital age*. Cambridge, MA: MIT Press.

- [7] Klehe, U. C., & Anderson, N. (2007). Working hard and working smart: Motivation and ability during typical and maximum performance. *Journal of Applied Psychology*, 92(4), 978-992.
- [8] Sackett, P. R., Zedeck, S., & Fogli, L. (1988). Relations between measures of typical and maximum job performance. *Journal of Applied Psychology*, 73(3), 482-486.
- [9] Cutumisu, M., Blair, K.P, Schwartz, D.L., & Chin, D.B. (2015). *A game-based assessment of children's choices to seek feedback and to revise*. Paper to be presented at the American Educational Research Association, Chicago, IL.