Social Comparison in MOOCs: Perceived SES, Opinion, and Message Formality

Heeryung Choi
University of Michigan
Ann Arbor, Michigan
heeryung@umich.edu

Christopher Brooks
University of Michigan
Ann Arbor, Michigan
brooksch@umich.edu

Nia Dowell
University of Michigan
Ann Arbor, Michigan
ndowell@umich.edu

Stephanie Teasley
University of Michigan
Ann Arbor, Michigan
steasley@umich.edu

ABSTRACT

There has been limited research on how perceptions of socioeconomic status (SES) and opinion difference could influence peer feedback in Massive Open Online Courses (MOOCs). Using social comparison theory [12], we investigated the influence of ability and opinion-related factors on peer feedback text in a data science MOOC. Perceived SES of peers and the formality of written responses were used as the ability-related factor, while agreement between learners represented the opinion-related factor. We focused on understanding the behaviors of those learners who are most prevalent in MOOCs; those from high socioeconomic countries. Through two studies, we found a strong and repeated influence of agreement on affect and formality in feedback to peers. While a mediation effect of perceived SES was found, a significant effect of formality was not. This work contributes to an understanding of how social comparison theory can be operationalized in online peer writing environments.

CCS CONCEPTS

• **Applied computing** → **Distance learning**; **E-learning**; **Collaborative learning**;

KEYWORDS

Social comparison, peer feedback, MOOC

ACM Reference Format:

Heeryung Choi, Nia Dowell, Christopher Brooks, and Stephanie Teasley. 2019. Social Comparison in MOOCs: Perceived SES, Opinion, and Message Formality. In *The 9th International Learning Analytics Knowledge Conference (LAK19)*, March 4–8, 2019, Tempe, AZ, USA. ACM, New York, NY, USA, Article 4, 10 pages. <https://doi.org/10.1145/3303772.3303773>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

LAK19, March 4–8, 2019, Tempe, AZ, USA

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6256-6/19/03...\$15.00

<https://doi.org/10.1145/3303772.3303773>

1 INTRODUCTION

In large online courses, such as Massive Open Online Courses (MOOCs), peer assessment is often used as a practical solution to the issue of grading assignments of thousands of students, especially when those assignments are too complex to be automatically graded [18]. Peer feedback, which is a type of peer assessment, has been adopted often in MOOCs due to its immediacy and opportunities for diversity of feedback, compared to teacher feedback [33, 35]. To maximize the benefits of peer feedback activity, researchers have made an effort to scale peer feedback such as focusing on features such as the huge size of classes and diversity of students [40]. Peer assessment is inherently as social and collaborative process that is subject to conscious and unconscious biases [39]. To ensure a more effective and inclusive learning experience, there is a need to understand how differences among learners are perceived and subsequently influence peer assessment.

Social comparison theory [12] claims that people prefer to compare themselves with similar others, and suggests how diversity attributes including socioeconomic status (SES) or opinion on a certain issue may influence learner interactions with peers. For instance, if a learner is matched to a peer with a dissimilar SES or opinion, the learner is more likely to feel further psychological distance and try to reduce the distance by either changing themselves or changing peers. This feeling of distance and efforts to reduce it can affect learners' responses to their peers [8, 30]. However, these previous studies have been conducted in traditional classrooms whose population sizes were not as large as MOOCs and where learner demographics were more homogeneous than those found in MOOCs.

In this work, we investigated how student writing in peer feedback activities in MOOCs can be considered using a social comparison theoretical lens operationalized through text analytic methods. In particular, we focus on how MOOC learners from countries which are generally considered as high-SES countries interact with peers showing a similarity or difference in opinions and in performance during an online feedback activity. To reflect a real-world MOOC context, we conducted the peer feedback field experiments in a data science MOOC course, a four week long data science course offered on Coursera.

Through this approach, we find that high-SES students' peer interaction patterns change after they are exposed to peers' opinion and SES. Our results provide novel evidence on how social comparisons occur in peer feedback activities of large scale online courses. This evidence contributes to a deepening understanding of how course instructors should design peer matching mechanism in global classrooms.

The subsequent sections of the paper are as follows: firstly, in section 2 we introduce previous studies on social comparison theory and studies investigating application of the theory in educational settings, and explain how we can identify social comparison behavior on MOOCs through natural language processing (NLP). Next, we elaborate how we expand study contexts from study 1 to study 2 with description of methodological features of the two studies. Finally, we conclude this paper with final discussion on results and general implication on MOOC design.

2 PREVIOUS WORK

2.1 Social Comparison Theory

Social comparison theory provides a useful theoretical framework to understand the potential relationship between learner diversity and peer feedback in large-scale online courses. One essential component of social comparison theory is that in the absence of objective criteria, people will choose to compare themselves with other people for self-evaluation. People conduct the comparison mainly on two components: opinions and abilities [12]. Research on social comparison demonstrates that both expectancy-related attributes, such as intensity of practice and performance on a particular subject, and ability-related attributes including gender, race, and SES are involved in interpersonal comparisons [1, 14, 30].

For instance, to study the influence of SES in an educational setting, Régner and Monteil [30] conducted an experiment with 110 students from two French middle schools, half of whom were classified as a low SES group and the rest of whom were classified as a high SES group based on data from the French National Institute of Economic and Statistical Information. During the experiment, after all the participants took a memory test, the researchers asked participants to indicate whom they would prefer to be compared with by numeric score. They found that a majority of students selected peers of similar SES, suggesting that SES can be considered as another ability factor of social comparison theory, and supporting Festinger's claim that an individual prefers to be compared with people with similar opinions and abilities [12].

Woods, Kurtz-Costes, and Rowley [41] also found evidence for SES as an ability factor; elementary school learners already considered other children with high SES academically better than other children with low SES. Cuddy et al. [5, 6] conducted studies across cultures and found that people with high SES and medium SES were thought to have competence while low SES were not. Interestingly, people with high SES are not only perceived more competent from others, but they also evaluate themselves highly due to their high self-esteem [13].

While these studies show evidence that learners would appreciate peers' SES as an ability factor, it is unclear how researchers can characterize behaviors resulted from students engaging in social comparison in particular settings. For example, if learners conduct

social comparison in a peer feedback activity, what types of learners behavior researchers can observe? In the next three sections, psychological distance will be introduced as one approach to understand social comparison behaviors. Then we will explain a particular educational setting and how we can identify the behaviors in the setting.

2.2 Psychological Distance

Perspectives on psychological distance offer insights on individuals' motivations for certain behavioral tendencies with respect to social comparison and feedback. Psychological distance is affected by internal and subjective experiences of objective distance [38]. During feedback activities in MOOCs, emotional intensity and formality of textual response are measurable behavioral proxies of psychological distance. Formality plays an important role in signaling social meaning, with low formality representing more interpersonal relationships, and closeness or lower psychological distance. [15, 32, 36]. For example, the informal language typically used between friends and family is associated with less psychological distance. Formal word choices could be an attempt to establish social distance, which is classified as tactics of polite avoidance rather than polite approach [25]. The similar finding was shown in an educational setting; further psychological distance between a sender and a recipient is expressed as higher formality of textual message in online classrooms [36]. Similarly, affect carries significant social meaning. For instance, emotional intensity is related to increased psychological distance in relationships [38]. In the educational context, this high formality and low emotional intensity in feedback can create discouraging learning experiences. Tu and McIsaac [36] demonstrated that higher formality of messages from peers made students feel less comfortable and more hesitant to respond which hampered learning.

Taken together, both social comparison and psychological distance suggest that people prefer to be compared with someone similar and, as such, psychologically closer to them. If learners are to be compared with peers who are psychologically distant (e.g. of dissimilar opinion), they will try to reduce the discrepancy, in this case, by persuading their peers. Therefore, detection of high formality and low affect can be interpreted as (a) greater psychological distance and (b) an effort to reduce the discrepancy through persuasion. That is, while feedback with high formality and low affect might hold more persuasive information which could be educationally beneficial, an extreme number of such feedback can harm learning experience of learners by discouraging them [15].

2.3 Peer Feedback in MOOCs

Typical classrooms, even with a low competitive and individualized atmosphere, are a site of frequent social comparison. This is because reward systems and peers' performance such as grades generate competitive atmosphere [8, 22, 27]. Previous work implies that MOOCs also offer an interesting context for social comparison. Davis et al. [7] is one of few studies showing the impact of social comparison in online courses. They designed a system offering personalized feedback message using progress information of successful peers from previous rounds of the same course. For

example, learners would get messages comparing their weekly forum contributions to those of their peers. Through the personalized social comparison feedback system, they found that achievement and engagement of learners increased as a result. The present study is different from Davis et al. [7] since this study is conducted in authentic peer feedback in consideration of precise social comparison factors, i.e., SES and opinion.

Peer feedback activities in MOOCs provide rich contexts for social comparison. Peer assessment where learners give and receive peer feedback is an inherently social process. Specifically, by providing and accepting feedback, and learning from each other, peer assessment with peer feedback can be viewed as a social process [39]. The peer feedback process should, therefore, be understood in the context of the elements that influence it, such as diversity in demographics, opinions, language, and learning behavior of MOOC learners who synchronously and asynchronously collaborate with other learners [16, 20, 21, 42]. Given the collaborative and social nature of peer feedback, it is important to study how interpersonal variables influence the collaborative learning activity such as how learners' responses change in the presence of various peer diversity attributes. These previous studies on social comparison theory, psychological distance, and peer feedback coupled with the amplified learner diversity in MOOCs motivate the following research question: In MOOCs, does formality and affect of peer feedback change based on the abilities and opinion of the peer?

There are various features in peer feedback activities on MOOCs which could shape social comparison differently compared to typical offline classrooms. MOOC courses typically have a large number of students and offer opportunities to work with peers who are different in terms of SES, writing skill, and opinion [16, 20, 21, 42]. That is, learners have more peers to compete with and compare themselves with, while they might have a harder time identifying similar peers who they would prefer to be compared with. Furthermore, features of computer-mediated communication (CMC) on MOOCs can influence intensity of social comparison. For example, a lack of nonverbal cues in CMC can lead MOOC learners to form impressions on peers heavily based on peers' social identities instead of peers' personal identities. This distinct way of forming impressions in CMC can give stronger motivation to learners for conducting social comparison based on social category indicators such as SES [34]. In this study, the formality of each learner's position is a personal ability factor while their SES is a social ability factor. Considering that learners will asynchronously interact with their peers with no nonverbal cues, will learners' impression about peers rely more on the perceived SES than a formality of a post? One contribution of this study is to reveal how social comparison would turn out in online education settings where nonverbal cues are barely present.

Another novel contribution of this study is how we match peers for feedback activity. Despite the importance of interpersonal variables in peer matching, peer feedback activities have generally been assigned using randomized peer matching without consideration of interpersonal variables [33]. Randomized peer matching may not be the ideal method for learners, considering that studies on peer assessment and diversity indicate that challenges for peer assessment lie in interpersonal variables, such as how assessors and assesses perceived each other and themselves [33, 39]. Studies on

how influential those interpersonal variables are on other learners' learning behavior in peer feedback activities could facilitate a potentially more beneficial student matching framework, rather than the random approach that is currently used in MOOCs.

2.4 Characterizing Social Comparison Behavior

Natural language processing (NLP) gives a way to identify social comparison-related behaviors from learners' responses to their peers. We used two NLP tools were adopted in this study: Coh-metrix [24], which is a text analysis tool that provides indices for linguistic and discourse characteristics and Linguistic Inquiry Word Count (LIWC) which is a computerized text analysis application with various language categories such as positive emotions and negative emotion [26]. We use these two tool kits to measure the following attributes:

2.4.1 Affect. Affect is a sum of scores of words representing positive and negative emotion. Higher affect is a proxy for emotional intensity in textual responses, which is a linguistic cue for less psychological distance [38]. We note here explicitly that we are considering the total amount of affect, and not the amount of positive affect alone. That is, critiques which are extremely negative (e.g. "I really despised how you used that ridiculous turn of phrase") would be considered high affect, even though the substance of the comment is affectively negative. This suggested NLP method was tried in previous related literature as an early pilot study [4].

2.4.2 Formality. The formality measure used in the current research was first introduced by Graesser et al. [15]. It is a composite measure derived from Coh-Metrix's five principal components, Narrativity, Deep Cohesion, Referential Cohesion, Syntax Complexity, and Word Concreteness. Specifically, this metric of formality increases with abstractness of words, syntactic complexity, cohesion (referential and deep), and the informational expository genre (as opposed to narrative). At the other end of the continuum, informal discourse tends to have concrete words, simple syntax, low cohesion (because knowledge-based inferences can fill the gaps), and high narrativity. We therefore computed a composite score of formality that integrated the five major dimensions of Coh-Metrix, wherein five dimensions were weighted equally [15]. This formality measure has been used to explore learners' social and cognitive processes in MOOCs [10, 11, 19], and has proven useful in building an understanding of individuals behavior in digital learning environments more broadly [9].

The degree of formality versus informality in discourse has been shown to play an important role in signaling socially significant meaning, such as authority, persuasion, politeness, solidarity, and interpersonal relationships closeness [2, 3, 32]. In the current study, formality was used as an ability-related factor and measure of psychological distance.

In summary, high formality and less affect can give more room for learning due to its elaborating, persuasive writing style, while

it can harm further interaction by increasing psychological distance. Therefore, extreme imbalance of such attributions of feedback should be managed both for more learning opportunities and interaction encouragement.

Taken altogether, in this work, we built hypotheses as follows: learners from high-SES countries will write less affective and more formal feedback when

- (1) H1. responding to cases of disagreement with peers compared to cases of agreement.
- (2) H2. responding to lower SES peers compared to when responding to the similar SES peers.
- (3) H3. responding to more formal peer position compared to less formal peer position.

We test these hypotheses through two studies: Study 1 was conducted through a controlled peer feedback activity and Study 2 in an authentic peer feedback activity. With the better understanding on how interpersonal variables affect learning behavior in peer feedback activities, MOOCs could be more equally beneficial to learners with different opinion and SES background.

3 STUDY 1

3.1 Methodology

3.1.1 Participants. Participants ($N = 456$) were recruited through an optional activity in the data science MOOC taught by Christopher Brooks. IRB oversight was sought under University of Michigan study ID HUM00120884.

3.1.2 Experimental Design. Study 1 was a $2 \times 2 \times 2$ experimental design (Figure 1). Learner's current country was obtained through their internet protocol (IP) address. Based on the country location data and the United Nations Development Programme Human Development Report [37], learners were classified as coming from high, medium, or low-SES countries. This classification was made based on whether each learner was in the first, second or third, or fourth quartiles of the Human Development Report respectively. In this work we focus on the actions of only those learners who were identified as coming from high-SES countries, since these learners are the majority of MOOC learners.

Learners started the activity by reading question prompts asking whether they agreed with the non-technical issues related to data science, and then completed a short writing position (Figure 2). The following question prompt was used, which focuses on ethical decision making based on the article by Poor and Davidson (2016):

Nathaniel Poor and Roei Davidson from the Data and Society Research Institute were faced with an ethical dilemma when the data they wanted to use for research but couldn't (because of logistics issues) was released by hackers. The data they were interested in was publicly available, but difficult to get at requiring an expensive manual process. The hacked dataset included this public data as well as private messages which they were not interested in. Read the case study. Do you agree with the author's decision to not use the hacked data, and state

any arguments they haven't considered for or against the use of this data? [29]

Next, learners were asked to leave feedback to three peer positions. Figure 3 shows the interface learners interacted with during the activities. Students were asked to write at least 90 words in their own response to the article, and they were allowed to write as many words as they wanted in response to a peer. In Figure 3 showing user interface of the peer feedback activity platform, the disclosure of country (a proxy for abilities based on perceived SES) is labeled as A, and a measurement of agreement to the given question (a proxy for opinion) measured as B in the written response. The actual text of the learner response is filled in in text area C, and is the main unit of analysis for all studies.

These positions each participant was assigned were comprised of one controlled peer feedback activity and two authentic peer activities. Among the three peer feedback activities, the data from the controlled activity was collected and analyzed in this study. This controlled activity was designed as a controlled setting. That is, instead of matching learners to other learners, learners were randomly shown one of four learner positions which was already written by participants from the pilot studies [4]. These learner positions represented each condition of agreement and disagreement, as well as those of high formality and low formality. The formalities of these responses were measured and z-scored as 1.23 and -0.23 (for the agreeing responses) and as 1.77 and -0.15 (for the disagreeing responses) respectively. The Study 1 was designed as a controlled setting to make a more clear observation on how the ability factors and opinion factor shape learners' interaction with peers.

Here are examples of the chosen learners' positions respectively with high formality and low formality:

High formality: Scientific research has a special status in a society because it creates and disseminates knowledge, thus providing informational externalities. Though there is no absolute guarantee that a larger knowledge base would necessarily benefit any society under every circumstance, I take for granted the empirical moral assumption common in the scientific community that publicly available data can be used, provided that no *additional* damage is originated. It is a researcher's responsibility to avoid any kind of dissemination of sensible personal data, but I can hardly see why, post factum, publicly available data cannot be employed for any serious scientific inquiry.

Low formality: This is ethics only for ethic's sake. Using the data wouldn't hurt anyone, certainly not the hackers nor the web site, but it could potentially help people with crowdfunding. Legitimizing criminal activity argument is silly, hackers don't care and would be punished if caught. Users privacy is already compromised. Researchers should anonymize data anyway even from ethically obtained data

Figure 1: Overview of the study design. Study 1 dataset was collected from the controlled peer feedback activities which are colored in this figure

Your Opinion

Nathaniel Poor and Roei Davidson from the Data and Society Research Institute were faced with an ethical dilemma when the data they wanted to use for research but couldn't (because of logistics issues) was released by hackers. The data they were interested in was publicly available, but difficult to get at requiring an expensive manual process. The hacked dataset included this public data as well as private messages which they were not interested in.

Read the case study [here](#).

Do you agree with the author's decision to not use the hacked data, and state any arguments they haven't considered for or against the use of this data?

Your thought?*

Explain / Describe (minimum = 90 words)*

Word count: 0

Submit

Figure 2: Interface for the student position prompt

Peer Feedback : Peer Feedback (Step 1 of 3)

A student from United States agrees with you.

I totally agree with the author's decision not to use the hacked data because it would intrude the privacy of many users even though the data was to be used for a research. Furthermore, the research could be done using other means and medium of data which would have been legal whereas using this data would set an example that might lead to increase hacking of personal data in the name of research. The research might just be very biased rather than generalizable because no one can confirm the validity of hacked data.

Do you find their opinion convincing?*

Yes Maybe No

Critique your peer's response. (Comments will be sent to peer anonymously.)*

Submit

Figure 3: Interface for the peer feedback prompt

which should be standard procedure. The argument of consent is silly anyway and reminds me of the Streisand effect, any party with malicious intentions could easily obtain it. The horse has already left the stable if your research is important to use the data if it isn't find something important to research.

The decision to use authentic learner positions was motivated by issues that were identified in the pilot studies [4]. Although the pilot studies explored a similar hypotheses, the studies did not show a clear conclusion. One possible reason for this lack of convergence could be the use of instructor-written positions instead positions written by learners. In the pilot studies, learners' positions ranged from a formality score of $-.57$ to 3.48 ($M = .69$, $sd = .55$), while the two instructors' positions had formality score of $.37$

and $.68$ ($M = .52$). There were two issues regarding the instructor-written positions. Firstly, there was difference in formality between the given instructor-written position. Considering distribution of formality of learners' positions, the difference should not be neglected. Secondly, there was a difference in formality between one of instructor-written positions and actual learner positions. That is, differences between formality of one instructor-written positions ($.37$) was comparably lower than the average formality of learner positions ($.69$). These differences could have caused the effect of perceived SES to be washed out, since higher formality could have been perceived as more persuasive to the educated learner populations [15, 28]. To address the issue, the next section replicates the design of the previous work, but the instructor-written positions were replaced with the selected actual learner positions to increase authenticity.

Figure 4: The structural model of mediation analysis with estimates and p-value (* $p < .05$, ** $p < .01$, * $p < .001$). The dependent variable in (a) is affect of feedback, while the dependent variable in (b) is formality of feedback.**

Table 1: Pearson Correlations in Study 1

-	SES	Agreement	Formality
SES	1.00	-	-
Agreement	-.17*	1.00	-
Formality	.03	.02	1.00

Note. Correlation $df=454$; * $p < .05$, ** $p < .01$, *** $p < .001$

3.2 Results

As a reminder, the principal independent variable is the country of origin of the peer (see Figure 3, Label A) which we are using as a proxy for perceived SES. We used two data cleansing approaches. First, one of the authors read all student responses and removed any that were deemed to be gaming the system or constituted invalid entries (e.g., repeated characters to fulfill a perceived word length equivalent). Secondly, as the primary focus was the relationship of agreement as a measurement of opinion, responses with "maybe" to the question "Do you find their opinion convincing?" were considered unclear signals of agreement and were removed from the dataset.

All of two text features, formality and affect, in the section were measured with all values converted into z-scores and rounded to the nearest hundredth for ease of comparison. A significant correlation was found between peer SES and agreement (Table 1). Therefore, a mediation analysis was conducted to see whether there were causal relationships where an independent variable influences a dependent variable through an intervening variable, the mediator. The impact of an independent variable on a dependent variable through a mediator is referred as the indirect effect, while the impact of an independent variable on a dependent variable without a mediator is called the direct effect [17]. Mediation analysis demonstrates indirect effect which could not be captured by regression analysis or ANOVA. In this study, the R package Lavaan was used for the mediation analyses [31]. Since the peer SES was displayed to

Table 2: ANOVA Test on Study 1 Results

	Formality Model		F-value
	Formal M (SD)	Informal M (SD)	
Affect	.03 (1.06)	-.03 (.94)	.31
Formality	-.05 (1.04)	.05 (.96)	1.16

Note. Correlation $df=452$; * $p < .05$, ** $p < .01$, *** $p < .001$

students but agreement was measured from students in the experiment setting, the mediation model was designed with peer SES as independent variable and agreement as a mediator (Figure 4).

In Figure 4, both affect in (a) and formality of feedback on (b) were dependent variables. Both peer SES and agreement were considered as categorical values. In both (a) and (b), peer SES was coded as low = 0 and high = 1 and agreement was coded as disagreement = 0 and agreement = 1.

H1 predicting the effect of disagreement on affect and formality was supported in this study. Regardless of the formality of the positions and SES levels, high-SES learners wrote significantly less emotional and more formal feedback to peers with whom they disagreed (agreement to affect: estimate .49, $p < .001$, agreement to formality of feedback: estimate -.32, $p < .001$). For H2 predicting an effect of peer SES on affect and formality, low-SES learners received more emotional and less formal writing considering that high-SES learners were more likely to agree more with low-SES learners (peer SES on agreement: estimate -.13, $p < .05$). This is contrary to H2 and partially opposite to findings from pilot study [4] - where it was revealed that low SES peers received significantly more formal writing. The study did not provide support for H3 (Table 2).

3.3 Discussion

The empirical findings in the Study 1 provide a new understanding of SES as an ability factor in a mediation model. The previous pilot study [4] proposed that peer SES might have been diminished due to a distinctive difference in the formality of the instructor-written position and decided to replace instructor-written position with authentic learner positions. The current study not only addressed the issue on gaps in formality scores but also showed a new mediation effect holding a significant effects of peer SES. Despite the statistically significant relationship, the correlations between peer perceived SES and dependent variables (affect and formality) was not consistent with the proposed H2; the relationship between peer perceived SES and dependent variables was revealed to be opposite to what was proposed. That is, high perceived SES learners are more likely to write *less* formal and *more* affective response to low perceived SES peers compared to high perceived SES peers.

With the results from the Study 1, the next study was designed to extend experimental environments by replacing the controlled dataset with the authentic peer interaction dataset. In the Study 1 formality was operationalized in the discrete manner as high and low. However it is entirely possible the phenomenon under this study could vary along a continuum of formality. Thus, in the

Figure 5: Study 2 dataset was collected from the authentic peer feedback activities which are colored in this figure.

Figure 6: The structural model of mediation analysis with estimates and p-value for the Study 2 (* $p < .05$, ** $p < .01$, *** $p < .001$). The dependent variable in (a) is affect of feedback and the dependent variable in (b) is formality of feedback.

subsequent follow up Study 2, we explored H3 along continuous formality in student responses.

4 STUDY 2

4.1 Methodology

4.1.1 Participants. Participants for this study are the same participants involved in Study 1. However, in the Study 2, we focused on the two authentic peer feedback activity (Figure 5). Since each learner was matched to two authentic peers, 456 learners generated 912 feedback responses. Among 912 feedback responses to peers, 695 feedback responses were produced by high-SES learners who were matched with high SES peers. 205 feedback responses were generated by high-SES students who were matched with medium SES peers. Due to the small number, feedback responses to low-SES learners ($n = 6$) were dropped from the dataset for analysis. In total, 906 feedback responses were included in the dataset for the Study 2.

4.1.2 Experimental Design. Study 2 replicated a $2 \times 2 \times 2$ experimental design of the Study 1: peer SES, agreement, and peer

formality. A difference between Study 1 and Study 2 in experimental design lies in dataset; Study 2 gathered data from authentic peer activities while Study 1 collected data from controlled peer feedback.

4.2 Results

We cleaned data in the same way from the Study 1. There were significant Pearson correlations between (a) agreement and peer SES and (b) peer formality and peer SES (Table 3). As the next step of the significant Pearson correlation, a mediation analysis was conducted with the R package Lavaan [31]. Through the meditation analysis, we found a mediation effect of agreement on peer SES (Figure 6).

The model (a) and the model (b) of Figure 6 show agreement as an effective mediator. Both affect and formality of feedback, which were dependent variables, were z-scored. Peer SES and agreement were considered as categorical values. Peer SES was coded as medium = 1, and high = 2 and agreement was coded as disagreement = 0, and agreement = 1. There was no mediation effect between peer SES and peer position formality, since the effect of peer formality on affect and formality of feedback was not significant (agreement to affect: estimate -.06, $p > .05$, agreement to formality of feedback: estimate .07, $p > .05$).

H1 on the impact of agreement on affect and formality of feedback was significant (agreement to affect: estimate .57, $p < .01$, agreement to formality of feedback: estimate -.15, $p < .01$). The results support H1: disagreement lead to less affective and more formal feedback. Evidence supporting H2 was not directly shown in this study. Yet, considering that medium-SES learners were disagreed with more by high-SES learners, the significant results found are further evidence for H2. In terms of H3, formality did not show significant impact. There was no significant difference in formality of responses written by medium SES and high-SES learners (estimate -1.49, $p > .05$).

4.3 Discussion

The main contribution of Study 2 is evidence from authentic peer matching data which confirms both H1 and H2. In Study 1, high-SES

Table 3: Pearson Correlations in Study 2

-	SES	Agreement	Formality
SES	1.00	-	-
Agreement	.09**	1.00	-
Formality	-.08*	-.06	1.00

Note. Correlation $df=904$; * $p<.05$, ** $p<.01$, *** $p<.001$

learners were more likely to agree with low-SES learners, which is opposite to H1. Yet, Study 2 demonstrated that high-SES students were more likely to disagree with medium-SES students. Taken altogether, these results from Study 1 and Study 2 suggest that high-SES learners react differently to peers from medium SES vs. peers from low SES. The psychological distance between high-SES learners and medium-SES learners may make peer feedback experience uncomfortable for lower SES students.

5 DISCUSSION AND LIMITATIONS

This work has explored the influence of social comparison on peer feedback in scaled online learning environments. Social comparison theory aims to explain student comparison of themselves with others, using ability-related factors and opinion-related factors as two comparison dimensions. Based on social comparison theory and previous studies, hypotheses were proposed predicting how high-SES learners' responses would change based on the effect of an opinion-related factor, agreement and ability-related factors, SES and formality: learners from high-SES countries will write less affective and more formal feedback when they (H1) respond to peers they disagree with, (H2) respond to peers from lower SES countries, and (H3) respond to peers where the peer holds a more formal position (Figure 6).

We investigated these hypotheses through two studies. Study 1 was conducted in the semi-controlled setting where learners were exposed to four researcher-chosen peer positions varying in agreement (agree vs. disagree), and linguistic formality (low vs. high). Study 2 replicated the design of the Study 1 but did so in an authentic peer interaction environment, showing actual student writing.

Agreement with peers' opinions was found to significantly influence how learners constructed their written responses, which confirms the role of agreement as an opinion factor in the context of MOOC peer feedback activity. In Study 1, learners provided less affective and more formal feedback when they found that they disagree with instructor positions (H1), and this was replicated in the Study 2. Another contribution is that agreement was revealed to be a mediator between SES and response features which are formality and affect. Traditional social comparison theory suggests that agreement of opinion and ability are two primary dimensions of social comparison. But this work shows that agreement of opinion operates as mediator between the ability-factor (in this case perceived SES) and the social comparison outputs (affect and formality). Thus a contribution of this work is a contextualization of how opinion and perceived ability work within social comparison in peer feedback activities of global classrooms.

Results from study 1 and study 2 revealed the differential relationship in perceived SES (H2). Specifically, high-SES learners agreed more with low-SES learners and thus tend to write more affective and less formal feedback, while they were more likely to disagree with medium-SES learners which caused less affective and more formal feedback. These findings suggest that higher SES students experience greater psychological distance from medium-SES students, a surprising result. Considering that text with higher formality requires more effort and is more persuasive [15], such high formality coupled with lower affective expression could be interpreted as higher SES students' effort to reduce the psychological distance with medium-SES students by changing their peers' opinion.

One explanation for why high-SES learners treated medium SES peers and low SES peers differently is that high-SES learners' expectation of medium-SES students may be higher than that of low-SES learners. Further, high-SES students become more open and accommodating to peers from the same SES group, while they may have higher expectations for the quality of responses from high SES peers than on the low SES peers. In addition, high-SES learners might have perceived medium-SES learners as familiar enough to build more concrete level of expectation, but at a (psychological) distance far enough to be less preferable to be compared with. This explanation fits with previous studies [5, 6] showing that people attribute higher competence to medium-SES learners and high-SES learners but not low-SES learners.

The interpersonal power and behavior model [23] also supports this finding and might be another explanation. This model claims that people with higher authority, power, or status are perceived as being more assertive, less pliable, and less persuadable. Therefore, high-SES learners would expect it to be more difficult to persuade their peers when they were matched with medium-SES learners in contrast to low-SES learners. This differential impression of interpersonal power attributed to peers' SES might cause high-SES learners to feel the need to assert their authority by increasing psychological distance. Since interaction in this study was text-based, psychological distance was demonstrated through less affect and more formality. In short, perceived SES as a proxy for an ability-related factor seemed to influence the high-SES learners' expectations in a non-linear fashion.

The studies did not show an impact of formality of peer comments on how high-SES learners wrote feedback. A possible explanation is that learners did not pay much attention to peers' formality, which is an ability factor based on a personal identity indicator. As discussed earlier, because of a lack of nonverbal cues in the online environment, social identity indicators such as SES may influence peers more strongly than personal identity indicators such as writing ability do. Although social comparison theory does not comment on a hierarchy among different ability factors, online learning environments may lead learners to impose a greater importance on SES than formality which is another ability factor. This result would be best tested in more controlled settings, comparing online and traditional peer feedback activities with respect to perceived SES.

Another explanation for non-significant effects could be the lack of English comprehension skills among learners leading to a hard time differentiating well-written instructor responses and

poorly-written peer responses. Indeed, there were many comments from learners to this effect: “Please take into account that it is not always easy to write in English for a lot of students.” Thus, a natural progression of our work would be to include a measurement of English comprehension skill (e.g. from an admissions survey or self-declaration) or expand the experiment to a more controlled setting.

The findings of the current study imply that high SES learners intentionally, or unintentionally, write feedback which reduces learning opportunities of low SES peers, or leads medium SES peers to be hesitant in peer interaction. This is a significant issue for MOOCs, where learners are sometimes shown indicators of peers’ countries of origin and may be able to use (rightly or wrongly) these as an indicator of SES. Learners from high SES countries make up the majority of learners in MOOCs, and may harm peer learning experience of medium SES and low SES peers, while the negative effect does not influence other high SES peers.

Instructors and course designers can alleviate this imbalance in peer learning experiences. Firstly, they can review which information of learners to be visible to other learners. For instance, particular information such as names or profile pictures might be cues with respect to peer SES and might lead learners to shape biased feedback. Instructor teams and course designers can hide this information during peer feedback. Secondly, course designers can adjust the population of peers giving feedback to a learner by avoiding using random matching. Through careful matching methods, one can ensure that medium and low SES learners get exposed to a breadth of their peers, and thus potentially remediate extremely biased feedback from the majority group.

The study presented was not designed to evaluate whether there was an ordering effect of peer positions and how that might influence how students construct peer feedback. For example, if a student was exposed to well-written positions twice in a row, then it is possible that they would judge the third position they see against these high standards. By contrast, if acceptable but not particularly well-written positions were shown first, the critique to a third position might be different just because of the ordering. As assignments were randomized, we expect that any ordering effects would be evenly distributed, but we did not control nor measure the impact that such effects might have had.

In this study we grouped individuals from multiple countries into three different categories – high, medium, and low SES – based on the UNDP report [37]. An individuals’ specific SES within their country could be influential on how other learners react to them and how they react to others. For example, even though peers are from the same medium-SES country, a high SES learners might show different reaction to a high-income peer and an average-income peer. Further, we make no claims on the distribution of SES within the learner base of a single country in the Coursera system, either between or within a given country. Instead, our study relies on the UNDP HDI [37] as a proxy for perception of SES by others. A more nuanced understanding of how income equity and other biases related to specific countries and how these might affect learners’ behavior would be valuable in the experimental design of future work. For instance, it may not be that being from a low-SES country has as much of an affect as being from a country in a different region of the world (e.g. hemisphere) or a country with a different primary

language. Our work here looked specifically at perceptions of SES, and we expect that there are other dimensions of social comparison between MOOC learners which deal with stereotypes and biases, and that understanding these through peer feedback writing may be valuable.

Finally, it is important to acknowledge that we considered the issue of learner writing from the position of how learners from high-SES countries responded to their peers, and not how learners from lower SES countries might perceive high SES peers. This was due in part to the low (relatively) sampling rate of actual students from low-SES countries, a known challenge in MOOC research. Understanding how low-SES students react to social comparisons upward has important implications for creating inclusive learning environments in the global online classroom.

6 CONCLUSION

Through the lens of the social comparison theory, we investigated how perceived SES and opinion of peers affect learners’ feedback to peers in MOOC courses. We found evidence showing learners from low-SES countries were more likely to be agreed with by learners from high-SES countries, and thus received less formal and more affective peer feedback. Further, learners from medium-SES countries were more likely to be disagreed with by high-SES learners, and therefore were received more formal and less affective peer feedback. The difference in the results between the studies implies that high-SES learners hold different levels of expectations and openness when engaging with perceived low-SES learners as compared to when they engage with perceived medium-SES learners.

The results of this study suggest several design implications for peer learning activities in the global classroom. The current practice is for learners to be randomly matched to one another for feedback activities [33]. Our work has shown that matching learners based on their opinion (agree or disagree) and perceived SES will substantively change the feedback the learners produce as measured by emotion and formality. As current MOOC user interfaces may provide clues for SES such as learners’ name and profile pictures, instructors and course designers should consider explicitly the attributes of a learner which are used (and made available in the user interface) for peer critique and feedback activities so that responses to learners are most useful. For example, an extremely low affect coupled with highly formal feedback may harm smooth interaction and learning experiences in peer feedback activity, while the extremely high affect and low formal feedback can harm learning experiences also by not producing informative, persuasive feedback which may enhance learning opportunities. Therefore, it is necessary to carefully match learners considering perceived SES and opinions.

7 ACKNOWLEDGMENTS

This work was supported in part by the Michigan Institute of Data Science (MIDAS) learning analytics challenge initiative.

REFERENCES

- [1] Hart Blanton, Bram P Buunk, Frederick X Gibbons, and Hans Kuyper. 1999. When better-than-others compare upward: Choice of comparison and comparative

- evaluation as independent predictors of academic performance. *J. Pers. Soc. Psychol.* 76, 3 (March 1999), 420.
- [2] S E Brennan and H H Clark. 1996. Conceptual pacts and lexical choice in conversation. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 6 (Nov. 1996), 1482–1493.
 - [3] Penelope Brown and Stephen C Levinson. 1978. Universals in language usage: Politeness phenomena. In *Questions and politeness: Strategies in social interaction*. Cambridge University Press, 56–311.
 - [4] Heeryung Choi, Nia Dowell, and Christopher Brooks. 2018. Social Comparison Theory as Applied to MOOC Student Writing: Constructs for Opinion and Ability. In *13th International Conference of the Learning Sciences (ICLS)*, Judy Kay and Rosemary Luckin (Eds.). International Society of the Learning Sciences, 1421–1422.
 - [5] Amy J C Cuddy, Susan T Fiske, and Peter Glick. 2007. The BIAS map: behaviors from intergroup affect and stereotypes. *J. Pers. Soc. Psychol.* 92, 4 (April 2007), 631–648.
 - [6] Amy J C Cuddy, Susan T Fiske, Virginia S Y Kwan, Peter Glick, Stéphanie Demoulin, Jacques-Philippe Leyens, Michael Harris Bond, Jean-Claude Croizet, Naomi Ellemers, Ed Sleebos, and Others. 2009. Is the stereotype content model culture-bound? A cross-cultural comparison reveals systematic similarities and differences. (2009).
 - [7] Dan Davis, Ioana Jivet, René F Kizilcec, Guanliang Chen, Claudia Hauff, and Geert-Jan Houben. 2017. Follow the successful crowd: raising MOOC completion rates through social comparison at scale. In *LAK*. rene.kizilcec.com, 454–463.
 - [8] Pieterneel Dijkstra, Hans Kuypers, Greetje van der Werf, Abraham P Buunk, and Yvonne G van der Zee. 2008. Social Comparison in the Classroom: A Review. *Rev. Educ. Res.* 78, 4 (Dec. 2008), 828–879.
 - [9] Nia M Dowell, Whitney L Cade, Yla Tausczik, James Pennebaker, and Arthur C Graesser. 2014. What Works: Creating Adaptive and Intelligent Systems for Collaborative Learning Support. In *Intelligent Tutoring Systems*. Springer International Publishing, 124–133.
 - [10] Nia M Dowell, Oleksandra Skrypnik, Srećko Joksimović, Arthur C Graesser, Shane Dawson, Dragan Gašević, Thieme A Hennis, Pieter de Vries, and Vitomir Kovanovic. 2015. Modeling Learners' Social Centrality and Performance through Language and Discourse. *International Educational Data Mining Society* (2015).
 - [11] Nia M M Dowell, Christopher Brooks, Vitomir Kovanovic, Srećko Joksimović, and Dragan Gašević. 2017. The Changing Patterns of MOOC Discourse. In *Proceedings of the Fourth (2017) ACM Conference on Learning @ Scale (L@S '17)*. ACM, New York, NY, USA, 283–286.
 - [12] Leon Festinger. 1954. A Theory of Social Comparison Processes. *Hum. Relat.* 7, 2 (May 1954), 117–140.
 - [13] S T Fiske. 2010. Interpersonal Stratification: Status, Power, and Subordination. In *Handbook of Social Psychology*, Susan T Fiske, Daniel T Gilbert, and Gardner Lindzey (Eds.). Society for Personality and Social Psychology, Tampa, Vol. 27. John Wiley & Sons, Inc., Hoboken, NJ, USA, 165.
 - [14] Gail Anna Golden and Frances Cherry. 1982. Test performance and social comparison choices of high school men and women. *Sex Roles* 8, 7 (July 1982), 761–772.
 - [15] Arthur C Graesser, Danielle S McNamara, Zhiqiang Cai, Mark Conley, Haiying Li, and James Pennebaker. 2014. Coh-Metrix Measures Text Characteristics at Multiple Levels of Language and Discourse. *Elem. Sch. J.* 115, 2 (2014), 210–229.
 - [16] Philip J Guo and Katharina Reinecke. 2014. Demographic differences in how students navigate through MOOCs. In *L@S '14: Proceedings of the first ACM conference on Learning @ scale conference*. ACM Request Permissions, New York, New York, USA, 21–30.
 - [17] Andrew F Hayes. 2009. Beyond Baron and Kenny: Statistical Mediation Analysis in the New Millennium. *Commun. Monogr.* 76, 4 (Dec. 2009), 408–420.
 - [18] M A Hearst. 2000. The debate on automated essay grading. *IEEE Intelligent Systems and their Applications* 15, 5 (2000), 22–37.
 - [19] Srećko Joksimović, Nia Dowell, Oleksandra Poquet, Vitomir Kovanovic, Dragan Gašević, Shane Dawson, and Arthur C Graesser. 2018. Exploring development of social capital in a CMOOC through language and discourse. *The Internet and Higher Education* 36 (Jan. 2018), 54–64.
 - [20] René F Kizilcec and Christopher Brooks. 2017. Diverse big data and randomized field experiments in massive open online courses: Opportunities for advancing learning research. *G. siemens & c. lang (eds.), handbook on learning analytics & educational data mining* (2017).
 - [21] René F Kizilcec, Andrew J Saltarelli, Justin Reich, and Geoffrey L Cohen. 2017. Closing global achievement gaps in MOOCs. *Science* 355, 6322 (Jan. 2017), 251–252.
 - [22] J M Levine. 1983. Social Comparison and Education. <http://eric.ed.gov/?id=ED236469>. Accessed: 2017-4-30.
 - [23] Marianne Schmid Mast. 2010. Interpersonal power and social perception in a hierarchy: The interpersonal power and behaviour model. *European Review of Social Psychology* 21, 1 (March 2010), 1–33.
 - [24] Danielle S McNamara, Arthur C Graesser, Philip M McCarthy, and Zhiqiang Cai. 2014. *Automated Evaluation of Text and Discourse with Coh-Metrix*. Cambridge University Press.
 - [25] David A Morand. 2000. Language and power: An empirical analysis of linguistic strategies used in superior–subordinate communication. *Journal of organizational behavior* 21, 3 (2000), 235–248.
 - [26] James W Pennebaker, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. 2015. *The development and psychometric properties of LIWC2015*. Technical Report.
 - [27] Emmy A Pepitone. 1972. Comparison Behavior in Elementary School Children. *Am. Educ. Res. J.* 9, 1 (1972), 45–63.
 - [28] Richard E Petty, John T Cacioppo, Alan J Strathman, and Joseph R Priester. 2005. To think or not to think. *Persuasion: Psychological insights and perspectives* (2005), 81–116.
 - [29] Nathaniel Poor and Roei Davidson. 2016. The Ethics of Using Hacked Data: Patreon's Data Hack and Academic Data Standards. (2016).
 - [30] Isabelle Régner and Jean-Marc Monteil. 2007. Low-and high-socioeconomic status students preference for ingroup comparisons and their underpinning ability expectations. *Revue internationale de psychologie sociale* 20, 1 (2007), 87–104.
 - [31] Yves Rosseel. 2012. lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, Articles* 48, 2 (2012), 1–36.
 - [32] Elena Stephan, Nira Liberman, and Yaacov Trope. 2010. Politeness and psychological distance: a construal level perspective. *J. Pers. Soc. Psychol.* 98, 2 (Feb. 2010), 268–280.
 - [33] Hoi K Suen. 2014. Peer assessment for massive open online courses (MOOCs). *The International Review of Research in Open and Distributed Learning* 15, 3 (June 2014).
 - [34] Lisa Collins Tidwell and Joseph B Walther. 2002. Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations: Getting to know one another a bit at a time. *Hum. Commun. Res.* 28, 3 (2002), 317–348.
 - [35] Keith J Topping. 2009. Peer Assessment. *Theory Pract.* 48, 1 (2009), 20–27.
 - [36] Chih-Hsiung Tu and Marina McIsaac. 2002. The Relationship of Social Presence and Interaction in Online Classes. *Am. J. Distance Educ.* 16, 3 (Sept. 2002), 131–150.
 - [37] United Nations Development Programme. 2017. Human Development Reports. <http://hdr.undp.org/en/composite/HDI>. Accessed: 2017-5-30.
 - [38] Leaf Van Boven, Joanne Kane, A Peter McGraw, and Jeannette Dale. 2010. Feeling close: emotional intensity reduces perceived psychological distance. *J. Pers. Soc. Psychol.* 98, 6 (June 2010), 872–885.
 - [39] Nanine A E van Gennip, Mien S R Segers, and Harm H Tillema. 2010. Peer assessment as a collaborative learning activity: The role of interpersonal variables and conceptions. *Learning and Instruction* 20, 4 (2010), 280–290.
 - [40] Andrii Vozniuk, Adrian Holzer, and Denis Gillet. 2014. Peer Assessment Based on Ratings in a Social Media Course. In *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge (LAK '14)*. ACM, New York, NY, USA, 133–137.
 - [41] Taniesha A Woods, Beth Kurtz-Costes, and Stephanie J Rowley. 2005. The Development of Stereotypes About the Rich and Poor: Age, Race, and Family Income Differences in Beliefs. *J. Youth Adolesc.* 34, 5 (Oct. 2005), 437–445.
 - [42] Saijing Zheng, Mary Beth Rosson, Patrick C Shih, and John M Carroll. 2015. Understanding Student Motivation, Behaviors and Perceptions in MOOCs. In *CSCW '15: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM Request Permissions, New York, New York, USA, 1882–1895.