

“Hide Your Video, Show Your Action!” Investigating a New Video Conferencing Interface for Virtual Studying

Soobin Cho*

Soft Berry, Seoul, Republic of Korea, soobincho0930@gmail.com

Seoul National University, Seoul, Republic of Korea

Bongwon Suh

Seoul National University, Seoul, Republic of Korea, bongwon@snu.ac.kr

Joongseek Lee

Seoul National University, Seoul, Republic of Korea, joonlee8@snu.ac.kr

With the advent of COVID-19, new virtual social activities arose. These activities include virtual studying, which is studying while joining a video conference. Virtual studying is different from most virtual activities in that users try to minimize unnecessary interactions while sharing their presence through video streaming. Here, video streaming that runs in the background can cause problems such as invasion of privacy and excessive self-awareness. In this paper, we aim to investigate whether a new video conferencing interface that reduces video explicitness but detects important actions can mitigate the problems of video streaming and still deliver users' presence in virtual studying. To this end, we designed a research prototype in three versions: blurred video version, small video version, and no video version. All versions were provided with an activity recognizer that detects absence, leaning, and using a smartphone. To evaluate the feasibility of the design, we conducted a user study where four virtual studying teams used all three versions of the prototype and participated in an interview. Our study explored the effects of new design strategies for virtual studying, which is a new virtual activity that focuses on sharing presence.

CCS CONCEPTS • Human-centered computing • Collaborative and social computing • Collaborative and social computing design and evaluation methods

Additional Keywords and Phrases: Virtual studying, Video streaming, Presence, Awareness

ACM Reference Format:

First Author's Name, Initials, and Last Name, Second Author's Name, Initials, and Last Name, and Third Author's Name, Initials, and Last Name. 2018. The Title of the Paper: ACM Conference Proceedings Manuscript Submission Template: This is the subtitle of the paper, this document both explains and embodies the submission format for authors using Word. In Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY. ACM, New York, NY, USA, 10 pages. NOTE: This block will be automatically generated when manuscripts are processed after acceptance.

* This research stemmed from the author's work at the previous affiliation, Seoul National University.

1 INTRODUCTION

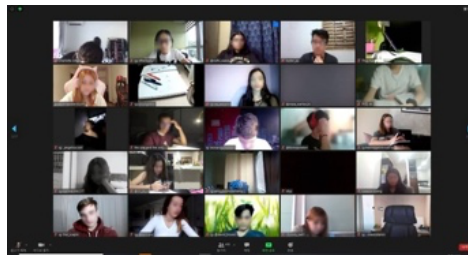


Figure 1: A screenshot of an open virtual study room, 'StudyStream' [20]

New virtual social activities emerged since the COVID-19 pandemic. A few of these activities involve video conferencing, including virtual eating, virtual exercising, and virtual studying. Among them, virtual studying is an activity where users study while joining a video conference (Figure 1). By joining a video conference where everyone studies, users aim to create a library-like studying environment at home. As with the libraries, unnecessary interactions are minimized in virtual studying. Instead, the users focus on sharing presence and awareness through video streaming.

One problem is that video streaming may cause inconveniences such as invasion of privacy and excessive self-awareness. This may especially be the case for virtual studying where video streaming is only used for sharing presence. In virtual studying, the camera runs in the background while users mind their own business, making it difficult for the users to care for and control what appears on the screen.

In this study, we aim to explore whether a new video conferencing interface that reduces video explicitness but detects important actions can resolve the problems of video streaming while delivering the presence of others in virtual studying. To this end, we designed a research prototype and conducted a user study with four virtual studying teams. The research prototype was designed in three versions: blurred video version, small video version, and no video version. All versions were provided with an activity recognizer that detects absence, leaning, and using a smartphone. For the user study, the virtual studying teams used all three versions of the prototype and participated in an interview asking whether the presence was maintained, whether the problems of video streaming were resolved, and their preference for each version.

The results of the study indicated that the presence was maintained in the blurred video and the small video versions of the prototype, and the problems of video streaming were resolved in all three versions. Among the three versions, the blurred video version was most preferred. This paper explored the effects of new design strategies for a new virtual activity that focuses on sharing presence.

2 RELATED WORKS

Virtual studying uses video streaming to share presence and awareness while studying. This resembles the use of video streaming in the early studies of media space. Media space is a tool connecting physically remote groups, and its early attempts used always-on video streaming to provide informal awareness of the counterparts' presence and activities [1, 12, 14, 16, 23]. The previous studies on media space point out the problems of video streaming and explore various design strategies to minimize the explicitness of the video.

The two major problems occurring due to video streaming were privacy concerns [2, 3, 9, 10, 17, 19, 22, 24] and excessive self-awareness [2, 8]. The privacy concerns arose as the users of media space had a lack of control over their privacy. When they video streamed themselves, they could not control who is watching them and what is shown on the video. Excessive self-awareness occurred as the users cared about how they appeared on the video and became

uncomfortable with the camera. Our previous study on virtual studying also revealed that the users were overly aware of how they appeared on the video and wanted to care less about how they looked [18]. This study explores whether these problems could be mitigated with the new interface.

Design strategies to minimize the explicitness of the video in media space included showing only captured video images [1, 2, 14, 24], simply blocking certain parts of the video [5, 6, 8, 21], and using video filters [2, 4, 7, 8, 11, 13, 15, 17]. Among them, the most common strategy was using blur or pixel filters [2, 4, 7, 8, 11, 13, 15]. In this paper, we investigate whether the strategies to minimize video explicitness can also be effectively applied to virtual studying.

3 METHOD

To investigate whether a new video conferencing interface can mitigate the problems of video streaming while still delivering users' presence, we designed a research prototype and conducted a user study with actual virtual studying teams.

3.1 Designing Research Prototype

We designed the research prototype in three versions: blurred video version, small video version, and no video version. The blurred video was chosen following the previous studies on media space. The small video was chosen to explore whether a small video size could also minimize video explicitness in virtual studying. No video was chosen as the very opposite of showing videos. The size of the blurred video was 700*700 pixels and the small video was 200*200 pixels.

All three versions of the prototype included an activity recognizer that detects absence, leaning, and using a smartphone. These activities were chosen as they were considered important when judging one's studying attitude in virtual studying, according to our previous study on virtual studying [18]. The activity recognizer utilized an image classification model that was made with Google Teachable Machine. It had five classes: upright, upright_with_phone, leaning, leaning_with_phone, and absent, and each class was trained with over 2,000 images collected by filming ourselves studying. The result of the classification appeared on the screen in the form of a traffic signal with red and green circles.

We created three websites containing each version of the prototype for the user study. Screenshots of the three prototype websites are shown in Figure 2.

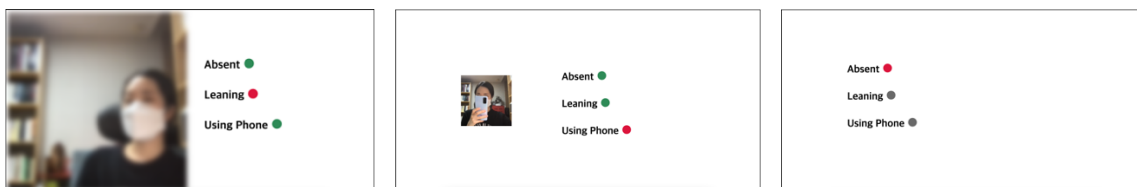


Figure 2: Screenshots of the prototype websites in the order of blurred video version, small video version, and no video version

3.2 User Study

For the user study, four virtual studying teams used all three versions of the prototype and participated in an interview.

The virtual studying teams were recruited through the online communities of various colleges. Participation criteria were limited to already-existing teams. A total of four teams composed of two to three users participated in the study. They were all female, and ages ranged from 24 to 35 with a median age of 26.5. Table 1 shows the demographic information of the study participants along with the order of the prototype versions used by each team.

Table 1: Demographic information and the order of the prototype versions used

Team	ID	Gender	Age	Occupation	Assigned Order
Team 1	P1	Female	27	Job Seeker	blur > small > none
	P2	Female	26	Job Seeker	
	P3	Female	25	Graduate Student	
Team 2	P4	Female	24	Undergraduate Student	small > none > blur
	P5	Female	24	Undergraduate Student	
	P6	Female	25	Job Seeker	
Team 3	P7	Female	28	Graduate Student	none > blur > small
	P8	Female	29	Office Worker	
Team 4	P9	Female	31	Undergraduate Student	blur > none > small
	P10	Female	35	Undergraduate Student	

The teams were first asked to read the purpose and the procedure of the study and how the information will be used. After agreeing to participate, they were given the links to the three prototype websites and were informed about the order of which version to use first, second, and third. They were then asked to use Discord, a video conferencing tool that allows multi-user screen share, to show each other's prototype website in a gallery view. When they were ready to study, one teammate sent us a picture before starting the studying session (Figure 3). For their 30-minute studying session, they were asked to study as usual while using each version for 10 minutes.

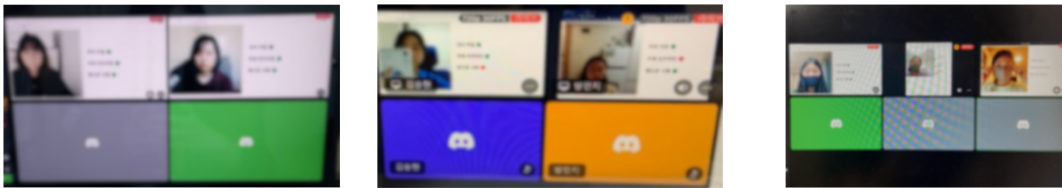


Figure 3: Picture of team 1, 2, and 3's virtual studying screen using the blurred video version of the research prototype

After the studying session, an interview was conducted asking whether the presence was maintained in each version, whether the problems of video streaming were mitigated in each version, and their preference for each version. The interview lasted for about 20 minutes. All interviews were recorded with the consent of the participants and were transcribed for analysis. After the interview, the participants were compensated with a voucher worth 30,000 KRW (about 23 USD).

The interview data were first quantitatively analyzed to measure how many participants answered that the presence was maintained and the problems were mitigated in each version, and to measure the rankings of the three versions. The reasons behind their answers were utilized to understand the relationships between the design strategy, maintaining presence, and resolving the problems of video streaming.

4 RESULTS AND DISCUSSION

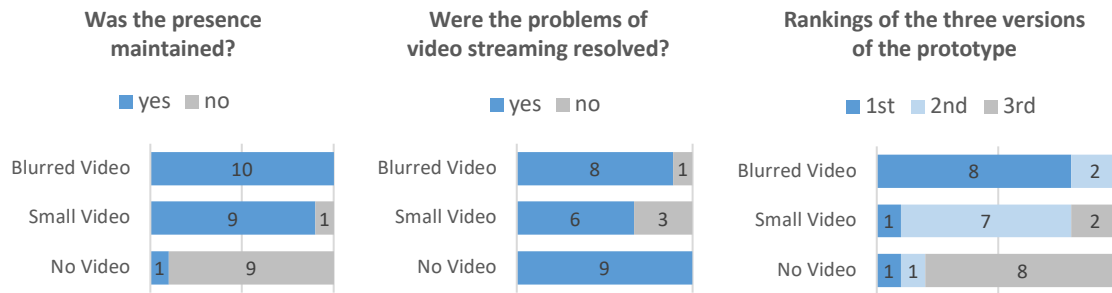


Figure 4: A summary of the answers to the interview questions

According to the interview results, nine to ten participants answered that the presence was maintained in the blurred video and the small video version. On the other hand, only one participant answered that the presence was maintained in the no video version. Regarding the problems of video streaming, six participants answered that the problems were resolved in all three versions. The rankings of the three versions of the prototype showed that the blurred video version was most preferred, followed by the small video version, and the no video version was least preferred (Figure 4). Additional insights regarding the relationships between the design strategy, maintaining presence, and resolving the problems of video streaming are as follows.

Presence and the Video Types. The interview results indicated that the presence was maintained when there is a video, even when the video is filtered or small. However, the presence did not seem to be delivered when there is no video at all. P10 answered, “The first (blurred version) and the third (small version) one felt like we were studying together because there was the video, but I couldn’t feel my friend’s presence when using the one with only the text because I couldn’t see her at all.” This may be because the main role of video streaming in virtual studying is to deliver presence. In that sense, video is important to feel the presence of others but a clear view of the counterpart is not necessary. Similarly, P7 said, “I liked the blurred version. I can see that someone’s there, I’m not trying to see their face, I only need the presence.” P9 also added, “I could still see my partner with the blur filter, just enough to make me feel like we’re studying together.”

Presence and the Activity Recognizer. Activity recognizer was initially designed to deliver the presence of others, but instead it was used to feel one’s own studying presence. While nine participants answered that they could not feel the presence of others without the video even with the constantly-changing signal of the activity recognizer, all ten participants used the recognizer for their self-surveillance effect. According to P10, “The absent signal made me stay in my seat, and trying not to leave my seat helped me to concentrate better.” P4 also said, “The red signal made me realize I was leaning and I liked how it alerts me,” and P5 added, “I was trying not to use my phone because the red signal was really strong.”

Problems of Video Streaming and the Video Types. Six out of nine participants answered that all three versions mitigated the problems of video streaming. One participant (P2) was not counted as she thought that video streaming does not cause problems when studying virtually. While the no video version most effectively solved the problems by eliminating the video, the blurred video also seemed to be a good solution since it “adequately hides (P9)” what might be “embarrassing to show to others, like my face and how I look (P3)”. Still, one participant (P8) answered that the problems remain as long as there is a video regardless of the filter and size. Aside from P8, the two participants (P7, P10) who also answered that the small video did not resolve the problems of video streaming explained that they were uncomfortable

because of the video clarity. P7 said, “Although it’s small, it was too clear and I was uncomfortable with it. They could see my face clearly and I was kind of concerned about how I appear on the screen.”

Presence and the Problems of Video Streaming. While maintaining the presence of others and resolving the problems of video streaming are both important, participants’ answers regarding the rankings indicated that they care more about maintaining the presence. The no video version completely removed the problems of video streaming but was ranked 3rd by eight participants, and all eight of them pointed out its lack of ability to deliver others’ presence. Between the blurred video and the small video version, the blurred video version was ranked 1st by more participants because the video was “bigger (P5)”, thus delivering “greater presence (P7)” and making the users “feel more like we’re together (P4)”. Yet, the blurred video, although it is filtered, is still a form of video streaming and was a burden for one participant. Thus, an ideal design strategy should eliminate the problems of video streaming like the no video version while delivering presence like the versions with the videos.

5 CONCLUSION AND FUTURE WORK

In this paper, we addressed virtual studying which is an activity focused on sharing presence through video streaming while minimizing any unnecessary interactions. We designed a new video conferencing interface and explored the effects and limitations of the design strategies for this new virtual activity.

However, our user study had three limitations and requires further research to confirm the applicability of the insights. First, the participants were all female and may have been more sensitive to the problems of video streaming. Second, 30 minutes of virtual studying session with 10 minutes for each version is insufficient, considering that the usual duration of a virtual studying session is two to three hours according to our previous research [18]. Lastly, the activity recognizer may not have correctly detected every move, possibly affecting the perceived effectiveness of the design. Based on this pilot study, our future study will explore additional design strategies and conduct research involving different groups of users in a more realistic setting of a virtual studying environment.

REFERENCES

- [1] Alan Borning and Michael Travers. 1991. Two approaches to casual interaction over computer and video networks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '91)*. Association for Computing Machinery, New York, NY, USA, 13–19. <https://doi.org/10.1145/108844.108847>
- [2] Alison Lee, Andreas Girgensohn, and Kevin Schlueter. 1997. NYNEX portholes: initial user reactions and redesign implications. In *Proceedings of the international ACM SIGGROUP conference on Supporting group work: the integration challenge (GROUP '97)*. Association for Computing Machinery, New York, NY, USA, 385–394. <https://doi.org/10.1145/266838.267359>
- [3] Beverly L. Harrison. 1993. The application of media space technology in high-technology corporate workplaces. *SIGOIS Bull.* 14, 1 (July 1993), 12. <https://doi.org/10.1145/155748.155752>
- [4] Carman Neustaedter and Saul Greenberg. 2003. The design of a context-aware home media space for balancing privacy and awareness. In *UbiComp 2003: Ubiquitous Computing, Lecture Notes in Computer Science*. Springer, Berlin, Heidelberg, 297-314. https://doi.org/10.1007/978-3-540-39653-6_24
- [5] Carman Neustaedter and Tejinder K. Judge. 2010. Peek-A-Boo: the design of a mobile family media space. In *Proceedings of the 12th ACM international conference adjunct papers on Ubiquitous computing - Adjunct (UbiComp '10 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 449–450. <https://doi.org/10.1145/1864431.1864482>
- [6] Joëlle Coutaz, François Bérard, Eric Carraux, and James L. Crowley. 1998. Early experience with the media space CoMedi. In *Proceedings of the IFIP TC2/TC13 WG2.7/WG13.4 Seventh Working Conference on Engineering for Human-Computer Interaction*. Kluwer Academic Publishers, USA, 57–72. https://doi.org/10.1007/978-0-387-35349-4_22
- [7] Joëlle Coutaz, François Bérard, Eric Carraux, William Astier, and James L. Crowley. 1999. CoMedi: using computer vision to support awareness and privacy in media spaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*. Association for Computing Machinery, New York, NY, USA, 13-14. <https://doi.org/10.1145/632716.632726>
- [8] Jose Eurico de Vasconcelos Filho, Kori M. Inkpen, and Mary Czerwinski. 2009. Image, appearance and vanity in the use of media spaces and video conference systems. In *Proceedings of the ACM 2009 international conference on Supporting group work (GROUP '09)*. Association for Computing Machinery, New York, NY, USA, 253–262. <https://doi.org/10.1145/1531674.1531712>

- [9] Margrethe H. Olson and Sara A. Bly. 1991. The Portland experience: a report on a distributed research group. *International Journal of Man-Machine Studies*. 34, 2 (Feb. 1991), 211–228. [https://doi.org/10.1016/0020-7373\(91\)90042-6](https://doi.org/10.1016/0020-7373(91)90042-6)
- [10] Marilyn M. Mantei, Ronald M. Baecker, Abigail J. Sellen, William A. S. Buxton, Thomas Milligan, and Barry Wellman. 1991. Experiences in the use of a media space. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '91)*. Association for Computing Machinery, New York, NY, USA, 203–208. <https://doi.org/10.1145/108844.108888>
- [11] Michael Boyle, Christopher Edwards, and Saul Greenberg. 2000. The effects of filtered video on awareness and privacy. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work (CSCW '00)*. Association for Computing Machinery, New York, NY, USA, 1–10. <https://doi.org/10.1145/358916.358935>
- [12] Michael Boyle and Saul Greenberg. 2005. The language of privacy: Learning from video media space analysis and design. *ACM Trans. Comput.-Hum. Interact.* 12, 2 (June 2005), 328–370. <https://doi.org/10.1145/1067860.1067868>
- [13] Michael Boyle, Carman Neustaedter, and Saul Greenberg. 2009. Privacy factors in video-based media spaces. In *Harrison, S. (eds) Media Space 20 + Years of Mediated Life*. Springer, London, 97-122. https://doi.org/10.1007/978-1-84882-483-6_7
- [14] Paul Dourish and Sara Bly. 1992. Portholes: supporting awareness in a distributed work group. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '92)*. Association for Computing Machinery, New York, NY, USA, 541–547. <https://doi.org/10.1145/142750.142982>
- [15] Qiang Alex Zhao and John T. Stasko. 1998. Evaluating image filtering based techniques in media space applications. In *Proceedings of the 1998 ACM conference on Computer supported cooperative work (CSCW '98)*. Association for Computing Machinery, New York, NY, USA, 11–18. <https://doi.org/10.1145/289444.289450>
- [16] Sara A. Bly, Steve R. Harrison, and Susan Irwin. 1993. Media spaces: bringing people together in a video, audio, and computing environment. *Commun. ACM*. 36, 1 (Jan. 1993), 28–46. <https://doi.org/10.1145/151233.151235>
- [17] Scott E. Hudson and Ian Smith. 1996. Techniques for addressing fundamental privacy and disruption tradeoffs in awareness support systems. In *Proceedings of the 1996 ACM conference on Computer supported cooperative work (CSCW '96)*. Association for Computing Machinery, New York, NY, USA, 248–257. <https://doi.org/10.1145/240080.240295>
- [18] Soobin Cho. 2022. Study on Remote Studying Activity with Video Conferencing Tool to Feel the Presence of Others: Focusing on ‘Revealing’ and ‘Hiding’ Themselves through Videos [Master’s thesis, Seoul National University].
- [19] Steve Harrison and Paul Dourish. 1996. Re-place-ing space: the roles of place and space in collaborative systems. In *Proceedings of the 1996 ACM conference on Computer supported cooperative work (CSCW '96)*. Association for Computing Machinery, New York, NY, USA, 67–76. <https://doi.org/10.1145/240080.240193>
- [20] StudyStream. accessed May 2022. <https://www.studystream.live/focus-room>
- [21] Tejinder K. Judge, Carman Neustaedter, and Andrew F. Kurtz. 2010. The family window: the design and evaluation of a domestic media space. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. Association for Computing Machinery, New York, NY, USA, 2361–2370. <https://doi.org/10.1145/1753326.1753682>
- [22] Victoria Bellotti and Abigail Sellen. 1993. Design for privacy in ubiquitous computing environments. In *Proceedings of the third conference on European Conference on Computer-Supported Cooperative Work (ECSCW'93)*. Kluwer Academic Publishers, USA, 77–92.
- [23] Victoria Bellotti, Robert Fish, Robert Kraut, Paul Dourish, Bill Gaver, Annette Adler, Sara Bly, Marilyn Mantei, and Gale Moore. 1994. Debating the media space design space. In *Conference Companion on Human Factors in Computing Systems (CHI '94)*. Association for Computing Machinery, New York, NY, USA, 193–194. <https://doi.org/10.1145/259963.260249>
- [24] William A.S. Buxton. 1997. Living in augmented reality: ubiquitous media and reactive environments. In K. Finn, A. Sellen & S. Wilber (Eds.). *Video Mediated Communication*. Hillsdale, N.J.: Erlbaum, 363-384. An earlier version of this chapter also appears in *Proceedings of Imagina '95*, 215-229.