

Effective Feedback Depends on the Emotional Weight of a Problem

Elliott Hedman
mPath
Boulder, USA
hedman@media.mit.edu

Christine Zanchi
Curriculum Associates
Billerica, USA
czanchi@cainc.com

Garrett Hedman
Northwestern University
Chicago, USA
GarrettHedman2022@u.northwestern.edu

I. PURPOSE

This study explored engagement in i-Ready lessons, which are online instructional modules for students in Grades K-8. Over the course of a three year exploratory research study, we researched how to engage students with digital lessons, prototyping alternative methods to traditional instruction delivery such as lecturing, which can be disengaging (Hedman, 2018). Feedback is considered one of the most powerful ways to increase learning (Hattie & Gan, 2011). How does feedback affect children's emotional engagement with content, and, inversely, how does emotional engagement affect children's attention to feedback? In our research we asked two questions:

- When a student answers a question incorrectly, what challenges can prevent students from paying attention and fully engaging with feedback?
- What software features lead students to pay attention and emotionally engage with digital feedback?

II. THEORETICAL FRAMEWORK

Emotions play a key role in student's engagement and attention around achievement. Emotions such as enjoyment of learning, hope, pride, anxiety, shame, and boredom affect student's performance and identity as a learner (Schutz & Pekrun, 2007). A substantial focus has been on test anxiety (Hembree, 1988; Zeidner, 1998, 2007). On the other end of text anxiety – boredom and apathy – the research is more scant (Pekrun et al., 2010). Additionally, psychometric scales mostly measure the feeling of achievement outcomes – how one feels when they receive a good or bad grade (e.g. Weiner, 1985). Research has not focused on feelings during a lesson, such as the feelings when receiving feedback (Perkun et al., 2009). One exception to this is digital tutors, which do measure emotions during lessons (e.g. D'mello & Graesser, 2012). However, this research remains high level (Which emotions occur during a lesson?) and has not delved into granular relationships such as what is the direct effect of feedback on emotional state.

Feedback is one of the most powerful tools digital lessons can use to teach a student (Hattie & Gan, 2011). When

students are given feedback on an item they can confirm or alter their knowledge (Mory, 2004). For our study, we focused exclusively on elaborated, immediate feedback (Shute, 2008). We found four meta-analysis on the role of feedback in digital education (Azevedo & Bernard, 1995; Jaehnig & Miller, 2007; Van der Kleij et al., 2011; van der Kleij et al., 2015). Like emotion research, feedback research was focused on outcomes – how much does feedback affect scores? Independent variables tested included feedback type, feedback timing, education level, and subject. Student's emotional reactions and motivation were not taken into consideration in these tests. This article aims to show that the in-lesson, emotional experience students have with feedback should be considered when evaluating feedback effectiveness.

III. METHODS

151 students participated in this study, with the majority of participants coming from Title 1 schools in Hartford and Denver. At the Boys and Girls Club of Metro Denver locations, 19.5% of students were African American and 64.5% of students were Hispanic (Omni Institute, 2015). At the Hartford school, 41% of students were African American and 28% of students were Hispanic. A range of students with varying skills, from 2nd to 8th grade, were tested. Most students were at least one year below grade level in reading and math. In the baseline, students received i-Ready lessons that covered reading comprehension, vocabulary, and math. During these lessons, students were given quizzes that determined the lesson score. For our prototype, we gave students a 4th grade reading assignment, level appropriate multiplication problems, and lessons on binary and Roman numerals.

Students were asked to either use current i-Ready software (baseline) or a prototype where we iteratively altered functionality to improve feedback effectiveness. Wireless eye tracking glasses helped assess whether students were visually paying attention and watching animated lectures. We used wireless skin conductance sensors to measure when students were emotionally reacting to content, a correlate to engagement (Hedman, 2014). Electrodes were dry and applied to the middle and index finger's distal phalanx. We had a well-trained ethnographer observe and interview the students about their experience with i-Ready. When possible,

researchers left the child's field of view as an attempt to remove observational effect.

IV. EVIDENCE

A. *Factors that can make feedback ineffective from the student perspective*

Feedback is a critical tool for the students we observed in lessons. As a student attempted to solve the area of the triangle with the same formula as the area of a square, we saw his frustration, and his clear need for a moment of feedback: "Remember to multiply by one half!" The literature confirms this critical need for quality feedback (Hattie & Gan, 2011). But what surprised us was the negative frustrations students can have with feedback. Not because the feedback is unhelpful, but because students have additional goals beyond "knowing how to answer the problem right." In this article we share student perspectives on feedback and suggest ways educational technology may address student frustrations. Below we list four challenges with feedback from the students' perspectives.

1) *Feedback "gets in the way" of immediately trying again*

Our initial baseline research suggested feedback in the digital lessons was viewed as a negative experience for many students. As the program delivered unsolicited feedback, students ignored and often avoided it. When asked about digital lessons, one student said, "They just keep talking." Another student pointed out that he knew what the answer was after the first incorrect answer, but was now forced to listen to the feedback before he could click on the answer. From our eye tracking videos, we saw multiple students hover their mouse over their next answer as they waited for the feedback to finish. When a 2nd grade boy answered a problem wrong, he quit and re-logged in so he could answer the question without having to hear the feedback. For many students answering a problem correctly is their main goal. When they answer a problem wrong, they want to show they know the answer even more. Feedback that prevents students from answering when they feel ready is viewed as an obstruction of the student's main goal, which is to show she knows the answer.

2) *Guessing is a legitimate strategy*

For many students, guessing was a legitimate strategy to answering the question. After needing to guess three or four times on a four-question multiple choice, some students told us they answered the question correctly and their "correct" answer showed they knew it. When the software gave the answer after a second incorrect guess, students would complain that they were "robbed" of showing they knew it. As one fifth grade student pointed out, "When you get a problem wrong two times, it gives you the answer, and it's not fair." In our prototypes we let students guess multiple times as a way to boost confidence. Consequentially, students do not see a need or benefit for feedback in these cases – they can always just guess again.

3) *Students want to show what they know, without help*

Furthermore, students have reported to us that listening to feedback is like cheating. In some of our prototypes we implemented an "I don't know" button so students could ask for feedback. No student pushed that button, explaining to us they could figure it out on their own (typically by guessing). Asking for a hint was like saying, "I can't do this and need help," which most students actively avoided. When we asked two 4th grade girls how they felt about feedback they began wagging their fingers and mimicking a high toned, lecturing voice; for them feedback was like a parent telling them they are incapable. Students not only wanted to show they knew the material, but that they could answer the questions on their own.

4) *If you get something wrong, the computer's response is viewed as a "punishment"*

Some students view any form of response after they get a problem wrong as a punishment. In one math prototype we told students because they answered a problem wrong they were going to do extra problems. Reluctant students were frustrated with these extra questions – "Do I have to?" One could imagine the student thinking, "If I got the problem right, I wouldn't have to do this." Regardless of their previous score, students were going to do more problems, but framing the new problems as a consequence of failure disengaged reluctant students. From these insights, we removed all mentions of "because you got this wrong" from future prototypes. However, feedback occurring right after a wrong answer can still be viewed as the "punishment."

B. *Consequence of children disliking feedback*

When feedback was given, students often did not pay full attention. During reading assignments where words and sentences became highlighted, students seldom looked back at those highlights, instead focusing on their next guess. This lack of attention occurred for both high and low achieving students. Other times we saw students drift off during feedback looking around the room, waiting for the talking to stop. Some students clicked on the digital marker tool and started drawing on the screen while the feedback finished. When feedback is viewed negatively, it can become ineffective as students do not engage and pay attention to the material.

Below is the skin conductance of a 6th grade boy who does not emotionally respond to feedback in a math lesson about finding the median.

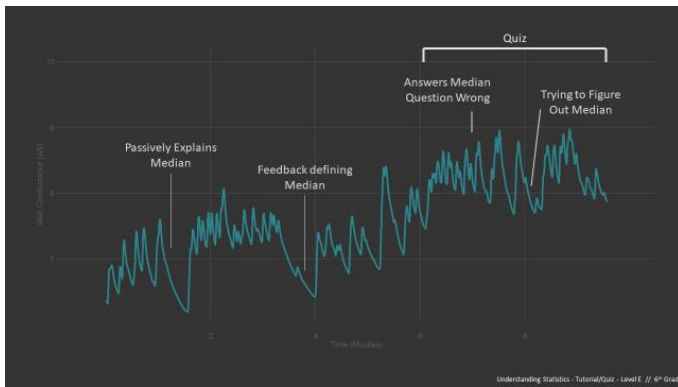


Fig. 1 Skin conductance of a boy learning about median

When he is first told what median is and when he is given feedback about how to find the median, his skin conductance trails downward, an indicator that he is not emotionally engaging. This boy cared about passing the quiz; however, ungraded feedback during the lesson did not grab his attention. Consequentially, he did not know how to find the median during the quiz.

C. Addressing challenges to design effective feedback

1) Adding “weight” to problems creates demand for feedback

As we noticed the above challenges across students, we attributed the distaste for feedback, in large part, to the lack of “weight” of problems. Because, from the students’ perspective, there was “no consequence” for answering a problem right or wrong, the problem had no “weight.” In elementary school, the end lesson score did not matter and was not acknowledged. In middle school, answers during lessons (before the quiz) did not affect student’s score. Kids could answer these questions wrong, and, from their perspective, there would be no effect.

Through multiple iterations, we attempted to create “weight” in problems via small rewards, like a star, egg, or penguin. Students needed to correctly answer 100% of the questions in order to receive these rewards. We gave the child two attempts on each question, and as such, the feedback became meaningful: “This will help save you.” Children’s behavior substantially changed as they began forming goals around earning these small acknowledgements.

As an example, consider one middle school boy who unenthusiastically did a reading lesson. With our eye tracking glasses, we saw him skip over all of the reading and start guessing at the questions. He then noticed a dotted line for a coin and asked us what that was used for. We explained that correct answers would earn him coins. Upon realizing there was a consequence – meaningful to him - he went back to the reading and then read the article and carefully considered each answer. For this student and many others in this study, small acknowledgments of correct answers added weight to the questions.

Adding weight to our lessons dramatically changed student’s behavior in regards to feedback. Multiple students began raising their hands and asking us for help on questions – a behavior we had not previously observed. In response, we

created a paper button that said “break it down” so the students could ask for help when they needed it, and was “clicked” during most future lessons.

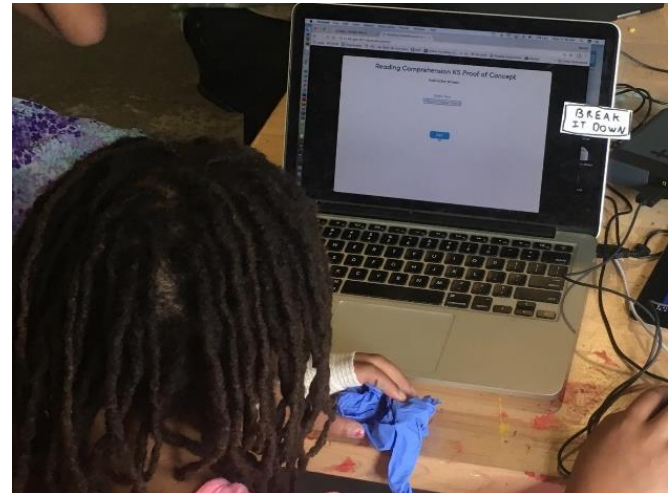


Fig. 2 We add a “break it down” button when students start asking for help to earn an acknowledgement.

In some iterations of “break it down” students began complaining that feedback was not helpful enough – a good sign they were paying attention to feedback. In our current lesson design we now begin with the design principle “Create lesson frameworks where students actively ask for feedback,” which typically leads to designing clear acknowledgements and markers of success such as coins and stars.

In our latest iteration with small rewards, we successfully had 34 students at the Boys and Girls Club learn how to read Roman numerals with feedback as the only mechanism of receiving information. Below is the skin conductance response of a 4th grade girl during the lesson.

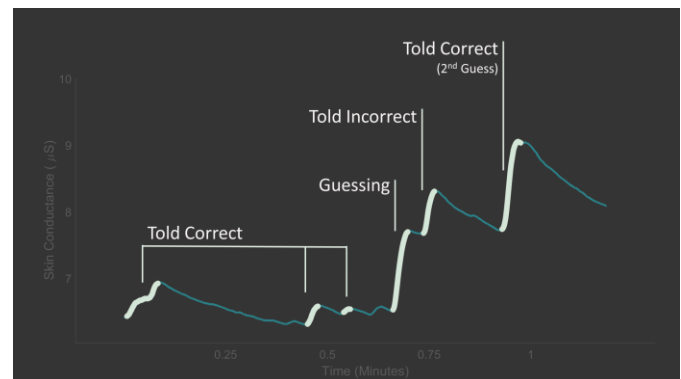


Fig. 3 Skin conductance response of a 4th grade girl reacting to questions that lead to a star

After answering four correct questions this child received a star. As she is working her way towards earning that star, she reacts each time she answers a question right, and even more when she answers incorrectly. When she guessed and answered incorrectly her skin conductance first rose 1.2 μS with 0.6 μS more when she found out she was incorrect. When her second answer (last chance) is marked

correct, her skin conductance increases 1.3 μS . The child then turned to her friend and snickered. As her psychophysiological response shows, earning a star created substantial weight for each question.

2) In middle school, quizzes and grades add the most weight

When the middle school boy from Figure 1 was given feedback during the lesson answering correctly had little weight: lesson questions did not directly affect his classroom grade, so why should he seek help on answering correctly? During the quiz, when his grade was affected, the boy's skin conductance increased 1.5 μS and he reacted more: answering correctly affected the score he cared about.

Similarly, this 6th grade boy from the same class reacts more during the quiz.

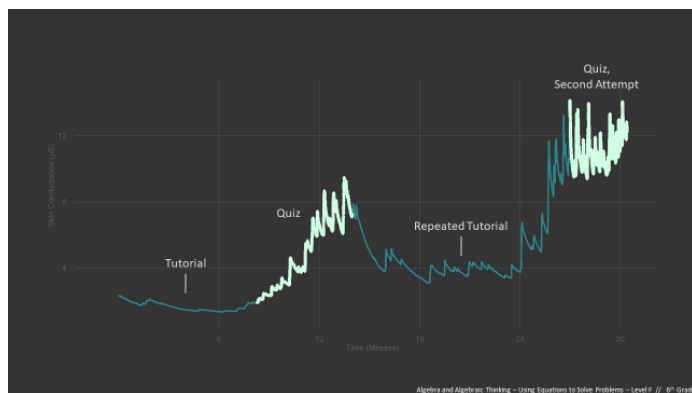


Fig. 4 Skin conductance of a middle school boy reacting to the quiz.

During the initial tutorial, he does not react to the feedback (0 μS) as answering correctly did not affect his grade. During the quiz the child started reacting substantially more (8 μS). After failing, he took the quiz again with an even stronger response.

3) Feedback can be most effective in the context of classroom grades

Middle school students in our study tend to care most about receiving a good score during the quiz, which puts test developers in an interesting bind with regard to feedback. Typically, quizzes are ways of evaluating student's knowledge so students cannot receive feedback or guidance. But these graded moments are when students are most engaged in answering correctly, and therefore to learn and seek feedback.

4) In delivering feedback, the details matter: don't force it

In one attempt for a reading comprehension prototype, we tried to make feedback more interactive and concise using single-sentence, simple task directions. Students were instructed to click on a word in the text and heard feedback about the clicked-on sentence. This interactive solution did engage children's attention: students started listening. But, because they could not skip the feedback, they became frustrated with the "forced" delivery of the feedback. For our other prototypes, most students did not mention feedback issues without the interviewer prompting. However, almost

every child complained about the "forced" feedback in this iteration. They were more aware of their incorrect answer and annoyed at having to listen to feedback, which felt like a penalty. We addressed this problem by framing feedback as a "new question" rather than a response to the incorrect guess.

V. LIMITATIONS AND FUTURE WORK

In response to this research, we continue to improve feedback and engagement through an iterative design process. We are now developing ways for students to receive feedback during quizzes while still accurately assessing their knowledge. In other prototypes, we changed the grading model where all parts of a lesson are scored: not just the quiz. We've prototyped the ability for students to try again during lessons, so they can fail and still receive 100%. We are now considering new models where grades are tied to demonstrating knowledge rather than answering a set of questions. This work is continuously evolving as we strive to deeply understand the relationship between student engagement, feedback and better learning outcomes for all students.

This work was exploratory. Early on, we observed passive lecturing content often disengaged, but we did not initially consider feedback as disengaging. Issues with attention feedback naturally emerged as students reflected on our iterative prototypes where feedback was added, emphasized, and altered. Findings were observed across multiple students at different grade levels and abilities. However, due to the natural limitations of qualitative studies, researchers should not generalize our results across all students.

Future controlled, quantitative research should consider asking: What percentage of students feel negatively about feedback? How does a child's test performance affect their desire for feedback? Our prototypes had students opting into (and not ignoring) feedback more often, but these prototypes need further validation in real classrooms with controlled comparisons. Our current findings describe influential variables to consider across feedback research. Future controlled, AB-test research would help bring these findings to more generalizable conclusions.

VI. SCHOLARLY SIGNIFICANCE

Conclusion 1: Some forms of digital feedback are undesirable and viewed as punishment by students, leading students to disengage and not pay attention to feedback. As researchers implement learning software, they should ask, "Are students paying attention and engaged with this feedback?" If the student is not seeking feedback, low scores may be due to student inattention rather than the feedback strategy itself.

Conclusion 2: The importance of answering a question correctly, its weight, is a strong determinant of effective feedback. If designers create assessments where students value answering questions correctly, feedback will likely be more effective. When considering feedback, researchers should document how invested students are in answering correctly

and look at ways of increasing that importance. How can the natural weight of quizzes, grades, and rewards tie into the feedback students are given? We envision a day where evaluation, feedback, and learning are more integrally tied together and students are invested and care about learning throughout the entirety of a digital lesson.

REFERENCES

- Azevedo, R., & Bernard, R. M. (1995). A Meta-Analysis of the Effects of Feedback in Computer-Based Instruction. *Journal of Educational Computing Research*, 13(2), 111–127. <https://doi.org/10.2190/9LMD-3U28-3A0G-FTQT>
- D'mello, S., & Graesser, A. (2012). AutoTutor and affective AutoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 2(4), 23.
- Hattie, J., & Gan, M. (2011). Instruction based on feedback. *Handbook of Research on Learning and Instruction*, 249–271.
- Hedman, E. (2014). Thick psychophysiology for empathic design. Retrieved from <https://dspace.mit.edu/handle/1721.1/91436>
- Hembree, R. (1988). Correlates, Causes, Effects, and Treatment of Test Anxiety. *Review of Educational Research*, 58(1), 47–77. <https://doi.org/10.3102/00346543058001047>
- Institute, O. (2015). *2014 – 2015 Evaluation of the Boys & Girls Clubs of Metro Denver*. Denver. Retrieved from <http://www.bgcmd.org/wp-content/uploads/2016/10/2014-15-BGCMD-Evaluation-Report.pdf>
- Jaehnig, W., & Miller, M. L. (2007). Feedback Types in Programmed Instruction: A Systematic Review. *The Psychological Record*, 57(2), 219–232. <https://doi.org/10.1007/BF03395573>
- Mory, E. H. (2004). Feedback research revisited. *Handbook of Research on Educational Communications and Technology*, 2, 745–783.
- Pekrun, Reinhard and Goetz, Thomas and Daniels, Lia M and Stupnisky, Robert H and Perry, R. P. (2010). Boredom in achievement settings: Exploring control–value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology*, 102(3), 531. Retrieved from <http://doi.apa.org/journals/edu/102/3/531.html>
- Pekrun, R., Elliot, A. J., & Maier, M. A. (2009). Achievement goals and achievement emotions: Testing a model of their joint relations with academic performance. *Journal of Educational Psychology*, 101(1), 115.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78, 153–189
- Schutz, P., & Pekrun, R. (2007). *Emotion in education*. Retrieved from <http://psycnet.apa.org/record/2007-04736-000>
- Van der Kleij, F. M., Feskens, R. C. W., & Eggen, T. J. H. M. (2015). Effects of Feedback in a Computer-Based Learning Environment on Students' Learning Outcomes. *Review of Educational Research*, 85(4), 475–511. <https://doi.org/10.3102/0034654314564881>
- Van der Kleij, F. M., Timmers, C. F., & Eggen, T. J. H. M. (2011). The effectiveness of methods for providing written feedback through a computer-based assessment for learning: A systematic review. *Cadmo*.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92(4), 548–573. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/3903815>
- Zeidner, M. (1998). *Test anxiety: The state of the art*. Retrieved from <https://books.google.com/books?hl=en&lr=&id=oMO6tnkjc5EC&oi=fnd&pg=PR3&dq=zeidner+1998+test+anxiety&ots=uiW3LXmxR&sig=KEO3AWAGdVsqJYyrEQcatszr1IU>
- Zeidner, M. (2007). Test anxiety in educational contexts: Concepts, findings, and future directions. In *Emotion in education* (pp. 165–184). Elsevier.