

Visual Peer Review: Examining Students' Parts of Speech Comments vs. Grade

Kevin Hawley
The Zimmerman School of Advertising & Mass Communication
University of South Florida, USA
kevinhawley@usf.edu

Alon Friedman
School of Information
University of South Florida, USA
alonfriedman@usf.edu

Abstract: Increasingly, educators are using peer review as a pedagogical tool to engage students in higher education. Studies of the effectiveness of student peer review have found that thinking critically about the work of peers improves student learning (Mulder et al. 2013). Separately, while visual literacy communication courses, in which educators teach effective visualization practices, are gaining in popularity, few options exist to measure student engagement with the course material. Recently, however, Friedman and Rosen (2017) developed a student peer review model of visual assessment to provide insight into student engagement in this area. In this study, we build upon that work using sentiment-based methods to analyze student peer review comments and relate elements of the text to assignment and final project grades. As part of this effort, we developed a new data dictionary to support natural language processing in visual literacy in the liberal arts. The results of the textual analysis show that students who received high grade scores from the instructor used nouns extensively. However, further examination revealed that the correlation between grades and speech was moderate ($R = 0.3$). Future studies should examine whether behavioral nudges may promote more complete sentences by students performing peer review.

Introduction

The growing popularity of social media and the use of visualizations to communicate scientific discoveries, and messages in both politics and the humanities, has generated heightened demand for visual communication courses in higher education. Although many educators in the field of engineering have proposed methods for the evaluation of visualizations in the classroom, to date none have offered a practical way to assess student engagement associated with learning. In the liberal arts, most existing techniques and algorithms for evaluating visualizations are difficult to implement in the classroom, because learning goals are focused more on understanding why than how. Friedman and Rosen (2017) proposed a new model for student peer review of visualizations based on Moxley's (2010) model from the field of writing studies, which stated that the use of peer review technologies had changed the ecology of assessment and student motivation (Moxley 2010, p. 21). Friedman and Rosen (2017) modified Moxley's model and outlined five categories of assessment, each of which emphasized a different aspect of visualization education: algorithm design, interaction design, visual design, design consideration, and visualization narrative. Using Friedman and Rosen's visual peer review model, Beasley et al. (2021) used sentiment analysis to match positive and negative sentiment-bearing keywords to peer review comments from two different visualization courses (in engineering and liberal arts). In this study, we build upon this previous research to investigate what insights may be drawn by conducting a parts-of-speech analysis (including nouns, adjectives, and adverbs) of the peer review comments of liberal art students attending a visual literacy course at an American public university and correlating these speech elements with the grades of the reviewers. Specifically, this study asked the following research questions:

- (1) What does a parts-of-speech analysis reveal about student grades in visualization education?
- (2) What insights might these data provide to an instructor regarding the use of visual peer review as a pedagogical methodology?

We first review visualization education and the use of peer review as an assessment strategy. Next, we outline our methods and results for measuring elements of speech from student comments and correlating these elements with student grades. Finally, we summarize our findings and provide recommendations for future research.

Visualization Education

The use of technology in higher education has grown at an unprecedented rate in recent years. Computers and mobile devices are now integral to the development and delivery of instruction in higher education. For example, during the coronavirus pandemic many institutions of higher learning in the United States pivoted to remote learning for students through online courses. However, the rise in online education is associated with renewed concern about how best to promote student engagement in coursework. In an educational context the term ‘engagement’ refers to the time, energy, and resources that students devote to class activities that are designed to enhance learning. As social media continues to weigh on society, Junco (2011) reported on the relationship between the frequency of Facebook use and student engagement in class. The study found Facebook use was significantly negatively predictive of engagement. According to Bransford et al. (2000), student assessment is a core component of measuring student engagement and effective learning.

Student assessment may take many forms, but peer review is one highly effective assessment strategy that is commonly used in higher education—especially in the liberal and creative arts. In student peer review, students may provide both formative (feedback) and summative (grading) assessments other students’ work. Peer review is also an established and essential component of many professional practices, including the scholarly publication process; during professional peer review, domain experts may appraise professional performance, creativity, or the quality of work produced by others in their area of study. Ultimately however, peer review is the evaluation of work by one or more people of similar competence to the producer(s) of the work (Spier, 2002).

The goal of this study was to gain insight into the use of student peer review in visual literacy education. Previous research in this area has primarily centered on two fields of education: engineering and liberal arts. Recently, Beasley et al. (2021) analyzed the use of student peer review in courses from both the College of Engineering and the College of Arts and Sciences at a public university. However, the authors did not compare differences between the two courses or include the average grades the students received in their analysis. In this study, we focus our analysis on a course offered by a College of Arts and Sciences titled ‘Visual Literacy.’

Visual Literacy Communication

John Debes coined the term visual literacy, stating that “visual literacy refers to a group of vision-competencies a human being can develop by seeing and at the same time having and integrating other sensory experiences. The development of these competencies is fundamental to normal human learning” (Debes 1969 pp.115-118). Hortin (1983) added the division of education to the concept, stating that “visual literacy is the ability to understand (read) and use (write) images, and to think and learn in terms of images, ie, to think visually” (Hortin, 1983, 99). In their review of the term, Avgerinou & Ericson (1997) also stressed the importance of education to the terminology.

A visualization course in the College of Arts and Sciences at USF

The University of South Florida has developed an interdisciplinary visual communication curriculum within the College of Arts and Sciences, as part of a larger university-led effort to improve visualization skills and access to visualization equipment. In the College of Arts and Sciences, students are exposed to various techniques for creating images, diagrams, and/or animations using multiple platforms and programming languages. As part of this effort, we offer a survey course called Visual Literacy that imparts to students a foundational knowledge of the many forms that visual communication can take and provides hands-on opportunities to develop projects in multiple areas. The assignments for this class are based on different rubrics that allow students to apply theoretical concepts to design and development projects.

The visual peer review model in visual literacy class education

Friedman and Rosen (2017) developed a new model for student visual peer review in response to the lack of availability of standard rubrics for the assessment of student engagement in visual literacy courses in the liberal arts. The model was based on the following five major assessment categories: algorithm design, interaction design, visual design, design consideration, and visualization narrative. The algorithm design category is concerned with algorithm

selection and implementation. Interaction design is concerned with how the user interacts with the visualization. Visual design is concerned with the technical aspects of how data are placed in the visualization. Design consideration focuses on the composition and aesthetic aspects of the visualization; aspects such as embellishments are also considered. The final category, visualization narrative, allows for assessment in projects in which the story associated with the visualization is as important as the visualization itself. This category allows for critical assessment of the extent to which the visualization supports the story and vice versa. In this study, we focused on only one category of assessment—visualization narrative—which was most closely related to the course content of the visual literacy course. For this course, we modified the category to better align with course objectives aimed at applying basic design principles to the planning, organization, and creation of visual communications. Figure 1 shows an example of a student-generated peer review, including both the visual work and student review comments. (Fig. 1).

Methodology

With recent growing data in educational technology, supervised machine learning is frequently used and is commonly employed for natural language processing (NLP) and sentiment analysis. In a recent review of education technology, Kastrati et al. (2021) considered different approaches for analyzing student peer reviews using the supervised machine as the core framework. The study report on a lack of a global scheme that could be applied to all the different educational domains and areas. Beasley et al. (2021) used NLP to draw insights from student peer review comments and found that students received more benefit from reviewing and providing feedback to others than from self-reflection or feedback received from peers. These findings were consistent with those of a previous study conducted in an engineering class (Garousi, 2009).

Most previous studies have been undertaken in the domain of engineering. Thus, in this study, we developed a specific dictionary-based NLP algorithm focused on liberal arts terminology to analyze student peer review comments and evaluate student engagement in a visual literacy course. Our analysis was based on an algorithm inspired by Beasley (2020) that generated unique metrics for assessment, including the overall sentiment of the review, counts of parts of speech (nouns, adjectives, adverbs), and average comment length. In this algorithm, a word extractor scanned the text in a sliding window and produced a list of essential words (nouns), near sentiment words (adjectives), and their preposition relationships. We compared these words to Friedman and Rosen's rubric to determine the extent to which students focused on a topic covered in this module and correlated those quantitative variables with the grade that the students received on this particular assignment in the class. This measurement is course content-based material. To match the rubric, our development dictionary was explicitly tuned to the visual terminology used in the class, wherein the algorithm matched visualization keywords used in the class, and more specific general nouns, verbs, adverbs, and adjectives.

In this study student peer review feedback comments were used to better understand and improve the effectiveness of an online visual literacy course. Our goals were to synthesize aspects of student motivation and improve student learning outcomes via improved understanding of the components of, and changes in, student engagement in the humanities. To meet these goals, we analyzed student peer review comments collected over a three-year period via metrics extracted from a modified visualization assessment rubric.

Data Collection

Student performance in the course was assessed through weekly assignments that included peer reviews. The original visual peer review rubric included assessments of label of the work, data-ink ratio, chart junk data density, and Gestalt principles. In this study, we modified the rubric to better reflect the data visualization topics covered in this course, and included assessments of color scheme, color value (lightness or darkness of a color), visual signs, space, and quality of labeling. 'Color scheme' required a student to evaluate a peer's visualization with respect to the topic based on the colors chosen for the work. 'Value' referred to the lightness or darkness of a color. Color values are often visualized as a gradient, in which variations in a hue are arranged from lightest to the darkest. Rose (1962) defined visual signs as part of culture and rhetoric; thus, the next category of assessment was 'Visual signs,' which included assessment of the semiotics of signs in the work. The assessment category 'Space' required students to evaluate the use of white space or negative space in visual design, which is the space between or around the focal point of an image (Golombisky and Hagen 2013). For the last category, 'Labeling,' students evaluated the effectiveness of titles and labels based on Tufte's principles of design (Tufte, 1990). We collected peer-review data

from the students' fourth course assignment, and from the final project. Data collection took place from 2020 to 2022, during fall semester.

The participants in this course were undergraduate students with a declared concentration in visual literacy. The course is offered in fall and spring semesters and is fully online. A total of 350 students participated in the study between 2020 and 2022, with an average of 116 students per semester. Many of the participants who enrolled in the course had declared majors in the humanities, including mass communication, library science, and English. Students were not required to enroll to the class. The course was developed as part of the university's Enhanced General Education Curriculum and certified for Information and Data Literacy (Intellectual and Practical Skills). It was designed in collaboration with the university's Innovative Education department to meet university-mandated High Quality online course design standards.

Results

Our first step was to analyze students' comments by looking at their adjectives, adverbs, and nouns from the comments collected over three years. Parts-of-speech analysis can provide valuable insight into student engagement because adjectives connote sentiment (e.g., "excited" or "proud"), nouns are aspects toward which the sentiment is applied (e.g., "design" or "presentation"), and adverbs are enhancers (e.g., "clearly" or "inappropriately"). According to Kahl and Venette (2010), greater use of these parts of speech can indicate increased engagement. Figure 2 summarizes our findings. The students invested more effort in reviews of the final project (9.03 total adjectives, adverbs, and nouns per review; mean \pm SD = 3.5 ± 1.3) than in the fourth assignment. The average number of words per review increased between the fourth assignment to the final project reviews (from 11.28 to 20.35). A temporal increase (from the fourth assignment to the final project) in the use of nouns was also observed over the three semesters of data collection. (Fig. 2).

The results of our analysis comparing the text in comments made on the fourth weekly assignment and the final project revealed changes in student motivation and participation over the semester. Next, we examined the relationship between review metrics (engagement) and student grades (outcomes).

We also evaluated several metrics with respect to association with the course grade of the reviewer. These metrics included the number of reviews submitted, the average number of words per sentence and number of words per review, and the average numbers of adjectives/adverbs/nouns used per review. The student engagement metrics evaluated in this study are shown in Table 1, along with the overall grade of the reviewing student. (Tab. 1).

First, we determined whether the types of part of speech used by students conducting reviews varied across the distribution of overall course grades. Course grades were non-normally distributed, with most students in the course earning an A (307), followed by a B (64), C (36), D (25), or F (18). The number of nouns varied linearly across the grade distribution, with students who received an 'A' using more nouns, on average, than students who received lower grades. One possible explanation for the finding is that students who received a high grade in the course included more nouns in their reviews to critically describe the work and demonstrate their level of understanding to the course instructor.

To evaluate trends in student review metrics over time, we also examined the extent to which parts-of-speech varied across the 3-year study period (Fig. 3). We found that the most diverse parts-of-speech metrics occurred in fall 2020, the first semester after the impact of COVID-19. We found decreases in the subsequent years for all parts of speech—adjectives, adverbs, and nouns. Students wrote an average of 21.14 words per review in 2020, 17.17 words per review in 2021, and 16.6 words per review in 2022. Accordingly, students wrote approximately 47% more in 2020 than in 2022. These results indicate greater student engagement in Fall 2020 than in 2021 and 2022. Although no changes were made to course content or delivery that semester, the instructor noted an increase in casual conversation and personal interaction during office hours. Increasing student-instructor interaction may have contributed to increased peer-review engagement earlier in the study period. (Figure 3).

Last, we conducted a traditional correlational analysis following Beasley (2020). In education, correlation analysis permits the evaluation of relationships among variables, such as a positive relationship between parts-of-speech and grades in the classroom (Zopluoğlu 2012). A correlation analysis makes no assumption about whether one variable is dependent on another (Sproull 2002), but rather provides an estimate of the degree of association between two variables—in our case, parts-of-speech and student grades. We analyzed a sample of the data with a 94% confidence level and a 12.3 confidence interval. Means and standard deviations (SDs) were computed for parts-of-speech and student grades. The mean for parts-of-speech was 4.51 and the grade was 0.34, with an SD of 1.63, and the score of grades with a sample was 4.21 and sample standard deviation of 1.41. Correlation analysis evaluating the

relative independence of parts-of-speech and grade revealed a correlation of $r = 0.34$, $p = <0.001$, with $R^2 = 0.29$. This result indicates that the relationship between type of speech and grade is moderate.

Recommendations

Peer review of visual course assignments was used in this study to engage students taking an online course. Similar instructor assessments of student engagement using data extracted from reviews, as outlined in this study, may be generalized to any visual literacy class. Using clear rubrics, sentiment analysis can be used for quantitative analysis of student engagement, and correlation analysis can be used to directly assess the variables at hand. The inclusion of ongoing feedback or additional review stages during peer review may further strengthen student engagement and core concept learning. Overall, visual peer review should be considered as a complementary method of assessment to existing educational approaches, with continued feedback and oversight from the instructor.

Summary

To gain insight into student engagement in visual literacy classes, this study analyzed textual feedback from peer review comments collected from a total of 350 students enrolled in an online visual literacy class between 2020-2022 in the College of Arts and Science at a public university. Trends in student engagement were illuminated through a parts-of-speech analysis, in which the instructor analyzed student comments by looking at the use of adverbs, adjectives, and nouns across assignments and over semesters. Students who received a grade of A used more nouns in their reviews. However, the correlation between students' parts-of-speech and grades was moderate, suggesting that further studies of this relationship are warranted. Since technology is increasingly encouraging users to interact on social media, where studies found that nudges technology have a direct effect on user interaction found on Facebook (Purohit et al., 2022), future studies should examine how students' peer review comments are influenced by consumer behavior and nudge theory. As well, studies should also examine behavioral nudges through student comments through the lens of part-of-speech analysis.

References

- Avgerinou, M., & Ericson, J. (1997). A review of the concept of visual literacy. *British Journal of Educational Technology* 28.4 (1997), 280–291.
- Beasley, Z. J., & Piegl, L. A. (2021). A domain-dependent lexicon to augment CAD peer review. *Computer-Aided Design & Applications* 18(1), 186-198.
- Beasley, Z. J. (2020). Sentiment analysis in peer review. Doctoral dissertation, University of South Florida.
- Beasley, Z., Friedman, A., & Rosen, P. (2021). Through the looking glass: insights into visualization pedagogy through sentiment analysis of peer review text. *IEEE Computer Graphics and Applications* 45.
- Bransford, J., Brown, A. L., Cocking, R. R., & National Research Council (U.S.). (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.
- Chen, M., Ebert, D., Hagen, H., Laramée, R. S., van Liere, R., Ma, K.-L., Ribarsky, W., Scheuermann, G., & Silver, D. (2009). Data, information, and knowledge in visualization. *IEEE Computer Graphics and Applications* 29(1), 12-19.
- Debes, J. L. (1969). Some hows and whys of visual literacy. *Educational Screen and Audiovisual Guide*, 14-15.
- Golombisky, K., & Hagen, R. (2013). *White space is not your enemy: a beginner's guide to communicating visually through graphic, web and multimedia design*. Routledge.
- Friedman, A., & Rosen, P. (2021). Leveraging peer review in visualization education: A proposal for a new model. arXiv preprint arXiv:2101.07708.

- Garousi, V. (2009). Applying peer reviews in software engineering education: An experiment and lessons learned. *IEEE Transactions on Education* 53(2), 182-193.
- Junco, R. (2012). The relationship between frequency of Facebook use, participation in Facebook activities, and student engagement. *Computers and Education* 58(1), 162-171.
- Kahl Jr, D. H., & Venette, S. (2010). To lecture or let go: A comparative analysis of student speech outlines from teacher-centered and learner-centered classrooms. *Communication Teacher* 24(3), 178-186.
- Kastrati, Z., Dalipi, F., Imran, A. S., Pireva Nuci, K., & Wani, M. A. (2021). Sentiment analysis of students' feedback with NLP and deep learning: A systematic mapping study. *Applied Sciences* 11(9), 3986.
- Moxley, J. (2010) Afterword: disciplinarity and the future of creative writing studies. *The writing workshop model: Is it still working?* Ed. Donnelly, Dianne. London: Multilingual Matters 2010. 230-237.
- Mowery, D. C., & Oxley, J. E. (1995). Inward technology transfer and competitiveness: The role of national innovation systems. *Cambridge Journal of Economics* 19(1), 67-93.
- Mulder, R. A., Pearce, J. M., & Baik, C. (2014). Peer review in higher education: Student perceptions before and after participation. *Active Learning in Higher Education*, 15(2), 157-171.
- Spier, R. (2002). The history of the peer-review process. *Trends in Biotechnology* 20(8), 357–358.
- Purohit, A. K., Barclay, L., & Holzer, A. (2020). Designing for digital detox: Making social media less addictive with digital nudges. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-9).
- Tufte, E. R. (1990) *Envisioning Information*. Cheshire, Conn. (P.O. Box 430, Cheshire 06410): Graphics Press.
- Zopluoğlu, C. (2012). A cross-national comparison of intra-class correlation coefficient in educational achievement outcomes. *Journal of Measurement and Evaluation in Education and Psychology* 3(1), 242-278.

Acknowledgement

This work was funded by the NSF Division of Undergraduate Education (DUE) and Directorate for STEM Education (EDU). Award: Using Behavioral Nudges in Peer Review to Improve Critical Analysis in STEM Courses. See: https://www.nsf.gov/awardsearch/showAward?AWD_ID=2216227

Figures and Tables

Grade	Reviews	Words/Sentence	Adj	Adv	Noun
A	307	21.44	1.65	0.99	8.99
B	64	18.66	1.78	1.19	6.03
C	36	15.88	1.34	0.56	5.78
D	25	6.11	0.33	0.33	5.21
F	18	5.86	0.24	0.33	5.11

Table 1. Text metrics of peer reviews classified by overall grade counted.

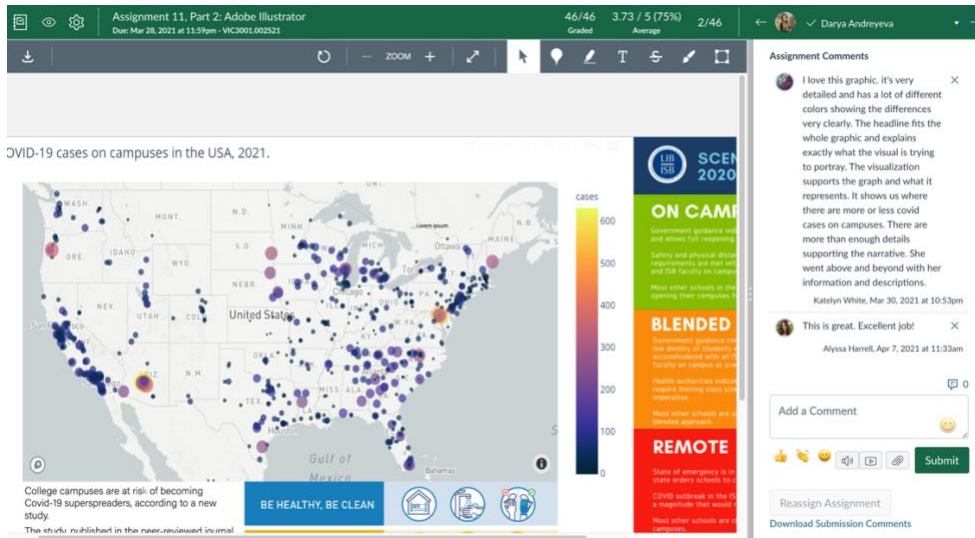
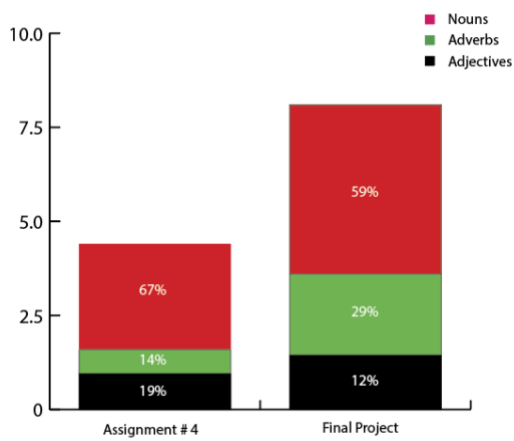


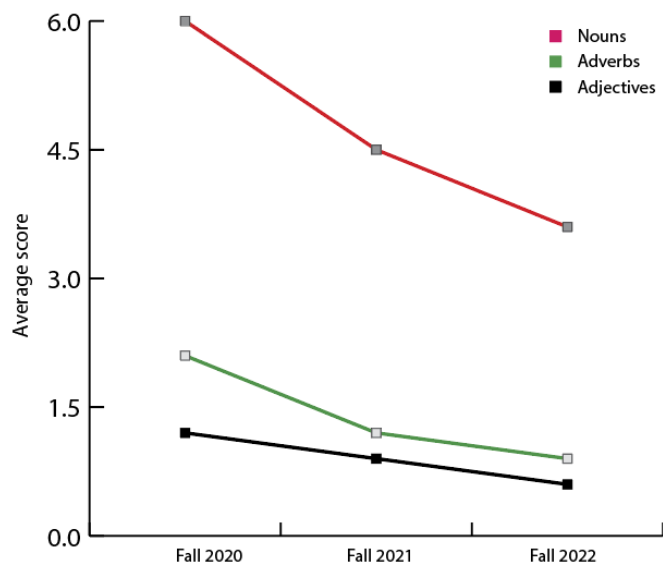
Illustration of student work and on the right-side student comments.

Figure 1. Student work and peer review



Average part-of-speech usage per project over all three years of data collection, showing the number of adjectives, adverbs, and nouns per review for the fourth weekly assignment and the final course project.

Figure 2. Average part-of-speech usage per project



Temporal trends in language use across three semesters. Data are the average numbers of parts-of-speech (adjectives, adverbs, and nouns) per review.

Figure 3. Temporal trends in language use