# Choosing Negative Feedback Improves Learning for Students of All Ages: A Game-based Assessment of Seeking Negative Feedback and Revising

Maria Cutumisu
*cutumisu@stanford.edu*
*Stanford University, United States*

Daniel L. Schwartz
*danls@stanford.edu*
*Stanford University, United States*

## Abstract

*This research examines the effect of age on students' choices to seek negative feedback and to revise, as well as on students' learning based on these choices. We designed Posterlet, an assessment game that measures the choices to seek negative feedback and to revise. In this study, 764 students played Posterlet, in which they designed posters and learned graphical design principles from feedback. Results showed that seeking negative feedback correlated with revision and with learning of graphical design principles. Notably, age did not influence the frequency of seeking negative feedback and it did not moderate the relation between seeking negative feedback and learning. We demonstrate that it is possible to measure learning choices, and we provide evidence that such behaviors are worth measuring and, perhaps, teaching to all ages.*

## 1. Introduction

Yackel and Cobb [1] stated, "the main purpose of education is autonomy." Our goal is to prepare students to be independent learners who can make good choices about their learning when they leave school. Free choice is a *sine qua non* condition for autonomy and, if that is our educational goal, we need assessments that can measure our success in achieving this goal. However, traditional assessments measure declarative or procedural knowledge, expecting a right/wrong response. Additionally, they offer a retrospective measure of students' learning, without capturing students' ability to continue learning on their own. Vygotsky [2] indicated that measuring knowledge at the end of instruction does not offer an insight into students' learning processes. Instead of only summative assessments, we need dynamic assessments to determine if students are prepared to learn. That way, we can understand the conditions that can help students' evolution as independent learners. For independent learning, we need assessment tools that can measure students' abilities and willingness to make good learning choices on their own [3]. Presently, there are automated environments, such as intelligent tutors [4], that embed assessments into the learning environment. However, they impose the steps that students take while learning, leaving students little choice. An alternative way to measure whether students are effective independent learners is to examine their free, not right or wrong, choices in a choice-abundant environment that offers learning opportunities and encourages students' typical learning behaviors, as opposed to their test-taking behaviors [5]. Our solution is to create a novel kind of assessment that measures students' learning choices and whether these choices improve learning.

In this paper, we focus on students' choices and abilities to learn from negative feedback. We designed the *Posterlet* game environment to measure two behaviors important for learning: students' choices to seek negative feedback and to revise. The game offers choices, opportunities to learn, and an enjoyable environment in which students are more likely to express their natural inclinations toward making choices. In preliminary research [6], we found that seeking negative feedback correlated with standardized achievement scores for middle-school students. Here, we further explore the relation between learning choices and learning outcomes, and whether our assessment is effective for students of all ages. We combined six datasets to examine the effects of age and we posed three research questions:

- Does age correlate with learning choices and learning outcomes?
- Do learning choices correlate with learning outcomes?
- Does the relation between learning choices and learning outcomes vary systematically with age?

## 2. Theoretical Framework

We introduce a game-based assessment approach focused on the construct of *choice*. We refer to choice as an opportunity to decide what and when to learn [7]. Our theoretical framework is centered on constructivist assessments [8], specifically *choice-*

*based* assessments [7,9], which measure not only students' knowledge but also their choices about learning. In contrast to using choice for motivation and learning during instruction [10], we consider choice to be the outcome of learning in its own right. Measuring choice as an outcome of students' interest is important for students' learning in the future, when they are not told what to do and must learn on their own. The Posterlet game is an instance of a *preparation for future learning* [3,11] assessment designed to measure two likely important choices for learning: seeking *negative feedback* and *revising*.

**Measurement Constructs**. We measured the choices to seek negative feedback and to revise, because we assume that they are desirable learning behaviors. While negative feedback is more effective for learning than positive feedback [12], it may also trigger an ego threat [13], which may cause people to shut down instead of heeding the feedback. In prior feedback research, students receive feedback without exercising a choice [14]. However, there is no evidence to date whether the *choice* of feedback is important for learning. One reasonable prediction is that students who feel threatened by negative feedback may not seek it and, therefore, learn less. Revising can also be valuable, but this choice has not been investigated either. Thus, we aim to investigate the impact of the *choice of seeking negative feedback* and of the *choice of revising* on learning. Additionally, we focus on the impact of age on students' learning choices and learning outcomes. Prior research [15] found age differences in negative feedback processing, but it tested the participants' ability to estimate their likelihood of experiencing adverse life events and then to re-estimate it from memory based on feedback, focusing on the valence of feedback rather than on the choice between positive and negative feedback.

## 3. Posterlet: An Assessment Game

Posterlet is a game in which players design posters for different booths at a Fun Fair (Figure 1).

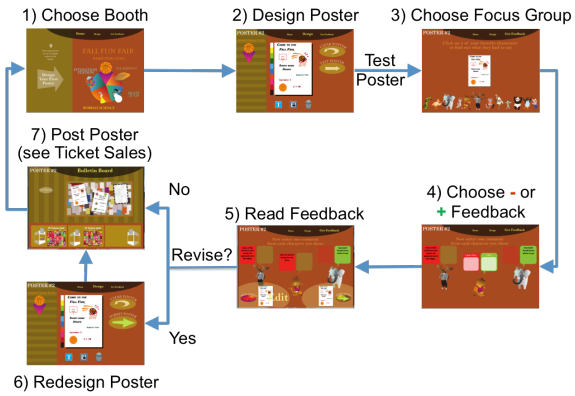


**Figure 1. Posterlet game flow**

Posterlet offers students the opportunity to learn graphical design principles through designing posters, seeking feedback, and revising their posters. The game enables players to choose positive or negative feedback for their poster designs and to revise their posters. Positive feedback (e.g., "Your poster has big letters. Really easy to read.") and negative feedback (e.g., "People need to be able to read it. Some of your words are too small.") carry equivalent information. There are 21 graphical design principles that our system evaluates and uses to provide feedback (e.g., text is readable, colors make reading difficult, etc.). Posterlet evaluates each poster based on technical features of graphical design, not based on students' artistic flair.

On each level, players complete their initial poster design using graphical design tools that enable them to select and place predefined text and images on the poster canvas for their chosen booth theme. Players can change the font, alignment, and color of the text, the color of the background, etc. Then they choose three characters from a focus group to provide feedback about the poster. For each character, the players choose either positive or negative feedback. The game's graphical analysis system generates feedback customized for the poster, according to the player's request for positive or negative feedback and the poster's adherence to the 21 graphical design principles. After reading the feedback, players choose to revise or submit the poster. Finally, the game displays the number of tickets sold at each booth, so that players have a general sense of their posters' quality. The number of tickets sold is based on the game's graphical analysis system. Students create three posters, with nine choices to seek negative feedback and three choices to revise.

## 4. Method

### 4.1. Participants and Procedures

Participants were N=764 students from several US public middle schools, a high school, and a community college, located in the states of California, Illinois, and New York (Table 1).

**Table 1. School and participant information**

| School | Grade | Age (US) | Played Game | Post test | Gender | |
|---|---|---|---|---|---|---|
| | | | N | N | F | M |
| Central | 6 | 11-12 | 22 | 22 | 13 | 9 |
| Chicago | 6-8 | 11-14 | 203 | 194 | 43 | 30 |
| NYC | 6-9 | 11-15 | 278 | 231 | 44 | 97 |
| Chi/NY | -- | -- | 36 | 29 | -- | -- |
| Hillview | 8 | 13-14 | 66 | 63 | 35 | 31 |
| Sequoia | 10 | 15-16 | 50 | 45 | 18 | 30 |
| Foothill | 13-18 | 15-52 | 109 | 103 | 63 | 45 |
| **Total** | **6-18** | **11-52** | **764** | **687** | **216** | **241** |

Students played Posterlet (M=15 min, SD=0.3 min) individually, followed by an individual posttest (M=4 min, SD=0.1 min). The assessment was carried out in the classroom by middle-school and high-school students and at a location of their choice by college students. Not all students completed the posttest. Therefore, some analyses could not use the full sample. Similarly, 36 middle-school students from NYC and Chicago did not indicate their grade level, being excluded from grade-level analyses.

## 4.2. Measures

Learning Choices:
*Negative Feedback*: measures the number of times students choose negative feedback ("I don't like…" choice) out of a possible maximum of nine.
*Revision*: measures the number of times students choose to revise out of a possible maximum of three.
Learning Outcomes:
*Poster Quality*: measures the *in-game performance* on the posters. Posterlet evaluated each poster using 21 graphical design principles and produced a poster score per level (-21 to 21). The Poster Quality measure is the sum of the poster scores (-63 to 63).
*Posttest:* measures the *in-game learning* of the design principles from feedback. Students were shown a model poster and they answered two questions by selecting items from a checklist of design principles about what was good and bad, respectively, about that poster. The Posttest measure is the sum of the scores for the two multiple-choice posttest questions (-10 to 10). We scored the answers by assigning a point for each item checked correctly and subtracted a point for each item checked incorrectly. The correct answers to the two questions are opposites.
Demographics:
*Grade Group*: We use Grade Group as a proxy for age, because we received grade (not age) information for most of the datasets. We define this measure as a number ranging from 1 to 5, as follows: 1 = 6$^{th}$ grade (ages 11-12), 2 = 7$^{th}$ grade (ages 12-13), 3 = 8$^{th}$ grade (ages 13-14), 4 = high-school students (grades 9, 10, and 12; ages 14-18), and 5 = community college students (grades 13 to 18, ages 15-52). This reflects our grouping by educational stages (middle school to college).

## 4.3. Results

In the previous section, we outlined our three major classes of measures: (1) learning choices: frequency of choosing negative feedback and frequency of choosing to revise, (2) learning outcomes: poster performance and learning of the graphical design principles, and (3) demographics: group grade, which constitutes our substitute for age. Note that Poster Quality improved across levels of the game [Level 1 = 8.76 (SD=5.61), Level 2 = 9.87 (SD=5.26), Level 3 = 10.22 (SD=5.22); $F(2,762)=32.47$, $p<.001$], so we consider poster performance as a measure of learning.

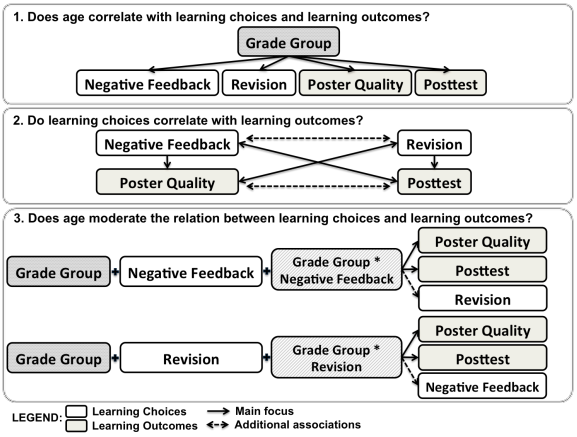


**Figure 2. Measures and analysis strategy**

Our analysis strategy (Figure 2) explores the relations between (1) age and learning choices and learning outcomes, (2) learning choices and learning outcomes, and (3) learning choices and learning outcomes, including the age by choice interaction.

First, we investigated whether age predicted learning choices and learning outcomes (Panel 1 of Figure 2). We correlated Grade Group with learning choices (Negative Feedback or Revision) and learning outcomes (Poster Quality or Posttest). Table 2 shows the results of all correlations corresponding to the arrows in Panel 1. Figure 3 shows the mean learning choices and outcomes per Grade Group, along with the sample size of each Grade Group. Grade Group correlates with both learning outcome measures, it does not correlate with Negative Feedback, and it inversely and modestly correlates with Revision (older students revised less). While both learning measures increase with age, the choice to seek negative feedback is not influenced by age and the choice to revise decreases modestly with age.

**Table 2. Correlations of Grade Group with learning choices and outcomes**

| Measure | Negative Feedback N=764 | Revision N=764 | Poster Quality N=764 | Posttest N=687 |
|---|---|---|---|---|
| GR | .03 | -.08* | .10** | .13** |

Note: $^{**}p<.01$, $^{*}p<.05$, GR: Grade Group.

Second, we investigated whether learning choices correlated with learning outcomes (Panel 2 of Figure 2). We correlated each choice (Negative Feedback and Revision) with each learning outcome (Poster Quality and Posttest) across the entire dataset. These analyses correspond to the solid arrows in Panel 2. Table 3 shows the correlations of learning choices with outcomes.

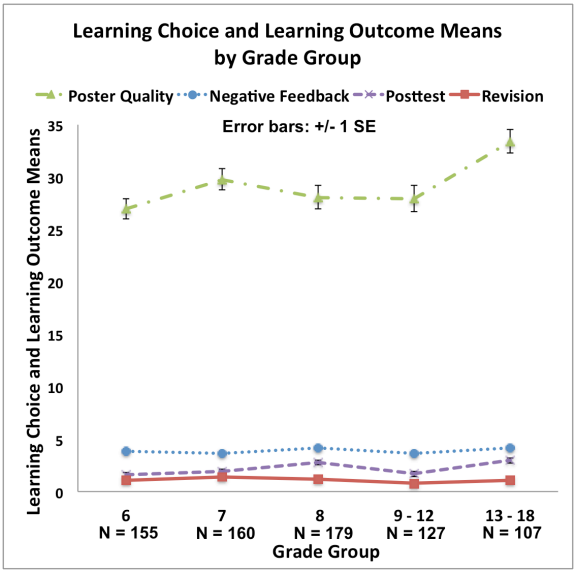


**Figure 3. Means of the learning choice and learning outcome measures by Grade Group**

We also included the correlations between learning choices and between learning outcomes (dashed arrows in Panel 2). Choices correlate with each other and with both learning outcomes. Additionally, learning outcomes correlate with each other. Thus, the choices to seek negative feedback and to revise are effective choices for learning.

**Table 3. Correlations between choices, learning outcomes, and Grade Group**

| Measure | NF | Rev N=764 | PQ N=764 | PT N=687 |
|---|---|---|---|---|
| NF | -- | .48** | .27** | .19** |
| Rev | | -- | .34** | .16** |
| PQ | | | -- | .34** |

Note: $^{**}p<.01$, NF: Negative Feedback, Rev: Revision, PQ: Poster Quality, PT: Posttest.

Third, we investigated the effect of age on the relation between learning choices and outcomes (Panel 3 of Figure 2). We were interested whether the relation between seeking negative feedback and learning differed by age. We regressed Grade Group, choice, and the interaction term Grade Group by choice on learning outcomes to explore the simultaneous effect of age, choice, and age by choice on learning outcomes (solid arrows in Panel 3). We also regressed Grade Group, Choice, and the interaction term Grade Group by Choice on the other choice. Prediction models are significant, but there are no interactions (dashed arrows in Panel 3). The value of negative feedback and revision for learning does not differ by age. Tables 4 and 5 show the percentage of the variance in the dependent variable accounted for by the overall model (Adj. $R^2$) and the $t\text{-value}_{GR*c}$ of the age by choice interaction term, where GR is the Grade Group and c is the choice.

**Table 4. Effect of Grade Group, Negative Feedback, and Grade Group by Negative Feedback on outcomes and Revision**

| Regression Models | | | |
|---|---|---|---|
| IV | Grade Group + Negative Feedback + Grade Group * Negative Feedback | | |
| DV | Poster Quality N=724 | Posttest N=654 | Revision N=724 |
| **Model Fit** Adj. $R^2$ | $.08^{***}$ | $.05^{***}$ | $.26^{***}$ |
| **Interaction** $t\text{-value}_{GR*NF}$ | -.08 | -.83 | .47 |

Note: $^{***}p<.001$, $t\text{-value}_{GR*NF}$: $t$-value of the interaction term Grade Group * Negative Feedback.

**Table 5. Effect of Grade Group, Revision, and Grade Group by Revision on outcomes and Negative Feedback**

| Regression Models | | | |
|---|---|---|---|
| IV | Grade Group + Revision + Grade Group * Revision | | |
| DV | Poster Quality N=724 | Posttest N=654 | Neg. Fbk. N=724 |
| **Model Fit** Adj. $R^2$ | $.14^{***}$ | $.05^{***}$ | $.26^{***}$ |
| **Interaction** $t\text{-value}_{GR*Rev}$ | 1.20 | 1.54 | .08 |

Note: $^{***}p<.001$, $t\text{-value}_{GR*Rev}$: $t$-value of the interaction term Grade Group * Revision.

## 5. Discussion and Future Work

We investigated students' choices to seek negative feedback and to revise depending on their age. Seeking negative feedback correlated with students' learning during the assessment. Revising correlated with learning outcomes and seeking negative feedback. This is a demonstration that *choosing* negative feedback yields better learning to the same degree for all ages investigated.

**Age vs. Learning Choices and Outcomes**: Older students learn more, but their choices do not change substantially with age. The choice to seek negative feedback is age-independent and older students revise slightly less than younger students. Older learners outperform younger learners on poster performance and learning of the graphical principles. Students of different ages exhibit the same patterns of choosing to seek negative feedback. Even if there are individual differences in our sample, we did not find differences in choosing negative feedback across age. Students from different cities and different schools with a wide range of income (e.g., individual parental income in the school neighborhood varied from $57,717 to $204,250) show the same pattern of choosing negative feedback, a remarkable result for education.

**Learning Choices vs. Learning Outcomes**: Both learning choices correlate with learning outcomes and strongly with each other. Additionally, both learning outcomes strongly correlate, providing internal, convergent validity for our measures. Although choosing to seek negative feedback and to revise strongly correlate, we do not know whether negative feedback caused revisions or whether those who seek negative feedback are more likely to revise. We will conduct a causal study to determine this. Choosing negative feedback correlates with in-game and posttest learning measures, possibly due to differential processing between positive and negative feedback, psychological factors, or the element of surprise stemming from different expectations when reading negative feedback. Positive feedback may highlight design principles that students already know and hence use well, whereas negative feedback may bring forward new knowledge related to design. More research is necessary to elucidate this.

Revising after feedback also correlates with the learning measures, though to a weaker degree on the posttest learning. Revision improves performance on the poster designs, presumably because by revising, one may fix any potential issues and avoid repeating them. Revision improves, albeit to a lesser extent, learning of the design principles, perhaps because by revising, one may revisit some principles and recognize them readily when they are presented on the posttest. Some of the principles on the posttest were never presented as feedback, since students had limited feedback opportunities and unique poster designs. Thus, revision is good for improving performance, but not necessarily to the same extent for learning. Merely performing the behavior of revising does not guarantee learning, but it may improve performance. We will pursue the issue of the effect of revision on performance and learning.

**Age, Learning Choice, and the Interaction between Age and Learning Choice vs. Outcomes**: The relation between learning choices and learning outcomes does not vary by age. The relation between learning choices does not vary by age and neither does the relation between learning outcomes. Although seeking negative feedback is a good predictor and age is a modest predictor of revising, there is no interaction between age and seeking negative feedback in predicting revision. Therefore, students choosing negative feedback are more likely to revise, but age is not moderating the relation between seeking negative feedback and revising. Students choosing negative feedback are more likely to perform better on the poster tasks, but age does not moderate the relation between seeking negative feedback and poster performance. Similarly, students choosing to revise are more likely to perform better on the posters, but age does not moderate the relation between revising and poster performance. We found similar results for the posttest. Although research [16] shows differential learning from positive and negative feedback across development when *receiving* feedback (learning from negative feedback increases with age), we did not find age differences in learning when *choosing* negative feedback. This has great implications for education, because it paves the way to generalized learning approaches.

**Limitations**. Our research was not causal. The modest differences in revision across different ages could be a function of students' different cities, parental income, school curricula, teachers, climate, or other factors. It is also possible that there is a latent variable (e.g., persistence) that drives the correlations between feedback, revision, and posttest behavior. Moreover, it would have been preferable to have a broader set of measures that would enable us to demonstrate divergent validity, because we are measuring something that is not captured by other common assessments. It might be useful to know whether seeking negative feedback shows a different pattern of correlation with learning outcomes than other relevant predictors (e.g., self-efficacy, fixed mindset). Now that we developed an assessment to measure choices and demonstrated that choices are related to learning in the game, it will be possible for researchers to determine why students of different ages seek negative feedback or revise, and for educators to evaluate whether a curriculum prepares students to make such independent learning choices.

We did not receive academic achievement scores for all students, but we intend to examine the relations between learning choices and both in-game and in-school learning outcomes to provide external validation of our assessment. In preliminary research, we found that both choices correlated with standardized mathematics and reading scores in two states (NY, Illinois) for $6^{th}$-$9^{th}$ graders.

Finally, we intend to examine the psychological differences in ego threat between making free choices (e.g., *choosing* to seek negative feedback) and being assigned choice outcomes (e.g., *receiving* negative feedback without requesting it) as they relate to learning and age differences. For example, letting patients choose their level of pain medication led to lower doses than when the doses were prescribed by medical staff [17]. Similarly, choosing negative feedback may diffuse ego threat, while being assigned negative feedback by the game characters may lead to less learning than enabling students to choose negative feedback. This issue is relevant to many instructional technologies.

We end with a pressing question that should be addressed by all assessment efforts. Given a measure of learning, do we also have a way to improve outcomes by this measure? We do not have an answer to the question of the best way to help students learn to choose negative feedback and to revise. However, now that we have a measure of these choices, it should be possible to investigate

how to inculcate an attitude that embraces negative feedback as a chance to learn rather than a reflection of one's personal worth. Meanwhile, our research provides a strong promise for engaging in further study of the choice to seek negative feedback. Assessing choices provides a new approach for evaluating process skills elusive to more traditional testing, but of great interest to many educators [7]. Instead of a source of motivation and self-selection, we view choice as an important outcome of learning.

## 6. Conclusions

This paper provides a first-of-kind demonstration that choosing negative feedback predicts better learning regardless of students' age. It presents empirical evidence that the choice-based assessment environment, Posterlet, measures two important learning choices. Surprisingly, the frequency of seeking negative feedback did not vary by age. Although learning outcomes improve with age, the relation between learning choices and outcomes does not vary with age. Finally, the relation between learning choices is also age-independent. The results of our novel examination of feedback and revision choices suggest that we should add these choices to the repertoire of independent learners of any ages, who will likely make such choices beyond school. In our game, students exercise free choices of feedback or revision, having the same experiences and learning opportunities regardless of their choices. By capturing choices, we approach our goal of measuring students' propensity for independent learning. We aim to build a suite of choice-based assessments measuring whether educational experiences foster independent learners who can make good learning choices.

## 7. Acknowledgments

This work was supported by the NSF under Grant EHR-1228831 and the MacArthur Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the granting agencies. We thank Neil Levine for his artwork and list of design principles, as well as the students and teachers who participated.

## 8. References

[1] E. Yackel and P. Cobb, "Sociomathematical norms, argumentation, and autonomy in mathematics". *Journal for Research in Mathematics Education*, 1996, pp. 458-477.

[2] L.S. Vygotsky, *The collected works of LS Vygotsky: Problems of the theory and history of psychology*, *3*, 1934.

[3] J.D. Bransford and D.L. Schwartz, "Rethinking transfer: A simple proposal with multiple implications" *Review of research in education*, 1999, pp. 61-100.

[4] K. Koedinger, J. Anderson, W. Hadley, and M. Mark, "Intelligent tutoring goes to school in the big city" *International Journal of AI in Education*, *8*, 1997, pp. 30-43.

[5] U.C. Klehe and N. Anderson, "Working hard and working smart: motivation and ability during typical and maximum performance". *Journal of Applied Psychology*, *92*(4), 2007, pp. 978.

[6] M. Cutumisu, D.B. Chin, and D.L. Schwartz, "A game-based assessment of students' choices to seek feedback and to revise", *Proc. 11th Intl. Conference on Cognition and Exploratory Learning in Digital Age*, Portugal, 2014, 8 ms.

[7] D.L. Schwartz and D. Arena, *Measuring what matters most: Choice-based assessments for the digital age*, 2013.

[8] D.L. Schwartz, R. Lindgren, and S. Lewis, "Constructivism in an age of non-constructivist assessments". *Constructivist Instruction*, 2009, pp. 34-61.

[9] D.L. Schwartz and D. Arena, "Choice-based assessments for the digital age." *MacArthur 21st Century Learning and Assessment Project*, 2009.

[10] S.S. Iyengar and M.R. Lepper, "Rethinking the value of choice: a cultural perspective on intrinsic motivation". *Journal of Personality and Social Psychology, 76*(3), 1999, pp. 349.

[11] D.L. Schwartz and J.D. Bransford, "A time for telling". *Cognition and Instruction*, *16*, 1998, pp. 475–522.

[12] A. Kluger and A. DeNisi, "Feedback interventions: toward the understanding of a double-edged sword" *Current Directions in Psychological Science*, *7*, 1998, pp. 67-72.

[13] J. Hattie and H. Timperley, "The power of feedback", *Review of Educational Research*, *77*(1), 2007, pp. 81-112.

[14] I. Roll, V. Aleven, B. McLaren, and K. Koedinger, "Improving students' help-seeking skills using metacognitive feedback in an intelligent tutoring system". *Learning and Instruction*, *21*(2), 2011, pp. 267-280.

[15] C. Moutsiana, N. Garrett, R.C. Clarke, R.B. Lotto, S.J. Blakemore, and T. Sharot, "Human development of the ability to learn from bad news". *Proceedings of the National Academy of Sciences*, *110*(41), 2013, 16396-16401.

[16] S. Peters, B. Braams, M. Raijmakers, P. Koolschijn, and E. Crone, "The Neural Coding of Feedback Learning across Child and Adolescent Development", *Journal of Cognitive Neuroscience 26*(8), 2014, pp. 1705-1720.

[17] M. Haydon, D. Larson, E. Reed, V. Shrivastava, C. Preslicka, and M. Nageotte, "Obstetric outcomes and maternal satisfaction in nulliparous women using patient-controlled epidural analgesia". *American Journal of Obstetrics and Gynecology*, *205*(3), 2011, pp. 271-e1.