

Overload in Asynchronous Online Courses: a Scan Rate Analysis

Daniel Zingaro, Murat Oztok, Joseph Zingaro, Jim Hewitt

Paper presented at AERA 2013

Abstract: Students often report a feeling of information overload in asynchronous computer-conferencing courses. To supplement such qualitative perceptions, this study uses a scan rate metric to quantitatively investigate the reading behaviors of students relative to the number of messages they read per-session. We find that students more often skim messages in sessions where they read many messages compared to sessions where they read fewer messages. Yet, we do not find that such skimming increases with the length of sessions. We suggest the exploration of several hypotheses for interpreting this latter finding.

1 Purpose and Perspective

Asynchronous computer-mediated communication (CMC) courses are now the dominant form of distance higher education (Johnson & Aragon, 2003). By reading and writing notes (messages), students exchange information and make progress toward belonging to a supportive online community of learners (Hiltz & Wellman, 1997). Much literature espouses the benefits of asynchronous CMC compared to both synchronous CMC (e.g. chat) and face-to-face courses, including time-independent access, increased participation quality, and greater peer-to-peer interaction (Morse, 2003).

But CMC courses are not without their problems. In particular, much literature over the past 25 years describes the insidious nature of information overload experienced by online students. It has been defined in many similar ways, all relating to the presenting of information too quickly to be of benefit (Hiltz & Turoff, 1985; Paulo, 1999). Its importance is highlighted by its inclusion in best-practice principles (Johnson & Aragon, 2003) and its observable impacts on public online discourse. For example, in response to information overload in online newsgroups, users preferentially reply to smaller rather than larger messages, and are more likely to end participation (Jones, Ravid, & Rafaeli, 2004). Hiltz and Wellman (1997) suggest that the problem is particularly pronounced in large classes involving many communication partners; perhaps overload may be implicated in the low levels of cognition found in many response posts (Zingaro, 2012).

Peters and Hewitt (2010) interviewed graduate students about their online practices and found overload to be a dominant theme. Students were overwhelmed by the volume of messages and threads, especially when other commitments precluded regular logins. Students reacted negatively to notes of little substance, long notes, and notes rife with academic language; often focused on single threads; and sometimes systematically ignored notes written by certain authors. When students **did** read notes while feeling overloaded, they admitted to skimming those notes until they found something of interest. As partial solutions to the problem, researchers have suggested using small task-oriented groups (Hiltz & Turoff, 1985; Hewitt & Brett, 2008) and purposefully limiting the number of activities and amount of information (Johnson & Aragon, 2003).

The purpose of the current study is to quantitatively investigate perceptual claims from Peters and Hewitt (2010) about the relationship between overload and skimming practices. We define a **session** to begin when a student logs into the environment, and to end when a student logs out or remains inactive for at least one hour. We are specifically interested in the number of notes read in each session, and assume that overload occurs when a significant number of notes is read in a single sitting.

Our perspective is based on the application of cognitive load theory (CLT) to CMC environments (Danilenko, 2010). CLT argues that as short-term memory becomes exhausted, students experience overload and may be unable to integrate new knowledge into their long-term organizational schema. We suggest that as more and more distinct ideas (notes) are read in a single sitting, perceptual systems will become exhausted and students' behavior will change as they actively cope with too much information.

How do skimming rates change with the number of notes read per session? To operationalize skimming, we use the scan rate metric (Hewitt, Brett, & Peters, 2007). Hewitt et al. (2007) argues that reading at a rate of more than

eight words-per-second on a computer screen is indicative of skimming. Therefore, a student is said to **scan** a note when they read that note at eight words-per-second or more, and the scan rate is the percentage of notes opened by a student that were scanned. For example, if a student opens (reads) 20 notes, and reads 15 of those at a rate faster than 8 words-per-second, their scan rate would be 75 percent. To be sure, scan rate provides only a lower bound estimate on the percentage of notes that were opened and closed too quickly to have been read thoroughly. For example, a student might open a short note for several minutes (indicative of a thorough reading) but nevertheless not read the note at all (e.g. they switched to another application on their computer). That is, scan rate may “miss” some scans, but accurately captures all true scans. Compared to other approaches such as eye-tracking, the scan rate is non-invasive, produces a measure derived from natural student reading practices, and provides a lower-bound on actual amount of scanning. If we conclude something about student reading practices using scan rate, we can be confident that we have done so with only a lower-bound estimate on the “real” effect.

Previous research has found significant positive correlations between scan rate and note size, as well as scan rate and class size. In general, long notes lead to more scanning, as do large classes where large numbers of notes are posted (Hewitt et al., 2007). Here, we are interested in scan rates within sessions, rather than the effect of note or class structure on scan rate. The **session scan rate** is calculated for each session as the number of notes scanned divided by the total number of notes read.

2 Methods and Data Sources

Our data source consists of 26 fully- or partly-online graduate-level education courses offered at a large North American research university in the 2010-2011 academic year. All asynchronous discussions took place in the same threaded CMC environment developed at our institution. Each week, students are expected to prepare for discussion through readings, then read and write notes in the current conference space. Some courses organized students in small weekly discussion groups; others relied on full-class discussion.

There are two modes that students can use to read notes: **note view** and **contents view**. Note view ensures that one note at a time is displayed on-screen. Our environment keeps timestamped logs of user actions, so we can calculate the length of time the note was visible in order to calculate scan rate.

The contents view, on the other hand, is problematic for computing scan rates: once all notes are displayed simultaneously in a large window, we are unable to calculate the length of time spent reading each note. However, of all 16045 sessions where at least one note was read, only 4.84% of those (816) used the contents view. Furthermore, of the 585 students who read notes in at least two different sessions (eliminating from consideration those students who signed-in just once prior to dropping the course), 94% used the contents view less than 5% of the time. We therefore decided to eliminate all sessions where contents view was used, understanding that this potentially misrepresents the practices of a small number of students.

For each student, we examined scan rates on three classes of sessions: those where 1-2 notes were read (small sessions), those where 3-7 notes were read (medium sessions), and those where 8-45 notes were read (large sessions). These ranges divide all sessions into approximate thirds; we cut the third range at 45 in order to eliminate a small number of outlier sessions with very high scan rates.

3 Results

We ran a repeated-measures ANOVA (alpha level 0.05) with a categorical, three-level independent variable (small, medium, and large session size) and a continuous dependent variable (scan rate). 519 students logged at least one session of each type and are therefore included in this analysis. The repeated-measures sphericity assumption was violated (Mauchly’s test for sphericity yielded $p = 0$); p-values are therefore Greenhouse-Geisser-corrected (Lawrence, 2011).

We found a significant effect of session size on scan rate ($F(2, 1036) = 98.5, p = 0$). Generalized eta squared (GES) effect size was 0.049, indicating a small-to-medium effect (Bakeman, 2005).¹

The mean and standard deviation of scan rates for the three session sizes are as given in Table 1. Fisher’s Least Significant Difference is 1.07, indicating that scan rates in large sessions are different from scan rates in at least one of the other two session categories.

¹Hewitt et al. (2007) do not provide effect sizes of independent variables on scan rate. To add meaning to our effect sizes, we replicated their analyses on our current data. Size-of-note yielded the largest effect size: we found a GES of 0.31 with a three-level independent variable (notes of

Table 1: Mean Scan Rates for Small, Medium and Large Sessions

Session Type	Mean (%)	sd (%)
1-2 Notes	23.57	13.33
3-7 Notes	25.28	12.54
8-45 Notes	30.92	15.54

Table 2: Mean Scan Rates when reading during the first fifteen minutes, and after the first fifteen minutes

Time	Mean (%)	sd (%)
0-15 mins	32.43	15.83
After 15 mins	27.66	15.75

As follow-up, we wondered whether scan rates remained relatively constant throughout a session, or whether there was a measurable impact of time on scanning. We chose fifteen minutes as our cut-off, based on its folk status as the number of minutes for which students can focus on learning in a lecture setting (Wilson & Korn, 2007). For each student, we compared scan rates exhibited during the first fifteen minutes of sessions to scan rates occurring after those fifteen minutes. Our independent variable therefore contained two levels (first fifteen minutes and after first fifteen minutes) and our dependent variable was again scan rate. 554 students logged at least one session of each type and are therefore included in this analysis.

The independent variable had a significant effect ($F(1, 552) = 108.83, p = 0$). GES effect size was 0.022, indicating a comparatively small effect. The mean scan rates for the two levels of the independent variable are as given in Table 2.

4 Significance and Discussion

We have shown that scan rate increases as the number of notes read per session increases. Prior literature shows that students' note-readings are almost invariably of notes that they have not previously read (Hewitt, 2005). We therefore hypothesize that students' feeling of overload is related to the number of unread notes they see when logging-in to the environment, and therefore to the number of notes they feel they must read to stay current in the course. Specifically, when students log-in and read only a few notes, they exhibit reading patterns suggestive of thorough reading. When they log-in and read many notes, these same students spend comparably less time reading each note.

Our results suggest that reading many notes in a single sitting may not be as educationally-valuable as reading fewer notes over a larger number of sessions. Hewitt et al. (2007) hesitated to ascribe meaning to high scan rates, arguing that skilled readers might adeptly use scanning to negotiate a large body of information. However, in this study, we have found that within-student scan rates increase based on the amount they are reading. At minimum, we argue that students are "acting differently" when there is "too much to read". Investigating whether this behavior-change leads to more superficial learning or retention would be valuable future work.

Our second result is perhaps more counterintuitive. In addition to an increased scan rate due to the presence of many notes, we expected that scanning would also increase from the start of a session to the end of that session. We instead found that scan rate decreases slightly after the first fifteen minutes of lengthy sessions. How can we square our two results?

One hypothesis is that students begin their sessions by scanning, quickly looking for material of relevance. The first few minutes of a session thus demonstrate an elevated scan rate, which drops as students begin to engage more deeply with what they have found. Fruitfully investigating this hypothesis would require student perspectives in addition to quantitative data analysis. A second hypothesis is that students gauge their per-session required reading at the outset of each session and adjust their scanning accordingly. After clicking on the current conference in our environment, students see a list of notes along with an indicator of whether they are read or unread. It is plausible that students use this information, along with the amount of time they are willing to spend online, to set a more-or-less constant "note-reading velocity" for each session.

5 Conclusion

When students feel that they have too many notes to read, it is likely that they engage in quick, shallow reading of many of those notes. Our results mesh with prior qualitative research suggesting that information overload is a widespread concern for students in asynchronous computer-mediated environments. The same students exhibit widely-differing reading patterns depending on the number of notes that they read in a session. We suggest continued use of scan rate as a measure of student skimming practices, and further investigation into the relationship between scanning and time-on-task.

References

- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, 37(3), 379-384.
- Danilenko, E. (2010). *The relationship of scaffolding on cognitive load in an online self-regulated learning environment* (Unpublished doctoral dissertation). University of Minnesota.
- Hewitt, J. (2005). Toward an understanding of how threads die in asynchronous computer conferences. *Journal of the Learning Sciences*, 14(4), 567-589.
- Hewitt, J., & Brett, C. (2008). The relationship between class size and online activity patterns in asynchronous computer conferencing environments. *Computers & Education*, 49(4), 1258-1271.
- Hewitt, J., Brett, C., & Peters, V. L. (2007). Scan rate: A new metric for the analysis of reading behaviors in asynchronous computer conferencing environments. *American Journal of Distance Education*, 21(4), 215-231.
- Hiltz, R., & Turoff, M. (1985). Structuring computer-mediated communication systems to avoid information overload. *Communications of the ACM*, 28(7), 680-689.
- Hiltz, R., & Wellman, B. (1997). Asynchronous learning networks as a virtual classroom. *Communications of the ACM*, 40(9), 44-49.
- Johnson, S. D., & Aragon, S. R. (2003). An instructional strategy framework for online learning environments. *New Directions for Adult and Continuing Education*, 2003(100), 31-43.
- Jones, Q., Ravid, G., & Rafaeli, S. (2004). Information overload and the message dynamics of online interaction spaces: A theoretical model and empirical exploration. *Information Systems Research*, 15(2), 194-210.
- Lawrence, M. A. (2011). ez: Easy analysis and visualization of factorial experiments [Computer software manual].
- Morse, K. (2003). Does one size fit all? exploring asynchronous learning in a multicultural environment. *Journal of Asynchronous Learning Networks*, 7(1), 37-55.
- Paulo, H. (1999). *Information overload in computer-mediated communication and education: is there really too much information? implications for distance education* (Unpublished master's thesis). University of Toronto.
- Peters, V. L., & Hewitt, J. (2010). An investigation of student practices in asynchronous computer conferencing courses. *Computers & Education*, 54(10), 951-961.
- Wilson, K., & Korn, J. (2007). Attention during lectures: Beyond ten minutes. *Teaching of Psychology*, 34(2), 85-89.
- Zingaro, D. (2012). Student moderators in asynchronous online discussion: a question of questions. *Journal of Online Learning and Teaching*, 8, 159-173. Retrieved from http://jolt.merlot.org/vol18no3/zingaro_0912.htm