

Talk with Me: Student Behaviours and Pronoun Use as Indicators of Discourse Health across Facilitation Methods

Carrie Demmans Epp

Learning Research & Development Center, University of Pittsburgh, USA EdTeKLA Group, Dept. of Computing Science, University of Alberta, Canada cdemmansepp@ualberta.ca

Krystle Phirangee

Ontario Institute for Studies in Education
University of Toronto, Canada

Jim Hewitt

Ontario Institute for Studies in Education University of Toronto, Canada

ABSTRACT: Identifying which online behaviours and interactions are associated with student perceptions of being supported will enable a deeper understanding of how those activities contribute to learning experiences. Student language is one aspect of their interaction in need of greater exploration within discourse-based online learning environments. As a result, personal pronoun usage is explored within online courses where the primary learning mechanism is student discourse. Student posting behaviours and features of their language, specifically their use of personal pronouns and academic vocabulary, are explored from the perspective of how supported students felt in their courses as well as how their communication behaviours and pronoun usage changed over time. Since prior work has shown the relationship between students' sense of support and how instructors choose to manage online discourse, the facilitation methods employed within courses were considered. Findings show those taking instructor-facilitated courses exhibited higher interaction levels, behaved more consistently from week to week, and used personal pronouns more. They also felt more supported. In contrast, those enrolled in peer-facilitated courses felt less supported, were less active, and used fewer personal pronouns. These differences in behaviour, language use, and experience suggest the potential for temporal analytics of student behaviours and pronoun use.

Keywords: Online learning, temporal analytics, learning analytics, student discourse, academic vocabulary, pronouns

NOTES FOR PRACTICE

 While many aspects of how students interact in discussion-based online courses are understood, we do not yet understand how student language use changes over time.



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

- This study shows how student behaviours and language use vary across the term and how these differences are associated with students' feelings of being supported as well as how their online discussion was facilitated.
- Differences in student behaviours and language use were associated with the facilitation method instructors adopted and imply that instructors should facilitate student discussions.
- Data suggest other forms of activity (e.g., reading) should be considered alongside student writing as an indicator of their participation in online courses.
- Differences in student use of pronouns suggest it may be worth monitoring their use of personal pronouns to ensure that they are talking with one another and acknowledging each other's ideas.
- Improvements in student communication patterns following instructor feedback suggests instructors should provide feedback early in the term (ideally within the first week or two).

1. INTRODUCTION

For several decades, researchers have experimented with different types of metrics for monitoring the health of online discussions. Most early efforts focused on tools that attempted to quantify participation, engagement, or interaction using simple counts, such as the amount of time spent online (Shih, Munoz, & Sanchez, 2006); numbers of posts (e.g., Davie, 1988; Guzdial, 1997; Hiltz, 1986, 1994; Hammond, 1999; Vrasidas & McIsaac, 1999); or number of replies over time (e.g., Ahern, Peck, & Laycock, 1992; Davie, 1988; Hiltz, 1994). These limited metrics provide a rough indication of the level of student involvement but they only consider the frequency of discourse events rather than the content of the discourse, which limits their usefulness.

To more deeply understand the pedagogical processes taking place in online courses, researchers have begun to study how the text of online discussions might be processed to identify potentially useful information. Recent studies have revealed that the text used in online discussions and the discursive moves made in different situations, can be tied to student learning processes (Akyol & Garrison, 2008; Baker, 2010; Wise, Hsiao, Marbouti, & Zhao, 2012; Wise, Hausknecht, & Zhao, 2014), student experiences (Baker, 2010; Joksimović et al., 2015), and learning outcomes (Dowell et al., 2015). Certain features of online text may also yield information relating to community building and online relationships since the language that we use when interacting with others influences how they perceive us and how they relate to us as speakers (Jaworski & Coupland, 1999). For example, Graham (2007) found the verbal behaviour of moderators affects student language use and participation, and Dowell et al. (2015) showed language use relates to student position and performance within a massive open online course. At a deeper level, Daniel, Schweir, and Ross (2005) used manual coding to examine how student reflection, information sharing, argumentation, and inquiry supported learning in online communities of practice. Such studies have extended our understanding of how text might be analyzed to uncover indicators of student experience and learning in online courses. Unfortunately, many of these methods rely on extensive



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

manual analysis, making it difficult to turn them into automated tools that provide timely diagnostic information.

One of the challenges of automating the analysis of online discussions is that different online courses use different subject-matter vocabularies. Danescu-Niculescu-Mizil, Lee, Pang, and Kleinberg (2012) observed that "finding domain-independent language-based signals has been a challenge" (p. 699). Attempts to use basic discourse markers have shown some promise, but they generally require the inclusion of domain-specific language to achieve accuracies that would allow the techniques to be used in active online learning environments (OLE) (Chandrasekaran, Demmans Epp, Kan, & Litman, 2017; Friedberg, Litman, & Paletz, 2012). However, there is a domain-independent linguistic feature that holds the potential to indicate aspects of group processes, learning experiences, and social presence. That feature is student use of pronouns (Baker, 2010; Friedberg et al., 2012).

Using personal pronoun data for diagnostic purposes has intuitive appeal. Personal pronouns (e.g., you, she, they, and us) tend to refer to other actors within an online community and thus denote mutual awareness and engagement. Personal pronoun use has been found to indicate which elements students are attending to, their cognitive state, and their social position within an OLE (Kacewicz, Pennebaker, Davis, Jeon, & Graesser, 2014). Thompson, Kennedy-Clark, Kelly, and Wheeler (2013) suggest it may be possible to use pronoun metrics as indicators of progress in online collaborative work. They posit that increases and decreases in the use of we, a first-person plural pronoun, can be indicative of changes in focus (i.e., focusing on individual vs. group concerns). Thus, differences in patterns of pronoun use may indicate differences in the quality of collaboration. Others have argued that the use of inclusive personal pronouns (e.g., we) may reduce the psychological and social distance between participants (Baker, 2010; Royai & Barnum, 2007) by contributing to student perceptions of sharing a common space (Kramer, Oh, & Fussell, 2006), which can lead to improved learning experiences. The use of inclusive pronouns (e.g., we, our) is further argued to contribute to a sense of group membership (Arbaugh, 2004; Michinov, Michinov, & Toczek-Capelle, 2004) and cohesion (Ubon & Kimble, 2004; Conaway, Easton, & Schmidt, 2005). Similarly, the use of possessive personal pronouns may indicate community development (Hughes, Ventura, & Dando, 2007), and some classes of personal pronouns (e.g., collective pronouns: we, they) may indicate the amount of support students received when other discourse behaviours, such as number of posts and discourse facilitation procedures, are taken into account (Demmans Epp, Phirangee, & Hewitt, 2017).

While all of the above work increases our understanding of how pronouns might influence student learning and learning experiences, it does not address the inherently temporal nature of language. Rather, the developed analysis procedures investigate learning outcomes or aspects of the learning process following course completion. This lack of automated temporal analytics that describe how student discourse changes over a term, in response to the learning environment, means our understanding of student language use and its relationship to learning experience could be better understood. This lack of



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

understanding of how student language use changes over time is of concern when considering how people automatically adjust their use of language in response to the persons with whom they are communicating (Doyle, Yurovsky, & Frank, 2016). This adjustment in language use occurs rapidly and predicts task success when conversational partners increase the similarity in the language they use (Friedberg et al., 2012). These types of measures are temporal in nature because they look for change over time, but their use to monitor discourse in online courses has yet to be studied since the ability to measure these changes has only been validated for small groups (e.g., dyads), with some work exploring language use in groups with as many as four members (Rahimi, Kumar, Litman, Paletz, & Yu, 2017). Taking guidance from those who have explored student participation in online discourse, this research aims to explore student use of pronouns as the course unfolds. Like work that has identified term-wide behavioural patterns in student discourse participation (e.g., Vassileva et al., 2001) or material use (e.g., Brooks, Demmans Epp, Logan, & Greer, 2011), we aim to study changes in system usage behaviours in addition to identifying how language use varies over time with respect to the manner in which the discourse environment is facilitated or managed.

To study how student behaviours and language use might be employed as an analytic tool to inform ongoing course activities, we posed the following questions:

- RQ1. How does student pronoun use and participation differ based on the facilitation methods used in their online courses?
- RQ2. How does student participation and pronoun use vary over time?
- RQ3. How do changes in student pronoun use relate to their assigned learning tasks?
- RQ4. How do other features of student language use (e.g., post length and use of academic vocabulary) differ based on the facilitation methods used in their online courses?

Answering these questions will enable a deeper understanding of how student contributions to online discourse, student behaviours, and instructor choices for managing online discourse interact to enable learning. It will also provide insight into how these factors influence learning experiences so analytics can be developed to enable instructor monitoring of student contributions to determine whether students are behaving in a manner that is consistent with pedagogical goals. The temporal nature of the measures required to answer these questions will enable ongoing monitoring that should allow instructors to respond in a timely manner.

2. METHODS

This study adopted an explanatory sequential design (Creswell & Plano Clark, 2011). This two-phase approach begins with a quantitative phase that is followed by a qualitative phase. The purpose of the qualitative phase is to explain and further inform the quantitative results. Using this design allowed trends and relationships to be assessed using both quantitative and qualitative data to gain a conceptual and statistical understanding of student pronoun use and behaviours within discourse-based online courses.



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

2.1 Participants

Participating graduate students (N = 90) were taking online courses to fulfill program requirements. The five selected courses were offered through a college of education at a research-intensive North American university. Courses were selected according to specific criteria: fully online seminar-style courses, using the same online learning environment (Pepper), offered in the same term, and with different instructors. Lastly, the instructors and all students from each course were contacted to participate in interviews to ensure both perspectives were captured for each course.

2.2 Online Learning Environment

Pepper is a discourse-based OLE that aims to support the cognitive aspects of learners' online interactions as well as the largely ignored socio-emotional aspects of learning (Kreijns, Kirschner, & Vermeulen, 2013). Pepper resembles many other OLEs at a surface level. It features course-wide discussion boards, an internal mail system, directed private chat, a learning resource area, private spaces for group discussions, calendars, announcements, assignment drop-boxes, and integrated analytic tools that allow instructors to monitor student activity in the course. In addition to these conventional tools, Pepper contains features designed to foster community and encourage students to support one another socially. Many of these features were inspired by social media environments like Facebook. These features include publically shared personal profiles, a "like" button, profile photos, emoticon support in the editor, and the ability to see other Pepper users who are co-present in the same virtual space. Upon entering a course, the user is notified of any social activity relating to their posts since their last login (e.g., Eli replied to your post; Zoe created a link to your post; or Joshua liked your post). In addition, a tool that displays a network of the learner's closest associates in each course allows students to monitor their frequent and infrequent collaborators. By merging the functionality of an OLE with some of the functionality of social media environments, Pepper aspires to better support both the cognitive and socio-emotional dimensions of online learning.

2.3 Data Collection and Analysis

The majority of data were collected through the OLE. Pepper's automated logging provided behavioural data. Student posts (notes) were extracted and measures of linguistic features characterizing student notes were obtained by applying natural language processing (NLP) techniques. Measures of pronoun use are unitized as a percentage rather than the number of pronouns per 1000 words because instructors and students are more accustomed to working with percentages and they are the intended users of analytics. The following measures were used:

- Participation: students who performed any activity (e.g., reading, writing, liking, sending a private message) were considered to have participated
- Number of notes written
- Number of notes read per student



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

- Likes per note
- Words per note
- AWL: the percent of words in each note that can be found in Coxhead's (2000) academic word list.
- Pronouns: the percent of words in each note that are personal pronouns
- I-class pronouns: the percent of first-person singular pronouns in a note (e.g., I've, me, or mine)
- You-class pronouns: the percent of second-person plural or singular pronouns in a note (e.g., yours or you)
- We-class pronouns: the percent of first-person plural pronouns in a note (e.g., us, let's, or we're)
- S/He-class pronouns: the percent of third-person singular pronouns in a note (e.g., she, oneself, or he'd)
- They-class pronouns: the percent of third-person plural pronouns in a note (e.g., their, they, or they'll)

These measures were calculated for the term as a whole and each week in the term. It is worth noting the week in the term is based on calendar dates. This definition of week may not align perfectly with a course's week. For example, an instructor might assign a discussion deadline of Wednesday, meaning that each week in the syllabus crosses two calendar weeks. However, this calendar definition of a week is sufficient when considering general trends in student behaviour as long as the potential for increased variability around the beginning and end of term as well as the term break is considered.

To obtain information about student perceptions of the amount of support they received, Rovai's (2002) classroom community scale (CCS) was distributed because it assesses the level of social and learning support that students feel they experienced. This widely used 20-item instrument uses a five-point Likert-scale to elicit student responses to items such as "I feel that it is hard to get help when I have a question" and "I feel I can rely on others in this course." Based on the Cronbach's alpha values reported in the original validation study (Rovai, 2002), it appears to be a reliable measure of classroom community (α = .93) that consists of two subscales that measure social support (α = .92) and learning support (α = .87).

All measures were analyzed using standard statistical procedures. Mean (M) or median (Mdn) are reported as measures of central tendency and standard deviation (SD) or interquartile range (IQR) are reported as measures of variability based on the characteristics of the data, which includes its distribution. Since few measures met the assumptions of parametric tests (e.g., normality), Mann-Whitney tests were used to perform comparisons between groups instead of t-tests.

Since the manner in which a course is organized and managed influences student behaviour (Phirangee, Demmans Epp, & Hewitt, 2016), courses were categorized into groups. Given the courses were discourse-based seminars, the manner in which student discourse was facilitated became the primary lens through which course design was considered. Hew's (2015) definitions of facilitation methods were used to determine which approach was used in each course. According to Hew (2015), instructor facilitation involves "keeping the discussion on track, establishing ground rules and good discussant behaviour, helping students overcome technical problems, and asking questions to help participants understand a



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

particular issue or topic, or drawing students' attention to opposing perspectives" (Hew, 2015, p. 20). Whereas, peer facilitation involves "the use of students as peer facilitators in an online discussion," which allows students to collaboratively control the discussions (Hew, 2015, p. 21). Grouping courses in this way made the groups large and varied enough to enable meaningful comparisons.

As a complement to the data collected through the scale and system, interviews were conducted with a student from each course and all instructors. These semi-structured interviews lasted approximately 1 hour and focused on how student online behaviours and interactions supported or hindered their perceived comfort with online learning. Following transcription, interviews were analyzed using a constructivist grounded theory approach (Charmaz, 2010). Using the *inductive approach*, codes were combined to identify themes, from both instructors and students, to accurately depict the data. The inductive approach was used because it is data-driven, which means the data was coded without forcing it to fit researcher biases (Braun & Clarke, 2006). We also ensured that themes were semantic: they were identified based on what was written and only reflect surface meanings (Braun & Clarke, 2006). This approach helped ensure appropriate themes were selected because it prevents the researcher from misinterpreting participant views. Thus, this approach enabled us to obtain themes that accurately reflected and explained participant experiences.

3. RESULTS

Overall, the facilitation method employed in participating online courses was indicative of the amount of support students felt they received, their levels and type of participation, and their use of language. Closer inspection of student behaviours and artefacts from each week in the term revealed additional differences between the two groups (peer-facilitated and instructor-facilitated).

3.1 Perceived Support

The questionnaire was reliable (α = .901) and had a response rate of 54.4%: it was completed by 21 of the 39 students from peer-facilitated courses and 28 of the 51 students from instructor-facilitated courses.

Students in instructor-facilitated courses felt they received greater amounts of learning and social support than those in peer-facilitated courses (Table 1). As can be seen by the information in the Mann-Whitney columns, these differences are measurable and are unlikely to be due to chance (column p). The effect size (column r) indicates these differences were moderate (around .3) to large (around .5). This perception of being supported was reflected in the group's score for the CCS, which measures the health of their learning community, with those in instructor-facilitated courses reporting scores around 10% higher than those in peer-facilitated courses.

Similar to the differences observed for the CCS, those in instructor-facilitated courses experienced higher levels of both social and learning support. The biggest between-group difference was the amount of learning support students received. The above differences between the support students felt they



received in peer- and instructor-facilitated courses are consistent with prior research (Phirangee et al., 2016; Demmans Epp et al., 2017). As such, we continue to analyze the data using these groups with the understanding that students in instructor-facilitated courses felt more supported than those in peer-facilitated courses.

	Pe	er	Instru	uctor	Mann-Whitney			
	М	SD	M	SD	U	р	r	
Social Support	48	14.0	58	15.3	195.0	.045	.29	
Learning Support	70	9.8	83	13.5	125.0	.001	.49	
CCS	59	10.5	70	13.3	149.0	.003	.42	

3.2 Learner Participation

Consistent with student perceptions of receiving higher levels of social and learning support, Figure 1 shows a higher proportion of students from instructor-facilitated courses actively participating in the forum from week to week. Their participation levels were also more stable (above 80%) when class was in session (RQ1 and RQ2). Those in peer-facilitated courses appear to have had similar participation levels to those in instructor-facilitated courses during the first week of the term. Their participation levels then appear to follow a downward trend and remain below 80%. We also see greater variability in these peer-facilitated courses as time goes on: the coloured bar surrounding the line becomes wider. This reduction in participation occurs sooner than previously found (Akyol & Garrison, 2008). The one time when students in peer-facilitated courses participated more was during the term break, where around 80% of those in instructor-facilitated courses stopped participating. The section titled "Learner Artefacts: Week by Week" presents information about why and how students chose to participate across the term.

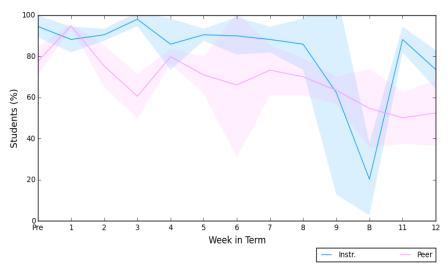


Figure 1. The percent of students from each group who performed some type of learning activity in a week (as *M* and *SD*). Pre = the week before classes started; B = the week of the term break.



3.3 Learner Artefacts: The Entire Term

Students in peer-facilitated courses wrote a total of 1,140 notes in contrast to the 4,075 notes written by those taking instructor-facilitated courses (RQ1; Table 2). The notes written by those taking peer-facilitated courses were moderately longer (RQ4). These students were also less likely to read peer contributions to the course (notes read/student row in Table 2). This large difference is characterized by students from peer-facilitated courses reading less than half of their peers' notes. In contrast, those in instructor-facilitated courses read more than half of their peers' notes even though their fellow students had posted considerably more. These behavioural differences resulted in those from peer-facilitated courses reading around 20 fewer notes than those in instructor-facilitated courses. This difference in reading habits means that those in peer-facilitated courses had less influence on their peers since the choice to not read fellow students' notes limited their exposure to the ideas and language presented in those notes.

Table 2. Student Writing Activities and Pronoun Use (as Percentage of Words) for the Whole Term

	Instructor			Peer				Mann-Whitney			
Measures	Mdn	IQR	Min	Max	Mdn	IQR	Min	Max	U	р	r
Text Composition											
Notes Written	55.0	49.0	4.0	166.0	21.0	25.0	12.0	80.0	303.0	< .001	.59
Words per Note	116.0	165.0	0	13,355.0	165.0	313.0	0.0	4,700	525.0	< .001	.40
Likes per Note	0	1.0	0	10.0	0	0	0	6.0	524.5	.001	.35
AWL (%)	7.0	5.8	0	50.0	7.4	5.0	0	33.3	866.0	.680	.04
Notes Read/Student	29.6	17.0	3.9	61.0	10.7	16.7	2.4	48.2	351.0	< .001	.57
Pronoun Usage	6.7	5.6	0	50.0	6.5	5.5	0	40	677.0	.010	.27
I-words	2.4	3.8	0	25.0	2.9	3.1	0	25	460.0	< .001	.46
You-words	0.5	2.0	0	50.0	0.9	2.2	0	33.3	814.0	.142	.15
S/He-words	0	0	0	18.4	0	0.3	0	13.2	723.0	.027	.23
We-words	0	1.2	0	20.0	0	0.9	0	8.9	716.0	.023	.25
They-words	0.5	1.9	0	25.0	0.8	1.7	0	8.9	700.0	.016	.25

Students in peer-facilitated courses were also moderately less likely to acknowledge peer contributions to the discussion by using the "like" feature. This is not to say that students from instructor-facilitated courses indiscriminately liked one another's notes; approximately one-quarter of student notes were liked and interviews with students revealed they enjoyed receiving likes from their peers because it helped to build a supportive environment and provided a quick way to value each other's contributions. As one student stated,

The first step is to create empathy. So you have to like the other ones, you have to thank the other ones, you have to share your experiences and your materials, websites or activities ... These types of behaviours I liked from my peers because it showed that they cared not only about their learning but mine too.



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

This behavioural difference cannot be easily attributed to student attitudes towards the liking feature since a student from the peer-facilitated group expressed a similar attitude. This student said:

Sometimes it was just reading a really good conversation that other people were having, so I would be liking some of those [notes]. Because they were really good and I was thinking you guys have this covered ... But they just had an amazing conversation and you learnt a lot from that and I don't necessarily have anything to add, so I would just like it.

Although students from instructor-facilitated courses and those from peer-facilitated courses differed moderately in the amount of liking they performed, they used it for similar reasons: to establish a supportive environment and value their peers' work. As the above quotes indicate, the "like" feature seems to provide an explicit indicator that students were listening when they did not speak.

Delving into the content of student-written notes, small to medium differences were detected (see Table 2) in their overall use of pronouns and student use of specific types of pronouns (RQ1). These differences should be considered in light of the weak negative relationship between note length and pronoun use (r =-.15, p < .001), which indicates students did not reduce their use of pronouns when writing shorter notes, as is sometimes seen in social media forums. Rather, those taking instructor-facilitated courses wrote shorter notes while using a greater proportion of pronouns in those notes. Those in instructorfacilitated courses also used inclusive pronouns slightly more (We-words) and used slightly fewer Theywords; the use of which would indicate someone was not a part of the group. Students' proportional use of pronoun classes also varied. Those taking instructor-facilitated courses dedicated 36% of their pronouns to I-words and 7% to They-words, whereas those taking peer-facilitated courses dedicated 45% of their pronouns to I-words and 12% to They-words. These patterns of pronoun use are consistent with other recent work which found differences in the strength and direction of relationships between student pronoun use and their CCS scores when using term-aggregate data (Demmans Epp et al., 2017). Given that people are known to adjust their behaviours and language to those in their surroundings (Doyle et al., 2016; Goffman, 1959), these factors were further investigated from the perspective of the consistency in student actions and language usage over time.

3.4 Learner Artefacts: Week by Week

To explore these term-wide differences further, we analyzed student activities and the artefacts they created within each week of the term. Similar patterns were found to those using the term-wide aggregate data (Figure 2), with students in instructor-facilitated courses writing a greater number of notes ($p \le .001$) each week when class was in (RQ2). These large to medium differences (r = .23 - .59) began in the week preceding the term and continued every week until week 9, re-appearing after the term break. The term break was the only week where students in peer-facilitated courses wrote more notes (U = .757.50, P = .031, V = .23). This behaviour was linked to a peer-facilitated course expectation that each student would reply to any note that someone had posted in response to a note they had authored themselves.



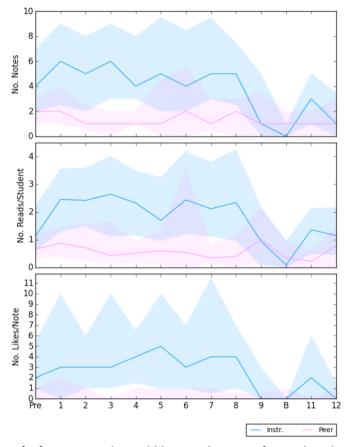


Figure 2. The number (No.) of notes, reads, and likes students performed each week (as *Mdn* and *IQR*).

Pre = the week before classes started; B = the week of the term break.

We can additionally see writing contributions to instructor-facilitated courses were consistent from week to week, excluding the period around the term break; that is, the blue band in Figure 2 remains relatively horizontal, with students typically writing between 2 and 8 notes each week. These students emphasized they consistently participated in online discussions because they wanted to make connections with their peers and learn from others' experiences. As one student stated,

I think for people to learn online together it's important to make connections so that could be connecting with ideas and also connecting with experiences. So, when we're teaching and everyone's a teacher we have similar experiences and that's one way we can connect to the ideas and the theories that we're learning about but also to each other.

In contrast, those in peer-facilitated courses wrote fewer notes each week but continued to post content during the term break. There was less variability in posting habits between students in peer-facilitated courses, where students typically wrote between 1 and 3 notes each week. These students emphasized that they became frustrated at times with the posting requirements given by the instructor because some



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

students only cared about meeting those requirements rather than creating meaningful notes. As a result, they opted not to reply to those peers even if it was required. Consider the following student statement,

There's some students where they post two weeks late for everything and you're like what are you doing? And at that point, they just care about filling the requirements, but even at that point most it's so late that they're getting zero marks ... You can tell who participates regularly. So, in the community people will respond to the regulars but the ones what are playing catch-up or do it out of habit nobody in the community will respond. Because it becomes frustrating, and it's a little disrespectful ... You're just coming in at the end and sticking your opinion in and not actually doing the work and are just faking it and so nobody responded to their post even if it was a requirement by the prof to do so.

Thus, students in instructor-facilitated online courses are more likely to post something when they want to connect with their classmates, which involved posting when they had something meaningful to contribute. Students in peer-facilitated courses produced fewer notes because they became frustrated with both the note requirements implemented by their instructor and some of their peers' approaches to meeting these requirements.

As Figure 2 shows, students taking instructor-facilitated courses read one another's notes more (p < .001; $.43 \le r \le .63$) and reduced their participation around the term break, at which point there was no measurable difference between group reading activities (p = .576 in week 9 and p = .924 in week 10). Consistent with the identified differences in student note writing and reading behaviours (RQ3), the higher explicit expression of support evidenced through student use of the liking function is present the entire time classes are in for those taking instructor-facilitated courses: p-values ranged from .023 to less than .001 and effect sizes were in the medium to large range ($.24 \le r \le .72$).

Those in peer-facilitated courses provided few behavioural indicators of explicit support. Interestingly, one of the peer facilitated course instructors revealed that she specifically spoke about how she used the "like" button so her students understood the role it played for her. She stated,

Normally every single night I read all posts everywhere ... Sometimes what I try to do is I click on the like button. I tell them the way I use the like button is different than the way they might be using it ... clicking like shows them that I've at least just read it. So, that's how I use the like button. I tell them that in advance.

Perhaps, knowing how an instructor used the "like" button and seeing this interaction unknowingly discouraged students in peer-facilitated courses from using the "like" button amongst each other. Students may have struggled with the idea of whether using the "like" button as they do on social media platforms would be academic, especially since their instructor was harnessing the "like" button for a different purpose.



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

While students in instructor-facilitated courses had more varied habits with respect to the number of notes they wrote, the length of their notes and the percentage of pronouns they used were more consistent than those written by students in peer-facilitated courses (Figure 3). The narrow blue band (Figure 3 - top) shows that students in instructor-facilitated courses typically wrote notes that ranged from 100 to 200 words in length, which is short enough to read on a smartphone screen without scrolling, making it unlikely that note length was a deterrent for others to read their posts.

The variability in the length of notes and use of pronouns in peer-facilitated courses from week to week does not allow them to be easily characterized (RQ2 and RQ4). However, we can see student notes are often far longer than those from instructor-facilitated courses (Figure 3 – top), with moderate to large differences being found in week 1 (U = 398.5, p < .001, r = .51), week 4 (U = 547.0, p < .001, r = .38), week 7 (U = 717.0, p = .024, r = .24), and the break (U = 647.5, p = .002, r = .33). Interestingly, the length of peer-facilitated student notes was shorter in week 11 (U = 670.0, p = .007, r = .28) than that observed in the instructor-facilitated courses; this difference is due to a decrease in the length of peer-facilitated student notes as the end of term approached.

These inconsistent differences indicate how assigned activities may have differentially influenced student writing behaviours in peer-facilitated courses (RQ3). It may also indicate some students felt a need to assume more of a leadership role when instructor involvement was lacking. Students from peer-facilitated courses emphasized that they were more focused on trying to meet their professor's note requirements when composing their notes. One student stated, "If I'm writing an initial note, like not a feedback note, then mainly I try to fulfill what's required and express my perspective." Similarly, another student said, "for full marks because I definitely want full marks [laughs], but I also think it's an opportunity to reflect on the readings or whatever the assignment and course requirements [are]." For students in peer-facilitated courses, meeting these requirements and receiving full participation marks strongly influenced their note writing behaviours. Perhaps, having strict note requirements pushed some students to spend more time producing more academic sounding notes rather than notes that supported student social and learning needs since this goal was not obviously and consistently encouraged by the professor.

As Figure 3 (middle) shows, few differences were found between the groups' academic vocabulary use: p-values for most of the term ranged from .071 to .841, largely because of the high within-group variability observed for peer-facilitated courses (RQ4). The variability in their academic vocabulary use also appears to increase as the term progresses. This variability may be indicative of some students struggling to determine whether they should use a more academic writing style, where personal pronoun use is often discouraged. As with many of the other measures, we see more within-group consistency from those taking instructor-facilitated courses. The only weeks where differences were found between facilitation methods were week 6 (U = 790.5, p = .049, r = .21), the break (U = 801.0, p = .035, r = .22), and week 11 (U = 645.0, p = .002, r = .33).



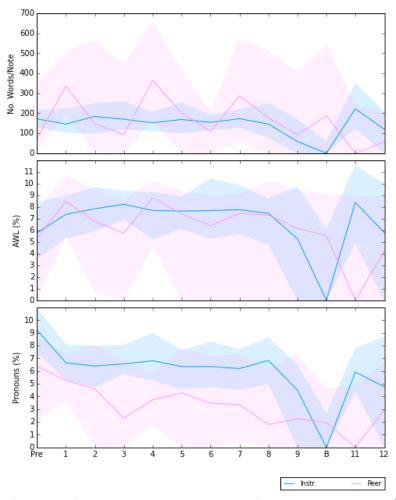


Figure 3. Linguistic features of student notes: the percent of academic vocabulary (AWL), percent of words that were pronouns, and length of their notes in number (No.) of words (as *Mdn* and *IQR*). Pre = the week before classes started; B = the week of the term break.

As Figure 3 (bottom) shows, the differences in student pronoun use are consistent with other measures of their behaviour (RQ2). The moderate to large difference in the level of pronoun use in week Pre (U = 534.0, p < .001, r = .40), week 2 (U = 742.0, p = .039, r = .22), week 3 (U = 532.0, p < .001, r = .43), week 4 (U = 497.0, p < .001, r = .0), week 6 (U = 661.0, p = .006, r = .29), week 7 (U = 717.0, p = .024, r = .24), week 8 (U = 584.0, p = .001, r = .35), and week 11 (U = 489.5, p < .001., r = .43) may be indicative of instructor involvement, encouraging students to reflect on their peers' thoughts and explicitly address the ideas of others within their notes. It may also be due to the activities they were assigned (RQ3), especially in the absence of direct instructor involvement. For instance, in one peer-facilitated course syllabus the instructor explicitly emphasized, "I also expect you to respond to classmates who have responded to your responses." The interview with the instructor revealed that he required students to respond to those who commented on their contributions to help students see the value in learning from each other. He stated,



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

My intention is to get students to learn what I think they need to learn and to do this by interacting with one another. So, in a sense I kind of force them to interact with one another, that they are required to make so many responses to other responses ... If someone comments on theirs ideas, they have to respond to them.

This requirement may have contributed to shallow student responses. Had deep, interactive responses been provided, we would expect to see more pronoun usage in the peer-facilitated courses because pronouns would be used to reference others' ideas and indicate author positions.

The higher use of personal pronouns observed in instructor-facilitated courses may be partly due to instructor guidance that was provided through the syllabus: "This will be a discussion-based course using online readings and resources in an asynchronous conferencing format. Collaborative learning, and particularly collaborative learning in a community of practice as a way to develop our understanding." Communities of practice (CoP) are defined as "groups of people who share a concern or a passion for something they do and learn how to do it better as they interact regularly" (Wenger, 2008, p. 1). CoPs have three characteristics: the domain, the community, and the practice (Wenger, 2008). The domain is the group identity, defined as the members' shared interests and common goals (Wenger, 2008; Eckert, 2006). The community refers to members engaging in joint activities, discussions, and information sharing while pursuing their interests, and practice refers to members developing "a shared repertoire of resources: experiences, stories, tools, ways of addressing recurring problems" (Wenger, 2008, p. 2) to learn and develop knowledge. According to Wenger, McDermott, and Snyder (2002), "people participate in communities for different reasons — some because the community directly provides value, some for the personal connection, and others for the opportunity to improve skills" (p. 27). Since this professor focused on collaborative learning within a CoP, it is important to recognize that there was an emphasis on reflecting on practice, which indirectly encourages the use of personal pronouns.

As part of the CoP process, students in this course were encouraged to contribute regularly to an individual learning journal. In the course outline, the instructor highlighted the importance of the learning journal, because it allows students to document, reflect on, and further develop their understanding; it gives them the opportunity to explore ideas not only in relation to the course but also as a graduate student and a professional. The journal is a combination of individual voice and community connections, and it is intended to enrich one's learning. As the instructor explained,

It makes people uncomfortable at first and sometimes people ask me if they could not do that. And I say sure and then over time I notice they start to and they do it at a point when they feel comfortable. When that happens, then other people comment on their work and so there is rather than enforced collaboration I see more spontaneous interactions around course artefacts, notes, and what have you that are meaningful to those people talking about it.

Although instructors from both instructor-facilitated and peer-facilitated courses encouraged student participation differently, all did so hoping that students would be willing to learn from each other by



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

sharing knowledge, experiences, and ideas. The differences in their requirements and how they communicated those requirements appears to have encouraged students from instructor-facilitated courses to use a higher proportion of personal pronouns and to more carefully attend to the content of their peers' notes.

Decomposing student pronoun use into the different types of pronouns (Figure 4) also shows differences between student language across facilitation methods (RQ1). The level of pronoun use is expected to vary with the intent of the post so a wide range of pronoun usage (minimum, mean, and maximum) is explored within each time unit.

If we consider the minimum percentage of pronouns that students used in each week (Figure 4 – top), we can see both groups' pronoun usage was quite low with neither group exceeding 4%, making the minimum usage statistic inappropriate for monitoring discourse health. On this end of the spectrum, I and You were the most used classes of pronouns in all courses, especially during the initial weeks where introducing oneself is a common activity. In peer-facilitated courses, They-class pronouns were used at similar levels to I- and You-class pronouns, indicating less of a focus on members of their course. This focus is consistent with student interviews, which indicated these students struggled with building a community in their online course because they did not understand its purpose. That is, they did not understand how having community would support their social and academic needs or whether building a community was academic. One student stated,

What's the value in a community and I think a lot of times how I see value and build a community and getting to know people is not included in what you're getting marked on. So, you're worried if it will conflict... Is the professor going to be like "what are you doing, you completely derailed from the purpose of the assignment." So, yeah you're building community but you're not talking about the readings anymore.

Similarly, another student expressed frustration over participating in a community because she never fully experienced it. She stated, "For me, having a community doesn't make a difference. Community doesn't change how and what I learn. [Pause] I don't get it. It doesn't matter for me or change how I feel."

All things considered, students in peer-facilitated courses opted to shift the way they communicated with each other when struggling with an idea, concept, or course reading (such as fostering a community) due to a perceived lack of guidance, involvement, and feedback from the instructor (RQ1).

Considering the more typical or mean use of pronouns (Figure 4 – middle), the general shape in student pronoun use was similar across facilitation methods (RQ1), which would be expected given the similarities in course structure. However, their levels of pronoun use differed on several dimensions. The first was the higher use of personal pronouns in instructor-facilitated courses. In this case, student pronoun use approached and occasionally met the 10% marker that the literature suggests should be expected during conversation (Doyle et al., 2016). Given these students felt supported, the data confirms this as an



approximate threshold that is indicative of students feeling supported while engaging in interactive discourse. In contrast, those who felt less supported were in peer-facilitated courses, where median pronoun use typically remained below the 5% mark.

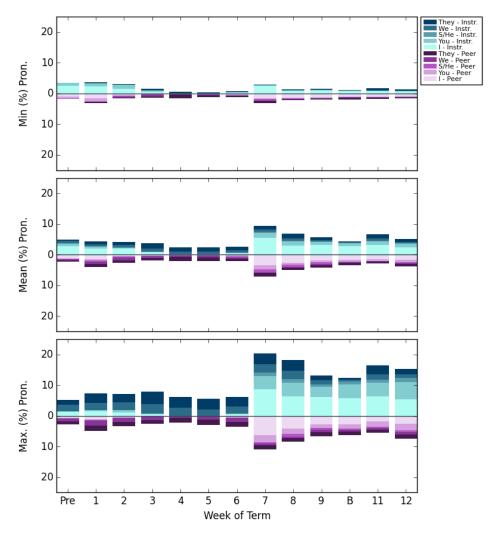


Figure 4. The lowest, mean, and highest percentage of each pronoun class used across the term.¹ Pre = the week before classes started; B = the week of the term break.

Beyond total pronoun level differences, the predominant differences in mean pronoun-class use between groups was the balance of pronoun classes being used. S/He class pronouns were hardly used within instructor-facilitated courses, especially at the beginning and end of the term: moderate differences were

¹ Note this chart represents the mean usage rather than the median even though the distribution of pronoun usage is non-normal.



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

found in the week before classes started (U = 758.0, p = .024, r = .24), week 3 (U = 670.0, p = .005, r = .29), week 9 (U = 759.0, p = .022, r = .24), the break (U = 798.5, p = .022, r = .24), and week 11 (U = 634.0, p = .001, r = .35). The chart also shows that those in instructor-facilitated courses were more consistent from week to week in directly addressing their peers using You-class pronouns. These students used a moderate to much ($.24 \le r \le .45$) greater proportion ($.001 \ge p \le .024$) of You-class pronouns from week 7 through the end of the course (RQ2). These differences between groups are indicative of students from instructor-facilitated courses attempting to communicate with others in their course. This observed linguistic feature is also consistent with instructor reports of trying to encourage collaboration rather than force responses.

On the top end of the pronoun usage spectrum (Figure 4 – bottom), a wider variety of pronoun classes was used from week to week and the proportional use of pronouns in instructor-facilitated courses increased more in week 7 (RQ2). At the top end of their pronoun use, these students approached or exceeded the 10% threshold whereas those in peer-facilitated courses only approached the 10% mark after instructors provided feedback around weeks 6 and 7. It is worth noting that the maximal use of pronouns by those in peer-facilitated courses closely resembled that of instructor-facilitated students' mean pronoun use. The high differentiability between groups suggests the measure of maximum pronoun use may be a better indicator, provided other analytics are also being used to support decision-making.

Beyond these level differences, students in instructor-facilitated courses used more You- and They-class pronouns than those in peer-facilitated courses, which is consistent with the increase in I-class pronoun usage and the linguistic alignment that is expected when people converse (Doyle et al., 2016): high self-reference by one speaker should be matched by the conversational partner making use of you. This difference is also consistent with the perceptions of students from peer-facilitated courses, where they felt a substantial portion of their peers did not truly engage in the discussion. The difference (.001 $\geq p \leq$.024) in They-class pronoun usage was present in all but week Pre (U = 792.0, p = .091, r = .17), week 9 (U = 517.5, p = .666, r = .41), and week 12 (U = 824.0, p = .146, r = .15), with moderate to large effect sizes (.22 $\leq r \leq$.49). Similarly large (.28 $\leq r \leq$.49) differences (.001 $\geq p \leq$.007) in You-class pronoun usage were present from week 9 forward.

Even though the levels of pronoun use vary, the overall patterns across these three visualizations remain similar with higher proportions of personal pronoun use and a greater variety of pronoun types being used in instructor-facilitated courses (RQ1). Looking across the Mean and Max sub-charts of Figure 4, an increase in pronoun use is noticeable at week 7 (RQ2). Closer inspection reveals this increase is driven by additional use of I-class pronouns.

Analysis of the syllabus from each course suggests students were working on a major assignment while participating in the online discussions. These major assignments were due in week 7. Instructors encouraged students to continue to share their experiences and relate those experiences to course materials in week 6, because it would help students to complete their assignments. During this week, instructors also provided feedback to students on their work thus far in the course. Such feedback seemed



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

to influence student participation in week 7, encouraging an increased focus on their personal experiences in relation to the course (RQ3). In addition, instructor interviews revealed that they were consistently emphasizing the importance of providing feedback to each other and accepting feedback from the instructor in hopes of making the assignments more personal. For example, one instructor stated,

So that it is more personal and I did ask upfront their background and alerting them that you all have similar backgrounds. It's okay to if you don't have a lot of experience doing self-assessment or program evaluation. So, I try to make it personal and comfortable.

Similarly, another instructor stated,

In the videos, the introductory videos, I try to talk about what my expectations are about the course, and how I understand individual differences and how this is something [pause] that their progress in the course is really about them. So, where they start and where they end, rather than having a single end point where everybody has to reach because some of them are doctoral students and some of them are MEd and some MA.

The interviews revealed the importance of sharing and including personal experiences in posts, thus encouraging pronoun use, to build commonality among peers. Students felt sharing was important because they believe it allows them to go deeper with the content and learn from each other. As one student from an instructor-facilitated course stated,

It's so important for that community to have something in common and for a community to develop. ... If we don't have something in common, we're not going to naturally go into deeper levels of communication and I think that's very important.

A student from a peer-facilitated course echoed the difficulty of meeting all of the posting requirements while emphasizing the need to balance these requirements and include personal experiences as much as possible. This student stated,

I've taken a few online classes now and usually in the discussion it's about building from your personal experience, evidence-based, and bring something from theory, and it's really easy to bring something from your personal experience but when you're trying to balance those three in every post you don't want to pull too much from one. It's like you don't want to pull too much from personal experience because they're more concerned about evidence-based or something like that but it's needed, you need to include it to learn other things.

This struggle to balance academic, experiential, and interpersonal aspects of online discourse may partially explain why language differences are seen between students' minimum, mean, and maximum pronoun use. However, the sharing of personal experiences allowed students from instructor-facilitated and peer-facilitated courses to make more meaningful connections to the readings while also learning from peer experiences. As one student explained,



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

People would summarize the theories or readings but what I liked to do was make a personal connection to what we were reading so I'm trying to make not only a connection to the theory but also use a personal example like from my teaching because I want to make it applicable and something other people can relate to.

Therefore, when students opted to share their personal experiences, they focused on doing so to make that personal connection to the course readings as well as provide a more meaningful perspective for their peers to relate to and learn from. These personal experiences allowed students to contextualize the readings and further nurture a more meaningful dialogue about course content with their peers.

Taking into consideration the syllabus and interview data, during weeks 6 and 7 we see students shifting their focus onto their assignments and instructors encouraging students, through feedback, to personalize their assignments by relating their personal experiences to the course. In instructor-facilitated courses, we also see increased student comfort with writing their personal learning journals over time, which would contribute to an increased use of I-class pronouns. Therefore, the additional use of personal pronouns in week 7 can be partially attributed to students focusing on their assignments and instructors encouraging students to tap into their personal experiences more.

4. LEARNING EXPERIENCE IMPLICATIONS

Students from instructor-facilitated courses experienced greater levels of social and learning support while consistently indicating their activities or explicitly supporting their peers through the liking feature. These types of indicators provide a proxy for peer listening behaviours when students do not necessarily have something to add to an ongoing conversation or one that is approaching completion. This type of recognition (Vassileva, McCalla, & Greer, 2016) and "group awareness is very important to successful collaboration [and] means that collaborators ought to be aware of each other's actions" (Brooks, Panesar, & Greer, 2006, p. 38), with the use of social awareness features being indicative of a well-functioning group (Brooks, Hansen, & Greer, 2006). The use of the "like" button enables students to monitor their fellow learners' participation so they can decide who to respond to in a similar way to the monitoring and exclusion of classmates reported by those in peer-facilitated courses: they used posting information to determine who was invested in the course and stopped replying to those who they felt were gaming the system. The use of this feature by those in instructor-facilitated courses may have increased student awareness of their peers' less visible participation, helping them to feel supported. This type of recognition and awareness are argued to decrease learner attrition, increase learner satisfaction, and contribute to their sense of being socially supported within online courses (Brooks, Panesar, & Greer, 2006).

The absence of student use of the "like" feature in peer-facilitated courses indicates that some form of training with respect to how this feature should be used in academic environments is needed. It is not enough to assume students will feel comfortable using a feature borrowed from social media in an



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

academic environment since the academic context can result in changes to the feature's semantics. This change in semantics is similar to how a physical location or technology can be repurposed to support different user needs (Dourish, 2003, 2006). The system feature or location does not determine how it can be used but the social configurations surrounding that feature and place interact to determine how the feature will be used (Dourish, 2006). Evidence of this conflict was seen in the interviews from peerfacilitated courses, which may partially explain why they used this feature so little. This conflict also highlights how it is important to maintain some of the semantics of the original feature. Otherwise, the repurposing of the "like" feature runs the risk of confusing students and their avoiding its use. Training could alleviate this problem and could come in the form of explicitly communicated expectations or instructor modelling of desired behaviours. This training will be needed in many cases and may need to be repeated over time, as has been recommended in other technology integration projects (Demmans Epp, 2016).

The positive changes in student communication patterns (see Figure 4), as represented through their use of different types of pronouns, following the receipt of instructor feedback more than a month into the term indicates the importance of providing that feedback. We would additionally argue that this feedback should be provided sooner so students can adjust their discourse. The finding that students in instructor-facilitated courses felt more supported while a greater proportion of them consistently participated from week to week (see Figure 1) confirms early findings from OLEs, where

knowledge investment by the "authorities" in a course seems to be an important influence on usage. In all three deployments to date, the participation of the instructor and of paid tutors (who initially contribute learning materials not available elsewhere, answer questions promptly, and make themselves available) is critical at the outset of a course to stimulate usage of the system. (Vassileva et al., 2001, p. 419)

This instructor involvement has been credited with contributing to the culture of use that develops within each course (Vassileva et al., 2001) and has been argued to engender increased student engagement (Vassileva et al., 2016) in a manner consistent with our findings. This includes reports of peer-facilitated students not responding to those who participated irregularly.

The need for this type of expectation setting early in the course (e.g., week 1 or the week before classes start) is further supported by research into how people communicate. People's use of pronouns and other linguistic features align quickly and people tend to align their language to the person who exercises power (Doyle et al., 2016). The consistency of pronoun and AWL vocabulary use by those in instructor-facilitated courses may be indicative of these known language-alignment patterns (RQ4). This type of vocabulary alignment or matching has also been found to distinguish successful groups from unsuccessful ones when collaborative learning approaches are used (Friedberg et al., 2012). The high variability observed in peer-facilitated courses may be indicative of a power vacuum being created by the instructor's absence.



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

Whether real or perceived, this absence means clear expectations about subtle aspects of communication were not set; meaning students had to navigate this space in the absence of a clear path or directions.

When considered alongside the literature, the behaviours and language use observed in this study imply instructor actions and language use will be at least partially replicated. This expected replication suggests that instructors who employ inclusive language and explicit references to the group or its members (e.g., by using personal pronouns) will implicitly encourage this behaviour. The associations found among the higher and more consistent use of pronouns, liking, instructor-facilitation, and a higher student perception of being supported suggest that instructor involvement will result in better learning experiences for students in online courses. The speed with which people align their language further suggests that instructor intervention could result in changes to the student experience within a short period, making students' mean and maximum use of pronouns a potentially interesting and useful analytic for instructors to monitor since it can be easily tracked and influenced through task design and instructor participation.

5. ANALYTICS IMPLICATIONS

From the perspective of how the applied measures might be used as analytics, several avenues have been expanded by this work.

The simplest measure that might be useful is a different form of one already employed in online discoursebased environments: student participation. In many systems, this is viewed through the lens of how many posts a student has written (Wang & Newlin, 2002) rather than the broader definition of participation used here. As our data shows, consistently high participation levels are associated with students feeling more supported regardless of the motivator behind this participation. Student reports from peerfacilitated courses also indicate online learners are more likely to interact with those who participate regularly. Similar to prior findings (Vassileva et al., 2001), student participation in high-activity courses did not fluctuate alongside assignment due dates, whereas student participation in low-activity courses seemed to more closely follow assigned activities. Interestingly, the requirement to reply to others' notes in peer-facilitated courses did not result in consistently high participation levels even though student grades were tied to their participation. In contrast, those in instructor-facilitated courses participated consistently even though no explicit requirements were communicated. This type of consistent learner activity has been linked to higher student grades across a variety of instructional domains (Brooks, Erickson, Greer, & Gutwin, 2014; Koedinger, Kim, Jia, McLaughlin, & Bier, 2015), which suggests its potential usefulness beyond indicating students' sense of being supported. Variability accompanied by a downward trend in weekly participation in peer-facilitated courses seems to be indicative of students not feeling supported. This pattern is consistent with tentative prior findings, where the listening patterns of students who disregarded peer contributions also followed a downward trend (Wise et al., 2014). However, further study is required to determine the amount of variability that indicates a degradation in the support students feel they have received. Methods for encouraging consistent and quality support also need further development and examination.



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

Along with a broad definition of participation, the data suggests that analytics describing student pronoun use are a reasonable proxy for the health of the online community. When using pronouns to monitor the health of a community, two aspects must be considered: the overall proportion of pronouns being used and the balance between the types of pronouns used. Looking at mean pronoun usage, an instructor would want to see pronoun levels approaching 10%. A better indicator might be for students' highest pronoun levels to exceed 10% with their mean use of pronouns above 5%. From the perspective of the balance of pronouns being used, we know that You- and I-class pronouns align differently than other types of pronouns (Doyle et al., 2016): when talking about yourself, you tend to use self-references (e.g., I or me) that encourage your speaking partner to respond using the second person (e.g., you). This alignment suggests an instructor should be concerned about the health of the discourse within a course if the proportion of I-class pronoun use is vastly different from that of You-class pronouns. The use of other classes of pronouns should be similar across students since they are known to align more directly. A potentially interesting analytic might highlight those who are not aligning with the group and whose activity patterns confirm the student's status as an outsider (Phirangee et al., 2016). The analytic could serve to alert instructors so targeted interventions could be performed.

In addition to these basic considerations, instructors using pronouns to monitor discourse in their online courses need to remember the interpretation of this analytic should be sensitive to the pedagogy employed (Chen, Wise, Knight, & Cheng, 2016). That is, student use of pronouns should be consistent with assigned tasks and course goals. For example, if one of the newly assigned tasks is for students to perform reflections, an instructor should expect to see an increase in the proportion of I- and We-class pronoun use. In a manner consistent with the data-assisted approach (Brooks, Greer, & Gutwin, 2014), the presentation of these analytics could involve supports that encourage instructor sensemaking. To ensure that instructors fully consider their pedagogy when interpreting the analytics, we suggest asking them to explicitly state their expectations before they are given access to the analytics so they can interpret them appropriately and respond in a manner that is consistent with their intentions and informed by both the data and the theory underlying their pedagogy.

6. LIMITATIONS

Given the relatively small scale of this study and the complex nature of conducting research in classroom settings, these results may not generalize to other contexts. While there is no evidence of differences in device usage across groups, it is possible that these or other factors contributed to some of the behavioural differences observed. However, the methods employed and information captured enabled a more in depth understanding of student behaviour and language use in online courses, which allows the exploration of how well these findings generalize to other contexts.



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

7. CONCLUSIONS

This investigation of student behaviour and language use across the term found that the facilitation methods employed in discourse-based online learning environments were associated with different learner behaviours and language use. These differences are further linked to the amount of support students felt they received, with those receiving higher levels of support making more conversational moves, acknowledging fellow student contributions to the learning community, and referencing one another using pronouns. The patterns of pronoun use and participation observed from week to week seem to be partly due to the facilitation method employed within the online course (RQ1 and RQ2): those taking instructor-facilitated courses participated more consistently from week to week, used proportionally more self-references (I-class pronouns) from week to week, and used proportionally more references to their peers (You-class pronouns) from the middle of the term to the end of the term. We also saw changes in student language use that appear to be related to the tasks assigned within their courses (RQ3): assignment due dates coincided with increases in student pronoun use, and the use of learning journals seemed to encourage the sharing of personal information, which all but requires the use of self-references (I-class pronouns).

The above analyses of student behavioural traces and student-created artefacts indicate the importance of monitoring student language use. These analyses also show the potential usefulness of simple measures of student language use. Now that we can see the relationship between students' feeling supported in online courses and their use of simple linguistic features (e.g., pronouns and AWL vocabulary), we can employ these measures as a proxy for student connectedness. This proxy is grounded in different theory than that reported by Cade and colleagues (2014), making it a potential complement to other approaches to measuring student connectedness. Moving forward, measures of changes in student vocabulary alignment (e.g., convergence) can be developed and tested to determine whether they provide additional information about the health of the discourse taking place in online courses.

Perhaps more importantly, these measures need to be integrated into an active system so we can explore how to support instructor use of simple temporal analytics. Ideally, instructors would be able to see the downward trend in student participation and intervene in a manner that encourages students to participate and relate with one another rather than reply to each other's posts for the sake of replying. Instructors might also see that students are not relating to each other, referencing one another's posts, or sharing personal experiences (as indicated through their use of pronouns) and assign activities that would indirectly encourage these behaviours. It may be helpful to allow instructors whose students are exhibiting productive discourse behaviours to anonymously share their approach to supporting student discourse with other instructors. Other instructors could then consult this bank of approaches when their course analytics indicate their students are not interacting effectively. Regardless of the approach taken, additional study of the influence that providing these analytics has on instructor praxis, student learning, and student-learning experiences is needed.



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

REFERENCES

- Ahern, T., Peck, K., & Laycock, M. (1992). The effects of teacher discourse in computer-mediated discussion. *Journal of Educational Computing Research*, 8(3), 291–309. http://dx.doi.org/10.2190/HFPW-JYR3-YMBE-0J7D
- Akyol, Z., & Garrison, D. R. (2008). The development of a community of inquiry over time in an online course: Understanding the progression and integration of social, cognitive and teaching presence. Journal of Asynchronous Learning Networks, 12(3–4), 3–22.
- Arbaugh, J. B. (2004). Learning to learn online: A study of perceptual changes between multiple online course experiences. *The Internet and Higher Education*, *7*(3), 169–182. http://dx.doi.org/10.1016/j.iheduc.2004.06.001
- Baker, C. (2010). The impact of instructor immediacy and presence for online student affective learning, cognition, and motivation. *Journal of Educators Online*, 7(1), n1.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. http://dx.doi.org/10.1191/1478088706qp063oa
- Brooks, C., Demmans Epp, C., Logan, G., & Greer, J. (2011). The who, what, when, and why of lecture capture. *Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (LAK '11), 27 February–1 March 2011, Banff, AB, Canada (pp. 86–92). New York: ACM. http://dx.doi.org/10.1145/2090116.2090128
- Brooks, C., Erickson, G., Greer, J., & Gutwin, C. (2014). Modelling and quantifying the behaviours of students in lecture capture environments. *Computers & Education*, 75, 282–292. http://dx.doi.org/10.1016/j.compedu.2014.03.002
- Brooks, C., Greer, J., & Gutwin, C. (2014). The data-assisted approach to building intelligent technology-enhanced learning environments. In J. A. Larusson & B. White (Eds.), *Learning Analytics* (pp. 123–156). New York: Springer. http://dx.doi.org/10.1007/978-1-4614-3305-7 7
- Brooks, C., Hansen, C., & Greer, J. (2006). Social awareness in the iHelp courses learning content management system. Paper presented at the Workshop on Social Navigation and Community-Based Adaptation Technologies at the 4th International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH 2006), 21–23 June 2006, Dublin, Ireland. Retrieved from http://www.pitt.edu/~paws/SNC_BAT06/crc/brooks.pdf
- Brooks, C., Panesar, R., & Greer, J. (2006). Awareness and collaboration in the iHelp courses content management system. In W. Nejdl & K. Tochtermann (Eds.), *Innovative approaches for learning and knowledge sharing* (pp. 34–44). Springer Berlin Heidelberg. http://dx.doi.org/10.1007/11876663 5
- Cade, W., Dowell, N., Graesser, A., Tausczik, Y., & Pennebaker, J. (2014). Modeling student socioaffective responses to group interactions in a collaborative online chat environment. In J. Stamper et al. (Eds.), *Proceedings of the 7th International Conference on Educational Data Mining* (EDM 2014), 4–7 July 2014, London, UK (pp. 399–400). International Educational Data Mining Society. Retrieved



http://educationaldatamining.org/EDM2014/uploads/procs2014/posters/75_EDM-2014-Poster.pdf

- Chandrasekaran, M. K., Demmans Epp, C., Kan, M.-Y., & Litman, D. (2017). Using discourse signals for robust instructor intervention prediction. *Proceedings of the 31st Conference on Artificial Intelligence* (AAAI-17), 4–9 February 2017, San Francisco, CA, USA (pp. 3415–3421). Palo Alto, CA: AAAI Press.
- Charmaz, K. (2010). Constructing grounded theory. London: SAGE Publications.
- Chen, B., Wise, A. F., Knight, S., & Cheng, B. H. (2016). Putting temporal analytics into practice: The 5th international workshop on temporality in learning data. *Proceedings of the 6th International Conference on Learning Analytics and Knowledge* (LAK '16), 25–29 April 2016, Edinburgh, UK (pp. 488–489). New York: ACM. http://dx.doi.org/10.1145/2883851.2883865
- Coxhead, A. (2000). A new academic word list. TESOL Quarterly, 34(2), 213-238. doi:10.2307/3587951
- Conaway, R. N., Easton, S. S., & Schmidt, W. V. (2005). Strategies for enhancing student interaction and immediacy in online courses. *Business Communication Quarterly*, *68*(1), 23–35. http://dx.doi.org/10.1177/1080569904273300
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research*. Thousand Oaks, CA: SAGE Publications.
- Danescu-Niculescu-Mizil, C., Lee, L., Pang, B., & Kleinberg, J. (2012). Echoes of power: Language effects and power differences in social interaction. *Proceedings of the 21st International Conference on World Wide Web* (WWW 2012), 16–20 April 2012, Lyon, France (pp. 699–708). New York: ACM. http://dx.doi.org/10.1145/2187836.2187931
- Daniel, B., Schweir, R., & Ross, H. (2005). Intentional and incidental discourse variables in a virtual learning community (pp. 1953–1965). Presented at the World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education (E-Learn), Chesapeake, VA: AACE. Retrieved from http://www.editlib.org/p/21483
- Davie, L. (1988). Facilitating adult learning through computer-mediated distance education. *Journal of Distance Education*, *3*(2), 55–69.
- Demmans Epp, C. (2016). Supporting English language learners with an adaptive mobile application (Doctoral thesis). University of Toronto, Toronto, ON, Canada.
- Demmans Epp, C., Phirangee, K., & Hewitt, J. (2017). The interplay between students' usage of pronouns and community levels in online courses. Presented at the American Educational Research Association (AERA) Annual Meeting. San Antonio, TX, USA.
- Dourish, P. (2003). The appropriation of interactive technologies: Some lessons from placeless documents. *Computer Supported Cooperative Work (CSCW)*, 12(4), 465–490. http://dx.doi.org/10.1023/A:1026149119426
- Dourish, P. (2006). Re-space-ing place: "Place" and "space" ten years on. *Proceedings of the 2006 Conference on Computer Supported Cooperative Work* (CSCW '06), 4–8 November 2006, Banff, Alberta, Canada (pp. 299–308). New York: ACM. http://dx.doi.org/10.1145/1180875.1180921



- (2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4
- Dowell, N., Krypnyk, O., Joksimović, S., Graesser, A. C., Dawson, S., Gašević, D., ... Kovanović, V. (2015). Modeling learners' social centrality and performance through language and discourse. In O. C. Santos et al. (Eds.), *Proceedings of the 8th International Conference on Educational Data Mining* (EDM 2015), 26–29 June 2015, Madrid, Spain. International Educational Data Mining Society. http://www.educationaldatamining.org/EDM2015/uploads/papers/paper 211.pdf
- Doyle, G., Yurovsky, D., & Frank, M. C. (2016). A robust framework for estimating linguistic alignment in Twitter conversations. *Proceedings of the 25th International Conference on World Wide Web* (WWW '16), 11–15 April 2016, Montreal, QC, Canada (pp. 637–648). New York: ACM. http://dx.doi.org/10.1145/2872427.2883091
- Eckert, P. (2006). Communities of practice. In E. K. Brown, R. E. Asher, & J. M. Y. Simpson (Eds.), Encyclopedia of language and linguistics (Vol. 2) (pp. 683–685). Boston, MA: Elsevier.
- Friedberg, H., Litman, D., & Paletz, S. B. F. (2012). Lexical entrainment and success in student engineering groups. *Proceedings of the Spoken Language Technology Workshop* (SLT 2012), 2–5 December 2012, Miami, FL, USA (pp. 404–409). http://dx.doi.org/10.1109/SLT.2012.6424258
- Goffman, E. (1959). The presentation of self in everyday life. London: Penguin.
- Graham, S. L. (2007). Disagreeing to agree: Conflict, (im)politeness and identity in a computer-mediated community. *Journal of Pragmatics*, *39*(4), 742–759. http://dx.doi.org/10.1016/j.pragma.2006.11.017
- Guzdial, M. (1997). Information ecology of collaborations in educational settings: Influence of tool. In R. Hall, N. Miyake, & N. Enyedy (Eds.), *Proceedings of the 1997 Conference on Computer Support for Collaborative Learning* (CSCL '97), 10–14 December 1997, Toronto, ON, Canada (pp. 83–90). International Society of the Learning Sciences.
- Hammond, M. (1999). Issues associated with participation in online forums: The case of the communicative learner. *Education and Information Technologies*, *4*, 353–367. http://dx.doi.org/10.1023/A:1009661512881
- Hew, K. F. (2015). Student perceptions of peer versus instructor facilitation of asynchronous online discussions: Further findings from three cases. *Instructional Science*, *43*(1), 19–38. http://dx.doi.org/10.1007/s11251-014-9329-2
- Hiltz, S. R. (1986). The "virtual classroom": Using computer-mediated communication for university teaching. *Journal of Communication*, *36*, 95–104. http://dx.doi.org/10.1111/j.1460-2466.1986.tb01427.x
- Hiltz, S. R. (1994). The virtual classroom: Learning without limits via computer networks. Norwood, NJ: Ablex.
- Hughes, M., Ventura, S., & Dando, M. (2007). Assessing social presence in online discussion groups: A replication study. *Innovations in Education and Teaching International*, 44(1), 17–29. http://dx.doi.org/10.1080/14703290601090366
- Jaworski, A., & Coupland, N. (Eds.). (1999). The discourse reader. London/New York: Routledge.
- Joksimović, S., Dowell, N., Skrypnyk, O., Kovanović, V., Gašević, D., Dawson, S., & Graesser, A. C. (2015). How do you connect? Analysis of social capital accumulation in connectivist MOOCs. *Proceedings*



- (2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4
 - of the 5th International Conference on Learning Analytics and Knowledge (LAK '15), 16–20 March 2015, Poughkeepsie, NY, USA (pp. 64–68). New York: ACM. http://dx.doi.org/10.1145/2723576.2723604
- 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*(2), 125–143. http://dx.doi.org/10.1177/0261927X13502654
- Koedinger, K. R., Kim, J., Jia, J. Z., McLaughlin, E. A., & Bier, N. L. (2015). Learning is not a spectator sport: Doing is better than watching for learning from a MOOC. *Proceedings of the 2nd ACM Conference on Learning @ Scale* (L@S 2015), 14–18 March 2015, Vancouver, BC, Canada (pp. 111–120). New York: ACM. http://dx.doi.org/10.1145/2724660.2724681
- Kramer, A. D., Oh, L. M., & Fussell, S. R. (2006). Using linguistic features to measure presence in computer-mediated communication. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '06), 22–26 April, Quebec City, QC, Canada (pp. 913–916). New York: ACM. http://dx.doi.org/10.1145/1124772.1124907
- Kreijns, K., Kirschner, P. A., & Vermeulen, M. (2013). Social aspects of CSCL environments: A research framework. *Educational Psychologist*, 48(4), 229–242. http://dx.doi.org/10.1080/00461520.2012.750225
- Michinov, N., Michinov, E., & Toczek-Capelle, M. C. (2004). Social identity, group processes, and performance in synchronous computer-mediated communication. *Group Dynamics: Theory, Research, and Practice*, 8(1), 27–39. http://psycnet.apa.org/doiLanding?doi=10.1037%2F1089-2699.8.1.27
- Phirangee, K., Demmans Epp, C., & Hewitt, J. (2016). Exploring the relationships between facilitation methods, students' sense of community and their online behaviours. Special Issue on Online Learning Analytics. *Online Learning Journal*, 20(2), 134–154. http://dx.doi.org/10.24059/olj.v20i2.775
- Rahimi, Z., Kumar, A., Litman, D., Paletz, S., & Yu, M. (2017). Entrainment in multi-party spoken dialogues at multiple linguistic levels. *Proceedings of the 18th Annual Conference of the International Speech Communication Association* (INTERSPEECH 2017), 20–24 August 2017, Stockholm, Sweden (pp. 1696–1700).
- Rovai, A. P. (2002). Development of an instrument to measure classroom community. *The Internet and Higher Education*, *5*(3), 197–211. http://dx.doi.org/10.1016/S1096-7516(02)00102-1
- Rovai, A. P., & Barnum, K. T. (2007). On-line course effectiveness: An analysis of student interactions and perceptions of learning. *International Journal of E-Learning & Distance Education*, 18(1), 57–73.
- Shih, P.-C., Munoz, D., & Sanchez, F. (2006). The effect of previous experience with information and communication technologies on performance in a Web-based learning program. *Computers in Human Behavior*, *22*, 962–970. http://dx.doi.org/10.1016/j.chb.2004.03.016
- Thompson, K., Kennedy-Clark, S., Kelly, N., & Wheeler, P. (2013). Using automated and fine-grained analysis of pronoun use as indicators of progress in an online collaborative project. In N. Rummel, M. Kapur, M. Nathan, & S. Puntambekar (Eds.), *Proceedings of the 10th International Conference*



(2017). Talk with me: Student behaviours and pronoun use as indicators of discourse health across facilitation methods. *Journal of Learning Analytics*, 4(3), 47–75. http://dx.doi.org/10.18608/jla.2017.43.4

- on Computer Supported Collaborative Learning (CSCL 2013), 15–19 June 2013, Madison, WI, USA (pp. 486–493). International Society of the Learning Sciences.
- Ubon, N. A., & Kimble, C. (2004). Exploring social presence in asynchronous text-based online learning communities (OLCS). *Proceedings of the 5th International Conference on Information Communication Technologies in Education* (ICICTE 2004), 1–3 July 2004, Samos Island, Greece (pp. 292–297).
- Vassileva, J., Deters, R., Greer, J., McCalla, G., Bull, S., & Kettel, L. (2001). Lessons learned in deploying a multi-agent learning support system: The I-Help experience. In J. D. Moore, C. L. Redfield, & W. L. Johnson (Eds.), *Proceedings of the 10th International Conference on Artificial Intelligence in Education* (AIED 2001), 19–23 May 2001, San Antonio, TX, USA (pp. 410–421). IOS Press. http://publicat.bham.ac.uk/7565/
- Vassileva, J., McCalla, G. I., & Greer, J. E. (2016). From small seeds grow fruitful trees: How the PHelpS peer help system stimulated a diverse and innovative research agenda over 15 years. *International Journal of Artificial Intelligence in Education*, 26(1), 431–447. http://dx.doi.org/10.1007/s40593-015-0073-9
- Vrasidas, C., & McIsaac, M. (1999). Factors influencing interaction in an online course. *The American Journal of Distance Education*, 13(3), 22–36. http://dx.doi.org/10.1080/08923649909527033
- Wang, A. Y., & Newlin, M. H. (2002). Predictors of performance in the virtual classroom: Identifying and helping at-risk cyber-students. *THE Journal (Technological Horizons in Education)*, 29(10), 21.
- Wenger, E. (2008). Communities of practice: A brief introduction. Retrieved from https://scholarsbank.uoregon.edu/xmlui/bitstream/handle/1794/11736/a%20brief%20introduction%20to%20cop.pdf?sequence=1
- Wenger, E., McDermott, R. A., & Snyder, W. (2002). *Cultivating communities of practice: A guide to managing knowledge*. Cambridge, MA: Harvard Business Press.
- Wise, A. F., Hausknecht, S. N., & Zhao, Y. (2014). Attending to others' posts in asynchronous discussions: Learners' online "listening" and its relationship to speaking. *International Journal of Computer-Supported Collaborative Learning*, 9(2), 185–209. http://dx.doi.org/10.1007/s11412-0
- Wise, A. F., Hsiao, Y.-T., Marbouti, F., & Zhao, Y. (2012). Tracing ideas and participation in an asynchronous online discussion across individual and group levels over time: Raw research notes and article annotations. In J. van Aalst, B. J. Reiser, C. Hmelo-Silver, & K. Thompson (Eds.), *The Future of Learning: Proceedings of the 10th International Conference of the Learning Sciences* (ICLS '12), 2–6 July 2012, Sydney, Australia (Vol. 2, pp. 431–435). International Society of the Learning Sciences.