Choosing versus Receiving Feedback: The Impact of Feedback Valence on Learning in an Assessment Game

Maria Cutumisu
University of Alberta
6-102 Education North, Edmonton, AB T6G 2G5
(780) 492 5211; cutumisu@ualberta.ca

Daniel L. Schwartz
Stanford University
485 Lasuen Mall, Stanford, CA 94305
(650) 725 5480; danls@stanford.edu

ABSTRACT
Studies examining feedback in educational settings have largely focused on feedback that is received, rather than chosen, by students. This study investigates whether adult participants learn more from choosing rather than receiving feedback from virtual characters in a digital poster design task. We employed a yoked study design and two versions of an online game-based assessment, Posterlet, to compare the learning outcomes of N=264 Mechanical Turk adults in two conditions: when they *chose* the feedback valence versus when they *received* the same feedback valence and order. In Posterlet, players design posters and learn graphic design principles from feedback. We found that the more the participants chose critical feedback, the more time they spent designing posters, but there were no differences in learning, revision, and time spent designing posters between conditions. In each condition, critical feedback correlated with performance and revision, suggesting that feedback valence is important for performance, regardless of being a choice.

Keywords
feedback valence, choice, assessment, game, learning

1. INTRODUCTION
A central goal of education is to prepare independent learners [16]. Previously, we operationalized this goal by a) identifying promising behaviors for autonomous learning that would reveal how students learned and b) creating novel choice-based digital assessment games that measured these behaviors. For instance, we measured students' choices to seek critical feedback and to revise, and we found that students who were more willing to seek critical feedback also learned more [4]. We examine learning choices (e.g., seeking social feedback), because such learning strategies can support ongoing learning, adapting to new challenges, and, ultimately, learning *how* to learn. These types of design thinking competencies, together with collaboration, persistence, and creativity, are crucial for 21st-century challenges, yet they are not formally assessed in schools [1, 21]. There are two main reasons why we need to measure learning behaviors. First, learning behaviors or attitudes enable learners to solve problems even when they do not have the domain knowledge skills to do so (e.g., collaborate with a partner from a different discipline). Second, current self-assessment techniques are not gender neutral: even though women and men scored similarly on a science exam (they had similar skills), women underestimated while men overestimated their performance (their attitudes did not match their skills; [7]). Such self-regulated learning behaviors [10] are worth investigating because revised self-assessment interventions may increase female representation in science, technology, engineering, and mathematics and could help create gender-inclusive 21st-century learning and assessment environments.

We previously examined the feedback valence (i.e., critical versus confirmatory) and its impact on performance and learning. In this study we examine for the first time the effect of feedback agency (i.e., choosing versus receiving). Our objective is to investigate the effect of choosing versus receiving feedback on learning, by comparing learning outcomes between participants who choose feedback and those who receive the same amount, valence, and order of feedback. We outline related work and theoretical perspectives that guide our research. Then, we describe our assessment environment, Posterlet, an online game designed to collect and assess participants' feedback and revision choices. We also created and presented a modified version of this game to accommodate the situation in which feedback is assigned to the learner in a principled way that mirrors the feedback chosen in the original Posterlet version. We then present evidence of the impact of choosing *versus* receiving feedback on learning outcomes, as well as theoretical and practical implications of this research.

We examine the impact of feedback choice and valence on learning by posing the following research questions:

1) Does critical feedback correlate with learning outcomes?
2) Are there learning outcome differences between choosing and receiving feedback?
3) Are there design duration differences between choosing and receiving feedback?
4) Are there gender differences on the measures by condition?

2. RELATED WORK
We distinguish several themes in the literature related to the theoretical perspectives that guide this research.

Choice-based Assessments. Traditional assessments measure learners' knowledge at the end of instruction, focusing on knowledge accuracy but providing little information about learners' readiness to learn new things. Vygotsky highlighted the importance of measuring learning processes [23], rather than only learning outcomes, to achieve deeper insights into students' potential to learn on their own. Schwartz and Bransford advocated *preparation for future learning* (PFL) assessments [19], which create learning opportunities during the assessment. Our research draws from work on *constructivist assessments* [20] and *choice-based assessments* [18]. Both these assessments build upon PFL assessments and measure not only learners' knowledge outcomes but also their learning processes (e.g., choices about what, when, and how to learn). For example, Posterlet [4], an online game that collects players' choices to seek critical feedback and to revise while they design posters, constitutes an instance of a choice-based assessment. The design of Posterlet is guided by the three core principles of choice-based assessments: *typical performance* (assessments need to capture every-day learning behaviors, not

test performance), *PFL* (assessments need to offer learning opportunities with measurable outcomes; [2]), and *choice* (assessments need to collect free learning choices that do not hinder the learners' ability to complete the assessments). Specifically, Posterlet provides players with a 10-15 minute fun game experience, with a chance to learn graphic design principles and to safely explore choices to seek critical feedback and revise, before applying them in more high-stakes situations. Concomitantly, Posterlet provides researchers with a way to track players' behaviors and learning outcomes to infer how prepared players are to learn on their own in new learning situations.

Confirmatory versus Critical Feedback. In educational contexts, feedback is defined as information related to a person's performance or understanding [11] and it is predominantly assigned by a teacher or a computer rather than chosen by the learner. There are some exceptions, but they pertain to help seeking [17] rather than specifically to feedback seeking. Here, we are mainly interested to investigate whether being given a choice about how to learn (i.e., choosing versus receiving feedback) has any impact on learning outcomes and other learning behaviors. In addition to feedback choice, the feedback literature provides some indication of the importance of feedback valence. For instance, critical feedback yields mixed results for performance [13], but studies of organizations show that most new ideas need critical constructive feedback to become successful [15]. A first challenge is that feedback is often absent from ideation environments. A second challenge is that critical feedback is even more elusive in such environments and it runs the risk of ego threat that causes people to reject instead of heed the feedback [11]. This suggests that attitudes towards seeking critical feedback are worth exploring. However, there is no evidence that the choice of critical feedback is as important as simply assigning critical feedback to the learner. Thus, we designed a variation of Posterlet and we employed a reduced-length game version for comparison to address this issue.

Choosing versus Receiving Feedback. Traditionally, most studies focused on supervised feedback, where the teacher assigned feedback to the student. However, in many situations, people need to actively seek feedback. Little is known about the implications of students' feedback choices on their learning or about variables that influence students' feedback choices, but researchers acknowledge the importance of the mechanisms underlying feedback for learning. For instance, Zimmerman [24] included "responsiveness to self-oriented feedback" among three critical features of students' self-regulated learning strategies. The effect of actively choosing rather than passively receiving critical feedback for learning raises interesting psychological questions. For example, patients who had control over their level of pain medication chose lower doses than those prescribed by medical staff [12]. Similarly, having a choice over critical feedback may act as a buffer against ego threat. Further, if learners are assigned critical feedback, would that lead to less learning than if they chose it? Consumer research provides corroborating evidence directly relevant to our prior research regarding the choice between confirmatory and critical feedback. Researchers found that novices sought confirmatory feedback more often, whereas experts sought critical feedback more often [9]. However, in contrast to our research, they did not measure learning outcomes.

3. POSTERLET

We employed two versions of the Posterlet game [4] to carry out our experiment. Participants playing the games assumed the identity of a school committee member in charge with designing a poster for each of the two booths advertising events for the school's Fun Fair. The effectiveness of each designed poster (i.e., the number of visitors attracted by the booth) is quantified by the number of tickets sold, which is displayed when the poster is submitted. Posterlet also measures the number of times critical feedback is chosen or received, depending on condition, and the player's choices to revise posters across the game. After designing each poster, the player chooses three virtual characters out of a focus group to find out what they think about the poster. In the Choose condition, the player clicks on one box ("I like" or "I don't like") above each character. For example, in Figure 1, a participant in the Choose condition has first selected critical feedback from the lion and then confirmatory feedback from the elephant, but no feedback from the panda yet.

Figure 1. In the Choose condition, the player has first chosen critical feedback from the lion, confirmatory feedback from the elephant, and no feedback from the panda yet.

In the Receive condition, the player clicks on the "Click for feedback" box to reveal a feedback valence assigned by the game. For example, in Figure 2, a Receive condition participant has first clicked on the elephant's "Click for feedback" box (revealing critical feedback), then on the ostrich's "Click for feedback" box (revealing confirmatory feedback). The amount of critical feedback chosen or assigned (depending on the condition) is Posterlet's first key measure. After reading the feedback, the player has a choice to revise or submit the poster. The number of revised posters is Posterlet's second key measure. The game's feedback system generates feedback by analyzing each poster against 21 graphic design principles provided by a graphic artist and organized into three broad categories: information (e.g., the poster should include the date of the event), readability (e.g., the color contrast between the text and the background should be high), and space use (e.g., the space used by images needs to be within 30% and 70% of the poster's surface).

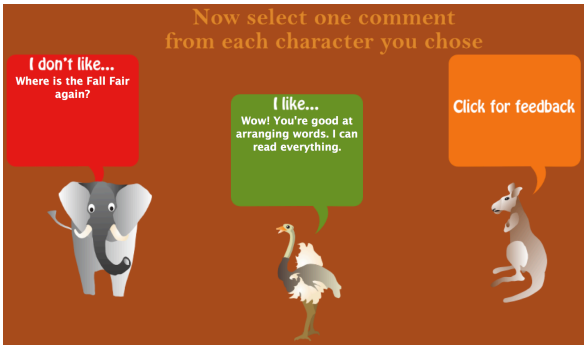

Figure 2. In the Receive condition, the player has first clicked on the elephant and received critical feedback, then on the ostrich and received confirmatory feedback.

It computes each poster's quality (i.e., the number of tickets sold) and it includes a priority scheme to ensure a balanced representation of these categories in the feedback. The critical and confirmatory feedback phrases are equivalent in length and informational content. For example, if a player omits the day of the fair, the critical feedback is: "You need to tell them what day the fair is." Otherwise, the confirmatory feedback is: "It's good you told them what day the fair is.", as shown in Figure 2.

4. METHOD

4.1 Participants, Procedures, Data Sources, and Experimental Overview

Participants (see Table 1) are N=264 Mechanical Turk adults randomly assigned to either the Choose or the Receive condition. Choose condition participants played a version of Posterlet that collected their feedback choices, while Receive condition participants played a modified Posterlet version that did not offer a feedback choice. In a one-to-one yoked experimental design, each participant in the Receive condition was assigned the feedback valence, number, and order of the feedback chosen by a matched Choose condition participant. Participants played a two-poster version of the Posterlet game individually, corresponding to their assigned condition, with a five-minute time limit on each poster or revision. Then, they completed an individual online posttest. The participants in the Choose condition were presented with a choice regarding the valence of their feedback. For instance, Figure 1 illustrates the feedback choices of a participant in the Choose condition: the participant chose a critical feedback from the lion and then a confirmatory feedback from the elephant. The Receive Condition participants were assigned the feedback valence of paired Choose condition participants, in the same order in which feedback was chosen by those paired participants. The game also collected participants' revision choices and computed the participants' poster performance (i.e., the quality of all their posters). Posterlet tracked the amount of critical feedback out of a maximum of 6 (3 feedback opportunities x 2 posters), as well as the amount of revisions out of a maximum of 2 (1 revision opportunity x 2 posters). A separate posttest measured the graphic design principles learned by participants in both conditions.

Table 1. Number of participants in each condition by gender

Cond.	Gender		Age Range	M_{age} (SD_{age})
	F	M		
Choose	54	78	19-69	32.26 (9.53)
Receive	61	71	19-63	33.30 (10.40)
Total	**115**	**149**	**19-69**	**32.78 (9.96)**

For instance, Figure 2 illustrates the feedback selection of a participant in the Receive condition: the participant was first assigned critical feedback and then confirmatory feedback, just like the participant in the Choose condition illustrated in Figure 1.

In the Choose condition, participants played Posterlet for an average of M=7 minutes (SD=3.11) and then completed the posttest for an average of M=6 minutes (SD=2.24). In the Receive condition, participants played Posterlet for an average of M=7 minutes (SD=2.91) and then completed the posttest for an average of M=7 minutes (SD=2.54). This study is correlational and experimental, aiming to determine whether having a choice about one's feedback valence aids in learning or in choosing to revise one's work. It compares adults who exercised a choice regarding the valence of their feedback (Choice condition) to adults who were assigned their feedback valence (Receive condition).

4.2 Dependent Measures

4.2.1 Feedback Valence and Revision Choices

Critical Feedback measures the number of "I don't like" boxes chosen or received by the player across the game (0-6). **Confirmatory Feedback** measures the number of "I like" boxes chosen or received, equivalent to 6 minus *Critical Feedback* (0-6), since there are six total feedback choices across the game. **Revision** measures the number of posters a player revised (0-2).

4.2.2 Design Duration

We measured the time a participant spent designing each poster, from the moment a booth theme was clicked to the moment the "Test" button was pressed.

4.2.3 Learning Outcomes

Poster Quality measures the poster performance, summing the poster quality across posters. The quality of each poster is the sum of the scores for each of the 21 features: 1 if a feature is always used correctly, 0 if a feature is not on the poster, and -1 if a feature is used incorrectly. Thus, the score of any individual poster ranges from -21 to 21, while *Poster Quality* from -42 to 42.

A posttest assessed learning of the graphic principles. The overall *Posttest* score represents the sum of the normalized scores of the *Recognition* and *Principle Selection* measures.

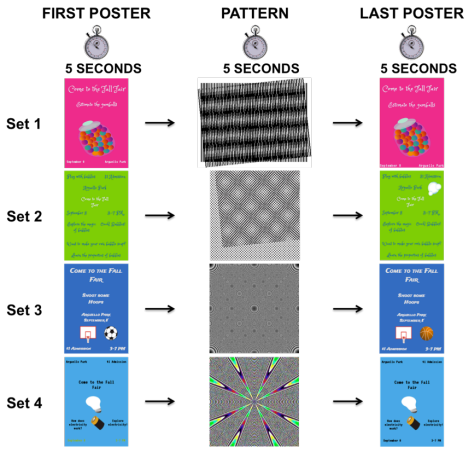

Figure 3. The *Recognition* posttest questions.

Recognition comprised four sets of posters (Figure 3). For each set, participants' task was to judge whether the quality of the second poster was the same/better/worse compared to the quality of the first poster and to provide a brief written explanation for their decision. A distractor image was inserted between the two posters to ensure that memory was not playing a role [22]. Participants were guided through a mini-tutorial and a trial poster comparison, in which pictures succeeded automatically on a five-second timer. Each correct answer is scored with one point, while each incorrect answer is scored with zero points. This measure sums up only the correct answers, thus ranging from zero to four. **Principle Selection** comprised two 10-item design principle checklist questions (Figure 4). A point was awarded/subtracted for each correct/incorrect answer and scores were summed up.

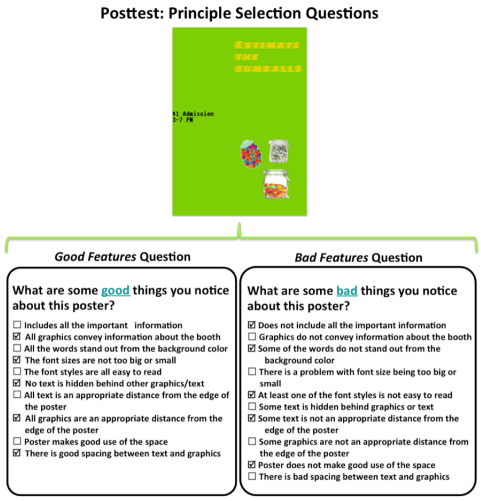

Figure 4. The Principle Selection posttest questions.

5. RESULTS

5.1 Does critical feedback correlate with learning outcomes?

We examined poster performance and design principle learning. Table 2 and Table 3 show the zero-order Pearson correlations by condition. Critical Feedback and Revision correlated with Poster Quality and strongly with each other. We consider Poster Quality a learning measure, due to participants' improvement across the game [*Choose*: $round_1$=10.64 (SD=5.0), $round_2$=11.76 (SD=4.5), Wilks' Lambda=.92, partial eta squared=.08, F(1,131)=11.67, p<.01; *Receive*: $round_1$=10.68 (SD=6.0), $round_2$=11.67 (SD=5.4), Wilks' Lambda=.96, partial eta squared=.04, F(1,131)=5.89, p<.05]. Revision correlated with Posttest and Design Duration. Poster Quality correlated with Posttest, supporting the learning measures' internal validity. In the Choose condition, Critical Feedback correlated with Design Duration.

Table 2: Correlations between critical feedback, revision, and learning outcomes for the *Choose* condition

Measures (N=132)	Revision	Poster Quality	Posttest	Design Duration
Critical Fb.	**.62****	**.25****	.08	**.32****
Revision	--	**.23****	**.21***	**.39****
PosterQuality		--	**.27****	**.39****

** p < .01, * p < .05

Table 3: Correlations between critical feedback, revision, and learning outcomes for the *Receive* condition

Measures (N=132)	Revision	Poster Quality	Posttest	Design Duration
Critical Fb.	**.58****	**.18***	.13	.16
Revision	--	**.24****	**.21***	**.36****
PosterQuality		--	**.21***	**.38****

** p < .01, * p < .05

We entered Critical Feedback and Revision in regressions to determine if they were independent predictors of the learning outcomes. In the Choose condition, for Poster Quality, the model was significant [F(2,129)=5.10, p<.01, R^2=.07, Adjusted R^2=.06], but Critical Feedback [t(129)=1.6, p=.11] and Revision [t(129)=1.6, p=.25] were not predictors. For Posttest, the model was significant [F(2,129)=3.33, p=.04, R^2=.05, Adjusted R^2=.03], Revision was a predictor: t(129)=2.38, p=.02, but Critical Feedback: t(129)=-.71, p=.48 was not. In the Receive condition, for Poster Quality, the model was significant [F(2,129)=4.23, p=.02, R^2=.06, Adjusted R^2=.05], Revision was a marginally significant predictor: t(129)=1.99, p<.05, but Critical Feedback: t(129)=.58, p=.56 was not. The Posttest model was not significant.

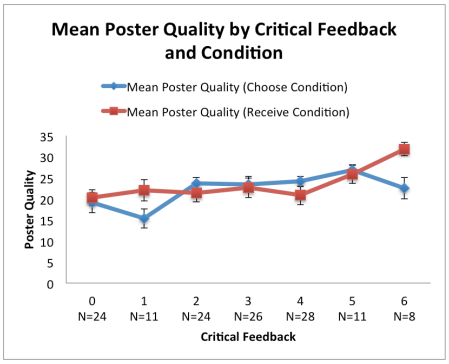

Figure 5. Poster Quality by Critical Feedback and condition.

5.2 Are there learning outcome differences between choosing and receiving feedback?

T-test analyses revealed no differences in Poster Quality [M_{Choose}=22.39 (SD=8.71), $M_{Receive}$=22.36 (SD=10.4), t(262)=.03, p=.97], Posttest [M_{Choose}=.10 (SD=1.53), $M_{Receive}$=.04 (SD=1.45), t(262)=.32, p=.75], and Revision [M_{Choose}=.80 (SD=.87), $M_{Receive}$=.93 (SD=.82), t(262)=-1.24, p=.22] between conditions. Figure 5, Figure 6, and Figure 7 plot our measures across the game as a function of critical feedback (from 0 to 6) by condition. Error bars represent one standard error. The x-axis shows the range of critical feedback and the number of participants for each amount of critical feedback (e.g., N=26 participants chose/received 3 pieces of critical feedback across all posters). Regressions of *critical feedback*, *condition*, and *critical feedback by condition* on learning and revision revealed no interactions of critical feedback and condition with our measures.

5.3 Are there design duration differences between choosing and receiving feedback?

A t-test analysis revealed no differences in Design Duration (time in seconds spent designing posters) between conditions [M_{Choose}=401.30 (SD=186.39) and $M_{Receive}$=394.44 (SD=174.94), t(262)=.31, p=.76]. Figure 8 plots participants' poster design time across the game as a function of critical feedback (from 0 to 6) by condition.

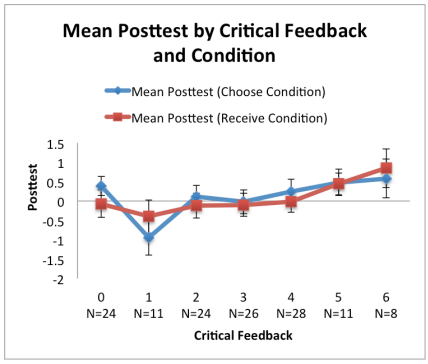

Figure 6. Posttest by Critical Feedback and condition.

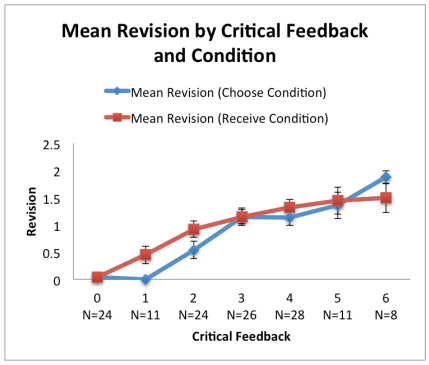

Figure 7. Revision by Critical Feedback and condition.

5.4 Are there any gender differences?

In the Receive condition, we found that females [M=433.28 (SD=176.84), $t(130)=2.41$, $p=.02$] spent more time designing posters than males [M=361.07 (SD=167.40)]. There were no gender differences by condition on any of the rest of the measures (Revision, Poster Quality, and Posttest).

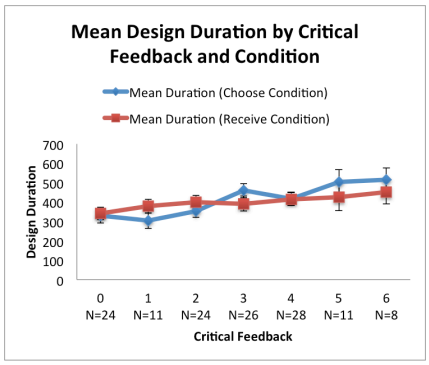

Figure 8. Design Duration by Critical Feedback and condition.

6. DISCUSSION

This is a first-of-kind examination of both the agency (choosing versus receiving) and the valence (critical versus confirmatory) of feedback and their impact on performance and learning. We found that, in each condition, the amount of critical feedback (either chosen or received) correlated with participants' performance on the poster design task. Consistent with our previous findings [3, 4], critical, rather than confirmatory, feedback seems beneficial for learning. Also, the choice to revise was beneficial for performance and learning outcomes and it strongly correlated with critical feedback (chosen or received). We found no differences between conditions in any of the measures outlined in this paper. These results held when we compared the measures by gender in each condition, although in the Receive condition, females spent more time designing posters than males. This indicates that these types of behavioral assessments of learning have the potential to be gender neutral. The next step would be to design more such dynamic assessments to evaluate other behaviors, such as self-assessment. Designing gender-neutral assessments that embed both skills and learning behaviors would bring us closer to determining the knowledge, skills, and delivery methods required to foster independent learners in the 21st century, as well as ways to ensure gender equality, especially when only 14.1% of North American computer science bachelor's degree graduates are female [25]. Our study points to critical, rather than confirmatory, feedback being beneficial for learning, regardless of being chosen or assigned. It also points to ways of designing assessments that measure learning behaviors equally regardless of gender. Finally, in the Choose condition, the more the participants chose critical feedback, the more time they spent designing posters. The relation between critical feedback and revision, as well as between critical feedback and poster quality, was stronger and more stable in the Choose condition, pointing to motivational factors of choosing versus receiving critical feedback for performance. More research is needed to elucidate this motivational aspect.

People's choices of critical feedback can be influenced by a wide range of factors. For instance, the perception of a trait as fixed may lead to avoidance of negative feedback [5]. Additionally, compared to a growth mindset (an incremental theory of intelligence - the belief that intelligence can be developed over time), a fixed mindset (an entity theory of intelligence – the belief that intelligence is fixed) was found to be associated with decreased attention to corrective feedback or errors [14]. However, the results of this study suggest that there is no underlying variable (e.g., desire to learn, self-confidence, growth mindset [6, 8], etc.) that drives the effect of critical feedback. People who choose critical feedback more often may exhibit one or more of these variables, yet, despite that, assigning the same amount of feedback leads to the same results as other factors that may causes them to choose critical feedback. Consequently, it seems that such factors (e.g., deep beliefs or personal attributions, such as "I am a learner") do not need to be changed to help people reap the benefits of constructive criticism. Learner beliefs do not mediate the benefits of receiving constructive criticism. One potential implication is the possibility to change people's beliefs about seeking critical feedback without having to change their broad beliefs about themselves as learners, which we also demonstrated in a separate study [3]: fairly straightforward instruction to seek social feedback (i.e., opinions of others) transferred to Posterlet and, consequently, students learned more.

Our study's limitations are associated with conducting Mechanical Turk experiments with a large population: (1) a maximum of five minutes allotted per poster, which may have hindered the discovery of some of the game's features (e.g., that the poster background color can be changed) and (2) a maximum of two game levels, which offered participants at most six pieces of feedback from which to learn graphic design principles, which may not have overlapped with the four principles included on the posttest (feedback content varied, depending on each participant's poster, but the posttest questions were the same for all participants). The latter is one possible explanation for the lack of correlation between critical feedback and posttest. Alternatively,

participants examined each poster for only five seconds and, if they missed one of the two posters in a set, they could not have accurately answered any of the questions about that set. Thus, we plan to compare this study's Choose condition data with data from the first two levels of previous three-level Posterlet game studies. That way, we may predict participant behaviors on the third game level, to potentially detect differences between conditions in our measures that are not apparent currently.

7. CONCLUSIONS

We modified a choice-based assessment game to measure learning when participants are offered a choice about the valence of their feedback and when they are assigned their feedback valence. The data enabled a novel examination of choosing versus receiving confirmatory *versus* critical feedback with regards to learning outcomes. We found that the more the participants chose critical feedback in the Choose condition, the more time they spent designing posters. There were no differences in learning outcomes (performance on the poster design task and learning of the graphic design principles), choice to revise, or time spent designing posters between participants who chose feedback and those who received the same amount, valence, and order of feedback. We plan a similar study with middle-school and college students to explore instruction and assessment implications. These studies could inform teachers to create environments in which students feel encouraged to engage more with critical feedback (proactively or reactively), even in open-ended tasks as digital poster design. The flexibility of such short assessments focused on specific choices (e.g., feedback seeking) enables the development and evaluation of a variety of instruction models. Concomitantly, researchers can design pedagogical interventions and learning environments that embed such assessments to empower all learners, regardless of gender, to be innovative, confident, and prepared for the challenges of the 21st century.

8. ACKNOWLEDGMENTS

We thank the Gordon and Betty Moore foundation and the NSF (Grant # 1228831), Jacob Haigh for assistance with the online setup, as well as the Mechanical Turk participants.

9. REFERENCES

[1] Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5–31.

[2] Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications, *Review of Research in Education, 24*, 61–100.

[3] Conlin, L., Chin, D. B., Blair, K. P., Cutumisu, M., & Schwartz, D. L. (2015). Guardian Angels of Our Better Nature: Finding Evidence of the Benefits of Design Thinking. In Proc. of *ASEE*, June 14-17, Seattle, WA, USA.

[4] Cutumisu, M., Blair, K. P., Chin, D. B., & Schwartz, D. L. (2015). Posterlet: A Game-Based Assessment of Children's Choices to Seek Feedback and to Revise. *Journal of Learning Analytics, 2*(1), 49-71.

[5] Dunning, D. (1995). Trait importance and modifiability as factors influencing self-assessment and self-enhancement motives. *Personality and Social Psychology Bulletin, 21*.

[6] Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*(2), 256-273.

[7] Ehrlinger, J., & Dunning, D. (2003). How chronic self-views influence (and potentially mislead) estimates of performance. *J of Personality & Social Psychology*, *84*(1), 5.

[8] Ehrlinger, J., Mitchum, A. L., & Dweck, C. S. (2016). Understanding overconfidence: Theories of intelligence, preferential attention, and distorted self-assessment. *Journal of Experimental Social Psychology*, *63*, 94-100.

[9] Finkelstein, S. R., & Fishbach, A. (2012). Tell me what I did wrong: experts seek and respond to negative feedback. *Journal of Consumer Research, 39*(1), 22–38.

[10] Garrison, D. R. (1997). Self-directed learning: Toward a comprehensive model. *Adult Ed. Quarterly*, *48*(1), 18–33.

[11] Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112.

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

[13] Kluger, A. N., & DeNisi, A. (1998). Feedback interventions: toward the understanding of a double-edged sword. *Current Directions in Psychological Science, 7*(3), 67–72.

[14] Mangels, J., Butterfield, B., Lamb, J., Good, C., & Dweck, C. (2006). Why do beliefs about intelligence influence learning success? A social cognitive neuroscience model. *Social Cognitive and Affective neuroscience*, *1*(2).

[15] March, J. (2008). *Explorations in organizations*. Stanford University Press.

[16] Piaget, J. (1964). Quoted by Eleanor Duckworth in "Piaget Rediscovered: A Report of the Conference on Cognitive Studies and Curriculum Development", *Cognitive Studies and Curriculum Development, New York*.

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

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

[19] Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, *16*, 475–522.

[20] Schwartz, D. L., Lindgren, R., & Lewis, S. (2009). Constructivism in an age of non-constructivist assessments. *Constructivist Instruction*, 34-61.

[21] Shute, V. J., & Ke, F. (2012). Games, learning, and assessment. In *Assessment in Game-Based Learning*, 43-58.

[22] Standing, L. (1973). Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology, 25*, 207-222.

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

[24] Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*(1), 3–17.

[25] Zweben, S., & Bizot, B. (2015). 2014 Taulbee Survey. *Computing Research News, May 2015, 27*(5), p. 20.