Hello? Who is posting, who is answering, and who is succeeding in Massive Open Online Courses

J. Miguel L. Andres-Bray
University of Pennsylvania
3700 Walnut Street
Philadelphia, PA 19103
miggyandresbray@gmail.com

Jaclyn L. Ocumpaugh
University of Pennsylvania
3700 Walnut Street
Philadelphia, PA 19103
jlocumpaugh@gmail.com

Ryan S. Baker
University of Pennsylvania
3700 Walnut Street
Philadelphia, PA 19103
ryanshaunbaker@gmail.com

ABSTRACT

Research has shown that participating in online discussion forums is tied to improved learning outcomes, but in the case of Massive Online Open Courses (MOOCs), very few students utilize this resource. This paper presents a preliminary investigation of demographic differences in the relationship between participation (using four different measures related to students production in the forum and the responses they receive) and course completion, demonstrating the need of further research into the potential causes for these differences.

Keywords

MOOC, MORF, demographic analysis, gender, race, production rules

1. INTRODUCTION

As educators seek to better support students in Massive Open Online Courses (MOOCs), research has shown that participation in MOOC discussion forums is tied to completion rates as well as other measures of learning [4, 9, 21, 26]. However, within the same body of research, we find that the vast majority of students never post even once [20].

These low rates of participation are perhaps not surprising. In addition to factors that can make it difficult for, say, a working adult to commit to a course [13], issues surrounding the social practices of both learning and language make the sheer size of a MOOC discussion forum difficult for people to navigate [2, 19]. Some researchers, concerned about the importance of having a social presence (see [22]) in online learning, have even suggested that we try to facilitate the community structure by encouraging MOOC participants enroll with a cohort of friends (e.g., [6]).

In addition to the social factors that influence motivation, encouraging students to enroll with friends may also improve students' opportunities to learn by giving them the opportunity to speak (or write) with people who use familiar communication practices. For example, we know that cultures differ in terms of several language practices that are likely relevant to learning contexts, including appropriate manners for giving advice [29], request strategies [30], expressing disagreement [31], including outsiders in the conversation [14] and even the interpretation of silence [10]. We also know that these differences can lead to cultural differences in broader interpretations of politeness (e.g., [7, 23]), and that sometimes these differences are even found across dialects of the same language (e.g., [18]) or across demographic groups, especially gender, within the same dialect (e.g., [17]).

Research on demographics in MOOCs has found that they can influence the level of student participation. For example, Huang et al. [11] found different rates of postings based on nationality and gender, with men posting more frequently than women and students from some countries posting at much higher rates than those from other countries. Likewise, Hodgson & Hui [12] also found differences based on nationality, with Chinese students posting at higher rates than those from other countries.

This study builds on these previous findings, exploring the role of demographic differences (namely race and gender) within the MOOC Replication Framework (MORF) [5]. Specifically, we are interested in two questions: 1) How do previous findings on forum posting behaviors and their relationship to course completion replicate across multiple MOOC sessions when broken down by reported gender and race? 2) Are these replications significantly different by race and gender?

2. METHODS

This research was conducted within the MOOC Replication Framework (MORF) [5], a platform designed to enhance large-scale replication efforts by representing previously published findings as production rules. In order to accomplish this, MORF uses a simple formalism that was previously employed in work to develop human-understandable computational theory in psychology and education [3, 16]. This approach allows findings to be represented in a fashion that human researchers and practitioners can easily understand.

All findings are converted into if-else production rules following the format, "If a student who is <attribute> does <operator>, then <outcome>." Attributes are pieces of information about a student, such as gender or race. Operators are actions a student does within the MOOC. Outcomes can represent a number of indicators of student success or failure including watching a majority of videos (e.g., [15, 24]) or publishing a scientific paper after participating in the MOOC (e.g., [26]). In the current study, we focus on the platform's most commonly-studied student outcome: whether or not the students in question completed the MOOC.

Each production rule analysis returns two counts: 1) the confidence [1], or the number of participants who fit the rule, i.e., meet both the if and the then statements, and 2) the conviction [8], the production rule's counterfactual, i.e., the number of participants who match the rule's then statement but not the rule's if statement. For example, in the production rule, "If a student started more threads than the session average, then they are more likely to complete the MOOC," the two counts returned are the number of participants that started more threads than the average and completed the MOOC, and the number of participants who started threads less than the average, but still completed the MOOC. As a result, for each MOOC, a confidence and a conviction for each production rule can be generated.

A chi-square test of independence can then be calculated comparing each confidence to each conviction. The chi-square test can determine whether the two values are significantly different from each other, and in doing so, determine whether the production rule or its counterfactual significantly generalized to the data set.

In this study, we tested the replicability of four previously published findings on discussion forum posting behavior, which can be put into two categories: initiating discussion and receiving uptake of discussion. The previously published findings can be found in Table 1 as production rules. These production rules use normalized measures of students forum activity (i.e., higher or lower than average) to investigate who is responsible for starting the most threads, posts more often overall, who gets more respondents on their own threads, and who posts more responses to others' threads.

Table 1. Categories of production rules used in the study

Condition	Source
Initiating Discussion:	
• If the student started more threads than the session average (Threads)	[4]
• If the student posts more frequently than the session average (Posts)	[9, 28]
Receiving Uptake of Discussion:	
• If the student has more respondents on their own threads than the session average (Respondents)	[21]
• If the student has more responses than the session average (Responses)	[27]

Note. The THEN clause of each production rule states, "then they are more likely to complete the course and earn a certificate."

These findings were turned into production rules, and executed in MORF against its data store of 100 MOOC sessions, comprised of 45 different courses offered by a University on Coursera. In integrating across MOOCs, we choose the conservative and straightforward Stouffer's [25] Z-score method to combine the results per finding across the multiple MOOC data sets, which we used to obtain a single statistical significance result across all MOOCs per reported gender and race.

Stouffer's Z-score method was used to compare resulting Z-scores across both genders and across the four races included in the study: White or Caucasian, Black or African American, Asian, and LatinX. In this study, we included only students who had reported their gender or their race in Coursra's optional demographics survey. Furthermore, only students from the U.S. were included in order to avoid complications of racial/ethnic categories that do not always translate in a straightforward manner across national boundaries. Table 2 summarizes the number of students in each category.

Table 2. Summary of research subjects by race and gender categories; all included reported being U.S. students

Demographic	N
Female	10,813
Male	14,394
White or Caucasian (W)	12,802
Black or African American (B)	1,089
Asian (A)	5,299
LatinX (L)	3,960

3. RESULTS

3.1 Gender Differences

The relationship between gender and the four discussion forum conditions included in this study can be found in Table 3. Each row represents the result of testing that condition, or MORF production rule, across the full set of MOOCs, reported by gender. For example, "Posts" refers to the production rule, "If the student's total number of posts was higher than the session average, then they are more likely to complete the course and earn a completion certificate." The *Female* and *Male* columns list the cumulative Z-scores per reported gender, and can be interpreted as how well each rule replicated across MORF's data. The *Diff* column lists the resulting Z-score when comparing whether the *Female* and *Male* scores are significantly different.

Table 3. Results of production rule analysis and gender differences in posting behavior (* $p < 0.001$)

Feature	Female	Male	Diff
Threads	16.36*	26.57*	7.22*
Posts	19.66*	29.89*	7.23*
Respondents	15.19*	21.69*	4.60*
Responses	14.60*	19.29*	3.32*

As seen in Table 3, all four production rules replicated significantly across all MOOC sessions in both genders. That is, thread, posts, respondents, and responses conditions all show positive relationships with MOOC completion. Moreover, as the *Diff* column shows, the positive relationship between forum participation and course completion is significantly more likely to replicate on men than women across all four measures.

3.2 Racial Differences

The relationship between race and the four discussion forum conditions included in this study can be found in Tables 4 and 5. In Table 4, each row represents the result of testing each production rule across the full set of MOOCs, broken down by reported race. As in Table 3, the *Feature* column lists the important key words of the rule's IF statement, and the other columns list the cumulative Z-scores per reported race. Like in Table 3, these values can be interpreted as significance values of how well each rule replicated across MORF's datasets. Table 4 shows that all four production rules replicated significantly across all MOOC sessions in all reported races included in this study.

Table 4. Results of production rule analysis broken down by reported race (* *p* < 0.01, ** *p* < 0.001)

Feature	White	Black	Asian	LatinX
Threads	19.57**	4.58**	8.67**	8.64**
Posts	22.78**	3.59**	9.42**	8.58**
Respondents	14.73**	3.12*	8.18**	7.11**
Responses	14.31**	4.27**	8.11**	6.31**

Table 5 reports the differences by race in the resulting Z-score. Here, it is important to note that the racial category listed first is the category with the higher Z-score value. As the table shows, nearly all of the comparisons were statistically significant, with notable exceptions found primarily in the comparison between Asian and LatinX students (the final column), although the difference between LatinX students and Black/African American students also had one condition (responses) that was not significant.

Table 5. Racial differences in posting behavior; W = White or Caucasian, B = Black or African American, A = Asian, L = LatinX (* *p* < 0.01, ** *p* < 0.001)

Feature	W > B	W > A	W > L	L > B	A > B	A > L
Threads	10.6**	7.7**	7.7**	2.9*	2.9*	0.0
Posts	13.6**	9.5**	10.**	3.5**	4.1**	0.6
Respondents	8.2**	4.6**	5.4**	2.8*	3.6**	0.8
Responses	7.1**	4.4**	5.7**	1.4	2.7*	1.3

4. DISCUSSION & CONCLUSIONS

In order to interpret these results, it is useful to consider the semantics of these conditions, two of which (posts and threads started) look at the initiation of discussion, measuring a student's production of discussion posts and two of which (respondents and responses) measure the uptake a student receives from those in his or her cohort. Our results show all four production rules significantly replicated across all MOOC datasets, and these findings are consistent no matter the demographic category considered.

At the same time, these results show that there are demographic differences both in who chooses to post and in whose posts receive the most attention. Namely, men are more likely than women to complete the course and earn a certificate based on initiating discussion and receiving uptake, results which may, in fact, influence one another.

We found the same kind of discrepancies when we looked at racial categories. Specifically, White/Caucasian students were more likely to succeed in MOOCs based on posting and receiving uptake than any other racial category. However, while Asian or LatinX students were more likely to succeed than Black/African American students based on forum behavior, the magnitudes of these differences were smaller than those between White/Caucasian students and other categories. Finally, Black/African American students were the least likely to succeed across all four measures.

More research is needed to better understand what might be causing these communication differences, but this preliminary study highlights the importance of examining these differences, as it shows they are important factors for predicting student success in MOOC platforms. If demographic differences exist at this level of analyses, it is likely that sociocultural strategies related to politeness or other communication practices are at play.

In future work, we hope to offer more sophisticated analyses of these issues, explicating the specific practices that may be contributing to these disparities. However, in the meantime, we hope this preliminary work serves to demonstrate a potential importance of these considerations when trying to support online learning.

5. REFERENCES

[1] Agrawal, R., Imielinski, T., & Swami, A. (1993). Mining Associations between Sets of Items in Massive Databases. In Proceedings of the ACM-SIGMOD Int'l Conference on Management of Data (pp. 207-216).

[2] Anderson, T., & Kanuka, H. (1998). Online social interchange, discord, and knowledge construction.

[3] Anderson, J. R., Matessa, M., & Lebiere, C. (1997). ACT-R: A theory of higher level cognition and its relation to visual attention. Human-Computer Interaction, 12(4), 439-462.

[4] Andres, J.M.L., Baker, R.S., Gašević, D., Siemens, G., Crossley, S.A. and Joksimović, S. (2018), March. Studying MOOC completion at scale using the MOOC replication framework. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (pp. 71-78). ACM.

[5] Andres, J. M. L., Baker, R. S., Siemens, G., Gašević, D., & Spann, C. A. (2016). Replicating 21 findings on student success in online learning. *Technology, Instruction, Cognition, and Learning*, 313-333.

[6] Bayeck, R. (2016). Exploratory study of MOOC learners' demographics and motivation: The case of students involved in groups. Open Praxis, 8(3), 223-233.

[7] Beebe, L. M., & Takahashi, T. (1989). Sociolinguistic variation in face-threatening speech acts. In The dynamic interlanguage(pp. 199-218). Springer, Boston, MA.

[8] Brin, S., Motwani, R., Ullman, J. D., & Tsur, S. (1997). Dynamic itemset counting and implication rules for market basket data. In ACM SIGMOD Record (Vol. 26, No. 2, pp. 255-264). ACM.

[9] Crossley, S., McNamara, D. S., Baker, R., Wang, Y., Paquette, L., Barnes, T., & Bergner, Y. (2015). Language to Completion: Success in an Educational Data Mining Massive Open Online Class. International Educational Data Mining Society.

[10] Duff, P. A. (2007). Problematising academic discourse socialisation. Learning discourses and the discourses of learning, 1(1), 1-18.

[11] Huang, J., Dasgupta, A., Ghosh, A., Manning, J., & Sanders, M. (2014). Superposter behavior in MOOC forums. In Proceedings of the first ACM conference on Learning@ scale conference (pp. 117-126). ACM.

[12] Hodgson, P., & Hui, B. (2017). The Social Presence of Active Learners across Five Countries through MOOCs: From Single to Super Postings.

[13] Kahan, T., Soffer, T., & Nachmias, R. (2017). Types of Participant Behavior in a Massive Open Online Course. The International Review of Research in Open and Distributed Learning, 18(6).

[14] Kitao, K. (1990). A Study of Japanese and American Perceptions of Politeness in Requests. Doshisha Studies in English, 50, 178-210.

[15] Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., & Miller, R. C. (2014). Understanding in-video dropouts and interaction peaks in online lecture videos. In Proceedings of the first ACM conference on Learning@ scale conference (pp. 31-40). ACM.

[16] Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). Soar: An architecture for general intelligence. Artificial intelligence, 33(1), 1-64.

[17] Mills, S. (2016). Gender and Performance Anxiety. Speaking Out: The Female Voice in Public Contexts, 61.

[18] Murphy, M. L., & De Felice, R. (2017). Routine politeness in American and British English requests: use and non-use of please. Journal of Politeness Research.

[19] Pfister, H. R., & Oehl, M. (2009). The impact of goal focus, task type and group size on synchronous net-based collaborative learning discourses. Journal of Computer Assisted Learning, 25(2), 161-176.

[20] Poquet, O, & Dawson, S. 2016. Untangling MOOC Learner Networks. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, New York, NY, USA, pp. 208–212

[21] Ramesh, A., Goldwasser, D., Huang, B., Daumé III, H. & Getoor, L. (2013, December). Modeling learner engagement in MOOCs using probabilistic soft logic. In NIPS Workshop on Data Driven Education (Vol. 2, pp. 1-7).

[22] Shelton, B. E., Hung, J. L., & Lowenthal, P. R. (2017). Predicting student success by modeling student interaction in asynchronous online courses. Distance Education, 38(1), 59-69.

[23] Sifianou, M., & Blitvich, G. C. (2017). (Im) politeness and Cultural Variation. In The Palgrave handbook of linguistic (im) politeness (pp. 571-599). Palgrave Macmillan, London.

[24] Sinha, T., Jermann, P., Li, N., & Dillenbourg, P. (2014). Your click decides your fate: Inferring information processing and attrition behavior from MOOC video clickstream interactions. arXiv preprint arXiv:1407.7131.

[25] Stouffer, S.A., Suchman, E.A., DeVinney, L.C., Star, S.A. & Williams, R.M. Jr. (1949). The American Soldier, Vol. 1: Adjustment during Army Life. Princeton University Press, Princeton.

[26] Wang, Y. Baker, R. (2015) Content or Platform: Why do students complete MOOCs? MERLOT Journal of Online Learning and Teaching, 11 (1), 17-30.

[27] Yang, D., Sinha, T., Adamson, D., & Rose, C. P. (2013). Turn on, tune in, drop out: Anticipating student dropouts in massive open online courses. In Proceedings of the 2013 NIPS Data-driven education Workshop (Vol. 11, p. 14)

[28] Yang, D., Wen, M., Howley, I., Kraut, R., & Rose, C. (2015). Exploring the effect of confusion in discussion forums of massive open online courses. In Proceedings of the Second (2015) ACM Conference on Learning@ Scale (pp. 121-130). ACM.

[29] Chun, M. Y. (2005). Culture-specific Concept of Politeness in Offering Advice. proceedings of the 10th Pan-Pasific Association of Applied Linguistics, Japan. PAAL Journal, 287-296.

[30] Veysi, E., & Abbaszadeh, F. (2016). A Comparative Study on Intonational Elements and Commissive Illocutionary Force Interface with Respect to Contextual Aspects and Cultural Background. Modern Journal of Language Teaching Methods, 6(4), 387.

[31] Zhu, W., & Boxer, D. (2012). Disagreement and sociolinguistic variables. Pragmatic variation in first and second language contexts: Methodological issues, 31, 113.