STUDENT AND INSTRUCTOR EXPERIENCE USING COLLABORATIVE
ANNOTATION VIA PERUSALL IN UPPER YEAR AND GRADUATE
COURSES

Agnes G. d’Entremont and Adrianna Eyking
Department of Mechanical Engineering, University of British Columbia
adentremont@mech.ubc.ca, aeyking@mech.ubc.ca

Abstract – Perusall is a collaborative annotation platform designed for pre-readings in a flipped classroom, but can also be used for stand-alone, asynchronous reading discussion components of courses. We examine the use of Perusall as a social constructivist learning tool in two upper year/graduate courses in Mechanical Engineering. Perusall was used to replace in-class discussion of readings during the shift to online teaching.

Data was collected from student surveys and from the student and instructor annotations themselves. Annotations were coded for content, and examined for factors such as upvoting.

We found substantial engagement from students, with collaborative annotation providing opportunities for: correction of misunderstanding; linking concepts from the course and between readings; discussing larger issues around research and research writing; sharing background information among peers; and critically analyzing the readings. Students reported deeper learning than in typical in-class discussions of readings; however, they also noted that annotation required much more time.

Overall, collaborative annotation appears to be an effective method for course reading discussion.

Keywords: Collaborative annotation, online learning, social constructivism, assigned reading discussions, undergraduate education, graduate education, engineering technical elective, English for academic purposes.

1. INTRODUCTION

Perusall is a collaborative annotation system that allows students and instructors to add comments and questions to documents, images, and videos [1]. Past research has indicated that completing Perusall assignments can contribute to deep learning for students [2], and, among engineering graduate students, led to more self-reported learning [3]. The system has been often illustrated as a tool for pre-reading assignments in a flipped classroom [4], however it can also be used for completely asynchronous components of a course, whether online or in-person. This second usage, as a stand-alone tool for reading discussion, is the focus of this paper.

Our research questions were:
1. How did upper year and graduate students in engineering use Perusall for reading discussions?
2. What were student and faculty perspectives on using Perusall for discussions of readings?

2. BACKGROUND AND CONTEXT

2.1. Conceptual Framework and Literature

Perusall can be considered a tool for social constructivist learning, where students learn by creating meaning through interactions with each other [5]. Social learning contexts allow both instructors and students to “examine, question and build on” the thinking that is expressed [6]. Prawat and Floden describe the process of negotiation within a social constructivist learning context as “an opportunity to surface and clarify points of agreement and disagreement”, and that instructors who are comfortable with social constructivist approaches may “view uncertainty and conflict as natural and potentially growth producing” for learners [7]. This approach is distinct from cooperative learning, where learning negotiation is focused on compromise and getting learners to agree to something [7]. Barak argues that collaborative learning technology can enable learner-driven social constructivist curricula, which in turn can develop higher-order thinking through discourse among learners [8].

Past examinations of Persuall use for research paper discussion highlight connections to social constructivism. Lee and Yeong describe observing interactions among students where they “correct[ed] and buil[t] on concepts proposed by the earlier contributors” when using Perusall in an undergraduate biology class [9]. Adams and Wilson describe how Perusall facilitated the development of a social constructivist learning community in an asynchronous graduate course through peer interactions in addition to text interactions [10].
2.2. Context

Due to the COVID-19 pandemic, all engineering course activities at our institution were run online during the 2020-2021 school year. As a result, two instructors of upper year and graduate courses in mechanical engineering chose to convert classroom discussion of readings into asynchronous collaborative annotation of readings. This was partly because some students were expected to be living in time zones that were incompatible with attending live classes (recorded for later viewing). Holding live class discussions would disadvantage those students, since analyzing readings was an important part of each course.

2.3. Perusall Function

Perusall promotes student participation and engagement in reading assignments by making reading both a social and a graded exercise. The system supports annotation of text, images, videos, and podcasts; setting smaller discussion groups; and automatic grading.

To add content, instructors can upload PDF files; link to webpages; or have students purchase an e-textbook through Perusall. Instructors then assign readings, set due dates, and decide if student comments will be anonymous. Instructors can divide large classes into smaller groups, to reduce the number of comments visible to students.

Students earn points on readings by posting original comments or questions, or by engaging with others’ posts. To add reading annotations, students highlight a text portion or drop a pin and type their comments. Students can also reply to an existing thread or “upvote” comments or questions. Upvoting a comment indicates that it “help[ed] [the student’s] understanding,” while upvoting a question indicates that they “have the same question” [1]. Instructors can use the same annotation tools to add prompts to readings or to interact with students.

Based on each student’s engagement with the text and with their peers, Perusall automatically assigns a grade between 0 and 3 per reading. Instructors can customize the grading by changing the number of annotations required, adjusting late penalties, or shifting the weightings of various grading factors. These factors include: distribution of comments throughout the document, number of times the student opened the document, time spent with the document open, number of responses from other students, and number of upvotes given and received.

3. METHODS

3.1. Course Descriptions and Design of Perusall Activities

We examined the use of Perusall in two courses during the Fall of 2020. Both are offered by the department of Mechanical Engineering. One course was a co-listed 4th year and graduate technical elective in injury biomechanics (TECH) with 25 students enrolled. Past practice in the course was to assign about 5 readings throughout the term. Students would complete a short summary of each reading in advance of class (to ensure completion), and would discuss in the classroom. Reading completion was assigned either a low mark percentage or bonus marks only.

When shifted online, TECH used Perusall for reading research literature, with one graded reading per week (12 total). The marks value of the readings was increased to 15% (undergraduate) or 10% (graduate) of the course mark (best 10 of 12 readings), with the grading completed automatically by Perusall. Since both reading volume and time required per reading with collaborative annotation were increased, the instructor cancelled an hour of live class/tutorial time per week (reduced from 4 to 3 hours).

The TECH instructor seeded initial questions for most readings. Although there were more than the system-recommended 20 students, the instructor chose to place all 25 students in one common commenting group. The TECH instructor also used Persusall live in class three times.

The second course was a graduate-level research writing course for engineering (WRIT) with 13 students. In the past, students received paper handouts of assigned readings and discussion questions, to be completed before the next class. There was no grade assigned for completing the readings or the questions, but students were expected to participate in a verbal discussion in the following class.

After switching to an online format, WRIT used Perusall for research articles and textbook excerpts, and participation through Perusall was assigned a value of 5% of the student’s final grade. As with TECH, grading was completed automatically by Perusall.

A total of 11 readings (7 research articles and 4 textbook excerpts) were assigned in WRIT during the term. The instructor seeded only the first article with questions, and two of the textbook excerpts contained questions within the reading. The number of readings ranged from 0-3 per week, with most weeks in the first two months having 1-2 graded readings per week. In the final month of the course, when other assignments were heavier, Perusall was used only to facilitate in-class discussion, and for providing additional resources as “optional” (ungraded) readings. As there were fewer than 20 students in the course, all students were in the same commenting group.

Throughout both courses, the instructors regularly reviewed comments, responded to questions, and corrected misunderstandings via Perusall.

3.2. Data Collection

We collected several types of data to allow for triangulation of results. Students completed an anonymous survey about Perusall at the end of each course. Students could also optionally give us permission to quote their personal comments in this paper.

We also evaluated the actual annotations, both using summary statistics and by coding both instructor and student annotations for one representative reading from
each course. A common coding guide was created to characterize comments. However, each instructor coded their own course reading, as subject matter knowledge was required to characterize comments.

4. RESULTS

4.1. Usage

A total of 3166 (372 instructor and 2794 student) comments were made during the TECH course, of which 3076 (326 instructor and 2750 student) comments were from readings (with the remainder from three in-class activities). From this point, we will discuss usage statistics for the readings only. While participation dipped at times during the term, possibly aligned with midterm exams and end-of-term projects (Figure 1), engagement remained relatively steady. TECH had an average of 9.2 (SD 1.3) comments per student per reading and 229.1 (SD 33.1) total student comments per reading.

There were 1160 unique TECH comment threads, and 803 threads had more than one comment, with up to 12 comments per thread (average 2.7 (SD 1.8) comments per thread). 3031 upvotes were given, with up to 12 upvotes for the most popular comments, although the average was 0.98 (SD 1.5) per comment. (Table 1).

A total of 1009 comments were made in WRIT (110 instructor and 899 student), including optional readings. In the assigned readings, there were 968 total comments (89 instructor and 879 student). Again, we will consider only the assigned readings from this point. Similar to TECH, there was variability in student participation (Figure 2), likely due to variation in the number of readings assigned in each week. The course had an average of 6.1 (SD 1.6) comments per student per reading, and 79.9 (SD 20.6) total student comments per reading.

WRIT had 414 comment threads (239 with more than one comment) with an average of 2.4 (SD 1.9) comments per thread. There were 290 upvotes, with a mean of 0.3 (SD 0.6) upvotes per comment, and a maximum of 6 for one comment (Table 1).

4.2. Comment Content

We found comments that served a wide range of goals within the readings. Below is an example exchange of comments from a TECH reading. The thread relates to an aspect (drop height) of a test where crash test dummy heads with embedded accelerometers were dropped onto a steel plate to record accelerations at impact [11].

Student A: An even lower drop height than the study we saw in the lecture (600mm). But as stated below, the standard calls for an acceleration of 225 to 275 g's. This seems pretty high just from dropping the head from about a foot off the ground. I suppose that the head has quite a bit of mass and the steel plate contributes nearly zero damping (which will decrease the delta time of the impulse). [1 upvote (instructor)]

Student B: An average human head weighs only around 5kg, and mass doesn't play a role in how fast an object falls (on Earth) so it does seem quite excessive to see such a high acceleration under such a short drop distance. I wonder if peak acceleration might be rotational acceleration instead of linear, as if the head is rotating during impact. [1 upvote]
Student A: I would imagine a rotational acceleration to have different units than g's. Something like rad/s^2 would be more representative. In a few lines below it actually says that three orthogonal linear acceleration measurements were made and the resultant of the 3 vectors made up the head response.

Student C: Isn't the acceleration stated that of the dummy head deaccelerating upon impact with the plate? The acceleration leading up the impact, if the head was truly dropped, would only ever be one g, no? [4 upvotes (1 instructor)]

Student D: Below it says that "head response is the resultant of these acceleration measurements" so I think it is the 3D linear resultant acceleration following impact.

This exchange demonstrates: making a connection to course content (comparing to study discussed in lecture, considering what influences the delta-t time of an impact impulse); providing background information (mass of a human head); critically analyzing reading content (considering the reasonableness of the acceleration range and the cause of those results); and correcting misunderstanding of reading content (the acceleration applies to the impact and not the free fall, the measure is not rotational acceleration).

The benefits of the original comment extended beyond the primary theme of that comment or the student who posted it. A critical analysis of one aspect of the reading (drop height) led to correcting two other misunderstandings of the reading content. Most upvotes were from the instructor (2) and others posting in the thread (Students A (2) and D (1)), but an additional student who did not comment also upvoted, indicating that the comment from Student C helped their understanding of the reading. It is important to note that this exchange was student-initiated, and not based on a question posed by the instructor. In this way, students are able to direct their own learning through what they comment on. The instructor can also quickly highlight analyses closely linked to course learning goals or correct assessments by upvoting – the comment then gets both an upvote and a prominent “Upvoted by instructor” tag (Figure 3).

In addition to upvoting, both instructors provided text responses to students’ questions and comments. In the following exchange, the WRIT instructor responds to one student’s question and corrects another student’s misunderstanding (by referring them to a previous discussion) about the “essential tasks of a literature review for a dissertation proposal” [12].

Student A: how does this format/tasks differ from literature reviews that are independent of a dissertation?

Student B: [...] Wouldn't identifying a research gap require a substantial review of what has been done, at least in the recent research history of the particular field? I think this is necessary at least to know that one is not wasting one's time duplicating research that has already been performed by someone else in the world. […]

Instructor: There is some overlap, but in review articles, your review is the study, so the purpose/objective is quite different. It's not about “setting up” a study (except perhaps in the article's introduction 1c), as it is for a dissertation literature review. In response to the second point, refer to my comment in the Boote and Beile article about the difference between reviewing as reading and reviewing as writing. [1 upvote, from an additional student]

![Image](https://link.springer.com/article/10.1007/s00586-020-06597-z)

**Fig. 3.** Number of upvotes (2, in this example) and “Upvoted by instructor” tag on a comment.

The students’ comments allowed the instructor to identify student confusion and provided an opportunity for the instructor to clarify and make connections with other course readings (Boote and Beile [13], Swales [14]). In another reading, student comments alerted the instructor to her own incorrect assumptions about student familiarity with a particular topic (systematic review articles), allowing her to adjust lecture content accordingly.

Student A: I've never actually seen this in a review, however, we have talked about it quite a bit in class. Does anyone know of an example where this is shown? [1 upvote]

Student B: I have never seen a review paper with those pieces of information. […]

Student C: For example you can see this paper: [https://link.springer.com/article/10.1007/s00586-020-06597-z](https://link.springer.com/article/10.1007/s00586-020-06597-z) The first line of Material and Method section contains those kind of information.

Instructor: Interesting that some of you have not seen many of this type of review article! I would have included some additional examples in today's lecture if I had realized this. […]

The above also illustrates one of many cases in which a student provided information to help their peers. This practice was not envisioned by the instructors prior to the implementation of Perusall, but the students seemed to value it and it was present throughout all the readings. Such
comments included definitions of terms, images that clarified a concept or study detail, and information from other sources for context, background, and comparison. Some examples:

- Curbweight/kerbweight = total mass of a vehicle with standard equipment and all necessary operating consumables such as motor oil, transmission oil, brake fluid, coolant, air conditioning refrigerant, and sometimes a full tank of fuel, while not loaded with either passengers or cargo. [4 upvotes]
- its a rather developed model! [https://www.researchgate.net/figure/THUMS-Version-40-brain-model_fig2_316602404] [3 upvotes]
- […] You can use sites such as https://www.color-blindness.com/coblis-color-blindness-simulator/ to check your figures prior to publishing. [2 upvotes]
- I wasn't sure what "parochial" meant so I looked it up. For those who might not know, it means "having a limited or narrow outlook or scope." (Oxford Languages) [1 upvote]

The WRIT instructor also noted that students asked opinions of the larger group of peers specifically, which was unexpected. Some examples:

- […] I've found it helpful to write out what I would say in a pretend conversation with someone, followed by revising into a written format. […] What approaches do others use on a first draft? [2 upvotes, 3 replies]
- Let us take a position on self-citation practice. What is your idea? Is it acceptable for you? I personally believe that self-citation alone is not a bad practice. I think any citation is appropriate when it is 'relevant.' And self-citation can result from the fact that the writer is a major contributor in that area. [6 replies]
- I've come across many examples of these in literature, but I've often been a bit resistant to including them in my own references. The reason is that they often do not require detailed peer-review or supporting evidence. Does anyone else feel the same? [3 replies]

The above examples demonstrate how Perusall establishes a community in which students question, engage with, and assist other students, often without instructor mediation and in ways not commonly observed in the traditional classroom. Both instructors also noted that as students relayed concepts in their own words, there was much more opportunity to correct misunderstanding of the content than would typically exist in the courses.

4.2.1. Comment coding. For TECH, one representative paper, assigned in week 3, was chosen for coding [15]. It had 23 students participate with an average of 9.7 comments per student (242 student comments in total). There were 46 instructor comments, (higher than the mean of 27.2 instructor comments). For WRIT, a paper was chosen from week 7 [16]. It included comments from 11 students, with an average of 6.1 comments per student (79 student comments in total) and 5 instructor comments (slightly lower than the mean of 8.1 instructor comments).

In the TECH paper, the longest threads were responses to instructor-seeded questions: 13 seeded questions, and 83 comments in response (6.4 comments per question). This effect seemed to reduce in later readings. There were many student questions and the majority were answered, with most answers coming from students (74%) rather than the instructor (26%). There were many comments stating observations from or opinions/assumptions about the paper (43%), and a more modest amount of critical analysis. This may reflect a fundamental focus on understanding the content for many students. There were more comments explicitly agreeing than disagreeing with earlier comments in a thread. Finally, there were very few social comments, and off-topic comments were essentially non-existent; the closest was "Such a great acronym I love it", referring to a clever acronym for the primary measurement device in the paper. A full breakdown is shown in Table 2.

The WRIT paper chosen for coding contained no seeded questions and 45 student-initiated threads, 25 of which had more than one comment. As with TECH, student questions were often answered by other students (80%) rather than the instructor (20%). There was also quite a bit of engagement with other course content, with students making connections both to lecture material (19%) and to other readings (14%). Statements of opinion or observation (48%) and instances of critical analysis (30%) were comparable to TECH, as were the percentages in most other categories. One notable exception is that WRIT students provided far less background/contextual information (2%) and made more references to personal experience (19%). This may be because engineering graduate students have more personal experience with writing than they have background in the field of the article (English for Specific Purposes or genre studies). Conversely, TECH students would have background knowledge of engineering concepts that may be relevant to an article on injury biomechanics, but less personal experience with the subject matter (e.g. car crashes).

4.2.2. Comment grading. The automated comment grading was not particularly able to discern high-quality comments. The score per comment seemed to rely primarily on comment length, which does not necessarily correlate with insightfulness. In fact, over 15% of the instructor comments in TECH received the lowest score. In WRIT, the instructor found that the lowest scored comments were not particularly insightful. Some examples of good quality, low-scoring student comments (TECH):

- I think that's okay as long as you have a method of bounding your error (i.e.: "my error doesn't exceed x, and that's ok because..."))
- I'm wondering if conditions like osteoporosis has a role? I'd assume that some people have stronger bones...
The instructors both tried adjusting the scoring scheme, which includes 16 factors, most of which are sliders to change the percentage weight of metrics. Both instructors found that grading was relatively insensitive to changes in the scoring settings for most factors.

Table 2: Coding of comments in each course for one representative reading. Some comments were assigned multiple codes, so the sum is >100%. Note that the comment numbers/percentages include both student and instructor comments.

| Code description | TECH (n = 288) | WRIT (n = 84) |
|------------------|---------------|---------------|
| Asking question about understanding of reading | 87 | 15 | 18% |
| Asking question about relationship of reading to other course content | 8 | 3% | 2% |
| Making connection with course content (except other readings) | 32 | 11% | 18% |
| Making connection with another course reading | 2 | 1% | 12% |
| Making connection with personal experience | 4 | 1% | 16% |
| Providing background/contextual information for others (external to course) | 64 | 22% | 3% |
| Critically analyzing reading content | 78 | 27% | 25% |
| Stating observation, opinion, inference | 123 | 43% | 40% |
| Responding to instructor prompt/question | 83 | 29% | 0% |
| Responding to instructor comment | 11 | 4% | 0% |
| Answering student question | 62 | 22% | 15% |
| Disagreeing with comment | 11 | 4% | 7% |
| Agreeing with comment | 27 | 9% | 20% |
| Suggesting different approach to that taken in reading | 15 | 5% | 15% |
| Correcting misunderstanding of reading content | 10 | 3% | 7% |
| Correcting misunderstanding of other course content | 2 | 1% | 0% |
| Highlighting minor error in reading | 8 | 3% | 0% |
| Highlighting a new-to-commenter idea/expressing surprise | 9 | 3% | 6% |
| Social response (e.g. "that's interesting", "thank-you", etc.) | 10 | 3% | 5% |

4.3. Surveys

We obtained good survey response rates from both classes (TECH: n = 15, 60%; WRIT: n = 8, 62%). The results of the Likert scale questions were statistically similar between courses (Mann-Whitney U Test, p > 0.90 for each prompt), so those results have been combined for both courses (Figure 4). TECH students reported spending 2.2 (SD 0.8) hours per reading. WRIT students reported spending 2.0 (SD 0.6) hours per reading, excluding one outlier who reported they spent 20 hours per reading.

In open-ended answers about the most successful aspects of using Perusall, students from both courses reported that they got more out of the readings than they normally would have, and that the discussions during the readings were interesting and provided different perspectives that led to deeper understanding. “Thinking critically about the paper content and trying to come up with insightful or educational comments helped me absorb it a lot more. I liked being able to build on other's discussions.” And “[…] some really interesting discussions happened in the annotations, often more interesting than the readings themselves.” Two students in TECH mentioned that instructor comments/questions, specifically including the initial seeded questions, were useful.

Students reported taking more time to read and critically assess the content than they would normally. Several students reported doing extra research and sharing that with classmates: “I found that I was actually looking up things I didn’t understand to make a comment about it so others would understand.”

Students said that they liked the interactivity of the system, noting “[…] with the collaborative system it was more dynamic and other students addressed challenging portions I might have otherwise skimmed over.” The students reported that the upvoting motivated them: “When you see a comment helped others […] it motivated me to make more thoughtful comments/questions.” Students also noted the higher levels of participation (whole class) compared to a typical in-person discussion.

There were several comments on the benefits of anonymity and time to think during the Perusall discussions: “Allowing me to do some research and cross-reference notes/previous readings led to more intelligent and thoughtful comments and discussions. Compared to traditional discussion settings, I felt significantly more comfortable being anonymous. The most successful aspect of Perusall in my opinion is that my grade didn't hinge on how brave I felt to raise my hand that day, but on the actual quality of my comments.”

When asked about the least successful aspect of Perusall in the courses, students most frequently mentioned technological or platform challenges. For example, several students had difficulty selecting sections of text or images.

Students also noted that they were still unclear on how the grading worked or what expectations were. This included TECH students, where the instructor tried to clarify the grading after similar comments during the term. A few students also reported (or suspected others of) trying to game the system for reading marks: “I have seen many people are trying to make comments that are longer than necessary […] I think people are trying to make big
descriptions to satisfy the scoring rubric” and “At times I just left comments to get my grade up and didn't focus on understanding the content.”

Several students also mentioned the time required: “[…] If used in a course I would say should be weighted 10% or so due to the effort that does go in.” A few students across both courses suggested having fewer readings. And one TECH student mentioned they didn’t know how the material would appear on quizzes, or how to study it.

Fig. 4. End-of-term survey results (Likert responses) for both courses combined (n = 23).

5. DISCUSSION

In this study, we found that Perusall was a useful tool for engaging with readings in upper-level and graduate courses in engineering.

There were surprisingly few differences in both the coding data and the survey responses of students between the two courses, despite their different natures (technical versus writing). Both instructors felt that course learning goals related to readings were met by using Perusall for discussions. The success in two very different courses suggests that the platform is versatile.

It was clear from both the survey and the quality and volume of annotations that students generally put effort into their comments and understanding the readings. They also went out of their way to share additional information. Most student questions were answered by other students, rather than the instructor, and students explicitly solicited opinions from other students (WRIT). This analysis of the annotations aligns with the survey data, which suggests that the social nature of the readings led students to make more and better-quality comments. Overall, the results are consistent with the social constructivist framework, where interaction leads to learning through the generation of common understanding [5].

Instructors observed a dramatic increase in discussion participation over in-class discussions from past sessions. Some reasons for this may be the anonymity of the system and the time students had to consider the readings and existing annotations before commenting, as well as the larger grade weight. Research related to speaking in class has shown that students feel that they need to give the right answer immediately, and expect that peers will form opinions of them based on their contributions and their choices to challenge other students [17]. As such, choosing to speak up in class is a high-stakes endeavour [17]. The asynchronous and anonymous nature of Perusall reduces some of these pressures. When attempting to use Perusall for in-class annotation, these benefits were stripped away. This might explain why in-class participation in Perusall annotation was much lower.
Persuall allowed students some self-direction within their learning, possibly more than in an instructor-led class discussion. They could choose to focus on aspects of the reading that interested them. Anecdotally, the TECH instructor noted several students who returned to their own particular themes of interest over a number of readings. Self-determination theory suggests that both autonomy in learning and relatedness (social interaction with others) are associated with higher motivation for learning [18]. This theory may partly explain the observed high levels of high-quality participation in collaborative annotation.

Our results were consistent with those of Lee and Yeong, who found that undergraduate cell biology students reading research papers with Perusall provided a range of comments, from definitions to critical analyses [9]. They noted interactions where students corrected peer misunderstandings (and the opportunity for instructors to do the same) and exchanges going beyond the immediate content to larger questions about subjectivity and interpretation. They also noted the benefits of anonymity for students who would not otherwise participate in the classroom. Overall, they saw similar high engagement and substantial learning through interaction on Perusall [9].

The strengths of this work include gathering data from two courses with quite different foci. We used comments, comment coding, and survey data, to allow triangulation of results. Limitations include the small class sizes, a population of upper-year undergraduate and/or graduate students only, coding only one paper per course, and the use of the system in an online-only scenario. It is not clear if the results would be similar with different class sizes, year levels, or in-person classes. Survey results suggest that students would be open to using Persuall once we return to in-person classes, which both instructors plan to do.

Overall, our findings suggest that collaborative annotation of readings is an effective class discussion approach in upper year/graduate engineering courses.

Acknowledgements

The collection of data about these courses is considered to be program evaluation by the research ethics board at the University of British Columbia and therefore was exempt from requirements for formal ethics review. The authors followed ethical principles in conducting data collection, including asking for consent, providing information about data usage, and maintaining confidentiality.

This work was completed on the traditional, ancestral, and unceded territory of the Musqueam, Squamish, and Tsleil-Waututh people.

References

[1] “Perusall website.” [Online]. Available: https://perusall.com. [Accessed: 26-Nov-2020].
[2] C. Suhre, K. Winnips, V. De Boer, P. Valdivia, and H. Beldhuis, “Students’ experiences with the use of a social annotation tool to improve learning in flipped classrooms,” in 5th International Conference on Higher Education Advances (HEAD’19), 2019, pp. 955–962.
[3] M. W. Liberatore, “Annotations and discussions of textbooks and papers using a web-based system (work in progress),” ASEE Annu. Conf. Expo. Conf. Proc., 2017.
[4] K. Miller, B. Lukoff, G. King, and E. Mazur, “Use of a Social Annotation Platform for Pre-Class Reading Assignments in a Flipped Introductory Physics Class,” Front. Educ., vol. 3, no. March, pp. 1–12, 2018.
[5] B. Kim, “Social Constructivism,” Emerg. Perspect. Learn. Teach. Technol., no. 2006, 2006.
[6] National Research Council, Education for life and work: Developing transferable knowledge and skills in the 21st century. National Academies Press, 2012.
[7] R. S. Prawat and R. E. Floden, “Philosophical Perspectives on Constructivist Views of Learning,” Educ. Psychol., vol. 29, no. 1, pp. 37–48, 1994.
[8] M. Barak, “Science Teacher Education in the Twenty-First Century: a Pedagogical Framework for Technology-Integrated Social Constructivism,” Res. Sci. Educ., vol. 47, no. 2, pp. 283–303, 2017.
[9] S. C. Lee and F. M. Yeong, “Fostering student engagement using online, collaborative reading assignments mediated by Perusall,” Asia Pacific Sch., vol. 3, no. 3, pp. 46–48, 2018.
[10] B. Adams and N. S. Wilson, “Building Community in Asynchronous Online Higher Education Courses Through Collaborative Annotation,” J. Educ. Technol. Syst., vol. 49, no. 2, pp. 250–261, 2020.
[11] J. K. Foster, J. O. Kortge, and M. J. Wolanin, “Hybrid III — A Biomechanically-Based Crash Test Dummy,” SAE Trans., vol. 86, no. 4, pp. 3268–3283, 1977.
[12] J. A. Maxwell, “Literature Review of, and for, Educational Research: A Commentary on Boote and Beile’s ‘Scholars Before Researchers,’” Educ. Res., vol. 35, no. 9, pp. 28–31, 2006.
[13] D. N. Boote and P. Beile, “Scholars Before Researchers: On the Centrality of the Dissertation Literature Review in Research Preparation,” Educ. Res., vol. 34, no. 6, pp. 3–15, 2005.
[14] J. M. Swales, Genre Analysis: English in academic and research settings. New York, NY: Cambridge University Press, 1990.
[15] M. Hitosugi, T. Koseki, T. Hariya, G. Maeda, S. Moriguchi, and S. Hiraizumi, “Shorter pregnant women restrained in the rear seat of a car are at risk for serious neck injuries: Biomechanical analysis using a pregnant crash test dummy,” Forensic Sci. Int., vol. 291, pp. 133–137, 2018.
[16] B. Kanoksilapatham, “Distinguishing textual features characterizing structural variation in research articles across three engineering sub-discipline corpora,” English Specif. Purp., vol. 37, no. 1, pp. 74–86, 2015.
[17] M. M. Reda, “What’s the Problem with Quiet Students? Anyone? Anyone?,” Chron. High. Educ., vol. 57, no. 3, p. A68, 2010.
[18] E. L. Deci, R. J. Vallerand, L. G. Pelletier, and R. M. Ryan, “Motivation and Education: The Self-Determination Perspective,” Educ. Psychol., vol. 26, no. 3&4, pp. 325–346, 1991.