Is screencast feedback better than text feedback for student learning in Higher Education? A Systematic Review

Simon Penn, The Open University, England
Nicola Brown, St Mary’s University, England

Abstract: Although the positive affordances of screencast feedback for student learning have been identified, the Higher Education (HE) system remains tied to electronic written (text) feedback. One factor in this stasis is the lack of robust evidence regarding the relative effectiveness of screencast feedback. This paper addresses this research gap. A systematic review of empirical research examined the effect of the medium of feedback (screencast vs. text) on student learning in HE. Fifteen of 502 studies were selected for analysis. Eleven studies indicated overwhelming positive perceptions of screencast feedback highlighting that it is more personal, supportive, detailed, helpful, and easier to understand in comparison to text feedback. Lecturers also offered overwhelming support for this medium stating that it was more engaging and impactful because it offered more detail, more teaching opportunities and increased the connection between the student and lecturer. Despite positive perceptions, seven studies employing direct measurements of student learning presented inconclusive evidence to determine if screencast feedback is better than text feedback for student learning in HE. Therefore, there is currently insufficient evidence to warrant adopting screencast feedback in HE.

Keywords: Screencast, Feedback, Learning, Higher Education

Introduction

Assignment feedback has the potential to be one of the most influential factors on student learning (Hattie and Timperley 2007). However, according to both students and lecturers, feedback in Higher Education (HE) is failing to achieve this potential (Office for Students 2019; Winstone and Carless 2019). Since the advent of the National Student Survey in 2005, students have consistently reported lower satisfaction scores for assessment and feedback than other areas (e.g., teaching, learning resources and academic support) of the survey (Office for Students, 2019). In addition to student dissatisfaction with the effectiveness of feedback, HE lecturers (i.e., academics involved in teaching, assessment and providing feedback) perceive that feedback is underused and unappreciated by students (Winstone and Carless 2019). To encourage more engagement and greater satisfaction with feedback, many lecturers and researchers have trialled alternative mediums of feedback (e.g., audio, video and screencast) in place of electronic written (text) feedback (Ice et al. 2007; Mahoney, Ajjawi, and Macfarlane 2019; Vincelette and Bostic 2013). Text feedback, typically consisting of annotations and summary comments, remains the dominant feedback medium used in HE (Orlando 2016; Ryan, Henderson, and Phillips 2019). Screencast feedback combines the strengths of text, audio, and video feedback by allowing a lecturer to record their on-screen actions simultaneously with spoken comments during the marking and feedback process. Students have indicated a preference for screencast feedback because it is richer, more detailed and easier to understand than text feedback (Vincelette and Bostic, 2013). However, to encourage lecturers to adopt a new feedback practice, there must be strong evidence that it is superior. A review of the literature is
needed to determine if there is sufficient empirical evidence to warrant replacing text with screencast feedback in HE.

**The Role of Feedback in Student Learning**

In HE, student learning typically signifies that a student has developed from the tuition. For written assignments, student learning can be shown by a progression of grades, assignment revisions, and assessment rubric scores focused on learning objectives (Suskie 2018). These objective and direct measurements of student learning can also be supported by indirect measures including retention rates, graduation rates, student reflections and student perception responses to course evaluations (Suskie 2018). Student satisfaction is also important because a more satisfied student is more likely to engage in learning and persist with studies (York, Gibson and Rankin 2015). However, this indirect measure alone would not signify learning and should be accompanied by more direct methods of learning measurement (Suskie 2018).

Traditionally, feedback has been viewed as a product, delivered by an expert (a lecturer), with the goal/aim of evaluating performance and telling students what they need to do to improve (Boud and Molloy 2013). This telling approach of feedback situates the student as a passive participant in their own learning. However, the social constructivist theory of learning suggests that cognitive development stems from social interaction with others (Vygotsky 1978). Students learn through collaboratively constructing meanings and artefacts together to assimilate understanding (Thurlings et al. 2013). Social interaction with more knowledgeable others allows learners to develop the skills and knowledge to solve more complex problems that they are unable to accomplish individually (Vygotsky 1978). The student becomes more active in their learning and the lecturer’s role is facilitative rather than instructive. Following the social constructivist theory of learning, the role of feedback has been reconceptualized as a social process where students learn through interaction and dialogue with peers and lecturers (Winstone and Carless 2019). Feedback in HE has transcended from a didactic performance evaluation to a means of empowering a student to develop their evaluative skills and academic identity to help them succeed beyond education (Boud and Molloy 2013; Carless 2019). Carless and Boud (2018, p. 1) defined feedback as “a process through which learners make sense of information from various sources and use it to enhance their work or learning strategies”.

For feedback to be ‘effective’ in developing student learning, it must encourage students to engage in the feedback process and stimulate them to enhance their ability to evaluate and learn from their work (student agency) and understand and utilize feedback (feedback literacy) (Carless and Boud 2018; O’Donovan, Rust, and Price 2016). This more agentic approach in the feedback process includes collecting, reading, understanding, and acting on feedback in addition to identifying learning opportunities (Price, Handley and Millar 2011). Due to the multiple points of engagement required during the feedback process, numerous variables can influence student engagement. The student’s perception of feedback, their understanding of the value of feedback, their ability to implement and act on feedback and their motivation or willingness to receive and develop from feedback are all potential barriers for engagement with feedback (Winstone et al. 2017). The medium used to deliver feedback is a potential mechanism
for reducing these barriers (Winstone et al. 2017). Therefore, determining the most effective method of delivering feedback to promote engagement, student feedback literacy and create more self-regulated learners is needed.

Multiple reviews have assessed the effectiveness of feedback in HE and suggested that for feedback to provide an impact on student learning, it should be timely, easy to understand, dialogic in nature and useful (Lefevre and Cox 2017; Rand 2017; Dawson et al. 2019; Gibbs and Simpson 2004; Evans 2013; Li and De Luca 2014). Although students and lecturers suggest that feedback should be received quickly (Lefevre and Cox 2017; Dawson et al. 2019), a delay in feedback benefits the constructivist approach to feedback by allowing time for reflection and self-regulation (Mullet et al. 2014). Yet, if the feedback is received too late, it may lose its relevance and students may be less likely to engage with it (Mulliner and Tucker 2017; Dawson et al. 2019). If feedback language is too academic or authoritative, a student is less likely to understand and engage with the feedback (Jonsson 2012; O’Donovan, Rust, and Price 2016). An authoritative tone is very formal and indicates a telling approach placing the student as a passive receiver of information. Contrastingly, a more conversational, dialogic tone appears more personal and follows the social constructivist theory of learning by encouraging cooperation in the construction of learning (Winstone and Carless 2019).

The Medium of Feedback

Text feedback has become a mainstay of feedback in HE (Orlando 2016; Ryan, Henderson, and Phillips 2019). However, the efficacy of text feedback in promoting student engagement and development has been questioned (Cavanaugh and Song 2014; Nicol 2010). Written annotations on student coursework are useful for correcting errors or highlighting strengths. However, a didactic directive approach of error correction is unlikely to encourage independent learning (Brown and Glover 2006). Furthermore, written annotations often fail to explain an error/strength or justify the correction, which are vital learning strategies (Brown and Glover 2006). Students are also often overwhelmed by the number of text comments, fail to engage with text feedback, or misunderstand the lecturer’s comments (Mathieson 2012; Vincelette and Bostic 2013). Text feedback is commonly provided in an academic discourse which requires a good level of assessment and feedback literacy to interpret (Leibold and Schwarz 2015; Orlando 2016). Therefore, text feedback can be delivered promptly, but the formal, didactic approach may reduce the impact and usefulness of this medium.

To address the limitations of text feedback, many lecturers and researchers have trialled audio, video and screencast feedback (Parkes and Fletcher 2019; Vincelette and Bostic 2013). Audio feedback, the recording of a lecturer’s voice explaining positive and constructive elements of the student’s work, is more detailed, personal and easier to understand than text feedback (Ice et al. 2007; Lunt and Curran 2010). More detail and in-depth feedback can be provided because a lecturer can talk much quicker than they can type (Ice et al. 2007; Lunt and Curran 2010). Additionally, the greater depth of detail afforded by this audio feedback approach provides more teaching opportunities (i.e., explanations of comments and examples of good practice) and more supportive feedback (Parkes and Fletcher 2019). Audio feedback is also considered easier to
understand because there is an increased ability to understand nuances in the lecturer’s voice that may be lost in written communication (Parkes and Fletcher 2019). Furthermore, as verbal communication cues are available through audio feedback, this medium is considered more informal, conversational and personal, thus increasing the connection between assessor and student (Ice et al. 2007; Parkes and Fletcher 2019).

In addition to the benefits of audio feedback, video feedback, where the lecturer records their face discussing feedback (i.e., talking head), offers further visual communication cues (e.g., gestures, expressions) which may increase the supportive personal nature of the feedback (Mahoney, Ajjawi, and Macfarlane 2019). Seeing the lecturer’s facial expressions and hearing their voice resembles face-to-face feedback and increases the dialogic nature of feedback, despite no dialogue taking place (Harper, Green, and Fernandez-Toro 2018; Mahoney, Ajjawi, and Macfarlane 2019). Following the social constructivist theory of learning, creating a dialogic relationship should improve the evaluative nature of the student and increase student engagement with feedback (Carless and Boud 2018). Despite the recognized benefits of audio and video feedback, a common limitation of these mediums is that when the lecturer is explaining a feedback point, the student may have difficulty determining the exact point the feedback is referring to (Henderson and Phillips 2015; Parkes and Fletcher 2019). To eliminate this limitation, screencast feedback enables lecturers to record on-screen actions (text feedback) simultaneously with spoken comments (verbal feedback) and can include webcam footage of the lecturer discussing the feedback (video feedback).

Screencast feedback offers the unique opportunity to incorporate the benefits of text, audio and video feedback. Students receive video files with running commentary of their coursework being highlighted, amended, and discussed (Brick and Holmes 2008). The additional visual presentation of the student’s work afforded by screencast feedback is considered to provide more extensive and informative detail than text, audio and video feedback (Harper, Green, and Fernandez-Toro 2018; Mathieson 2012). This feedback approach gives the presence of the lecturer and is therefore considered more personable, richer and easier to understand (Harper, Green, and Fernandez-Toro 2018; Vincelette and Bostic 2013). Despite these potential benefits of screencast feedback identified by research, HE academia remain reluctant to change to this new form of feedback (Orlando 2016; Ryan, Henderson, and Phillips 2019). This reluctance may stem from a high number of action research and case study studies which have limitations including small sample sizes, no control groups and inefficient measures to determine a significant effect (Alvira 2016; Denton 2014; Ghosn-Chelala and Al-Chibani 2018; Marriott and Teoh 2012; West and Turner 2016). In 2019, Mahoney, Ajjawi and Macfarlane provided a qualitative synthesis of video feedback (i.e., talking head, screencast, and combination screencast). However, screencast feedback was not evaluated independently from video feedback and the synthesis was not restricted to empirical evidence. To provide robust evidence on the effectiveness of screencast feedback, alternative modes of video feedback that do not include the affordances of screencast feedback (i.e., talking head) must be excluded from the analysis.

To determine if screencast feedback is more effective than text feedback in enhancing student learning in HE, a systematic review of current empirical evidence is required to provide academics with a robust critical evaluation and research-informed summary. The purpose of this systematic review was to answer the following research question: Is
screencast feedback more effective than text feedback in enhancing student learning in HE?

Materials and Method

Procedure

Following the procedure outlined by Petticrew and Roberts (2006) for systematic reviews in the Social Sciences, three stages were followed. In stage one, the research question, databases, search terms and inclusion criteria were identified before performing the literature search to identify and select eligible studies for inclusion. In stage two, data were collated consistently across each study (e.g., characteristics, methodology, results) using a data extraction form. The data form also assessed study rigor to determine the strength of each study’s conclusion. In stage three, a systematic qualitative synthesis of the studies’ conclusions was conducted to answer the study research question.

Databases and Search Strategy

A systematic search of the literature was performed using the database EbscoHost incorporating the most searched databases for educational publications: Academic Search Complete, British Education Index (BEI), Education Abstracts, Education Research Complete (ERC) and Education Resources Information Center (ERIC). A thesaurus and scoping search were used to identify the most relevant keywords for screencast, feedback and HE. To encapsulate all available studies researching screencast feedback, the term ‘student learning’ was not included in the search process. To ensure that the search identified research studies with a clear focus on screencast feedback, only study titles, abstracts and keywords were searched using the search string (screencast* OR “screen cast*” OR “audio-visual” OR “screen capture”) AND (student* OR "higher education" OR college OR universit* OR "post-secondary" OR postsecondary OR undergrad* OR postgrad* OR tertiary) AND (feedback OR “feedback”). The wildcard symbol (*) was used to retrieve studies containing any derivational affixes of the words ‘screencast’, ‘student’ and ‘university’. The term ‘audio-visual’ was included because it is commonly used to describe feedback provided by a combination of audio and visual material like screencasting.

Although ScreenCam, the first computer program capable of screencasts, appeared in 1994, the ‘screencasting’ term was not coined until 2004 (Udell 2004). Additionally, a scoping search revealed that the first screencasting based article was published in 2005, although this was a magazine article (Udell 2005). Therefore, the search was limited from 2005 using the year filter on EbscoHost. To ensure that only peer reviewed academic journal articles were identified, the search was also limited to ‘Scholarly (Peer Reviewed) Journals’ and ‘Academic Journals’ source types. Therefore, unpublished dissertations and conference proceedings that do not undergo a peer review appraisal were excluded. To maximize the scope of the literature search, no geographical limit was placed on the search. However, only studies published in English language were included. Searching was concluded in August 2020.
Inclusion Criteria

Studies were only eligible for inclusion if they included screencast feedback on an individual assignment and a direct measurement (grade score gains, assignment revisions, and assessment rubrics), indirect measurement (i.e., grades, retention/graduation rates, student reflections or perceptions), or influential factor (i.e., tutor perception and feedback delivery) on student learning. To narrow the search and focus on the effect of screencast feedback on written coursework, studies were only included if the feedback delivered in studies needed to relate to a form of written HE assessment (formative or summative) to enable comparison between studies. During the infancy of screencast research, many action research studies were completed and lacked a comparative or control group (Alvira 2016; Denton 2014; Ghosn-Chelala and Al-Chibani 2018; Vincelette and Bostic 2013). Using an action research or case study approach with only one population group, assignment and environment limit the validity, reliability, and generalizability of the results. Therefore, to allow the study to determine the effect of screencast feedback, only more rigorous studies that included a comparative or control group were included in the review.

Selection Process

All studies identified were exported to Mendeley Desktop (Foeckler, Henning, and Reichelt, 2019). Then, using the inclusion criteria, the lead author completed three rounds of screening. The primary screening round assessed study titles against the inclusion criteria. The second round of screening followed the same criteria to assess abstracts to confirm study eligibility. If study suitability could not be confirmed by screening stages one or two, the full text was obtained and screened against the inclusion criteria. If all inclusion criteria were met, the study was included in the review. To assess inclusion criteria reliability, the first five articles identified from EBSCOhost were double screened by the lead and second author. The two authors discounted the same two studies at screening round one (Agerskov et al. 2017; Ahrens et al. 2019), the same two studies at screening round two (Alharbi 2017; Alvira 2016) and included the same article for consideration (Ali 2016).

Data Extraction and Appraisal of Studies

For every article selected, a data extraction form was completed recording information on the population (type, age, no.), method (screencast program, timings, type/s of feedback, type/s of assessment), results and conclusions. The data form also assessed the methodological quality of each study, using the revised 2018 Mixed Methods Appraisal Tool (MMAT; Hong et al. 2018). The MMAT allowed critical appraisal of the different types of study design (qualitative, quantitative, and mixed methods) included in the review; this provided an indication of quality rather than an overall score for comparison. Initially, two screening questions are answered before rating each study on five criteria related to their study design as either ‘yes’, ‘no’ or ‘can’t tell’. Study quality score can range from failing to meet any of the five criteria (zero) to meeting all five criteria (five). To assess the reliability of researchers coming to the same conclusions from the data extraction form, the lead and second researchers independently extracted the data and appraised five of the studies. After data extraction
and appraisal, both researchers highlighted the same strengths and limitations of the five studies and agreed on the quality of each study. The lead researcher then performed data extraction and appraisal of the remaining 10 studies.

![Review flowchart indicating the study selection process](source: Penn)

**Results**

The primary systematic literature search of the common educational databases identified 502 articles for review. After removing 192 duplicates, a total of 310 articles were screened to assess their suitability for inclusion in the review. In the first round and second of screening, a total of 275 articles were excluded. Full texts of the remaining 34 articles were reviewed and 19 articles were excluded for no screen casting intervention, no control or comparative group and feedback provided to a group rather than individualized (see Figure 1). A final selection of 15 articles was included for data analysis.

Two of the 15 studies included the results of two studies within one article. Two studies were highlighted for failing to include a comparison group in one part of their studies and the results from these sections were excluded from review (Boone and Carlson 2011; Jones, Georghiades, and Gunson 2012). The first study by Boone and Carlson (2011) did not include a comparative or control group and was excluded from the review. Their second study did include a comparative group and was therefore
included (Boone and Carlson 2011). Similarly, no comparative or control group was included in the second study by Jones, Georgiades, and Gunson (2012). Therefore, only their first study was included in the review.

**Study Appraisal**

The MMAT (Hong et al. 2018) was used to critically appraise the methodological quality of each study. To assess inter-rater agreement, both researchers independently appraised five studies (Boone and Carlson 2011; Cunningham 2019a, 2019b; Mathieson 2012; Silva 2012). Using SPSS statistical package version 22 (IBM Corp 2013), Cohen's κ determined there was substantial agreement (κ = .714, p < 0.05) between the two researchers on the methodological quality of each study (Landis and Koch 1997). Minimal discrepancies were resolved through discussion before the second researcher appraised all remaining studies. Eleven studies were categorized as high quality after meeting all core criteria for their study design (Ali 2016; Anson et al. 2016; Bakla 2020; Cavaleri et al. 2019; Cunningham 2019a, 2019b; Elola and Oskoz 2016; Grigoryan 2017; Harper, Green, and Fernandez-Toro 2018; Jones, Georgiades, and Gunson 2012; Kim 2018). The remaining four studies did not meet all quality criteria indicating low methodological quality (Hong et al. 2018). The major limitation of all four studies was that the descriptive data was unsubstantiated by statistical analysis (Boone and Carlson 2011; Hope 2011; Mathieson 2012; Silva 2012). For example, Boone and Carlson (2011) only provided descriptive data and no statistical analysis was performed to determine if final grades significantly differed between groups. Additionally, the demographic details of the two comparative groups were not provided to clarify the homogeneity of groups, highlighting the lower quality of this study identified by the MMAT appraisal.

**Study Characteristics**

Included studies were published between 2011 and 2020. Most were completed in the United States of America (n = 8), and the remaining studies conducted in the United Kingdom (n = 3), Turkey (n = 1), Australia (n = 1), Egypt (n = 1) and Korea (n = 1) (see Table 1). Six studies were conducted in universities offering traditional face to face (F2F) tuition, four were performed in a distance learning setting, two incorporated a mixture of F2F, distance-learning or blended learning and three did not report their style of teaching. All undergraduate (UG) levels (1st year (freshmen) to 4th year (senior)) were represented. Sample sizes ranged from 3 to 156 (median = 33) and a variety of study subjects were sampled across the studies (see Table 1).

**Study Interventions**

Table 2 shows that many studies compared text feedback to a group receiving both text and screencast feedback (n = 6), screencast feedback with highlighted text or minimal marginal comments (n = 3) or to screencast feedback alone (n = 6). The primary method of providing text feedback was Microsoft Word (n = 13) using either the track change function, the comment bubbles, or a combination of both. Jing, a screencast software
program that limits screencasts to five minutes was the most popular program for providing screencast feedback (n = 8). The average screencast length was 3 to 15 minutes, although not all studies reported this data. No study reported on the use of a webcam to show the instructors face. One study identified that within the screencast, students were introduced to a program called Grammarly (Grammarly 2020) which helps correct and teach grammatical errors within the text (Kim 2018).

Table 1: Population Demographics in Studies Selected for Analysis

<table>
<thead>
<tr>
<th>Author/s (Year)</th>
<th>Location</th>
<th>Sample Size</th>
<th>UG Level</th>
<th>Subject</th>
<th>Mode of teaching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ali (2016)</td>
<td>Egypt</td>
<td>63 students</td>
<td>1st year</td>
<td>Management Sciences</td>
<td>F2F</td>
</tr>
<tr>
<td>Anson et al. (2016)</td>
<td>USA</td>
<td>141 students</td>
<td>1st- 3rd year</td>
<td>Sciences, Social Sciences and Humanities</td>
<td>F2F and Distance</td>
</tr>
<tr>
<td>Bakla (2020)</td>
<td>Turkey</td>
<td>33 students</td>
<td>1st year</td>
<td>English</td>
<td>NR</td>
</tr>
<tr>
<td>Boone and Carlson (2011)</td>
<td>USA</td>
<td>157 students</td>
<td>Not reported</td>
<td>Developmental Writing</td>
<td>Distance</td>
</tr>
<tr>
<td>Cavaleri et al. (2019)</td>
<td>Australia</td>
<td>20 students</td>
<td>1st year</td>
<td>Not reported</td>
<td>F2F and Distance</td>
</tr>
<tr>
<td>Cunningham (2019a)</td>
<td>USA</td>
<td>3 lecturers</td>
<td>Not reported</td>
<td>Academic ESL writing</td>
<td>NR</td>
</tr>
<tr>
<td>Cunningham (2019b)</td>
<td>USA</td>
<td>12 students</td>
<td>Not reported</td>
<td>Academic ESL writing</td>
<td>F2F</td>
</tr>
<tr>
<td>Elola and Oskoz (2016)</td>
<td>USA</td>
<td>4 students</td>
<td>3rd-4th year</td>
<td>Spanish</td>
<td>F2F</td>
</tr>
<tr>
<td>Grigoryan (2017)</td>
<td>USA</td>
<td>50 students</td>
<td>1st year</td>
<td>Composition courses</td>
<td>Distance</td>
</tr>
<tr>
<td>Harper, Green, and Fernandez-Toro (2018)</td>
<td>UK</td>
<td>54 students 9 lecturers</td>
<td>Not reported</td>
<td>Spanish and German</td>
<td>Distance</td>
</tr>
<tr>
<td>Hope (2011)</td>
<td>UK</td>
<td>90 students</td>
<td>2nd year</td>
<td>Human Genetics</td>
<td>F2F</td>
</tr>
<tr>
<td></td>
<td></td>
<td>55 students</td>
<td>1st year</td>
<td>Forensic Science</td>
<td></td>
</tr>
<tr>
<td>Jones, Georghiades, and Gunson (2012)</td>
<td>UK</td>
<td>75 students 20 lecturers</td>
<td>1st-3rd year</td>
<td>Information Technology and Business</td>
<td>F2F</td>
</tr>
<tr>
<td>Mathieson (2012)</td>
<td>USA</td>
<td>15 students</td>
<td>Post-Professional</td>
<td>Health Sciences</td>
<td>Distance</td>
</tr>
<tr>
<td>Silva (2012)</td>
<td>USA</td>
<td>19 students</td>
<td>1st - 2nd year</td>
<td>Engineering</td>
<td>NR</td>
</tr>
</tbody>
</table>
In six studies, the researcher was the lecturer who provided the feedback to study participants (Ali 2016; Bakla 2020; Cavalieri et al. 2019; Cunningham 2019b; Kim 2018; Mathieson 2012). Five studies failed to report the involvement of the researcher (Elola and Oskoz 2016; Harper, Green, and Fernandez-Toro 2018; Hope 2011; Jones, Georghiades, and Gunson 2012; Silva 2012) and in the remaining four studies, the researcher was not included within the delivery of feedback (Anson et al. 2016; Boone and Carlson 2011; Cunningham 2019a; Grigoryan 2017). Six studies reported that lecturers who provided feedback were experienced in delivering screencast feedback (Boone and Carlson 2011; Elola and Oskoz 2016) or that they received some training to promote a consistent feedback approach across lecturers (Cunningham, 2019a, 2019b; Grigoryan, 2017; Harper, Green and Fernandez-Toro, 2018). However, in studies employing more than one lecturer (Anson et al. 2016; Boone and Carlson 2011; Cunningham 2019a; Harper, Green, and Fernandez-Toro 2018; Jones, Georghiades, and Gunson 2012), only one study reported that they provided clear guidance and instruction on the method and purpose of feedback to lecturers to promote consistency across groups (Grigoryan 2017). Only two studies reported that students were provided with some training on accessing, interpreting, and using this new form of digital feedback (Ali 2016; Bakla 2020).

<table>
<thead>
<tr>
<th>Author/s (Year)</th>
<th>Feedback Interventions</th>
<th>Screencast Program</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ali (2016)</td>
<td>Text vs. (Text + Screencast)</td>
<td>Jing</td>
</tr>
<tr>
<td>Anson et al. (2016)</td>
<td>Text vs. (Screencast + Text)</td>
<td>Jing</td>
</tr>
<tr>
<td>Bakla (2020)</td>
<td>Text vs. Screencast vs. Audio</td>
<td>Not reported</td>
</tr>
<tr>
<td>Boone and Carlson (2011)</td>
<td>Text vs. (Screencast + Text)</td>
<td>Jing</td>
</tr>
<tr>
<td>Cavalieri et al. (2019)</td>
<td>Text vs. (Screencast + Text)</td>
<td>Jing</td>
</tr>
<tr>
<td>Cunningham (2019a)</td>
<td>Text vs. Screencast</td>
<td>SnagIt</td>
</tr>
<tr>
<td>Cunningham (2019b)</td>
<td>Text vs. Screencast</td>
<td>SnagIt</td>
</tr>
<tr>
<td>Elola and Oskoz (2016)</td>
<td>Text vs. (Text + Screencast)</td>
<td>Screencast-O-Matic</td>
</tr>
<tr>
<td>Grigoryan (2017)</td>
<td>Text vs. (Text + Screencast)</td>
<td>Jing</td>
</tr>
<tr>
<td>Harper, Green, and Fernandez-Toro (2018)</td>
<td>Text vs. (Text + Screencast)</td>
<td>Jing</td>
</tr>
<tr>
<td>Hope (2011)</td>
<td>Text vs. Screencast</td>
<td>Jing and Camtasia</td>
</tr>
<tr>
<td>Jones, Georghiades, and Gunson (2012)</td>
<td>Text vs. (Text + Screencast)</td>
<td>Windows media encoder</td>
</tr>
<tr>
<td>Mathieson (2012)</td>
<td>Text vs. (Text + Screencast)</td>
<td>Jing</td>
</tr>
<tr>
<td>Silva (2012)</td>
<td>Text vs. Screencast</td>
<td>Camtasia</td>
</tr>
</tbody>
</table>
Study Research Instruments

In 11 of the 15 studies, non-validated questionnaires were used to assess student and tutor perceptions of the feedback mediums (see Table 3), although two studies assessed face and content validity (Mathieson 2012; Ali 2016), with one reporting a Cronbach alpha of 0.93 (Ali 2016). Three studies supported the questionnaire with interviews (Anson et al. 2016; Cunningham 2019b; Jones, Georgiades, and Gunson 2012) to gain in-depth detail and validate questionnaire responses. Similarly, for studies assessing student learning using academic literacy and writing ability measurements (n = 8), six employed and reported good interrater and/or intercoder reliability measures of coding, revisions and scoring (Ali 2016; Anson et al. 2016; Bakla 2020; Cunningham 2019a, 2019b; Elola and Oskoz 2016; Grigoryan 2017; Kim 2018). Cohen’s kappa measures ranged between 0.87 to 0.97 (Ali 2016; Anson et al. 2016; Bakla 2020; Elola and Oskoz 2016) and intraclass correlation coefficient ranged between 0.73 to 0.966 (Grigoryan 2017; Kim 2018).

Study Results and Conclusions

Table 3 shows that the primary outcomes of the included studies focused on direct, indirect and influential factors on student learning. Direct measures including assignment rubrics, essay revisions and student grades. Indirect measures of student learning included student satisfaction and motivation. Influencing factors on student learning included tutor perceptions and delivery of screencast feedback.

Direct measures of student learning

Seven studies assessed the impact of feedback modality on direct measures of student learning: academic writing skill, essay revisions and academic grades (Bakla 2020; Boone and Carlson 2011; Cavaleri et al. 2019; Elola and Oskoz 2016; Grigoryan 2017; Kim 2018). Using a very small population of four language students, Elola and Oskoz (2016) reported no differences in the modality of feedback on the number and type of essay revisions on draft essays. Similarly, using a larger population group of 50 first-year students on writing composition courses, Grigoryan (2017) reported no statistically significant differences in the number or type of revisions between feedback modalities. Both studies compared text feedback against text feedback plus screencast feedback.

Cavaleri et al. (2019) reported that 84% of screencast feedback (plus minimal marginal comments) comments led to positive changes while text-only feedback only achieved 77%. Logistic regression demonstrated that the odds of a successful revision were 1.59 times greater with the screencast feedback. Students with lower English language proficiency levels demonstrated greater improvement in comparison to students with higher levels of proficiency. Similarly, Kim (2018) demonstrated greater revisions when a mixture of 67 undergraduate level students rewrote business letters following screencast feedback alone in comparison to text feedback. Contrastingly, Bakla (2020) found no differences between text, audio, and screencast feedback methods on 33 students revising a piece of undergraduate first-year English writing.
Table 3: Outcome measures and instruments implemented by studies selected for analysis

<table>
<thead>
<tr>
<th>Author/s (Year)</th>
<th>Primary Outcome/s</th>
<th>Instrument/s</th>
<th>Assignment/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ali (2016)</td>
<td>Writing skill/accuracy, student perceptions</td>
<td>Writing test, questionnaire</td>
<td>Non-academic 150-225-word Essays</td>
</tr>
<tr>
<td>Anson et al. (2016)</td>
<td>Student perceptions, mediation of face, construction of identities</td>
<td>Questionnaire, interviews</td>
<td>3–5-page Essays</td>
</tr>
<tr>
<td>Bakla (2020)</td>
<td>Essay revisions, student perceptions/engagement</td>
<td>Essay scores, interviews, Observational recordings</td>
<td>Essays</td>
</tr>
<tr>
<td>Boone and Carlson (2011)</td>
<td>Student perceptions, grade point comparison</td>
<td>Questionnaire, grade comparison</td>
<td>Essays</td>
</tr>
<tr>
<td>Cavalieri et al. (2019)</td>
<td>Student engagement, feedback quality/quantity</td>
<td>Revisions made, interviews, lecturer feedback review</td>
<td>Various academic assignments</td>
</tr>
<tr>
<td>Cunningham (2019a)</td>
<td>Lecturer language</td>
<td>UAM corpus tool</td>
<td>Essays</td>
</tr>
<tr>
<td>Cunningham (2019b)</td>
<td>Student perceptions, engagement, feedback production time</td>
<td>Questionnaire, interviews, observations, feedback production time</td>
<td>Non-academic 200-300-word Essays</td>
</tr>
<tr>
<td>Elola and Oskoz (2016)</td>
<td>Essay revisions, student perceptions, feedback quality/quantity</td>
<td>Number and type of errors, revisions made, lecturer comments, questionnaire</td>
<td>Essays</td>
</tr>
<tr>
<td>Grigoryan (2017)</td>
<td>Essay revisions and writing skill</td>
<td>Revisions made</td>
<td>1000–1300-word Essays</td>
</tr>
<tr>
<td>Harper, Green, and Fernandez-Toro (2018)</td>
<td>Student and lecturer perceptions, feedback quality/quantity</td>
<td>Questionnaires, lecturer feedback review</td>
<td>Essays</td>
</tr>
<tr>
<td>Hope (2011)</td>
<td>Student perceptions</td>
<td>Questionnaire</td>
<td>1500–2000-word Essays, Laboratory Reports</td>
</tr>
<tr>
<td>Jones, Georgiades, and Gunson (2012)</td>
<td>Student and lecturer perceptions</td>
<td>Interviews, questionnaire</td>
<td>500-word Essays</td>
</tr>
<tr>
<td>Mathieson (2012)</td>
<td>Student perceptions</td>
<td>Questionnaire</td>
<td>Statistics Assignments</td>
</tr>
<tr>
<td>Silva (2012)</td>
<td>Student perceptions</td>
<td>Questionnaire</td>
<td>3-10-page Essays</td>
</tr>
</tbody>
</table>
In addition to overall essay quality, Grigoryan (2017) evaluated a range of higher and lower order writing concerns including overall essay quality, task fulfilment, content, essay organization, paragraph organization, essay purpose, audience, and mechanics to determine if these aspects of student writing had improved across drafts. Despite reporting no effect on essay revisions, Grigoryan (2017) demonstrated that students who received screencast feedback in addition to text feedback on their initial drafts scored significantly higher scores for essay purpose and audience. However, no statistically significant differences were found on essay overall grades or for all other writing concerns (Grigoryan 2017). Contrastingly, using very short non-academic essays (150-250 words) for 63 first-year students on a Management Sciences course, Ali (2016) reported significantly better writing (content, organization and structure) scores and grades when students received screencast feedback as well as text feedback in comparison to text feedback only. However, there was no effect on writing accuracy scores. Kim (2018) demonstrated significantly higher grades when rewriting business letters following screencast feedback (M = 13.60, SD = 1.015) in comparison to text feedback (M = 11.44, SD = 1.307). Boone and Carlson (2011) also identified greater scores when comparing text feedback to screencast feedback (plus some text comments) on essay writing, but no statistical analysis was performed to identify if the scores were significantly greater.

**Indirect measures of student learning**

Eleven studies examined student perceptions of screencast feedback (Ali 2016; Anson et al. 2016; Bakla 2020; Cunningham 2019a; Elola and Oskoz 2016; Harper, Green, and Fernandez-Toro 2018; Hope 2011; Jones, Georgiades, and Gunson 2012; Kim 2018; Mathieson 2012; Silva 2012) and one study investigated the effect of screencast feedback on student motivation and self-efficacy (Kim 2018). Using the Motivated Strategies for Learning Questionnaire (Pintrich et al. 1991), Kim (2018) reported that student motivation and self-efficacy were significantly increased following screencast feedback, which may be a contributing factor to the increase in student grades and indicate learning following this modality of feedback.

There were overwhelming positive responses to screencast feedback in the studies (Ali 2016; Anson et al. 2016; Cunningham 2019a; Elola and Oskoz 2016; Harper, Green, and Fernandez-Toro 2018; Hope 2011; Kim 2018; Mathieson 2012). Screencast feedback was perceived to be more personal, supportive, detailed, more helpful, and easier to understand in comparison to text feedback. Screencast feedback was considered more personal because students perceived that the marker was talking directly to them and likened the experience to being in a face-to-face conference with the lecturer (Ali 2016; Anson et al. 2016; Cunningham 2019b; Elola and Oskoz 2016; Harper, Green, and Fernandez-Toro 2018; Hope 2011; Jones, Georgiades, and Gunson 2012; Mathieson 2012; Silva 2012). The notion of increased dialogue enhanced lecturer:student interaction because students felt more connected with their lecturer and felt more comfortable to contact them to ask follow up questions following their feedback (Anson et al., 2016; Cunningham, 2019b; Elola and Oskoz, 2016; Harper, Green and Fernandez-Toro, 2018; Hope, 2011; Jones, Georgiades, and Gunson 2012; Kim, 2018; Silva, 2012). Screencast feedback was perceived to be more helpful and
easier to understand than text feedback because students could follow the cursor and see corrections made on the screen (Ali, 2016; Cunningham, 2019b; Harper, Green and Fernandez-Toro, 2018; Hope, 2011; Jones, Georghiades, and Gunson 2012; Kim, 2018; Mathieson, 2012; Silva, 2012). The quantity and quality of feedback were also perceived to be higher than text feedback as more words and examples of good practice were provided in the screencast feedback affording more teaching opportunities (Ali 2016; Anson et al. 2016; Elola and Oskoz 2016; Harper, Green, and Fernandez-Toro 2018; Hope 2011; Jones, Georghiades, and Gunson 2012; Kim 2018; Mathieson 2012).

Despite many affordances of screencast feedback, students also identified some limitations of this feedback medium. Some considered screencasts as time-consuming because downloading the video files took time. Additionally, some screencasts were considered too long and when completing revisions, watching the whole video was too linear and took longer than making immediate corrections from text comments (Ali 2016, Silva 2012). Initially, before becoming accustomed to this new form of feedback, some students were nervous about opening their feedback and felt anxious listening to their lecturer without seeing them (Ali 2016; Elola and Oskoz 2016; Hope 2011). Some students stated that screencast feedback was inconvenient because a quiet room or headphones were required to watch the video (Elola and Oskoz 2016). Additionally, some perceived it to be more difficult to understand because you need to be proficient at listening, watching, and comprehending the feedback simultaneously and for students with English as their second language, or students who perceived themselves as visual learners, this was often more difficult than interpreting text comments (Elola and Oskoz 2016; Kim 2018; Mathieson 2012). Additionally, in one study the screencasts were limited by poor audio and video quality (Ali 2016).

From the 11 studies examining student perceptions, only five asked students to confirm whether they preferred screencast or text feedback (Bakla 2020; Cunningham 2019a; Elola and Oskoz 2016; Hope 2011; Silva 2012). Three studies demonstrated a student preference for screencast feedback (Cunningham 2019b; Hope 2011; Silva 2012). Elola and Oskoz (2016) reported that students preferred a combination of both mediums and Bakla (2020) did not indicate a significant preference of medium.

**Influential factors on student learning**

Two studies assessed lecturer perceptions of screencast feedback (Harper, Green, and Fernandez-Toro 2018; Jones, Georghiades, and Gunson 2012) and five examined the impact on tutor feedback language (Anson et al. 2016; Cavalieri et al. 2019; Cunningham 2019a; Elola and Oskoz 2016; Harper, Green, and Fernandez-Toro 2018). Lecturers stated that they were satisfied with the increased detail of feedback that they could provide through screencasts (Harper, Green, and Fernandez-Toro 2018; Jones, Georghiades, and Gunson 2012). The ability to talk through the corrections was perceived as an additional teaching opportunity and lecturers believed this form of feedback was less overwhelming, more engaging and impactful than text feedback, leading to more satisfied students (Harper, Green, and Fernandez-Toro 2018; Jones, Georghiades, and Gunson 2012). As a result of increased interaction with feedback, lecturers suggested that more contact was received from students to discuss the feedback (Harper, Green, and Fernandez-Toro 2018) and this also led to an improved lecturer:student relationship (Harper, Green, and Fernandez-Toro 2018; Jones,
Georghiades, and Gunson 2012). Lecturers also suggested that this form of feedback was beneficial for students and lecturers with learning difficulties such as dyslexia because it was easier to listen or talk the information rather than deal with the written word (Harper, Green, and Fernandez-Toro 2018; Jones, Georghiades, and Gunson 2012).

In contrast to perceived benefits of screencast feedback, lecturers also highlighted some potential limitations of adopting a new form of feedback. Some lecturers suffered from initial anxiety of learning a new tool and talking to students (Harper, Green, and Fernandez-Toro 2018; Jones, Georghiades, and Gunson 2012). Lecturers were also concerned with the time needed to learn and become efficient with the new feedback medium and its effect on their workload. Harper, Green and Fernandez-Toro (2018) stated it took approximately one hour to become familiar with the tool and then each recording took between 15-30 minutes. Some lecturers took longer if they made mistakes during the recording and felt that they needed to restart the recording (Harper, Green, and Fernandez-Toro 2018). Despite these perceived inconveniences, no lecturers suggested any negative consequences of using screencast feedback on student learning.

Five studies examined the effect of the feedback modality on the language of lecturer feedback and its effectiveness on increasing the connectedness between the lecturer and student (Anson et al. 2016; Cavalieri et al. 2019; Cunningham 2019a; Elola and Oskoz 2016; Harper, Green and Fernandez-Toro, 2018). Increased connectedness of lecturer and student should enhance student engagement and thus the student’s opportunity to learn. Analyzing the feedback comments of three English as a second language lecturers across multiple assignments, Cunningham, (2019a) showed that the modality of feedback impacted the identity of the lecturer as text feedback focused more on the deficiencies of the writing and placed the lecturer as an authoritative figure. The screencast feedback was more suggestive and advising which placed the lecturer's feedback as one of many possible opinions and therefore encouraged student autonomy (Cunningham 2019a). In an advanced Spanish writing course, Elola and Oskoz (2016) reported that text feedback was explicit and focused on form whereas screencast feedback provided more, and lengthier comments on content, structure, and organization. Similarly, in another undergraduate language course, Harper, Green, and Fernandez-Toro (2018) showed that the combination of audio and visual feedback allowed language lecturers to increase student exposure to the spoken language. In this study, the screencast feedback was provided in addition to text feedback. The screencast video acted as a summary because the lecturer focused on the key areas of the text feedback. This provided a very personalized nature to the feedback and the tone of the lecturer’s language was supportive and encouraging. However, analysis of the language and tone of the text feedback was not performed for comparison (Harper, Green, and Fernandez-Toro 2018). Through analysis of lecturer feedback comments and student interviews, Cavalieri et al. (2019) confirmed that screencast feedback was more personal and conversational nature, highlighting that the verbal explanations and hearing the lecturer’s feelings were possible mechanisms for increasing student engagement in comparison to text feedback (Cavalieri et al. 2019).
Discussion

Screencasting is a relatively new, innovative form of providing feedback and there has been a surge of research interest over the last decade to determine its effectiveness in HE. However, research has been fragmented and lacking in rigor. This systematic review aimed to evaluate the empirical literature on screencast feedback in HE to determine if this medium of feedback is more effective than text feedback at enhancing student learning. This review shows overwhelming student and lecturer perceptions of screencast feedback but there are currently insufficient high-quality studies using direct, objective measurements to denote its effectiveness on student learning. The current review does indicate that screencast feedback may offer some improvements in these areas, but more rigorous research studies are required to confirm this.

Twelve of the 15 studies included in this review focused on perceptions of screencast feedback. Most students suggested a preference for screencast feedback because it was considered more personal, supportive, detailed, helpful, and easier to understand than text feedback (Ali 2016; Anson et al. 2016; Cunningham 2019b; Elola and Oskoz 2016; Harper, Green, and Fernandez-Toro 2018; Hope 2011; Jones, Georgiades, and Gunson 2012; Kim 2018; Mathieson 2012; Silva 2012). Lecturers agreed that screencast feedback was more detailed, focused and allowed more developmental opportunities because they could talk through corrections (Harper, Green, and Fernandez-Toro 2018; Jones, Georgiades, and Gunson 2012). Both lecturers and students stated that the more conversational tone of screencast feedback was more engaging, impactful and created greater connections between the student and lecturer (Anson et al. 2016; Cavaleri et al. 2019; Harper, Green, and Fernandez-Toro 2018). More satisfying, engaging, and impactful feedback suggests that this medium could be more effective at enhancing student learning (Henrie, Halverson and Graham 2015; York, Gibson and Rankin 2015).

Students and tutors suggest that screencast feedback is more dialogic, which aligns with the social constructivist theory of learning where the lecturer is more facilitative than instructional (Vygotsky 1978). Increased dialogue and a reduced power differential between lecturer and tutor should enhance the opportunity for student learning. However, the perception of increased dialogue may be misplaced because like text feedback, screencast feedback is still a one-directional method of feedback and there is no immediate opportunity for the student to ask questions. Despite largely positive perceptions of screencast feedback, some students and lecturers voiced concerns over the use of this medium. Initial anxiety of watching recordings, the need for proficient listening, watching, and reading skills and technical issues such as audio quality, internet speeds and download time were key frustrations identified by students (Ali 2016; Elola and Oskoz 2016; Hope 2011). Additionally, text feedback was preferred by some students because of these problems and because text feedback felt easier and quicker to make corrections to (Ali 2016; Silva 2012). Lecturers highlighted the additional workload needed and anxiety of producing effective videos as key concerns for adopting this new medium of feedback (Harper, Green, and Fernandez-Toro 2018; Jones, Georgiades, and Gunson 2012).

Student and lecturer perceptions are only indirect measures of student learning and they need to be substantiated with more quantitative measures of student learning. For example, students have stated that screencast feedback is easier to understand and more
useful, but this has not been tested in the research. Only four from seven studies employing quantitative, direct methods of assessment for student learning reported significantly better academic results (Ali 2016; Boone and Carlson 2011) or essay revisions (Ali 2016; Cavaleri et al. 2019; Kim 2018) following screencast feedback. The methodological differences and limitations of these studies may have led to the inconsistent results. More empirical studies employing validated quantitative measures are needed to understand the influence of screencast feedback on student learning.

The research methods employed in all studies of this review were diverse and sometimes limited which restricted the opportunity to effectively combine or compare the study conclusions. Limitations in study populations, feedback provision, comparative groups and data collection/analysis were highlighted by the review. Despite a variety of population groups represented in the literature (i.e., study year, subject and nationality), the observational/case study approach and convenience sampling strategy adopted by many studies led to small sample sizes, often restricted to one population group (i.e., subject, university level), limiting the generalizability of the study results (Cavaleri et al. 2019; Cunningham 2019b; Elola and Oskoz 2016; Harper, Green, and Fernandez-Toro 2018; Mathieson 2012; Silva 2012). Similarly, many studies failed to report the age range of participants (n = 12). First-year students often require more support adapting to the tertiary environment (O’Keefe 2013) and the more personal, conversational nature of screencast feedback may provide this. At later levels, this additional support may not be needed. Therefore, future research should examine the effect of screencast feedback simultaneously across multiple population groups and on varying university levels of students.

In many studies, the researcher was embedded within the research design and was the provider of feedback (Ali 2016; Bakla 2020; Cavaleri et al. 2019; Cunningham 2019b; Kim 2018; Mathieson 2012). Having one assessor may offer a more homogenous approach to the delivery of feedback but the risk of assessor bias reduces the viability of study conclusions. Additionally, in most studies, limited or no detail was reported on the experience of the feedback provider in delivering screencasts or any training given. Harper, Green and Fernandez-Toro (2018) provided examples of screencasts but did not offer any further guidance. Comparing a lecturer with years of experience providing text feedback against someone new to screencasts may affect the delivery of feedback and its results. Future studies should assess the impact of lecturer experience on the delivery of different mediums of feedback to determine if any potential impact is caused by more experience or the medium itself. Training on delivering screencast feedback should help to reduce the bias when comparing the feedback from lecturers with years of experience delivering text feedback to limited or no experience of screencast feedback. Similarly, studies did not fully report the training of students in receiving video feedback. Students may need training in how to download, engage with and learn from this new modality of feedback. Future studies should use multiple independent lecturers and provide training for both lecturers and studies.

The diversity in the delivery of feedback (i.e., comparative group, screencast program, length of the screencast, and the number of screencast interventions) within this review also limits the comparison and validity of results. Six of the 15 studies compared text feedback to text plus screencast feedback which did not offer a valid comparison of the different affordances of each medium because additional, rather than alternative
feedback was provided (Ali 2016; Elola and Oskoz 2016; Grigoryan 2017; Harper, Green and Fernandez-Toro 2018; Jones, Georgiades, and Gunson 2012 Mathieson 2012). Most studies (n = 8) used a screencast program that was limited to five minutes whereas there was no limit on the text feedback provided (Ali 2016; Anson et al. 2016; Boone and Carlson 2011; Cavalari et al. 2019; Grigoryan 2017; Harper, Green, and Fernandez-Toro 2018; Hope 2011; Mathieson 2012). Sixteen students from one study suggested that videos of 10 minutes were too long but they would watch it if it was pertinent to their progression (Silva 2012). No study provided an analysis of different screencast durations to determine the time effect. Therefore, although more research is required on the optimal duration of feedback videos to maximize student engagement and learning, the duration of feedback in the studies in this review may have been a confounding factor in their results. Future research should assess the effect of different screencast programs and the length of screencasts on student learning outcomes.

This review is limited by the small sample sizes and methodological quality of the included studies. Systematic reviews offer a transparent analysis of available literature in a field, but the rigorous selection criteria employed by this study may have limited the number of studies for analysis. Additionally, many of the included studies suffered from low sample sizes, researcher bias and diverse methods common with media comparison studies. Even in well-designed media comparison studies, there are too many confounding variables for the conclusions to be considered valid (Surry and Ensminger 2001). Future research should focus on the affordances of different media, and how people can learn from them, rather than simply comparing them.

**Conclusion**

Currently, there is insufficient empirical evidence to suggest that screencast feedback is more effective than text feedback at enhancing student learning in HE. Therefore, the results of this review do not warrant the adoption of screencast feedback in HE. More empirical comparisons employing independent lecturers who are experienced or trained to deliver screencast feedback across multiple population groups is required. Studies need to compare different screencast programs, durations and adopt valid and reliable direct outcome measures to confirm if this more personal, conversational, and supportive form of feedback can be more effective than text feedback at improving student learning.
REFERENCES


Silva, Mary L. 2012. “Camtasia in the Classroom: Student Attitudes and Preferences for Video Commentary or Microsoft Word Comments During the Revision


ABOUT THE AUTHORS

Simon Penn: Lecturer, Faculty of Wellbeing, Education and Language Studies, The Open University, Milton Keynes, England.

Nicola Brown: Associate Professor, Faculty of Sport, Allied Health and Performance Sciences, St Mary’s University, Twickenham, England