

Integrating Learning Analytics to Measure Message Quality in Large Online Conversations

Evren Eryilmaz¹, Brian Thoms², Zafor Ahmed¹, Avneet Sandhu¹

¹California State University Sacramento, ²California State University Channel Islands,
 evren.eryilmaz@csus.edu, brian.thoms@csuci.edu, zafor.ahmed@csus.edu, avneetsandhu3@csus.edu

Abstract

Research on computer-supported collaborative learning often employs content analysis as an approach to investigate message quality in asynchronous online discussions using systematic message-coding schemas. Although this approach helps researchers count the frequencies by which students engage in different socio-cognitive actions, it does not explain how students articulate their ideas in categorized messages. This study investigates the effects of a recommender system on the quality of students' messages from voluminous discussions. We employ learning analytics to produce a quasi-quality index score for each message. Moreover, we examine the relationship between this score and the phases of a popular message-coding schema. Empirical findings show that a custom CSCL environment extended by a recommender system supports students to explore different viewpoints and modify interpretations with higher quasi-quality index scores than students assigned to the control software. Theoretical and practical implications are also discussed.

1. Introduction

Big data and business analytics have generated tremendous excitement in both academic and business communities over the past two decades. Today, these terms penetrate all areas of life including e-commerce, e-government, healthcare, finance, and education. Big data is typically characterized by a focus on very large, unstructured, and fast-moving data that comes from a variety of sources including the internet of things [1]. Businesses employ analytics to leverage opportunities presented by large datasets in many critical and high impact application areas. Examples of big data applications include driving recommendation engines, identifying patterns of customer behaviors, and developing advertisement campaigns [2].

With the increasing quantity and availability of analyzable educational data, the application of business analytics principles to learning processes in

technology-enhanced learning have led to the emergence of two specialized subfields: learning analytics and educational data-mining. There are numerous definitions of learning analytics [3]. This paper employs the definition emerged at the First International Conference on Learning Analytics and Knowledge in 2011 and adopted by the Society for Learning Analytics Research (SoLAR): “the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs” [4]. Educational data-mining is defined by Romero and Ventura as “developing, researching, and applying computerized methods to detect patterns in large collections of educational data that would otherwise be hard or impossible to analyze due to the enormous volume of data within which they exist” [5, p.12].

Both subfields have explored common research objectives including performance prediction, attrition risk detection, recommendation systems, skill estimation, and behavior detection with a strong emphasis on theory in the learning sciences [6]. Thus, although overlaps between the two subfields exist, learning analytics emphasizes a more holistic understanding of the relationship between students and learning environments, while educational data-mining concerns itself more with automated processing of large learning-related data in order to model specific constructs and the relationships between them [7].

Online collaboration continues to play an integral role in learning systems. As noted by Bause et al. [8], collaboration can create synergy effects that go beyond what any individual student could achieve alone. Computer-supported collaborative learning (CSCL) environments afford students the opportunity to build novel ideas for the collective benefit of a community. Asynchronous online discussions (AODs) are popular CSCL tools commonly used in blended and fully online courses, such as massive open online courses. The advantages of AODs have been associated with their time flexibility, which enables students, including less-assertive ones, to prepare,

reflect, and search for additional information before contributing to a discussion [9].

Despite their affordances, the actual benefits of AODs in large groups (i.e., 30 students or more as defined in Hiltz [10]) are not clear. One reason considers disorientation. We define disorientation as students' difficulties in discovering interesting and relevant information due to overwhelming quantity and disorganization of existing messages. This disorientation can decrease the quality of students' messages. For example, in Eryilmaz et al. [11] it was found that disorientation led students to produce many navigational uncertainty markers related to what they have read, have not read, and where to find the relevant information when they encounter voluminous discussions. These uncertainties can take them out of their comfort zone with negative effects on learning. For example, students can quickly agree on existing ideas (termed cumulative talk by Mercer [12]) instead of creating new ones that no one had prior to collaboration, or the reverse, where students get stuck in disagreements and cannot resolve gaps and inconsistencies (termed dispositional talk by Mercer [12]). From an instructional design standpoint, instructors can take different pedagogical actions to remedy these issues by interpreting and evaluating the quality of students' ideas. However, it may be impossible for instructors to analyze the quality of students' messages in a timely manner when they are buried under hundreds of messages [7].

Consequently, this research employs learning analytics to produce a message quasi-quality index score automatically that would otherwise be time-consuming to obtain. Moreover, we examine the relationship between this score and the phases of the interaction analysis model developed in Gunawardena et al. [13] to understand how students capitalize on each other's ideas to gradually refine ambiguous, figurative, and partial understandings. The following high-level research question guides our investigation: What are the effects of a recommender system on the quality of students' messages in large AODs? To answer this question, we conduct an experiment that compares two versions of a custom CSCL environment: one with a recommender system and one without it. Empirical findings show that a custom CSCL environment extended by a recommender system supports students explore different viewpoints and modify interpretations with higher quasi-quality index scores than students assigned to the control software.

2. Theoretical background

There are numerous theories on how and why collaboration can increase group performance and individual learning outcomes. This study is influenced by theories of group cognition [14] and knowledge building discourse [9]. Group cognition underscores that the production and continual improvement of ideas from an AOD cannot be meaningfully or completely traced back to a single individual because ideas arise through interactions among students [14]. Group cognition can occur during knowledge building discourse, which views ideas as continually improved public knowledge objects (e.g., new ways of thinking or solutions to problems) valuable to a community in educational and organizational settings [9, 15]. Thus, we can view the deliberate effort of coming up with ideas and reshaping them in the light of different viewpoints as indications of learning taking place among students from AODs.

However, students have a finite set of resources (i.e. time, effort, attention) for AODs [16]. The premise of the disorientation problem is that the effort required for searching interesting and relevant ideas with respect to needs and preferences from a potentially overwhelming number of messages is greater than the effort students are willing to invest [11]. Disorientation represents a devastating issue in AODs because students may not be aware of different viewpoints and divergent expertise, which can lead to erroneous knowledge building or misunderstandings [17]. This notion of awareness on the individual level is a constituting element for common ground in CSCL [18]. Disorientation can lead to a situation where students fail to reconcile inconsistencies, which can prevent participants from filling understanding gaps. For example, using heat map analysis, prior research shows that students navigate randomly and interact superficially with ideas in order to complete instructional tasks when they suffer from disorientation [11, 19, 20].

As described by Chen et al. [21], there are four potential contributors to the disorientation problem in large AOD group settings. The first potential contributor is limited student readiness, which refers to a lack of technical skills for participating in AODs and inadequate prior subject knowledge, both of which may cause students to be more susceptible to disorientation than others. A second potential contributor concerns the quantity of information, which considers the excessive amount of messages over a diversity of topics, which students must sort through during their collaboration process. This obstacle can exacerbate feelings of insecurity and lead students to drop a course or participate less or late in

AODs. A third potential contributor is quality of information, which underscores the inherent difficulty of instructional materials. When students perceive instructional materials to be difficult, they may develop the habit of scanning for points in AODs where they can most easily contribute with simple responses, rather than diagnosing and revising misunderstandings on difficult topics. A final potential contributor is poor user interfaces, which can fragment students' ideas across the non-linear threading structure of many AOD systems.

Educational recommender systems can recommend a broad range of items such as books, lecture notes, test items, assignments, or semester schedule plans. To the best of our knowledge, only few CSCL specific AODs were extended by recommender systems and evaluated through user studies beyond accuracy, recall, and precision measures [22]. For example, Eryilmaz et al. [11] found that a recommender system decreased the number of navigational uncertainty markers (e.g., "I don't remember well but I have seen some arguments elsewhere about digital divide you brought up in your message.") related to what students have read, have not read, and where to find the relevant information in their messages. Furthermore, Reynolds and Wang [23] demonstrated that a recommender system encouraged students to ask questions and clarify interpretations. However, these findings do not explain how students articulate their ideas in categorized messages.

Message quality is not always explicitly defined and its measurement varies in literature [24]. Consistent with De Wever et al. [25], we view message quality in AODs as the merits of circulated ideas for fostering creativity. For example, messages can be insightful or elaborate on the one hand and shallow or trivial on the other. Most CSCL studies employ the content analysis approach to investigate message quality in AODs via systematic message-coding schemas [for an overview, see 25]. Currently, systematic message-coding schemas are combined with learning analytics to develop a more profound understanding of message quality. For example, Dyke et al. [26] demonstrated that off-topic messages are more harmful to discussions that focus on learning basic facts than during discussions of problem-solving activities. Moreover, Wise et al. [27] showed a relationship between the time students take to read and re-read existing messages and the quality of new messages in AODs. The amount of information and writing styles have also been employed as cues for evaluating message quality. Regarding the amount of information, Matuk and Linn [28] found that students who generated more redundant ideas in AODs constructed more coherent explanations, while

students who generated more unique ideas constructed less coherent explanations. Regarding writing styles, Gunawardena et al [29] found that the sentiment of a message does not tie to a specific phase in the interaction analysis model [13].

3. CSCL environment

Our CSCL environment is the modular and flexible anchored AOD system developed in Eryilmaz et al. [30] and extended by the addition of a recommender system in Eryilmaz et al. [11]. When compared to AODs in existing learning management systems, this environment binds the instructional material and its related discussion in a single window to prevent students from toggling back and forth across multiple windows. At the heart of this environment are two open-source programs: (1) Poppler PDF rendering library; (2) Marginalia browser independent JavaScript program. Poppler PDF rendering library converts PDF-based instructional materials to a more flexible HTML format, which serves as the basis for the Marginalia that enables fine-grained annotations.

Marginalia has two features conducive to creating a tight-coupling between the instructional material and its related discussion. The first feature distinguishes which discussion thread corresponds to which annotated passage by lighting up both elements in red when either element is under the cursor. This representation allows students to recover the portion of a conversation that is concerned with a given part of a text. The second feature embeds a student's key idea (i.e., justification for making an annotation) in the direct context that elicited it by inserting a pop-up sticky-note that appears only when the cursor is on an annotated passage. This design decision prevents sticky-notes from interfering with students' task-oriented reading of a text. Taken together, both features present students with an intuitive means to collaboratively process academic literature.

Prior research demonstrates that the above-mentioned tight coupling facilitates a close spatial proximity between an instructional material and its associated discussion, which increases the communicative efficiency of AODs [30]. Moreover, this increase in communicative efficiency allows this system to produce a larger number of messages than AODs in existing learning management systems.

3.1. Recommender system

The purpose of our CSCL environment's recommender system is help students identify the most useful messages from a potentially overwhelming number of messages with respect to their preferences.

Among possible design approaches, the recommender system centers on collaborative filtering because this approach resonates with the notion of group cognition [14], which underscores that recommendations cannot be traced back to the behavior of any user. Put another way, the neighborhood of messages is responsible for a recommendation. The recommender system collects users' preferences in the form of a star ratings system. Although this mechanism alters users' regular navigation and reading patterns (i.e. they have to stop and rate items), it gives possibilities for diagnosing and resolving common and pertinent problems of understanding [31]. Thus, the recommender system aims to promote collaboration among like-minded peers in order to improve learning. The overall steps for recommending annotations to users in large discussions are as follows.

First, the recommender system employs the constrained Pearson correlation coefficient (PCC) similarity metric to compute similarity scores among users. Second, in order to boost up the task of finding the closest k -neighbors for a given user, the recommender system employs the K -nearest-neighbor (KNN) classification method. KNN classification uses distance measures to discard poor correlations for decreasing noise and improving the quality of recommendations. The current design employs the Euclidian distance as a common measure to enhance the generalizability of our findings. We set the value of nearest neighbors (k) at 3 because this value provided the highest classification achievement compared to other values. Third, the recommender system calculates a user's preference score for each item based on the best neighbors' preferences by using the weighted averaging equation [11]. Finally, the

recommender system generates a top- N recommendation list based on the highest prediction scores for a given user. We fixed the number of recommendations (N) to 4 because users in real world applications view only the very first recommended items. If a user gives a low rating to a recommendation, a new neighborhood may emerge, resulting in adjusted recommendations.

Figure 1 illustrates the user interface with a special color schema for navigation. There are two window panes in the user interface. The left window pane displays personalized annotation recommendations and threaded discussions. Through Marginalia, clicking on a recommendation in the left window pane navigates the text in the right window pane to the referenced position. Furthermore, as depicted in Figure 1, Marginalia highlights both the selected recommendation and relevant passage from the text in red to help students read annotation recommendations that match their preferences. In Figure 2, the student who received a recommendation moves the cursor over a recommended annotation on text. Employing this movement as input, Marginalia navigates the discussion to the pertinent thread and draws a red border around that thread. This design consideration highlights all messages related to the annotation affording students a sense of the complete discussion without extra navigational effort. Moreover, Marginalia uses the same input to display small pop-up boxes adjacent to each annotation. These boxes aim to prevent students from altering their regular navigation and reading patterns to rate annotations. Additionally, these boxes display a student's key idea for posting a comment, community members' average rating for that idea, and a star ratings system.

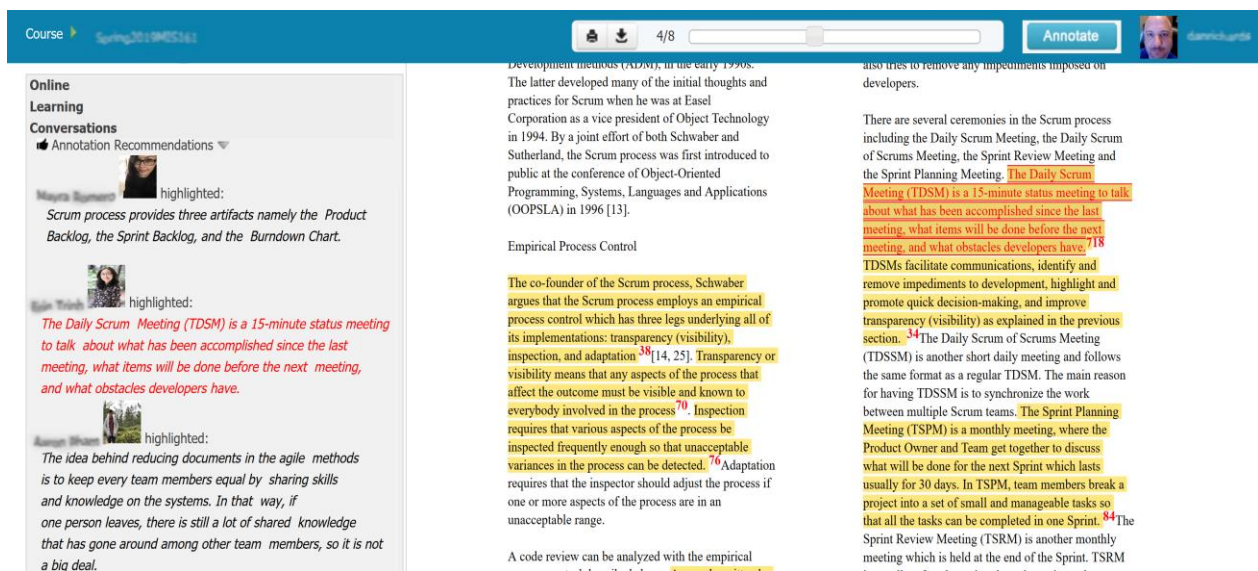


Figure 1. Screenshot of a top- N recommendation list

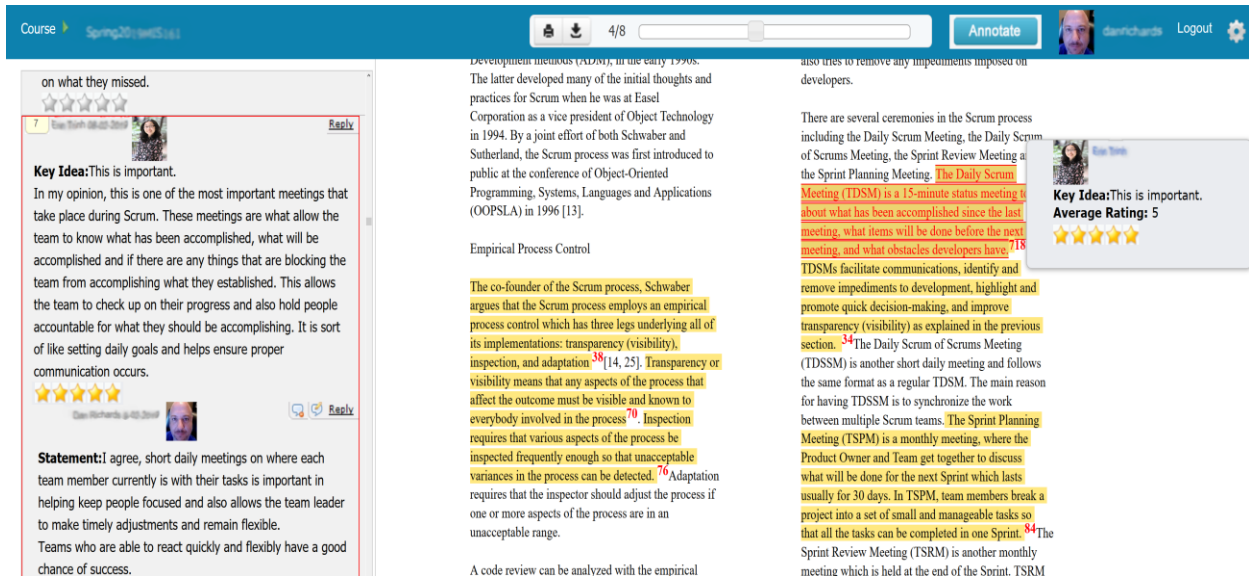


Figure 2. Screenshot of a discussion thread associated with a recommendation from previous figure

Prior research demonstrated that the design of this recommender system reduced students' disorientation and afforded them the opportunity to become better aware of interesting and relevant information based on their needs and preferences without heavy costs (i.e., time and effort) from large online conversations [11].

3.2. Control system

To isolate the effects of the recommender system, control software was implemented. The control system includes the same ratings system, but without the recommender system. Figure 3 shows the interface of the control system.

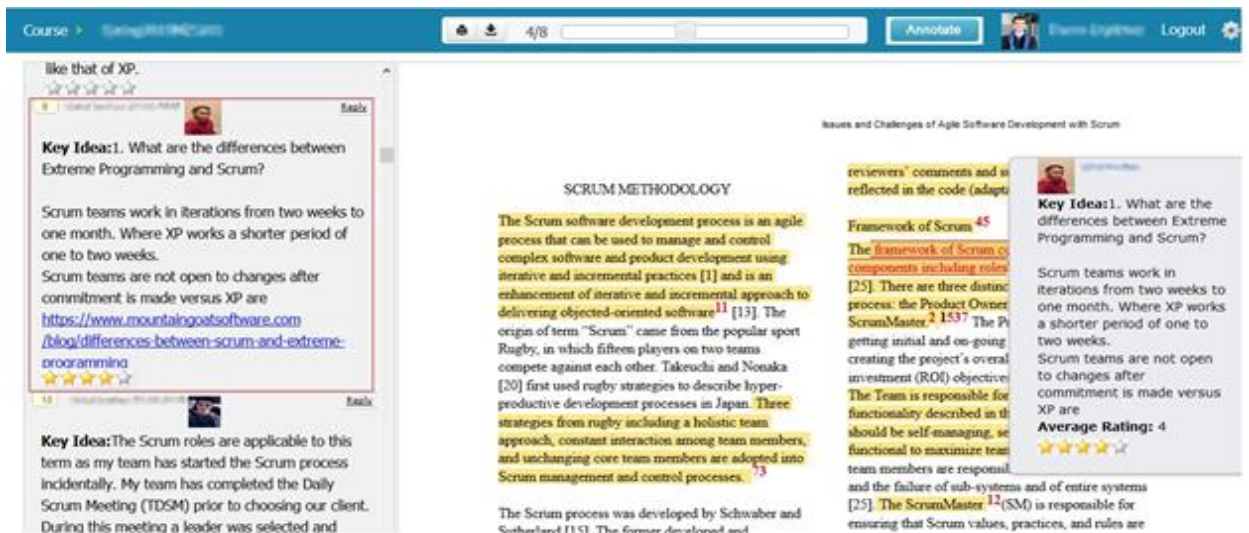


Figure 3. Screenshot of the control system

4. Research questions and methodology

The high-level research question of this paper asks: What are the effects of the recommender system on the quality of students' messages in large AODs? This

question can be divided into the following sub-questions:

1. What are the effects of the recommender system on the phases of the interaction analysis model developed by Gunawardena et al. [13]?

2. What are the effects of the recommender system on message quasi-quality index scores per student?
3. Is there a relationship between message quasi-quality index scores and the phases of the interaction analysis model developed by Gunawardena et al. [13]?

To answer these research questions, an experiment was conducted across two sections of a system analysis and design course required for information system majors. The learning objective of the course was to help students understand how to analyze, design, and develop business information systems to solve real world problems. Participants were 70 sophomore undergraduate students majoring in information systems. Of the 70 participants, 48% were females and 52% were males. The mean age of the participants was 20.3 ($SD = 1.01$). All participants were split into two sections of the same course. Each section had 35 students. Both sections were taught by the same instructor and followed the same schedules to eliminate confounding factors. We randomly assigned one section to the recommender system and the other to the control system. Prior to the experiment, we provided training in a face-to-face class session to ensure that all students would be able to work with the respective system. The instructional topic for the purpose of this experiment was the scrum methodology. This topic included the research paper, "Issues and Challenges of Agile Software Development with Scrum" [32]. The paper was covered during a two-week online discussion period.

The learning task for both groups included two discussion activities. The first discussion activity asked students to annotate important topics of interest from this paper by constructing their own explanations based on evidence and reasoning. The second discussion activity asked students to refine each other's ambiguous, figurative, and partial explanations in order to enhance their conceptual understanding of the instructional topic. Participation in online discussions was required and part of students' regular curriculum. All students were required at minimum to make two annotations and provide focused feedback as well as ratings to at least two fellow students' explanations. To keep conditions equal, students use of the recommendation system was voluntary.

4.1 Quality measures

All instruments were adopted from existing literature to increase validity. As a popular research methodology in CSCL, content analysis allows concise and generalizable categorization of AOD

messages based on systematic coding schemas [25]. One of the most popular coding schemas is the interaction analysis model developed in Gunawardena et al. [13]. This systematic coding schema identifies five phases in knowledge building discourse. The first phase, sharing information, denotes statements of initial interpretations on a topic. The second phase, exploring dissonance, represents identification of areas of disagreement among interpretations. The third phase, negotiating meaning, underscores modification of initial interpretations or clarification of different viewpoints. The fourth phase, testing proposed synthesis, involves evaluation of proposed syntheses against received facts, personal experience, or other sources. Finally, the fifth phase, agreeing on new knowledge, demonstrates summarization of agreement(s) on refined interpretations.

But, the interaction analysis model [13, 29] does not go beyond the categorization of AOD messages. Within technical disciplines, where technical expressions are continuously evolving, clarifying reasoning processes with subject matter keywords can lead students to higher levels of thinking [29]. For example, what subject matter keywords drive the knowledge building discourse? Moreover, Thoms et al. [33] showed that the lexical complexity of a message (i.e., how difficult a message is to read) affects the number of replies in AODs. Taken together, both subject matter keyword usage and lexical complexity are common features employed in automatic essay grading systems (for a review, see [33]).

Thus, the interaction analysis model [13] offers only a partial view of the knowledge building discourse. To fill this gap, we adopted the quasi-quality index (QQI) developed and validated in Thoms et al. [33]. Represented below, QQI considers the sum of average lexical complexity and weighted scores for subject matter keyword usage.

$$QQI = \left(\frac{1}{n} \sum_{i=1}^n Xi \right) + \left(\frac{d}{\frac{1}{n} \sum_{i=1}^n d_i} * \frac{u}{\frac{1}{n} \sum_{i=1}^n u_i} \right)$$

n = Total elements	$d = \frac{k}{W-S}$
x = Post readability score	W = Total post words
d = Post keyword density	S = Total post stopwords
u = Total post non-stopwords	k = Total post keywords

Lexical complexity in the formula above employs Automated Readability Index, Coleman-Liau Index, Flesch-Kincaid Readability Test, and the Gunning-Fog Index. Linear mapping normalized these readability metrics to a 0-to 8-point scale. We identified subject matter keywords through a combination of relevant texts using the keyword generator in Thoms et al. [33] to extract pertinent

keywords as well as expert-driven keyword identification and inclusion. All keywords generated were reviewed by experts in the field for their relevance to this experiment's learning task. For easier assessment, QQI scores in the results section are represented out of 100. Thoms et al. [33] evaluated the model above with 2,157 AOD messages from various computer science courses. Their results show that the model recognized incomplete sentences, poor punctuation, complex words, and penalized scores if students did not use subject matter keywords to externalize their ideas. Moreover, they demonstrated that higher quality messages received more replies than lower quality messages. Finally, they found that only few messages in their study achieved a perfect score (100).

5. Results

Three independent coders with no knowledge of the study's purpose were trained to use the interaction analysis model [15] with a random sample of 50 messages. The unit of content analysis was each complete message because students' messages were rather short and mainly consistent of only one stage in knowledge building discourse. After training, each coder independently coded all remaining messages in the data set. 149 messages were recorded ($M = 4.26$, $SD = 0.44$) in the recommender system and 132

messages ($M = 4.14$, $SD = 0.84$) in the control software. In total, 70 users posted 281 messages, excluding the randomly selected messages for training purposes. The inter-coder Krippendorff's alpha reliability was 0.76, which indicates a satisfactory agreement beyond chance. All disagreements between coders were resolved by discussion after the Krippendorff's alpha measurement.

5.1 Phases of the interaction analysis model results

To answer the first sub-question, the effects of the recommender system on the phases of the interaction analysis model [13], six message scores were created for each student based upon data from content analysis. Message scores were computed as the proportion of students' posts in each message type. For example, if a student posted a total of 4 messages, and 2 of those messages were coded as sharing information, the sharing information message score for the student was 2/4 or 0.50. Table 1 presents descriptive statistics and the results of independent samples t-tests. According to Table 1, students assigned to the recommender system posted more messages coded as exploring dissonance and negotiating meaning, but fewer messages coded as sharing information than students assigned to the control software.

Table 1. Phases of the interaction analysis model results

Phase	Control Software (n=35)		Recommender System (n=35)		Test Statistics	
	M	SD	M	SD	<i>p</i> value	Cohen's <i>d</i>
Sharing information	0.43	0.22	0.23	0.18	<0.001	-0.99
Exploring dissonance	0.24	0.13	0.32	0.14	0.02	0.60
Negotiating meaning	0.14	0.13	0.21	0.12	0.03	0.56
Testing proposed synthesis	0.07	0.11	0.10	0.13	0.33	0.25
Agreeing on new knowledge	0.06	0.12	0.08	0.13	0.54	0.16
Off-topic messages	0.07	0.11	0.05	0.09	0.41	-0.20

5.2 Message quasi-quality index score results

Table 2 presents descriptive statistics per student and the results of independent samples t-tests for

quasi-quality index scores. According to Table 2, students assigned to the recommender system posted messages with higher quasi-quality index scores than students assigned to the control software.

Table 2. Quasi-quality index score results

	Control Software (n=35)		Recommender System (n=35)		Test Statistics	
	M	SD	M	SD	<i>p</i> value	Cohen's <i>d</i>
Quasi-quality index scores	67.84	8.87	73.12	7.97	0.01	0.63

5.3 Relationship between quasi-quality index scores and phases of the interaction analysis model

Table 3 presents descriptive statistics per student and the results of independent samples t-tests for the relationship between quasi-quality index scores and the phases of the interaction analysis model [13].

Accordingly, students assigned to the recommender system posted messages coded as exploring dissonance and negotiating meaning with higher quasi-quality index scores than students assigned to the control software. Moreover, we found that messages coded as testing proposed synthesis had the highest average quasi-quality index score in both groups.

Table 3. Relationship between quasi-quality index scores and the phases of the interaction analysis model

Quasi-quality index score for phase	Control Software (n=35)		Recommender System (n=26)		Test Statistics	
	M	SD	M	SD	p value	Cohen's d
Sharing information	67.10	11.90	68.83	15.07	0.62	0.13
	Control Software (n=30)		Recommender System (n=33)		Test Statistics	
Exploring dissonance	M	SD	M	SD	p value	Cohen's d
	67.13	12.58	74.35	13.89	0.04	0.55
Negotiating meaning	Control Software (n=21)		Recommender System (n=28)		Test Statistics	
	M	SD	M	SD	p value	Cohen's d
Testing proposed synthesis	70.38	12.85	80.27	14.24	0.02	0.72
	Control Software (n=10)		Recommender System (n=13)		Test Statistics	
Agreeing on new knowledge	M	SD	M	SD	p value	Cohen's d
	80.30	4.19	81.73	7.42	0.59	0.24
	Control Software (n=11)		Recommender System (n=11)		Test Statistics	
	M	SD	M	SD	p value	Cohen's d
	63.54	13.19	68.18	16.79	0.48	0.31

6. Discussion

This research employed learning analytics to produce a message quasi-quality index score automatically that would otherwise be time-consuming to obtain. Moreover, we examined the relationship between this score and the phases of the interaction analysis model developed in Gunawardena et al. [13] to understand how students capitalized on each other's ideas to gradually refine ambiguous, figurative, and partial understandings. In this section, we will interpret our findings in light of the theoretical background.

Regarding our first sub-question, phases of the interaction analysis model [13], results show that students assigned to the recommender system posted more messages coded as exploring dissonance and negotiating meaning, but fewer messages coded as sharing information than students assigned to the control software. From the lens of knowledge building discourse [9], we can consider messages coded as exploring dissonance and negotiating meaning

indications of gap filling learning behaviors. Perhaps the key insight these findings offer, compared to prior research [22,23], is that the recommender system increased students' awareness of missing or low prior knowledge based on their needs and preferences, thereby prompting students to fill understanding gaps. Thus, it can be reasonably inferred that the recommender system genuinely facilitated a knowledge-advancing online discussion. Accordingly, the high number of messages coded as sharing information in the online discussion facilitated by the control software corroborates Scardamalia and Bereiter's [9] remark that if a community is unaware of ambiguous, figurative, and partial understandings, higher phases of knowledge construction cannot be realized. Extending prior research [11], this important finding suggests that disorientation problem (i.e., students' difficulties in discovering interesting and relevant information based on their needs and preferences from voluminous discussions) constrained students assigned to the control software from modifying their interpretations.

Regarding the second sub-question, message quasi-quality index scores per student, we found that students assigned to the recommender system posted messages with higher quasi-quality index scores than students assigned to the control software. Turning back to the quasi-quality index algorithm, this finding indicates that messages in the online discussion facilitated by the recommender system were highly readable and students assigned to this system integrated more subject matter keywords into their messages. This is an important finding because, as noted by Matuk and Linn [28], the increase message readability can improve students' online social and communication skills. Furthermore, when these highly readable messages contain more subject matter keywords, they may act as catalyst for facilitating greater conceptual understanding as students discuss, integrate, and explain instructional materials in collaborative scenarios. A possible explanation for this finding is that students were active in reading and re-reading recommended materials. This explanation is consistent with the heat map analysis results reported in prior research [11].

Lastly, regarding the third sub-question, relationship between message quasi-quality index scores and the phases of the interaction analysis model, we found that students assigned to the recommender system posted messages coded as exploring dissonance and negotiating meaning with higher quasi-quality index scores than students assigned to the control software. Extending prior research that examined the effects of recommender systems on the quality of AODs with content analysis methodology [23], these findings suggest that when students identify interesting and relevant information based on their needs and preferences from voluminous discussions, they can ask questions and modify interpretations with substantive evidence and reasoning. From the lens of knowledge building discourse [9], these are important findings because they increase the likelihood that information will be understood and retained.

All of this said, we recognize several limitations of this study. First, because content analysis of large online conversations is time-consuming, our experiment focused on one AOD per group. Second, because user contributions are vital in many online communities, all students were required to write and evaluate at least two messages during the experiment. Future research can extend the time span of the experiment while keeping student contributions voluntary to enrich the literature. Third, we manually generated keywords employed by the quasi-quality index algorithm for our experiment's instructional topic. Therefore, another fruitful direction for future

research is to automate the keyword suggestion process.

In conclusion, this study addresses an important gap in content analysis literature [e.g., 13, 25, 29] in terms of understanding how students articulate their ideas in categorized AOD messages. Furthermore, it extends the literature on user studies that evaluate AOD based recommender systems beyond accuracy, recall, and precision measures [22, 23]. Our findings are both timely and important for the information systems (IS) curriculum because online collaboration is becoming more prevalent as IS projects become increasingly dispersed, whether as a result of outsourcing or open-source development model.

7. References

- [1] Davenport, T. (2014). *Big data at work: dispelling the myths, uncovering the opportunities*. Harvard Business Review Press.
- [2] Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS quarterly*, 36(4).
- [3] Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE review*, 46(5), 30.
- [4] SoLAR.(2011). Open Learning Analytics: An Integrated & Modularized Platform. WhitePaper. Society for Learning Analytics Research. Retrieved from <https://solaresearch.org/wp-content/uploads/2011/12/OpenLearningAnalytics.pdf>
- [5] Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1), 12-27.
- [6] Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Journal of Educational Technology & Society*, 17(4), 49-64.
- [7] Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5/6), 304-317.
- [8] Bause, I. M., Brich, I. R., Wesslein, A. K., & Hesse, F. W. (2018). Using technological functions on a multi-touch table and their affordances to counteract biases and foster collaborative problem solving. *International Journal of Computer-Supported Collaborative Learning*, 1-27.
- [9] Scardamalia, M., & Bereiter, C. (2006). Knowledge building theory, pedagogy, and technology. *The Cambridge Handbook of the Learning Sciences*, R. K. Sawyer, ed., Cambridge University Press, 97-115.

- [10] Hiltz, S. R. (1997). Impacts of college-level courses via asynchronous learning networks: Some preliminary results. *Journal of Asynchronous Learning Networks*, 1(2), 1-19.
- [11] Eryilmaz, E., Thoms, B., Ahmed, Z., & Lee, K. H. (2019). Affordances of Recommender Systems for Disorientation in Large Online Conversations. *Journal of Computer Information Systems*, 1-11.
- [12] Mercer, N. (2000). *Words and minds: How we use language to think together*. Routledge.
- [13] Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of educational computing research*, 17(4), 397-431.
- [14] Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: The MIT Press.
- [15] Nonaka, I. (2008). *The knowledge-creating company*. Harvard Business Review Press.
- [16] Eryilmaz, E., van der Pol, J., Ryan, T., Clark, P. M., & Mary, J. (2013). Enhancing student knowledge acquisition from online learning conversations. *International Journal of Computer-Supported Collaborative Learning*, 8(1), 113-144.
- [17] Ruttun, R. D., & Macredie, R. D. (2012). The effects of individual differences and visual instructional aids on disorientation, learning performance and attitudes in a Hypermedia Learning System. *Computers in Human Behavior*, 28(6), 2182-2198.
- [18] Engelmann, T., Kozlov, M. D., Kolodziej, R., & Clariana, R. B. (2014). Fostering group norm development and orientation while creating awareness contents for improving net-based collaborative problem solving. *Computers in Human Behavior*, 37, 298-306.
- [19] Eryilmaz, E., Thoms, B., Mary, J., Kim, R., & van der Pol, J. (2015). Instructor versus peer attention guidance in online learning conversations. *AIS Transactions on Human-Computer Interaction*, 7(4), 234-268.
- [20] Eryilmaz, E., Chiu, M. M., Thoms, B., Mary, J., & Kim, R. (2014). Design and evaluation of instructor-based and peer-oriented attention guidance functionalities in an open source anchored discussion system. *Computers & Education*, 71, 303-321.
- [21] Chen, C. Y., Pedersen, S., & Murphy, K. L. (2012). The influence of perceived information overload on student participation and knowledge construction in computer-mediated communication. *Instructional Science*, 40(2), 325-349.
- [22] Erdt, M., Fernandez, A., & Rensing, C. (2015). Evaluating ecommender systems for technology enhanced learning: a quantitative survey. *IEEE Transactions on Learning Technologies*, 8(4), 326-344.
- [23] Reynolds, B. L., & Wang, S. L. (2014). An investigation of the role of article commendation and criticism in Taiwanese university students' heavy BBS usage. *Computers & Education*, 78, 210-226.
- [24] Peters, V. L., & Hewitt, J. (2010). An investigation of student practices in asynchronous computer conferencing courses. *Computers & Education*, 54(4), 951-961.
- [25] De Wever, B., Schellens, T., Valcke, M., & Van Keer, H. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers & education*, 46(1), 6-28.
- [26] Dyke, G., Howley, I., Adamson, D., Kumar, R., & Rosé, C. P. (2013). Towards academically productive talk supported by conversational agents. In *Productive multivocality in the analysis of group interactions* (pp. 459-476). Springer, Boston, MA.
- [27] Wise, A. F., Hausknecht, S. N., & Zhao, Y. (2014). Attending to others' posts in asynchronous discussions: Learners' online "listening" and its relationship to speaking. *International Journal of Computer-Supported Collaborative Learning*, 9(2), 185-209.
- [28] Matuk, C., & Linn, M. C. (2018). Why and how do middle school students exchange ideas during science inquiry?. *International Journal of Computer-Supported Collaborative Learning*, 13(3), 263-299.
- [29] Gunawardena, C. N., Flor, N. V., Gómez, D., & Sánchez, D. (2016). Analyzing social construction of knowledge online by employing interaction analysis, learning analytics, and social network analysis. *Quarterly Review of Distance Education*, 17(3), 35.
- [30] Eryilmaz, E., Thoms, B., & Canelon, J. (2018). How design science research helps improve learning efficiency in online conversations. *CAIS*, 42, 21.
- [31] Thoms, B., Garrett, N., & Ryan, T. (2010, January). The design and evaluation of a peer ratings system for online learning communities. In *2010 43rd Hawaii International Conference on System Sciences* (pp. 1-10). IEEE.
- [32] Cho, J. (2008). Issues and challenges of agile software development with SCRUM. *Issues in Information Systems*, 9(2), 188-195.
- [33] Thoms, B., Eryilmaz, E., Dubin, N., Hernandez, R., & Colon-Cerezo, S. (2018, December). Dynamic visualization of quality in online conversations. In *2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI)* (pp. 329-335). IEEE.