Work in Progress: Minority Bias in Peer Evaluations at a Freshman-level Engineering Cornerstone Course

Prof. Catalina Cortázar, Pontificia Universidad Católica de Chile

Catalina Cortázar is an Adjunct Assistant Professor at DiLab the Design initiative at the School of Engineering at Pontificia Universidad Católica de Chile. Catalina holds a bachelor degree in Civil Engineering, with concentration in Structural Design. After graduating and working at an Engineering firm in Chile, Catalina completed a master's degree in media studies at The New School, and a MFA in Design and Technology @ Parsons The New School for Design, New York. At DiLab Catalina teaches and coordinates the Engineering Challenges course which aims to initiate freshmen students in to engineering design practices by encouraging students to develop a project following a user-centered design process. She also teaches Visual Thinking, the exploratory course of the Major in Engineering, Design, and Innovation. This course addresses the theories and ideas that sustain the visual thinking process as well as methodologies and practical implementation of visual representation through infographics, computer graphics, and physical computing. The course focus on representing the narrative of the findings using visual tools. Catalina has been directing FabLabUC since 2015. FabLabUC is a fabrication laboratory located at the Innovation Center, PUC . Currently she is pursuing a PhD in Computer Sciences with a research focus on Engineering Education at PUC.

Ing. Isabel Hilliger, Pontificia Universidad Catholica de Chile

Isabel Hilliger is the Associate Director for Assessment and Evaluation at the Engineering Education Division in Pontificia Universidad Católica de Chile (UC). Isabel received a BEng from UC and an MA in Education Policy from Stanford University. She is currently a PhD Candidate in Computer Science at UC-Engineering. Her research theme is the use of methodologies and analytical tools for continuous curriculum improvement in Higher Education. She has created qualitative and quantitative instruments for outcome assessment in enginering education. She has also evaluated policy efforts towards engineering diversity and undergraduate research.

WIP: Using a Peer Evaluation tool to explores minority bias at a Freshman-level Engineering cornerstone course

This Work in Progress (WIP) paper explores the use of a peer evaluation tool to analyze if minority biases exist when students evaluate their peers at a Freshman-level Engineering Cornerstone course. In 2002, this course was created partly in response to the particular emphasis in the use of methodologies that favor collaborative learning in engineering courses [1]. Organizations such as the Accreditation Board for Engineering and Technology (ABET) have encouraged team projects to develop skills such as leadership, effective communication, and conflict resolution [2]. Therefore, these learning instances are essential to prepare the future engineers to face the world of real work [1].

However, the benefits of cooperative learning are not immediate or automatic [3]. There are certain considerations to be taken into account in order to avoid non-participation, social loafing, and minority bias. By non-participation, the literature refers to situations when team members do not attend team meetings, show no interest in teamwork at an academic context, or refuse to do anything by themselves [4]. Other issue is denominated "Social Loafing"; behaviors under which some team members do not contribute their fair share to the project. According to the literature, peer evaluations are one of the mechanisms that can reduce the presence of social loafing by identifying individual contributions to teamwork [5].

Research has shown that students receive peer evaluation scores that correlate positively with their average academic performance [6]. However, students' attitudes towards peer evaluations are mixed. Many students are concerned that shared assessments will be skewed [7]. Other concerns are about the lack of self-criticism in self-assessment, collusion among team members to avoid conflict, and finally, gender bias, race bias, and prejudice that could influence voting [6]. Millis and Cottell [9] propose four components to be evaluated by peers: regular attendance to meetings, completion of assignments, contributions of each member and / or support within the group when necessary

This study examines the development and use of a peer evaluation tool in a freshman core-curriculum course at a School of Engineering in Chile (CL-Engineering). We are analyzing different peer evaluation strategies used during the first semesters of 2014, 2015, and 2016 to explore if bias exists when students evaluate minorities (i.e. female students, alternative admission students, and student from other geographical areas). We are interested in receiving feedback from this community on the different ways we can use these data in order to answer the following potential research question: Does bias exist when students evaluate minorities in the context of teamwork in an Engineering cornerstone course?

Course Description

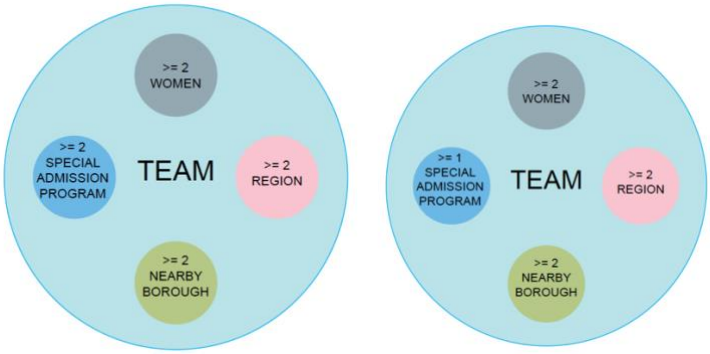
The overall aim of this cornerstone course is to introduce freshmen to engineering design practices by encouraging students to develop a project following a user-centered design process. Regarding course teaching methods, it's a project-based course where students must work on teams to design a product that solves a real-world problem, for example: Urban cyclist, Reduce Reuse, Recycle in Chile, and Firefighters.

Concerning assessment methods, the course measures students' performance at an individual and team level. At an individual level there are homework and an exam. At a team level, students are required to submit different project deliverables as the course progresses. At the end of the course, there is a technological fair were teams exhibit their final prototypes and present their findings to contextual stakeholders. Students are peer evaluated within the semester in order to grade their individual contribution to the teamwork. In this study we examined peer evaluation practices in this course and its evolution through the first semesters of 2014 (1-2014), 2015 (1-2015), and 2016 (1-2016).

Class And Team Demographic

In the context of this course, the Office for Undergraduate Studies of the School of Engineering made the working teams using the criteria shown in Figure 1. This team building strategy was established in order to avoid isolating students in the minority subgroups described above.

Figure 1. Team demographic for 2014 & 2016 (Left) & 2015 (Right). Adapted from a presentation by the Office for Undergraduate Studies



As shown in Figure 1, the criteria used in 2014 and 2016 paired students belonging to minority subgroups: female, students who came from outside the Metropolitan Region, and students who entered through alternative admission programs. Students living in nearby borough were paired, while students that come from the same high school were separated. During 2015, because of external reasons, students who enrolled through an alternative admission program were left one per group.

Table 1 shows data of course enrollment in the course 1-2014, 1-2015, and 1-2016. Besides describing course enrollment data, it presents data of minorities' representation, and the standard deviation of the University Selection Test as academic indicator.

Table 1. Class demographic for the first semester of 2014, 2015, and 2016.

	1-2014	1-2015	1-2016
Course Enrollment			
Number of Sections (n)	7	8	9
Total number of students that completed the course (n)	724	734	758
Number of teams (n)	70	76	108
Number of students per team	11-10	10-9	8-7
Minorities			
Females that completed the course	197 (27%)	180 (25%)	211 (28%)
Students registered by alternative admission program that completed the course.	108 (15%)	81 (11%)	112 (15%)
Students from outside the Metropolitan Region (RM) that completed the course**	188 (26%)	132 (18%)	182 (24%)
Total number of minority students that completed the course	493	393	505

Notes:
** *Students from outside the Metropolitan Region (RM)* are the students who graduated from high school at institutions outside RM.

Regarding minorities representation, we have considered gender, admission process, and students from other regions. Participation of foreign students was left out of the scope of the study because this subgroup does not represent a significant population in this Engineer student body. Even though immigration in Chile has increased in the last 10 years from 1% to 2,7% of the total population [10], foreign students in CL-Engineering are less than 1%.

In what respects to gender, CL-Engineering female students range between 25% and 28%. When reviewing the enrollment rate by gender in the scientific careers in Chile, it is observed that the percentage of women enrolled to study careers related to the sciences and technology is only 23.7% [11], so it is important to analyze the experience of this subgroup.

Concerning the admission process, the indicators that lead to admission in Chilean Universities are high-school grades and a Standardized University Admission test known as PSU (University Selection Test). Because of disparities in test preparation between students from high- and low-income background the system has historically privileged the entry of upper-class students to the university system [12], [13]. Consequently, the school of engineering has implemented an alternative admission program to ensure the inclusion of talented students from lower socioeconomic background [14].

Students that come from regions outside the Metropolitan Area also represent a minority in the Chilean context conserving that the latest census of 2012 indicates that the Metropolitan Region brings together 40.33% of Chileans [15].

Peer Evaluation Practices

Students evaluate their peers answering an online questionnaire. During 2014 students they were asked to evaluate their peers in two instances during the semester, responding to the following question: *Taking into account Availability, Participation, Creativity and Effort. How many points would you give your peers?*

In 2015 student assessed their peers in 2 instances and in 2016 after each course delivery (5 times). Both years a system of 3 questions was implemented:

1- *Individual contribution outside meetings and team activities:* Taking into account if he/she performs the tasks assigned by the team within the stipulated time and if his/her work is rigorous enhancing the team: *How many points* would you give your teammates?

2- *Contribution to team meetings:* Taking into account availability and participation in team meetings and activities: *How many points* would you give your teammates?

3- *Working environment:* Taking into account if he/she promotes a constructive team environment by transmitting a positive and respectful attitude toward work and team members: *How many points* would you give your peers?

In order to grade their peers every team member is given an amount of points to distribute within their teammates. Students are not obliged to use all of their points.

Current Research and Data Collection

This research is part of a larger study to explore the use of peer evaluations in the context of a core cornerstone course. To examine if there are biases towards minorities, we will use peer evaluation results from the first semesters of 2014, 2015 and 2016, along with individual grades and demographical data obtained from administrative sources. 238 teams were analyzed, which were composed by 2.057 freshmen students. So far, we have analyzed the correlation between the peer evaluations and individual grade averages by conducting linear regressions in STATA SE/14, using peer evaluations average as dependent variable and the robust option due to heteroscedasticity. As we conducted linear regressions, we controlled for:

- The logarithm of the scores associated to students high school GPA (Ln(High School GPA)) (logarithm function used to transform variable due to skewed scores according [16]; see variable distribution in the following link: http://bit.ly/2HxPIpd)
- Gender as a categorical variable (Female=1, Male=0),
- Admission (Alternative Admission=1, Regular Admission=0),
- Region of origin (Outside Metropolitan Region=1, Metropolitan Region=0).

Tables 2, 3 and 4 show the results obtained for the first semesters of 2014, 2015 and 2016 respectively. For these three periods, results confirm that students receive peer evaluation scores that correlate positively with their average individual grades [6]. However, results also show that some collusion, bias or prejudice might have affected scores in 2015 and 2016. In 2015, students

who were admitted through the alternative program scored on average two decimals less than those who were admitted regularly (with a statistical significance of 95% for Model 2 and 90% for Models 3 and 4). Besides, students who came from outside the metropolitan region scored on average one and a half less than those who came from this region (with a statistical significance of 99%). This statistically significant difference is sustained in 2016 with the same statistical significance.

Table 2. Linear regression results obtained from using peer evaluation averages as dependent variable for explore bias towards minority students in the first semester of 2014

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
Av. Individual grade	0.550***	0.551***	0.549***	0.549***
	(0.0722)	(0.0730)	(0.0739)	(0.0739)
Ln(High School GPA)	1.502**	1.518**	1.580***	1.378**
	(0.605)	(0.595)	(0.598)	(0.638)
Alt. Admission = 1		0.0150	0.0121	0.00320
		(0.0970)	(0.0969)	(0.0958)
Outside M. Region = 1			-0.0616	-0.0580
			(0.0528)	(0.0530)
Female = 1				0.0854
				(0.0572)
Constant	-8.417**	-8.532**	-8.916**	-7.597*
	(3.847)	(3.778)	(3.793)	(4.051)
Observations	602	602	602	602
R-squared	0.244	0.244	0.246	0.249

Notes:
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Difference in the number of observations with the numbers presented in Table 1 are due to missing data concerning the scores associated to High School GPA, besides pending individual grades.

Table 3. Linear regression results obtained from using peer evaluation averages as dependent variable for explore bias towards minority students in the first semester of 2015

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
Av. Individual grade	0.330***	0.324***	0.318***	0.314***
	(0.0614)	(0.0618)	(0.0608)	(0.0624)
Ln(High School GPA)	0.510	0.412	0.522	0.474
	(0.512)	(0.500)	(0.499)	(0.504)
Alt. Admission = 1		-0.188**	-0.167*	-0.175*

		(0.0922)	(0.0926)	(0.0938)
Outside M. Region = 1			-0.147***	-0.147***
			(0.0557)	(0.0557)
Female = 1				0.0326
				(0.0472)
Constant	-0.733	-0.0367	-0.710	-0.371
	(3.289)	(3.194)	(3.198)	(3.240)
Observations	706	706	706	706
R-squared	0.115	0.122	0.131	0.132

Notes:
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Difference in the number of observations with the numbers presented in Table 1 are due to missing data concerning the scores associated to High School GPA, besides pending individual grades.

Table 4. Linear regression results obtained from using peer evaluation averages as dependent variable for explore bias towards minority students in the first semester of 2016

	(1)	(2)	(3)	(4)
VARIABLES	Model 1	Model 2	Model 3	Model 4
Av. Individual grade	0.226***	0.224***	0.221***	0.221***
	(0.0432)	(0.0433)	(0.0434)	(0.0436)
Ln(High School GPA)	0.573	0.525	0.594*	0.599*
	(0.349)	(0.354)	(0.360)	(0.363)
Alt. Admission = 1		-0.0440	-0.0555	-0.0549
		(0.0548)	(0.0542)	(0.0547)
Outside M. Region = 1			-0.102***	-0.102***
			(0.0355)	(0.0353)
Female = 1				-0.00255
				(0.0312)
Constant	-0.612	-0.275	-0.690	-0.725
	(2.235)	(2.278)	(2.310)	(2.342)
Observations	710	710	710	710
R-squared	0.134	0.135	0.146	0.146

Notes:
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Difference in the number of observations with the numbers presented in Table 1 are due to missing data concerning the scores associated to High School GPA, besides pending individual grades.

Discussion

Although the preliminary results in this paper contribute to a better understanding of peer

evaluation scoring at a CL-engineering cornerstone course by presenting the course and the type of instrument applied, we still need to address some issues to understand students' attitudes towards minorities. First, Ramsey test in Stata indicated that the models estimated had omitted variables. Second, we are aware that observing statistically significant coefficients is not enough to examine issues associated to bias toward students coming from a geographical area further from CL-Engineering campus or being admitted through an alternative program. Future work should address these issues in detail by disaggregating peer evaluation scores at a question level, besides exploring institutional dynamics that might affect students' perceptions in a teamwork setting. We are interested in receiving feedback from this community on other issues that we should consider in future studies.

Acknowledgement

Reserved for blind review.

References

[1] M. Borrego, J. Karlin, L. D. Mcnair, & K. Beddoes, "Team effectiveness theory from industrial and organizational psychology applied to engineering student project teams: A research review". Journal of Engineering Education, 2013. Available: https://doi.org/10.1002/jee.20023

[2] A. Patil, & G. Codner, "Accreditation of engineering education: review, observations and proposal for global accreditation". European Journal of Engineering Education, 2007. Available: https://doi.org/10.1080/03043790701520594

[3] Chin-Min Hsiung. "The Effectiveness of Cooperative Learning". Journal of Engineering Education, 2012. Available: https://doi.org/10.1002/j.2168-9830.2012.tb00044.x

[4] R. Felder, & R. Brent. "Effective strategies for cooperative learning". Journal of Cooperation & Collaboration, 2001. Available: https://doi.org/http://dx.doi.org/10.1016/S0742-051X(96)00045-5

[5] K. H. Price, D. A. Harrison, & J. H. Gavin. "Withholding inputs in team contexts: Member composition, interaction processes, evaluation structure, and social loafing". Journal of Applied Psychology, 2006. Available: https://doi.org/10.1037/0021-9010.91.6.1375

[6] Kaufman, D. B., Felder, R. M., & Fuller, H. "Accounting for individual effort in cooperative learning teams". Journal of Engineering Education, 2000. Available: https://doi.org/10.1002/j.2168-9830.2000.tb00507.x [7] M. W. Ohland, R. Layton, M. L. Loughry, & G. Yuhasz, "Effects of behavioral anchors on peer evaluation reliability". Journal of Engineering Education, 2005.

[9] B. J. Millis, & P. G. Cottell, "Cooperative Learning for Higher Education Faculty. Series on Higher Education". *American Council on EducationOryx Press series on higher education*, 1998.

[10] OECD, International Migration Outlook 2017, *OECD Publishing*, Paris. p.178, 2017. Available: http://dx.doi.org/10.1787/migr_outlook-2017-en

[11] Consejo Nacional de Educación Distribución de la matrícula total por área del conocimiento y género Univ, IP, CFT. Santiago, Chile, 2013. Available: https://www.cned.cl/

[12] A. Bernasconi, & F. Rojas, "Informe sobre la Educación Superior en Chile: 1980-2003". New York, 2003.

[13] A. Matear, "Barriers to equitable access: Higher education policy and practice in Chile since 1990". *Higher Education Policy*, 2006.

[14] I. Hilliger, C. Gelmi, L. Cifuentes, M. Bennett & J. C., De la Llera. "Design and implementation of an alternative admission program to engineering: Talent and Inclusion". *Studies in Higher Education*, V. 43, pp. 1454-1467, 2018.

[15] Ministry of the Interior and Public Security. Información Geográfica. Santiago, Chile: Autor. 2013. Available: http://intendenciametropolitana.gov.cl

[16] Osborne, Jason. "Notes on the use of data transformations." *Practical Assessment, Research & Evaluation*, 8(6). Retrieved December 2, 2005 from: http://PAREonline.net/getvn.asp?v=8&n=6 .