Gender and Gendered Discourse in Two Online Science College Courses

Genevieve M. Henricks, Michelle Perry, Suma Bhat

University of Illinois at Urbana-Champaign

Corresponding author: henrick2@illinois.edu

Abstract. Discussion forums are important components of online courses because of the collaboration and community they foster, and the language used within the discussion forums may be influential in this development. In particular, studying the gendered language patterns of discussion forums can help gain insight into students' state of mind and propensity to form a community and thus ultimately may help explain men and women's differential success in online courses. This study describes the state of gendered language use in two online science courses. The results reveal that women and men do not differ in their language use along traditionally gendered lines, which holds promise for women in online courses. This means that the language that women use does not overtly mark them as female, which has the potential to help subvert the typical result of the negative outcomes associated with the female marker in science courses.

Henricks, G., Perry, M., Suma, B., (2021) Gender and Gendered Discourse in Two Online Science College Courses. Computer-Based Learning in Context, 3 (1), 1-16.

Keywords: Distance education, Gendered language, Higher education, Online course, Online discussion, Science education, STEM education

1 Introduction

The need for STEM (Science, Technology, Engineering, and Mathematics)-field graduates is greater than ever. STEM jobs have grown by 14% since 2008, compared to 1.4% for non-STEM jobs, and they are expected to grow another 8.9% by 2024. Along with the greater need comes greater salaries: STEM workers earn 29% more than non-STEM workers (Noonan, 2017). Although these factors should make STEM careers tempting to students, STEM positions in the United States are going unfilled due to gaps in individuals' knowledge and skill sets necessary to perform these jobs [National Math and Science Initiative (NMSI), 2014].

Women, in particular, are shying away from STEM positions: they make up only 24% of the STEM workforce (Noonan, 2017) and are nearly twice as likely to work part-time in STEM positions as compared to men (National Science Foundation, National Center for Science and Engineering Statistics, 2019). To ensure that the STEM workforce not only has enough participants but also is diversified (thereby fostering the possibilities that multiple and diverse perspectives are available to STEM enterprises), it is of utmost importance to focus on recruiting, retaining, and educating underrepresented students in STEM fields. Online courses are one possibility for solving this problem because these courses currently are widespread in STEM programs that lead to baccalaureate degrees and allow for students to control more of when, where, and how they participate, thereby potentially mitigating some of the barriers to success for underrepresented STEM students in face-to-face STEM courses. However, as Huang, Hood, and Yoo (2013) pointed out, there is a gender divide when it comes to engaging with some computer applications for learning, including blogs, wikis, online games, and immersive virtual environments.

Women's anxieties with computers may fuel this divide, but such anxieties may be mitigated in situations that are heavily dependent on the social networking aspects of interacting through computers because of women's needs for and thriving in situations that depend on affiliation (e.g., Drescher & Schultheiss, 2016). Huang, Hood, and Yoo's (2013) study also suggests this, as they found that unlike other computer applications, women did not have more anxiety than men when using social networking and online video sharing tools, which are very collaborative in nature.

This study will examine one potential mechanism affecting women's involvement in one subset of STEM courses: online courses in the physical sciences. Specifically, it will analyze the type of language used in class discussion forums. Students' language in class discussion forums was chosen as the key mechanism because this can be a marker of social status (i.e., gender), and thus these markers have the potential to impact how and with whom they share information in the online environment (Cho, Gay, Davidson, & Ingraffea, 2007).

In this study, we aim to advance an understanding of how gendered language is used by men and women in ways that potentially help or hinder community building in two online courses from different science disciplines chosen for their diverse topic areas within the physical sciences and their varying target audiences. We will do so by examining the language that students use, with special attention to whether men and women use traditionally gendered language, in the categories that we have identified. We ask:

1. To what extent do men use a language consistent with a report style of communicating and do women use language consistent with a rapport style of communicating in an online chemistry course?
2. To what extent do men use a report style of communicating and do women use a rapport style of communicating in an online astronomy course?
3. When comparing the two courses at large, to what extent are they similar and different in terms of the general styles of communicating?

2 Literature Review

2.1 **Women in STEM Fields.** The demographically disproportionate enrollment of greater numbers of men than women in STEM courses is a worrisome and challenging problem facing United States colleges and universities. Except for the biological and medical sciences, women are significantly underrepresented in STEM fields (Corbett & Hill, 2015). The breadth of contexts in which women are underrepresented in STEM is apparent from the plethora of studies investigating attrition from the STEM academic pipeline (e.g., George-Jackson, 2011; Heilbronner, 2013; Jackson et al., 2013; Myers & Pavel, 2011; Xu, 2008).

The attrition rate for women in STEM has been attributed to several causes (e.g., Blickenstaff, 2005; Hill et al., 2010). Some explanations point to women assessing themselves lower in their STEM abilities and having higher expectations for success than their male counterparts (e.g., Correll, 2001). When this is the case, it is easy for women to opt out because they get defeated more easily than men. Such assessments may be because women believe they are cognitively different from their male peers, but such an explanation ignores the strong interplay between sociocultural factors like stereotypes and socialization on cognitive performance (e.g., Miller et al., 2015; Wai et al., 2010). Other researchers (e.g., Margolis & Fisher, 2002) reason about a non-inclusive culture in STEM classes, leading to a sense of isolation and a lack of confidence. These two explanations can also conspire to work against women staying and succeeding in STEM: a non-inclusive culture, isolation, and lack of confidence puts women at risk of not having the necessary supports when they feel like they have not succeeded at the level of their male peers.

Because of the increased risk for women in online courses (of vulnerability to feelings of isolation, leading to feelings of failure, leading to women's attrition in STEM) on the one hand and, on the other, the draw of women to online courses (given the convenience and increasing ubiquity of online offerings), this study will examine a potential mechanism that could hinder or help women's sense of belonging and confidence in the online space: the language used in the online courses' discussion forums.

2.2 **Women in Online STEM Courses.** Although firm numbers are not tracked well, online course drop-out rates tend to be high, with some studies showing about 10% (Tan & Shao, 2015), others finding 30-40% (Tyler-Smith, 2006), and yet others reporting 40%-80% (Smith, 2010). The results to do not improve when homing in on women in online STEM courses. In their study of 3,600 community college students, Wladis et al. (2015) found that women were overrepresented in online STEM courses, but they also found that women were more likely to fail and withdraw than men when compared to their face-to-face counterparts. Likewise, Cochran et al. (2014) found that women who were majoring in math or science fields withdrew from online courses more so than men in these fields. Although Wladis and colleagues (e.g., Hachey et al., 2015; Wladis et al., 2015) are advancing the development of online models of success for STEM courses in community colleges, research on how women perform in STEM online courses at four-year institutions, in particular, is lacking. To shed light on this problem, this study will contribute to advance an understanding of how women participate in online STEM courses at a four-year institution.

Feeling a part of a community is one important component to success in any course, and online courses are no exception. Because it has long been known that students learn from each other, as well as from the instructor (e.g., Bell et al., 1985; Forman & Cazden, 1985), and often feel more connected to the course when they see that other students are having comparable experiences (Freeman et al., 2007), the absence of a community along with feelings of isolation can be detrimental to success in a course, especially in an online course.

Discussion forums may be helpful for women in particular because they provide an outlet for community building and interacting, which is important given that women have relatively higher needs for affiliation (Drescher & Schultheiss, 2016). The report of the 2015 Programme for International Student Assessment (PISA) echoes this, stating that “girls in almost every country and economy tend to value relationships more than boys, meaning that girls agree more often than boys that they are good listeners, enjoy seeing their classmates be successful, take into account what others are interested in, and enjoy considering different perspectives” (Gurria, 2016, p.24). Research on Massive Open Online Courses (MOOCs) suggest that women’s participation in online discussion forums is indeed a crucial key to their success in those courses. For example, Crues et al. (2018) found that students who participated in forums had greater rates of persistence in a computer programming MOOC, and women participated in the forums at greater rates than men. These outcomes were despite the fact that women were significantly outnumbered by men in this class. Bayeck’s (2016) analysis of an engineering MOOC also reiterated the positive effects of providing collaborative opportunities to women in an online STEM course: specifically, women enrolled at an unusually high rate of 1.57 times that of men.

Simply offering the discussion forums does not necessarily guarantee community formation or other positive outcomes for either men or women; instructors need to implement them thoughtfully to maximize outcomes (e.g., Salter & Conneely, 2015). And, once implemented, students need to engage with them productively. The ways in which students engage with one another may preclude them from maximizing learning outcomes, especially if there is exposure to language that is off-putting or disparaging (Herring et al., 1994). Negative outcomes, like women withdrawing from participation, may certainly arise from the use of gendered language (e.g., authoritative language for men, signaling that the men have power, and polite or demure language for women, signaling that they are relatively subservient) in online courses (Herring, 1996).

2.3 **Language: A Potential Mechanism for Online Course Success.** A Given that students taking up and owning ideas from others is an important aspect of learning (e.g., Barron, 2000, 2003), examining the language used to share those ideas is crucial. Numerous fields, including psychology, linguistics, communications, anthropology, and sociology, have investigated the ways in which self-identified males and self-identified females use language. The consensus is that, although more similar than not, there are measurable differences in men and women’s use of certain features of language (Canary & Dindia, 2009). Patterned differences in words, phrases, and sentences have led researchers to categorize men’s communication style, compared to women’s, as generally dominant and aggressive and packed with information (a “report” style of communicating) and women’s communication style, compared to men’s, as generally submissive and affiliative (a “rapport” style of communicating) (Tannen, 2007). Importantly, these language styles signify power differences, leading to real-world power differentials between men and women in both the private and public spheres (Hall, 2004; McConnell-Ginet, 2004).

These differences continue into the online realm. Although, in theory, computer-mediated communication mitigates inequality because of the relative lack of social markers (Kiesler et al., 1984) it is clear that social markers exist, with men’s language characterized as more authoritative and women’s language characterized as more supportive (Guiller & Durndell, 2007). Herring’s (1993) work on computer-mediated communication details how language contains cues about the communicator’s gender, finding that women are more supportive and men more adversarial (Herring et al., 1994). Men’s language and interactions tend to delegitimize what women have to say; this makes them feel unwelcomed in the community of ideas, ultimately making women feel uncomfortable and lessening their participation (Herring, 1996). Yates (1997), too, agreed that social markers exist within computer-mediated communication, but she also pointed out that these markers have the potential to be manipulated, making room for gender identities to be reconstructed, thus mitigating the negative effects that gendered language may promote.

The digital landscape has rapidly changed since the 1990s when much of the seminal work on gender and computer mediated communication was published; since then, software programs like Coh-Metrix (McNamara & Graesser, 2012) have allowed for studying these issues at scale and at very granular, linguistic levels by automatically detecting cognitive, social, and affective processes through linguistic cues. Developing these tools and effectively using their results requires continuously refining theoretical understandings of what processes are good and why, and it also involves exploring how such processes evolve. The deeper insights that follow allow researchers and developers to practically apply language’s role in, for example, promoting or hindering community-building, which is an important aspect of online courses (Tu & Corry, 2001). Thus, this study investigates the extent to which gendered discourse is used in two online science courses’ discussion forums to understand if there is differential use that has the potential, in turn, to affect a sense of community and a safe place for learning.

Additionally, this study conducts its investigation in the same vein of Cade et al. (2014); it has the potential to model students’ outlooks on the group in a nuanced way, beyond emotive language, which overtly marks one’s feelings and perspectives. The aspects of language chosen for this study were derived from Tannen’s classic (1990) work on gendered discourse because Tannen essentially laid the important groundwork on how men’s and women’s language styles promote different social agendas. Furthermore, these categories were used to shape Newman et al.’s

(2008) analysis of gendered language in a variety of contexts. Borrowing directly from Newman et al.'s analytic approach (also see Table 1), this study examines:

(a) Pronouns	Women tend to use more pronouns than men. The use of any pronoun implicitly requires shared understanding and meaning (e.g., if a student writes we or they, that student gives the reader credit for knowing to whom the writer is referring), thus highlighting a closeness between discussants and building community among students. The use of first-plural pronouns (we, us), in particular, convey a sense of community.
(b) Politeness	Women's language tends to be more polite than men's and politeness provides a welcoming atmosphere for others to join in.
(c) Hedging	Vague language that avoids definitiveness tends to be associated with women. This sort of uncertainty may make others think that you have less to contribute, thereby diminishing that person's contribution; however, hedging may also signal to others to help clarify and to join in to support the community.
(d) Personal and interpersonal queries	Women tend to focus more on others than men when communicating. By explicitly addressing others, this sort of language can build warmth and community.
(e) Information giving	Men tend to focus on relaying facts more so than women. This places the speaker in a position of authority, and may discourage others from engaging with or challenging the speaker, thereby diminishing community and potentially leaving women more reluctant to join in.
(f) Confidence	Men tend to portray more confidence in their language than women. When confidence is detected, like in information giving, the speaker is afforded authority, which may discourage others from engaging with or challenging the speaker, thereby diminishing community and potentially leaving women less reluctant to join in.

In addition to examining the aforementioned linguistic categories that Newman et al. (2008) investigated, this study will explore whether there are Word Count differences, which Newman et al. did not find. Other studies have found differences. In their review of gender differences in computer-mediated communication and computer-supported-collaborative learning, Prinsen et al.'s (2007) reported conflicting results: some found that men tend to have higher word counts and other studies found that women tend to have higher word counts. Word Count is an important feature to examine because it is generally an indicator of engagement levels (Tausczik & Pennebaker, 2010), which directly relates to community formation (Liu et al., 2007).

Table 1: Examples of each LIWC Category, from Chemistry and from Astronomy

Previous Findings[1]	Relevant LIWC Categories	Finding reported by:
Word Count M=F	Word Count	Newman et al., 2008
Pronouns F>M	Total Pronouns	Newman et al., 2008
Politeness F>M	Discrepancies	Tannen, 1990
Hedging F>M	Tentative	Tannen, 1990
Personal/interpersonal queries F>M	Social Processes; Authenticity; Affiliation	Tannen, 1990
Information Giving M>F	Numbers; Analytic	Tannen, 1990
Confidence M>F	Clout; Certainty	Tannen, 1990

3 Methods and Data Sources

3.1 **Data Set - Chemistry.** Data were collected from the discussion assignments from the Fall 2015, Spring 2016, Summer 2016, and Fall 2016 semesters of an introductory online chemistry course offered at a large university in the

[1] M>F indicates that this category is more typical for men than for women, and F>M indicates that this category is more typical for women than for men, as reported by the authors in the far-right column.

Midwest. The same instructor taught each of the semesters, and the syllabi did not change across semesters. The course was taught entirely asynchronously, and in addition to the students using an e-textbook, the instructor provided: recorded video lectures that were accompanied by pre-lecture assignments and followed by homework problems; biweekly quizzes, and three exams as well as a cumulative final exam. Germain to this study are the 13 discussion assignments, one for each week of material; students had to participate each week if they wanted to earn the full 5% of the final grade that was allocated for discussion assignments. The assignments entailed the following: Each week, the instructor created 4-5 discussions forums, with each forum consisting of an exam-like homework problem. The students could choose to post to any of the of the available forums for that week, and then they had to (a) post a solution to the problem, (b) post a question, or (c) answer a question. Students were required to do only one of the aforementioned activities and on only one of the forums each week, although they were welcome to participate more if they chose to do so. The instructor's intended goal of this assignment, as posted on the syllabus, was to have students "learn how to approach challenging problems from other student explanations, and by teaching other students."

3.2 **Data Set - Astronomy.** Data were also collected from the discussion forums from the Fall 2016 semester of an upper-level online astronomy course. This course was taught entirely asynchronously; the instructor did not provide any form of video lectures and relied heavily on the e-textbook. Students had weekly quizzes, a large group project, and a final exam. Like the chemistry course, weekly forum participation in the discussion forum was required. Participation in at least 10 of the forums constituted 25% of students' grades (compared to only 5% of students' grades in Chemistry). The stated goal of the forums, according to the syllabus, was "to discuss class facts to better understand the science…" Students were required to post a response to a topic and post at least two responses to other students' postings each week.

3.3 **Participants.** Accounting for all four semesters of chemistry, there were 368 total students enrolled, only 345 of whom were unique enrollees (18 students had enrolled in two semesters and 3 students had enrolled in three semesters). There were 74 total drops, but because there were 8 students who had dropped the course twice across the four semesters, there were only 66 unique students who dropped (37 women and 29 men). This left 271 unique students who were enrolled and completed the course. We had incomplete data for 24 of the students, leaving a total of 247 students for analysis (132 women and 115 men). The astronomy course began with a total of 221 students (53 women, 164 men, and 4 students of unknown or non-binary gender, who were not included in the analysis). A total of 8 students (2 women and 6 men) dropped, leaving 51 women and 158 men enrolled by the end of the semesters.

A Center for Innovation in Teaching and Learning (CITL) staff member with clearance to access FERPA-protected data took the discussion posts and replaced any personally identifiable information with a random hash and created a key detailing the gender that corresponded with each of the random hashes. (Note that this was done after the courses were complete; students enrolled in the courses were able to see the names of who posted at all times.) The chemistry students generated 3,121 unique posts throughout the four semesters under investigation. Women created 1,677 of those posts, and men created 1,444 of those posts. The astronomy students generated 13,335 unique posts. Women created 3,375 of those posts, and men created 9,960 of those posts. Each students' postings were aggregated, and we conducted analyses at the level of the student.

3.4 **Text Analysis.** To analyze the gendered language employed in the discussion forums, we used Linguistic Inquiry and Word Count 2015 (LIWC; Pennebaker et al., 2015), a computerized text analysis program in its third iteration that outputs the percentage of words in a given text that fall into one or more of over 80 linguistic (e.g., pronouns, conjunctions), psychological (e.g., anger, achievement), and topical categories (e.g., health, religion). Its corpus consists of more than 500,000 texts that range from tweets to novels. A major addition to the 2015 version of LIWC is the inclusion of four summary variables empirically developed in the research labs of the LIWC team: Analytical Thinking (Pennebaker et al., 2014), Clout (Kacewicz et al., 2012), Authenticity (Newman et al., 2003), and Emotional Tone (Cohn et al., 2004). Based on the relation to previous research, this study focused on Analytic Thinking (e.g., critical thinking and logical thinking), Clout (e.g., confidence and expertise), and Authenticity (e.g., openness and honesty). Analytic Thinking is used as a means of indicating information sharing; Authenticity is used to denote personal and interpersonal inquiries; and Clout is used to understand confidence.

LIWC generates a rating for each of these summary categories. We include examples from each category, from each course, in Table 2. We chose this application, at least in part, because of its wide-spread usage, because exploring trends within studies of gendered language is often difficult due to the varied meanings and understandings of what constitutes gendered language (Newman et al., 2008). A program as ubiquitously used and updated as LIWC can help systematize the study of gendered language. As such, this study builds on Newman et al.'s (2008) use of LIWC 2001 to analyze gendered language. Note that there is nothing innately gendered about the language that comprises this study's definition of gendered language; rather the summation of the categories represents patterns of communication that have been correlated to men and women. To clarify the relation between LIWC categories explored in this investigation (Table 2) and previous findings for use of gendered language, we present the relationship between the gendered language categories explored in previous research examined by Newman et al. (2008) and Tannen (1990) and the LIWC categories we explored in this investigation, in Table 1.

Table 2: Examples of each LIWC Category, from Chemistry and from Astronomy

LIWC Category (promoting report style or promoting rapport style)	Chemistry Example	Astronomy Example
Analytic[2] (report)	For 29, you have to use the equation $lnk = -Ea/R(1/T) + lnA$. Slope is equal to $-Ea/R$ and the intercept is equal to lnA. Thus, your equation should look like this now: $lnk = -917(1/T) + -.441$ Next, you need to [find] k by plugging in the temperature given in the problem. Once you have k, then you can plug it into the differential rate. The rate of the equation depends on which order your problem is in. Hope this helps!	L gives an estimate of how long advanced civilization can sustain interstellar communication. Humans have been involved into communication with aliens for the past 75 years. Taking into account factors like mortality rate, lack of resources the upper limit of L can be set to that year. I think that we do not have the resources to sustain interstellar communicate. I estimate that L should be around 200.
Clout[2] (report)	I solved for K, which was 0.2963 after rounding Lastly I plugged everything in to the $DeltaG$ equation (CONVERT $-2.00kJ$ into Joules) $DeltaG = (-2000) + (8.314)(298K)\,In\,(0.2963) = -5013.68$, divide by 1000 to convert back into kJ DeltaG = -5.014kJ which equals maximum work able to be put in.	My Estimate for the drake equation is $18 * 0.7 * .148 * 0.8 * 0.1 * 0.7 * 250 = 26.1072$. My number is on the pessimistic side. According to the textbook, the average is over 800,000. This number is a little less than my original estimation. However, after actually thinking about the values more critically, I do believe that this value is a lot more accurate. This number leads to the nearest civilization being approximately 9800 light years away, according to the book.
Authenticity[2] (rapport)	What kind of tripped me up at first was trying to figure out what to do with the amount of	I'm curious to know what variable changed the most for you from the beginning of the course to the end. Your estimates are interesting because it is the

[2] these are summary variables, which are automatically determined by LIWC's proprietary algorithm, so we provide LIWC-generated examples in their entirety.

	water we're given. You have to go back in your brain and remember that $molarity = moles/liters$, and that the concentration of $H+$ is molarity…by multiplying the concentration by the mL of water given, you can get moles.	opposite of what I got. My first week estimate was significantly higher than my final. However, I find your estimate optimistic. I believe that life can exist and I'm sure it does, but I think the obstacles for that life to become intelligent is just too much. Like you said the only evidence we have is on earth, and we just have to keep looking and hoping!
Pronoun (rapport)	Almost this same question was on the recent quiz, yet we weren't given DG standard. …	…Mine was about 1000, but everyone else seems to be hovering nearly $10,000$ so I think we should go with about $8,000$. Anyone disagree?
Number (report)	$G = Go + RT\ lnQ$ BUT since its equilibrium lnK is used (which I forgot). $A <--> B = 1.0M$ $G = -1.85kJ$ $Go = -5.15kJ$	My NP estimate was: 0.5 My FS estimate was : 0.2 MY FL estimate was: 0.5…My L estimate was: 10000 Using Drakes equation, $10 * 1 * 0.5 * 0.2 * 0.5 * 0.1 * 0.1 * 10000$, my estimate was 50.
Social Processes (rapport)	Nice to meet you! If you need help with anything, I'm always willing to help! I can't imagine having a language barrier. I would do anything I can to help!	Hey, I just joined! I can provide WhatsApp, groupme, Facebook, or phone number for communication…This way we can form our group and start the process for the video presentation.
Discrepancy (rapport)	I think the concentration of oxygen would have to play a role if we dipped it into liquid oxygen…We could also increase the temperature to increase the rate of the burning of the cheeto also.	I think your estimate is on the low side, having only 28 other civilizations in the galaxy would make for a very lonely place. I can see how you could get this number especially if your life span estimate is low and who knows, given the volatile nature of intelligent species there could only be 28.
Tentative (rapport)	I don't think you can use the equation because there is an acid and a base but the conjugate is not present. I think you just figure out what is left over and figure the pH or depending on what species is left. Hope this helps!	I am not sure if this is one of those discussion posts…I think we need someone to manage what the group is doing (leader sorts).
Certainty (report)	Yes Kw is always neutral for water… The fact that it is the constant for water means that it will always be neutral even the PH number is slightly below 7 or above.	This was drastically smaller than my original estimate and I still think that this is very optimistic. If I were to do another Drake Equation I would definitely go with more pessimistic answers. Our estimates are completely different but I really like yours.
Affiliation (rapport)	Instagram and Facebook (and Snapchat) are how I keep my family and friends up to date on travel. I totally agree with you.	There shouldn't be any communication problems in the group as each of us at least know a friend or two on the team. …

4 Results

The LIWC output generated counts for Word Count, ratings for the Authentic, Analytic, and Clout categories, and percentages for the Pronouns, Discrepancies, Tentative, Social Processes, Numbers, and Certainty categories (to perform analyses on percentages, we transformed these into counts). What follows is an exploration of men's and women's use of the targeted LIWC categories for each course.

4.1 **LIWC Categories and Gender in Chemistry and Astronomy.** To answer the first research question—To what extent do men use a report style of communicating and do women use a rapport style of communicating in an online chemistry course? —we examined differences between men and women in the LIWC categories identified. In a posthoc analysis of a two-sided Wilcoxon-Mann-Whitney test calculation provided by $G * Power$ 3.1 (Faul et al., 2009), we reached a computed power of 0.97 with $\alpha = 0.05$, effect size $d = 0.5$ given $N_1 = 132$ and $N_2 = 115$. Because the data were not normally distributed, we analyzed the data using the Mann-Whitney U test, which is a nonparametric alternative to a t-test that compares medians. We found no significant differences between men and women in any of the summary (Analytic Thinking, Authenticity, and Clout), linguistic (Pronoun, Number, and Word Count), or psychological (Social processes, Discrepancy, Tentative, Certainty, and Affiliation) categories. Table 3 contains the median values for each of the LIWC categories by gender across semesters for chemistry. Because medians, rather than means, are compared, Table 3 gives the interquartile range as a way to understand the spread instead of the standard deviation or standard error.

Next, we analyzed whether these same findings held for another STEM course: an upper-level astronomy class. To answer the second research question—To what extent do men use a report style of communicating and do women use a rapport style of communicating in an online astronomy course? —we performed the same analyses that we conducted on the data from the online chemistry course. In a posthoc analysis of a two-sided Wilcoxon-Mann-Whitney test calculation provided by $G * Power$ 3.1 (Faul et al., 2009), we reached a computed power of 0.86 with $\alpha = 0.05$, effect size $d = 0.5$ given $N_1 = 51$ and $N_2 = 158$. Table 3 contains the median values for each of the LIWC categories by gender for astronomy. With the exception of Tentative language use, which men used more than women, we found no differences between men's and women's use of language in any of the categories we examined, which parallels the results found for the chemistry course.

Table 3: Median Counts of LIWC Categories by Gender and Course

	LIWC Category for C (Chemistry) or A (Astronomy)	*Median*	Mann-Whitney U	Z	IQR	*p* value
Summary Categories	**Analytic (C)**		6849.00	-1.323		.187
	Men	86.5			18.15	
	Women	84.4			29.25	
	Analytic (A)		3827.00	-.538		.591
	Men	67.6			20.08	
	Women	65.7			23.04	
	Clout (C)		7255.0	-.598		.358
	Men	48.4			15.97	
	Women	49.7			13.48	
	Clout (A)		3471.00	-1.486		.137
	Men	58.2			26.23	
	Women	60.5			25.80	
	Authenticity (C)		7430.0	-.285		.776
	Men	9.5			21.59	
	Women	9.5			37.47	
	Authenticity (A)		3997.0	-.085		.932
	Men	49.3			30.39	
	Women	49.0			31.58	
Linguistic Categories	**Pronoun (C)**		7259.0	-.591		.554
	Men	29.0			54.0	
	Women	20.5			56.0	
	Pronoun (A)		3825.0	-.556		.578
	Men	3.0			2.0	
	Women	3.0			1.0	
	Number (C)		7201.5	-.694		.488
	Men	63.0			81	
	Women	60.0			0.0	
	Number (A)		4015.5	-.045		.964
	Men	1.0			0.0	
	Women	1.0			0.0	
	Word Count (C)		7157.00	-.773		.439
	Men	451.0			601	
	Women	424.0			593	
	Word Count (A)		3795.5	-.622		.534
	Men	500.0			290	
	Women	540.0			265	
Psychological Categories	**Social Processes (C)**		7493.00	-.173		.862
	Men	12.0			24.0	
	Women	12.0			25.0	
	Social Processes (A)		3792.5	-.700		.484
	Men	1.0			1.0	
	Women	1.0			1.0	

Discrepancy (C)		7429.0	-.289		.773
Men	4.0			8.0	
Women	4.0			8.0	
Discrepancy (A)		3657.0	-1.413		.158
Men	0.0			0.0	
Women	0.0			0.0	
Tentative (C)		7565.6	-.044		.965
Men	7.0			10.0	
Women	6.0			12.0	
Tentative (A)		3366.0	-2.017		.044*
Men	1.0			1.0	
Women	0.0			1.0	
Certainty (C)		7139.0	-.810		.418
Men	4.0			7.0	
Women	4.0			7.0	
Certainty (A)		3838.5	-.902		.367
Men	0.0			0.0	
Women	0.0			0.0	
Affiliation (C)		7202.5	-.697		.486
Men	6.0			10.0	
Women	4.0			11.0	
Affiliation (A)		3561.5	-1.429		.153
Men	1.0			1.0	
Women	1.0			0.0	

4.2 **Differences in Gendered Language Use, Between Courses.** Because we did not find differences between men's and women's use of gendered language in either chemistry or astronomy, we wondered whether the use of gendered language as a whole might be different between these two courses. This was possible given the differences in content, in how the discussion forums were organized, and how much weight the discussion forum posts had toward students' final grades. This led to the third research question: When comparing the two courses, to what extent is the language used similar and to what extent is it different? Table 4 compares the median counts of the LIWC Categories for each of the courses.

Table 4: Median Counts of LIWC Categories by Course

LIWC Category	*Median*	*Mann-Whitney U*	Z	IQR	*p* value
Analytic		12489.5	-2.769		.000
Chemistry	85.8			21.69	
Astronomy	67.1			21.01	
Clout		16252.5	-6.818		.000
Chemistry	48.9			14.71	
Astronomy	59.9			25.1	
Authenticity		6829.0	-13.541		.000
Chemistry	9.5			24.93	
Astronomy	49.2			31.22	
Pronoun		8057.0	-12.714		.000

Chemistry	23.0			55.0	
Astronomy	3.0			2.0	
Number		3472.5	-16.267		.000
Chemistry	60.0			83.0	
Astronomy	1.0			0.0	
Word Count		21928.5	-2.769		.006
Chemistry	433.0			595	
Astronomy	512.0			285	
Social Processes		9159.0	-12.065		.000
Chemistry	12.0			23.0	
Astronomy	1.0			1.0	
Discrepancy		5993.0	-14.897		.000
Chemistry	4.0			8.0	
Astronomy	0.0			0.0	
Tentative		8022.0	-12.984		.000
Chemistry	6.0			11.0	
Astronomy	1.0			1.0	
Certainty		6041.0	-15.149		.000
Chemistry	4.0			7.0	
Astronomy	0.0			0.0	
Affiliation		13563.5	-8.994		.000
Chemistry	5.0			11.0	
Astronomy	1.0			1.0	

In a posthoc analysis of a two-sided Wilcoxon-Mann-Whitney test calculation provided by $G * Power$ 3.1 (Faul et al., 2009), we reached a computed power of 0.88 with $\alpha = 0.5$, effect size $d = 0.3$ given $N_1 = 247$ and $N_2 = 209$. Using the Mann-Whitney U test, we found that all of the summary categories were significantly different between the two courses. Although the posts in the chemistry course were significantly more Analytic, the posts in the astronomy course displayed significantly more Authenticity and Clout. The analyses also revealed a significant difference in the Word Counts between the two courses. Specifically, the average word count of the posts in the astronomy discussion forums was significantly higher than that in the chemistry forums. Likewise, both Pronouns and Numbers were used significantly more frequently in the chemistry forum posts than in the astronomy forum posts. Additionally, all of the psychological categories were significant. Students in the chemistry course used language pertaining to the categories of Social Process, Affiliation, Certainty, Discrepancy, and Tentative significantly more in their forum posts than students in the astronomy course.

Although Research Questions 1 and 2 revealed no significant differences between men and women's use of gendered language, Research Question 3 showed that gendered language was still used. Students within each course used gendered language in very different ways, however, which we discuss in the following section.

5 Discussion

This study revealed that women and men did not differ in their language use along traditionally gendered lines. There are many possible explanations for this. It may be the case that online discussion forums provide a context in which patterns of gendered language do not thrive; similarly, it is possible that using gendered language along gendered lines has dwindled in both online and in-person contexts over the decades. It is also possible that both genders are intentionally using markers typically associated with the other gender to their perceived advantage, as Yates (1997) suggested could happen.

In any case, this finding is very promising for women in online courses; when the language that women use does not overtly mark them as a woman, they may be able to subvert the typical result of the negative outcomes associated with that marker. Additionally, by not strictly adhering to using language traditionally associated with women, women are more likely to be heard and have their ideas considered (Tannen, 1995). This attention to women's input and insights has the ability to influence learning positively for all, so it is promising that the online environment might serve as a conduit for not downplaying women's contributions, notably in STEM environments, thus positioning women to have their ideas taken seriously. Although instructors may still need to account for other aspects of gendered discourse (like interrupting, using pregnant pauses, etc.) in synchronous discussions, this research shows promise that instructors do not need to worry that women in particular are unfairly being targeted because of the language that they use on discussion forums. Rather, these findings suggest that women can more freely form a sense of community.

Although not used strictly along gendered lines, gendered language was used within each course; the extent to which students used gendered language varied by course, though. Given the differing topics, course products, and grade allocations, there are many possible reasons why the courses were so different in their use of gendered language; future studies should investigate the effect of various course structures in order to understand what aspects are developing or hindering the use of more community-oriented language.

Of note is that students' posts in the chemistry course had higher Analytic ratings and greater use of Numbers than students in the astronomy course. We suspect that this finding is related to the assignments for the forum posts, where students often had to report their solutions.

Students in the chemistry course also used more community-building language than students in the astronomy course, by way of using more Social and Affiliation words as well as politeness indicators through the Discrepancy category. This indicates that students in the chemistry course may have felt more connected to one another than the students in the astronomy course. This is particularly interesting since the astronomy course had a greater emphasis on group work, with a group project constituting 37% of the final grade.

We also found differences in the average word length per post between the two courses, which implies a difference in the amount of information being shared. Although it is not possible for us to know why this was the case, we can hypothesize that because the astronomy course put more weight on discussion posts (25%) than the chemistry course (5%), this may have influenced students to write more. This can be tested in the future with other online courses that assign different weights to the discussion posts. It will be interesting to see whether the importance implied by the effect on final grade impacts students' length of posts in other courses in ways similar to those documented here. In addition, we are curious about the relation between length of post and student learning. This issue, too, will have to be taken up in future investigations. Work from Dowell et al. (2014) signals that there is a positive relationship, as they have found that using more complex, robust syntax and expository styles of writing when collaborating online leads to better learning outcomes.

Although we have no confirmatory data, we can hypothesize that, because the chemistry course and astronomy course were structured very differently, the course structures may have influenced differential use of gendered language. For example, requiring students to respond to others' posts in the astronomy course (while only making it an option in the chemistry course) may have generated more dialogue and conversation between participants. Future investigations that have access to a larger number of courses with structures similar to the chemistry course and astronomy course could begin to unravel the relation between how course structures influence students' use of gendered language forms.

From this study, it is unclear whether particular language choices were more productive (in terms of final grade, course retention, feelings of belonging, etc.); it is worth asking whether and how students' language could be tailored to the discussion forum's structure. For example, it may be advantageous to use more communal language and let one's guard down when posting questions for others to answer than when posting replies to others' questions. Using phrases like "I think" and "maybe" imply uncertainty, but, when asking questions, students indeed are often uncertain and that may be why they are seeking help. By using language that promotes them coming across as rather humble, these students may attract more students who are willing to answer their questions, which likely bodes well for productive dialogue and community building. On the other hand, using words associated with certainty may be useful when posting a solution or an answer to someone's question. Providing a post with certainty may give the content more credibility to the other students who may be learning from it, doing so may encourage those students to continue to ask questions when they know they will be getting solid, confident responses. This, in turn, may promote dialogue and community. Thus, future work should investigate the extent to which differences in language use are related to the ways in which the discussion forums are structured, as well as the extent to which certain types of language forms (e.g., hedging vs. certainty) are related to outcomes of success (e.g., final grade, course retention, feelings of belonging, etc.) within and between various discussion forum structures. Doing so can inform the development of language feedback systems, which can be used in real-time to alter group dynamics for the better (Tausczik & Pennebaker, 2013).

Of course, the structure of the course may be only loosely related or even unrelated to language choices. Rather, the course itself may be more relevant for dictating the language used. This study compared an introductory-level science course (chemistry), which is required for an array of majors, to a more advanced-level science course (astronomy), which is an elective for several STEM majors. Compared to the astronomy course, the chemistry course focused on more foundational material that potentially covered material that some students may have learned in high school. Thus, many students in the chemistry course may have been more confident in their knowledge of the material and thus may have been more willing to share their knowledge in a helpful, communal, collegial way. Future work should compare several introductory-level STEM courses with one another and several upper-level STEM courses with one another, to see if the language varies as much as it did in this study, to begin to tease out whether the differences observed in this investigation might be a function of the level of the course.

Furthermore, examining the interplay of students' backgrounds (e.g., previous coursework, reasons for taking the course, ACT score, GPA, experience with technology, etc.) will also be valuable, as these variables may very well moderate language use. For example, students who have taken a similar course previously likely may engage in more information sharing. These factors all contribute to each students' relative status, and the pairing of different statuses when collaborating can yield very different outcomes. It also affects who responds to whom and in what ways, questions of which are important extensions of this work that future studies should explore. In addition to these types of background factors, of note is that there are other markers of gender besides language (e.g., one's name or a picture) that may be present and thus can affect interactions.

The ratio of men and women enrolled in a course is also important to track in future work. The sample from the chemistry course in this study was fairly balanced in terms of the number of men vs. women enrolled; the astronomy course, however, had very few women compared to men. Even so, it seems that gendered language was not used along gendered lines; seeing if this pattern holds for all types of course dynamics is important. Even if the gendered patterns of gendered language hold, there could be other implications for the discussion forums, such as getting questions answered or feeling vulnerable enough to ask questions in the first place.

Along these lines, it is important to note that this study is limited in that it treats gender in a binary manner. This is partially due to how the data were collected. Background data was taken from students' registration paperwork with the university; when registering, students may opt to identify as male, female, or neither. There were only four students across both data sets who did so, making their identities impossible to conceal and yielding too small of a sample from which to derive conclusions. Future studies should actively investigate all gender identities.

Lastly, it is important to note that the median counts for the psychosocial categories were quite low, particularly for the astronomy course. Seemingly words associated with social processes, discrepancy, tentativeness, certainty, and affiliation do not occur naturally in this context. As such, instructors' promotion of the intentional use of language associated with community building, especially social processes and affiliation, may aid in community building efforts.

5 Conclusion

That men and women did not use gendered language across gendered lines is somewhat surprising but is also promising. The lack of differential use of gendered language between men and women means, at least for these courses, that this marker of gender did not differentiate men from women, and thus did not advantage or disadvantage students based on their gender in the online environment. This suggests that the context of computer-based learning might provide some protection from threats that are experienced by women in face-to-face educational contexts. By requiring students to participate, instructors remove barriers that may be difficult to overcome without this requirement and because the participation is posting to an online forum, women may feel more anonymous in this relatively neutral context compared to face-to-face STEM classroom contexts, when they may be in the visible minority. Furthermore, it is possible that, without gendered language to distinguish men from women, at least one potential barrier for women to succeed in the online environment was absent.

Strikingly, even though there were no gender differences in use of gendered language, we found differences between the two courses examined here. The language differences between the two courses highlight the need to study the role of language when analyzing online course discussion forums; researchers and practitioners alike need a better understanding of language's influence in the online environment so that discussion forums, which are the area of the online environment most crucial to community building, can be as productive as possible for all students.

Acknowledgements

This work was supported by the Institute of Education Sciences [grant number R305A180211].

References

Barron, B. (2000). Achieving coordination in collaborative problem-solving groups. *The Journal of the Learning Sciences*, *9*, 403-436.

Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences*, 12, 307-359.

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

Bell, N., Grossen, M., & Perret-Clermont, A. N. (1985). Sociocognitive conflict and intellectual growth. *New Directions for Child and Adolescent Development*, 29, 41-54.

Blickenstaff, J. C. (2005). Women and science careers: Leaky pipeline or gender filter? *Gender and Education*, 17, 369-386.

Buchner, A., Erdfelder, E., Faul, F.,*Lang, A.-G*. (2021). G*Power. https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower.

Cade, W. L., Dowell, N. M., Tausczik, Y. R., Pennebaker, J. W., & Graesser, A. C. (2014). Modeling students socioaffective responses to group interactions in a collaborative online chat environment. In J. Stamper, Z. Pardos, M. Mavrikis, B. M. McLaren (Eds.), *Proceedings of the 7th International Conference on Educational Data Mining* (pp.399-400). Worcester, MA: International Educational Data Mining Society.

Canary, D. J., & Dindia, K. (2009). Sex differences and similarities in communication. *New York: Routledge*.

Cho, H., Gay, G., Davidson, B., & Ingraffea, A. (2007). Social networks, communication styles, and learning performance in a CSCL community. *Computers & Education*, 49(2), 309–329.

Cochran, J. D., Campbell, S. M., Baker, H. M., & Leeds, E. M. (2014). The role of student characteristics in predicting retention in online courses. *Research in Higher Education*, 55, 27-48.

Cohn, M. A., Mehl, M. R., & Pennebaker, J. W. (2004). Linguistic markers of psychological change surrounding September 11, 2001. *Psychological Science*, 15, 687-693

Corbett, C., & Hill, C. (2015). Solving the equation: The variables for women's success in engineering and computing. *DC: AAUW*.

Correll, S. J. (2001). Gender and the career choice process: The role of biased self-assessments. *American Journal of Sociology*, 106, 1691-1730.

Crues, R.W., Henricks, G.M., Perry, M., Bhat, S., Anderson, C.J., Shaik, N., & Angrave, L., (2018). How do gender, learning goals, and forum participation predict persistence in a computer science MOOC? *ACM Transactions on Computing Education*, 18.

Dowell, N. M. M., & Graesser, A. C. (2014). Modeling Learners' Cognitive, Affective, and Social Processes through Language and Discourse. *Journal of Learning Analytics*, 1(3), 183-186. https://doi.org/10.18608/jla.2014.13.18

Drescher, A., & Schultheiss, O. C. (2016). Meta-analytic evidence for higher implicit affiliation and intimacy motivation scores in women, compared to men. *Journal of Research in Personality*, 64, 1-10.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149-1160.

Foreman, E. A., & Cazden, C. B. (1985). Exploring Vygotskian perspectives in education. In J.V. Wertsch (Ed.), Culture, communication, and cognition: *Vygotskian perspectives*, 323-347.

Freeman, T. M., Anderman, L. H., & Jensen, J. M. (2007). Sense of belonging in college freshmen at the classroom and campus levels. *The Journal of Experimental Education*, 75, 203-220.

George-Jackson, C.E. (2011). STEM switching: Examining "departures" of undergraduate women in STEM fields. *Journal of Women and Minorities in Science and Engineering*, 17, 149-171.

Guiller, J., & Durndell, A. (2007). Students' linguistic behaviour in online discussion groups: Does gender matter? *Computers in Human Behavior*, 23, 2240-2255.

Gurria, A. (2016). PISA 2015 results in focus. *PISA in Focus*, (67), 1.

Hachey, A. C., Wladis, C., & Conway, K. (2015). Prior online course experience and GPA as predictors of subsequent online STEM course outcomes. *The Internet and Higher Education*, 25, 11-17.

Hall, J. A., & Matsumoto, D. (2004). Gender differences in judgments of multiple emotions from facial expressions. *Emotion,* 4, 201-206.

Hechtman, L. A., Moore, N. P., Schulkey, C. E., Miklos, A. C., Calcagno, A. M., Aragon, R., & Greenberg, J. H. (2018). NIH funding longevity by gender. *Proceedings of the National Academy of Sciences*, 115, 7943-7948.

Heilbronner, N. N. (2013). The STEM pathway for women: What has changed? Gifted Child Quarterly, 57, 39-55.

Hill, C., Corbett, C., St. Rose, A., & American Association of University Women (2010). Why so few? Women in science, technology, engineering, and mathematics. *American Association of University Women*. 1111 Sixteenth Street NW, Washington, DC 20036.

Herring, S., Bucholtz, M., Liang, A., Suttoj, L., & Hines, C. (1994). Politeness in computer culture: Why women thank and men flame. Cultural performances: *Proceedings of the Third Berkeley Women and Language Conference*, 278–294.

Herring, S. (1996). Gender and democracy in computer-mediated communication. In Rob King (Ed.), Computerization and controversy: *Value conflicts and social choices (2nd ed.)* (pp. 476-489). San Diego: Academic Press.

Huang, W. H. D., Hood, D. W., & Yoo, S. J. (2013). Gender divide and acceptance of collaborative Web 2.0 applications for learning in higher education. *The Internet and Higher Education*, 16, 57-65.

Jackson, D. L., Starobin, S. S., & Laanan, F. S. (2013). The shared experiences: Facilitating successful transfer of women and underrepresented minorities in STEM fields. *New Directions for Higher Education*, 162, 69-76.

Kacewicz, E., Pennebaker, J. W., Davis, M., Jeon, M., & Graesser, A. C. (2014). Pronoun use reflects standings in social hierarchies. *Journal of Language and Social Psychology*, 33, 125-143.

Kiesler, S., Siegel, J., & McGuire, T. W. (1984). Social psychological aspects of computer-mediated communication. *American Psychologist*, 39, 1123-1134.

Liu, X., Magjuka, R. J., Bonk, C. J., & Lee, S. H. (2007). Does sense of community matter? An examination of participants' perceptions of building learning communities in online courses. *Quarterly Review of Distance Education*, 8, 9-24.

Margolis, J. & Fisher, A. (2002). Unlocking the clubhouse: *Women in computing*. Cambridge, MA: MIT Press.

McConnell-Ginet, S. (2010). Gender, sexuality, and meaning: Linguistic practice and politics. *Oxford University Press*.

McNamara, D. S. and Graesser, A. C. (2012). Coh-Metrix: An automated tool for theoretical and applied natural language processing. In P. M. McCarthy and C. Boonthum (Eds.), *Applied natural language processing: Identification, investigation, and resolution*. IGI Global, Hershey, Pennsylvania.

Miller, D. I., Eagly, A. H., & Linn, M. C. (2015). Women's representation in science predicts national gender-science stereotypes: Evidence from 66 nations. *Journal of Educational Psychology*, 107(3), 631.

Myers, C. B., & Pavel, D. M. (2011). Underrepresented students in STEM: The transition from undergraduate to graduate programs. *Journal of Diversity in Higher Education*, 4, 90-105.National Science Foundation, National Center for Science and Engineering Statistics (2019).

Women, Minorities, and Persons with Disabilities in Science and Engineering: 2019. *Special Report NSF 19-304. Alexandria*, VA. Available at https://www.nsf.gov/statistics/wmpd.

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*, 29, 665-675.

Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008). Gender differences in language use: An analysis of 14,000 text samples. Discourse Processes, 45, 211-236.*National Math and Science Initiative*. (2014). National Math and Science Initiative Annual Report 2014. Retrieved from https://www.nms.org/getmedia/5d1446e4-2cd6-408e-b411-5b32471e22e9/2014-Annual-Report.pdf.aspx?ext=.pdf

Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008). Gender differences in language use: An analysis of 14,000 text samples. *Discourse Processes*, 45, 211-236.

Noonan, R. (2017). STEM Jobs: 2017 Update (*ESA Issue Brief# 02-17).* Washington, DC: US Department of Commerce, Economics and Statistics Administration, Office of the Chief Economist.

Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., & Beaver, D. I. (2014). When small words foretell academic success: The case of college admissions essays. *PloS One*, 9, 110.

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). The development and psychometric properties of *LIWC2015*.

Prinsen, F.R., Volman, M.L.L, & Terwel, J. (2007). Gender-related differences in computer-mediated communication and computer-supported collaborative learning. *Journal of Computer Assisted Learning*, 23, 393-409.

Salter, N. P., & Conneely, M. R. (2015). Structured and unstructured discussion forums as tools for student engagement. *Computers in Human Behavior*, 46, 18-25.

Smith, B. (2010). E-learning technologies: A comparative study of adult learners enrolled on blended and online campuses engaging in a virtual classroom (Doctoral dissertation). Retrieved from *ProQuest Dissertations and Theses database.*

Tan, M. & Shao, P. (2015). Prediction of student dropout in E-learning program through the use of machine learning method. *International Journal of Emerging Technologies in Learning*, 10(1), 11–17.

Tannen, D. (1990). You just don't understand: Men and women in conversation. *New York: Morrow.*

Tannen, D. (1995). The power of talk: Who gets heard and why. *Harvard Business Review*, 73, 138-148.

Tannen, D. (2007). Talking voices: Repetition, dialogue, and imagery in conversational discourse. *Cambridge University Press*.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29, 24-54.

Tausczik, Y. R., & Pennebaker, J. W. (2013, April). Improving teamwork using real-time language feedback. In Proceedings of *the SIGCHI Conference on Human Factors in Computing Systems* (pp. 459-468).

Tu, C. H., & Corry, M. (2001). A paradigm shift for online community research. *Distance Education*, 22, 245-263.

Tyler-Smith, K. (2006). Early attrition among online e-learners: a review of factors that contribute to dropout, withdrawal and non-completion rates of adult learners undertaking e-learning programs. *Journal of Online Learning and Technology*, 2(2).

Wai, J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2010). Accomplishment in science, technology, engineering, and mathematics (STEM) and its relation to STEM educational dose: A 25-year longitudinal study. *Journal of Educational Psychology*, 102(4), 860.

Wladis, C., Conway, K. M., & Hachey, A. C. (2015). The online STEM classroom—who succeeds? An exploration of the impact of ethnicity, gender, and non-traditional student characteristics in the community college context. *Community College Review*, 43, 142-164.

Xu, Y. J. (2008). Gender disparity in STEM disciplines: A study of faculty attrition and turnover intentions. *Research in Higher Education*, 49, 607-624.

Xu, D., & Jaggars, S. S. (2014). Performance gaps between online and face-to-face courses: Differences across types of students and academic subject areas. *Journal of Higher Education*, 85, 633-659.

Yates, S. J. (1997). Gender, identity and CMC. *Journal of Computer Assisted Learning*, 13, 281–290.

Yuan, J., & Kim, C. (2014). Guidelines for facilitating the development of learning communities in online courses. *Journal of Computer Assisted Learning*, 30, 220-232.