

Promoting Active Learning Through Collaborative Annotation Tools

Tuiloa, S. H., Arnesen, K. T., Guo, Q., West, R. E., & Graham, C. R.

Active Learning

Annotation

Collaboration

collaborative annotation

content interaction

deep reading

digital annotation

learner engagement

social annotation

Student Interaction

In this study, we report findings on how digital social annotation can generally be an effective means of helping students engage deeply with text, although assignment requirements and classroom dynamics may also play a factor. Various learning strategies are widely used as a means to kindle active learning and promote higher levels of content and peer interaction in online education. One strategy growing in popularity is the use of digital collaborative annotation tools which allow students to mark and annotate digital text as well as communicate with and learn from the annotations of their peers. This paper discusses a multiple-case study which evaluated the effectiveness of Hypothesis, a digital collaborative annotation tool, in helping students to engage actively with reading assignments and with their peers.

Introduction

Helping students actively engage with both course content and their peers is a priority in online classes. Various learning strategies can help students engage in active learning and encourage higher levels of learning and comprehension in the online setting ([Bean, 2011](#); [Garrison, 2016](#); [Young, 2014](#)). While online discussion boards are among the most common of these strategies ([Garrison, 2016](#)), one method increasing in popularity is the use of collaborative annotation tools which allow students to highlight and annotate text and communicate directly with their peers within the context of the text ([Zyrowski, 2015](#)).

Since the growing integration of digital texts in education has led to decreased reading comprehension and understanding, along with challenges in engaging in deep reading ([Ben-Yehudah & Eshet-Alkalai, 2014](#)), studies indicate effective use of collaborative annotation tools can foster deeper learning ([de Oliveira Neto & Dobslaw, 2024](#); [Hanč et al., 2023](#)). However, further exploration of how to implement collaborative annotation tools to enhance educational experiences is needed ([Morales, 2022](#)). In order to fill this gap, this study employs a multiple case-study approach to evaluate the effectiveness of this method in two university classes, to demonstrate ways to implement the collaborative annotation tools, and to explore the value of using collaborative annotation to aid students in deep reading of texts while also engaging with peers.

Literature Review

Modern educational theory is often based on social constructivism, which has two key components. First, students must actively construct their own learning, and second, they do this through social interactions ([Polly et al., 2018](#)). In conjunction with this, current educational research regarding online learning underscores the necessity of both social and content interactions in helping students achieve the higher levels of thinking required in university-level courses ([Anderson & Dron, 2011](#); [Dawson et al., 2011](#); [Garrison et al., 2010](#)). Garrison's ([2016](#)) review of the literature primarily pointed to online (usually asynchronous) group discussions as a strategic means to accomplish both interactions with content and with learning community members. One strategy to help students connect with content is deep reading—a means whereby students actively construct their own learning through deliberate and contemplative reading which goes beyond a superficial reading of a text ([Birkerts, 2006](#)).

If higher levels of learning and reading comprehension are to be achieved in online environments, students need to expand past a surface-level understanding. Unfortunately, increasing the use of digital texts within education has resulted in lower levels of reading comprehension and understanding, as well as an inability to engage in deep reading ([Ben-Yehudah & Eshet-Alkalai, 2014](#)). For example, some studies have shown that text comprehension on digital screens is inferior to text comprehension in print ([Ackerman & Goldsmith, 2011](#); [Ackerman & Lauterman, 2012](#)). However, other studies dispute this. For example, one study used both print and online texts from various newspaper publications and tested participants on how well they could retain and comprehend information from both types of sources. While students generally preferred reading a printed text and had a general distrust of online texts, reading comprehension was not significantly different ([Young, 2014](#)). There seems to be a consensus that various learning strategies can be implemented to increase a student's ability to engage in deep reading and comprehension, including engaging online with peers ([Ben-Yehudah & Eshet-Alkalai, 2014](#); [Young, 2014](#); [Zyrowski, 2015](#)).

One strategy to increase this deep reading engagement is the growing use of online digital annotation tools as an alternative to online asynchronous discussions ([Zyrowski, 2015](#)), which can help students actively construct their learning ([Gao, 2013](#); [Peng, 2017](#)), increase their motivation to read ([Clinton-Lisell, 2023](#); [Cui & Wang, 2023](#)), and improve students' academic performance ([Clinton-Lisell, 2023](#); [de Oliveira & Dobslaw, 2024](#)). Clinton-Lisell's (2023) quantitative study revealed students reported higher motivation levels when using social annotation compared to traditional quizzes. This increased motivation was linked to more active reading of textbooks, suggesting social annotation encourages deeper engagement with course

materials. The same study also found a positive correlation between the use of social annotation and course grades (Clinton-Lisell, 2023). Students who engaged in social annotation tended to achieve higher grades, suggesting social annotation tools not only enhance engagement but also contribute to improved academic performance.

Annotation is a long-held practice referring to the act of jotting brief notes and summarizing main points in the text's margins (Nist & Simpson, 1988). Those interested in deeply understanding a piece of text often seek out annotated versions to support understanding. In a few instances, an author's own annotations have been known to be just as important as the text itself, as the annotations have provided further insight into a writer's perspective (Liu, 2006).

Traditional annotation can be done in many ways (Allington, 1977), and various studies have argued that annotation is a useful comprehension tool as it forces readers to actively engage with the text, making them participants in the learning process and, therefore, achieving higher levels of comprehension and understanding (Fowler & Barker, 1974; Leutner et al., 2007). Digital annotation tools allow students to mark and annotate a digital text similar to reading/annotating a printed version of text and tend to produce the same results.

While annotation—on either digital or paper—does seem to aid students in engaging more deeply with the content, digital annotation tools can also increase collaboration to promote active peer-to-peer interaction (de Oliveira & Dobslaw, 2024; Zyrowski, 2015). Many digital annotation tools also have collaborative features, which allow individuals to engage with each other in addition to marking the text. These tools allow students to see each other's annotations as well as their own and engage in conversations with each other directly within the context of the text. Annotation can go beyond helping an individual reader summarize the main ideas; it can also help readers communicate their thoughts effectively to others (Zyrowski, 2015). Zyrowski (2015) noted this approach “takes annotation to a collaborative, multi-modal level” (p.21) and can help students become active agents in a social learning process. In their mixed-method-research, de Oliveira and Dobslaw (2024) implemented Perusall—a social annotation tool—in an introductory programming course. Students reported this collaborative approach increased their interest in the subject matter. The findings highlighted the role of social annotation in fostering a collaborative learning environment. Other research discovered that digital annotation's role in facilitating students' discussion and collaboration did not inherently promote collaboration or change the learning experience (Chan & Pow, 2020).

Despite these mixed results, online social annotation of readings could prove to be beneficial as a supplement to or replacement for online asynchronous discussions. While research shows that the proper use of asynchronous online communication can kindle higher levels of learning (Garrison, 2016), there is still much to be understood about implementing online learning tools and activities to create valuable educational experiences. For example, some research indicates online group discussions result in trivial conversations that fail to reach beyond agreeing, sharing, and comparing within a group conversation (Klemm & Snell, 1996). A better understanding of how online collaborative annotation tools can effectively facilitate this kind of learning can aid in mitigating the trivial and surface-level reading comprehension occurring in the online space and promote expectations for higher levels of learning.

In an effort to better understand the effectiveness of collaborative annotation tools to help students engage in actively constructing their learning through deep reading and meaningful peer-to-peer interaction, we studied how university students engaged with academic texts and with each other when provided a digital annotation tool. We also asked students what they perceived to be the strengths and weaknesses of using a collaborative annotation tool for engaging with academic texts and their peers.

Methods

We utilized a multiple case study approach to investigate the effectiveness of Hypothesis, a collaborative digital annotation tool, in helping students engage with academic texts and each other in one undergraduate-level university course and one

graduate-level university course. Studying two separate cases helped us better understand if this method could be successful across various education levels, courses, and assignment types. The research questions for this study were:

1. How do university students engage with academic texts and with each other when using collaborative annotation tools?
2. What do students perceive as the strengths and limitations of using collaborative annotation tools for engaging with academic texts and their peers?

Research Settings

We utilized a learning management system (LMS) integration version of the online collaborative digital annotation tool Hypothesis. This collaborative annotation tool is an open-resource web extension that anyone on the internet can download, create an account for, and use to annotate a digital text online (Hypothesis, 2020). Users have the ability to make their annotations private or publicly viewable to other Hypothesis users. The LMS integration tool allows teachers to use this tool for assignments, readings, and collaboration with the students in their courses. The Hypothesis team claims their collaborative annotation tool provides a more active and social reading experience which allows students to engage at a deeper level and in more meaningful ways with the text, their peers, and instructors (Hypothesis, 2020). The LMS integration works with a number of LMSs including Brightspace, Canvas, Sakai, Blackboard, Moodle, and Schooner (Hypothesis, 2020). Both cases we studied were piloting Hypothesis for their course and utilized the integration within the Canvas LMS.

Participants

Both the undergraduate-level course and the graduate-level course we studied took place in the school of education at a large private university in the West/Midwest. We selected these two cases because both courses were piloting Hypothesis, and the instructors were seeking to understand the effectiveness of this new teaching and learning method.

Graduate Students. The graduate-level course was a three-credit introductory research class consisting of nine students taught during a six-week spring term. The class met twice a week via Zoom for three hours per class session.

Undergraduate Students. The undergraduate course taught preservice education students how to teach and facilitate online and blended learning classrooms. The one-hour course consisted of 50 students and was taught completely online. Because this version of the course was completed during a six-week spring term, the class met with the instructor twice a week for 50 minutes per session.

Data Collection

The data collected was different for each course due to the way each course used social annotation. See Table 1 for assignment specificities.

Table 1

Undergraduate and Graduate Course Hypothesis Assignments and Reflections

Data Collected	Description
Graduate Course Reading Annotation Assignments	Annotations and peer comments from 24 reading annotation assignments, comprising 1,004 total annotations analyzed. The description of the assignment was as follows: • This assignment is designed to help you slow down and engage deeply with the concepts in the chapters and articles you will read in this course.

- You should make at least two annotations in every article that extend the reading and provide additional insight for your peers, or that extend the discussion.
- Suggestions include: mark unfamiliar terms & provide definitions; link to additional resources with summary; pose a question; add an insight; respond to another student; note patterns or themes; agree or disagree with a concept or claim in the text or one posted by a classmate.

Undergraduate Course Reading Annotation Assignments	Annotations and peer comments from five reading annotation assignments, comprising 1,332 total annotations analyzed. The description of the assignment was as follows: • Create five qualifying annotations within this chapter. • Each annotation should be tagged with the level of Bloom's taxonomy in which you are engaging by creating the post. • At least two of five qualifying annotations must be at higher levels of Bloom's Taxonomy (#Analyze, #Evaluate, or #Create).
Undergraduate Course Reflection Questions	As part of their class assignments, we collected a weekly reflection from undergraduate students about their experiences using Hypothesis, as well as a cumulative reflection at the end of the course. Weekly Reflection Question: • What is working well and not working well for you in the Hypothesis assignments? • Do you feel that Hypothesis is helping you to read differently than you normally would? If yes, how so? End-of-Course Reflection Questions: • Did this tool help you to better engage with the text? If so, how? • How do you feel using this tool as a means of assessment compared to taking quizzes to test your knowledge of the material or discussing the material with classmates via a discussion board? • What are the strengths and/or weaknesses of this new method? Explain. • Do you think you might incorporate something similar or alter this method in your future classroom? If so, how? If not, why not?

Data Analysis

We conducted descriptive statistics and qualitative thematic analysis for both undergraduate and graduate students' annotations and peer comments to answer research question 1. We also conducted a qualitative thematic analysis of students' reflections on social annotation. Since the graduate course did not ask the students to evaluate the tool, we used only undergraduate students' reflection responses to answer research question 2. See Table 2 for an overview of our Data Analysis.

Table 2

Research Questions, Data Collected, and Data Analysis Methods

Research Question	Participants	Data	Analysis
-------------------	--------------	------	----------

RQ #1 How do university students engage with academic texts and with each other when using collaborative annotation tools?	Graduate and undergraduate students	Student annotations and peer comments in texts from course assignments.	Descriptive statistics including number of replies to student annotations and number of annotations in each section of the text. Qualitative thematic analysis with a priori codes described in Table 3.
RQ#2 What do students perceive as the strengths and limitations of using collaborative annotation tools for engaging with academic texts and their peers?	Undergraduate students	Student reflections about social annotation.	Qualitative thematic analysis of reflection responses.

Descriptive Statistics

We calculated descriptive statistics in both classes to gain an overall understanding of how students were interacting with different sections of the text and with each other. We divided each text into three equal sections and noted the total number of annotations in each section of the text to see if students were engaging with the text more at the beginning of a text or the end. We also counted the number of replies on an original annotation to see how often students interacted with each other.

Thematic Analysis of Annotations

Our qualitative thematic analysis process helped us understand what types of annotations students were making and determine if they were actively engaging with the text by adding insights beyond what the text already said. Using the same method for both courses, we established a list of a priori codes which were adapted from Fahy et al., (2001) and Zhu (1996), who both look at patterns of interaction online. We established all of the codes prior to coding and refined a few definitions during our initial coding process. To ensure reliability, we calculated interrater reliability percentage to measure agreement between coders. For every article, two researchers independently coded the annotations. After initial coding, we compared results and identified discrepancies. Disagreements were discussed and resolved through iterative dialogue, referring back to the codebook and making minor refinements as needed to maintain consistency. If consensus could not be reached, a third researcher was consulted to finalize the coding decision. See Table 3 for the list and definitions of the a priori codes, along with examples of each category in both the undergraduate and graduate courses.

Table 3

A Priori Codes for Qualitative Analysis of Annotations

Code	Definition	Examples
Information-Seeking Questions	Asking questions which seek further information and clarification with a specific answer.	Undergraduate: "So, why do blended courses cost money?" Graduate: "Does anyone know the definition of 'familial privatism'?"

Discussion-Promoting Questions	Asking questions which seek to start a dialogue with peers and/or get opinions from experts. Answers to these questions may not have a clear answer.	Undergraduate: “This is a really good idea but I question what this will look like in the classroom. How can we co-design learning experiences with our students when there are 20+ students in each class?” Graduate: “I honestly love this idea of giving students more agency and power in their learning! But, how do we determine grades and meet university standards if there are no assessments?”
Reflection	Personal thoughts, judgments, and opinions. Reveals one’s personal values, beliefs, and ideas. (Includes rhetorical questions).	Undergraduate: “I wish more teachers would do this! I did not feel like my high school prepared me for college or my future. Looking back, I definitely see the things I missed out on that would have better prepared me for college and my career. I was often on my own to make decisions regarding college, classes, and my major. This makes me want to be the kind of teacher that will help students with this!” Graduate: “I think this is a really important distinction to make when we are considering the ethical issues surrounding data mining. People may not consider the many ways their public data could be used. There definitely needs to be a process put in place to approve the use of this type of data for research.”
Explanation/Scaffolding	Providing an explanation/definition/outline of what is going on in the text to scaffold the learning of others reading the article. For example, it might be an outline of the article, or a definition of a keyword, or an explanation of the idea.	Undergraduate: “The AAA process helps teachers work effectively with data: 1. Ask- what is the question you want to ask of the data? 2. Analyze- analyze the data for patterns to answer your question. 3. Act- this could include adjusting learning activities or ways of gathering new data.” Graduate: “Metacognition - ‘awareness and understanding of one’s own thought processes.’”
New/ Outside Ideas	Introduce new ideas not covered in the text and/or quoting or referencing an outside source.	Undergraduate: “This article goes over how to teach students to give good feedback to their peers: “https://www.edutopia.org/article/teaching-students-give-peer-feedback” Graduate: “I think one of the earliest epistemology studies may have been [Plato’s Allegory of the Cave] (https://en.wikipedia.org/wiki/Allegory_of_the_cave).”
Evaluation	Evaluating the given text. Agree/disagree with the text.	Undergraduate: “I agree with the statement that just because there are non-negotiable outcomes, it doesn’t necessarily mean that instruction can’t be personalized, or at least interesting.” Graduate: “I disagree with this. Don’t we ask questions like this? There are plenty of studies that have addressed these types of questions. Individual studies on specific perspectives could be

looked at in aggregate or a meta-analysis performed to address this.”

Generic Comments	General comments which do not add additional meaning or insight and/or do not promote further conversation.	Undergraduate: “I love that you chose a meme, those are my favorites. I couldn't even get a picture to paste onto the spot.” Graduate: “Amen!”
------------------	---	---

Thematic Analysis Process of Reflection Questions

In addition to the coding of the annotations themselves, we also coded the weekly and final exam reflections of the undergraduate students, which explored their perceptions of studying a text using the annotation tool. During the two weeks the students read the textbook, they submitted weekly reflections twice. Some of these reflections (n=48) included their thoughts and opinions about Hypothesis. The final exam asked questions about students’ experiences with the tool, their opinions about the tool as an assessment, especially when compared to quizzes or discussion boards, and their plans for using the same or a similar tool in their own teaching.

The second author and a research assistant independently coded these reflections, building two different coding structures based on the raw data. Once they finished coding, they met together to compare the codes and coding structures and to create a new coding structure that both felt accurately represented the richness and variety of the data. The second author used this new coding structure to recode the data, which the lead author then reviewed. They discussed any discrepancies and established a consensus.

Findings

We organized our findings based on the two research questions addressed in the study.

Question 1: How did Students Engage with each other and the text using Collaborative Annotation?

Graduate Course

The texts used in the graduate course were journal articles that introduced fundamental principles of research. Each annotation was coded for the primary purpose of the annotation. We classified the main content of an annotation or comment as the primary purpose of the annotation and the secondary expression as the secondary purpose of the annotation. The results are summarized in Table 4.

Table 4
Graduate Course—Number of Each Type of Annotation (n=1,004)

Type of Annotation	Primary Purpose Code n=1004	Secondary Purpose Code n=40	Total codes n=1044
--------------------	--------------------------------	--------------------------------	-----------------------

Discussion-promoting Questions	54 (5.4%)	1 (2.5%)	55 (5.3%)
Evaluation	87 (8.7%)	2 (5%)	89 (8.5%)
Explanation/Scaffolding	221 (22.0%)	3 (7.5%)	224 (21.4%)
Generic	128 (12.7%)	0 (0%)	128 (12.3%)
Information-seeking Questions	75 (7.5%)	6 (15%)	81 (7.7%)
Introduction of New/outside Ideas	108 (10.7%)	18 (45%)	126 (12.1%)
Reflection	331 (33.0%)	10 (25%)	341 (32.7%)

Occasionally, students expressed more than one type of comment in their annotation or peer comments. In 4% of the total annotations, researchers identified a secondary purpose in the student annotations. Table 4 shows the most common secondary purpose was the introduction of an outside idea (18) followed by a reflection (10), then information-seeking questions (6), explanation/scaffolding (3), evaluation (2), and discussion-promoting questions (1).

In the graduate course, the highest number of annotations (33.0%) were in the reflection category: personal thoughts, judgments, and opinions. Interestingly, the categories with the lowest percentages—discussion-prompting and information-seeking, with a combined primary purpose code total of 12.85%—were the only two types of codes that inherently invited responses or feedback from others in the class. However, out of the total number of annotations (1,004), students responded to another student's post about three times that much (37%). This means they frequently chose to respond to other types of annotations as well as those that implicitly seemed to call for responses. Simply put, even when original posts were not inviting discussion with each other, students chose to respond, discuss, and engage with each other through the collaborative annotation tool. Out of all the annotations made, 633 (63%) were original annotations, while 371 (37%) were responses to original annotations. Surprisingly, over half of annotations (61.1%) did not receive a reply. While the majority of original annotations with a response only had one response (23.2%), some annotations received up to four responses, with one annotation receiving five responses (see Table 5).

Table 5

Number of Graduate Responses and Original Annotation Received (n=633)

# of original annotations	# of replies received	Percentage of annotations
387	0	61.1%
147	1	23.2%
74	2	11.7%
21	3	3.3%

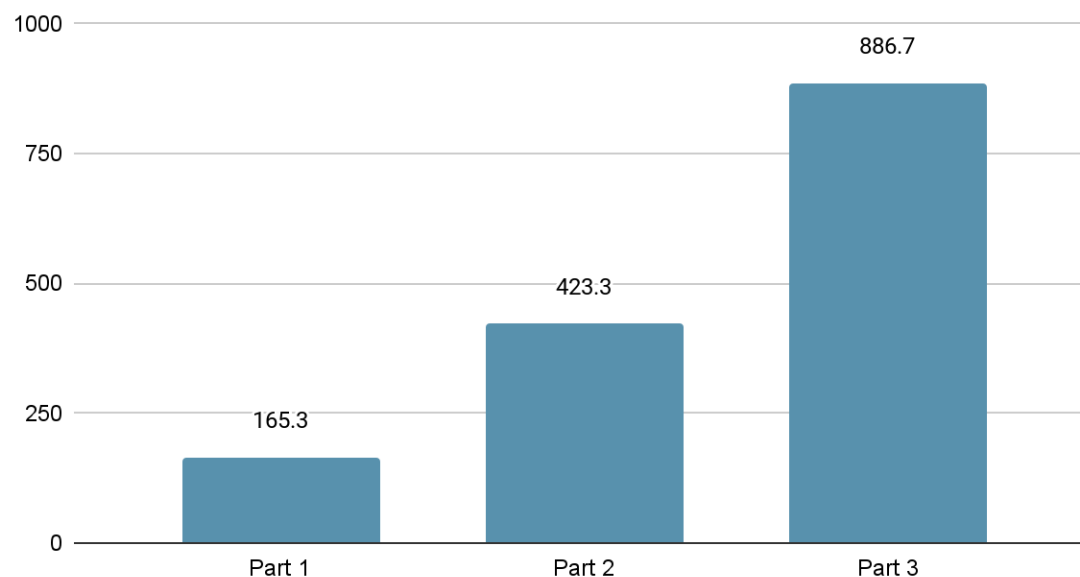
3	4	0.5%
1	5	0.2%
0	6	0.0%

In addition to the types of annotations students wrote, the location of annotations within the reading also highlighted how students were engaged throughout the entirety of the text. In almost every single reading, there were significantly more student annotations in the first third of the text, and the number of annotations dropped significantly by the end of the text. Figure 1 displays the average number of words between each annotation in the text and shows that the later parts of the texts had fewer annotations as there was increasingly more text between each annotation.

Figure 1

Graduate: Average Number of Words Between Posts for Each Part of the Text

Graduate Average Number of Words Between Posts



Undergraduate Course

The undergraduate students used the annotation tool as they read *K-12 Blended Teaching: A Guide to Personalized Learning and Online Integration*, the required text for the class. The text introduced and developed concepts and principles of blended teaching, which were both theoretical and practical. Each annotation was coded for the primary type of annotation category it belonged to. The results are summarized in Table 6.

Table 6

Undergraduate Number of Each Type of Annotation (n=1,332)

Type of Annotation	Primary Purpose Code n=1,332	Secondary Purpose Code n=89	Total codes
--------------------	------------------------------	-----------------------------	-------------

n=1,421

Discussion-promoting Questions	174 (13.1%)	25 (28.1%)	199 (14.0%)
Evaluation	125 (9.4%)	16 (18.0%)	141 (9.9%)
Explanation/Scaffolding	79 (5.9%)	2 (2.2%)	81 (5.7%)
Generic	8 (0.6%)	0 (0.0%)	8 (0.6%)
Information-seeking Questions	27 (2.0%)	3 (3.4%)	30 (2.1%)
Introduction of New/outside Ideas	102 (7.7%)	5 (5.6%)	107 (7.5%)
Reflection	817 (61.3%)	38 (42.7%)	855 (60.2%)

A substantial majority (61.3%) of the first-order codes were in the reflection category: personal thoughts, judgments, and opinions. When the 38 secondary reflection category codes were added, the percentage of total reflection annotations rose to 60.2%. In contrast, only 16.1% of the annotations (including secondary purpose codes) invited a response by either seeking information or promoting discussions, which were the only two categories that inherently promoted discussion. These numbers could indicate undergraduate students did not see the use of the tool (or the specifics of the assignment) as something to promote dialogue. The number of original annotations versus responses seems to support this interpretation as 76.4% of all annotations were original annotations, while only 23.6% of all annotations were a response to an original annotation. Of the original annotations, only 131 (12.9%) received even one response, with only 78 (7.6%) receiving more than one response (see Table 7).

However, a common theme in student responses to the assignment was that there were too many comments in the chapters, and it was difficult to find time to even read all of them. Responding to the comments was not a priority. For example, each of the 50 students was supposed to comment five times in each of the five chapters. Thus, each chapter would have 250 comments. Several students mentioned they hid all but their own annotations because of the volume of comments. Interestingly, 21.5% of the students made connections between the text and their own experiences, with 286 annotations including personal experiences.

Table 7

Number of Undergraduate Responses and Original Annotation Received (n=1,018)

# of original annotations	# of replies received	Percentage of annotations
809	0	79.5%
131	1	12.9%
56	2	5.5%
16	3	1.5%
3	4	0.3%

2	5	0.2%
1	6	0.1%

As with the graduate students, more undergraduate students' annotations were in the reflection category than in any other category. However, these annotations were a much higher percentage of the overall comments. Approximately two-thirds of the undergraduate annotations were reflections, while just under one-third of the graduate students' annotations were reflections.

Graduate students demonstrated a broader diversity in their annotation types compared to undergraduates. In undergraduate courses (Table 6), annotation types ranged from a low of 0.6% (Generic) to a high of 61.3% (Reflection), a spread of 60.7%. In contrast, graduate students' annotations (Table 4) ranged from 5.4% (Discussion-promoting Questions) to 33% (Reflection), with a narrower spread of 27.6%.

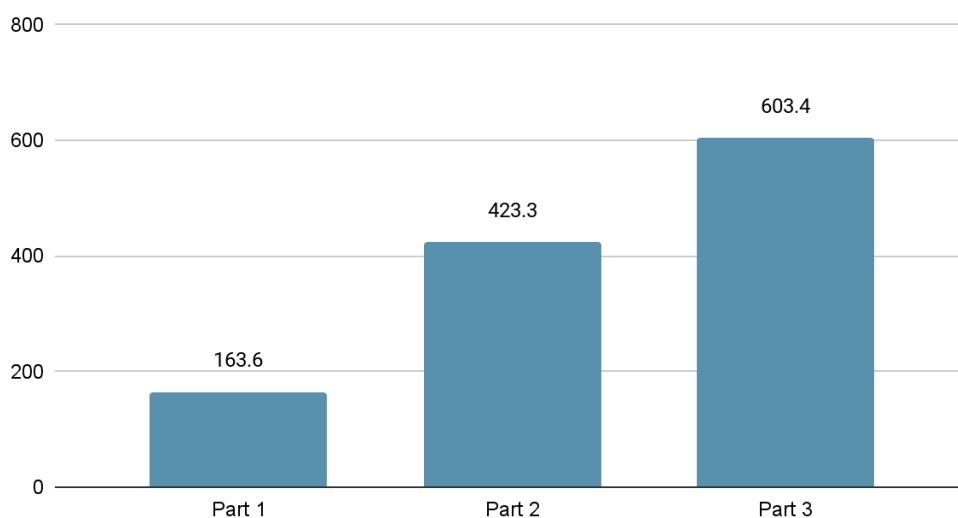
In addition, the graduate students tended to make more annotations than the undergraduate students. The undergraduate students made an average of 26.64 (an average of 1.64 more than the 25 required) during the term, while graduate students made an average of 111.56 annotations (an average of 63.6 more than the 48 required) during the term, and 4.19% more than the undergraduates.

The undergraduate students followed a similar pattern as the graduate students with their levels of engagement throughout the entirety of a text. The frequency of student annotations on the text noticeably decreased once they had created the required five annotations (see Figure 2).

Figure 2

Undergraduate: Average Number of Words Between Posts for Each Part of the Chapters

Undergraduate Average Number of Words Between Posts



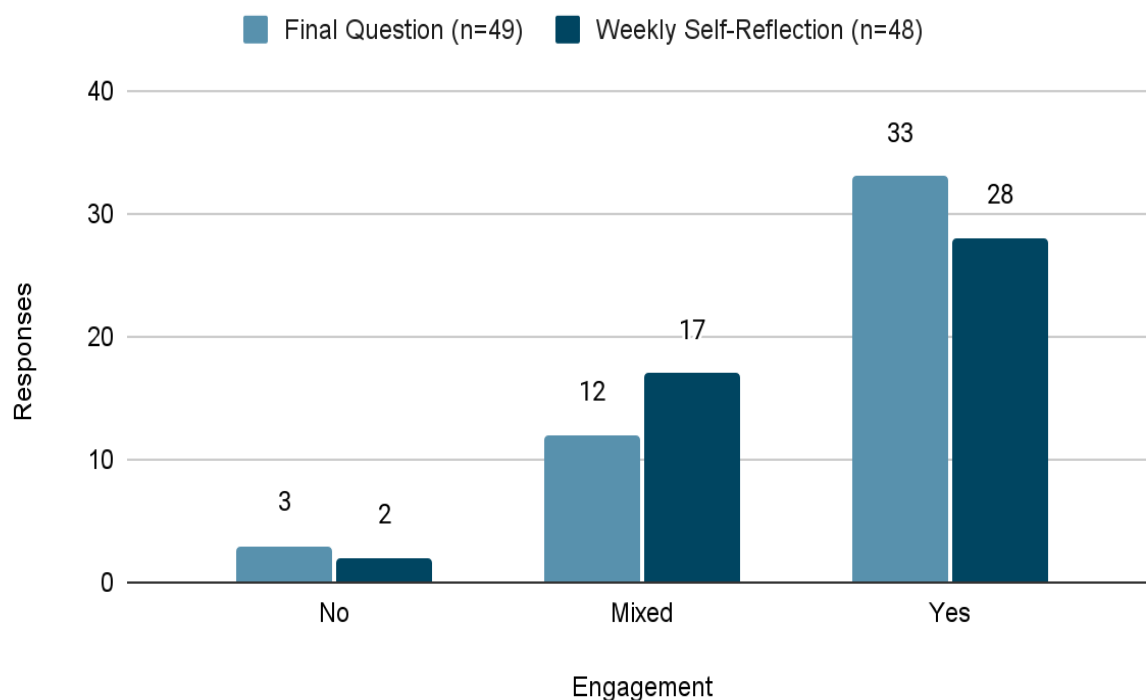
Question 2: Student Perceptions of Collaborative Annotation Tools

Students commented on the effectiveness of the tool on their engagement, their preference for using Hypothesis, a quiz, or a discussion board to assess their reading, and the possibilities of using an annotation tool with their own future students. In both the weekly and final reflections, the students indicated that the Hypothesis helped them engage with the text. Figure 3 shows the results.

Figure 3

Undergraduate Responses to Hypothesis Use

Undergraduate Engagement with the Text



Note. This graph shows student responses to a question asked on their weekly self-reflection and a reflection question on their final exam regarding engagement: “Did this tool help you better engage with the text? If so, how?”

During the two weeks of reading assignments, undergraduate students submitted weekly reflections on their perceptions of their learning and work for that week. Forty-eight students included their perceptions of the Hypothesis tool. Of these 48 students, 29 felt the tool had increased their engagement with the text, 17 had mixed comments, and two felt it had negatively influenced their engagement. Positive comments included that it was “working well,” “it is definitely helping me read differently than normal,” and that students thought “more deeply” about the text because they were annotating it. Although an uncommon response, one student really enjoyed looking for answers to “questions that my classmates have posed. That’s my favorite part of the whole thing. I think that posting questions based on the text is a good use of collaboration and it feels really genuine!”

Most of the mixed comments focused on three themes: students were confused with how their comments would be graded and with having to tag their comments with a level of Bloom’s taxonomy; Hypothesis interfered with links in the text, making viewing the embedded videos difficult; and students felt overwhelmed by the number of people commenting on the text.

The opinions of the two students who felt the tool negatively influenced their engagement could be summed up in one of their comments: “I am honestly not a fan of the Hypothesis assignments.”

The students' reflections in their final exam took a more balanced and nuanced view of how Hypothesis influenced their reading. In these reflections about two-thirds of the students (33 of 49) reported Hypothesis had positive effects on their engagement with the text. They noted using the tool and having to make annotations helped them attend better to the text and think more deeply about it. One reported having to annotate the text "definitely" helped her "record a different type of thought than I would by just reading or even reading and taking notes." Another explained, "I really had to comprehend what I was reading so that I could make my comments." Another found she interacted more with the text than she normally would, "looking up related ideas, articles, and videos than I usually do to learn more. I also spent more time thinking about how what I was learning related to my other classes."

An additional 17 students found the Hypothesis tool increased their engagement with the text but had a few caveats which tempered their feelings about it. Some mentioned they read the text and thought deeply about it only until they had completed their five annotations: "Once I hit the required 5 comments, I tended to start skimming." Others found the sheer volume of comments from their peers overwhelming and often hid them, or they felt requirements for the types of comments they had to make hindered them from making annotations which were meaningful to them. One student explained the problem, saying: "I would have preferred to simply read the text and comment on what I found interesting, make notes where I wanted to, etc. I didn't like having to meet requirements for using Hypothesis [annotating the comment with a level of Bloom's taxonomy] because I found it to be restricting and less effective."

Three students found no value in using the tool. One said, "I felt like I was generally annoyed by this tool. . . . I already take notes when I'm reading, so this was just an extra and less convenient way to do it." Another was indifferent, feeling it did not change "how I interacted with the text at all," and the other felt "it made my reading comprehension worse."

Although students generally felt using Hypothesis to annotate the text helped them attend better to the text, the frequency of their comments on the text noticeably decreased once they had created the required five annotations (see Figure 2 above). Each reading consisted of three parts, and the number of words between posts almost quadrupled from the first section of each chapter to the last section. As suggested by some of their reflections, students seemed much more engaged at the beginning of their reading than at the end. Their engagement may not have extended beyond the five required annotations.

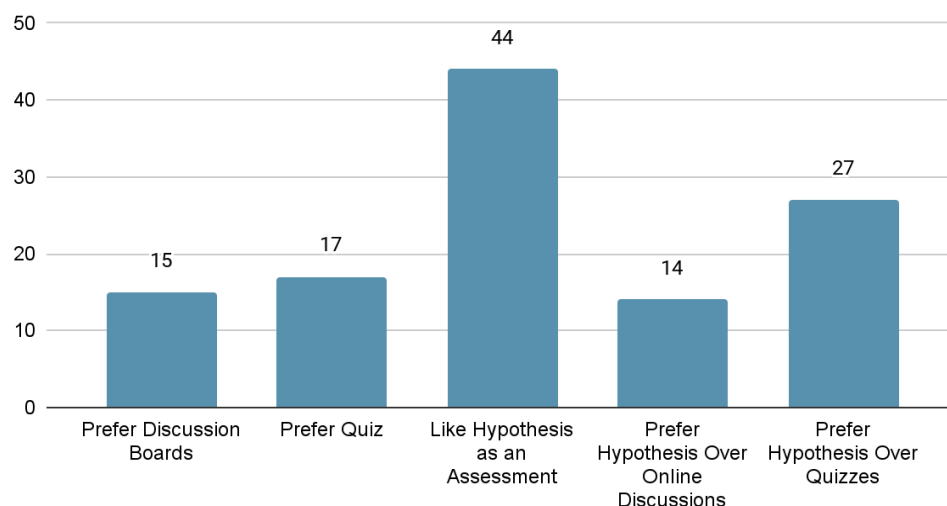
Preferences

As part of their final exam, the undergraduate students answered questions about their assessment preferences, expressing opinions about the use of the annotation tool, mastery quizzes, or online discussion boards. The results are shown in Figure 4.

Figure 4

Student Reading Assessment Preferences (n=47)

Student Assessment Preferences



Of the 47 students who answered the question on the final exam, 44 expressed that they liked the Hypothesis tool and felt it was an effective, more personalized way of assessing reading, with 14 specifically stating that they liked the annotations better than an online discussion and 27 preferring it over mastery quizzes. However, not everyone agreed. While these students may have liked Hypothesis, about one-third preferred an online discussion (15) or a quiz (17) over Hypothesis.

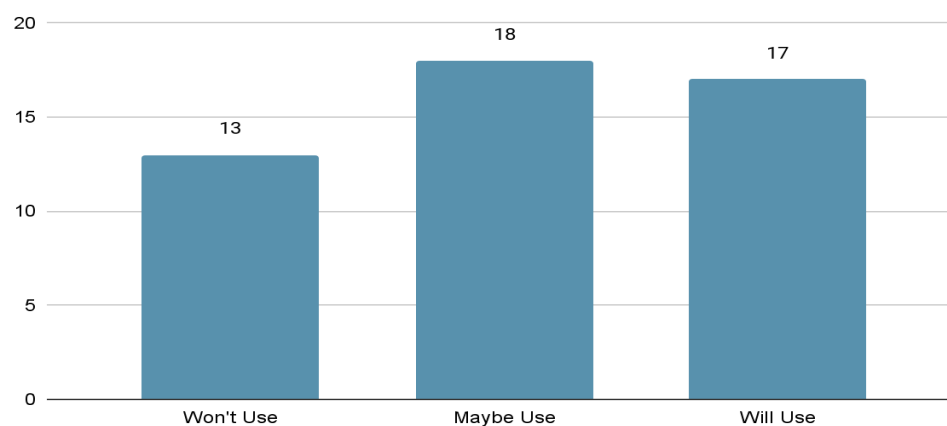
Future Use

In the final exam, these preservice teachers also commented on whether they would use a similar tool in their classrooms. They were generally positive about the use of the tool for their own learning and open to using the tool with their own students. Figure 5 shows the students' thoughts about the use of the tool.

Figure 5

Probability of Preservice Teachers Using an Annotation Tool in Their Classrooms

Probability of Future Use in Classroom



Students' comments seemed to be divided along pedagogical, age, and subject matter lines. Preservice teachers who did not feel comfortable using textbooks and felt texts were not a good teaching tool would not consider using any kind of annotation

feature. Other teachers, who could see some value in different kinds of texts, were more open. The age of their future students was another dividing line. Those who would be teaching younger children (generally third grade or younger) did not feel that such a tool would be useful for their students, although one said that she wanted to teach her young students how to take notes. Another dividing line was subject matter. English and social science teachers were favorable to an annotation tool and felt it could really enhance their students' learning. Science and physical education teachers felt the opposite.

Students who favored the tool saw value in tracking reading and in using reading as a means to think more deeply and to interact with classmates. But even those who favored the tool suggested modifications, especially in the size of the groups working on a single document. They overwhelmingly advocated smaller groups. Other modifications included offering an annotation instrument as a tool which could be used to augment learning but not required as an assignment, using video or audio responses instead of written ones (especially for younger children), using the tool sparingly so students did not get tired of it or did not just read to annotate and not to learn, and including it as one of many choices for learning from a text.

Discussion

We divided the findings of our research questions into three primary themes that can be helpful for practitioners when considering the value this or a similar tool may have in an online classroom.

How Well Does This Foster Deep Reading of Texts?

The results of our study point to digital annotation as an effective means to encourage students to engage more actively with the text through deep reading, especially when prompted to do so by the assignment requirements. As noted in the findings, only 12.7% of the graduate student comments and 0.6% of the undergraduate student comments were categorized as "Generic," meaning they did not add additional meaning or insight to the text or promote further conversation. The remaining 87.3% of undergraduate student comments and 99.4% of graduate student comments were quality comments which went beyond a surface-level understanding of the text as students applied what they were reading through reflection, asking questions, analyzing the text, and referencing additional related outside information.

This finding is in line with research showing annotation and highlighting of text as a beneficial method for actively engaging with text, having better retention of text, and engaging in the learning process at higher levels of reading ([Fowler & Barker, 1974](#); [Leutner et al., 2007](#)). Additionally, many of the student reflection responses in the undergraduate course were positive as students saw the value of using this tool to better understand what they were reading.

The results of the descriptive statistics, however, reveal this tool was only beneficial to promote deep reading inasmuch as students completed assignment requirements. In both the graduate and undergraduate courses, the number of annotations at the beginning of each text was significantly greater than the number of annotations near the end of any given text. Many students completed their annotation requirements as quickly as possible and did not annotate after reaching the requirement. It seems there needs to be some motivation or assignments which directs students to continue annotating throughout the entire reading.

Will This Foster Good Collaboration in my Classroom?

Collaborative and social learning have their roots in constructivism. Anderson and Dron ([2011](#)) portrayed constructivist models as having several similar themes, including viewing learning as an active process which uses language to record what is being learned. It includes the ability for learners to evaluate their learning, see paradigms from a variety of perspectives, and be willing to subject their learning to social discussion. Collaborative annotation tools such as Hypothesis are uniquely designed to foster these processes, especially in distance education.

However, while the annotation tool did seem to help students engage more thoroughly with the text, it did not have as pronounced an effect in helping students engage with each other, especially in the undergraduate course. As noted in the findings, only 20% of annotations in the undergraduate course received even one response from a peer. Many students commented in their reflections on the assignment that the high number of annotations in a given selection was distracting, and they hid all but their own annotations. Only one student mentioned enjoying reading her classmates' comments. This contradicts Clinton-Lisell's (2023) study, which found that students enjoyed peer interactions and valued how peer-annotated content enhanced their comprehension of the course material.

However, the students' reaction may not have been a result of the tool itself. By their own admission, the number of people was the main reason for their not interacting. There were too many comments to track. The lack of interest could also have resulted from a lack of experience with interacting academically about a text. In addition, students may have seen text as being more didactic than dialectic. Perhaps the structure of the assignment (having to identify their annotation with a level of Bloom's taxonomy) shifted their focus away from interacting with class members. They may also have just wanted to get finished, or their interest fell after annotating the required five passages. Since the Canvas LMS integration with Hypothesis cannot support separate group annotations, instructors should brainstorm ways to work around this barrier, such as assigning different readings to different groups as a way to disperse a potentially large amount of annotations. This aligns with Lee et al. (2019), which proposed that grouping students can foster peer interactions and increase engagement levels despite large class sizes.

These suppositions gain legitimacy when comparing them to the graduate class. Of the graduate students' annotations, more than one-third (36.95%) were responses to other student's annotations. Although still not indicative of a robust conversation, these graduate students more than doubled the undergraduate interactive responses.

What Does This Teach Us About Patterns in How Students Read Texts?

According to the number and range of annotations that undergraduate students and graduate students made, we can see that both groups preferred to use social annotation to reflect their thoughts and understanding of the reading texts. Graduate students tended to make more annotations and the types of annotations were more diverse compared with undergraduate students.

From findings, we discovered that Graduate students demonstrated a broader diversity in their annotation types compared to undergraduates. The possible reasons for this diversity could be (1) the maturity of the students, (2) the class size and credit hours, and (3) the assignment prompts and specifications. The design of the undergraduate course focused on identifying their annotations on the levels of Bloom's Taxonomy, which dragged students' attention to finishing their assignments as required instead of commenting on each other's annotations. This is partially aligned with a previous study that concluded that collaborative annotation tools can not support collaboration by themselves ([Chan & Pow, 2020](#)). Under this situation, instructors have more control over designing the annotation assignments rather than the credit hours, class sizes and students' personalities. To support either learning outcome, instructors could add the related requirements or prompts when designing social annotation assignments.

Moreover, the results indicated that Hypothesis is an effective tool to strengthen students' deep reading. Most student responses to others' annotations were quality comments that went beyond the surface understanding of contents, and many undergraduate students perceived that using Hypothesis helped them read more deeply. Many students also completed more than the minimum required annotations for their assignments, supporting previous studies ([Gao, 2013](#); [Peng, 2017](#)). Without using Hypothesis or a similar annotation tool, instructors have no way of ensuring students really read the required texts, or how much and how deeply they read. When using a tool such as Hypothesis, instructors can more easily discern which students are involved in deep reading and which students might be misunderstanding or skipping over reading. Combined with the first pattern, we infer that deep reading and peer collaboration are two learning outcomes that Hypothesis could support.

Instructors can choose to use Hypothesis to focus more on deep reading and/or peer collaboration when designing assignments.

Limitations and Future Direction

This study includes some limitations. Firstly, because students from only two courses participated in this study, the results may not be representative of a larger population. Similarly, the design of the Hypothesis assignments as well as the different number of students using Hypothesis could also change the results of this study and make it unrepresentative. Another limitation is that we surveyed only the undergraduate students and lacked graduate students' perceptions. The research could have given more discrete and interesting results if students had more options than no, mixed, and yes as answers to the question "Did this tool help you better engage with the text?" A Likert type scale, for example, could give more nuanced answers. Follow up interviews could explore the students experiences and perceptions at a deeper level.

Future studies could modify the surveys so they collect both quantitative and categorical data, thus allowing for statistical modeling. Such research could involve more participants from more diverse courses and use participants' academic performance data to test the effectiveness of collaborative annotation tools in students' learning outcomes. It could also involve different types of assignments to promote student engagement throughout the text and to explore more closely how assignment requirements and specifications influence students' interaction with the text and with each other. Further, future research could collect both graduate and undergraduate students' perceptions from surveys and interviews.

Conclusion

This study investigated the effectiveness of Hypothesis, a digital annotation tool, in facilitating students' deep reading of texts and interaction with peers through a multiple-case study. The findings revealed that digital collaborative annotation can be an effective alternative to discussion boards or an additional supplement in an online course in aiding students in their ability to actively engage with the text and promote discussion. However, our study also discovered Hypothesis as a tool itself did not foster peer collaboration as expected. Instructors have control in deciding whether to encourage peer collaboration or foster deep reading in using Hypothesis. By designing the requirements and prompts of assignments, instructors could realize either learning outcomes. Finally, instructors should be aware of the potential risk that students might be overwhelmed by a large amount of annotations in designing annotation assignments. Along with appropriate assignment design, digital collaborative annotation tools have the potential to become a means for helping students engage deeper with texts and with their peers online.

References

-
- Ackerman, R., & Goldsmith, M. (2011). Metacognitive regulation of text learning: On screen versus on paper. *Journal of Experimental Psychology: Applied*, 17(1), 18-32. <https://doi.org/10.1037/a0022086>
- Ackerman, R., & Lauterman, T. (2012). Taking reading comprehension exams on screen or on paper? A metacognitive analysis of learning texts under time pressure. *Computers in Human Behavior*, 28(5), 1816–1828. <https://doi.org/10.1016/j.chb.2012.04.023>
- Allington, R. L. (1977). If they don't read much, how they ever gonna get good? *Journal of Reading*, 21(1), 57–61. <https://www.jstor.org/stable/40009862>

- Anderson, T., & Dron, J. (2011). Three generations of distance education pedagogy. *International Review of Research in Open and Distributed Learning*, 12(3), 80–97. <https://doi.org/10.19173/irrodl.v12i3.890>
- Bean, J. C. (2011). *Engaging ideas: The professor's guide to integrating writing, critical thinking, and active learning in the classroom*. Wiley.
- Ben-Yehudah, G., & Eshet-Alkalai, Y. (2014). The influence of text annotation tools on print and digital reading comprehension. *Proceedings of the 9th Chais Conference for Innovation in Learning Technologies*, 28–35.
- Birkerts, S. (2006). *The Gutenberg elegies: The fate of reading in an electronic age*. Farrar, Straus and Giroux.
- Chan, J. W., & Pow, J. W. (2020). The role of social annotation in facilitating collaborative inquiry-based learning. *Computers & Education*, 147, 103787. <https://doi.org/10.1016/j.compedu.2019.103787>
- Clinton-Lisell, V. (2023). Social annotation: what are students' perceptions and how does social annotation relate to grades? *Research in Learning Technology*, 31. <https://doi.org/10.25304/rlt.v31.3050>
- Cui, T., & Wang, J. (2024). Empowering active learning: A social annotation tool for improving student engagement. *British Journal of Educational Technology*, 55(2), 712-730. <https://doi.org/10.1111/bjet.13403>
- Dawson, S., Tan, J. P. L., & McWilliam, E. (2011). Measuring creative potential: Using social network analysis to monitor a learners' creative capacity. *Australasian Journal of Educational Technology*, 27(6). <https://doi.org/10.14742/ajet.921>
- de Oliveira Neto, F. G., & Dobsław, F. (2024). Building collaborative learning: Exploring social annotation in introductory programming. In *Proceedings of the 46th International Conference on Software Engineering: Software Engineering Education and Training* (pp. 12-21). <https://doi.org/10.1145/3639474.3640063>
- Fahy, P. J., Crawford, G., & Ally, M. (2001). Patterns of interaction in a computer conference transcript. *International Review of Research in Open and Distributed Learning*, 2(1), 1–24. <https://doi.org/10.19173/irrodl.v2i1.36>
- Fowler, R. L., & Barker, A. S. (1974). Effectiveness of highlighting for retention of text material. *Journal of Applied Psychology*, 59(3), 358-364. <https://doi.org/10.1037/h0036750>
- Gao, F. (2013). A case study of using a social annotation tool to support collaboratively learning. *The Internet and Higher Education*, 17, 76–83. <https://doi.org/10.1016/j.iheduc.2012.11.002>
- Garrison, D. R. (2016). *E-learning in the 21st century: A community of inquiry framework for research and practice*. Routledge.
- Garrison, D. R., Cleveland-Innes, M., & Fung, T. S. (2010). Exploring causal relationships among teaching, cognitive and social presence: Student perceptions of the community of inquiry framework. *The Internet and Higher Education*, 13(1–2), 31–36. <https://doi.org/10.1016/j.iheduc.2009.10.002>
- Hypothesis. (2020, October 3). Hypothesis for education. <https://web.hypothes.is/education/>
- Hanč, J., Hančová, M., & Borovský, D. (2023). Social reader perusal—a highly effective tool and source of formative assessment data. arXiv preprint arXiv:2308.07188. <https://doi.org/10.48550/arXiv.2308.07188>
- Klemm, W. R., & Snell, J. R. (1996). Enriching computer-mediated group learning by coupling constructivism with collaborative learning. *Journal of Instructional Science and Technology*, 1(2), 1–11.

- Lee, S. C., Lee, Z. W., & Yeong, F. M. (2019). Using social annotations to support collaborative learning in a Life Sciences module. *ASCILITE Publications*, 487-492. <https://doi.org/10.14742/apubs.2019.317>
- Leutner, D., Leopold, C., & den Elzen-Rump, V. (2007). Self-regulated learning with a text-highlighting strategy. *Zeitschrift Für Psychologie/Journal of Psychology*, 215(3), 174–182. <https://doi.org/10.1027/0044-3409.215.3.174>
- Liu, K. (2006). Annotation as an index to critical writing. *Urban Education*, 41(2), 192–207.
- Morales, E., Kalir, J. H., Fleerackers, A., & Alperin, J. P. (2022). Using social annotation to construct knowledge with others: A case study across undergraduate courses. *F1000Research*, 11. <https://doi.org/10.12688/f1000research.109525.2>
- Nist, S. L., & Simpson, M. L. (1988). The effectiveness and efficiency of training college students to annotate and underline text. National Reading Conference Yearbook.
- Peng, R. (2017). Promoting active learning of graduate student by deep reading in biochemistry and microbiology pharmacy curriculum. *Biochemistry and Molecular Biology Education*, 45(4), 305–312.
- Polly, D., Allman, B., Casto, A. R., & Norwood, J. (2018). Sociocultural perspectives of learning. In *Foundations of Learning and Instructional Design Technology*. https://edtechbooks.org/lidtfoundations/sociocultural_perspectives_of_learning
- Young, J. (2014). A study of print and computer-based reading to measure and compare rates of comprehension and retention. New Library World.
- Zhu, E. (1996). Meaning negotiation, knowledge construction, and mentoring in a distance learning course. Available at <https://files.eric.ed.gov/fulltext/ED397849.pdf>
- Zyrowski, M. L. (2015). Digital annotation tools and comprehension [PhD Thesis]. Oakland University.



This work is released under a CC BY license, which means that you are free to do with it as you please as long as you properly attribute it.



