

Korero: Facilitating Complex Referencing of Visual Materials in Asynchronous Discussion Interface

SOON HAU CHUA, NUS-HCI Lab, National University of Singapore

TONI-JAN KEITH MONSERRAT, Institute of Computer Science, University of the Philippines Los Baños

DONGWOOK YOON, Computer Science, University of British Columbia

JUHO KIM, School of Computing, KAIST

SHENGDONG ZHAO, NUS-HCI Lab, National University of Singapore

ABSTRACT

In asynchronous online discussions, users actively reference visual materials (e.g., video, document) to provide supporting evidence and additional context. However, creating and comprehending complex references can be challenging, especially when there are multiple referents to refer, or when a referent is highly specific (e.g., specific sentences in a paper rather than the paper as a whole). To identify users' challenges in making references with *multiple and specific referents* while using existing discussion tools, we conducted an observational study and a preliminary interview. Based on the design lessons, we built Korero, a discussion interface that aims to facilitate complex referencing actions. For evaluation, we compared Korero against conventional interfaces in two user studies with referencing tasks of different referential difficulty. We found that Korero not only significantly reduces the time and effort in making references with multiple and specific referents, but also shows potential in increasing users' engagement with the discussion and referent materials.

CCS CONCEPTS • Human-centered computing~Graphical user interfaces

KEYWORDS: Asynchronous discussion interface; referencing

ACM Reference format:

Soon Hau Chua, Toni-Jan Keith Monserrat, Dongwook Yoon, Juho Kim and Shengdong Zhao. 2017. Korero: Facilitating Complex Referencing of Visual Materials in Asynchronous Discussion Interface. *Proc. ACM Hum.-Comput. Interact.* 1, 2, Article 34 (November 2017), 20 pages.

DOI: <https://doi.org/10.1145/3134669>

INTRODUCTION

“Bob and Susan are having a face-to-face learning discussion. As Susan is making her point, *she uses her finger to refer to a passage in the textbook and a concept map she created in her notes to support her explanation.*”

In the scenario above, Susan is building common ground with Bob by making references to specific information from different visual materials. Instead of describing the passage and concept-map verbally, non-verbal methods such as referencing make common ground easier and more efficient to build between interlocutors [15].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org. 2573-0142/2017/November - Article 34 \$15.00

Copyright is held by the owner/author(s). Publication rights licensed to ACM. <https://doi.org/10.1145/3134669>

Recent trends have suggested that the scenario above is becoming more prevalent in computer-based discussions. As online learning platforms attract millions of people to learn together at on a global scale, asynchronous online discussions have become an important part of the overall experience [10]. At the same time, there is an increasing number of people who prefer learning by connecting knowledge from different materials on the web instead of relying on a single source [30]. In a recent study, students were found to refer multiple learning elements in a collaborative learning activity that involved asynchronous discussion [36]. The confluence of these two trends suggests that the effectiveness of an online knowledge-building discourse can be influenced by users' ability to make references to information from multiple visual materials in the discussion interface [1,6].

Common tools for asynchronous discussion include threaded forums and anchored discussion interfaces (ADI), which allow annotating and referencing visual material with comments on the side [4,13]. However, their capacity to facilitate complex referencing, such as the scenario mentioned earlier, has not been directly studied. To observe how users create complex references using these interfaces, we conducted an observational study. We learned that existing interfaces provide limited support for complex referencing needs, especially those involving *multiple* (e.g., linking various external materials) and *specific* (e.g., a particular passage in a document) referents¹. In threaded forums, the hyperlinking and embedding functionalities require more writing effort and can make discussion posts longer while referring specific objects in visual materials. On the other hand, ADI only allows one object in the material to be anchored for each thread, and one material to be collocated in the interface. Referring more than one object or material in ADI requires manually conceiving additional deixes (words that point to a referent) for each additional referent, which can be cumbersome for the users. These referencing needs are also supported by an interview study conducted with learners experienced in online discussions. In this study, we discovered that there exists a greater need of more visual context along with increased use of external resources during the discussion.

From these qualitative inquiries, we derived three design requirements that drive our iterative design process. The requirements inform three key referencing actions that should be supported by asynchronous discussion interface: 1) Specify referents of varying granularities in visual materials with minimum deixes 2) View referents alongside discussion in the UI and 3) Visualize all the referents and choose which to focus on. We designed and built Korero, an asynchronous discussion interface with two novel UI components, namely, contextual activity window and on-demand widget with two action views. In this manner, Korero provides the necessary visual space and awareness to facilitate complex referencing actions.

To test the efficacy of Korero against conventional forum and ADI, we conducted two lab studies to evaluate two key aspects of referencing: establishing and comprehending references with four referencing tasks of different referential difficulty [6,19]. The findings suggest that Korero is significantly less cumbersome and requires less time and effort than the forum and ADI to create and comprehend references with multiple and/or specific referents. As Korero provides better awareness and visualization of the referents in the discussion, participants felt that it could potentially increase their engagement in the discussion as well as the visual materials being referred.

We believe the lessons from this research will have rich implications on future work intending to support expressive referencing of visual materials in other applications and usage contexts.

This paper makes the following contributions:

- A preliminary investigation that observed how users make complex references to visual materials with existing interfaces, and the challenges they face.

¹ In this paper, we define referents as the objects being referred. As we focus on visual materials such as documents and videos [11], a referent can be the entire material, or any unit of visual information in it.

- An interface prototype designed and implemented to facilitate complex referencing actions identified in the design requirements.
- A two-study evaluation of our prototype that showed a significant reduction in cumbersome as well as the time and effort required to establish and comprehend complex references when compared to existing interfaces.

BACKGROUND AND RELATED WORK

We first review frameworks on making references and building common ground to understand the underlying affordances and behaviors, followed by related work on asynchronous discussion interfaces.

Referring Visual Information to Build Common Ground

Building common ground is an essential activity to improve the efficiency of communication [5]. In a collaborative learning scenario, good common ground facilitates the knowledge building process and improves the understanding of the materials at hand [1]. Common ground can be built verbally or non-verbally [15]. Non-verbal methods, which include gesturing (e.g., pointing), artifacts and spaces, make grounding more efficient by reducing verbal effort and its associated complexity [15]. As people tend to spend the least amount of effort to convey their messages [5], the ability to leverage non-verbal communication becomes crucial.

This paper investigates one non-verbal aspect of building common ground: making references to visual information (see footnote 1 on page 2) [15,22]. The difficulty of making references is contingent on the complexity of the referencing task and the medium of communication [19,22]. There are two general dimensions to referential complexity: codability and discriminability [19]. Codability refers to the linguistic expressivity of the referents, while discriminability refers to the difficulty in which the referent(s) can be differentiated from other visual targets in the referencing context [19]. For example, in a face-to-face setting, we use pointing gestures in shared visual spaces to make references to visual information that is difficult to express with words, such as a *specific* passage in a book (codability), as well as refer to *multiple* target (discriminability). In the example above, a *specific* passage in a book can be pointed out directly instead of having to describe its location and characteristics verbally, and *multiple pointings* can be done to refer various passages in the same book (or different books) to differentiate them from other passages.

A large body of work has studied remote gestures [12,20,21] and awareness widgets [9,15] in different computer-supported cooperative work contexts, including authoring [9,24,32,33], groupware [12,16,17] and many others. In this paper, we focus on facilitating referencing tasks with *multiple* and *specific* referents in the asynchronous discussion context.

Asynchronous Discussion Interfaces that Support Grounding and Referencing

Threaded forums and ADI are two main interfaces commonly used to facilitate asynchronous learning discussions [4,18,26]. Originated from newsgroups, threaded forums have evolved from text-only discussion to hyperlinking of external URLs and embedding of materials (e.g., images and programming code [35]) directly in a discussion post, making grounding and referencing easier to achieve.

ADI adopts the document annotation model by anchoring discussions onto the side of a learning material within the same interface. With ADI, the user can start a discussion by leaving a comment on a specific part of the material. This design allows users to get an awareness of the learning material collocated in the UI while also being able to refer to a specific part of the same material. Discussions on ADI were also found to be more sustained and on-topic in different e-learning contexts [4,8,18,36]. However, some studies found that users prefer forum over ADI for high-level

or general discussion [4,25,36], which raises the question of whether two different discussion interfaces are needed in an online learning environment to support different types of discussions.

Even though much work has been done on the functionalities of ADI and threaded forums, little research has investigated their referencing capabilities with *multiple* and *specific* referents. This is important as it allows users to build common ground more efficiently with referencing actions when there are multiple and/or specific information or visual materials involved. Rather than looking at discussion interfaces as a whole, we investigate design elements that support referencing actions in asynchronous discussion interfaces.

PRELIMINARY INVESTIGATION

Observational Study

We conducted a formative observational study to investigate the referencing capabilities of existing asynchronous discussion interfaces and understand the pain points of their referencing features. We learned how users make references with multiple and specific referents with existing interfaces, which informed our design requirements and process.

Methods. Participants were asked to refer document-based materials while writing a discussion post with a forum interface and ADI. For the forum condition, we used the forum interface in Coursera (<http://www.coursera.org>) as it represents the conventional forum interface in online courses. It allows users to provide hyperlinks in the post and embed images directly. For the ADI condition, we used the NB interface developed by Zyto et al. [36]. NB allows a PDF document to be collocated with the discussion in the same interface. Such collocation is not supported in the forum.

Due to the exploratory nature of our study, we limited the number of referents in each task to a maximum of two to simplify the study process. We chose two documents of similar length (6 pages in PDF format) as the visual materials, one designated as the primary document, and the other as the secondary document. We defined “primary document” as the document collocated with the discussions in ADI, and “secondary document” as the document that is not collocated.

We designed 5 tasks with different numbers and types (primary and/or secondary) of referents. In task 1, participants were asked to refer one specific object (e.g., paragraph, figure) in the primary document. In task 2, they were asked to do the same, but on the secondary document. In task 3, participants were asked to refer two specific objects in the primary document. In task 4, they were asked to refer one object in each of the the primary and secondary documents. We included a control condition with no references to make in the discussion post.

Four participants (2 females, mean age = 27.5) with prior forum usage experiences were recruited from the university. We adopted a within-subject design where each participant used both forum and ADI to attempt each task, but with different stimuli.

Before the study, we provided a short practice of both the forum and ADI interfaces to familiarize participants with the referencing features and general UI elements. We asked participants to think aloud while performing the tasks to facilitate the noting down of any interesting actions and pain points that were raised. We followed up with a brief interview after all tasks had been conducted. All data were collected and analyzed based on the interface and task conditions to inform any similarities and contrasts.

Results. With the forum interface, participants found it easy to refer multiple documents easily (e.g., participants can create multiple links consecutively in their post, such as “[*link 1*], [*link2*]”). However, referring a specific object that is difficult to express with words, such as a particular passage in the document, was perceived to be cumbersome and effortful. We observed that participants had to conceive the right deices (e.g., “first paragraph of...”, “page 6”, etc.) to point out the object’s exact location in the document. This was more difficult in task 3 and 4, where there were two objects to refer. Moreover, since these dexies can be quite long, the resulting increase in words could make the posts more difficult to read. While participants could avoid making the

reference at all by copy and pasting that specific passage (or embedding for images) directly into the post, this would also make the discussion post longer.

For the ADI condition, we observed that making reference to a specific object in the primary document is easier than the forum. However, it fell short when there was more than one referent to refer, or when the referent was located at the secondary document that was not collocated in the ADI.

There were three conditions where the ADI was limited. (1) If the second referent is another object in the primary document, the user had to write additional deixes to point to this object. This is because each discussion thread in ADI can only link to one object in the primary document with its anchoring interface. Moreover, this linkage could only be associated with the entire thread and not individual discourse elements (e.g., words, sentences, paragraphs, etc.) in the discussion posts. (2) If the second referent is the secondary document, the user had to conceive the deixes and include its URL in the discussion post in order to make the reference. (3) Lastly, if the second referent is a specific object in the secondary document, additional referential terms had to be written on top of the deixes in (2) in order to point out the location of that object in the secondary document.

In summary, the study suggested that forum and ADI have different limitations on the “*multiple*” and “*specific*” axes, and neither were able to facilitate referencing with these two characteristics at the same time.

Preliminary Interviews

On top of the observational study, we also conducted preliminary interviews to understand how users utilize and interact with visual materials in an asynchronous learning discussion setting.

Method. We recruited another group of 6 interviewees (5 males, mean age = 27.6) who had used discussion forums in Massive Open Online Courses (MOOC), where all learning materials are given digitally on the platform. We screened the participants before the interview to ensure all had completed at least one MOOC and were frequent forum users in MOOCs. Recruitments were conducted in authors’ network via snowball sampling.

All interviews were conducted in a semi-structured fashion. We started each interview with a self-introduction, followed by asking what MOOCs they had taken and why they took them. Then, we progressed to the main questions by first asking how they used the forum and what role it played in their learning. We then segued into the issues they had encountered on the forum, how they took part in discussions related to the instructional activities (e.g., learning video, quiz, assignment), and how they made reference to them in the discussion. We concluded the session by asking what features they would like to have in the forum, specifically those related to learning materials and instructional activities.

We collected field notes and transcribed the audio recordings for further analysis. We adopted the inductive thematic analysis approach to code the data and identify common themes in the transcripts [3].

Main Findings. We identified 3 main themes in our data. They not only informed us how our interviewees used and interacted with learning materials and activities in the discussion but also presented some interesting design opportunities and motivations.

Need for more context in the discussion. As a lot of discussions in the forum are related to the instructional activities, most of our interviewees raised the need for more information on those activities in the discussion interface. One interviewee felt that the ability to incorporate this information could provide more context to his posts. Another expressed that being able to closely relate the content to the discussion could simplify his explanation and allow others to better understand his idea. Existing research also echoed similar sentiments [7,27], recommending a

tighter integration of the discussion with the instructional activities to bring more learning benefits for the students.

Limitations of text-only discussion. Interviewees also brought up the limitations of using only text in the discussion. Some noted that it can be difficult to express their ideas with text and wished that they could use and incorporate other modalities seamlessly, such as videos and audios. Results from a recent deployment of a multimodal annotation system also echoed similar needs to support modalities beyond text [34].

Use of external resources in the discussion. When our interviewees encountered questions in MOOCs, apart from reviewing the instructional activities or the learning materials provided, most would also search for external learning resources on the Internet to supplement their learning and discussion.

DESIGN REQUIREMENTS

Based on our literature review, observational study (OS) and interviews (IN), we came up with the following design requirements by examining the referencing actions in establishing and comprehending references that our interface should facilitate.

DR1: Users can refer to varying granularities of referents, from *specific* to *general*, with minimum deixes. [25, OS].

DR2: Users can view the referent materials/objects alongside the discussion [21, IN].

DR3: Users can visualize all the referents and choose which to focus on [6,23, OS].

We summarized the capabilities of the forum and ADI through the lenses of our design requirements in Table 1.

Table 1 Capabilities of the discussion interfaces with respect to the design requirements.

Design Requirements	Forum	ADI	Korero
DR1: Specific referencing	No	Only 1	Multiple
DR2: View referents alongside the discussion	No	Limited	Yes
DR3: Visualize referents and choose which to focus on	No	No	Yes

FACILITATING COMPLEX REFERENCING IN ASYNCHRONOUS DISCUSSION INTERFACE

The Korero (means discussion in Māori language) interface was designed to facilitate referencing with multiple and specific visual material referents in asynchronous discussion. In this section, we describe its key features and components, together with our design process and rationales.

General Interface Features

To collocate discussions with visual material, Korero consists of two main windows - the activity window and the discussion window (Figure 1a), to keep both in sight simultaneously [19,23,IN].

To facilitate referencing actions informed by our design requirements, users need to refer multiple referents to any discourse element in the discussion [OS]. Instead of using traditional hyperlinks, which only support 1:1 linking, Korero supports multi-linking to create 1:N links between the referential term and the referents (Figure 1b). Unlike ADI that only allows assigning a single referent to the entire thread, in Korero any sub-texts of a post can be the referential term.

Providing Visual Space to Support Contextual Actions

Unlike face-to-face discussion, there is limited visual space in an asynchronous discussion interface to allow interlocutors to point to a specific object in the visual material, or visualize the referent, while still keeping the discussion in context. To solve this problem, we draw on the fact that users

do not necessarily have to view the main activity/material while specifying or browsing the referents in the interface. Thus, we designed the activity window (left window in Figure 1a) to be flexible in being replaced by the materials or referents associated with the ongoing referencing actions.

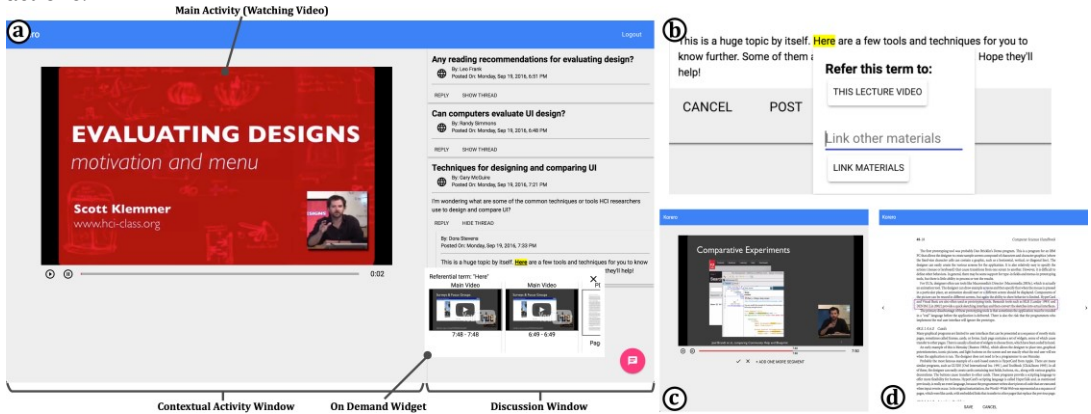


Figure 1. (a) Overview of our Korero interface. The contextual activity window shows the materials (e.g., video or documents). Discussion window on the right shows all the discussion threads on the main material. (b) Multi-linking pop-up. After selecting any texts (referential term) in the post, the pop-up would appear beneath. (c) Specifying referents in video material and (d) text material after the material is selected via the multi-linking pop-up.

The **contextual activity window** supports two key referencing actions. The first is specifying different granularities of referents in the visual material (**DR1**). These granularities can be: (1) The entire material, (2) One section/object of the material or (3) Multiple sections/objects of the material. After linking a material with the multi-linking pop-up (Figure 1b), the material is shown in the activity window temporarily.

For video-based material, users can specify a time frame or a video segment with the scrubber UI (see Figure 1c). To refer more than one referent in the same video, users can click the “Add One More Segment” button after specifying the first referent. For text-based material, users can draw a bounding box in the document to specify a section, or multiple boxes if there are multiple referents to refer in the document (Figure 1d). If no object is specified by the user, the entire material will be automatically recognized as the referent, which allows users to make references to the materials in general.

The second referencing action supported by the contextual activity window is visualizing the referents within the UI (**DR2**). This visualization functionality is important for both the posters and the readers, which provides them the necessary visual space to see the referent clearly and in detail. With the contextual activity window, the posters can review the referents they refer to while writing their discussion posts. For the readers, the window allows them to comprehend the referents without leaving the discussion interface, which provides both focus and context side-by-side. For an object-level referent (e.g., video timestamp, passage in document), Korero visualizes the object directly instead of showing the entire material.

Accessing and Providing Awareness of the Referents

While the contextual activity window could provide the needed visual space for certain referencing actions, it does not (1) provide users with an awareness of all the referents of a referential term, nor does it allow them to (2) get a quick glimpse at the referents and choose which to focus on. In our

design process, we realized that existing interfaces lack the necessary feature to facilitate these actions.

After a few iterations, we created an **on-demand widget with two action views** to provide the necessary awareness, as well as the ability to view and act on the referents of a referential term easily (see Figure 1a and Figure 2). A thumbnail view of the referents is provided in the widget, and the user can invoke the widget on-demand (Figure 2). We also designed the prototype to handle different mouse events (hovering and clicking) on both the widget and the thumbnail to enable two different action views – quick glancing and multitasking.

The on-demand widget supports two key referencing actions. The first is editing the referents of a referential term (**DR3**). After the posters have specified referent(s) in the material (Figure 1c and 1d), thumbnail(s) of the material are shown in the widget, with each thumbnail representing one referent (see Figure 2). Beneath each thumbnail, the location of the referent in the material is shown (e.g., in-video timestamp or page number). Hovering on the thumbnail enables a quick preview of the referent in the activity window. Clicking the thumbnail, on the other hand, allows users to edit and specify the location of the referent in the material again. To allow users to refer other materials to the same referential term, we include a “Refer More” button to bring up the multi-linking pop-up again for the users (Figure 2).

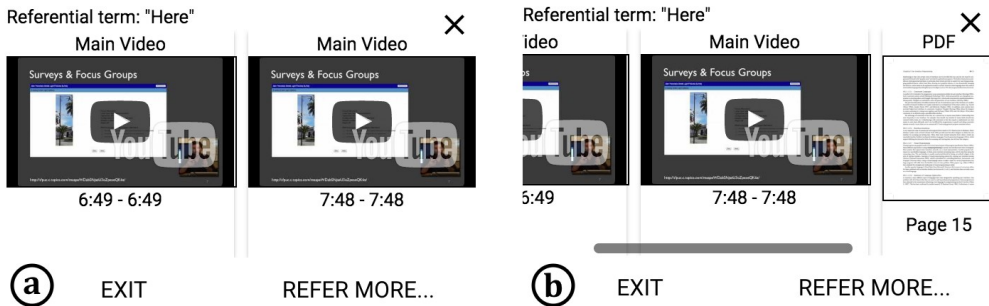


Figure 2. (a) On-demand widget. (b) To accommodate more thumbnails, we include a horizontal scrollbar to provide additional horizontal space.

The second referencing action supported by the widget is comprehending the referents/references (**DR3**). When the user hovers on a referential term in an existing post, the widget appears directly beneath the term (Figure 3). In the widget, the user can hover on each thumbnail to glimpse and visualize the referent in the contextual activity window. By hovering on different thumbnails located side-by-side (see Figure 3), the user can get a quick glance of each referent in the activity window. Holding the hover over a particular thumbnail allows users to focus on that particular referent in the activity window until moved beyond the thumbnail.

While a hovering action enables quick glancing of the widget and the referents, a clicking action, on the other hand, anchors them in the interface so that users can interact with other interface elements while keeping the selected widget or referent in users' view within the interface. For example, clicking the referential term fixes its referent widget at the bottom of the discussion window so that users can hover on other referential terms to see their referents while keeping the referents of the earlier term in sight. Clicking the thumbnail, on the other hand, fixes its referent in the activity window so that users can interact with the discussions or other referential terms while comprehending the referent's content in the activity window simultaneously.

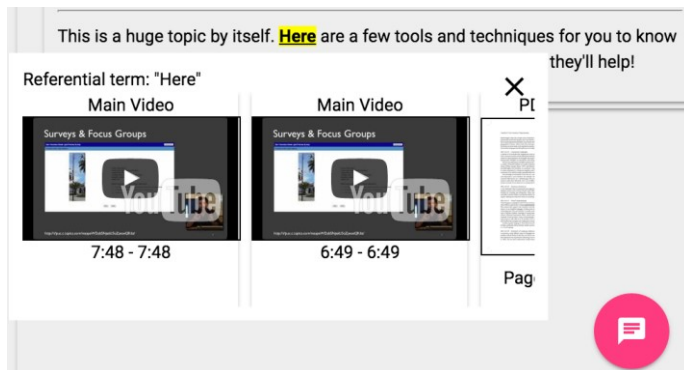


Figure 3. Comprehending the referents of a referential term. In the widget, the user can hover on the different thumbnails side-by-side to get a quick glance of each referent in the activity window.

Implementation

We packaged the interface components into a web-based prototype. We implemented Korero using Google's Polymer web component library (<https://www.polymer-project.org>) with Firebase running as the backend database. Due to Cross-origin resource sharing (CORS) restriction, our implementation focused on linking PDF documents stored in the Firebase storage and avoided linking HTML pages outside the site. Videos were embedded into our interface from Youtube via the Youtube API (<https://www.youtube.com/yt/dev/api-resources.html>).

In our prototype, each video/main content interface retrieves an array of discussion threads. Each discussion thread has the following attributes: 1) A unique identifier 2) The video/main content interface identifier where the thread is located 3) Text body (in HTML format) of the first post of the thread 4) An array of referential terms (and their referents) in the text body of the first post of the thread 5) Metadata such as user ID and posting date & time 6) Thread title 7) An array of discussion post IDs the thread has. If the thread is created in relation to a particular timestamp in the video, it also includes the video timestamp information. Each discussion post in a thread, sans the first post, has the same attributes (1, 2, 3, 4, 5) as a discussion thread described above, plus the thread's ID where the post belongs to.

Each referential term has a unique identifier and the thread/post's ID where the referential term is used. It also has an array of referents, which holds either a video referent or PDF referent. Each referent has a unique identifier and the referential term ID where the referent is linked. It also has a type attribute to know if the referent is a video or a PDF document. For PDF referent, it has the URL of the PDF document, page number where the referent is located, a bounding box object that highlights the referent in the page, and a thumbnail image of the PDF for the on-demand widget. For video referent, it has a Youtube ID and a timestamp attribute, which links the referent to a particular timestamp in the video. If no timestamp is assigned, the referent will be the whole video. Unlike the PDF referent, it is not necessary to save a thumbnail image of the video referent as the smaller version of the video player can serve as the thumbnail itself. These referents are connected to the text body by encompassing referential terms with span tags that hold the referent IDs. The system loads the referent when the user hovers the referential term inside the span tag.

All data are loaded from the database and pushed to the interface in an on-demand fashion (i.e., when the referential term is hovered upon) to optimize the prototype's performance.

EVALUATION

To test the referencing efficacy of the Korero interface, we conducted a comparative evaluation of Korero against the forum interface and ADI on two key aspects of referencing: the establishment and comprehension of the references [6]. We evaluated the establishment and comprehension processes independently in two separate studies instead of combining them together in one conversational study, as users in asynchronous discussion cannot repair and coordinate the references in real time due to its asynchronous nature. This usage context is reflected in our evaluation.

To make the visual design and baseline UI elements (e.g., video scrubber, button) consistent across the study conditions, we implemented the forum and ADI using the same interface framework as Korero. Besides, to make a fairer comparison of the core referencing model of each interface, we also added hyperlinking functionality into our ADI prototype, as previous systems like NB did not support [36].

We included both video-based and text document-based visual materials in our evaluation as they are commonly used in many online courses. As in most online learning interfaces, the video was designated as the material anchored onto Korero and ADI, while documents have to be referred externally from the discussion interface. In the forum interface, both videos and documents have to be referred externally as there is no side-by-side view of the material and discussion in the forum.

Our evaluation seeks to answer the following research questions (RQs):

RQ1: Does Korero facilitate efficient and easy establishment and comprehension of references with multiple and specific referents?

RQ2: How do users create references with multiple and specific referents? What can we learn from their referencing behaviors and preferences to support rich and expressive referencing?

RQ3: What are the benefits of Korero for simpler references (singular or non-specific referents)? How could Korero influence behaviors around the referencing actions, such as users' engagement with the materials being referred?

Designing Referencing Tasks

We answer the RQs above by comparing the efficacy of different interfaces for supporting the establishment and comprehension tasks. We devised a set of different referencing tasks with varying levels of difficulty. The parameters that formulate the difficulty of a task include media types, specificity, and the number of referents. It is worth noting that the focus of our comparative evaluation was on the differences between the three *interfaces*, not the task contexts. Thus, the diverse referencing contexts served as a set of lenses to crystalize the comparative task setting rather than making a direct comparison between the contexts.

We delineate the difficulty of referencing multiple and specific referents in asynchronous discussion context with the theoretical basis explicated in the referential complexity framework described earlier (see Related Work) [19]. In regards to discriminability, a referencing task is harder when referents are specific than general, and multiple than singular. For example, there is a higher target ambiguity when referring to objects (e.g., timestamps of a video or passages of a document) than referring to materials, such as the video or the document *per se*. As for codability, material-level referents have higher codability than object-level referents in general as material-level referents can be referred with simple labels such as “this video” or “the document”.

We devised four tasks to cover different levels of task difficulty in terms of the specificity and the numbers of referents (see Table 2). Referencing task 1 (RT1) has less specific (material level) but multiple referents. RT2 has a singular but specific referent (timestamp) in the video since video material is anchored onto the interface in the Korero and ADI conditions. RT3 has two specific referents in the video. Finally, RT4 goes one step further from RT3 in terms of difficulty by adding two specific referents in a document to refer, on top of the two in the video (see Table 2). It not only

increases the number of referents, but also lowers the codability of the referents as we deliberately designed the specific referents in the document to be arbitrary passages of 3-5 lines, which is more difficult to refer than video timestamp since there are no straightforward labels for those passages. We did not include a task with singular and non-specific referent since it would be trivial and there would be no difference between the interfaces for a simple referencing task with a singular and non-specific referent.

Table 2. The four referencing tasks and the research questions (RQs) they address.

Referencing Tasks (RT)	Related RQs		
	1	2	3
RT1: Refer to 1 video and 1 document	-	√	√
RT2: Refer to 1 timestamp in a video	-	√	√
RT3: Refer to 2 timestamps in a video	√	√	-
RT4: Refer to 2 timestamps in a video and 2 arbitrary passages in a document	√	√	-

Apparatus

Both studies in our evaluation were conducted on a 15-inch MacBook Pro (2013 model) running OS X El Capitan. A 27-inch monitor was connected to the laptop as the secondary display. The experimental software was operated on the Google Chrome browser. Visual materials used in our evaluation were in the same subject area (Cryptography and Information Theory) derived from the Computer Science courses on Khan Academy (<https://www.khanacademy.org/computing/computer-science>). In our recruitment, we ensured the participants had similar background on those subject areas.

STUDY 1: ESTABLISHING REFERENCES

In the first study, we investigated how participants established references in a discussion post.

Participants

Twelve participants (10 females, mean age = 20.75) were recruited from the host university. All were proficient in written English at the advanced level and above. Participants were given course credits for their time.

Procedures

In each trial, the participants were given a simulated discussion posting scenario with one of the RTs in Table 2 as the references condition of the trial. In the scenario, they read a discussion post and decided that they could answer the question in the post. The questions were designed in the following format: *Which sections of the materials are [adjective] (e.g., straightforward, easy, etc.) for you?* To reply the post, participants were instructed to use the given interface in the trial to refer to the materials or objects specified in the instructions of the trial, and state that the materials/objects were [adjective] for them. The trial ended after the participants posted their reply.

The instructions, which show the specifications of the materials/objects to refer, were shown on the secondary display, while the actual post writing tasks were conducted on the primary display (laptop). In each trial, the instruction was shown before the commencement of the actual task (with the interface) to allow participants to understand the specifications beforehand. While the instruction was shown, we did not reveal the interface that would be used in the trial to prevent participants from developing a referencing strategy at this stage, which could be a potential confounding factor. After the instruction and specification were understood, the participants clicked a button on the instruction page to reveal the interface and start the task.

The specifications consisted of a visual representation of the materials/objects to refer and their URLs. We did not include textual representations nor detailed descriptions of the objects' location to prevent direct usage of the deixis. For video-based objects (timestamp), a screenshot of the frame was shown in the specification, together with its approximate location (10 seconds range) in the video. During the task, participants were told to identify and refer to the exact timestamp shown in the video frame. For document-based objects (passages of 3-5 lines), a screenshot of the page with the passage(s) highlighted was shown in the specifications, together with the page number. In the task, the participants were told to refer to the passage(s) as accurately as possible.

Design

Since the 4 referencing tasks in Table 2 were chosen independently to represent references conditions of interest to our RQs, we did not conduct a multivariate analysis across tasks. For each referencing task, a repeated-measures within-subject design was used, with discussion interface (*forum*, *ADI*, *Korero*) as the independent variable. Each interface was repeated twice (2 trials), and each trial used visual materials with different content to prevent learning effect. The trials were randomized to prevent participants from anticipating the interface of the coming trial while reading the instructions. The sequence of referencing tasks blocks was counterbalanced across participants using Latin Square.

The study session lasted between 90 minutes to 2 hours for each participant. After filling in the pre-study form, a demo of the discussion interfaces and their referencing features was given. Before each referencing task block, the experimenter first explained the task and the general strategies that could be used to accomplish them, followed by a set of practice trials to familiarize participants with the UI and the tasks. After each block was completed, the participants could take voluntary breaks.

Results

Due to space constraints, we included the data, F-value, effect size (η^2), Chi-square values and mean values in the respective figures, and only reported the differences that were significant.

Completion Time of Establishing the References. We measured completion time as the time taken to establish the reference and post the reply. The data were analyzed using repeated-measures ANOVA with Greenhouse-Geisser corrections for sphericity violation, and pairwise t-tests with Bonferroni correction for post-hoc analysis.

We found a significant main effect of the UI on completion time in RT2, RT3, and RT4 (all $p < .001$, Figure 4), but not in RT1. In RT2-RT4, where the referents are specific on the object level, Korero was significantly faster than ADI and forum, and ADI was notably faster than the forum (all $p < .05$).

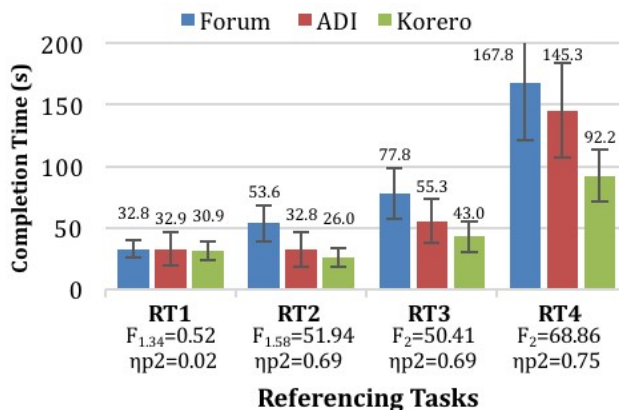


Figure 4. Completion time of each UI in the referencing tasks in Study 1. Data labels are mean values, and error bars represent standard deviation.

Cumbersomeness and Writing Effort of Establishing the References. After each trial, participants rated the cumbersomeness and writing effort in establishing the references on a 7-point scale (1: strongly disagree, 7: strongly agree). Such measures have been used in previous studies [2,23,28]. We analyzed the data using Friedman’s ANOVA and Wilcoxon signed-rank tests with Bonferroni correction for post-hoc analysis.

There was a significant main effect of the UI on cumbersomeness in all 4 RTs (all $p < .05$, Figure 5a). In RT1, where its referents are on the material-level, ADI was more cumbersome than the forum ($p = .011$). Whereas in RT2 (refer a timestamp in the video), the forum was more cumbersome than ADI and Korero (both $p < .001$). In RT3 and RT4, where there are multiple and specific referents, both forum and ADI were significantly more cumbersome than Korero (all $p < .001$).

Writing effort wise, we also saw a significant main effect of the UI in all tasks (all $p < .05$, Figure 5b), with Korero taking the least effort in overall. Apart from RT2 where there was only one specific, object-level referent, in other tasks Korero required significantly less writing effort than ADI to establish the reference (all $p < .01$). The significance was even higher when comparing Korero with the forum in RT2-4 (all $p < .001$). Lastly, in RT2 and RT3, ADI’s writing effort was also significantly lesser than the forum (all $p < .01$).

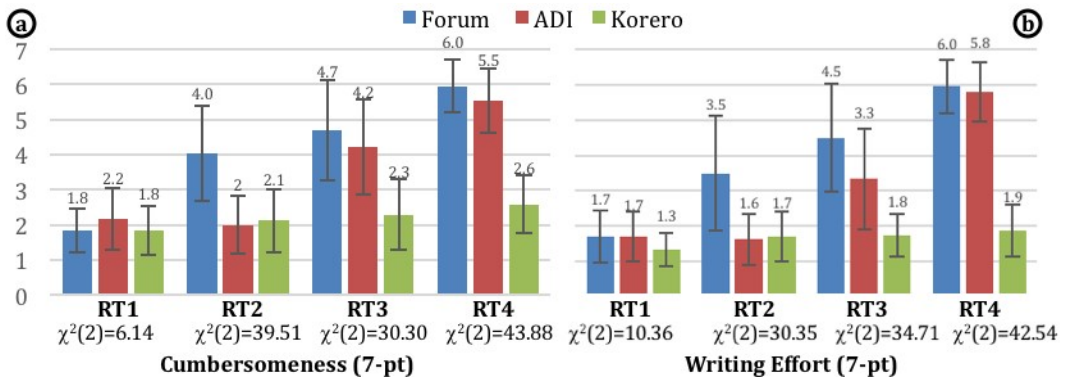


Figure 5. (a) Cumbersomeness and (b) writing effort of each UI in the referencing tasks in Study 1. Data labels are mean values, and error bars represent standard deviation.

Overall Preferences. At the end of each task block, we asked the participants to assign a preference score of 1 to 7 to each interface, with 7 as the most preferred and 1 as the least preferred. In RT1, forum (5.3) and Korero (5.5) received a similar score that is slightly higher than ADI (4.9). In RT2, ADI (6.2) and Korero (5.8) got a higher score than the forum (3.6). In RT3, forum (3.6) and ADI (3.9) received a similar score that is lower than Korero (6.4), and the same trend was also observed in RT4 (forum: 2.9, ADI: 3.3, Korero: 6.3).

STUDY 2: COMPREHENDING REFERENCES

The second study looked at how participants comprehend references in a discussion post to extract information.

Participants

Twelve participants (6 females, mean age = 23.41) were recruited from the host university with the same recruiting requirements used in Study 1. None of them had taken part in Study 1.

Procedures

In each trial, the participants were given a simulated discussion reading scenario with one of the RTs in Table 2 as the references condition of the trial. In the scenario, they read a post that contains references to visual materials or objects that they want to examine further. To motivate the examining behavior in each trial, we designed a question related to the referred materials or objects and asked the participants to answer them. To answer the questions, the participants first had to comprehend the references established in the post, locate the materials/objects, and extract the relevant information. We designed the questions in a way that it does not require participants to have any prior knowledge on the subject matter to extract the information. For example, one question asked for the color of the bar chart that appears on the video timestamp referred. The trial ended after the participants inputted their answers into a dialog box in the task window.

We adopted the same trial setup used in Study 1 for Study 2. The instructions, which now show the questions related to the materials/objects being referred in that trial (instead of the referents' specifications in Study 1), were displayed on the secondary display (with the interface) before the actual task started. Apart from the trial setup, we also extended Study 2's experimental context from Study 1 to ensure the referencing topics and difficulties were consistent. To do so, we derived the stimuli postings in Study 2 from the discussion replies created by the participants in Study 1. We first analyzed all the replies from Study 1 and selected those that did not contain acronyms and grammar/sentence issues to ensure they are understandable. After that, we chose replies with decent communicative efficiency by hiring 2 external coders to count the number of deixes written in each reply and select those with the least deixes [14]. Interrater reliability was good, with Intraclass Correlation Coefficient (ICC) (3,2) = 0.93 [29]. Lastly, we picked replies that were more consistent in wording and structure by selecting postings from as fewer participants as possible.

Design

The design of Study 2 was identical to that of Study 1 (see Study 1: Design section for details). Study 2 lasted between 1 hour to 90 minutes for each participant.

Results

Due to space constraints, we included the data, F-value, effect size (ηp^2), Chi-square values and mean values in the respective figures, and only reported the differences that were significant.

Completion Time of Comprehending the References. We measured completion time as the time to comprehend the references and answer the questions related to the referents. It was analyzed with the same statistical test used in Study 1.

There was a significant main effect of the UI on completion time in all 4 RTs (all $p < .001$, Figure 6). In RT1, RT2, and RT3 (refer to Table 2 for descriptions of each RT), we saw the same trend where Korero and ADI were significantly faster than the forum (all $p < .01$). The differences between Korero and ADI in RT1-3 were not significant. In RT4, however, where the reference refers to 2 specific objects in both the video and document, participants took significantly less time to complete the task with Korero compared to both ADI and forum (both $p < .001$).

Cumbersomeness of Locating the Referents and Mental Effort of Comprehending the References. After each trial, we asked the participants to rate the cumbersomeness of locating the referents and the mental effort of comprehending the references with the same 7-point scale used in Study 1. The data were analyzed with the same statistical tests used in Study 1.

We saw a significant main effect of the UI on cumbersomeness in all 4 RTs (all $p < .001$, Figure 7a). In RT1, RT3, and RT4, where there are multiple referents, Korero was rated significantly less cumbersome than the forum and ADI in locating the referents (all $p < .05$). ADI was also significantly less cumbersome than the forum in RT2 and RT3 that refer to timestamp(s) in the video (both $p < .01$).

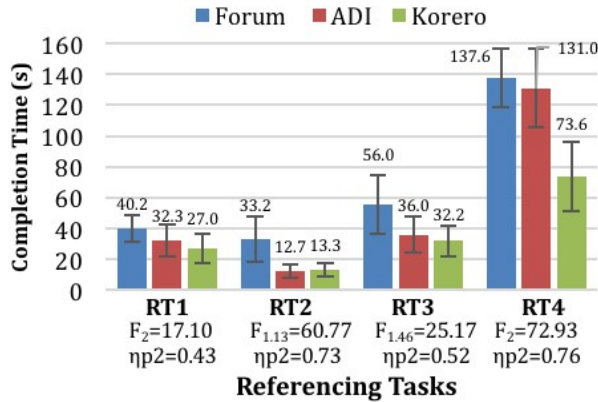


Figure 6. Completion time of each UI in the referencing tasks in Study 2. Data labels are mean values, and error bars represent standard deviation.

For the mental effort of comprehending the references, a significant main effect of the UI was found in all 4 tasks (all $p < .01$, Figure 7b). In RT1, where the reference refers to both video and document in general, participants reported needing significantly less mental effort to comprehend the references with Korero compared to the forum ($p < .05$). In RT2, RT3, and RT4, where the references refer to specific objects in the materials, both ADI and Korero were rated significantly less effortful than the forum (all $p < .05$). Lastly, between Korero and ADI, significances were only found in references with multiple and specific referents (RT3 and RT4, both $p < .01$), with Korero being rated requiring lesser mental effort than ADI.

Overall Preferences. At the end of each task block, we asked the participants to assign a preference score of 1 to 7 to each interface. In RT1, Korero was the most preferred (6.3), followed by ADI (5.1) and the forum (4.1). In RT2, ADI (6.3) and Korero (6.1) were similarly more preferred than the forum (3.7). In RT3 and RT4, where the references have multiple and specific referents, Korero (RT3: 6.3, RT4: 6.5) was more preferred than ADI (4.7, 4.0) and the forum (3.0, 2.3).

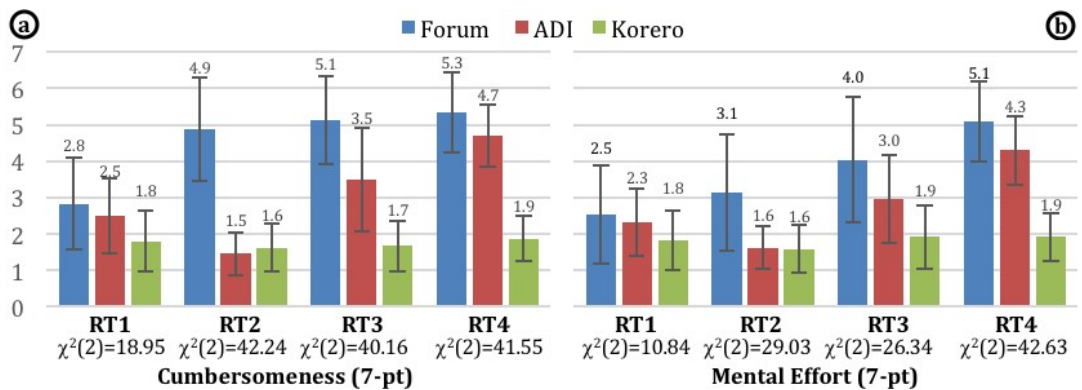


Figure 7. (a) Cumbersomeness and (b) mental effort of each UI in the referencing tasks in Study 2. Data labels are mean values, and error bars represent standard deviation.

DISCUSSION

We discuss the results from Study 1 and Study 2, as well as the subjective comments from the participants in post-study interviews, through the lenses of the three RQs we proposed. We also discuss the limitations of our studies.

RQ1. Establish and Comprehend Reference with Multiple & Specific Referents

In referencing tasks with multiple and specific referents (RT3 and RT4), we found that Korero was significantly faster, less cumbersome, less effort requiring and much preferred in most of the comparisons with ADI and the forum in establishing and comprehending references. This shows that the contextual activity window and the on-demand widget in our interface have successfully provided the necessary visual space and awareness for facilitating referencing actions needed for multiple and specific referents.

Participants' feedback on the design of our interface's main components were mostly positive. The contextual activity window and on-demand widget were deemed natural and not distracting. P9 in Study 1 felt that the display of different materials in the contextual activity window based on user's current action is intuitive, while P10 found the on-demand widget useful for checking what she has referred. While some found a slight learning curve with Korero, they all agreed that it was more convenient and efficient once they have learned it, to the extent that they can quickly adapt to the interface in the relatively short time of the user studies. In the post-study interviews, P3 and P12 from Study 1 voiced the need to support and facilitate elaborate references in certain scenarios, such as collaborative discussion in Google Docs and MOOCs.

RQ2. Users' Referencing Behaviors and Preferences

Our studies also revealed several interesting behaviors and insights on how users make references with multiple and/or specific referents. First, we found that ADI was not much better than the forum in referencing task (RT3) with two specific referents with high codability (video timestamp). In Study 1, 5 out of 12 participants gave forum a higher preference score than ADI in RT3. When asked why, they explained that ADI took more mental effort since they had to apply two different referencing strategies – direct anchoring with the first referent, and writing deixis with the second referent – in order to make the references. This shows the importance of a consistent referencing method/model that is adaptable for complex references while designing referencing features for asynchronous discussion interface.

Second, we found that most participants (9) preferred to use the features provided by the interface to make the references instead of writing the deixis manually. This is valid even in simpler referencing task (RT1), where half of our participants (6) preferred our interface over the forum. We observed that the key lies with how much of the referencing actions are being afforded by the interface features, and their usability. For example, in ADI, in order to refer to the video timestamp anchored to the thread, the users have to write "*the frame I referred in the thread*" to create an explicit connection to the anchor, especially when the thread contains a reference to two different video timestamps (RT3). With Korero, because the connections are already embedded in the design of the linkage between an on-demand widget and referential term, users would only have to write "*these frames*" to create the connection.

Third, our participants also contrasted the referencing actions in Korero with the forum while dealing with specific referents. In the forum, many would attempt to copy and paste the document passages directly into the discussion post. However, P3 from Study 1 and P5 from Study 2 felt that copy-paste not only requires more effort than Korero, but it could also introduce more text into the discussion post, which could overwhelm the readers and make the post harder to read. Another pitfall of copy-paste is the loss of context around the referents in the visual materials since copy-paste can only include the referent into the discussion post, not its surrounding contents in the material. This could lead to potential misunderstanding of the referent and introduce more frictions

for the readers to go through the materials in detail. The way referent is being visualized in the contextual activity window of Korero allows users to see the referent's surrounding context in the visual material easily.

RQ3. Implications on Making References and Beyond

Our studies suggest that Korero can benefit general referencing actions and beyond. First, Korero makes it easy for readers to see what other posters are referring to in the interface. Reflecting on her past experiences in using the forum, P4 from Study 2 recalled she often had to read deixes that were confusing and unclear to her, and she felt that Korero could mitigate this issue. From the poster's perspective, P12 from Study 1 thought that using Korero allows her to double-check her references easily before posting.

Second, Korero can potentially encourage readers to check out the materials referred in the discussions. When we asked Study 2's participants how often they checked out the materials referred in the forum they had used before, most told us that they rarely do so unless it is mandatory or referred by the course instructor. They cited a general reluctance to switch to a different tab or window to view the materials as one of the main impediments. After the study, P9 and P10 from Study 1 felt that Korero's capability to check out the referents within the same interface could encourage learners to check out the materials. This was echoed by most of our participants in Study 2, including P10 and P12, who brought up this sentiment themselves.

Limitations

Our evaluation has a few limitations that need to be taken into account while applying the lessons from this work. First, we did not study every possible referencing task within the codability and discriminability parameters in our study context, which could answer other interesting questions on referencing and discussions. For example, can our interface be used equally well in discussions with less need for making references? Can our interface handle a large number of referents in a usable manner? While our studies and post-study interviews have shed some hypothetical light on these questions, further study is necessary to answer them concretely. Second, even though we designed simulated discussion posting and reading tasks for our evaluation, the tasks were not conducted in a natural setting. Thus, it remains to be seen whether our interface has any drawbacks in actual usage. We plan to carry out an ecological validity investigation as our future work.

CONCLUSION AND FUTURE WORK

This paper investigates the design of an asynchronous discussion interface that facilitates the establishment and comprehension of references with multiple and specific referent materials. We evaluated our interface and found that it is less cumbersome and requires less time and effort compared to existing interfaces. Our evaluation also suggests potential benefits beyond referencing that we would like to explore in our future work. First, we plan to deploy our interface in actual courses with substantial learning discussions held online to investigate its efficacy in an ecologically valid setting. We are interested in studying its impact on students' engagement in the discussion activities and the learning materials, as well as investigating its impact on learning gained by administering pre-tests and post-tests. As our work targets a general feature of the discussion system commonly used for peer learning, the findings in this paper can be generalizable to learning activities that involve asynchronous discussion, such as group assignments and collaborative video watching, where multiple learning materials could be involved. Second, we plan to explore other meaningful interactions on the referent objects/materials to better utilize visual materials in the collaborative discourse. Third, we intend to support other forms of visual materials (e.g., webpages, images) on top of PDF documents and videos in the next iteration of our interface.

ACKNOWLEDGEMENTS

The authors would like to thank the course staff of NM2103 and CS4249 from the National University of Singapore for study advertisement and recruitment. Soon Hau Chua is supported by the NUS Graduate School for Integrative Sciences and Engineering (NGS) Scholarship. This work was supported by the National University of Singapore Academic Research Fund T1 251RES1617 and the Institute for Information & Communications Technology Promotion (IITP) grant by the Korea government (MSIT) (No.2017-0-00537, Development of Autonomous IoT Collaboration Framework for Space Intelligence).

REFERENCES

- [1] Michael Baker, Tia Hansen, Richard Joiner, and David Traum. 1999. The role of grounding in collaborative learning tasks. *Collaborative learning: Cognitive and computational approaches*: 31–63.
- [2] Aaron Bangor, Philip T Kortum, and James T Miller. 2008. An empirical evaluation of the system usability scale. *Intl Journal of Human-Computer Interaction* 24, 6: 574–594.
- [3] H Russell Bernard and Gery W Ryan. 2009. *Analyzing qualitative data: Systematic approaches*. SAGE publications.
- [4] A J Brush, David Barger, Jonathan Grudin, Alan Borning, and Anoop Gupta. 2002. Supporting interaction outside of class: anchored discussions vs. discussion boards. In *Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community*, 425–434.
- [5] Herbert H Clark and Susan E Brennan. 1991. Grounding in communication. *Perspectives on socially shared cognition* 13, 1991: 127–149.
- [6] Herbert H Clark and Deanna Wilkes-Gibbs. 1986. Referring as a collaborative process. *Cognition* 22, 1: 1–39.
- [7] Derrick Coetzee, Armando Fox, Marti A Hearst, and Bjoern Hartmann. 2014. Chatrooms in MOOCs: all talk and no action. In *In Proc. of ACM L@S '14*, 127–136.
- [8] Brian Dorn, Larissa B Schroeder, and Adam Stankiewicz. 2015. Piloting TrACE: Exploring spatiotemporal anchored collaboration in asynchronous learning. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 393–403.
- [9] Paul Dourish and Victoria Bellotti. 1992. Awareness and coordination in shared workspaces. In *Proceedings of the 1992 ACM conference on Computer-supported cooperative work*, 107–114.
- [10] Michael Eisenberg and Gerhard Fischer. 2014. MOOCs: A perspective from the learning sciences. *Proceedings of ICLS 2014*.
- [11] Susan R Fussell, Robert E Kraut, and Jane Siegel. 2000. Coordination of communication: Effects of shared visual context on collaborative work. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, 21–30.
- [12] Susan R Fussell, Leslie D Setlock, Jie Yang, Jiazhi Ou, Elizabeth Mauer, and Adam D I Kramer. 2004. Gestures over video streams to support remote collaboration on physical tasks. *Human-Computer Interaction* 19, 3: 273–309.
- [13] Fei Gao, Tianyi Zhang, and Teresa Franklin. 2013. Designing asynchronous online discussion environments: Recent progress and possible future directions. *British Journal of Educational Technology* 44, 3: 469–483.
- [14] Darren Gergle, Robert E Kraut, and Susan R Fussell. 2013. Using visual information for grounding and awareness in collaborative tasks. *Human-Computer Interaction* 28, 1: 1–39.
- [15] Carl Gutwin and Saul Greenberg. 2002. A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work (CSCW)* 11, 3–4: 411–446.
- [16] Carl Gutwin, Mark Roseman, and Saul Greenberg. 1996. A usability study of awareness widgets in a shared workspace groupware system. In *Proceedings of the 1996 ACM conference on Computer supported cooperative work*, 258–267.
- [17] Carl Gutwin, Gwen Stark, and Saul Greenberg. 1995. Support for workspace awareness in educational groupware. In *The first international conference on Computer support for collaborative learning*, 147–156.
- [18] Mark Guzdial and Jennifer Turns. 2000. Effective discussion through a computer-mediated anchored forum. *The journal of the learning sciences* 9, 4: 437–469.
- [19] Michel Hupet, Xavier Seron, and Yves Chantraine. 1991. The effects of the codability and discriminability of the referents on the collaborative referring procedure. *British Journal of Psychology* 82, 4: 449–462.
- [20] David S Kirk and Danaë Stanton Fraser. 2005. The effects of remote gesturing on distance instruction. In *Proceedings of the 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years!*, 301–310.
- [21] David Kirk and Danae Stanton Fraser. 2006. Comparing remote gesture technologies for supporting collaborative physical tasks. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, 1191–1200.
- [22] Robert E Kraut, Susan R Fussell, and Jane Siegel. 2003. Visual information as a conversational resource in collaborative physical tasks. *Human-computer interaction* 18, 1: 13–49.
- [23] Benjamin Lafreniere, Tovi Grossman, and George Fitzmaurice. 2013. Community enhanced tutorials: improving tutorials with multiple demonstrations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1779–1788.

- [24] David W McDonald, Chunhua Weng, and John H Gennari. 2004. The multiple views of inter-organizational authoring. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, 564–573.
- [25] Jakko van der Pol, Wilfried Admiraal, and P Robert-Jan Simons. 2006. The affordance of anchored discussion for the collaborative processing of academic texts. *International Journal of Computer-Supported Collaborative Learning* 1, 3: 339–357.
- [26] J Van Der Pol, W Admiraal, and R-J Simons. 2003. Grounding in electronic discussions: Standard (threaded) versus anchored discussion. In *Designing for change in networked learning environments*. Springer, 77–81.
- [27] Carolyn Penstein Rosé and Oliver Ferschke. 2016. Technology Support for Discussion Based Learning: From Computer Supported Collaborative Learning to the Future of Massive Open Online Courses. *International Journal of Artificial Intelligence in Education*: 1–19.
- [28] Susana Rubio, Eva D'íaz, Jesús Martín, and José M Puente. 2004. Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods. *Applied Psychology* 53, 1: 61–86.
- [29] Patrick E Shrout and Joseph L Fleiss. 1979. Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin* 86, 2: 420.
- [30] George Siemens. 2014. Connectivism: A learning theory for the digital age.
- [31] Michael Tsang, George W Fitzmaurice, Gordon Kurtenbach, Azam Khan, and Bill Buxton. 2002. Boom chameleon: simultaneous capture of 3D viewpoint, voice and gesture annotations on a spatially-aware display. In *Proceedings of the 15th annual ACM symposium on User interface software and technology*, 111–120.
- [32] Patricia G Wojahn, Christine M Neuwirth, and Barbara Bullock. 1998. Effects of interfaces for annotation on communication in a collaborative task. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 456–463.
- [33] Dongwook Yoon, Nicholas Chen, François Guimbretière, and Abigail Sellen. 2014. RichReview: blending ink, speech, and gesture to support collaborative document review. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*, 481–490.
- [34] Dongwook Yoon, Nicholas Chen, Bernie Randles, Amy Cheatle, Corinna E Löckenhoff, Steven J Jackson, Abigail Sellen, and François Guimbretière. 2016. RichReview++: Deployment of a Collaborative Multi-modal Annotation System for Instructor Feedback and Peer Discussion. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 195–205.
- [35] Joyce Zhu, Jeremy Warner, Mitchell Gordon, Jeffery White, Renan Zanelatto, and Philip J Guo. 2015. Toward a Domain-Specific Visual Discussion Forum for Learning Computer Programming: An Empirical Study of a Popular MOOC Forum.
- [36] Sacha Zyto, David Karger, Mark Ackerman, and Sanjoy Mahajan. 2012. Successful classroom deployment of a social document annotation system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1883–1892.

Received June 2017; revised August 2017; accepted November 2017.