Automated Feedback on Group Processes: An Experience Report

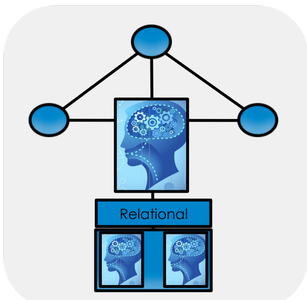
Marcela Borge
Pennsylvania State University
301C Keller Building
University Park, PA 16802
mborge@psu.edu

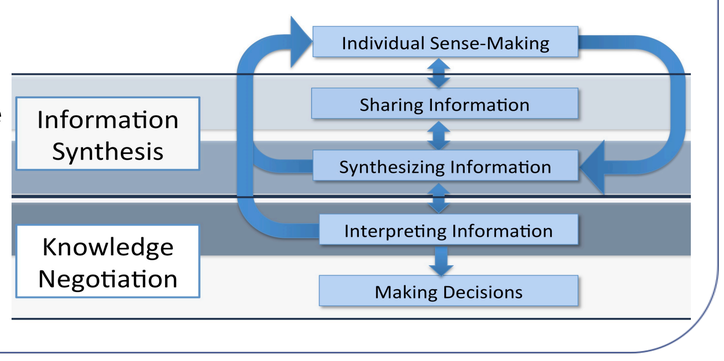
Carolyn P. Rosé
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
cprose@cs.cmu.edu

Introduction

Research suggests that developing socio-metacognitive expertise, the ability to understand, monitor, and regulate collective thinking processes that occur during collaboration, can help to mitigate group dysfunction and optimize collaborative interactions.

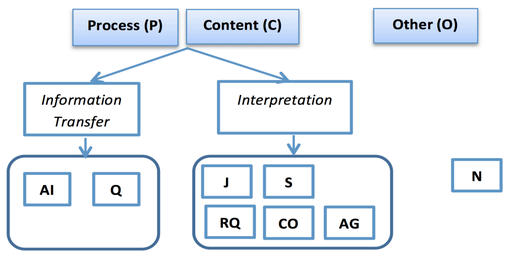
We have been working on design models to inform the design of Computer Supported Collaborative Learning (CSCL) systems to support socio-metacognitive development [1].

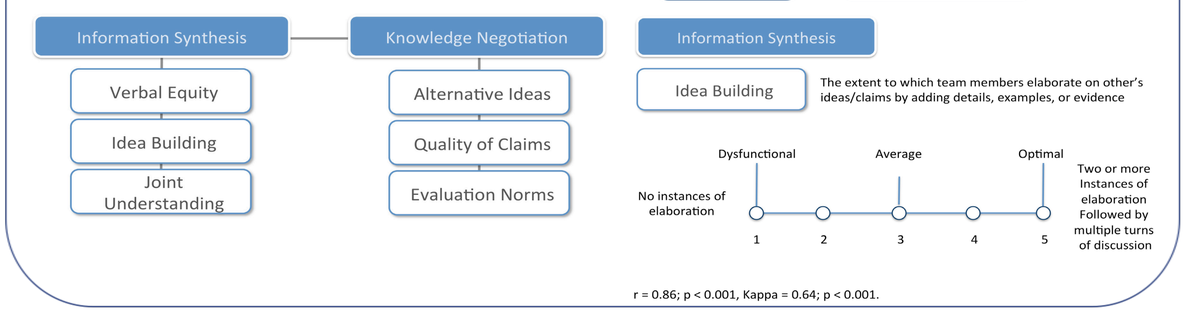




Assessment of Collaborative Discourse Quality

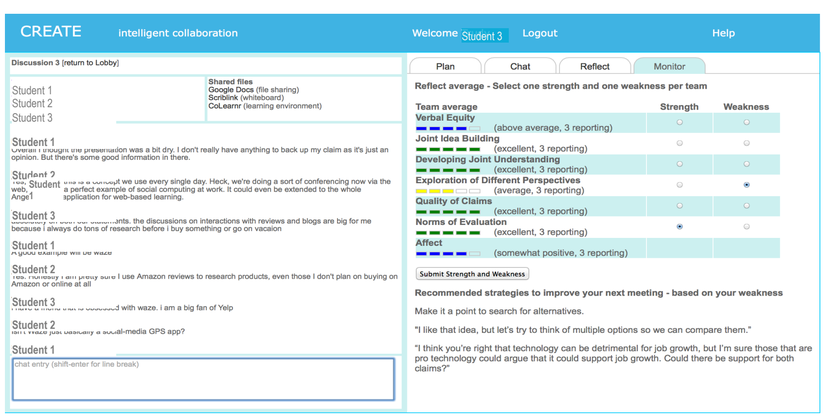
Information synthesis, consists of three categories of behavior: verbal participation, developing joint understanding, and joint idea building. Knowledge negotiation consists of contributing alternative ideas, quality of claims, and norms of evaluation.





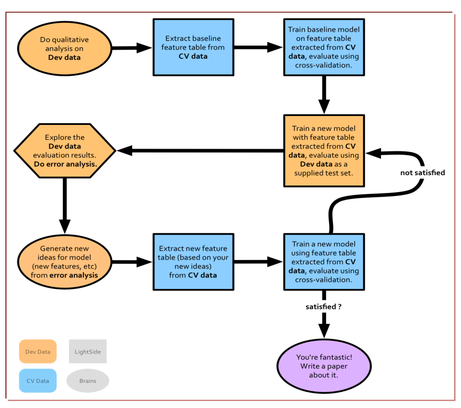
Study Context

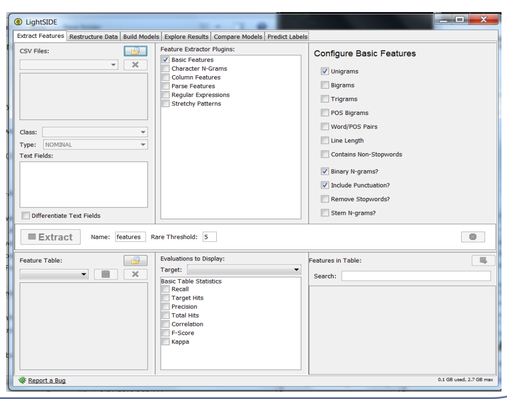
The study took place during a 16-week undergraduate, online course on information sciences and technology. 41 students participated in the study, each belonging to one of 14 groups. Students were required to read a chapter from the textbook or supplementary materials each week. In weeks five, seven, nine, eleven, and fourteen, students participated in a synchronous discussion related to the reading materials held in a collaborative workspace with chat capabilities called CREATE.



Automated Assessment

The automated process analysis models were trained using LightSIDE [2], using unigrams, bigrams, POS bigrams, and a line length feature, and a Logistic regression classifier with L2 regularization. In a cross-validation we achieved an accuracy of 86% and kappa of .77. In order to generate feedback for the six scales for the two core competencies we used the counts of predicted process codes per team to predict these six scales using a separate linear function trained using a simple linear regression for each scale.





Study Design

Across the five time-points, we compared the effect of four different feedback conditions on the quality of collaboration at the next time point. The study was run as a within-subject manipulation. The four conditions included: (1) no feedback, (2) expert feedback, (3) automated feedback, and (4) best practices. An assessment of group processes was conducted for each discussion. The first assessment was treated as a baseline. There was no significant effect of condition on collaboration quality or growth therein.

2 x 4 Mixed Factorial Design

Time 4 – Levels

Feedback Condition 4-levels	Time 1	Time 2	Time 3	Time 4	Time 5
None	Group 1	Group 4	Group 3	Group 2	
Expert	Group 2	Group 1	Group 4	Group 3	
Auto	Group 3	Group 2	Group 1	Group 4	
Best Practices	Group 4	Group 3	Group 2	Group 1	

Results & Conclusion

At each of four time points, each group was assigned a rating on a 5-point scale for each of the six dimensions. We evaluate the quality of the automated rating by computing a kappa with linear weighting between the sets of automated ratings and human ratings. At time point one, the automated ratings were assessed to be at random. By time point two, the weighted kappa was .19. It was .4 by time point 3. And finally, at time point four, it was up to .58. Altogether ratings for 10 sessions of the second course were needed to adapt the models and achieve a weighted kappa of .58.

References

[1] Borge, M., Ong Shiou, Y., & Rosé, C. 2015. Design models to Support the Development of High Quality Collaborative Reasoning in Online Settings. In *the Proceedings of the International Conference of Computer Supported Collaborative Learning (CSCL) 2015,* Volume 2, 427-434.

[2] Mayfield, E. & Rosé, C. P. (2013). LightSIDE: Open Source Machine Learning for Text Accessible to Non-Experts, *The Handbook of Automated Essay Grading*, Routledge Academic Press.