# THE IMPACT OF CRITICAL FEEDBACK CHOICE ON STUDENTS' REVISION, PERFORMANCE, LEARNING, AND MEMORY

Maria Cutumisu[a, *], Daniel L. Schwartz[b]

[a] Department of Educational Psychology, 6-102 Education North, University of Alberta, Edmonton, AB, T6G 2G5, Canada, cutumisu@ualberta.ca

[b] Stanford Graduate School of Education, 485 Lasuen Mall, Stanford, CA, 94305-3096, USA, danls@stanford.edu

[*] Corresponding author.

**ABSTRACT**

This article examines empirically the impact of students' critical feedback choices on their memory for feedback. It also examines the effect of choosing versus receiving feedback on learning outcomes. First, a correlational study was designed to collect the choices to seek critical feedback and to revise posters from a hundred and six Grade 8 middle-school students via Posterlet, a digital assessment game in which students design posters. Upon completing the game, students filled a post-test asking them to freely recall the feedback messages they encountered in Posterlet. Results show that, when they have a choice between critical and confirmatory feedback, students tend to remember critical feedback better than confirmatory feedback. Second, a yoked experiment was designed to compare the performance and learning of college students who chose and were assigned the same amount and order of critical feedback, respectively. Results show that 1) critical feedback and revision are positively associated with performance when students choose their feedback, while critical feedback is negatively associated with learning when students receive their feedback; 2) students who engaged with higher levels of critical feedback significantly outperformed the rest of the participants only if they had a choice over their feedback; and 3) students enjoy designing posters significantly more when they choose rather than receive their feedback. Ramifications for student learning are discussed. Future work will examine whether there are any differences in memory for feedback between students who choose and those who are assigned the same amount and order of critical feedback.

**KEYWORDS**

feedback, memory, assessment, choice, game, learning performance

# 1 INTRODUCTION

Feedback, defined as information regarding performance outcomes and learning processes (Hattie & Timperley, 2007), plays a major role in educational performance. It is an essential component of learning (Ackerman & Gross, 2010; Carless, Salter, Yang, & Lam, 2011; Evans, 2013; Hounsell, Slowey, & Watson, 2003; Jonsson, 2013; Mulliner & Tucker, 2017; Rietsche, Lehmann, Haas, & Söllner, 2017; Schwartz, Tsang, & Blair, 2016), the most powerful single influence on student achievement (Hattie & Timperley, 2007), and one of the most frequently applied psychological interventions (Attali, 2015; Kluger & DeNisi, 1998). For instance, a synthesis of over 800 meta-analyses related to student achievement revealed a 0.75 overall feedback effect size (Hattie, 2009), while a review of 12 meta-analyses of feedback revealed an average feedback effect size of 0.79 (Hattie, 2013). Increasingly, the psychology of giving and receiving high-quality feedback is becoming central to student attainment (Kulkarni, Bernstein, & Klemmer, 2015), persistence, retention, and completion (Kizilcec, Saltarelli, Reich, & Cohen, 2017), as well as satisfaction (Evans, 2013;

Wiliam, 2011) from K-12 to higher education (Hattie & Gan, 2011; Jonsson, 2013; Sargeant et al., 2011; Walsh, Nixon, Walker, & Doyle, 2015; Yorke, 2003).

Despite this, the specific mechanisms relating feedback to learning are still not well understood. Feedback research has yielded mixed results in the educational literature, with varying effect sizes of feedback on performance (Danielmeier & Ullsperger, 2011; Kluger & DeNisi, 1996; Kluger & DeNisi, 1998). In educational settings, feedback was generally found to improve performance (Hattie & Timperley, 2007), but studies have shown that feedback is not always helpful, making either no difference (Brummelman, Thomaes, de Castro, Overbeek, & Bushman, 2014; Burnett & Mandel, 2010; Harris, Brown, & Harnett, 2015; Hattie & Timperley, 2007), such as non-specific praise, or even hindering learning (Kluger & DeNisi, 1996; Kluger & DeNisi, 1998; Shute, 2008; Winstone, Nash, Parker, & Rowntree, 2017), such as disconfirming feedback. For instance, a meta-analysis examining 607 effect sizes and 23,663 observations found that feedback interventions improved performance on average, but that they decreased performance in a third of the feedback interventions examined (Kluger & DeNisi, 1996). Concomitantly, a more recent meta-analysis reviewing 58 effect sizes from 40 reports found that feedback effects varied widely with control for factors including the type of feedback or the type of instruction (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991).

There are many reasons for these mixed results, as there are many factors, both motivational and informational, that influence the effectiveness of feedback for performance. Kluger and DeNisi (1996) proposed a *Feedback Intervention Theory* (FIT) positing that the effectiveness of feedback interventions decreases as the learner's locus of attention moves up a hierarchy of levels of control, task learning, task motivation, and meta-tasks (including self-related) processes, closer to the self and away from the task. For example, task-directed feedback seems to be more helpful than person-directed feedback, such as praise or punishment unrelated to the task (Black & Wiliam, 1998; Hattie & Timperley, 2007). A research study tested the *regulatory focus theory* to explore the role of self-regulatory focus, which constitutes an orientation towards positive versus negative outcomes (Higgins, 1997), in the relationship between feedback and performance. Results revealed that promotion-focused individuals benefitted more from positive feedback in terms of performance and innovation, but that prevention-focused individuals' performance did not benefit from negative feedback (Sonnentag & Sparr, 2008). However, their data were based on feedback that was assigned and on self-reports, including peer questionnaires to assess individuals' performance and innovation.

Recent research findings point to other factors that influence feedback differentially, including feedback message construction and delivery of critical feedback (Gregory & Levy, 2015; Landis-Lewis, Brehaut, Hochheiser, Douglas, & Jacobson, 2015), credibility of the source (LaDonna, Hatala, Lingard, Voyer, & Watling, 2017), specificity and relevance (Roscoe, Wilson, Johnson, & Mayra, 2017), as well as feedback orientation that constitutes an individual's positive affect, interest, and engagement with feedback (Waller & Papi, 2017). Specifically, the role of feedback valence (i.e., confirmatory or critical) on performance is still a matter of debate (Fishbach, Eyal, & Finkelstein, 2010). Critical feedback seems to aid performance in some situations (Kluger & DeNisi, 1998), but individual factors (mindset, self-worth, etc.) may affect engagement with critical feedback. For example, feedback implying that performance is malleable, rather than due to an inherent ability, may protect learners against setbacks. Moreover, an incremental theory of intelligence promotes more openness towards critical feedback that leads individuals to a better understanding of their own abilities (Ehrlinger, Mitchum, & Dweck, 2016). In contrast, an entity theory promotes lessened attention to critical feedback (Mangels, Butterfield, Lamb, Good, & Dweck, 2006) and more preference for confirmatory feedback (Ehrlinger, Mitchum, & Dweck, 2016). Threats to self-worth that may hinder the engagement with critical feedback can be mitigated by self-affirmation techniques (Critcher, Dunning, & Armor, 2010).

The quality of the learners' engagement with feedback is believed to be a determinant factor of feedback effectiveness (Winstone, Nash, Parker, & Rowntree, 2017), yet few studies examine this aspect of feedback (Bounds et al., 2013). Furthermore, most research focuses on feedback being assigned to the learner. Moreover, becoming an autonomous learner requires a willingness and propensity for actively engaging with critical feedback (Williams & Ehrlinger, 2017). Although feedback seeking was examined in organizational settings (Anseel, Beatty, Shen, Lievens, & Sackett, 2015; Ashford, De Stobbeleir, & Nujella, 2016; De Stobbeleir, Ashford, & Buyens, 2011; Porath & Bateman, 2006), there is a paucity of research examining the effectiveness of feedback seeking in education (Evans, 2013). There are even fewer studies examining the impact of feedback on students' memory, especially when students choose the valence of their feedback.

In previous research, we investigated the impact of the choice of feedback valence on performance in educational settings. Results indicated that choosing critical feedback was positively associated with learning and performance (Cutumisu, Blair, Chin, & Schwartz, 2015). This suggests that both feedback agency and valence are worth exploring in relation with learning and performance. However, it is still not known whether it is the choice over feedback or the actual amount of critical feedback that impacts performance. Building on research that validated choices as predictors of learning (Cutumisu, Chin, & Schwartz, 2014; Cutumisu, Blair, Chin, & Schwartz, 2015), the present research aims to gain an insight into the mechanisms of feedback processing by focusing on feedback choices and memory for feedback to reveal important insights into the types of feedback students remember. It hypothesizes that students remember critical feedback better than confirmatory feedback when they have a choice between critical and confirmatory feedback. This manuscript makes the following contributions. First, instead of focusing on feedback that is assigned to the learner, it examines the mechanisms that unfold when students engage proactively with feedback by choosing between confirmatory (positive) and critical (negative) feedback. Second, it focuses for the first time on the lasting impact of choices between confirmatory and critical feedback on students' memory for feedback. Furthermore, many studies in the literature (Sargeant et al., 2011; Thurlings, Vermeulen, Bastiaens, & Stijnen, 2013) include small sample sizes (e.g., n = 13). In contrast, each of our studies sampled approximately a hundred students. Third, it also examines for the first time the effect of feedback agency (i.e., choosing versus receiving feedback) on college students' performance. Fourth, most results regarding students' learning behaviours as they engage with feedback are based on self-report measures representing the learners' views on their use of feedback collected via surveys and interviews (Mahfoodh, 2017; Mulliner & Tucker, 2017; Narciss, 2013; Sargeant et al., 2011). Authors of a recent comprehensive systematic review (Winstone, Nash, Parker, & Rowntree, 2017) found that, of the empirical studies they examined, only 19% measured actual feedback behaviours and 55% were based on surveys, with only 7% of the empirical studies using a psychometric approach. In contrast, our studies are based on behavioural data collected via an online assessment game. Our research is opportune, as it collects more objective measures and it explores feedback choices that may have a long-lasting effect on memory and learning. Finally, Winstone and colleagues (2017) also noted that most studies focused on university students. In contrast, this manuscript makes a contribution to the literature by assessing middle-school students' memory for feedback.

This topic has wide applicability. For example, in medical research, patients who had control over their level of pain medication chose lower doses than those prescribed by medical staff (Haydon et al., 2011). In educational research, an open research question of relevance for designing feedback environments is whether students who have control over their feedback valence choose critical feedback more often, remember the feedback content better and, consequently, learn more than students who receive (i.e., are assigned) feedback. Therefore, this research starts tackling this question by focusing on two orthogonal dimensions of feedback, choice (feedback is chosen or assigned) and valence (feedback is confirmatory or critical). It also focuses on the impact of the feedback valence that students choose, rather than receive, on memory, performance, and learning. Lastly, it focuses on task-directed, not self-directed, feedback that is customized to each student's performance outcomes.

The article begins by reviewing the literature relevant to this study. Then, it describes the Posterlet assessment instrument, a game that collects students' choices to seek critical feedback and to revise their work in the domain of digital poster design (Cutumisu, Blair, Chin, & Schwartz, 2015; Cutumisu, Blair, Chin, & Schwartz, 2016). It describes the two studies, together with their post-tests administered immediately after the game: a free-recall task in Study 1 to examine the impact of feedback choices on students' memory for feedback, and a perception task in Study 2 to gauge learning of graphic design principles from Posterlet. Then, it presents empirical evidence provided by two studies. Finally, it concludes with a discussion of the educational implications, limitations, and future research directions.

# 2   LITERATURE REVIEW

This section reviews the relevant literature exploring the links among feedback, performance, choice-based assessments, and memory for feedback.

## 2.1    Critical Feedback and Memory for Feedback

In this research, critical feedback denotes constructive criticism. In cognitive tasks, behaviours after error commission have been investigated thoroughly (Houtman & Notebaert, 2013). Research suggests that error (negative) feedback is better remembered than correct (positive) feedback (Van der Borght, Schouppe, & Notebaert, 2016). There are many accounts for this result. First, the enhanced memory for critical feedback is supported by the *attentional boost effect theory* (Sparato, Mulligan, & Rossi-Arnaud, 2013) showing that stimuli co-occurring with critical feedback are also better recognized (Swallow & Jiang, 2010; Swallow & Jiang, 2012). Second, *post-error slowing* (PES) is a prominent post-error adjustment process that denotes a longer reaction time after an error than after a correct response, but its mechanisms are not clear (Danielmeier & Ullsperger, 2011). Third, a new theoretical framework posits that errors are surprising (Houtman, Castellar, & Notebaert, 2012; Notebaert et al., 2009) and they cause a distraction or an *attentional orienting* from the task towards the error, thus leading to PES (Houtman & Notebaert, 2013; Steinborn, Flehmig, Bratzke, & Schröter, 2012). Moreover, findings show that surprising feedback (i.e., for high-confidence errors or low-confidence correct answers) in a general-knowledge test leads to enhanced memory for both the content and the surface features (i.e., appearance) of feedback (Fazio & Marsh, 2009). Fourth, the *hypercorrection effect* (Butterfield & Metcalfe, 2006) from the memory literature also posits that surprising feedback increases attention, showing that high-confidence errors are more likely to be corrected after feedback than low-confidence errors.

Memory for feedback was also linked to individual differences leading to information processing biases (e.g., increased processing of the error feedback, compared to correct feedback). Learners adopt many strategies to cope with self-threatening feedback that accurately highlights their weaknesses. For instance, they display an inferior recall for such feedback compared to other types of feedback (e.g., self-affirming feedback that highlights one's strengths). The *mnemic neglect theory* posits that such an effect is attenuated, triggering self-improvement motivation when feedback is perceived as referring to modifiable traits (Dauenheimer, Stahlberg, Spreemann, & Sedikides, 2002; Green, Pinter, & Sedikides, 2005; Sedikides, Green, Saunders, Skowronski, & Zengel, 2016). In a different study, participants remembered feedback more accurately when the feedback valence was congruent with their self-esteem (Story, 1998) or their preferred self-views (Sanitioso, Kunda, & Fong, 1990). However, when it was not, participants recalled feedback distortedly as being more congruent with their self-esteem than it was in actuality. Notably, the study also revealed that the effect of self-esteem on recall accuracy persisted regardless of participants' feedback dwell time or perceptions of the feedback's credibility.

Researchers found a biased evaluation of evidence leading individuals to readily accept positive feedback, while closely scrutinizing critical feedback (Dawson, Gilovich, & Regan, 2002). For instance, researchers probing predictions of memory biases stemming from cognitive theories of anxiety found that biased recognition of social feedback was associated with social anxiety (Cody & Teachman, 2010). Specifically, participants with high social anxiety remembered a peer's feedback more positively than their own feedback. Moreover, they remembered their negative feedback as worse than the low social anxiety participants and they also experienced diminished positive feedback over time. Finally, post-event processing (i.e., a type of rumination typical of social anxiety) mediated the relationship between social anxiety and memory for negative feedback. Additionally, Kovacs and Beck (1978) found memory deficits for positive feedback in participants suffering from depression.

In the light of the research literature support for the better retention of critical than of confirmatory feedback, this research hypothesizes that, when students can choose between critical and confirmatory feedback, they remember critical feedback better than confirmatory feedback. It also begins to elucidate whether this stems from the choice or from the valence of feedback.

## 2.2    Feedback Seeking and Memory for Feedback

Traditional instructor-controlled models (Askew & Lodge, 2000) of assessment and feedback have been morphing into learner-controlled models (Carless, Salter, Yang, & Lam, 2011; Chai, 2003; Geitz, Joosten-Ten Brinke, & Kirschner, 2016; Hounsell, Boud, & Falchikov, 2007; Hounsell, McCune, Hounsell, & Litjens, 2008; Merry, Price, Carless, & Taras, 2013; Sadler, 2010) that facilitate self-regulated learning (SRL) and lifelong learning skills (Butler & Winne, 1995), encouraging dialogic feedback practices between students and instructors (Hattie & Gan, 2011). Consequently, the focus in education is shifting to not only enhancing students' performance outcomes, but also their learning strategies, because assessing such

processes provides an insight into how well students are prepared to innovate in the 21st century (Clark, 2012). Researchers argue that first-year university feedback practices should explicitly focus on developing students' SRL skills (Beaumont, O'Doherty, & Shannon, 2011; Geitz, Joosten-Ten Brinke, & Kirschner, 2016) and that "feedback is inherent in and a prime determiner of processes that constitute SRL" (Butler & Winne, 1995). Knowing what we do not know and when to seek feedback to fill the gaps in our knowledge is a crucial metacognitive skill for learning (Ehrlinger, Mitchum, & Dweck, 2016). Examining feedback choices that enable students to play an active role in their learning is also important from the perspective of SRL. Butler and Winne (1995) emphasized that, by engaging proactively with their feedback, learners can develop effective self-assessment skills that enable them to better appraise their own performance (McDonnell & Curtis, 2014; Wakefield, Adie, Pitt, & Owens, 2014).

However, most feedback research and school practices overlook feedback choice, examining feedback that is assigned to students, although learners often need to seek feedback to improve their learning. Thus, there has been little examination of feedback seeking, most feedback being assigned to, and not sought by, the learner. In an eye-tracker study, a feedback model was employed to investigate if and how long the learner looked at the feedback chosen or received (Conati, Jaques, & Muir, 2013; D'Mello, Olney, Williams, & Hays, 2012), but the option presented to the learner was whether to receive feedback, not to choose its valence. Additionally, students do not always remember the feedback they receive (Gibbs & Simpson, 2005; Price, Handley, Millar, & O'Donovan, 2010), although research shows that feedback is beneficial for memory (Sitzman, Rhodes, & Kornell, 2016). Several models are proposed in the literature to describe feedback (Hattie & Timperley, 2007; Kluger & DeNisi, 1996) and memory for feedback (Butler, Karpicke, & Roediger, 2008). Few examined the feedback process with its cognitive and affective components to account for the differential effect of feedback on learning, memory, and learners. Most models share an important limitation in scope, as they focus almost exclusively on feedback that is assigned to, not chosen by, the learner (Geitz, Joosten-Ten Brinke, & Kirschner, 2016; Timms, DeVelle, & Lay, 2016). To date, little is known about learners who have a preference for choosing critical (negative) rather than confirmatory (positive) feedback. Research on the neural correlates of learning provides evidence that neural responses to feedback can predict future performance. For instance, the brain responses to feedback are predictive of whether university students will repeat mistakes or will learn from their mistakes (van der Helden, Boksem, & Blom, 2010). In contrast, the present article examines a different population (i.e., middle-school students).

Despite research showing that feedback information is rarely used in revision of work (Carless, 2006), we found that revision was strongly associated with willingness to choose critical feedback across many studies (Cutumisu, Chin, & Schwartz, 2014; Cutumisu, Blair, Chin, & Schwartz, 2015; Cutumisu, Blair, Chin, & Schwartz, 2016). Moreover, recent research shows a positive association of revision with feedback recall (Bolzer, Strijbos, & Fischer, 2015). The current research examines this association when students choose between critical and confirmatory feedback. In this research, critical feedback constitutes constructive criticism and not punishment, and students exercise a choice regarding their feedback valence. This research tests the hypothesis that students perform better when they have more agency (e.g., a choice) regarding their feedback valence.

# 3 THE ASSESSMENT GAME: POSTERLET

Assessing how students seek and remember feedback requires a new type of assessment that examines students' choices. Increasingly, educators aim to support learners in developing 21st-century skills that will prepare them to tackle complex problems (e.g., rapidly-spreading diseases). In 2012, the *Programme for International Student Assessment* (PISA) introduced items that collected information about students' attitudes towards problem solving for the first time since it started administering tests in 2000 (OECD, 2013). This trend is due in part to the focus of traditional assessments on outcome accuracy, rather than on the preparedness of students to perform well on new tasks. In contrast, choice-based assessments focus on the learning processes in which students engage when solving a new challenge and, thus, these types of novel assessments offer a glimpse into how prepared students are to learn on their own (Schwartz & Arena, 2013). In this study, a choice-based dynamic assessment game, Posterlet, is employed to collect and measure students' proactive choices to seek feedback and to revise as a way of capturing their preparedness to learn on their own. In addition to confirmatory and critical feedback choices, Posterlet offers players opportunities to learn graphic design principles while designing a digital poster. This computer-based assessment and

learning environment employs artificial intelligence in an educational context to track students' learning processes and performance, as well as to provide them with customized feedback as they play the game. Thus, Posterlet also examines the impact of students' choices to seek critical feedback on performance and learning. The design of Posterlet draws on constructivist, choice-based assessments (Schwartz, Lindgren, & Lewis, 2009; Schwartz & Arena, 2013) that emphasize learning during the assessment and that shift the assessment focus from the learning outcomes to the learning processes (e.g., choosing critical feedback) involved in solving a challenge (e.g., designing a poster). The next section presents Posterlet, the assessment instrument that collects students' choices to seek critical feedback and to revise their posters.

Posterlet is typical of a graphic design environment and it involves an artificial learning task (digital poster design). Although all students have equal opportunities to learn about graphic design principles while playing the game and engaging with feedback, Posterlet is mainly designed to assess students' choices rather than only to teach students about poster design. Posterlet is an instance of a constructivist assessment in which students can learn (i.e., graphic design principles) during the assessment (i.e., of learning choices and poster performance). This enables us to explore the potential relation between students' learning choices (e.g., seeking critical feedback and choosing to revise) and their poster task performance in the same assessment environment. This creative, open-ended task enables us to control the variables of interest. It also enables us to measure students' choices to seek feedback and to revise their posters, and, thus, it enables us to measure students' preparedness to learn on their own. Notably, Posterlet is more representative of most tasks students face in real learning and assessment environments than in school, especially because creativity is quintessential for innovation. Importantly, this task presents little variation in learners' prior experience. A thorough description of the Posterlet assessment game is included in our previous research (Cutumisu, Blair, Chin, & Schwartz, 2015).

In Posterlet, players take on the role of designing posters for booths at a fair. They choose a booth of interest (e.g., basketball toss) and they design a poster using a graphical user interface provided by the game. Posterlet measures two choices that a player makes upon completing a poster: 1) choose either confirmatory or critical feedback from three virtual characters about the poster and 2) choose to revise the poster after reading all three pieces of feedback. A variation of the Posterlet game was specifically designed to enable a performance outcome comparison between college students who choose and those who receive (i.e., are assigned) feedback. In this new version, feedback is assigned to the player in a principled way that mirrors the feedback chosen by a corresponding player of the original Posterlet version. The game enables students to design a poster on each of the game's three rounds and to perform two main choices at the end of each poster design task. Upon completing each poster, students choose either confirmatory (e.g., "It's good you told them what day the fair is.") or critical (e.g., "People need to be able to read it. Some of your words are too small.") feedback from three animal characters, as shown in Figure 1. In the Posterlet game, feedback messages of each valence alternate between informative (non-generic) elaboration feedback that pertains to the poster design task (e.g., "You did not mention the location of the fair." or "It is good that you mentioned the location of the fair.") or uninformative (generic feedback that does not pertain to the task or to the self) feedback ("I don't like fairs." or "I go to lots of fairs."). After reading the feedback, students choose whether to revise that poster. Posterlet tracks these two choices (seeking critical feedback and revising) and computes a poster score per poster round, as well as a cumulative poster score per game (poster performance).

Figure 1. In the Posterlet game, the students design posters and choose either confirmatory or critical feedback about their posters from each of the virtual animal characters of a focus group. (Reprinted from Cutumisu, Blair, Chin, and Schwartz, 2015.)

# 4  DESCRIPTION OF THE STUDIES

**Study 1** presents a novel examination of students' memory for feedback when they have a choice between critical and confirmatory feedback. An empirical study sampling a hundred and six Grade 8 middle-school students was designed to collect their choices to seek critical feedback via Posterlet, a digital assessment game in which students design posters, and to test their memory for the feedback they chose in Posterlet. This study tested the hypothesis that students remember critical feedback better than confirmatory feedback. **Study 2** examines the effect of choosing versus receiving feedback on the learning outcomes of ninety-eight post-secondary students from California on a digital poster design task. The study employs a yoked experimental design where college students were randomly assigned to play a choice-based assessment game, Posterlet, in one of two conditions, Choose or Receive. In the Choose condition, students chose confirmatory (i.e., positive) or critical (i.e., negative) feedback about their posters. In the Receive condition, students were assigned the same feedback valence in the same order that students in the Choose condition chose. The study tests the hypothesis that students' critical feedback and learning performance are associated positively when students choose their feedback and negatively when students receive their feedback. Additionally, differences between conditions are explored in terms of performance, the choice to revise, enjoyment, and time spent designing posters. Implications for designing automated feedback-rich environments are discussed.

# 5  STUDY 1

This study poses the following research question to assess how students' memory for critical feedback is impacted by feedback agency: Do students remember critical feedback more than confirmatory feedback when they have a choice between critical and confirmatory feedback?

## 5.1  Methods

### 5.1.1  Participants and procedure

Participants were n = 106 (60 females) Grade 8 students, aged 13-14, from a public middle school in California. All students played Posterlet and n = 86 of them filled an online memory for feedback post-test immediately after playing the game in May 2015. The assessments were administered during class time, but students were not rushed or given a specific time to complete the tasks. Some students were not able to finish all three posters or the post-test. Students who did not provide consent (n = 9) or did not complete all posters (n = 8) were excluded from analyses. Thus, the analyses include n = 89 students (50 females). In addition to some of the students not completing the post-test due to time constraints, some parents did not provide consent for sharing their children's standardized test scores, so students were removed from the analyses as needed.

First, students played the Posterlet game individually, in which they designed three posters (M = 14.76 minutes, SD = 4.07), as one of several assessments administered that day. They were not provided with any instruction beyond that included in the introduction of the Posterlet game. In the introductory portion of the game, students advanced through a series of web pages representing the steps needed to complete each poster and the rationale of the task: creating posters for the school's fun fair to attract visitors for each of their three poster booths. In Step 1, students were provided with the following prompt: "You're on the team to organize the school's Fall Fun Fair". In Step 2, students were assigned the task: "Your job: Design posters for the activity booths". In Step 3, students were provided with the following prompt: "Test your posters & get feedback". Lastly, in Step 4, students were given the final instructions: "Post your poster & see if people come to the booth". Due to time constraints, a post-test to measure students' learning of graphic design principles was not administered.

Second, immediately following the game, students were automatically directed to an online memory for feedback post-test (M = 2.33 minutes, SD = 1.09). The memory for feedback post-test asked students to recall as many feedback comments as they remembered from the Posterlet game, out of a maximum of nine (i.e., there are three opportunities to choose feedback for each of the three game rounds). Students were provided with the following prompt and a screenshot from the game, as illustrated in Figure 2: *If you played the Posterlet game in which you designed posters for a funfair, please list below as many **comments** as you can remember that you received from the animal characters in the game.*



**Figure 2. The memory for feedback post-test asking students to recall their feedback right after playing Posterlet.**

### 5.1.2  Measures and data sources

Students designed three posters in the Posterlet game with no time limit imposed on each poster. Posterlet measured students' choices to seek critical feedback. A post-test immediately following the game measured students' free recall of the feedback messages they chose in Posterlet.

### 5.1.2.1 Choices

*Critical Feedback* measures the total amount of critical ("I don't like…") feedback a student chose, ranging from zero (i.e., the student chose confirmatory feedback throughout the game) to nine (i.e., the student chose critical feedback across the game). *Confirmatory Feedback* is computed as nine minus Critical Feedback, as students can choose either critical or confirmatory feedback out of the nine feedback choices across the game. In this research, critical feedback constitutes negative, elaboration feedback that is constructive and not punitive. For instance, if the text on a poster is too small to be read from a distance and if the student chooses to receive critical feedback on that poster, the feedback message could be "People need to be able to read it. Some of your words are too small." *Revision* measures the total number of posters a student chose to revise, ranging from zero (i.e., the student did not revise any poster) to three (i.e., the student revised all posters).

### 5.1.2.2 Memory for feedback

The memory for feedback post-test was devised to collect the feedback messages students freely recalled after playing the game. Feedback alternated between informative and uninformative to avoid cognitive load for younger participants. The example presented in Table 1 of a student's post-test answers and scores shows that, of the seven feedback messages that the student remembered, four were critical (two informative and two uninformative), while three were confirmatory (all informative). The table also contains the actual amount of feedback chosen by this student per type of feedback. For example, out of the total of nine feedback choices, this student chose critical feedback six times and confirmatory feedback three times across the entire game, as shown in the last row of this table.

**Table 1. A sample of a student's answers and scoring for the memory for feedback post-test. The student did not provide answers for items 8 and 9, so these answers were scored as zero.**

| Order | Feedback Remembered | Cr. I. Rem. | Cr. U. Rem. | Co. I. Rem. | Co. U. Rem. | Cr. Rem. | Co. Rem. | Total |
|---|---|---|---|---|---|---|---|---|
| 1 | I don't like fairs | 0 | 1 | 0 | 0 | 1 | 0 | 1 |
| 2 | I like that the text does not cut off the page | 0 | 0 | 1 | 0 | 0 | 1 | 1 |
| 3 | I don't really go to fairs | 0 | 1 | 0 | 0 | 1 | 0 | 1 |
| 4 | I don't like that the text is too close together | 1 | 0 | 0 | 0 | 1 | 0 | 1 |
| 5 | I don't like that it doesn't have the admissions price | 1 | 0 | 0 | 0 | 1 | 0 | 1 |
| 6 | I like that is has the date and time | 0 | 0 | 1 | 0 | 0 | 1 | 1 |
| 7 | I like that it has the location | 0 | 0 | 1 | 0 | 0 | 1 | 1 |
| 8 | - | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | - | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Total** | | **2** | **2** | **3** | **0** | **4** | **3** | **7** |
| **Actual** | | **3** | **3** | **3** | **0** | **6** | **3** | **9** |

Note: Cr. = Critical, Co. = Confirmatory, I. = Informative, U. = Uninformative, Rem. = Remembered

*Critical Feedback Remembered* measures the number of critical feedback messages that each student recalled from Posterlet, while *Confirmatory Feedback Remembered* measures the number of confirmatory feedback messages that a student recalled. *Total Feedback Remembered* is the sum of *Critical Feedback Remembered* and *Confirmatory Feedback Remembered*.

*Percent of Critical Feedback Remembered* measures the portion of Critical Feedback Remembered out of the Critical Feedback chosen in Posterlet (four out of six for the example provided in Table 1), while *Percent of Confirmatory Feedback Remembered* measures the portion of Confirmatory Feedback Remembered out of the Confirmatory Feedback chosen in Posterlet (three out of three for the example provided in Table 1). It is possible that participants remember higher levels of critical feedback simply because they choose critical feedback more often. Thus, computing a percent of feedback remembered out of the feedback chosen for each feedback valence enables a comparison between critical and confirmatory feedback remembered, regardless of the amount and valence of the feedback chosen. Overall, these measures illustrate the quantity or completeness of students' relevant memory for feedback (i.e., how many relevant feedback messages of each valence students remembered). High percent values indicate that students remembered most of the relevant possible feedback messages they chose from Posterlet. A score of 1.0 indicates that the student remembered all the relevant feedback messages from Posterlet.

## 5.2  Data Analyses and Results

### 5.2.1  Do students remember critical feedback more than confirmatory feedback when they have a choice between critical and confirmatory feedback?

Figure 3 illustrates the average percent of confirmatory feedback remembered by levels of confirmatory feedback chosen (Figure 3a) and the average percent of critical feedback remembered by levels of critical feedback chosen (Figure 3b), respectively. In Figure 3a, the x-axis represents the confirmatory feedback levels chosen by students across the game and the number of students who chose each level (e.g., n = 14 students chose confirmatory feedback exactly four times across the game). In Figure 3b, the x-axis represents the critical feedback levels chosen by students across the game and the number of students who chose each level (e.g., n = 14 students chose critical feedback exactly five times across the entire game). In this study, all participants who completed the memory post-test made at least two critical feedback choices (i.e., the x-axis starts at two). In all figures, error bars represent one standard error and the y-axis represents the average memory for feedback measures. Also, critical feedback measures are represented in red, while the confirmatory feedback measures are represented in green.
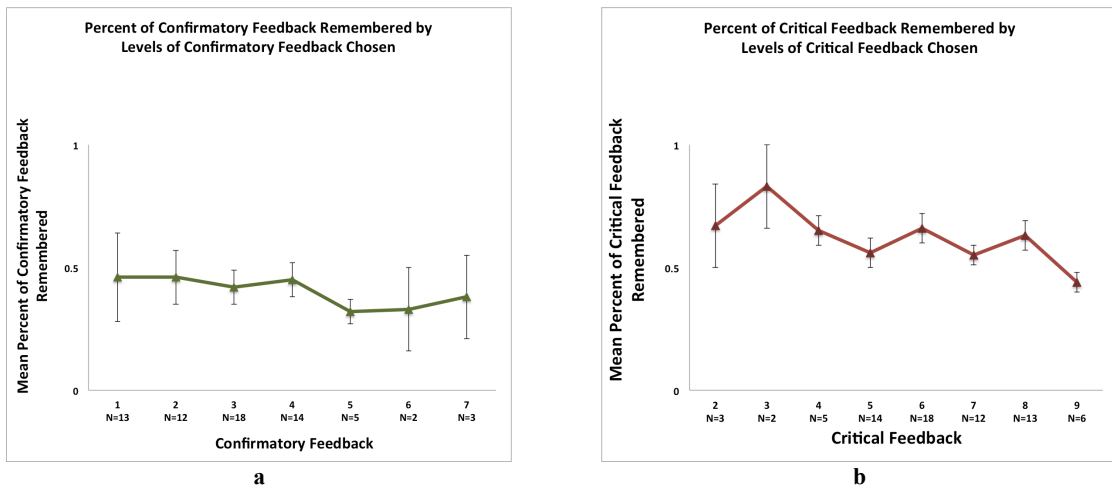


a                                                                          b

**Figure 3. Average percent of confirmatory (a) and critical (b) feedback remembered by levels of feedback chosen.**

Figure 4 illustrates the average percent of both confirmatory and critical feedback remembered, respectively, by levels of critical feedback remembered, suggesting that critical feedback is better remembered than confirmatory feedback. It shows that the average values of the memory for critical feedback were statistically significantly higher than the average values of the memory for confirmatory feedback for each level of critical feedback chosen, except for the first critical feedback level (i.e., students who made exactly two critical feedback choices and, consequently, seven confirmatory feedback choices across the game), where this relation was reversed.
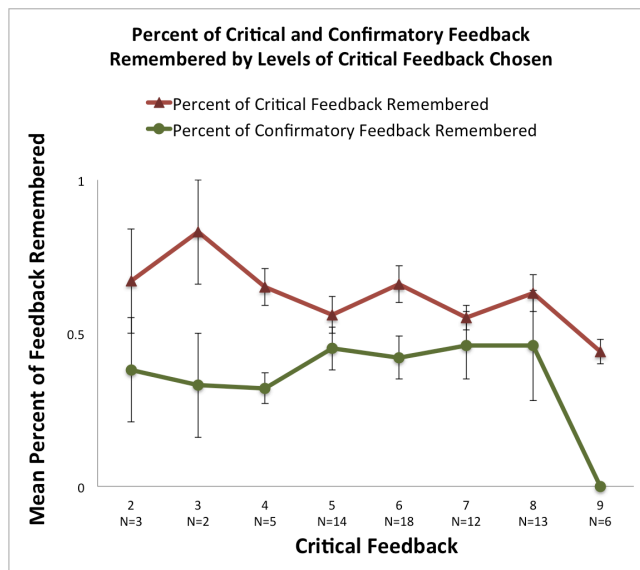
**Figure 4. Average percent of critical and confirmatory feedback remembered by levels of critical feedback chosen.**

A paired-samples t-test analysis comparing *Percent of Critical Feedback Remembered* with *Percent of Confirmatory Feedback Remembered* found that, indeed, students remembered significantly more critical than confirmatory feedback from the Posterlet game [$t(72) = 4.26$, $p < .001$]. On average, the *Percent of Critical Feedback Remembered* was higher with .21 points than the *Percent of Confirmatory Feedback Remembered* (95% CI [.11, .30], $\eta_p^2 = .20$). Furthermore, Figure 5 shows that students remembered a significantly larger amount of critical than of confirmatory feedback.
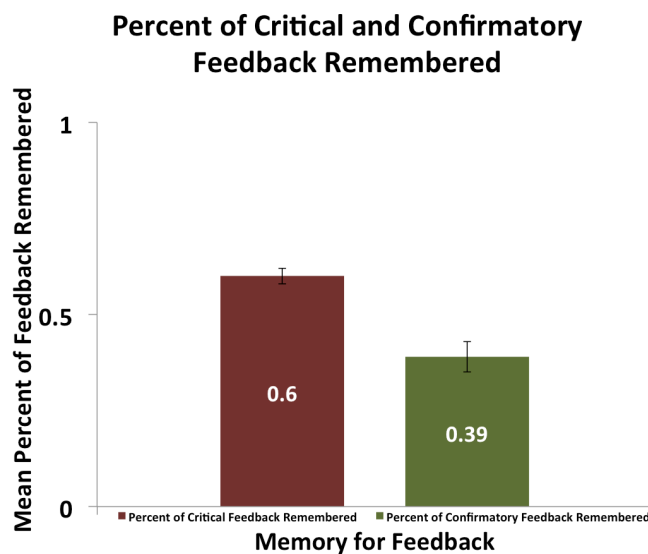


**Figure 5. Average percent of critical and confirmatory feedback remembered.**

In Figure 5, the x-axis represents the categories of critical and confirmatory feedback types remembered by students. The same pattern of results emerged in this figure as well: the percent of critical feedback remembered was statistically significantly higher than the percent of confirmatory feedback remembered.

## 5.3   Discussion, Limitations, and Future Work

The current study reveals that students remember critical feedback significantly more than confirmatory feedback, consistent with other results in the memory for feedback literature. This result also held at each

critical feedback level greater than two. This seems to support our hypothesis that critical feedback has a more lasting positive impact on students' memory for critical feedback when students have a choice over their feedback. It also suggests that students do not seem to discount critical feedback in hindsight. A possible explanation for this result is that individuals process critical feedback on a deeper level than confirmatory feedback when they have a choice over the valence of their feedback. To determine if this result is due to the critical valence of the feedback, to feedback agency, or to a combination of these factors, future research will explore whether students remember critical and confirmatory feedback differentially when they are assigned feedback rather than when they choose feedback.

The results of many feedback studies do not enable conclusions about the long-term effect of choice on the memory for feedback. One limitation of the current study is the immediate testing of students' recall following the Posterlet game. However, a related study, in which high-school students played Posterlet and, after a week's delay, they were asked to recall the feedback they chose in Posterlet, yielded consistent results with those of the current study (Cutumisu & Schwartz, 2016b). A follow-up study in which students' memory for feedback will be probed not only immediately after playing Posterlet, but also after several variable-length delays, will be conducted to determine any recall biases.

Future work will focus on the relation between attention and memory for critical feedback, as postulated by the orienting account, through exploring short-term and long-term memory for feedback effects. Thus, relations among Posterlet measures will be examined, including the feedback dwell time, and post-test measures to test the *attentional orienting* and *hypercorrection effect* accounts of *PES* by recording students' levels of surprise, confidence, and subsequent revision behaviour, respectively, after reading their feedback. The *attentional boost effect theory* (Spataro, Mulligan, & Rossi-Arnaud, 2013) will be tested by asking students to freely recall other game features co-occurring with critical feedback, such as the animal virtual characters that delivered this feedback.

## 5.4 Conclusion

This research examined the relation between *feedback valence choice* and *memory* by exploring the impact of students' feedback agency on their memory for feedback. An empirical study was designed to collect students' learning choices via an assessment game, Posterlet, and to assess their memory for feedback via a free recall post-test that immediately followed the game. Results provide evidence that students remember critical feedback better than confirmatory feedback when they have a choice over their feedback valence.

# 6 STUDY 2

This study examines the effect of choosing versus receiving feedback on the learning outcomes of n = 98 post-secondary students from a community college in California on a digital poster design task. The study employs a yoked experimental design where college students are randomly assigned to play Posterlet in one of two conditions, Choose or Receive. In the Choose condition, students choose confirmatory (i.e., positive) or critical (i.e., negative) feedback about their posters. In the Receive condition, students are assigned the same feedback valence and quantity that students in the Choose condition chose. Specifically, this research aims to examine the effect of feedback agency (i.e., choosing versus receiving feedback) on the performance of college students, by comparing learning outcomes between participants who choose feedback and those who are assigned the same amount, valence, and order of feedback. In addition to performance and time on task, students' reported enjoyment of designing posters is assessed, as well as their willingness to revise their posters (i.e., their choice to revise). An experimental study was designed to address the following research questions:

1) Does critical feedback correlate with performance outcomes by condition?
2) Are there any outcome differences between choosing and receiving feedback?
3) Are there any enjoyment differences between choosing and receiving feedback?

## 6.1 Methods

### 6.1.1 Participants and procedure

Participants were n = 98 (55 female) students aged 18 to 52, $M_{age}$ = 22.71 years (SD = 5.59) from a college in California. Both versions of the Posterlet game and a post-test were employed to collect data between Spring

2014 and Spring 2015. All students provided consent and received credit in their psychology courses for their participation in this study. They were randomly assigned to one of two conditions, Choose (n = 49 students aged 18 to 52, $M_{age}$ = 22.53 years, $SD_{age}$ = 5.96, 32 female) and Receive (n = 49 students aged 18 to 38, $M_{age}$ = 22.90 years, $SD_{age}$ = 5.26, 23 female), according to a yoked study design. These conditions correspond to two different game versions. In each game version, in contrast to Study 1, there were only two rounds (i.e., game levels) available. Both Posterlet versions shared a common introduction to the game in which participants were provided with the same instructions described in Study 1. In the Choose game version, students choose their feedback valence, as illustrated in Figure 6a, by selecting either critical or confirmatory feedback from each virtual character. In the Receive game version, students are assigned their feedback valence, as illustrated in Figure 6b. After designing posters in two rounds of the game for $M_{Choose}$ = 8.72 minutes (SD = 3.28) and $M_{Receive}$ = 7.54 minutes (SD = 3.84), all participants completed the same online post-test that measured their knowledge of graphic design principles and enjoyment of designing posters. After removing two outliers in the Receive condition, the time students spent on the post-test in each condition was $M_{Choose}$ = 6.56 minutes (SD = 1.44) and $M_{Receive}$ = 6.66 minutes (SD = 1.64), respectively.
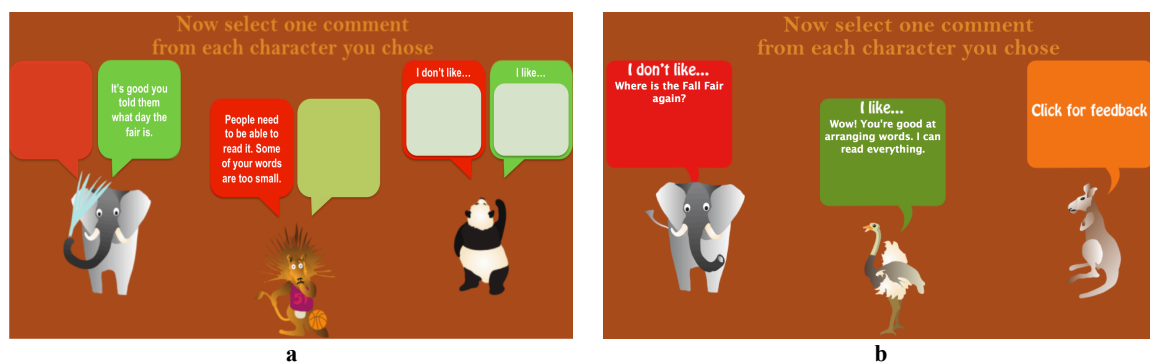


**Figure 6. In the Choose condition (a), the player chose critical feedback from the lion and then confirmatory feedback from the elephant. Correspondingly, in the Receive condition (b), the player first received critical feedback from the elephant and then confirmatory feedback from the ostrich.**

### 6.1.2 Measures and data sources

Students completed one of two versions of the game with a five-minute time limit per poster, with the primary difference being whether 1) they chose between critical and confirmatory feedback or 2) they were assigned a schedule of critical and confirmatory feedback. Posterlet measured students' behaviours, as well as their poster performance in the game. A post-test following the game measured students' learning of graphic design principles acquired through playing Posterlet.

#### 6.1.2.1 Choices

*Critical Feedback* measures the number of times that the student chose or received critical (i.e., *I don't like*) feedback, ranging from 0 (students chose/received only confirmatory feedback across the game) to 6 (students chose/received only critical feedback across the game). As in Study 1, feedback of both valences alternated between informative and uninformative.

*Revision* measures the number of times that the student chose to revise a poster, ranging from 0 (the student did not revise any posters) to 2 (the student revised both posters). All students had a choice to revise their posters, even though students in the Receive condition did not have a choice regarding the valence of their feedback. In contrast with Study 1, in this experiment, each game version consisted of only two game rounds, with each poster design being limited to five minutes.

#### 6.1.2.2 In-game performance

In this research, in-game performance constitutes the ability of students to design posters. Posterlet computes a *Poster Quality* score based on 21 design principles reflecting a student's performance across the game to measure in-game performance (i.e., the quality of the posters created by the student in Posterlet). The game evaluates each poster against a set of 21 graphic design rules provided by a graphic artist, illustrated in Figure 7 and described in detail in previous research (Cutumisu, Blair, Chin, & Schwartz, 2015). Then, every time

the student clicks on a feedback box in Posterlet, the game's feedback system selects a feedback message to return to the student according to the quality of the poster and a priority scheme. It alternates the selection of feedback among three main categories: crucial information, readability, and space use. Within each category (e.g., *readability*), it retrieves the feedback in a top-down fashion (e.g., *text size small* before *text style unreadable*).

| INFORMATION | READABILITY | SPACE USE |
|---|---|---|
| location absent | text cut off by bounding box | space used by graphics out of range |
| date absent | text size small | top half empty |
| time absent | text style unreadable | bottom half empty |
| ticket price absent | text & graphics overlap | graphics touching edge |
| booth description absent | text and other text overlap | text flush with edge |
| graphics irrelevant | graphics size small | text flush with other text |
| text absent | text contrast low | - |
| images absent | - | - |

**Figure 7. The 21 graphic design rules employed by the feedback system embedded in the Posterlet game. (Reprinted from Cutumisu, Blair, Chin, and Schwartz, 2015.)**

For each poster, the game evaluates each rule with 1, if the rule is always used correctly on that poster; 0, if the rule is not applicable on that poster; and -1, if the rule is used incorrectly on that poster. The score of any individual poster created by a student represents the sum of all the 21 rule scores, ranging from -21 to 21. Thus, in this study, *Poster Quality* represents the score sum across the game of the last individual poster on each of the two rounds, ranging from -42 to 42.

### 6.1.2.3  Time on task
*Design Duration* measures the amount of time students spent designing all posters, including revisions. Specifically, the game starts measuring the time a student spends from choosing a poster theme (e.g., basketball) until submitting the first poster draft. If the student chooses to revise the poster after reading the feedback, then this measure includes the additional time that the student spends updating that poster.

### 6.1.2.4  Learning
*Poster Ranking* measures a student's learning of design principles on a post-test independent of the game. After completing the game, the student is directed to an online post-test to assess four sets of posters. Each set contains two versions of a poster, featuring a design principle used correctly on one poster and incorrectly on the other poster. For each set, the student is shown in a five-second succession the first poster, a distractor pattern image, and a modified version of the first poster, as illustrated in Figure 8. Then, the student decides whether the second poster is the same, better, or worse than the first poster. Each answer is scored with 1, if it is correct and 0, if it is incorrect. Thus, *Poster Ranking* ranges from 0 to 4.

### 6.1.2.5  Enjoyment
*Enjoyment* measures students' self-reported enjoyment of designing posters on a 1-5 Likert-type response scale, where 1 = *none* and 5 = *a huge amount*. Students answered this question after completing the game and the post-test.
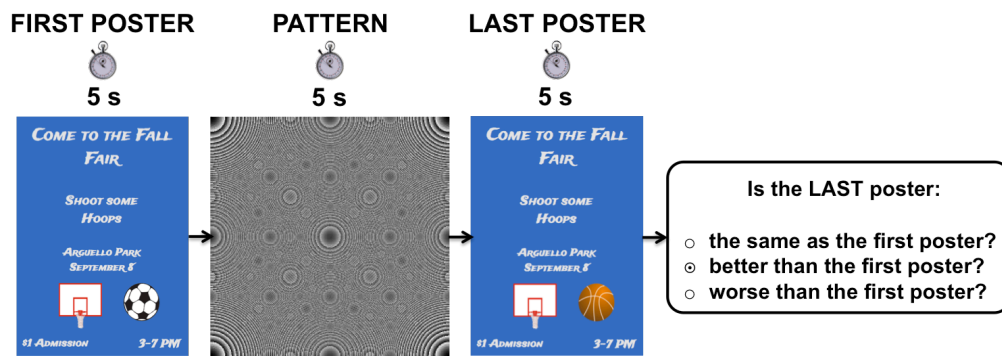
**Figure 8. One of the four post-test items measuring students' learning of graphic design principles. This item targets the rule of graphics relevance (e.g., a basket ball is more suited than a soccer ball) to the poster's theme (e.g., basketball).**

## 6.2 Data Analyses and Results

### 6.2.1 Does critical feedback correlate with learning outcomes by condition?

Spearman's rank correlation coefficient (*rho*) is reported in all the analyses included in this section, because the variables measured were not normally distributed. The relations between students' learning choices (critical feedback and revision) and learning outcomes (poster performance as measured by Poster Quality and learning of the graphic design principles as measured by Poster Ranking) are explored.

Although it measures in-game poster performance, Poster Quality can also be considered to be a learning measure, because students improved their performance from the first to the second game round. Poster 1 is considered to be the pretest, being the first poster designed by the player in the Posterlet game, before revision. Poster 2 is the last poster designed by the student in the Posterlet game, after a potential revision. A repeated-measures analysis of variance reveals that the poster quality of the students in the Choose condition increased significantly from Poster 1 (M = 9.59, SD = 5.69) to Poster 2 (M = 13.04, SD = 3.96); F(1, 48) = 24.67, $p < .001$, $\eta_p^2 = .34$. As well, in the Receive condition, poster quality increased significantly from Poster 1 (M = 9.33, SD = 6.93) to Poster 2 (M = 12.63, SD = 4.37); F(1, 48) = 14.72, $p < .001$, $\eta_p^2 = .23$. Critical Feedback also correlates positively with Poster Ranking, although not statistically significantly.

In the Choose condition, Critical Feedback correlates with performance on the posters measured by Poster Quality and strongly with Revision, as shown in Table 2. Values marked in bold font indicate statistically significant correlations.

**Table 2. Correlations between choices and learning outcomes (in-game and post-test) in the Choose condition.**

| Measures (n = 49) | Revision | Poster Quality | Poster Ranking |
|---|---|---|---|
| Critical Feedback | **.64\*\*** | **.28\*** | .22 |
| Revision | -- | .27 | .14 |
| Poster Quality | -- | -- | .18 |

Note: \*\* *p* < .01, \* *p* < .05

In the Receive condition, Critical Feedback correlates strongly with Revision and inversely with performance on the graphic design principles measured by Poster Ranking, as shown in Table 3. Although not statistically significantly, Critical Feedback also inversely correlates with poster quality in this condition.

**Table 3. Correlations between behaviours and learning outcomes (in-game and post-test) in the Receive condition.**

| Measures (n = 49) | Revision | Poster Quality | Poster Ranking |
|---|---|---|---|
| Critical Feedback | **.44\*\*** | -.16 | **-.34\*** |
| Revision | -- | .09 | .10 |
| Poster Quality | -- | -- | .23 |

Note: \*\* *p* < .01, \* *p* < .05

Taken together, these results imply that the revealed association between critical feedback and learning outcomes is positive when students exercise a feedback choice and negative otherwise (i.e., when they are assigned their feedback). The strength of the relation between critical feedback and the measures of learning was compared in more depth between the two conditions, Choose and Receive. Specifically, a *t*-test analysis was conducted to compare the correlation coefficients of Critical Feedback and Poster Quality between the Choose and Receive conditions to investigate whether the two correlation coefficients are significantly different from each other. Results indicated that the two correlation coefficients were statistically significantly different from each other, *z-score* = 2.15, *p* = .03 (Fisher, 1921; Soper, 2015). A comparison of the Critical Feedback and Poster Ranking correlation coefficients between conditions yielded similar results (*z-score* = 2.77, *p* < .01). Thus, the relation between critical feedback and the measures of learning differs statistically significantly between the two conditions.

## 6.2.2   Are there any outcome differences between choosing and receiving feedback?

In the light of the correlation discrepancies between critical feedback and learning outcomes between conditions, several two-way analyses of variance were conducted to examine whether the condition (Choose versus Receive) moderates the relations between critical feedback levels and poster performance, learning of graphic design principles, the choice to revise, and time on task, respectively. Participants were divided into two percentile groups according to the amount of critical feedback they encountered (Group 1: zero to two and Group 2: three to six). This demarcation also coincides with a turning point in outcomes at the critical feedback level three (i.e., when participants encountered three critical feedback messages across the entire game) in the figures presented in the next section, indicating a possible interaction between critical feedback and condition. Moreover, this is a valence equilibrium point, because participants encountering three critical feedback messages also encounter an equal number of confirmatory feedback messages (i.e., there are six feedback opportunities across the game in this study).

Results revealed an interaction of critical feedback and condition for both learning outcomes, Poster Quality [$F(1, 94) = 6.26$, $p < .05$, $\eta_p^2 = .06$] and Poster Ranking [$F(1, 94) = 6.79$, $p < .05$, $\eta_p^2 = .07$], but no main effects, as shown in Figure 9. Post-hoc comparisons with the Bonferroni adjustment revealed that there were no differences in Poster Quality between conditions, regardless whether participants encountered higher or lower levels of critical feedback. Also, participants who engaged with higher levels of critical feedback significantly designed better posters [Poster Quality: $F(1, 94) = 5.14$, $p < .05$, $\eta_p^2 = .05$] than the rest of the participants only if they chose their feedback (i.e., only in the Choose condition). There was no difference in poster designs between levels of critical feedback for participants in the Receive condition ($p = .21$). Post-hoc comparisons with the Bonferroni adjustment also revealed that participants in the Choose condition learned significantly more graphic design principles [Poster Ranking: $F(1, 94) = 4.71$, $p < .05$, $\eta_p^2 = .05$] than their counterparts in the Receive condition only when they engaged with higher (i.e., three or more) levels of critical feedback. However, there was no difference between conditions when participants engaged with lower (i.e., less than three) levels of critical feedback ($p = .14$). Also, participants who engaged with higher levels of critical feedback learned significantly less graphic design principles [Poster Ranking: $F(1, 94) = 5.76$, $p < .05$, $\eta_p^2 = .06$] than the rest of the participants only if they were assigned their feedback. There was no difference between levels of critical feedback in learning of graphic design principles for participants in the Choose condition ($p = .20$). There was also a main effect of critical feedback for Revision [$F(1, 94) = 18.91$, $p < .001$, $\eta_p^2 = .17$], but no interaction between critical feedback and condition. Lastly, no interaction or main effects were found for Design Duration.
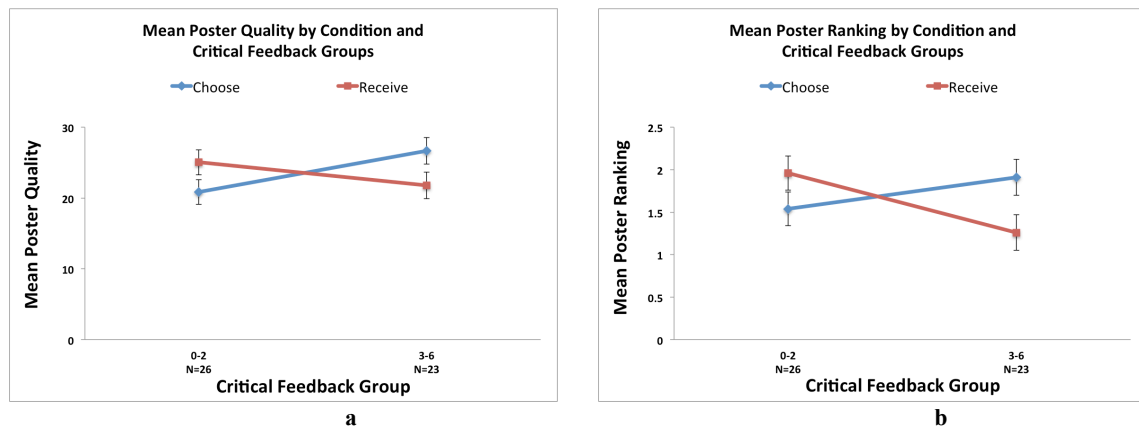
**Figure 9. Interactions between condition and critical feedback group in predicting the Poster Quality (a) and Poster Ranking (b) learning outcomes.**

In the next section, each of the learning strategies (e.g., choice to revise and time on task) and the learning outcomes (Poster Quality and Poster Ranking) were compared between conditions. Both parametric and non-parametric analyses of outcome differences were conducted and yielded the same results. The parametric analyses are presented next for ease of interpretation.

### 6.2.2.1 Revision

An independent-samples t-test analysis was conducted to compare Revision between students in the two conditions. There were no significant differences in Revision between students in the Choose (M = .76, SD = .83) and Receive conditions (M = .53, SD = .62); $t(88.56) = 1.52$, $p = .13$. Figure 10 shows the mean Revision plotted by levels of Critical Feedbacks (from zero to six) for each of the two conditions, suggesting that the choice to revise increases with the levels of critical feedback encountered by participants in each condition.
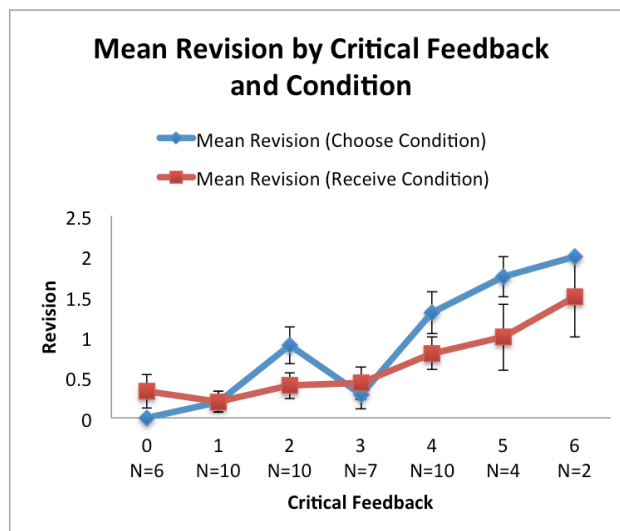


**Figure 10. Mean Revision by critical feedback and condition.**

### 6.2.2.2 Poster Quality

An independent-samples t-test analysis was conducted to compare Poster Quality between conditions. There were no significant differences in Poster Quality between students in the Choose (M = 23.57, SD = 8.85) and Receive conditions (M = 23.51, SD = 9.48); $t(96) = .03$, $p = .97$. Figure 11 shows the mean Poster Quality plotted by Critical Feedback for each of the two conditions, suggesting that participants who are assigned lower levels of critical feedback outperform those who choose the same levels of critical feedback, but that

there is a crossover in learning outcomes after participants encounter three critical feedback levels (i.e., the point when participants encounter the same number of critical and confirmatory feedback). This indicates that too much critical feedback may potentially hurt the performance of participants who are assigned feedback, while the opposite trend seems to occur for participants who choose their feedback. Another independent-samples t-test analysis was conducted to compare the quality of the Pretest (the first poster before revisions) between students in both conditions. There were also no significant differences in Pretest between students in the Choose (M = 9.59, SD = 5.69) and the Receive conditions (M = 9.33, SD = 6.93); $t(96) = .21$, $p = .84$.
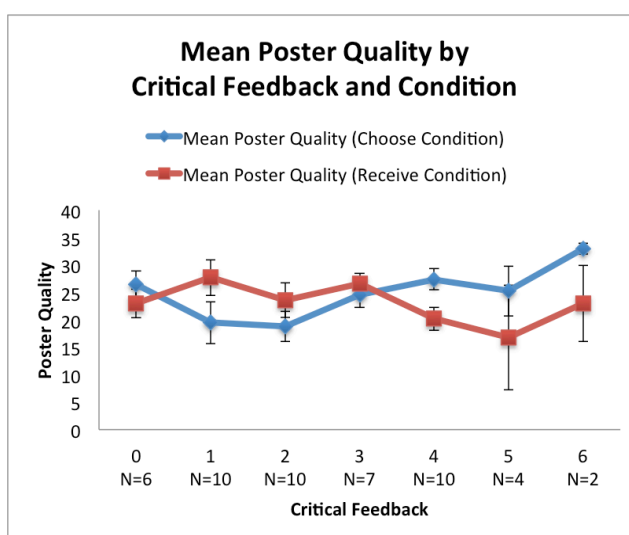


**Figure 11. Mean Poster Quality by critical feedback and condition.**

### 6.2.2.3 *Poster Ranking*

An independent-samples t-test analysis was conducted to compare Poster Ranking between students in the two conditions. There were no significant differences in Poster Ranking between students in the Choose (M = 1.71, SD = 1.02) and the Receive conditions (M = 1.63, SD = 1.07); $t(96) = .39$, $p = .70$. Figure 12 shows the mean Poster Ranking plotted by Critical Feedback for each of the two conditions, suggesting that participants who are assigned lower levels of critical feedback outperform those who choose the same levels of critical feedback, but that there is a crossover in learning outcomes after participants encounter two critical feedback levels. This indicates that too much critical feedback may potentially hurt the learning outcomes of participants who are assigned feedback, while the opposite trend seems to occur for participants who choose their feedback.
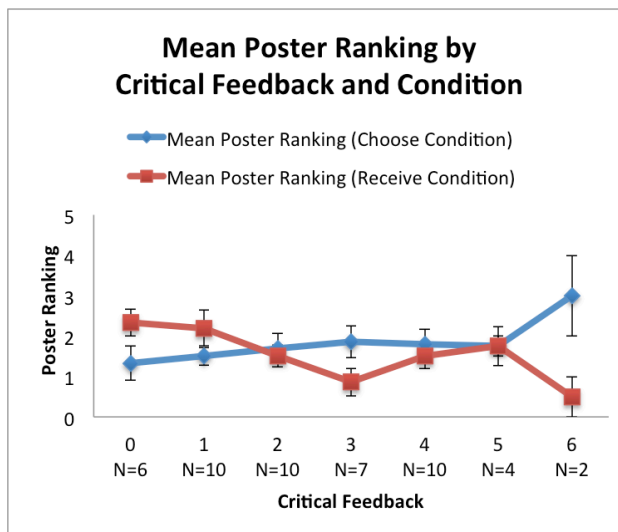
Figure 12. Mean Poster Ranking by critical feedback and condition.

### 6.2.2.4 Design Duration

An independent-samples t-test analysis was conducted to compare Design Duration between conditions. There were no significant differences in Design Duration between students in the Choose (M = 8.72 minutes, SD = 3.28) and the Receive conditions (M = 7.54 minutes, SD = 3.84); $t(96) = 1.63$, $p = .11$. Figure 13 shows the mean Design Duration measured in seconds plotted by Critical Feedback, suggesting that time on task tends to increase with higher levels of critical feedback only when participants choose their feedback and that it does not seem to be affected by critical feedback when participants are assigned their feedback.
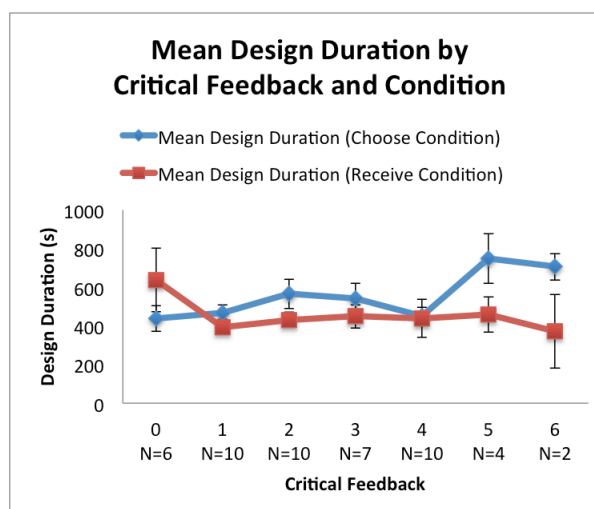


Figure 13. Mean Design Duration by critical feedback and condition.

### 6.2.3 Are there any enjoyment differences between choosing and receiving feedback?

An independent-samples t-test analysis was conducted to compare Enjoyment between conditions. Results show that students in the Choose condition (M = 4.03, SD = .80) enjoyed designing posters statistically significantly more than students in the Receive condition (M = 3.45, SD = .91); $t(81) = 2.99$, $p < .01$. However, the examination of Figure 14, which shows the mean Enjoyment plotted by critical feedback for each condition, suggests that enjoyment is hardly affected by the amount of critical feedback, except at the high end where two participants chose (and were being assigned, respectively) only critical feedback across the game.
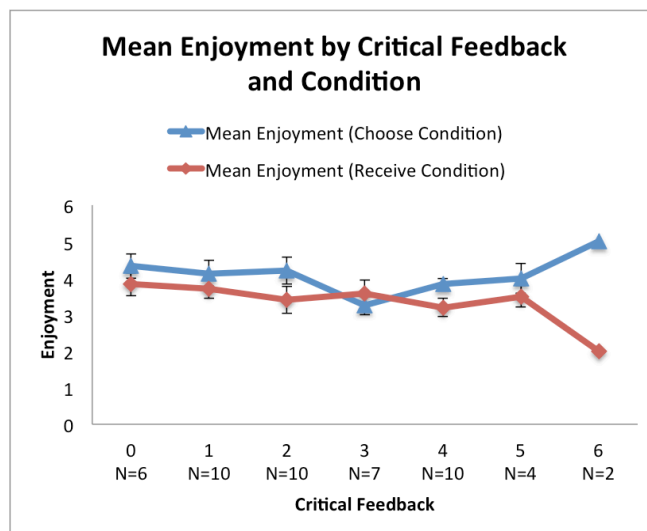
**Figure 14. Mean Enjoyment by critical feedback and condition.**

The data also revealed that, in the Choose condition, 76% of the students reported 4 and 5 levels of enjoyment, as opposed to only 51% of the students in the Receive condition. Moreover, a Spearman correlation analysis revealed that Enjoyment is positively associated with Design Duration ($rho = .37$, $p < .05$) only in the Choose condition.

## 6.3 Discussion, Limitations, and Future Work

### 6.3.1 Critical feedback and performance outcomes

Critical Feedback and Revision were strongly correlated and students improved their poster design performance as they played the game, regardless of condition. Moreover, the more the students chose critical feedback and revised, the better they performed on the poster design task. This is consistent with research reporting that providing children with a choice led to better information-seeking performance in a text search task (Reynolds & Symons, 2001). Concomitantly, the more the students were assigned critical feedback, the worse they performed on the post-test. However, the graphical representation of poster ranking by critical feedback per condition suggests that critical feedback could aid performance depending on the amount of critical feedback assigned. Participants in the Choose condition increase their learning outcomes (performance reached significance, but learning did not) with higher levels of critical feedback (i.e., three or more). In contrast, participants in the Receive condition experience a decline in learning outcomes (learning reached significance, but performance did not) with higher levels of critical feedback. Specifically, students who choose higher levels of critical feedback perform significantly better and learn slightly more than students who choose lower levels of critical feedback, but students who are assigned higher levels of critical feedback learn significantly less and perform slightly less than students who are assigned lower levels of critical feedback. A limitation of this study is the reduced amount of data points for each critical feedback value. For example, a situation in which students choose only critical or only confirmatory feedback is quite rare. More data will be collected in subsequent studies to better understand how students' performance relates to each amount of critical feedback.

Overall, results showed that critical feedback is associated with better performance when students choose their feedback valence, but with worse performance when they are assigned their feedback valence. One possible explanation for this outcome is that choice can be a source of motivation, which may lead students to engage more with their learning. For example, high-school students who chose which of their homework assignments to complete outperformed their peers who were not given this choice (Patall, Cooper, & Wynn, 2010). The next section provides a more general explanation for these results.

### 6.3.2 Outcome differences between choosing and receiving feedback

This research hypothesized that students who chose critical feedback performed better than students who were assigned the same levels of critical feedback. Results revealed no effect of feedback choice (i.e., condition), as there were no differences between conditions in the choice to revise, the performance on the posters and on the post-test, and the time spent designing posters. These results indicate that no underlying variable (e.g., mindset) drives the effect of critical feedback, since assigning the same amount of feedback leads to the same results as other factors that may cause students to choose critical feedback. A limitation of this study is that no additional information, such as students' theories of intelligence, was collected from the participants. For instance, students who display an entity belief are expected to revise and learn less than their incremental belief peers who chose or received same levels of critical feedback, given their lessened attention to critical feedback (Mangels, Butterfield, Lamb, Good, & Dweck, 2006). A further limitation of this study could serve as another explanation for this outcome: both game versions consisted of two rounds and five minutes for each poster design. Future research will include one more round in each condition and provide more time per round to assess students' performance in both conditions and perhaps note differences.

A limitation stemming from the yoked study design is that, by having a choice over their feedback, participants in the Choose condition may self-select into levels of critical feedback. Specifically, they may employ strategies that could help them perform and learn optimally (i.e., choosing critical feedback more often than confirmatory feedback). In contrast, such strategies may not necessarily be optimal for participants in the Receive condition who may in fact perform worse and learn less given the same amount of critical feedback as their Choose condition counterparts. However, this self-selection explanation cannot be generally true, because findings revealed an interaction of condition with critical feedback predicting learning outcomes, meaning that condition moderates the relation between critical feedback and learning outcomes. The interaction shows that once participants encounter critical feedback more often than confirmatory feedback, those who choose feedback outperform those who are assigned feedback. Specifically, when they encountered the same higher levels of critical feedback (i.e., three or more, equivalent with encountering critical feedback at the same rate or more often than confirmatory feedback across the game), participants who chose feedback performed slightly better and learned significantly more than those who were assigned feedback, but there was no difference between conditions in learning outcomes when students encountered lower levels of critical feedback (i.e., less than three, equivalent with encountering confirmatory feedback more often than critical feedback across the game).

An explanation for the interaction effect is that students who choose higher levels of critical feedback (i.e., choose critical feedback at the same rate or more often than confirmatory feedback) may be higher achievers, while students who are assigned the same higher levels of critical feedback may be a mix of higher and lower achievers. Thus, we would expect the Receive condition participants to perform worse and learn less than Choose condition participants in this case, which is what the results show for higher levels of critical feedback. Using the same argument, students in the Choose condition who opt for lower levels of critical feedback (i.e., less than three or encountering confirmatory feedback more often than critical feedback) may be lower-achieving students who would, therefore, generally perform and learn at lower levels than higher-achieving students. In contrast, Receive condition participants who are assigned the same lower levels of critical feedback (i.e., less than three) may be sampled from both higher-achieving and lower-achieving students. Thus, it is expected that, for the same lower levels of critical feedback, Receive participants perform better and learn more than their Choose condition counterparts. Although Receive participants slightly outperformed Choose participants for lower levels of critical feedback, this difference was not statistically significant.

However, if the differences between conditions were only due to Choose participants' self-selection into levels of feedback and not also due to critical feedback, the graph representing the mean learning outcomes in Figure 9 would show a flat line for the Receive condition participants (i.e., these students would perform and learn the same, no matter how much critical feedback they were assigned). This is not the case, as Receive participants perform slightly less and learn significantly less with higher than with lower levels of critical feedback. This indicates that, when critical feedback is assigned, higher levels (i.e., three or more) can hurt both performance and learning, while the opposite effect is detected when critical feedback is chosen. Further research will investigate the interaction between critical feedback and condition in predicting learning outcomes by examining other factors that may influence the differential processing of the same amount of critical feedback depending on condition.

### 6.3.3 Enjoyment differences between choosing and receiving feedback

Students in the Choose condition enjoyed designing posters significantly more than those in the Receive condition. The graph representing enjoyment per condition for every critical feedback value indicates that this effect could simply stem from the existence of choice, rather than from the actual valence of choice. However, more data is necessary to be able to significantly compare the outcomes of the two conditions for each amount of critical feedback ranging from zero to six. This is consistent with research showing that choices can promote a sense of autonomy that could be highly motivating for the students (Deci & Ryan, 1985). Although greater perceived autonomy is associated with higher levels of enjoyment and intrinsic motivation (Reeve, Bolt, & Cai, 1999), choice is not always a motivator (Katz & Assor, 2007). The finding that enjoyment correlates with the time the students took to design the posters only in the Choose condition supports this motivational hypothesis and is consistent with a previous Posterlet study that compared the outcomes of choosing versus receiving the same amount and order of critical feedback for Mechanical Turk adults (Cutumisu & Schwartz, 2016a).

The majority of the students reported high levels of enjoyment, which is encouraging for a game that is ultimately an assessment. Thus, games such as Posterlet could be enjoyable assessment environments for college students, especially when they provide feedback valence choices.

## 6.4 Conclusion

This experiment presented a preliminary comparison of the effect of choosing versus receiving feedback on college students' performance. Results revealed that when students chose higher levels of critical feedback, they significantly outperformed on both learning outcomes the students who received the same amount of critical feedback. The choice to revise is positively associated with performance, but only when students choose their feedback. Students who choose their feedback enjoy designing posters significantly more than students who receive their feedback. Moreover, students who have a choice regarding their feedback spend time designing posters directly proportionally to their enjoyment of designing posters. Future research in this area needs to address the theoretical underpinnings of feedback valence and choice, and further examine the impact of motivation, enjoyment, and critical feedback on students' performance and learning.

# 7 Educational Implications

This research indicates that the learning environment is important for performance and feedback retention. The results have implications for the design of digital assessment and learning environments that enable both giving and receiving feedback to maximize learning. Students may benefit from short, fun, low-stakes, and easily-administered game-based learning and assessment environments where they can engage proactively with feedback to improve their performance and their memory for the feedback content. They can also learn more about how to give and receive effective feedback through their exposure to such environments. In previous research, we showed that design thinking strategies, such as seeking critical feedback, can be taught and they can help students improve their performance (Conlin, Chin, Blair, Cutumisu, & Schwartz, 2015). Instructors may also benefit from assessment environments that integrate the measurement of both students' learning choices (e.g., willingness to seek critical feedback and to revise) and students' learning outcomes to evaluate their impact on students' memory for feedback. Researchers can administer such automated assessments of students' learning strategies at scale, even for open-ended, unstructured, creative tasks such as poster design to evaluate different programs of instruction, as well as to collect and assess behavioural measures in a more objective manner.

One lesson that can be taken away from these studies is that having a choice over the valence of feedback seems to play an important role for performance, the choice to revise, time on task, and enjoyment for college students, as well as for the memory for feedback of middle-school students. Thus, feedback valence choice (i.e., choice between confirmatory and critical feedback) should be considered as an important feature in instructional environments to maximize feedback's effectiveness in impacting performance, learning, and retention for these populations. A feedback choice-rich learning and assessment environment could also lead to student enjoyment and more time spent on task. For example, in a physical activity intervention drawing on self-determination theory (Deci & Ryan, 1985), students who were taught by autonomy-supportive teachers participated more frequently in leisure-time physical activities than students who were taught by less autonomy-supportive teachers (Chatzisarantis & Hagger, 2009). In the current study, even though no

significant differences in learning and performance outcomes were found overall between choosing and receiving feedback, students who chose higher levels of critical feedback significantly outperformed students who were assigned the same amount of critical feedback. Thus, existing assessment environments that do not provide choices could integrate dynamic ways of adjusting the amount of critical feedback available to students and evaluate their performance in each case.

Even though Posterlet is a relatively short 10-15 minute assessment of students' learning behaviours, it provides an indication of the learning strategies that students employ when they engage with an open-ended poster design task. Previous research revealed significant positive correlations between students' choices in Posterlet and students' academic achievement measured by standardized tests of Mathematics, Reading, and Science (Cutumisu, Blair, Chin, & Schwartz, 2015; Cutumisu, Blair, Chin, & Schwartz, 2016). Moreover, recent research suggests that performance on the first game round predicts students' learning choices across the game (Cutumisu & Schwartz, 2017). Thus, even a short game like Posterlet is a viable assessment of students' behaviours and it could be employed easily and quickly as a diagnostic tool for program evaluation.

Our results add to the feedback literature by clarifying the relations among feedback valence choice, revision choice, learning, performance, and memory for feedback. These findings could help explain why some forms of feedback are more effective than others and, thus, they may also aid researchers in gaining insights into the mechanisms of feedback processing and recall, and in comparing different feedback interventions.

# Acknowledgement

# References

Ackerman, D. S., & Gross, B. L. (2010). Instructor feedback: How much do students really want? *Journal of Marketing Education, 32*(2), 172-181.

Anseel, F., Beatty, A. S., Shen, W., Lievens, F., & Sackett, P. R. (2015). How are we doing after 30 years? A meta-analytic review of the antecedents and outcomes of feedback-seeking behavior. *Journal of Management, 41*(1), 318-348.

Ashford, S. J., De Stobbeleir, K., & Nujella, M. (2016). To Seek or Not to Seek: Is That the Only Question? Recent Developments in Feedback-Seeking Literature. *Annual Review of Organizational Psychology and Organizational Behavior, 3*, 213-239. Annual Reviews.

Askew, S., & Lodge, C. (2000). Gifts, ping-pong and loops-linking feedback and learning. In S. Askew (Ed.), *Feedback for learning* (1st ed.). pp. 1-17. London, UK: Routledge.

Attali, Y. (2015). Effects of multiple-try feedback and question type during mathematics problem solving on performance in similar problems. *Computers & Education, 86*, 260-267.

Bangert-Drowns, R., Kulik, C. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61*(2), 213-238.

Beaumont, C., O'Doherty, M., & Shannon, L. (2011). Reconceptualising assessment feedback: A key to improving student learning? *Studies in Higher Education, 36*(6), 671-687. doi:10.1080/03075071003731135.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7-74.

Bolzer, M., Strijbos, J., & Fischer, F. (2015). Inferring mindful cognitive-processing of peer-feedback via eye-tracking: Role of feedback-characteristics, fixation-durations and transitions. *Journal of Computer Assisted Learning, 31*, 422-434. doi:10.1111/jcal.12091.

Bounds, R., Bush, C., Aghera, A., Rodriguez, N., Stansfield, R. B., & Santen, S. A. (2013). Emergency medicine residents' self-assessments play a critical role when receiving feedback. *Academic Emergency Medicine, 20*(10), 1055-1061.

Brummelman, E., Thomaes, S., de Castro, B. O., Overbeek, G., & Bushman, B. J. (2014). "That's Not Just Beautiful - That's Incredibly Beautiful!" The Adverse Impact of Inflated Praise on Children With Low Self-Esteem. *Psychological Science, 25*(3), 728-735. Sage Publications.

Burnett, P. C., & Mandel, V. (2010). Praise and feedback in the primary classroom: Teachers' and students' perspectives. *Australian Journal of Educational & Developmental Psychology, 10*, 145-154.

Butler, A. C., Karpicke, J. D., & Roediger, H. L. (2008). Correcting a metacognitive error: Feedback increases retention of low-confidence correct responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(4), 918-928.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*(3), 245-281.

Butterfield, B., & Metcalfe, J. (2006). The correction of errors committed with high confidence. *Metacognition and Learning, 1*(1), 69-84.

Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education, 36*(4), 395-407.

Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher Education, 31*(2), 219-233.

Chai, L. (2003). To have or have not: An examination of feedback, learner control knowledge type in online learning. *Proceedings of the 36th Annual Hawaii International Conference on System Sciences,* 10 pp.

Chatzisarantis, N. L. D., & Hagger, M. S. (2009). Effects of an intervention based on self-determination theory on self-reported leisure-time physical activity participation. *Psychology and Health, 24*(1), 29-48.

Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review, 24*(2), 205-249.

Cody, M. W., & Teachman, B. A. (2010). Post-event processing and memory bias for performance feedback in social anxiety. *Journal of Anxiety Disorders, 24*(5), 468-479.

Conati, C., Jaques, N., & Muir, M. (2013). Understanding attention to adaptive hints in educational games: An eye-tracking study. *International Journal of Artificial Intelligence in Education, 23*(1-4), 136-161.

Conlin, L. D., 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. *Proc.of the 122nd American Society for Engineering Education (ASEE),* Seattle, WA. 14-17.

Critcher, C. R., Dunning, D., & Armor, D. A. (2010). When self-affirmations reduce defensiveness: Timing is key. *Personality and Social Psychology Bulletin, 36*(7), 947-959. doi:10.1177/0146167210369557.

Cutumisu, M., Blair, K. P., Chin, D. B., & Schwartz, D. L. (2016). Assessing whether students seek constructive criticism: The design of an automated feedback system for a graphic design task. *International Journal of Artificial Intelligence in Education, 27*(3), 419-447. doi:10.1007/s40593-016-0137-5.

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.

Cutumisu, M., Chin, D. B., & Schwartz, D. L. (2014). A game-based assessment of students' choices to seek feedback and to revise. *Proceedings of the 11th International Conference on Cognition and Exploratory Learning in Digital Age (CELDA),* Porto, Portugal. 17-24.

Cutumisu, M., & Schwartz, D. L. (2017). Assessing middle-school students' feedback and revision behavior in a computer-based task. *symposium: Behavior measurement - investigating performance tasks as measures of social and emotional learning skills. American Educational Research Association (AERA),* San Antonio, TX. 1-10.

Cutumisu, M., & Schwartz, D. L. (2016a). Choosing versus receiving feedback: The impact of feedback valence on learning in an assessment game. *Proc. of the 9th International Conference on Educational Data Mining (EDM),* Raleigh, NC. 341-346.

Cutumisu, M., & Schwartz, D. L. (2016b). The relation between high-school students' performance and feedback memory. *London International Conference on Education,* London, UK. 1-6.

Danielmeier, C., & Ullsperger, M. (2011). Post-error adjustments. *Frontiers in Psychology, 2*, 233.

Dauenheimer, D. G., Stahlberg, D., Spreemann, S., & Sedikides, C. (2002). Self-enhancement, self-verification, or self-assessment? The intricate role of trait modifiability in the self-evaluation process. *Revue internationale de psychologie sociale, 15*(3-4), 89-112. Presses universitaires de Grenoble.

Dawson, E., Gilovich, T., & Regan, D. T. (2002). Motivated reasoning and performance on the wason selection task. *Personality and Social Psychology Bulletin, 28*(10), 1379-1387.

De Stobbeleir, K. E. M., Ashford, S. J., & Buyens, D. (2011). Self-regulation of creativity at work: The role of feedback-seeking behavior in creative performance. *Academy of Management Journal, 54*(4), 811-831.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum Press.

D'Mello, S., Olney, A., Williams, C., & Hays, P. (2012). Gaze tutor: A gaze-reactive intelligent tutoring system. *International Journal of Human-Computer Studies, 70*(5), 377-398.

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.

Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of Educational Research, 83*(1), 70-120. doi:10.3102/0034654312474350.

Fazio, L. K., & Marsh, E. J. (2009). Surprising feedback improves later memory. *Psychonomic Bulletin & Review, 16*(1), 88-92.

Fishbach, A., Eyal, T., & Finkelstein, S. R. (2010). How positive and negative feedback motivate goal pursuit. *Social and Personality Psychology Compass, 4*(8), 517-530.

Fisher, R. A. (1921). On the probable error of a coefficient of correlation deduced from a small sample. *Metron, 1*, 3-32.

Geitz, G., Joosten-Ten Brinke, D., & Kirschner, P. A. (2016). Sustainable feedback: Students' and tutors' perceptions. *The Qualitative Report, 21*(11), 2103.

Gibbs, G., & Simpson, C. (2005). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education,* (1), 3-31.

Green, J. D., Pinter, B., & Sedikides, C. (2005). Mnemic neglect and self-threat: Trait modifiability moderates self-protection. *European Journal of Social Psychology, 35*(2), 225-235.

Gregory, J. B., & Levy, P. E. (2015). *Using feedback in organizational consulting.* Washington, D.C.: American Psychological Association.

Harris, L. R., Brown, G. T. L., & Harnett, J. A. (2015). Analysis of New Zealand primary and secondary student peer- and self-assessment comments: applying Hattie and Timperley's feedback model. *Assessment in Education: Principles, Policy & Practice, 22*(2), 265-281. doi:10.1080/0969594X.2014.976541

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* Abingdon: Routledge, Taylor and Francis Group.

Hattie, J. (2013). *Visible learning for teachers: Maximizing impact on learning* (2nd ed.). London, UK: Routledge.

Hattie, J., & Gan, M. (2011). Instruction based on feedback. In R. Mayer, & P. Alexander (Eds.), *Handbook of research on learning and instruction* (1st ed.). pp. 249-271. New York: Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. doi:10.3102/003465430298487.

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. e6.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52*(12), 1280.

Hounsell, D., Boud, D., & Falchikov, N. (2007). Towards more sustainable feedback to students. In D. Boud, & N. Falchikov (Eds.), *Rethinking assessment in higher education* (1st ed.). pp. 101-113. Abingdon, UK: Routledge London.

Hounsell, D., McCune, V., Hounsell, J., & Litjens, J. (2008). The quality of guidance and feedback to students. *Higher Education Research & Development, 27*(1), 55-67.

Hounsell, D., Slowey, M., & Watson, D. (2003). Student feedback, learning and development. In M. Slowey, & D. Watson (Eds.), *Higher education and the lifecourse* (1st ed.). pp. 67-78. Buckingham: SRHE & Open University Press.

Houtman, F., Castellar, E. N., & Notebaert, W. (2012). Orienting to errors with and without immediate feedback. *Journal of Cognitive Psychology, 24*(3), 278-285.

Houtman, F., & Notebaert, W. (2013). Blinded by an error. *Cognition, 128*(2), 228-236.

Jonsson, A. (2013). Facilitating productive use of feedback in higher education. *Active Learning in Higher Education, 14*(1), 63-76.

Katz, I., & Assor, A. (2007). When choice motivates and when it does not. *Educational Psychology Review, 19*(4), 429.

Kizilcec, R. F., Saltarelli, A. J., Reich, J., & Cohen, G. L. (2017). Closing global achievement gaps in MOOCs. *Science, 355*(6322), 251-252.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance. *Psychological Bulletin, 119*(2), 254-284. doi:10.1037/0033-2909.119.2.254.

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.

Kovacs, M., & Beck, A. T. (1978). Maladaptive cognitive structures in depression. *The American Journal of Psychiatry, 135*(5), 525-533. doi:10.1176/ajp.135.5.525.

Kulkarni, C. E., Bernstein, M. S., & Klemmer, S. R. (2015). PeerStudio: Rapid peer feedback emphasizes revision and improves performance. *Proceedings of the Second (2015) ACM Conference on Learning@ Scale,* Vancouver, Canada. 75-84. doi://dx.doi.org/10.1145/2724660.2724670

LaDonna, K. A., Hatala, R., Lingard, L., Voyer, S., & Watling, C. (2017). Staging a performance: Learners' perceptions about direct observation during residency. *Medical Education, 51*(5), 498-510.

Landis-Lewis, Z., Brehaut, J. C., Hochheiser, H., Douglas, G. P., & Jacobson, R. S. (2015). Computer-supported feedback message tailoring: Theory-informed adaptation of clinical audit and feedback for learning and behavior change. *Implementation Science, 10*(1), 12.

Mahfoodh, O. H. A. (2017). "I feel disappointed": EFL university students' emotional responses towards teacher written feedback. *Assessing Writing, 31*, 53-72.

Mangels, J. A., Butterfield, B., Lamb, J., Good, C., & Dweck, C. S. (2006). Why do beliefs about intelligence influence learning success? A social cognitive neuroscience model. *Social Cognitive and Affective Neuroscience, 1*(2), 75-86.

McDonnell, J., & Curtis, W. (2014). Making space for democracy through assessment and feedback in higher education: Thoughts from an action research project in education studies. *Assessment & Evaluation in Higher Education, 39*(8), 932-948.

Merry, S., Price, M., Carless, D., & Taras, M. (2013). *Reconceptualising feedback in higher education: Developing dialogue with students* (1st ed.). New York, NY: Routledge.

Mulliner, E., & Tucker, M. (2017). Feedback on feedback practice: Perceptions of students and academics. *Assessment & Evaluation in Higher Education, 42*(2), 266-288.

Narciss, S. (2013). Designing and evaluating tutoring feedback strategies for digital learning environments on the basis of the interactive tutoring feedback model. *Digital Education Review,* (23), 7-26.

Notebaert, W., Houtman, F., Van Opstal, F., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: An orienting account. *Cognition, 111*(2), 275-279.

OECD. (2013). *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy.* Paris, France: OECD Publishing. doi:http://dx.doi.org/10.1787/9789264190511-en

Patall, E. A., Cooper, H., & Wynn, S. R. (2010). The effectiveness and relative importance of choice in the classroom. *Journal of Educational Psychology, 102*(4), 896.

Porath, C. L., & Bateman, T. S. (2006). Self-regulation: From goal orientation to job performance. *Journal of Applied Psychology, 91*(1), 185.

Price, M., Handley, K., Millar, J., & O'Donovan, B. (2010). Feedback: All that effort, but what is the effect? *Assessment & Evaluation in Higher Education, 35*(3), 277-289.

Reeve, J., Bolt, E., & Cai, Y. (1999). Autonomy-supportive teachers: How they teach and motivate students. *Journal of Educational Psychology, 91*(3), 537.

Reynolds, P. L., & Symons, S. (2001). Motivational variables and children's text search. *Journal of Educational Psychology, 93*(1), 14.

Rietsche, R., Lehmann, K., Haas, P., & Söllner, M. (2017). The Twofold Value of IT-Based Peer Assessment in Management Information Systems Education. *13th International Conference on Wirtschaftsinformatik (WI),* 1-5. St. Gallen, Switzerland: University of St. Gallen, Institute of Information Management.

Roscoe, R. D., Wilson, J., Johnson, A. C., & Mayra, C. R. (2017). Presentation, expectations, and experience: Sources of student perceptions of automated writing evaluation. *Computers in Human Behavior, 70*(5), 207-221. doi:https://doi.org/10.1016/j.chb.2016.12.076.

Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education, 35*(5), 535-550.

Sanitioso, R., Kunda, Z., & Fong, G. T. (1990). Motivated recruitment of autobiographical memories. *Journal of Personality and Social Psychology, 59*(2), 229.

Sargeant, J., Mcnaughton, E., Mercer, S., Murphy, D., Sullivan, P., & Bruce, D. A. (2011). Providing feedback: Exploring a model (emotion, content, outcomes) for facilitating multisource feedback. *Medical Teacher, 33*(9), 744-749. doi:10.3109/0142159X.2011.577287.

Schwartz, D. L., & Arena, D. (2013). *Measuring what matters most: Choice-based assessments for the digital age.* Cambridge, Massachusetts: MIT Press.

Schwartz, D. L., Lindgren, R., & Lewis, S. (2009). Constructivism in an age of non-constructivist assessments. In T. M. Duffy, & S. Tobias (Eds.), *Constructivist Instruction: Success or Failure?* (1st ed.). pp. 34-61. New York: Routledge/Taylor & Francis Group.

Schwartz, D. L., Tsang, J. M., & Blair, K. P. (2016). *The ABCs of how we learn: 26 scientifically proven approaches, how they work, and when to use them.* US: W.W. Norton & Company.

Sedikides, C., Green, J. D., Saunders, J., Skowronski, J. J., & Zengel, B. (2016). Mnemic neglect: Selective amnesia of one's faults. *European Review of Social Psychology, 27*(1), 1-62.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153-189.

Sitzman, D. M., Rhodes, M. G., & Kornell, N. (2016). The influence of feedback on predictions of future memory performance. *Memory & Cognition, 44*(7), 1102-1113.

Sonnentag, S., & Sparr, J. L. (2008). Feedback and self-regulatory focus: Predicting affective reactions and performance in a field study. *Academy of Management Proceedings, 2008*(1), 1-6. doi:10.5465/AMBPP.2008.33643503.

Soper, D. (2015). Significance of the difference between two correlations calculator [Software]. Retrieved from http://www.danielsoper.com/statcalc

Spataro, P., Mulligan, N. W., & Rossi-Arnaud, C. (2013). Divided attention can enhance memory encoding: The attentional boost effect in implicit memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(4), 1223.

Steinborn, M. B., Flehmig, H. C., Bratzke, D., & Schröter, H. (2012). Error reactivity in self-paced performance: Highly-accurate individuals exhibit largest post-error slowing. *The Quarterly Journal of Experimental Psychology, 65*(4), 624-631.

Story, A. L. (1998). Self-esteem and memory for favorable and unfavorable personality feedback. *Personality and Social Psychology Bulletin, 24*(1), 51-64.

Swallow, K. M., & Jiang, Y. V. (2010). The attentional boost effect: Transient increases in attention to one task enhance performance in a second task. *Cognition, 115*(1), 118-132.

Swallow, K. M., & Jiang, Y. V. (2012). Goal-relevant events need not be rare to boost memory for concurrent images. *Attention, Perception, & Psychophysics, 74*(1), 70-82.

Thurlings, M., Vermeulen, M., Bastiaens, T., & Stijnen, S. (2013). Understanding feedback: A learning theory perspective. *Educational Research Review, 9*, 1-15.

Timms, M., DeVelle, S., & Lay, D. (2016). Towards a model of how learners process feedback: A deeper look at learning. *Australian Journal of Education, 60*(2), 128-145.

Van der Borght, L., Schouppe, N., & Notebaert, W. (2016). Improved memory for error feedback. *Psychological Research, 80*(6), 1049-1058.

van der Helden, J., Boksem, M. A., & Blom, J. H. (2010). The importance of failure: Feedback-related negativity predicts motor learning efficiency. *Cerebral Cortex (New York, N.Y.: 1991), 20*(7), 1596-1603. doi:10.1093/cercor/bhp224.

Wakefield, C., Adie, J., Pitt, E., & Owens, T. (2014). Feeding forward from summative assessment: The essay feedback checklist as a learning tool. *Assessment & Evaluation in Higher Education, 39*(2), 253-262.

Waller, L., & Papi, M. (2017). Motivation and feedback: How implicit theories of intelligence predict L2 writers' motivation and feedback orientation. *Journal of Second Language Writing, 35*, 54-65.

Walsh, B., Nixon, S., Walker, C., & Doyle, N. (2015). Using a clean feedback model to facilitate the learning process. *Creative Education, 6*(10), 953.

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation, 37*(1), 3-14.

Williams, E. F., & Ehrlinger, J. (2017). Failing to learn from feedback. In J. E. Ellingson, & R. A. Noe (Eds.), *Autonomous Learning in the Workplace* (1st ed.). pp. 263-286. London, UK: Routledge.

Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting learners' agentic engagement with feedback: A systematic review and a taxonomy of recipience processes. *Educational Psychologist, 1*, 1-21.

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education, 45*(4), 477-501.