

## **Comparing Teamwork Peer Evaluations Between Culturally Homogenous Teams and Culturally Diverse Teams**

### **Mr. siqing wei, Purdue University-Main Campus, West Lafayette (College of Engineering)**

Siqing Wei received bachelor degree in Electrical and Computer Engineering at Purdue University. He is in the dual program to obtain master degree in Electrical and Computer Engineering and Ph.D degree in Engineering Education at Purdue University. After years of experience of serving a peer teacher and a graduate teaching assistant in first year engineering courses, he is now interested in study of the existence, cause and interventions on international engineers' teaming behaviors.

### **Dr. Matthew W. Ohland, Purdue University-Main Campus, West Lafayette (College of Engineering)**

Matthew W. Ohland is Professor of Engineering Education at Purdue University. He has degrees from Swarthmore College, Rensselaer Polytechnic Institute, and the University of Florida. His research on the longitudinal study of engineering students, team assignment, peer evaluation, and active and collaborative teaching methods has been supported by the National Science Foundation and the Sloan Foundation and his team received Best Paper awards from the Journal of Engineering Education in 2008 and 2011 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.

### **Dr. Daniel M. Ferguson, Purdue University-Main Campus, West Lafayette (College of Engineering)**

Daniel M. Ferguson is CATME Managing Director and the recipient of several NSF awards for research in engineering education and a research associate at Purdue University. Prior to coming to Purdue he was Assistant Professor of Entrepreneurship at Ohio Northern University. Before assuming that position he was Associate Director of the Inter-Professional Studies Program [IPRO] and Senior Lecturer at Illinois Institute of Technology and involved in research in service learning, assessment processes and interventions aimed at improving learning objective attainment. Prior to his University assignments he was the Founder and CEO of The EDI Group, Ltd. and The EDI Group Canada, Ltd, independent professional services companies specializing in B2B electronic commerce and electronic data interchange. The EDI Group companies conducted syndicated market research, offered educational seminars and conferences and published The Journal of Electronic Commerce. He was also a Vice President at the First National Bank of Chicago [now J.P. Morgan Chase], where he founded and managed the bank's market leading professional Cash Management Consulting Group, initiated the bank's non-credit service product management organization and profit center profitability programs and was instrumental in the breakthrough EDI/EFT payment system implemented by General Motors. Dr. Ferguson is a graduate of Notre Dame, Stanford and Purdue Universities, a special edition editor of the Journal of Engineering Entrepreneurship and a member of Tau Beta Pi.

**Full Paper: Comparing Teamwork Peer Evaluations Between  
Cultural Homogenous Teams and Cultural Diverse Teams**

## **Abstract**

This teaching and learning experience paper examines cultural influences on engineering student teamwork and peer assessment behavior. Teamwork is considered as one major competency that engineers in both academy and industry. Dominant first year engineering programs create team-based fundamental engineering courses to develop students' ability to work in a team. The number of foreign students studying at American higher educational institutions is consistently increasing and they possess unique cultures influencing the study and experience of domestic students. In this study, we focus on teamwork and peer assessment behaviors of multi-cultural teams as compared to domestic [U.S.] teams in a large Midwestern first year engineering program. Our research question is: do teams containing one or more international student(s) have, on average, different peer rating behaviors from teams containing only domestic students? We find significant peer evaluation differences in three CATME dimensions: Contributing to the team's work; Interacting with teammates; Keeping team on track.

## **Introduction**

In the first year engineering class where we collected teamwork peer evaluation data, 55.56% of the total number of teams include at least one international students. As Joseph Distefano pointed out, diverse teams typically perform worse or better than homogeneous teams and better performance for diverse teams is conditioned on proper management and training because "Compared to homogeneous teams they (multicultural team) can be more creative, generate more and better alternatives to problems, and generate more and better criteria for evaluating alternatives". [1] It implies that if a team has more complexly cultural background, then the team has greater potential to produce better outcome. One important aspect to achieve better performance is that each team members have good teamwork behaviors as the prerequisite to ensure a functioning team. Identifying whether the differences of teamwork behaviors between diverse teams and homogenous team exists is the key and foundation to understand this phenomenon. To collect data and analyze students' teamwork behavior, we use CATME system, a validated behavior-anchored rating tool using Likert scale in five dimensions [2]. This paper compares the CATME peer ratings for three groups: (1) teams with all domestic students; (2) teams including one international students; (3) teams containing two international students from different nations. By comparing CATME rating scores for these three groups, we would like to examine the pattern of teamwork behavior differences.

## **Research Data**

We collect three rounds student peer evaluation data from an introductory engineering course in FYE department of a large Midwestern University in Fall 2017 for all 15 sections. All students are required to participate in rater practice training [3]- [4], from both online module videos and professors' lectures for course credit. All students are assigned into a team consisting of 3 or 4 persons by CATME Team Maker system [5]. We categorize all teams into the following groups: (1) teams with all domestic students; (2) teams having only one international student; (3) teams containing two international students from different nations; (4) teams consisting of two foreign students from different countries; (5) teams having three international students. For group (4)

and (5), there are only 3 and 9 teams respectively, so we exclude them outside this research because the small number of sample size leads to insignificant comparison result. We remove those teams where at least one student drops the course or failed to provide nationality. Group (1) has 164 teams containing 647 students; Group (2) has 136 teams with 542 students; Group (3) includes 57 teams with 214 students. All of these three groups have 997 male students and 355 female students and 51 students did not respond.

### Statistical Models

We use ANOVA test with Bonferroni correction to examine the dispersion pattern for both all five dimensions rating data and individual dimension data [6] and convergence pattern between self-rating and teammate-rating in terms of their standard deviation and average score [7].

We also embed a conceptual and analytic method called Social Relations Model (SRM) to examine interpersonal perceptions of the ratee and the rater [8]. For the purpose of this study, the target variance is the most important focus among three major variance component of SRM (actor, target and relationship variance). Target variance measures the extent that an individual is rated similarly by all other teammates and the larger value of target variance, the more consistent rating is [8].

### Analysis and finding

We perform ANOVA test for all CATME dimensions rating result for both average score and standard deviation and we do not find any significant difference for both across peer review time and across intervention groups. All across intervention comparisons do not show difference. By analyzing across peer review period, all intervention groups show significant difference in second round of peer evaluation compared with first round and no change for the third peer evaluation compared with second round. We are also not able to find difference for rating pattern for self-rating and other teammate rating for average score. However, as shown in table 1, we find various developed patterns for self-rating and peer-rating dispersion in standard deviation. All three intervention groups show significant difference in peer review 1 to 2. However, from peer evaluation round 2 to 3, ratings for all domestic student teams are statistically different (with the threshold of 5%), while ratings for teams containing international students appear converged. It suggests that teams with only domestic students tend to have less dispersion within their teams. One plausible reason could be there are relatively less distribution for this group compared with teams containing international students.

Table 1 Dispersion (Standard Deviation) Across All Dimensions AVONA Analysis

	Intervention	Peer Review Time	Intervention	Peer Review Time	Dispersion Difference	P-Value	Effect Size (Cohen's d)
Self-rating	All Domestic	1	All Domestic	2	-0.1856	<0.0001	0.6047
	All Domestic	2	All Domestic	3	-0.07244	0.0042	0.2221
	1 INTL student	1	1 INTL student	2	-0.1889	<0.0001	0.4915
	1 INTL student	2	1 INTL student	3	-0.05580	0.0983	0.1240
	2 INTL students from same nation	1	2 INTL students from same nation	2	-0.1785	<0.0001	0.5041

	2 INTL students from same nation	2	2 INTL students from same nation	3	-0.01983	1.0000	0.1730
Peer-rating	All Domestic	1	All Domestic	2	-0.7027	<.0001	1.8521
	All Domestic	2	All Domestic	3	-0.06670	0.0181	0.3430
	1 INTL student	1	1 INTL student	2	-0.7117	<.0001	1.2673
	1 INTL student	2	1 INTL student	3	-0.05559	0.1380	0.1242
	2 INTL students from same nation	1	2 INTL students from same nation	2	-0.6846	<.0001	1.3313
	2 INTL students from same nation	2	2 INTL students from same nation	3	-0.03948	1.0000	0.2370

We also conduct ANOVA analysis on the average difference of self-rating and peer-rating to see how one rate himself/herself is different from how teammates rate this individual. We also do not find any statistical difference for these three groups.

We then perform the divergence and convergence test for average score of CATME five individual dimensions (Contributing to the team's work, Having relative KSAs, Interacting with teammates, Expecting quality and Keeping the team on track). As shown in Table 2, for dimension C, I and K, we found that there exist some discrepancies for rating patterns across peer reviews. For those three mentioned dimensions, rating converges after second peer evaluation, however, for teams containing international students, it is not the case. They did not converge and possess some similar rating patterns because we cannot find significant difference from rating 1 to 2 and from rating 2 to 3. Compared with the other two dimensions (H and K), these three dimensions (C, I and K) are relatively more subjective so that it is reasonable that international students are hard to distinguish the difference of the grading scale and calibrate their grading scale to desired expert rating result.

Table 2 Mean Individual Dimensions ANOVA Divergence Analysis

	Intervention	Peer Review Time	Intervention	Peer Review Time	Mean Difference	P-Value	Effect Size (Cohen's d)
Contributing to the Team's Work	All Domestic	1	All Domestic	2	0.1147	0.0114	0.2221
	All Domestic	2	All Domestic	3	0.03315	1.0000	0.0764
	1 INTL student	1	1 INTL student	2	0.06477	0.7584	0.0808
	1 INTL student	2	1 INTL student	3	0.09019	0.1462	0.0039
	2 INTL students from same nation	1	2 INTL students from same nation	2	0.1182	0.4184	0.0959
	2 INTL students from same nation	2	2 INTL students from same nation	3	0.1045	0.7036	0.0922
Interacting with teammates	All Domestic	1	All Domestic	2	0.1239	0.0146	0.2267
	All Domestic	2	All Domestic	3	0.05217	1.0000	0.0707
	1 INTL student	1	1 INTL student	2	0.07646	0.5894	0.0808
	1 INTL student	2	1 INTL student	3	0.01214	1.0000	0.0142
	2 INTL students from same nation	1	2 INTL students from same nation	2	0.04242	1.0000	0.0896
	2 INTL students from same nation	2	2 INTL students from same nation	3	0.1152	0.7124	0.0892

Keeping the team on track	All Domestic	1	All Domestic	2	0.1473	0.0034	0.2511
	All Domestic	2	All Domestic	3	0.03370	1.0000	0.0410
	1 INTL student	1	1 INTL student	2	0.09527	0.2654	0.1274
	1 INTL student	2	1 INTL student	3	-0.01335	1.0000	0.0152
	2 INTL students from same nation	1	2 INTL students from same nation	2	0.1848	0.0678	0.1484
	2 INTL students from same nation	2	2 INTL students from same nation	3	0.03333	1.0000	0.1172

Table 3 shows the least squares means if difference between self and peer rating. We only show the last evaluation because several research conducted based on CATME system and data indicates that the rating ability and result tend to be stable for the last round of peer evaluation [3], [4], [7]. As a result, we could identify the difference caused by deep and inherent reasons with minimal external interferes. As shown in dimension C, I and K, mean differences increase from no international student teams to 2 international student teams. This phenomenon indicates that team members of smaller number of international students tend to give lower rating to themselves compared with their peers. So for teams containing international students, they either over estimate their teamwork performance or have difficulty distinguishing perception from behavior in teaming.

Table 3 Mean Individual Dimensions AVONA Convergence Analysis for Third Peer Review

Dim	All Domestic			1 INTL student			2 INTL students from same nation		
	Mean Difference	P-Value	Effect Size (Cohen's d)	Mean Difference	P-Value	Effect Size (Cohen's d)	Mean Difference	P-Value	Effect Size (Cohen's d)
C	-0.05869	0.1043	0.0302	-0.01251	0.7428	0.0028	0.08182	0.1750	0.0286
I	-0.01667	0.6613	0.0282	0.04733	0.2391	0.0231	0.05152	0.4173	0.0239
K	0.01304	0.7474	0.0528	0.06109	0.1536	0.0307	0.06762	0.7089	0.0561
H	-0.05290	0.1392	0.0235	-0.06634	0.0793	0.0038	-0.01212	0.8392	0.0394
E	0.004710	0.9026	0.0609	0.01497	0.7129	0.0407	0.03131	0.6263	0.0408

We analyze Social Relations Model for all dimension data across peer review and we do not find significant difference for that. However, rating difference of SRM analysis among three groups echoes in dimensions C, I and K as shown in figure 1-3. For dimension C, domestic teams have different development pattern for target variance compared with other two groups. And for peer evaluation round 3 specifically, teams only containing international students have highest target variance and teams containing two international students from the same nation have lowest value. It indicates that students from similar background tend to rate consistently for their teammates and vice versa. For dimension I, the situation becomes a little bit different because teams with one international students have the highest target variance while two international student teams remain the lowest variance. For dimension K, the difference of target variance is obvious across three rounds of peer reviews. Teams without domestic students consistently have lowest target variance and teams with two international students consistently have highest target variance.

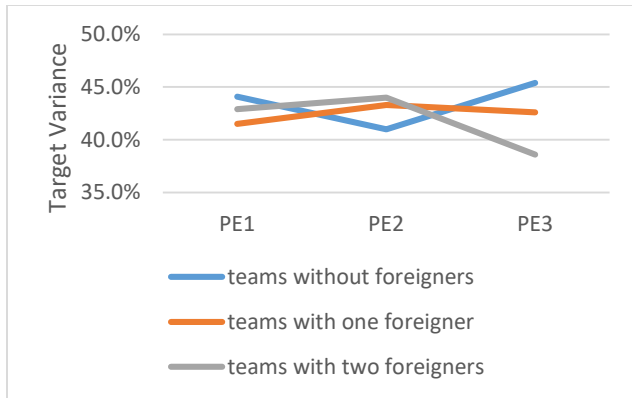


Figure 1 SRM Analysis for Individual Dimension (C)

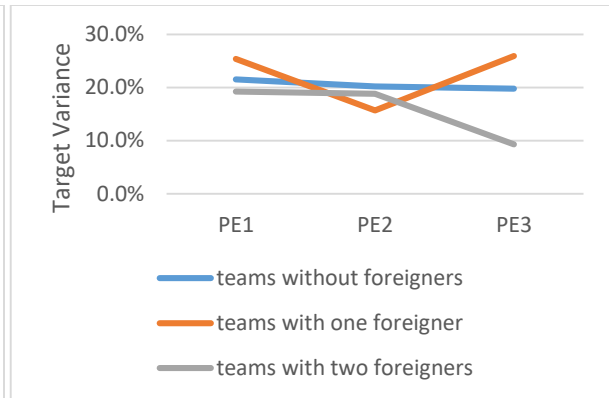


Figure 2 SRM Analysis for Individual Dimension (I)

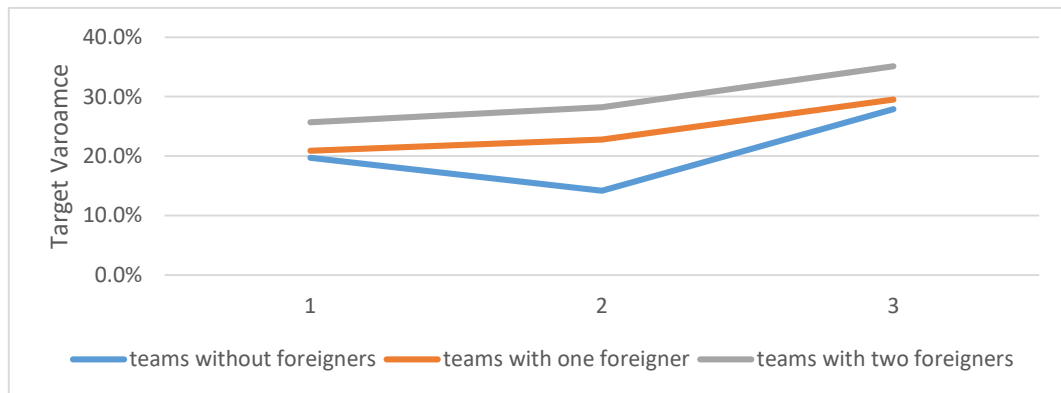


Figure 3 SRM Analysis for Individual Dimension (K)

## Limitation

Due to small number of teams for teams with two international students from different nations and teams of three international students, we are unable to analyze and compare their performance compared with other three intervention groups. We do expect to see differences especially from all domestic student teams. The way that we define international students is only based their nationality information collected when they enroll the university. We do not consider the cultural similarity between international student and domestic students. For example, Canadian students are culturally similar to domestic American students. We do not also consider the team demographics based on gender, which is also believed to influence team dynamics. We also do not consider cultural distance among students' ethnicity. We assume the students' rating is accurate for this study.

## Summary and conclusion

We perform both AVOVA test and SRM analysis on student peer evaluation results from all dimensions and individual dimension on team level comparing teams with 0, 1 and 2 international students from same nation. For all dimension analysis, we cannot find statistical difference for average rating score and standard deviation in dispersion and convergence. However, we identify significant difference for both regression and SRM analysis in three of five CATME dimensions: Contributing to the team's work; Interacting with teammates; Keeping team on track.

Based on ad-hoc experience of authors, we often observe that international students interact with teammates minimally and their discussions are usually limited to academic area. To some extent, the only international student in the team tends to be isolated spontaneously and this phenomenon leads to relatively stable team dynamics, which partially explains why in dimension I teams with one foreigners rate each other most consistently. However, when teams contain two international students from same nation, they tend to unite as a subgroup within the team and force the other two domestic students to eventually build up another subgroup within a team.

This study identifies the existing teamwork behavior rating difference between homogenous teams and cultural diverse teams. Interested researches could further investigate the team performance between these groups and develop optimal method to assign diverse students into teams.

## Reference

- [1] J. J. Distefano and M. L. Maznevski, "Creating value with Diverse Teams in Global Management." *Organizational Dynamics*, 29(1), 45-63, 2000.
- [2] M. W. Ohland, M. L. Loughry, D. J. Woehr, C. J. Finelli, L. G. Bullard, R. M. Felder, 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." *Academy of Management Learning & Education*, vol. 11, pp. 609-630, 2012.
- [3] E. Shu, A. Collins, Y. Cao, J. Napiokowski, D. Ferguson and M. W. Ohland, Poster: "Examining the Effect of a Game-like Practice Tool on the Quality of Peer Evaluations." *Proceeding of the American Society for Engineering Education Annual Conference & Exposition*, Salt Lake City, UT, 2018.
- [4] M. Benner and D. M. Ferguson, "Continued Examination of the Effect of a Game-like Practice Tool On the Quality of Student Peer Evaluations." Manuscript submitted for publication, 2019.
- [5] A. Layton, L. Loughry, M. Ohland and D. Ricco. "Design and validation of a web-based system for assigning members to teams using instructor-specified criteria." *Advances in Engineering Education*, vol. 2(1), pp. 1-28, 2010.
- [6] T. Poling, D. J. Woehr, L. M. Arciniega and A. Gorman, "The Impact of Personality and Value Diversity on team performance." *Annual Meeting for the Society for Industrial and Organizational Psychology*, Dallas, TX, 2006.
- [7] D. M. Ferguson, M.W. Ohland, C. Lally, H.I. Somnooma and Y. Cao, "Evaluating the Effect of Different Teamwork Training Interventions on the Quality of Peer Evaluations." *Proceedings of the Frontiers in Education Conference*, San Jose, CA, 2018.
- [8] A. Loignon, D. Woehr, J. Thmoas, M. Loughry, M. Ohland and D. Ferguson, "Facilitating Peer Evaluation in Team Contexts: The Impact of Frame-of-Reference Rater Training.", *Academy of Management Learning & Education*, vol. 16(4), pp. 562-578, 2017.