ASEE 2022 ANNUAL CONFERENCE
Excellence Through Diversity
MINNEAPOLIS, MINNESOTA, JUNE 26TH-29TH, 2022

Paper ID #37205

# Work-in-Progress: Using Latent Dirichlet Allocation to uncover themes in student comments from peer evaluations of teamwork


## Gaurav Nanda (Assistant Professor)

Dr. Gaurav Nanda is an Assistant Professor in the School of Engineering Technology at Purdue University with focus in Industrial Engineering Technology. His research interests include text mining, collaborative information systems, and intelligent decision support systems combining AI and human expertise. Dr. Nanda has worked as a postdoctoral researcher at Purdue University for two years and in the software industry for five years. He obtained his Ph.D. in Industrial Engineering from Purdue University and his Bachelors (B.Tech.) and Masters (M.Tech.) from Indian Institute of Technology Kharagpur, India.

## Siqing Wei

Siqing Wei received B.S. and M.S. in Electrical Engineering from Purdue University. He is currently pursuing a Ph.D. degree in Engineering Education program at Purdue University. After years of experience serving as a peer teacher and a graduate teaching assistant in first-year engineering courses, he has been a research assistant at CATME research group studying multicultural team dynamics and outcomes. The research interests span how cultural diversity impacts teamwork and how to help students improve intercultural competency and teamwork competency by interventions, counseling, pedagogy, and tool selection (such as how to use CATME Team-Maker to form inclusive and diversified teams) to promote DEI. In addition, he also works on many research-to-practice projects to enhance educational technology usage in engineering classrooms and educational research by various methods, such as natural language processing. In addition, he is also interested in the learning experiences of international students. Siqing also works as the technical development and support manager at the CATME research group.

## Andrew Katz (Assistant Professor)

## Christopher Greg Brinton

## Matthew W. Ohland (Dale and Suzi Gallagher Professor of Engineering Education)

Matthew W. Ohland is Associate Head and the Dale and Suzi Gallagher of Professor of Engineering Education at Purdue University. He has degrees from Swarthmore College, Rensselaer Polytechnic Institute, and the University of Florida. He studies the longitudinal study of engineering students and forming and managing student teams and with collaborators has been recognized for the best paper published in the Journal of Engineering Education in 2008, 2011, and 2019 and from the IEEE Transactions on Education in 2011 and 2015. Dr. Ohland is an ABET Program Evaluator for ASEE. He was the 2002–2006 President of Tau Beta Pi and is a Fellow of the ASEE, IEEE, and AAAS.

# Work-in-Progress: Using Latent Dirichlet Allocation to uncover themes in student comments from peer evaluations of teamwork

**Abstract:** This work-in-process research paper investigates common themes in peer-to-peer comments of teamwork behavior effectiveness collected with peer evaluations in engineering student teams in three time horizons – prior to COVID-19 pandemic, early phase of pandemic, and mature phase of pandemic. Constructive feedback is imperative to maintaining healthy team climate and dynamic, which facilitates positive individual and team learning outcomes. Asking engineering students to provide self- and peer-evaluation feedback in comments accomplishes multiple objectives. Students reflect on teammates' behavior and performance rather than relying on (potentially biased) general perceptions to provide evidence-based comments for the assessment period. Repeated practice giving feedback also tends to improve students' ability to provide constructive and insightful evaluations. To better understand what and how engineering students provide feedback in teamwork, the Comprehensive Assessment of Team-Member Effectiveness (CATME) peer evaluation tool suite was used to provide a framework to teach students about effective team behaviors using a behavioral-anchored rating scale. Using CATME also provided a mechanism for collecting self- and peer- evaluation survey data in both structured (the behavioral scale) and open-ended (comments) ways. Latent Dirichlet Allocation (LDA) was used as the classic method for topic modeling to analyze first-year engineering students' self- and peer- comments in the introductory engineering foundation courses in a large Midwestern R1 university. Topic Coherence measure (c_v) for topic quality was used to determine the optimal number of topics to represent the comment data. The themes of each of the topics identified were interpreted by thematic analysis of the most commonly used words and responses associated with each topic identified by the LDA model. The preliminary results showed that pre-pandemic themes closely matched the five behavioral dimensions of the CATME instrument. Data collected in Spring 2020 required more themes to capture the complexity of the transition to online learning. Comments from Spring 2021 required an even larger number of themes to describe the experience of teamwork during a fully virtual class implementation.

## Introduction

The use of teams to facilitate students' learning is widely adapted as one active learning pedagogy in engineering classrooms and labs and deeply integrated in engineering curriculum [1]. As one of the central competencies recognized by engineering education community, teamwork skills can be improved by practice and feedback, especially learning through peer evaluations [2] – [5]. In addition, the use of peer evaluations could also increase students' sense of autonomy, responsibility and motivation to contribute to team tasks [6] – [7].

Both instructors and students were provided detailed guidelines by many universities on how to survive and thrive through online teaching and learning, especially for team-based projects. Prior to the spread of COVID-19 pandemic, the dominant teaching modality was residential learning, where students could interact with each other in person. As Wut and Wu reported [8], under the virtual context, students' teamwork might have been impacted by a range of factors, including the collaboration tool availability (as the means to enhance communication), the familiarity among team members (as closer relationship are supposed to invoke timely and in-depth communication), and dissimilar levels of self-motivation for learning and collaboration (as the social presence could enhance the team cohesion). For engineering student teams that were assigned unstructured and complex problems, more challenges were likely to be faced, thus requiring richer resources to overcome them, as observed by Maznevski and Chudoba [9]. During the COVID-19 period, most energies of instructors and researchers were devoted to redesigning the course content and delivery, and the most common rapid-responded research on the impact of COVID19 is on how instructors and students experienced the event. The focus of this study was to analyze the impact of pandemic on student's team-based learning experiences in three time phases (prior to pandemic, early phase of pandemic, and mature phase of pandemic) as it underwent significant changes from in-person mode prior to pandemic, completely online during the early phase of the pandemic, and hybrid online and in-person during the mature phase of the pandemic. To this end, the research site and course highly incorporated teamwork into its core learning objectives, team-based projects and teamwork assessments via both manual grading of projects and peer evaluations of teamwork behaviors were collected in a similar manner during the three periods [10]. The CATME peer evaluation data, particularly the student peer-to-peer comments, enabled us to further investigate how the instructional modalities impacted the ways students comment on each other by the COVID-19 pandemic.

## Method

The student peer reviews collected from the CATME system were grouped into three data groups based on the timeline: 2017-19 (pre-pandemic), 2020 (beginning of pandemic), and 2021 (mature phase of pandemic). Students were enrolled in mandatory first-year engineering courses at the research site, where students were systematically assigned to teams for learning and projects. Students were asked to participate in the survey four times per course, roughly at the stages of the beginning of the course, the beginning, the middle and the end of the of the team final projects. Students were instructed to provide constructive feedback in writing to themselves and teammates via CATME interfaces on their teamwork behaviors.

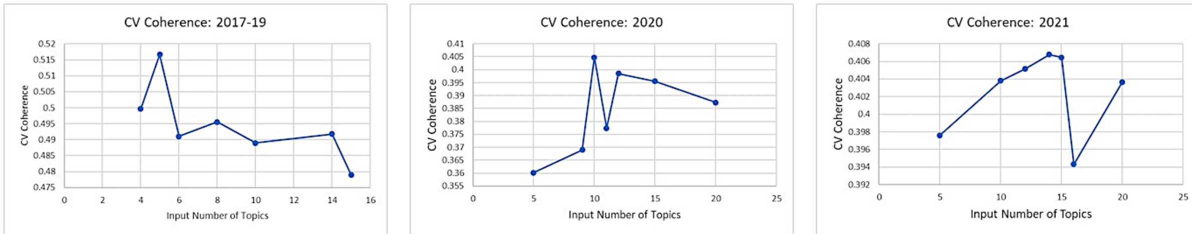
To analyze students' peer comments on teamwork behaviors, we utilized natural language processing and machine learning techniques, and qualitative analysis. Previous literature has demonstrated the feasibility to analyze qualitative data by those novel methods. For example, Wang et al. verified the existence structure of the student peer comments and CATME behavioral-anchored rating scale dimensions [10] - [11]; Wei et al. developed a pipeline tool to preprocess the peer comment data to be anonymized [12]. The specific method we are adopting in this study is called Latent Dirichlet Allocation (LDA) topic modelling.

These three groups of data were then analyzed using Latent Dirichlet Allocation (LDA) topic modeling approach to identify the prominent themes from each collection of peer evaluation comments. LDA topic model is an unsupervised machine learning method that assumes each document (review comment in our case) in the textual collection can be represented as a probabilistic distribution over underlying/latent topics and each topic is represented as a probability distribution over words present in the text collection [13] - [14]. LDA topic modeling has been found to be effective in educational research for efficiently analyzing large collection of text data such as survey responses, discussion forums, etc. [15] - [19].

As a novel technique applied to educational fields, the data analysis and interpretation paradigm has not been established. Because the analyzed data is qualitative, we still consider the use of LDA algorithms under the qualitative paradigm with positivism perspective. We assume the results shown as top topic words generated by LDA as solid clusters that distinguish with each other with specific connotations. Then, we inductively coded the topics into themes [20] by analyzing the top-20 topic words generated by the LDA model and qualitatively examining the most strongly associated comments for each topic in context of the CATME framework that was provided as a guiding material/metric to students for critically evaluating their peers' contributions and performance.

## Results

For each dataset (2017-19, Spring 2020, and Spring 2021), the optimum number of topics for the LDA topic model were determined using the CV Coherence measure. The CV Coherence measure was calculated for a range of input number of topics and the optimum number of topics was selected corresponding to the highest CV Coherence score. This optimization balances the complexity of the topic modeling and the unique information retained in the topics. The CV coherence values corresponding to the range of input number of topics for the three datasets are shown in Figure 1.




(a) (b) (c)

Figure 1: CV values by number of topics for (a) 2017-19, (b) Spring 2020, and (c) Spring 2021

As shown in Figure 1 (a-c), the maximum CV Coherence value was observed for 5 topics for the 2017-19 dataset (a), 10 topics for the 2020 dataset (b), and 14 topics for the 2021 dataset (c). The topics generated for each of the dataset along with their interpreted themes are presented in tables 1, 2, and 3 respectively, with the topic weight representing an estimated portion of the overall collection assigned to that topic, top-10 words representing the topic, and the contextually interpreted theme of the topic.

Table 1: Topics generated for dataset 2017-19

| **Topic Weight** | **Top-10 Topic Words** | **Interpreted Theme** |
|---|---|---|
| 0.53 | *work, team, good, time, group, quality, share, assignments, track, meetings* | CATME Dimension E, Expecting Quality. Also have connections to other dimensions of Keeping the Team on Track, Contributing to the Team, and Having Relevant Knowledge, Skills and Abilities. This topic thus has some representation of overall team-member effectiveness. |
| 0.39 | *good, ideas, coding, problem, team, matlab, problems, code, work, group,* | CATME dimension H, Having Relevant Knowledge, Skills and Abilities. Confirmed by reading several complete quotes in this grouping. |
| 0.31 | *team, work, great, good, group, member, teammate, working, project, coding* | CATME dimension I, Interacting with Teammates, as well as H, Having Relevant Knowledge, Skills and Abilities. Confirmed by reading several quotes in this grouping related to (helping behaviors, knowledge contribution, experience) |
| 0.19 | *work, class, time, meetings, team, group, meeting, project, doesnt, late,* | CATME Dimension K, Keeping the Team on Track – related to tracking of milestones and monitoring the work of others. |
| 0.13 | *code, algorithm, project, milestone, technical, contributed, helped, tasks, milestones, writing* | Context of the project. Comments in this grouping included detailed mentions of contributions related to statistics and algorithm design. |

In the dataset generated from 2017 to 2019, student peer comments are centered around the team effectiveness anchored on CATME behavioral-anchored rating scale dimensions (in the top four topic clusters of Table 1), including contribution to the team, interacting with teammates, expecting quality, keeping the team on track, and having relevant knowledge, skills, and abilities [10]. In addition, students also touched on the context of the team-based projects in programming, shown in the last topic cluster.

Table 2: Topics generated for dataset Spring 2020

| Topic Weight | Top-10 Topic Words | Theme |
|---|---|---|
| 0.49 | *work, team, quality, time, assignments, group, good, complete, assignment, high* | CATME Dimension E, Expecting Quality, with some hint of Contributing to the Team's Work. |
| 0.44 | *team, work, good, job, great, track, group, assignments, keeping, making* | Dimension K, Keeping the Team on Track. |
| 0.38 | *team, work, great, teammate, member, working, semester, hard, good, positive* | Dimension I, Interacting with Teammates, prominent, although some aspects don't appear. |
| 0.26 | *team, work, share, ideas, fair, good, meetings, member, group, teammates* | Contributing to the Team's Work, along with the feedback aspects of Interacting with Teammates. |
| 0.26 | *work, feel, project, online, working, semester, team, teammates, class, performance* | **Focused on negative aspects of connecting with a particular teammate after virtual instruction was initiated.** |
| 0.26 | *questions, group, good, problems, team, problem, great, understand, work, helpful* | Dimension H, Having Relevant Knowledge, Skills, and Abilities |
| 0.23 | *time, group, work, meetings, meeting, class, team, times, chat, due, zoom* | **Focused on positive aspects of a teammate's ability to engage after virtual instruction was initiated.** |
| 0.22 | *team, knowledge, skills, coding, matlab, assignments, improve, group, class, lot* | Higher levels of Dimension H, Having Relevant Knowledge, Skills, and Abilities, including learning new skills to benefit the team. |
| 0.11 | *code, project, coding, writing, technical, algorithm, helped, lot, job, contributed* | Dimension H, Having Relevant Knowledge, Skills, and Abilities, with a particular focus on the specific class context. |
| 0.09 | *assignment, code, function, main, team, attention, helped, make, made, part* | Expecting Quality – checking the team's work and correcting errors. |

In 2020 when the COVID-19 virus broke out, the LDA algorithm produced ten topic clusters shown in Table 2, covering not only CATME five dimensions, but also how the virtual instruction impacted student teamwork behaviors in both positive and negative ways (see cluster 5 and 7; the themes are highlighted). The negative reaction concerns about whether the team could still engage and perform as in residential setting. While the positive feeling relates to the relevant flexibility on team meeting scheduling.

Table 3: Topics generated for dataset Spring 2021

| Topic Weight | Top Topic Words | Theme |
|---|---|---|
| 0.44 | *good, job, team, work, great, group, track, making, keeping, improve, assignments* | Expecting Quality – improving assignments, helping others improve. |
| 0.39 | *team, work, assignments, assignment, time, make, complete, part, teammates, tasks,* | Dimension K, Keeping the Team on Track – planning, monitoring teammates and making sure their work is finished on time. |
| 0.31 | *team, great, work, teammate, group, member, makes, meetings, helps, hard, time* | Maintaining a positive attitude – so most closely related to Interacting with Teammates. |
| 0.30 | *ideas, team, good, work, great, group, meetings, teammate, member, discussion* | Contributing to the Team's Work – good ideas, sharing in discussion, giving input |

| | | |
|---|---|---|
| 0.29 | *feel, work, semester, group, team, teammates, performance, project, lot, improved* | Split between Contributing to the Team's Work (effort) and Interacting with Teammates (communication) |
| 0.29 | *coding, matlab, knowledge, skills, team, assignments, understand, understanding, good, class* | Dimension H, Having Relevant Knowledge, Skills, and Abilities, with a particular focus on the specific class context. |
| 0.29 | *work, share, team, fair, good, meetings, timely, time, manner, quality* | Split between Contributing to the Team's Work (completing a fair share on time) and Expecting Quality (inspiring teammates to improve the team's work) |
| 0.26 | *group, questions, assignments, class, good, answer, work, time, chat, asks, helpful* | Contributing to the Team's Work specifically in the context of virtual instruction during team time in Zoom breakout rooms. |
| 0.26 | *great, work, semester, working, teammate, team, project, future, forward, enjoyed* | **Strongly positive in a general way – not focused on a particular CATME dimension.** |
| 0.25 | *team, member, success, teams, skills, improve, valuable, members, important, assignments* | Interacting with Teammates, including multiple aspects of that dimension. |
| 0.24 | *time, meetings, class, meeting, work, times, team, group, late, good* | **Focused on negative aspects of participation and schedule issues with a particular teammate during virtual instruction.** |
| 0.19 | *problems, problem, ideas, solutions, solve, team, code, great, solving, good* | Dimension H, Having Relevant Knowledge, Skills, and Abilities, including learning new skills and teaching others to help the team. |
| 0.18 | *work, quality, high, team, time, teammate, great, group, good, makes* | Expecting Quality – more focused on high standards. |
| 0.17 | *project, code, coding, helped, algorithm, writing, lot, contributed, made, final* | Dimension H, Having Relevant Knowledge, Skills, and Abilities, with a particular focus on the specific class context. |

In the latest dataset (2021), student peer comments seem discussing a broader range of themes in teamwork shown in Table 3 (14 themes). Although some clusters converge into CATME dimensions, but they could also get mixed. Particularly, students provide some general comment to the teammates (cluster 9); and concern about the conflict in participation and scheduling with certain teammates (cluster 11).

**Discussion and Future Work**

Our preliminary results suggest that students could consistently provide comments anchored to CATME rubrics throughout the years from 2017 to 2021. With the impact of COVID-19 pandemic and the followed virtual instruction, students also comment on the associated effects, such as scheduling and participation. Given to the limited interaction modality bounded by virtual cooperation, students might not observe their peers' teamwork behaviors completely, which might explain the increased variety genres in 2020 and 2021. The method centering on LDA we propose here is novel; yet the paradigm of this particular research method is still in exploration. In future work we will apply an alternative approach to modeling students' teammate comments. That approach will use a transformer-based neural network architecture that embeds text from the comments in a high-dimensional embedding space. This is different from the approach described in the present paper in that it will not be a probabilistic model as is created with LDA.  Instead, these transformer models capture semantically similar statements by training on large corpora (on the order of 100s of GBs of text). We can then use these pre-trained models for embedding the text in order to perform subsequent mathematical operations such as clustering to identify more nuanced themes in student teammate comments that may not be captured by the broader topics identified by the LDA model and may also have been impacted by the difference in data size of peer comments for the three phases (2017-19, 2020, and 2021). Combined the results from different techniques, we aim to seek more complete understanding of the student peer comments to illustrate the development of student teamwork behaviors and skills under the impact of the COVID-19.

## References

[1] ABET. (2017). 2018-2019 Criteria for Accrediting Engineering Programs. Retrieved from https://www.abet.org/wp-content/uploads/2018/02/E001-18-19-EAC-Criteria-11-29-17.pdf

[2] A. C. Loignon, D. J. Woehr, J. S. Thomas, M. L. Loughry, M. W. Ohland, and D. M. Ferguson, "Facilitating peer evaluation in team contexts: The impact of frame-of-reference rater training," Acad. Manag. Learn. Educ., 2017.

[3] J. E. Bono and A. E. Colbert, "Understanding responses to multi-source feedback: The role of core self-evaluations," Pers. Psychol., 2005.

[4] G. Dai, K. P. De Meuse, and C. Peterson, "Impact of multi-source feedback on leadership competency development: A longitudinal field study," J. Manag. Issues, 2010.

[5] J. W. Smither, M. London, and R. R. Reilly, "Does performance improve following multisource feedback? a theoretical model, meta-analysis, and review of empirical findings," Pers. Psychol., 2005.

[6] S. A. Hernandez, "Team learning in a marketing principles course: Cooperative structures that facilitate active learning and higher-level thinking," J. Mark. Educ., 2002.

[7] B. J. Millis and P. G. Cottell, Cooperative Learning for Higher Education Faculty. Series on Higher Education. 1998.

[8] T. ming Wut and J. Xu, "Person-to-person interactions in online classroom settings under the impact of COVID-19: a social presence theory perspective," Asia Pacific Educ. Rev., vol. 22, no. 3, pp. 371–383, 2021, doi: 10.1007/s12564-021-09673-1.

[9] M. L. Maznevski and K. M. Chudoba, "Bridging space over time: Global virtual team dynamics and effectiveness," Organ. Sci., vol. 11, no. 5, pp. 473–492, 2000.

[10] M. W. Ohland, M. L. Loughry, D. J. Woehr, L. G. Bullard, C. J. Finelli, R. A. Layton, H. R. Pomeranz and D. G. Schmucker, "The comprehensive assessment of team member effectiveness: Development of a behaviorally anchored rating scale for self- and peer evaluation," Acad. Manag. Learn. Educ., vol. 11, no. 4, p. 609, 2012, doi: 10.5465/amle.2010.0177.

[11] R. Wang, S. Wei, M. W. Ohland, and D. M. Ferguson, "Natural language processing system for self-reflection and peer-evaluation," in the Fourth North American International Conference on Industrial Engineering and Operations Management, 2019, pp. 229–238.

[12] S. Wei, R. Wang, M. M. Ohland, and G. Nanda, "WIP: Automating anonymous processing of peer evaluation comments," in ASEE Annual Virtual Conference and Exposition, 2020.

[13] Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. the Journal of machine Learning research, 3, 993-1022.

[14] Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. Multimedia Tools and Applications, 78(11), 15169-15211.

[15] Nanda, G., Douglas, K. A., Waller, D. R., Merzdorf, H. E., & Goldwasser, D. (2021). Analyzing Large Collections of Open-Ended Feedback from MOOC Learners Using LDA Topic Modeling and Qualitative Analysis. IEEE Transactions on Learning Technologies.

[16] Daenekindt, S., & Huisman, J. (2020). Mapping the scattered field of research on higher education. A correlated topic model of 17,000 articles, 1991–2018. Higher Education, 80(3), 571-587.

[17] Zhang, Q., & Kozak, K. (2019, June). Quick Understanding Our Engineering Faculty Research Needs Using Topic Modeling. In 2019 ASEE Annual Conference & Exposition.

[18] Rolim, V., de Mello, R. F. L., Ferreira, M., Cavalcanti, A. P., & Lima, R. (2019, July). Identifying Students' Weaknesses and Strengths Based on Online Discussion using Topic Modeling. In 2019 ieee 19th international conference on advanced learning technologies (icalt) (Vol. 2161, pp. 63-65). IEEE.

[19] Ramesh, A., Goldwasser, D., Huang, B., Daumé III, H., & Getoor, L. (2014, June). Understanding MOOC discussion forums using seeded LDA. In Proceedings of the ninth workshop on innovative use of NLP for building educational applications (pp. 28-33).

[20] S. Merriam and E. Tisdell, “Qualitative data analysis,” in Qualitative research: A guide to design and implementation, San Francisco, CA: San Francisco, CA: Jossey-Bass, 2016, pp. 195–236.