Academic Integrity during the COVID-19 Pandemic: a Social Media Mining Study

Mohammad S. Parsa
University of Waterloo
mohammad.parsa@uwaterloo.ca

Lukasz Golab
University of Waterloo
lgolab@uwaterloo.ca

ABSTRACT

Academic integrity has been a frequently reported challenge in online education. Given the widespread transition to online program delivery during the COVID-19 pandemic, we ask the following question: *How do college students feel about online cheating?* Our analysis is based on academic discussions on the Reddit social curation platform in Fall 2020 and, for comparison, Fall 2019. We found more discussions related to cheating in 2020 than in 2019, and the topics have expanded from plagiarism in programming assignments to online assessments in general. Topic modelling of the Fall 2020 discussions revealed three concerns raised by students: that cheating inflates grades and forces instructors to increase the difficulty of assessments; that witnessing cheating go unpunished is demotivating; and that academic integrity policies are not always communicated clearly.

Keywords

academic integrity, online education, social media, text mining

1. INTRODUCTION

Recent studies have reported that online academic misconduct has increased during the COVID-19 pandemic [12, 6, 2, 4, 3, 18]. We therefore ask the following question in this paper: *How do college students feel about online cheating?* To answer this question, we turn the Reddit social curation platform (reddit.com). Reddit hosts over 100,000 user-created discussion communities refereed to as *subreddits*. Within a subreddit, users create posts that other users comment on. Subreddit names begin with "r/" and correspond to the subreddit topic, e.g., r/politics or r/relationship_advice.

Descriptive subreddit names make it easy to locate discussions about specific topics or discussions initiated by various kinds of users. Of interest to our study are over 80 subreddits corresponding to Canadian and U.S. universities, which we call *academic subreddits*. We collected all posts and comments on academic subreddits created during the Fall 2019 and Fall 2020 semesters (September through December inclusive) that match at least one keyword related to cheating, such as 'cheat' or 'misconduct'.

Our analysis consists of two steps. First, collecting data from the same time period in 2019 and 2020 allows us to compare cheating-oriented discussions from before the pandemic, when classes were held in person, and during the pandemic, with most courses delivered online. To do so, we train a logistic regression classifier to distinguish between Fall 2019 and Fall 2020 content based on the words used. Next, we analyze Fall 2020 discussions in detail. We apply the Non-negative Matrix Factorization algorithm [20], which clusters posts and comments based on the words used and allows us to identify common discussion topics.

Related Work: Social media have become a go-to source of public opinion on a variety of topics. In particular, academic subreddits have been analyzed in recent work on students' mental health [1, 16], but academic integrity was not discussed. The closest works to ours are those in [4] and [5], which interviewed a small set of undergraduate students and educators. The participants identified some positive aspects of online education, but expressed concerns about cheating and the level of difficulty of online assessments. Our social media analysis explores these and other concerns in detail.

2. DATA AND METHODS

Previous work on students' mental health [1, 16] identified 83 *academic* subreddits corresponding to major U.S. and Canadian universities. We analyze the same subreddits in this paper, listed in the first column of Table 1 (U.S.) and Table 2 (Canadian). We collected all posts and comments on these subreddits from the Fall 2019 semester, when classes and examinations were held in person, and the Fall 2020 semester, when most campuses moved to online delivery (September-December inclusive). We downloaded the data using a publicly-accessible Reddit interface at pushshift.io.

Next, we retain only those posts and comments that contain at least one of the following keywords: 'cheat', 'plagiari', and 'misconduct'. We perform *substring* matching, meaning that 'plagari' also matches 'plagiarize' and 'plagiarism'. Tables 1 and 2 report the number of posts ("P") and comments ("C") on each U.S. and Canadian academic subreddit, respectively, in Fall 2019 and Fall 2020. The "Before" numbers correspond to all posts and comments. The "After" numbers correspond to posts and comments that matched at least one cheating-related keyword; note that there are three times as many such posts and comments in 2020 than in 2019 (7,809 vs. 2,524) even though the total number of posts and comments on academic subreddits has not changed much from 2019 to 2020 (see the total "Before" numbers in the last row of Tables 1 and 2).

We then perform standard text pre-processing. Following previous work on Reddit topic modelling [10, 16], we remove posts and

Table 1: Number of posts and comments on U.S. academic subreddits in 2019 and 2020 before and after filtering to find cheating-related discussions (C: Comments, P: Posts).

Subreddits	2020				2019			
	Before		After		Before		After	
	C	P	C	P	C	P	C	P
UIUC	39974	6991	160	21	40556	6431	104	10
berkeley	37355	6343	365	69	28537	4637	114	17
Cornell	36235	8139	165	27	22562	3900	45	8
Purdue	34376	6317	148	15	33322	5273	42	11
UCSD	30589	5798	175	34	28214	5364	106	15
rutgers	29861	6622	269	69	44114	8902	122	16
UMD	21937	4225	206	28	25794	4631	97	6
SBU	20521	4301	163	20	28328	5373	63	13
uofm	19954	3174	79	14	13553	2213	44	5
udub	17867	3487	82	17	18187	3187	59	6
UWMadison	14870	2447	103	18	14236	2039	33	3
UTAustin	13620	3112	53	7	13866	2811	90	6
utdallas	12763	2235	74	7	20731	3109	25	5
PennStateUniversity	12345	1944	64	5	9620	1610	42	2
msu	12052	2104	86	10	15066	2329	23	6
NCSU	11653	1794	72	5	18943	2524	32	1
UVA	11627	2424	79	9	5071	1084	19	5
rit	11603	1577	48	2	10768	1643	6	1
nyu	11034	2952	37	7	5731	1438	10	2
UNCCharlotte	10132	1709	93	12	10700	1508	18	1
USC	9551	1958	82	15	6800	1419	17	4
Baruch	9370	2226	94	16	4851	1144	36	12
UPenn	8886	2083	55	10	4212	997	11	1
UNC	8347	1644	30	8	3800	790	6	2
byu	6951	707	39	2	3165	407	25	3
UGA	6637	1520	20	3	6852	1349	2	0
columbia	6496	1573	55	5	4699	708	22	3
RPI	5652	1220	70	0	7622	1343	5	0
uichicago	4880	894	46	4	6606	1009	84	1
SJSU	4661	1068	27	5	5108	1136	18	3
stanford	3944	1223	13	2	3782	882	10	0
bostoncollege	3493	1006	0	0	753	188	0	0
cmu	3388	657	27	2	2764	517	3	0
washu	3159	572	4	0	1134	259	0	0
Vanderbilt	2581	555	9	1	1447	311	0	0
Harvard	2219	634	1	1	2294	517	1	0
UMBC	2036	457	21	3	2479	464	4	0
duke	2020	469	2	1	1397	317	7	2
mit	1758	532	3	0	1651	373	4	0
BrownU	1363	438	2	1	1315	276	0	0
IndianaUniversity	1225	588	1	1	1797	543	9	1
Caltech	494	130	0	0	220	59	0	0
Total	509479	99849	3122	476	482647	85014	1358	171

comments with fewer than 40 or more than 4000 characters: short ones are unlikely to be meaningful (and may correspond to URLs), while long ones may mention more than one topic. We also remove stopwords and lemmatize the remaining words using the Python NLTK parser.

To distinguish between cheating-related discussions before and during the pandemic, we train a logistic regression classifier to predict whether a post or comment was written in Fall 2020 or Fall 2019. We use term frequency–inverse document frequency (TF-IDF) word scores as features in the model. We chose logistic regression due to its interpretable nature: words with positive coefficients represent Fall 2020 content and words with negative coefficients represent Fall 2019. Our model obtained a 10-fold cross-validation accuracy score of 73%, a precision of 76%, a recall of 96% and an F1-score of 86%.

(We also tested logistic regression models with additional features, including word bigrams, the sentiment of the post or comment (computed using the Valence Aware Dictionary and Sentiment Reasoner (VADER) [8]) and linguistic features computed using Linguistic Inquiry and Word Count (LIWC) [17]. After adding these features, accuracy improved by two percent to 75%. However, none of these additional features were assigned large coefficients and therefore are not considered further in the remainder of the paper.)

Finally, we apply the Non-negative Matrix Factorization (NMF) topic modelling algorithm [20], which was used in prior work on Reddit mining [14, 7, 11], on the Fall 2020 posts and comments that match at least one cheating-related keyword. We again represent each post and comment using the TF-IDF scores of the words occurring in it. NMF clusters documents into topics and assigns a list of representative terms called *topic descriptors* to each topic. NMF also calculates the "representativeness" score of each topic descriptor, and we report the top-10 highest-scoring descriptors for each topic. Moreover, we report top-10 frequent word n-grams (for n up to three, i.e., sequences of up to three consecutive words) for each topic.

Table 2: Number of posts and comments on Canadian academic subreddits in 2019 and 2020 before and after filtering to find cheating-related discussions (C: Comments, P: Posts).

Subreddits	2020				2019			
	Before		After		Before		After	
	C	P	C	P	C	P	C	P
uwaterloo	72244	8372	381	58	88996	9888	130	17
UofT	54343	8460	701	86	67649	9375	171	23
UBC	40058	5281	766	42	39416	5039	109	11
uAlberta	33265	7164	341	58	49494	8270	137	23
McMaster	24556	5188	219	45	14932	2638	27	3
mcgill	21380	3376	167	15	20852	3067	58	6
yorku	15671	4065	228	46	22078	3862	47	6
CarletonU	15455	2531	207	11	16874	2706	43	2
Concordia	10065	2394	192	27	10292	2185	27	7
uwo	9717	1856	122	10	11758	1764	35	2
wlu	8097	1788	97	16	5499	1203	13	4
uvic	7291	1178	85	3	4756	828	11	3
ryerson	6503	2282	87	6	14922	2927	37	8
queensuniversity	5234	1107	18	1	4758	824	6	2
umanitoba	4408	861	66	7	3183	717	3	1
uoguelph	3381	794	51	8	3691	693	5	2
Dalhousie	1807	401	21	4	2019	407	6	2
usask	1177	290	0	0	666	178	0	0
brocku	1007	366	2	0	1442	329	4	2
memorialuniversity	785	183	6	1	637	147	2	0
UdeM	422	90	1	0	174	48	0	0
lakeheadu	119	59	2	1	51	21	0	0
uleth	112	35	0	0	82	33	0	0
University_Of_Regina	96	30	1	0	8	11	0	0
AcadiaU	69	29	1	0	60	15	0	0
UQAM	67	22	0	0	48	17	0	0
uwinnipeg	65	24	2	1	15	10	0	0
unb	62	35	0	1	8	12	0	0
laurentian	33	16	0	0	9	4	0	0
stfx	32	12	0	0	0	1	0	0
SMUHalifax	24	17	0	0	21	9	0	0
nipissingu	13	8	0	0	3	4	0	0
UPEI	12	10	0	0	1	3	0	0
stthomas	6	4	0	0	0	3	0	0
BishopUniversity	5	2	0	0	0	4	0	0
UNBC	3	5	0	0	15	10	0	0
mta	1	0	0	0	6	6	0	0
cbu	0	2	0	0	3	1	0	0
MSVU	0	0	0	0	0	1	0	0
uottawa	0	0	0	0	83	43	0	0
usherbrooke	0	0	0	0	0	2	0	0
Total	337585	58337	3764	447	384501	57305	871	124

Additionally, NMF assigns a *closeness score* for each document-topic pair, indicating how close the document is to a topic. To obtain more information about the topics produced by NMF, for each topic, we manually inspect 5% of the posts and comments with the highest closeness scores.

NMF requires the number of topics as input. Following previous work [15], we run NMF to produce between 5 and 50 topics and compute the *coherence* score for each. Coherence measures the extent to which the top representative terms representing each topic are semantically related (higher is better). We obtained the highest scores for 5 and 20 topics. A preliminary analysis of the NMF output at five topics revealed that most topics consisted of several discussion themes. This observation suggested that a larger number of topics may be more appropriate, and thus we selected 20 topics.

3. RESULTS

We begin with the results of our logistic regression analysis, shown in Table 4 in the Appendix. The most positive coefficients, predicting Fall 2020 posts and comments, include ‘chegg’ (an online platform for answering college and high school questions), as well as words related to online proctoring such as ’proctor’, ’proctorio’, ’zoom’, ’camera’, ’webcam’ and ’privacy’. The most negative coefficients, predicting Fall 2019 posts and comments, suggest in-person examinations (‘cheat sheet’, ’bring’, ‘sit’) and programming assignments and projects (‘code’, ’program’, ’project’).

Next, we move to topic modelling. Table 3 shows the NMF topic descriptors, the frequent n-grams, and the percentage of posts and comments assigned to each topic. We group the topics into the following three categories based on the information in Table 3 and manual inspection of a sample of posts and comments.

First, about 40% of the posts and comments include concerns about cheating leading to grade inflation, which in turn leads to assessments becoming more difficult. Students have observed grade inflation (Topic 13) and expressed concerns that Fall 2020 examinations will be more difficult to reduce the class average (Topics 1 and 20). Moreover, students commented on various methods used by

Table 3: **Fall 2020 topic modelling results**

#	Topic descriptors	Frequent N-grams	%
1	work, really, time, way, learn, try, hard, help, school, good	*'feel like', 'work hard', 'first year', 'high school', 'office hour', 'mental health', 'learn material', 'get catch', 'make sure', 'in person'*	10.4
2	say, academic, email, integrity, case, code, worry, report, flag, mean	*'academic integrity', 'academic dishonesty', 'integrity violation', 'academic integrity violation', 'get flag', 'student conduct', 'academic offense', 'would say', 'get catch', 'even though'*	10.3
3	think, probably, pretty, fine, worry, fair, sure, reason, away, good	*'think would', 'think people', 'think get', 'like think', 'get away', 'make sure', 'really think', 'feel like', 'think go', 'think make'*	6.4
4	student, university, honest, case, punish, international, chinese, issue, school, conduct	*'international student', 'student get', 'many student', 'chinese student', 'honest student', 'academic integrity', 'student would', 'mental health', 'academic dishonesty', 'first year'*	5.7
5	know, want, let, wrong, happen, person, tell, need, mean, consequence	*'let know', 'want know', 'know people', 'get catch', 'know would', 'lot people', 'feel like', 'know know', 'know go', 'student know'*	5.5
6	prof, email, mark, ta, ask, tell, send, chance, midterm, try	*'prof make', 'first year', 'email prof', 'open book', 'feel like', 'prof say', 'prof ta', 'prof would', 'make sure', 'ask prof'*	5.4
7	question, answer, time, ask, quiz, look, minute, similar, wrong, google	*'answer question', 'go back', 'multiple choice', 'short answer', 'exam question', 'one question', 'look answer', 'question answer', 'question exam', 'choice question'*	5.1
8	test, open, book, note, close, online, tab, internet, easy, search	*'open book', 'open note', 'make test', 'take test', 'test open', 'close book', 'book exam', 'open book exam', 'exam open', 'book test'*	4.9
9	people, lot, stop, say, agree, mean, proctor, probably, maybe, care	*'people get', 'lot people', 'many people', 'people would', 'get catch', 'people like', 'mental health', 'people go', 'know people', 'feel like'*	4.8
10	like, feel, sound, look, yeah, lol, bad, thing, lot, shit	*'feel like', 'seem like', 'look like', 'sound like', 'something like', 'even though', 'would like', 'make feel', 'online school', 'like people'*	4.8
11	exam, proctor, final, online, open, book, sheet, time, hour, note	*'take exam', 'final exam', 'open book', 'online exam', 'make exam', 'proctor exam', 'write exam', 'take home', 'home exam', 'person exam'*	4.7
12	use, software, proctor, proctorio, computer, browser, note, flag, lockdown, webcam	*'lockdown browser', 'secondary device', 'make sure', 'proctor software', 'take exam', 'get flag', 'student use', 'use respondus', 'virtual machine', 'use note'*	4.5
13	course, year, average, math, midterm, final, assignment, fail, term, quiz	*'first year', 'take course', 'last year', 'math course', 'feel like', 'midterm final', 'year course', 'course average', 'final exam', 'class average'*	4.5
14	class, curve, online, semester, average, fail, homework, lot, easy, problem	*'take class', 'class average', 'online class', 'class get', 'one class', 'feel like', 'math class', 'class take', 'in person', 'make sure'*	4.4
15	grade, curve, average, semester, high, final, letter, higher, better, good	*'good grade', 'letter grade', 'final grade', 'get good', 'get good grade', 'grade get', 'get grade', 'grade inflation', 'grade curve', 'better grade'*	4.2
16	professor, happen, try, evidence, accuse, report, tell, prove, probably, email	*'professor make', 'take exam', 'make exam', 'professor would', 'professor might', 'make sure', 'student professor', 'professor try', 'in person', 'tell professor'*	4
17	catch, happen, wonder, lol, hear, dumb, expel, time, lmao, guy	*'get catch', 'people get', 'people get catch', 'first time', 'catch people', 'catch get', 'use chegg', 'get away', 'without get', 'without get catch'*	3.7
18	chegg, post, account, use, ip, information, address, answer, view, solution	*'use chegg', 'ip address', 'chegg account', 'get catch', 'post chegg', 'question chegg', 'post question', 'chegg exam', 'chegg answer', 'answer chegg*	2.8
19	group, chat, leave, join, share, report, quiz, snitch, post, want	*'group chat', 'share answer', 'get trouble', 'group member', 'join group', 'class group', 'leave group', 'academic integrity', 'group project', 'study group'*	2.5
20	make, harder, sure, sense, hard, easier, difficult, mistake, thing, pretty	*'make sure', 'make harder', 'make exam', 'make sense', 'harder make', 'want make', 'make mistake', 'make difficult', 'make feel', 'want make sure*	1.4

instructors to combat cheating and reduce grades, such as grading on a curve (Topics 14 and 15) and using anti-cheating and online proctoring software (Topics 9 and 11).

Next, students reported feeling demotivated when they know that cheating happens in examinations (Topics 4 and 5) and often goes unpunished (Topics 3, 10 and 17). Students discussed examples of cheating that instructors failed to identify, such as seeking answers on Google and question-answering websites such as Chegg (Topics 7, 8 and 18), and discussing solutions in online chat groups (Topic 19).

Finally, students reported concerns about new methods used to prevent cheating in online examinations. They worried that some legitimate actions may be misconstrued as cheating: looking away from the computer screen, accidentally pressing a button, or disconnecting from a video meeting due to internet connectivity issues (Topics 6 and 12). Furthermore, some students reported being accused of cheating during online examinations, but did not realize they did anything wrong (Topics 2 and 16).

4. CONCLUSIONS

Logistic regression analysis suggests that cheating-related discussions on academic subreddits have expanded from plagiarism in computer programming (representative of Fall 2019) to online assessments in general. The word 'chegg' was associated with Fall 2020 content, suggesting an increase in the use of Chegg and related websites, which is consistent with prior work [6, 3]. Furthermore, words indicating online proctoring were predictive of Fall 2020 content, e.g., 'camera', 'webcam' and 'record'. Inspection of the posts and comments containing these terms revealed students' concerns about their privacy during online examinations. Similar concerns were raised in recent work [4, 9].

Topic modelling analysis identified three discussion themes in Fall 2020. First, students believe that cheating causes grade inflation, which motivates instructors to make assessments harder and introduce strict anti-cheating protocols such as not being able to scroll back to a previous question on an online examination. Some of these concerns have been highlighted in previous work [18, 19, 2, 4, 13, 3], and our analysis reflects students' opinions on this topic. Second, unpunished cheating lowers students' morale and motivation. Students report feeling demotivated when classmates cheat and obtain high grades. Third, students report not knowing exactly what constitutes cheating and what is allowed, underscoring the importance of clear academic integrity policies. These concerns were often reported in the context of online examinations, with students unsure of how their actions are being monitored.

5. REFERENCES

[1] S. Bagroy, P. Kumaraguru, and M. D. Choudhury. A social media based index of mental well-being in college campuses. In *CHI*, pages 1634–1646, 2017.

[2] E. Bilen and A. Matros. Online cheating amid COVID-19. *Journal of Economic Behavior & Organization*, 182:196–211, Feb. 2021.

[3] T. M. Clark, C. S. Callam, N. M. Paul, M. W. Stoltzfus, and D. Turner. Testing in the time of COVID-19: A sudden transition to unproctored online exams. *Journal of Chemical Education*, 97(9):3413–3417, July 2020.

[4] J. R. Deters, M. C. Paretti, and J. M. Case. How implicit assumptions about engineering impacted teaching and learning during COVID-19. *Advances in Engineering Education*, 8(4):1–5, 2020.

[5] B. Dorić, M. Blagojević, M. Papic, and N. Stanković. Students' attitudes regarding online learning during COVID-19 pandemic. In *Information Technology and Education Development*, pages 157–160, 2020.

[6] K. A. Gamage, E. K. de Silva, and N. Gunawardhana. Online delivery and assessment during COVID-19: Safeguarding academic integrity. *Education Sciences*, 10(11):301, Oct. 2020.

[7] N. Gozzi, M. Tizzani, M. Starnini, F. Ciulla, D. Paolotti, A. Panisson, and N. Perra. Collective response to media coverage of the COVID-19 pandemic on reddit and wikipedia: Mixed-methods analysis. *Journal of Medical Internet Research*, 22(10):e21597, Oct. 2020.

[8] C. Hutto and E. Gilbert. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *ICWSM*, pages 216–225, 2015.

[9] M. V. Jamieson. Keeping a learning community and academic integrity intact after a mid-term shift to online learning in chemical engineering design during the COVID-19 pandemic. *Journal of Chemical Education*, 97(9):2768–2772, Aug. 2020.

[10] A. Khan and L. Golab. Reddit mining to understand gendered movements. In *Proc. EDBT Workshop on Data Analytics Solutions for Real-Life Applications*, pages 3:1–3:8, 2020.

[11] H. Liu, Q. Li, R. Yao, and D. D. Zeng. Analyzing topics of JUUL discussions on social media using a semantics-assisted NMF model. In *ISI*, pages 212–214, 2019.

[12] D. M. Low, L. Rumker, T. Talkar, J. Torous, G. Cecchi, and S. S. Ghosh. Natural language processing reveals vulnerable mental health support groups and heightened health anxiety on reddit during COVID-19: Observational study. *Journal of Medical Internet Research*, 22(10):e22635, Oct. 2020.

[13] C. K. C. Ng. Evaluation of academic integrity of online open book assessments implemented in an undergraduate medical radiation science course during COVID-19 pandemic. *Journal of Medical Imaging and Radiation Sciences*, 51(4):610–616, Dec. 2020.

[14] A. Nobles, C. Dreisbach, J. Keim-Malpass, and L. Barnes. Is this an STD? please help!: Online information seeking for sexually transmitted diseases on reddit. *ICWSM*, pages 660–663, 2018.

[15] D. O'callaghan, D. Greene, J. Carthy, and P. Cunningham. An analysis of the coherence of descriptors in topic modeling. *Expert Systems with Applications*, 42(13):5645–5657, 2015.

[16] M. Parsa and L. Golab. Social media mining to understand the impact of co-operative education on mental health. In *EDM*, pages 653–657, 2020.

[17] J. W. Pennebaker, R. L. Boyd, K. Jordan, and K. Blackburn. The development and psychometric properties of LIWC 2015. Technical report, 2015.

[18] D. M. Telles-Langdon. Transitioning university courses online in response to COVID-19. *Journal of Teaching and Learning*, 14(1):108–119, May 2020.

[19] N. A. A. Tuah. Is online assessment in higher education institutions during COVID-19 pandemic reliable? *Siriraj Medical Journal*, 73(1):61–68, Dec. 2020.

[20] W. Xu, X. Liu, and Y. Gong. Document clustering based on non-negative matrix factorization. In *SIGIR*, pages 267–273, 2003.

APPENDIX

Table 4: Words with the most positive and most negative logistic regression coefficients

Term	coefficient	Term	coefficient
chegg	2.19	sheet	-3
online	1.79	cheat sheet	-2.95
proctor	1.79	code	-1.87
open	1.62	project	-1.68
covid	1.55	plagiarism	-1.51
zoom	1.45	phone	-1.47
prof	1.37	plagiarize	-1.32
pandemic	1.25	relationship	-1.31
proctorio	1.11	sit	-1.1
flag	1.09	talk	-1.02
cheat	1.08	sexual	-0.98
chat	1.06	notice	-0.94
camera	1.03	bring	-0.93
internet	1	textbook	-0.93
privacy	1	international	-0.92
book	1	misconduct	-0.78
cheater	0.98	appeal	-0.78
webcam	0.95	program	-0.79
100	0.93	go	-0.79
format	0.92	front	-0.81
screen	0.9	report	-0.81
open book	0.89	try cheat	-0.81
sem	0.88	ask	-0.81
record	0.88	homework	-0.82
math	0.88	dean	-0.82
term	0.87	practice	-0.83
average	0.86	allow	-0.88
respondus	0.85	partner	-0.88
email	0.83	final	-0.89
semester	0.83	english	-0.9