研究ノート

A Tool to Extract Quantitative and Qualitative Data from Zoom Chat Transcripts from Online Classes

William Pellowe[1)]

■**Abstract**

The pandemic caused upheavals in tertiary education, but one positive outcome is that many universities now offer online makeup lessons. However, this convenience comes with the same set of challenges as before: students participating in Zoom classes report a loss of motivation, and some reportedly "ghost" classes, appearing in name only. Given these circumstances, it becomes crucial for teachers to find ways to keep students engaged, attentive, and active, especially in large classes. One way to accomplish this in Zoom classes is to periodically elicit student answers via DM (direct chat messages) in IRT (Initiation-Response-Feedback) question cycles. We can elicit a variety of answer types, including multiple choice and short answers. Yet how can teachers reliably and quickly track the frequency of student responses? In this paper, the author describes the development of a tool designed to analyze student responses through chat-based direct messaging.

Key Words; student engagement, online classes, Zoom chat transcripts

1. Introduction

The COVID-19 Pandemic caused an upheaval in tertiary education worldwide, with many classes either going online or becoming hybrid. Many courses, including those at Kindai University, used the Zoom platform for online classes. Students joined classes from home on their own devices. Teachers were able to share PowerPoint slides and other visual aids while delivering the class material. Many teachers made use of breakout rooms, which allowed students to break away into smaller groups for discussion or groupwork.

As the pandemic continued, students participating in classes through Zoom often reported "Zoom fatigue", which is when students experience more difficulty maintaining attention and focus in online classes, as well as a loss of interest and motivation (Shoshan & Wehrt, 2022).

Some students reportedly "ghost" the classes, appearing in name only. These students have logged in to the Zoom class, but their camaras are off, and they are engaged in other activities, such as completing work for other classes, checking their social media, or doing household chores (Wakefield, 2020).

Not all students ghosted classes or experienced Zoom fatigue, yet Peper et al. (2021) report that some students experience difficulty remembering (and therefore learning) the course content during online classes. Those researchers compare the passivity of watching Zoom presentations to watching TV and streaming videos, which conditions students to take in information passively. "Learning requires engagement, which means a shifting from passively watching and listening to being an active participant shareholder in synchronous online classes" (Peper et al., 2021, p. 51).

Post-pandemic, classes at Kindai University have returned to being in person. However, some classes (such as make-up lessons) are still being held through Zoom. Furthermore, this pandemic remains unpredictable; we may find ourselves back to online classes due to future circumstances. Therefore, it remains important to be prepared for online classes.

2. Increasing Engagement: Peer Instruction and the IRF Cycle

During in-person classes, one effective technique to increase student engagement is to periodically ask summative multiple-choice questions to the whole class, and have all students respond simultaneously either with response cards (Pellowe et al., 2015; Randolf, 2007)

1）近畿大学産業理工学部　教養・基礎教育部門　准教授　pellowe@fuk.kindai.ac.jp

or through an online answer system. These periodic questions are often called "Concept Tests" in a lecture method called Peer Instruction (Mazur, 1997), which was pioneered in physics classes at Harvard University. Peer Instruction has effective learning outcomes regardless of whether the instructor uses answer cards or clicker systems (Lasry, 2007). In a Peer Instruction model, if fewer than 35% of a class replies to a Concept Test question with the correct answer, the teacher revisits the previous content. However, if the class is evenly divided between correct and incorrect answers, the students are asked to talk with each other to try to clear up their misunderstanding before they "vote" again on the answers (see Fig. 1, from Lasry, 2007).

Concept tests are carefully selected to illustrate the key concept being discussed. In science classes, a single question often suffices to illustrate the scientific principle being taught, but learning a foreign language requires more examples. In language classes, a common way to check understanding while providing extra practice is the Initiation-Response-Feedback (IRF) cycle (Sinclair & Coulthard, 1975). The traditional method of responding in an IRF cycle is a choral answer, with all students saying the answer aloud. As an example, imagine the teacher has a series of slides with illustrations of situations meant to elicit an answer in passive continuous voice. The teacher shows a slide of someone stealing a bicycle, and then:

Teacher: What's happening in this picture? The bicycle…

All students: The bicycle is being stolen.

Teacher: Yes, the bicycle is being stolen.

In IRF cycle terms, the exchange above consists of these three stages:

1. Initiation: The teacher initiates the cycle by showing a picture on a slide, asking a question, and (in this case) providing the starting point of the answer to signal the type of answer being sought.
2. Response: Students respond aloud as a group.
3. Feedback: The teacher indicates if the answer is correct, and then models the correct answer aloud. This correct answer may then be revealed on the slide.

This IRF cycle part of the lesson may include three to six such IRF cycles in a row, with other images selected to elicit similar responses.

Instead of choral responses, teachers can opt to do multiple choice IRF cycles. In a multiple choice IRF cycle, the correct answer as well as one or more distractors are shown on the slide with the illustration of the situation. To re-use the previous example, the slide would include a few choices:

A: The bicycle is being stolen.

B: Someone is being stolen the bicycle.

C: The bicycle stole someone.

During the response stage, students would demonstrate their understanding with a response card or through a classroom response system.

In a Peer Instruction implementation of the IRF cycle, the feedback for a correct answer remains the same, but if the responses are about evenly divided between correct and incorrect answers, the teacher does not give immediate feedback. Instead, the teacher would say something like, "Now, please talk with two or three

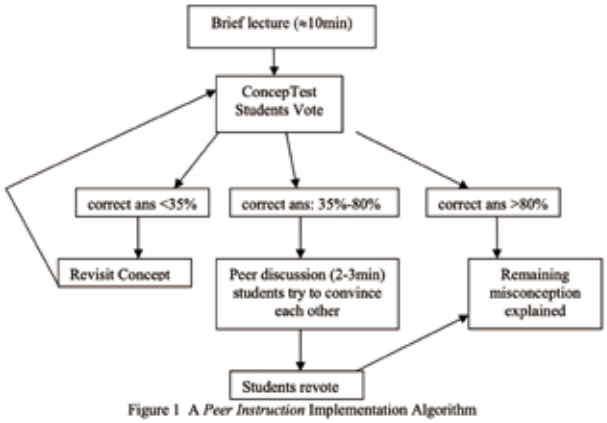

Figure 1 A *Peer Instruction* Implementation Algorithm

Figure 1 A Peer Instruction Implementation Algorithm

others around you to say what you think the answer is, and why you think your answer is correct. About two minutes later, I'll ask again."

3. Methods of eliciting student responses online

In online classes through Zoom, implementing the IRF cycle poses some challenges. Unlike in-person classes, students online cannot call out their answers chorally. In this section, I describe methods of eliciting responses.

3.1 Google Classroom, NearPod, or Kahoot

One option is for students to answer multiple questions outside of Zoom through a platform such as Google Quiz (which comes included with Google Classroom), NearPod, or Kahoot. One drawback to this option is that switching all students over to the outside platform causes delays and confusion. Also, even though these outside platforms can be very effective for many students, they are not effective for all students. The most effective approach with platforms such as Kahoot requires students to use a second device for answers. In the best-case scenario, a student looks at the quiz question on their computer screen while answering on their smartphone. However, financially disadvantaged students lack access to multiple devices, so this solution would not work with all the students in a class.

The Google Quiz system included with Google Classroom is very useful during other phases of instruction. This system is an effective way to create a quiz activity for students to do as a review, or as a way for students to report their answers to questions posed in their textbooks. In my own online classes, I have sent students to do an activity in Google Classroom, telling them to return to our Zoom lesson 10 minutes later. Google Classroom quizzes can accept written answers, too, which can be checked later by the instructor. However, for a quick answer during an IRF cycle, the Google Quiz option is not fast enough.

3.2 Answer cards

An option that I have made extensive use of in my own in-person classes is to have each student hold up an answer card (Pellowe et al., 2015). These cards have four answer options (A, B, C, and D), and in an in-person class, it is very easy to ask students to simultaneously hold the cards up near their faces to show their answers. A card-free option is to just have students hold up fingers (one for A, two for B, etc.). However, in a Zoom lesson, this option requires students to have their cameras on, which violates privacy policies at some universities. Even if university policy encourages students to keep their cameras on during lessons, having so many camera feeds streaming at the same time may cause bandwidth issues in large classes, resulting in Zoom "freezing" for any students who have slower internet connections. Furthermore, answer cards are ineffective when students are using virtual backgrounds, which many do for privacy, because the answer cards become difficult to see.

3.3 Zoom Polls

Zoom's polling feature seems on the surface to be a great option. With this option, the most effective use of time is for teachers to reuse a generic A, B, C, D poll for each question. When the teacher launches the question, the poll will automatically appear on every student's screen

Figure 2 Screenshot of a Zoom poll "launch" box appearing on-screen over a presentation slide (Pellowe, 2020)

until they choose an answer.

Polls have numerous advantages. As long as the teacher plans ahead, polls can be added to the session, and to allow for spontaneity, generic A-B-C-D polls can be re-used for several questions. Also, since the poll appears directly on the students' screens, they do not have to leave the presentation screen to answer. Finally, one option available with polls is to finish the poll by showing the answer distribution during the Feedback stage of the IRF cycle.

However, there are some drawbacks to this approach. First, if any participant is logged in twice (for example, if a student is accessing for video on one device and audio on a second device), polls will automatically be disabled. Also, standard poll data is anonymous, so without delving deep into the settings before a class and requiring students to log in to Zoom, the teacher has no record of which students are participating and which are not. Recently, however, Zoom has added advanced features to polls (Zoom, 2023) which seem to mitigate some of these disadvantages (for example, advanced polls can be created during the session, and reply data can be downloaded after the class session). However, if you wish to retain the data for each poll, you cannot re-use your polls each IRF cycle. Zoom recommends creating new polls for every polling session. In a standard IRF cycle of questions, this would require the teacher to create a new poll to launch for each response cycle. Also, the advanced polls require each participant to have upgraded their version of Zoom, which adds an additional hurdle when trying an activity with forty or more students in a class.

3.4 Zoom chat

The final option discussed here is for students to respond in the chat feature that is built into Zoom. This is an easy way to elicit a variety of answer types (multiple choice, words, or phrases). Zoom chat can be saved, which preserves records of each student's participation. Chat messages are motivating because the students know that the teacher sees their name alongside their answers. Furthermore, while these questions are best planned beforehand, chat responses can be elicited spontaneously. Finally, chat is relatively easy. Students on computers can easily access the chat input area while maintaining a full visual of the teacher's screen. On mobile devices, chat responses are a little more complicated, yet still doable.

On the surface, it may seem that students lose their privacy when they have to respond in Zoom chat. However, students can send Zoom chat messages to the teacher as direct messages (DM). DM is low risk; students' answers are invisible to the other students, so no student needs to feel embarrassed about their answers, and their privacy is preserved.

However, for teachers of large classes, it is very difficult to keep track of all the answers coming in through Zoom chat, and extremely difficult to notice which students are declining to participate. Chat transcripts can be saved, but manually checking the saved chat transcript after class is arduous, and too much to ask of busy teachers. To address this difficulty, the author decided to create a tool to analyze the chat transcript.

4. Zoom chat analysis tool

When chat transcripts are saved by Zoom, they are saved as a structured text file, which suggests that creating an automated way of analyzing each student's participation would be possible. Allan and Vahid (2021) created a Zoom Chat Log Analysis site online. The chat log is uploaded to the site, and the teacher is then provided with the number of messages per participant in a chat. This approach is a start, but is not sufficient for measuring the quality and true quantity of students' chat responses to the questions posed by the teachers in class. Also, Sustenance (2023) describes pasting the Zoom chat transcript into Word and performing a complex series of global search-and-replace functions to convert the transcript into a format suitable for exporting to Excel. My own Zoom Chat Analysis tool was created in Excel, and will automatically convert the transcript into data.

4.1 IRF Cycle in Zoom Chat

This chat transcript excerpt took place during an online lesson in an EIGO class at Kindai University. (Students' Zoom numbers have been changed to preserve student privacy.) The students were studying using "should have" + past participle to describe an alternative action in the past, such as, "He should have studied harder for the test." For the IRF cycle, students were shown a few short video clips from the movie *Mister Baseball* in which a new-to-Japan foreign player makes some cultural errors. They were then asked to describe the mistake.

In one scene, the character on screen walked into the locker room without taking off his boots. All of the players on his team yelled at him to go back and take his shoes off. In the chat excerpt below (Fig. 3), the IRF cycle begins with the Initiation: After students watch the video, the class was asked, "What should he have done?" In the chat, I wrote, "He should've __" to let students know the form of the answer.

```
13:32:12 From ﾍﾟﾛｳ(Pellowe) to Everyone:
        He should've ___
13:33:21 From 24B183 to ﾍﾟﾛｳ(Pellowe)(Direct Message):
        taken off his shoes.
13:33:24 From 24D300 to ﾍﾟﾛｳ(Pellowe)(Direct Message):
        taken his boot off
13:33:25 From 24Q017 to ﾍﾟﾛｳ(Pellowe)(Direct Message):
        taken off his shoes
13:33:36 From 24K153 to ﾍﾟﾛｳ(Pellowe)(Direct Message):
        He should've take off his boots.
13:33:40 From 24K132 to ﾍﾟﾛｳ(Pellowe)(Direct Message):
        He should've off shoes
13:34:12 From ﾍﾟﾛｳ(Pellowe) to Everyone:
        He should've taken his shoes off
```

Figure 3 Chat transcript excerpt

As students responded, I verbally replied with encouragement (such as "Yes, that's right"), which lets students know that others are writing answers (as other students cannot see direct messages which are not sent to them). After some responses, I provided the feedback. In chat, I typed an answer, and the on-screen slide showed the answer as well.

4.2 Creating the chat analysis tool

Regarding the structure of the file, as you can see in Figure 3, each contribution to the Zoom chat takes two lines. The first line contains a time stamp, the word "From", and the Zoom ID of the person making the post, followed by "to" and the Zoom ID of the recipient of the post. The second line starts with a tab space, followed by the response itself.

My Excel document to analyze Zoom transcripts has three sheets:

- "Sheet 1" tallies the number of contributions per student from the third sheet.
- "Transcript 1 Paste" is where teachers paste the text file of the Zoom transcript.
- "Transcript 1 Formatted" converts the raw text file into a formatted data sheet.
- (Each additional transcript would require its own "Paste" and "Formatted" sheets, with corresponding columns in "Sheet 1")

First, the teacher pastes all of the students' Zoom IDs into the first column of Sheet 1. Then, the teacher copies the compete Zoom transcript, and pastes it into the second sheet (Transcript 1 Paste). Figure 4 shows what the chat excerpt from Figure 3 looks like when pasted into the second sheet. Notice that the tab in the second line becomes an empty cell in Excel.

	A	B	C	D
55	13:32:12 From ﾍﾟﾛｳ(Pellowe) to Everyone:			
56		He should've ___		
57	13:33:21 From 24B183 to ﾍﾟﾛｳ(Pellowe)(Direct Message):			
58		taken off his shoes.		
59	13:33:24 From 24D300 to ﾍﾟﾛｳ(Pellowe)(Direct Message):			
60		taken his boot off		
61	13:33:25 From 24Q017 to ﾍﾟﾛｳ(Pellowe)(Direct Message):			
62		taken off his shoes		
63	13:33:36 From 24K153 to ﾍﾟﾛｳ(Pellowe)(Direct Message):			
64		He should've take off his boots.		
65	13:33:40 From 24K132 to ﾍﾟﾛｳ(Pellowe)(Direct Message):			
66		He should've off shoes		

Figure 4 Zoom Chat Transcript pasted into Excel

This raw transcript is formatted on the "Transcript 1 Formatted" sheet as seen in Figure 5.

	A	B	C	D	E	F	G	H
1	Student	Time	Answer	Points	Check (TRUE=This student is listed on Sheet1)			
56	ﾍﾟﾛｳ(P	13:32:12	He should've ___		FALSE			
57			0		FALSE			
58	24B183	13:33:21	taken off his shoes.		TRUE			
59			0		FALSE			
60	24D300	13:33:24	taken his boot off		TRUE			
61			0		FALSE			
62	24Q017	13:33:25	taken off his shoes		TRUE			
63			0		FALSE			
64	24K153	13:33:36	He should've take off his boots.		TRUE			
65			0		FALSE			
66	24K132	13:33:40	He should've off shoes		TRUE			

Figure 5 Formatted Chat

Note that every other row in Figure 5 is blank, which is an unfortunate formatting byproduct of the way that each contribution in Figure 4 takes up two lines. The "points" column is where the teacher awards points for chat messages that are on task, such as a response in an IRF cycle, and awards no points for miscellaneous or off-task contributions (greetings, technical or procedural questions, and so on). This allows the teacher to collect qualitative data separate from the quantitative data of the simple tally in the "Count" column.

Using row 58 (Fig. 5) an example, the formulas are as follows:

- Student column, row 58:
  =MID ('Transcript 1 Paste'! A59, 15, 6)
- Time column, row 58:

=LEFT('Transcript 1 Paste'! A59, 8)

- Answer column, row 58:
  ='Transcript 1 Paste'! B60
- Check column, row 58:
  =NOT (ISERROR (MATCH (A58, Sheet1! $A$2: $A$96, 0)))

The "Check" column (Fig. 5, row E) simply gives us "TRUE" if that student ID exists in our list of student IDs on our first sheet (Fig. 6). This helps alert us if a student has mistyped their Zoom ID.

After the Zoom chat transcript has been pasted, the number of contributions per student is tallied up on Sheet 1 (see Figure 6). If points are awarded (see Figure 5, column D), then these points are also tallied on Sheet 1 (Fig. 6, row F).

	A	B	C	F
1	Student ID	Student Name	T 1 Count	T 1 Points
2	24K177		12	11
3	24Q012		10	9
4	24Q017		10	9
5	24B179		7	6
6	24K132		7	6
7	24D300		6	5
8	24K157		6	5
9	24B172		4	3
10	24H256		4	3

Figure 6 Sheet 1 tallies up the number of contributions per student (sorted high to low)

Using row 2 (Fig. 6, cell C2) as an example, the formula to tally the student contributions is:

=COUNTIF ('Transcript 1 Formatted'! $A$1: $A$1396, VLOOKUP (A2, 'Transcript 1 Formatted'! A:C, 1, FALSE))

The formula to tally the points awarded to student contributions (Fig. 6, cell F2) is:

=IFERROR (SUM (FILTER ('Transcript 1 Formatted'!$D$1: $D$1396, 'Transcript 1 Formatted'!$A$1: $A$1396=A2)), 0)

As seen in Figure 6, some students made several contributions in the Zoom chat, while others made only a few. Some students made no contributions in chat at all, with no explanation, so they were most likely "ghosting" the class (as described in Wakefield, 2020). However, for the students who did contribute, their answers showed me in real time how well they understood the grammar, as well as informing me of how much review was needed.

5. Discussion: the efficacy of the Zoom Chat Analysis tool

As-is, the Zoom Chat Analysis tool can be easily distributed to other teachers who would like to use it (e.g., Pellowe 2023), but using it requires some work on their part. First, the student Zoom IDs are presumed to be 6 characters long. Teachers familiar with Excel can adjust that to their own requirements, but a cleaner method would be to have a master preferences sheet where teachers could type in the length of their students' Zoom IDs, and that number would be used in the relevant formulas.

Also, the Zoom Chat Analysis tool needs a cleaner-looking layout in the "Formatted" sheet. Currently, every other line is blank due to the copy paste of the formula. This makes it difficult to scan the sheet to find students whose Zoom IDs do not match the Zoom IDs they were supposed to use.

Furthermore, instead of having teachers award points for good contributions, I think it would be more convenient to automatically award a point for each contribution, and have teachers delete the points for irrelevant or duplicate contributions. After all, most contributions by students in the chat are on-task responses. Ideally, this could be an option for teacher to select in a preferences sheet.

Finally, rather than rely on Excel, developing this tool as software would eliminate many of the problems described above.

6. Conclusion

Maintaining and monitoring student engagement during online lessons remains a challenge for educators, especially in large classes. If we could ensure that all students were taking their classes on good quality computers, in an environment with robust internet connections, we would have many more options to engage students meaningfully with the class content. However, when we have large differences between students in terms of their equipment and Wi-Fi access, we must rely on the simplest solutions. When using Zoom, the simplest solution seems to be using the Zoom chat as a place where students can post answers to questions. This Zoom Analysis Tool adequately helps teachers access

qualitative and quantitative data about their students' Zoom chat contributions.

References

Allan, J.M & Vahid, F. (2021). *UCR - Zoom Chat Log Analysis*, Retrieved September 10, 2021 from https://www.cs.ucr.edu/~jalle010/zoom_chat_analysis/release/

Haque, A.u. (2021). "No One Wants to be a Host of Ghost Sessions: Techniques to Improve Student Engagement and Active Participation", In: Fayed, I., Cummings, J. (eds) *Teaching in the PostCOVID-19 Era*. Springer, Cham. https://doi.org/10.1007/978-3-030-74088-7_43

Lasry, N. (2007). "Peer Instruction: Comparing Clickers to Flashcards", *Physics Education*, arXiv:physics/0702186 [physics.ed-ph] Retrieved online September 10, 2021 from https://arxiv.org/abs/physics/0702186

Mazur, E. (1997). *Peer Instruction : A User's Manual*. Upper Saddle River, NJ: Prentice Hall.

Pellowe, W. (2020) "Active Participation in Online Lessons" Online video preview of an online presentation, *Online Teaching Japan (OTJ) Summer Sessions*, August 24, 2020. Available: https://onlineteachingjapan.com/summer-sessions-calendar/active-participation-in-online-lessons

Pellowe, W. (2023) "Using Zoom Text Chat to Increase Participation and Engagement in Online Language Classes", Online resource page for a conference presentation given at the Hokkaido JALT Winter Conference 2023, February 5, 2023). Available: http://billp.me/hok23/

Pellowe, W., Paton, S. & Shimizu, P. (2015) "Encouraging Active Participation With Response Cards", in *Sustainability: Making Teaching and Learning Last,* PanSIG Conference Proceedings, pp. 172-177.

Peper, E., Wilson, V., Martin, M., Rosegard, E. & Harvey, R. (2021) "Avoid Zoom Fatigue, Be Present and Learn". *NeuroRegulation*, 8(1) pp. 47-56 https://doi.org/10.15540/nr.8.1.47

Randolph, J. J. (2007). "Meta-Analysis of the Research on Response Cards: Effects on Test Achievement, Quiz Achievement, Participation, and Off-Task Behavior". *Journal of Positive Behavior Interventions,* 9(2), pp. 113-128.

Sustenance, S., "Converting Zoom Messages into Participation Points", presentation at *the 2023 JALT Hokkaido Language Teaching Conference: New Directions and New Techniques: Teaching in the New Normal,* Hokkai-Gakuen University, February 5, 2023.

Shoshan, H. N., Wehrt, W. (2022) "Understanding 'Zoom fatigue': A mixed-method approach". *Applied Psychology*, Volume 71, Issue 3, pp. 827-852 https://doi.org/10.1111/apps.12360

Sinclair, J., & Coulthard, M. (1975). *Towards an Analysis of Discourse*. Oxford, UK: Oxford University Press.

Wakefield, J. (2020) "Zoom Ghost Busting", *LX at UTS,* University of Technology Sydney, Online https://lx.uts.edu.au/blog/2020/05/07/zoom-ghost-busting

Zoom (2023) "Conducting advanced polls and quizzes in meetings" Retrieved May 20, 2023 from https://support.zoom.us/hc/en-us/articles/4412325214477-Conducting-advanced-polls-and-quizzes-in-meetings#h_01ETTJTP6HXJM3Z5PVQZWD8DWF